The role of histone phosphorylations in chromatin regulation during mitosis

Jennifer Louise Mitchell

Biosciences Institute
Newcastle University

Doctor of Philosophy

September 2024

Abstract

A network of epigenetic components tightly regulates chromatin structure and accessibility, including histone modifications, transcription factors and reader proteins. The epigenetic landscape changes dramatically during mitotic chromosome condensation. Notably, histone phosphorylations increase significantly, contributing to chromosome segregation and reader protein recruitment. However, the functional role of many mitotic histone phosphorylations remains unclear. Histone H3 Serine 10 phosphorylation (H3S10ph), the most frequently studied mitotic phosphorylation, begins to enrich at the centromeric regions in late G2 of the cell cycle, before spreading along chromosome arms by prometaphase. However, whether significant H3S10ph enrichment occurs at specific loci remains unclear, and different studies have reached contrasting conclusions regarding the functions of H3S10ph. One possible issue is that traditional chromatin immunoprecipitation techniques struggle to distinguish subtle enrichments of abundant marks, preventing understanding of the detailed distribution of mitotic histone phosphorylations. To address this, we develop the CUT+RUN technique for use in mitotic HeLa cells. We also validate a quantitative ChIP approach, MINUTE-ChIP, targeting H3S10ph and H3S28ph in mouse embryonic stem cells. We develop an integrative analysis approach, combining MINUTE-ChIP-sequencing with chromatin epigenetic state data, to produce a high-resolution map of mitotic H3S10ph and H3S28ph enrichment in the context of the chromatin regulatory landscape. For the first time, we reveal significant mitotic enrichment peaks in both H3S10ph and H3S28ph at promoters in specific regulatory states, and identify gene ontology functional pathways correlating with these enrichments. We also analyse ChIP-seq of H3 threonine 3 phosphorylation (H3T3ph) and identify significant H3T3ph peaks at promoters in the absence of inhibitory adjacent marker H3 lysine 4 trimethylation (H3K4me3). We conclude that histone phosphorylations are significantly enriched at promoters in mitosis, dependent on the local chromatin state, and propose that previous findings linking histone phosphorylation to interphase gene regulation should be expanded to consider the mitotic gene regulatory context.

Acknowledgements

This has been an incredible few years of my life, full of challenges, late nights in the lab, laughter and pride. I wouldn't have made it this far without the love, support and encouragement of so many wonderful people. To everyone who has helped me along the way, be it teaching me, listening to my worries, or being there with a hug and an invite for coffee – you have my eternal gratitude.

I am deeply grateful to have spent this PhD working with such fabulous supervisors, Jon and Daniel. Jon, you have always had your door open for my regular random questions, and I am so grateful for all your help, advice and general wisdom. You have helped shape me into a more rigorous, thoughtful scientist, and more importantly you have always shown me understanding and kindness. Thank you for your unwavering support and faith in me, even when I didn't feel that way myself. Daniel, thank you for being a friendly face, for always being ready with a creative new idea for analysis, and for always encouraging me. You have helped me become a much more confident researcher. You have both been pivotal forces in my journey through this PhD. Thank you both for the time, effort and passion you have put into guiding me.

To Simon Elsasser, Philip Yung and their colleagues at the Karolinska Institute, thank you so much for kindly providing us with your MINUTE-ChIP-seq dataset, as well as your expertise in the technique. We appreciate the tremendous work that had already been put into developing this technique and generating this dataset, and I am grateful for your advice and guidance over the last two years. I have truly enjoyed playing with this dataset, and I look forward to continuing our collaboration.

Thank you to Lisa Russel's team, particularly Lisa and Phoebe, for providing equipment and reagents, and valued opinions and ideas, for CUT+RUN troubleshooting. Phoebe, thank you for keeping me company and for commiserating when experiments refused to behave, and for your guidance with library preparations.

I also want to thank my panel review team for their advice and encouragement each year, and for taking the time to read through my reports which definitely could have done with some more concision.

I have been lucky to be part of two amazing, inspiring research groups, the Computational Epigenomics group, and the Higgins lab group. Thank you to everyone who has been a part of these teams over the last four years, you gave me a friendly, supportive and collaborative environment to grow in, and I have been inspired by all of your creativity and resilience. Special thanks to Rebecca Harris for her wisdom in the lab, teaching me the tricks of Bash scripting, and welcoming me into the group. Massive thank you also to Rebecca for her hard graft generating the H3T3ph ChIP-seq dataset for this analysis.

To my Mum and Dad, Emma and Brian Mitchell, I can never thank you enough and I will never stop trying. Your unwavering support and encouragement are felt every day. You always listen and give advice - even when the work I am ranting about is basically in another language. Dad, thanks for always taking an interest in my studies, especially the coding, for always being ready to drive north at a moment's notice, and for the Big Hug Bug Hugs. Mum, our video chats have often kept me sane on a long day, and your brilliance is truly inspiring. Thank you for always believing in me and reminding me that I can do anything I set my mind to. I love you both so much.

Thanks to all my amazing friends across the UK and beyond who keep me laughing and are always at the end of a phone. Lucy, Sam and Alix, thanks for all the lovely cards, the west end trips, and your unerring support and love. Roo, thanks for our hilarious debriefs and for always having more faith in me than me. Iz, thanks for all the movie nights, yelling at me when my brain was being stupid, coming with me to zoos when life got too tiring, and being the bestest fren. Juliana, I am so lucky to have started this PhD journey with you as my office buddy. You are one of the most joyful, strong, amazing humans I have ever known, your curiosity and passion for the world around us is infectious and I have loved exploring our corner of the world together. Lastly to Em, the Best of best friends, I do not know where I would be or who I would be without you by my side all these years. Thanks for always listening to the 10-minute voice notes, the board game and crochet nights, the epic hen do, and for just being a constant pillar of love and laughter in my life. Wuv oo.

Finally, to my wife Leah. I can never put into words how grateful I am that you are in my life. You show me joy, support and unwavering strength every day. You have been by my side for the vast majority of this PhD, through the midnight labs, the Sunday Scaries, and the occasional (okay fairly frequent) tearful days. You are always there with kindness and you're my biggest cheerleader, and I wouldn't have finished this PhD without you. It has been an amazing 3 years growing and learning together. Thank you for all the coffee, for our morning commutes being silly, for always picking up the phone, and for listening to my ramblings about this thesis for the last several months. I love you.

Publications

Harris, R.J., Heer, M., Levasseur, M.D., Cartwright, T.N., Weston, B., **Mitchell, J.L.,** Coxhead, J.M., Gaughan, L., Prendergast, L., Rico, D., Higgins, J.M.G. Release of Histone H3K4-reading transcription factors from chromosomes in mitosis is independent of adjacent H3 phosphorylation. *Nat Commun* **14**, 7243 (2023).

Table of Contents

List of Figures

Appendix Figures

List of Tables

Appendix Tables

Chapter 1: Introduction

1.1 Chromatin regulation during mitosis

The structure of the eukaryotic genome must be carefully regulated in order to control vital cell processes, including gene expression levels, protein recruitment, and correct chromosome segregation during cell division. Chromatin structure is maintained and regulated throughout the cell cycle through numerous epigenetic components including transcription factors, reader, writer and eraser proteins, and post-translational modifications. These epigenetic markers interact with each other creating an intricate chromatin regulatory network, which is then differentially modulated through the cell cycle, to allow crucial processes such as DNA replication during S phase, and chromosome condensation and segregation during mitosis. The extensive structural changes to chromatin during mitosis and mitotic exit require tight regulation. Errors in this process have severe consequences; for example, chromosome mis-segregation can cause damaged inherited chromosomes and/or abnormal chromosome copy numbers (aneuploidy), promoting tumorigenesis (Santaguida and Amon, 2015). However, many of the molecular mechanisms underpinning chromatin regulation remain poorly understood, and research continues to expand with the aim of elucidating the precise genomic distributions and functional contributions of epigenetic regulatory factors. The epigenetic regulatory network of chromatin, and how it changes through the cell cycle, is detailed below.

1.1.1 Chromatin structure and regulatory elements

The organisation of the eukaryotic genome is a long-studied topic. Broadly, within the nucleus, chromosomes fold into discrete territories, within which are compartments of high-density interactions, segregating the whole genome into active and inactive compartments (Lieberman-Aiden *et al.,* 2009). Discrete mega-base scale regions known as topologically associated domains (TADs) are composed of DNA extrusion loops mediated by the ring-like cohesin complex. CTCF binding at TAD boundaries and strong interactions within a given TAD insulate enhancer activity, making TADs fundamental components of structural genome regulation (Dixon et al., 2012; Narendra et al., 2015; Nora et al., 2012; Sexton et al., 2012; Sikorska & Sexton, 2020; Symmons et al., 2014). These genomic structures within the nucleus

are illustrated in **Figure 1.1A**. The genome is organised into a structure known as chromatin, in which 146 base pairs (bp) of DNA are wound around an octamer of proteins called histones. This histone octamer together forms a nucleosome, shown in **Figure 1.1B** (crystal structure first reported by Luger et al., 1997). Arrays of repeating nucleosome units form short and long-range interactions, condensing DNA to form the chromatin structure illustrated in **Figure 1.1B**. Briefly, short-range interactions between neighbouring nucleosomes create chromatin fibres, and further short-range interactions between individual nucleosomes in different fibres can further condense fibres into less accessible chromatin regions referred to as heterochromatin (Luger et al., 2012). Spacing between nucleosomes varies depending on DNA sequence, chromatin remodelling enzymes, and DNA binding factors (e.g. transcription factors), and secondary and tertiary structure of chromatin fibres is stabilised by architectural proteins (e.g. heterochromatin protein 1 (HP1)).

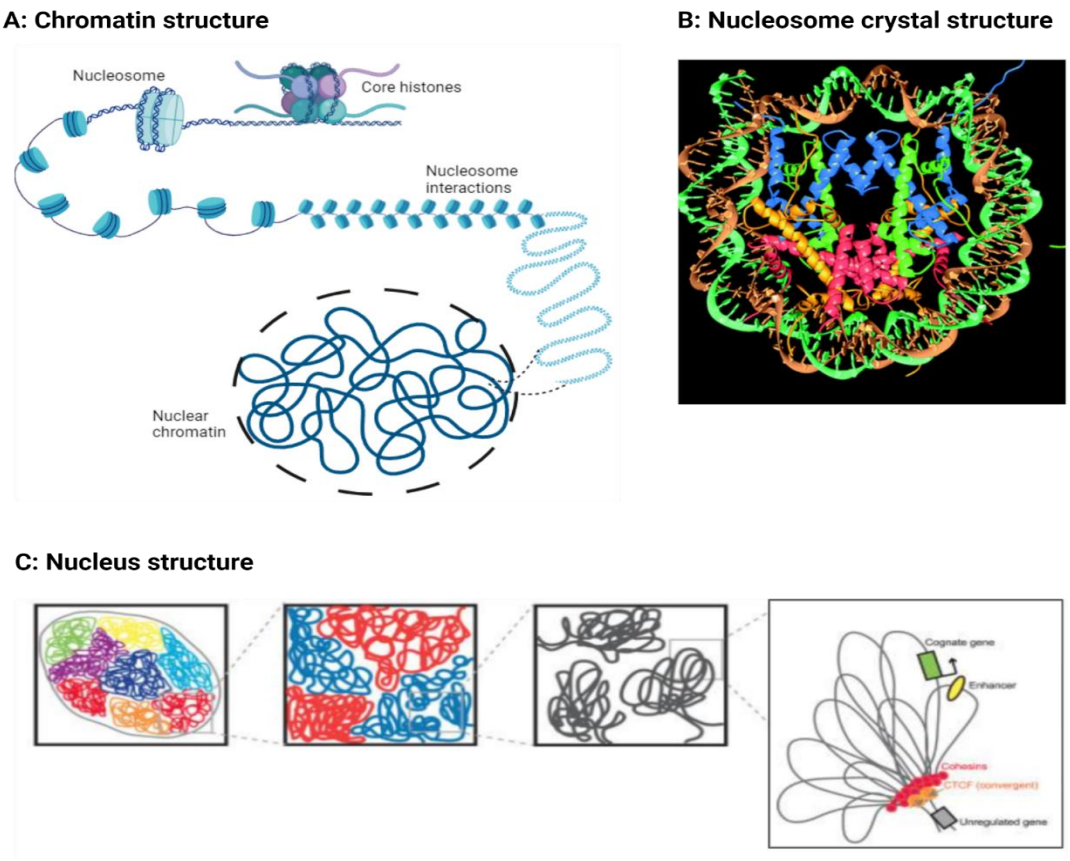


Figure 1. 1: Eukaryotic genome structure.

A: Structure of chromatin. *Nucleosomes are composed of an octamer of core histones (pink, purple, green) with protruding amino acid tails. 146 bp of DNA (dark blue string) is wound around each histone octamer. Nucleosomes interact with each other to form chromatin fibres.* ***B: Crystal structure of the nucleosome core.*** *The 146bp DNA phosphodiester backbone (brown and turquoise ribbon traces) is wound around eight histone proteins (main chains shown: H3(blue); H4(green); H2A(yellow); H2B(red).* ***C: Chromatin organisation in the nucleus.*** *The eukaryotic genome is separated into discrete territories. Each territory contains active and inactive compartments, and each compartment is composed of extrusion loops mediated by cohesin rings, known as topologically-associated domains (TADs). These TADs insulate enhancers. Figure contains adapted material from Luger et al,, 1997; Sikorska, 2019, and original illustrations.*

Chromatin structure and accessibility are tightly regulated by a range of epigenetic factors, including histone variants and histone post-translational modifications (PTMs), remodelling enzymes, and transcription factors (TFs), creating a complex transcription regulatory system (Woodcock & Dimitrov, 2001). This project focuses on the regulatory roles of histone post-translational modifications (PTMs).

Histones have terminal amino acid tails protruding from the nucleosome core, susceptible to extensive post-translational modification (Kouzarides, 2007). The most common PTMs of the four core histones (H3, H4, H2A and H2B) are methylation, acetylation, ubiquitination and phosphorylation. Histones can be mono-, di- or tri-methylated (me1, me2, me3) on lysine, or mono-, asymmetrically or symmetrically di-methylated (me, me2, me2a) on arginine. Ubiquitination and acetylation can also occur on lysine, while phosphorylations tend to occur on threonine and serine. These histone PTMs can be deposited by "writer" enzymes, removed by "eraser" enzymes, and recognised or bound by "reader" proteins, as illustrated in **Figure 1.2**. Histone PTMs form a complex epigenetic signalling "histone code" through crosstalk between modifications and/or interacting proteins (Strahl & Allis, 2000; Wang & Higgins, 2013).

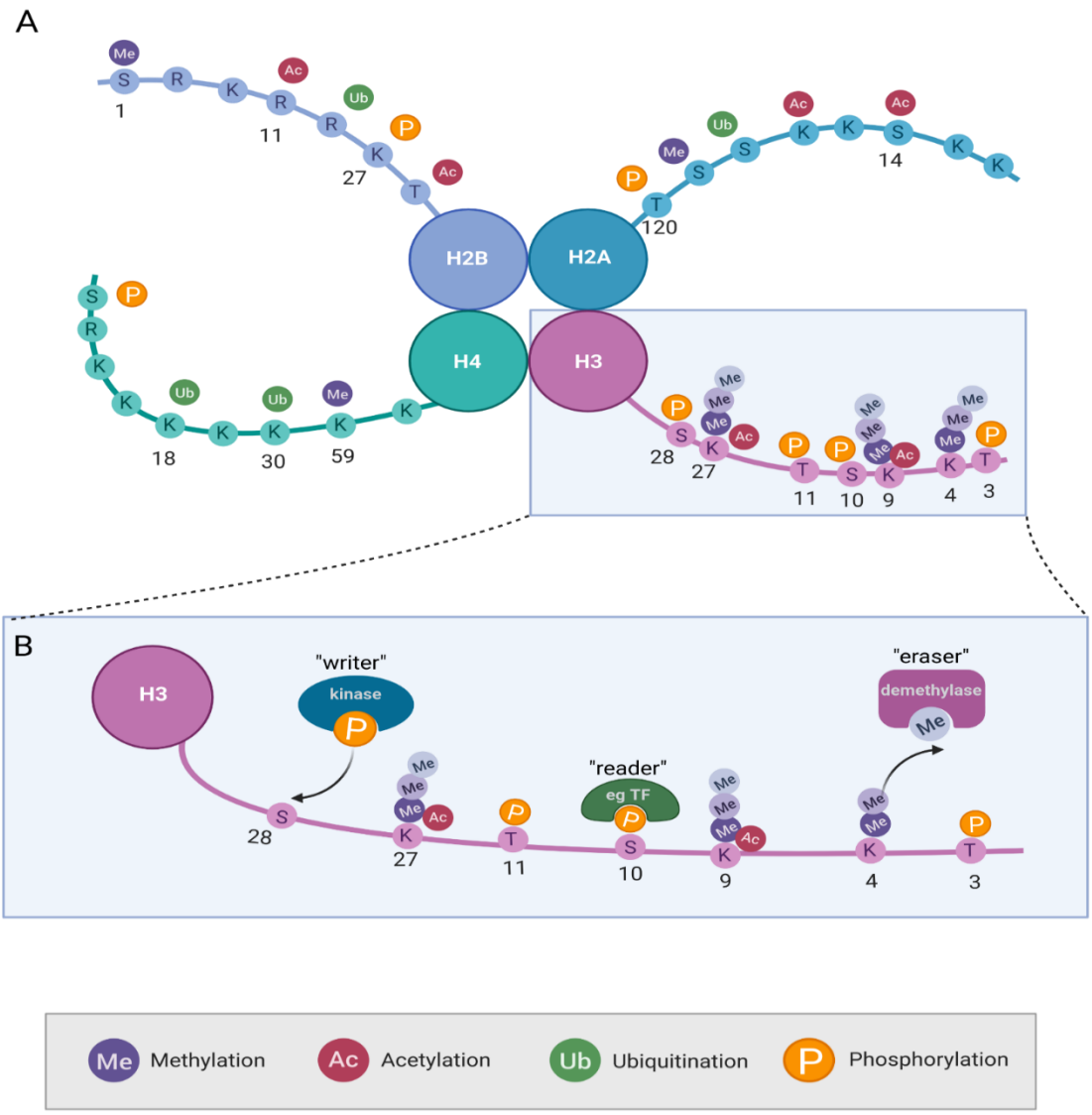


Figure 1. 2: The histone code.

The common modifications of the four core histones H2A, H2B, H3, H4: methylation (purple), acetylation (red), ubiquitination (green) and phosphorylation (orange). Key provided below. Amino acid residue numbers are labelled. "Writer" enzymes (left) such as phosphorylation kinases (blue) introduce chemical groups as post translational modifications (PTMs) to specific amino acids in the protruding tails of histones. "Eraser" enzymes such as demethylases (pink) remove these PTMs. "Reader" proteins such as transcription factors (TF, green) or regulatory proteins recognise specific histone PTMs and bind them. Figure is an original illustration.

1.1.2 Chromatin regulation during mitosis

Cell division presents a very interesting field for studying the epigenetic landscape. Long-range interactions and transcription activity are largely lost, while some epigenetic components are retained or even enriched.

During mitosis, chromatin structure is dramatically altered in order to faithfully segregate replicate chromosomes into daughter cells, as depicted in **Figure 1.3**. This condensed chromosome organisation has been observable for over a century; Carl Nageli unknowingly observed mitosis in the early 1800s, mistaking it for a dead cell abnormality, until 40 years later Flemming developed a way to stain chromosomes and observed mitosis in salamander embryos in 1879 (Flemming, 1879). Long-range chromatin interactions such as TADs and compartments are not seen in mitosis, supported by numerous studies (Dileep et al., 2015; Gibcus et al., 2018; Kang et al., 2020a; Naumova et al., 2013). As cells enter mitosis, beginning with prophase, the chromatin condenses extensively to form chromosomes made up of identical sister chromatids of replicate DNA. Sister chromatids then line up at the cell equator during metaphase and spindle microtubules form attachments between the kinetochore, a region at the centromere of each sister chromatid. The microtubules then contract and pull sister chromatids to opposite ends of the cell in anaphase, ensuring an identical set of chromosomes is segregated into each daughter cell. The cell then divides by cytokinesis, and the chromosomes decondense into interphase chromatin structure during telophase, and the daughter cells exit mitosis.

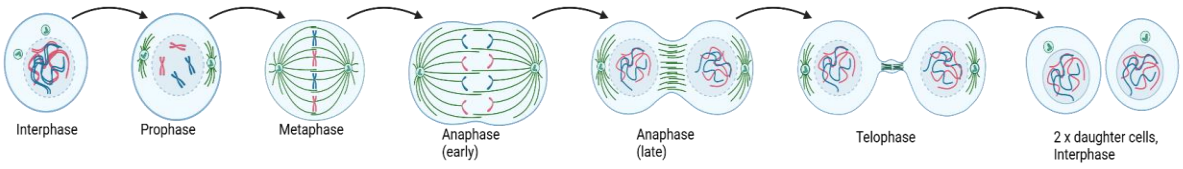


Figure 1. 3: Chromosomes during the stages of mitosis.

Interphase (left) and the stages of mitosis are shown: prophase, metaphase, anaphase, telophase. Chromosomes and chromatin = dark blue and red; spindle microtubules = green. Figure is an original illustration.

Alongside this change in 3D genome structure, gene expression levels are greatly reduced during mitosis. Until recently, transcription was thought to be almost entirely silenced during mitosis, among the proposed causes: transcription factor(TF)-DNA binding inactivation, and dispersal of TF and RNA polymerase II following nuclear envelope breakdown (Dileep et al., 2015; Gibcus et al., 2018; Gottesfeld & Forbes, 1997; Martínez-Balbás et al., 1995; Naumova et al., 2013; Parsons & Spencer, 1997; Prasanth et al., 2003; Prescott & Bender, 1962; Taylor, 1960). However, some studies have discussed that some transcription may occur during mitosis; one example study by Palozola et al. (2017) developed a method using nascent RNA labelling (EU-RNA-seq), which revealed low-level mitotic transcription remained (see **Figure 1.4A**), suggesting technical limitations had until then prevented detection of nascent RNA production during mitosis. EU-RNA-seq data from (Kang et al., 2020a) also observed transcriptional activity in a small subset of genes during mitosis (see **Figure 1.4B**). The mechanisms underlying low-level mitotic transcription and its regulation remain poorly understood.

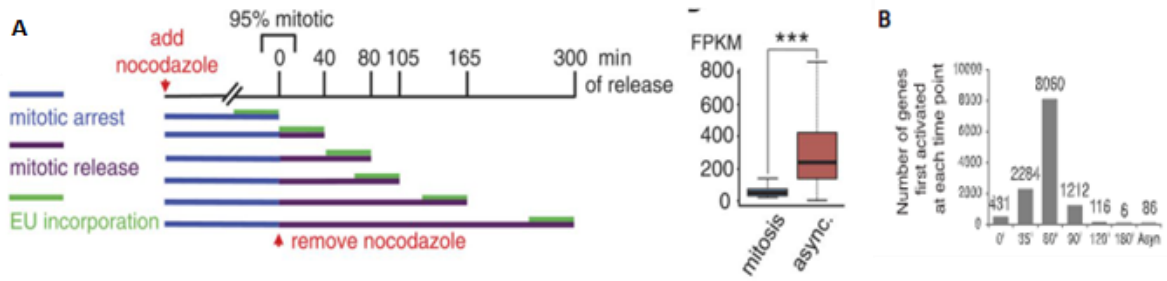


Figure 1. 4: Detection of mitotic nascent RNA.

(A) *In a study by (Palozola et al., 2017)**,** 5-ethynyluridine labelling (green) of nascent RNA transcripts was performed at times 0-300 mins post-mitotic release. 0 min extracts are 95% mitotic. Right plot shows Fragments per kilobase per million fragments mapped (FPKM) of mitotically expressed transcripts in mitosis and in asynchronous cells. Bar = mean, whiskers = quartiles, P<0.0001, n = 8074 transcripts.* ***(B)*** *Figure from (Kang, Maxim N Shokhirev, et al., 2020), showing gene activation at 0, 35, 60, 90, 120, 180 mins post mitotic release. 431 genes found to be active during mitosis (0'). Figure adapted from Palozola et al., 2017 and Kang et al 2020.*

On exit from mitosis, once sister chromatids have been segregated into the two daughter cells, 3D chromatin interactions including TADs and compartments are re-established rapidly. The Palozola et al. (2017) and Kang et al., (2020) studies shown above in **Figure 1.4** nicely demonstrate the sequential gene reactivation using data extraction at increasing times post mitotic release. Development in recent years of tools like EU-RNA-seq are helping to elucidate the exact timing of restoring the interphase epigenetic landscape. However, the molecular mechanisms by which interphase chromatin structure and gene expression are re-established remain unclear.

One proposed mechanism is that, rather than *de novo* regeneration of the interphase epigenetic landscape, a small subset of epigenetic factors/"bookmarkers" remain in place during mitosis, serving as a mitotic "memory" (first proposed by John and Workman, 1998; for review see Wang and Higgins, 2013; Gonzalez, Molliex and Navarro, 2021). Factors proposed to play mitotic bookmarking roles include histone variants, DNA methylation, noncoding RNA, transcription factors (TFs) and histone PTMs (Caravaca et al., 2013; Deluz et al., 2016; Egli et al., 2008; Festuccia et al., 2016; Kadauke et al., 2012; Moazed, 2011; Sarkies & Sale, 2012; Teves et al., 2016).

This project addresses the key research question: how do cells regulate the drastic structural and epigenetic changes that occur during mitosis, and then "remember" interphase chromatin regulatory landscape on mitotic exit?

1.2 The roles of histone phosphorylations in mitosis

1.2.1 Histone modifications in mitosis

To address the above key research question, this project focuses on histone modifications known to widely show increased and dynamic enrichment during mitosis: histone phosphorylations.

Many histone PTMs show mitosis-specific alterations. Historically, histone acetylation was reported to largely be lost during mitosis (D'Anna et al., 1977; Kruhlak et al., 2001). Although, more recent studies have identified retention of some acetylations across mitotic chromatin, including H3K14ac, H3K27ac, H3K122ac, H4K8ac and H4K16ac (Behera *et al.*, 2019; Bellec *et al.*, 2022). Similarly, ubiquitination of histones H2A and H2B appears to decrease during mitosis (Joo et al., 2007; Mueller et al., 1985). Contrastingly, some histone PTMs are retained or enriched during mitosis, as described in primarily immunofluorescence studies, going back over 30 years (Turner, 1989). Histone methylations are often retained during mitosis; for example H3K4me2/3 (histone H3 lysine 4 di/tri-methylation), associated with active gene promoters, remains constant through mitosis, potentially "bookmarking" genes for transcription reactivation (Sullivan and Karpen, 2004; Muramoto et al., 2010; Ribeiro et al., 2010; Bergmann et al., 2011; Javasky et al., 2018; Kang et al., 2020).

Recent techniques measuring mitotic transcription activity (e.g. EU-RNA-seq pulse-labelling nascent RNA transcripts) have allowed comparison of histone modification data with post-mitotic transcription reactivation (Kang et al., 2020a; Palozola et al., 2017, 2019). However, functional mechanisms linking histone modifications to mitotic gene expression remain poorly understood.

The most abundant and dynamic PTMs in mitosis are histone phosphorylations, discussed in the section below (Olsen et al., 2010a).

1.2.2 Histone phosphorylations in mitosis

Histone phosphorylations are of particular interest in the context of mitosis due to their significant enrichment in mitosis compared to interphase. Since the early 1970s, where

electrophoresis and densitometer mobility measurements of isolated histone samples were used to calculate histone phosphate content, histones H1 and particularly core histone H3 were shown to be hyper-phosphorylated during mitosis in correlation with chromosome condensation (Gurley et al., 1973). Since then, phosphorylation has been identified at numerous specific amino acids of the histone tails, many of which show specific interesting enrichment patterns along the chromosomes. For example, phosphorylation of serine 10 on histone H3 (H3S10ph) is observed to begin at the pericentromeric regions in G2 before spreading all along the chromosome arms by prophase of mitosis in tight correlation with mitotic chromosome condensation; this widespread serine 10 phosphorylation is highly conserved in eukaryotes (Gurley *et al.*, 1978; Hendzel *et al.*, 1997; Wei *et al.*, 1998). Serine 28 of histone H3 also shows widespread phosphorylation during mitosis across eukaryotes (Goto et al., 1999). Contrastingly, histone H3 threonine 3 phosphorylation (H3T3ph), and histone H2A threonine 120 phosphorylation (H2AT120ph) are found enriched at centromeric regions of the mitotic chromosomes (Aihara et al., 2004; Dai & Higgins, 2005; Kawashima et al., 2010). It is possible that these mitotically enriched histone phosphorylations could act as "bookmarks" to maintain and remember chromatin states through cell division (Hsiung et al., 2016; Kang et al., 2020a; Liang et al., 2015; Oomen et al., 2019a; Pelham-Webb et al., 2021; Valls et al., 2005).

These histone phosphorylations also have numerous demonstrated functional roles in regulating mitotic chromatin. Many histone phosphorylations are enriched at the centromeric regions during mitosis, with proposed functional roles in faithful spindle orientation, kinetochore attachment, and chromosome segregation. Other histone phosphorylations are seen spread along the chromosome arms during mitosis, although the precise locations, levels, and functions of many of these phosphorylations remain unclear. A few key histone phosphorylations of interest and evidence of their functional roles are discussed below.

H3T3ph and H2AT120ph: vital roles in accurate chromosome segregation

Both Histone H3 phosphorylation of threonine 3 (H3T3ph) and histone H2A phosphorylation at tyrosine 120 (H2AT120ph) are prime examples of histone phosphorylation playing critical roles in mitotic chromosome segregation. Kawashima et al. (2010) first found that H2AT120 was phosphorylated by the kinetochore-associated kinase Bub1, with immunofluorescent staining showing H2AT120ph localising at the centromeres immediately beneath kinetochores. This phosphorylation was found to be critical for recruitment of shugoshin protein (Sgo), a protector of centromere cohesion during early phases of mitosis (Kitajima et al., 2004). Sgo also acts as a recruiter of the chromosomal passenger complex (CPC) to the centromeres (Kawashima et al., 2007; Wang, Ulyanova, Van Der Waal, et al., 2011a). Composed of Aurora B and its chromatin targeting factors INCENP, Survivin and Borealin, this complex is required for correcting erroneous kinetochore-microtubule attachments and maintenance of checkpoint signalling to avoid chromosome mis-segregation (Schmitz et al., 2020). Kawashima *et al.* (2010) found that in fission yeast *Schizosaccharomyces pombe*, H2AT120ph mutants, as well as Bub1 kinase-dead mutants, lost centromeric shugoshin localisation, and lead to chromosomal instability. This demonstrates a critical role for H2A-T120ph/Bub1 in faithful chromosome segregation in mitosis, although it is noted that mutant or kinase-knockdown experiments such as these are near impossible in human cells, so whether these mechanisms are conserved across all eukaryotes is less clear. Interestingly, H3T3ph has been shown to play a similar role in faithful chromosome segregation in mammalian cells. H3T3 is phosphorylated during mitosis initially on the chromosome arms in prophase by the kinase haspin, before it appears to enrich at the inner centromeres by prometaphase-metaphase (Dai et al., 2006a; Dai & Higgins, 2005a; Markaki et al., 2009)**.** H3T3ph has been shown to directly bind the BIR domain of survivin, a subunit of the CPC (Kelly et al., 2010; Wang et al., 2010; Yamagishi et al., 2010)**.** These two pathways, H3T3ph and H2AT120ph, can recruit CPC independently, with loss of either phosphorylation alone causing partial de-localisation of the CPC, although loss of both phosphorylations does appear to have an additive effect with increased de-localisation detected (Broad et al., 2020; De Antoni et al., 2012; Hadders et al., 2020; Kelly et al., 2010; Wang et al., 2010; Wang, Ulyanova, Van Der Waal, et al., 2011b; Wang et al., 2012a; Yamagishi et al., 2010)**.** Thus, both H3T3ph and H2AT120ph have potential roles in CPC

recruitment, enabling correct kinetochore-spindle attachments and accurate chromosome segregation in mitosis.

While functional roles have been proposed for the above histone phosphorylations, their certainty and conservation across eukaryotes is less clear. Moreover, the subtle distributions and putative functions of other mitosis-enriched phosphorylations remain poorly understood. This can be in part attributed to limitations in experimental techniques, broad widespread enrichments making peak identification more challenging, and/or previous lack of strong, specific anti-phosphorylation antibodies. Two examples of such phosphorylations are discussed below.

H3S10ph and H3S28ph

Histone H3 serine 10 phosphorylation (H3S10ph), mediated by the kinase Aurora B, is one of the most studied mitotic histone modifications and is conserved from yeast to humans (Hendzel et al., 1997; Hsu et al., 2000; Prigent & Dimitrov, 2003). Soon after the identification of H3S10ph, it was discovered that other histone phosphorylations driven by Aurora B also occur. Phosphorylation of H3 serine 28 (H3S28ph) was identified by Goto et al. (2002) as a separate phosphorylation event to H3S10ph, initiating in prophase, later than H3S10ph's appearance in late G2. H3S10ph in particular provides a very interesting example of difficulties to-date in elucidating mitotic functional roles. Early studies showing H3S10ph correlating spatially and temporally with mitotic chromosome condensation across a range of eukaryotes sparked research interest in this mitotic histone phosphorylation *(Gurley et al., 1978; Wei et al., 1998)*. With the development of H3S10ph-specific antibodies, H3S10ph was then shown to appear first at the pericentromeric chromatin in early G2 phase and spreading to cover the chromosome arms by metaphase (Hendzel et al., 1997). However, while H3S10ph is arguably the most famous and most studied histone phosphorylation, the literature presents many often-contradictory findings as to its possible functional role(s).

Due to the strong correlation of H3S10ph deposition with chromosome condensation on entry into mitosis, some argue that H3S10ph is required for correct chromosome condensation, but evidence is often conflicting. In favour of a role for H3S10ph in chromosome condensation,

studies in *Tetrahymena* found an H3S10A mutation showed disrupted chromatin condensation during mitosis (Wei et al., 1999). In budding yeast, mutation of an Aurora B homologue Ipl1 and serine 10 mutations both showed disturbed condensation specifically in anaphase, of long chromosome arms (Neurohr et al., 2011). Inhibition or depletion of Aurora B also showed reduced condensation in anaphase in human cells and fission yeast respectively, however H3S10 was not the implicated substrate (Mora-Bermúdez et al., 2007; Tada et al., 2011). This highlights an important point that studies using Aurora B inhibition or depletion must be interpreted very carefully, as the kinase has many downstream substrates and is integral to many mitotic chromatin regulatory processes; therefore knockdown phenotypes could be caused by any number of substrates and pathway disturbances. Depletion of Aurora B also compromises the mitotic checkpoint which could lead to chromosome decondensation as a result of mitotic exit. In a more recent study, Meel *et al.*, (2024) instead used Serine 10 mutation in human HeLa cells. They found H3S10ph showed mitotic spreading in broad enrichment "islands", and found that pre-mitotic S10A mutation led to more open euchromatin in these regions, followed by more rapid transcription activation of these regions on mitotic exit. They used these findings to argue a role of H3S10ph in mitotic "bookmarking" through compaction (i.e.. condensation) of genes via these broad islands, in order to regulate their transcription activation timing after mitosis.

Contrastingly, numerous *in vitro* studies argue against a role of H3S10ph in condensation (e.g. de la Barre et al., 2001; Houston et al., 2008), and H3S10A mutation has no effect on condensation in budding yeast (Hsu et al., 2000). Depletion or mutation of kinase Aurora B has very little impact on chromosome condensation in human cells (Dai et al., 2006a; Ditchfield et al., 2003; Hauf et al., 2003).

Functional roles of H3S10ph have often been investigated through studies inhibiting Aurora B kinase; a few examples are described here. One proposed role of H3S10ph is regulation of heterochromatin protein 1 (HP1). HP1, recruited by H3K9me3, plays a key role in heterochromatin structure and accessibility, and dissociates from chromatin during mitosis (Murzina et al., 1999; Schmiedeberg et al., 2004; Sugimoto et al., 2001). Hirota et al. (2005) depleted H3S10ph by inhibiting Aurora B kinase using its inhibitor Hesperadin. They found that when Aurora B was inhibited, heterochromatin protein 1 (HP1) remained enriched at

pericentromeric chromatin. With consistent findings in immunoblotting *in vitro* experiments, and in a similar study removing H3S10ph by inhibiting Aurora B (Fischle et al., 2005), these researchers proposed a mechanism where H3S10 phosphorylation in mitosis leads to HP1 dissociating from the adjacent H3K9me3 recruit in a model known as the "methyl-phos switch". It is therefore conceivable that H3S10ph may contribute to dissociation of HP1 from chromosome arms during mitosis. It is also hypothesised that the reported role of H3S10ph in transcription activation in interphase could be due to an ability of H3S10ph to displace HP1 (Kouzarides, 2007; Wang & Higgins, 2013). A similar phospho-methyl switch model has also been suggested for H3S28ph, displacing polycomb group proteins from H3K27me3 heterochromatin regions in both mitosis and interphase (Gehani et al., 2010; Goto et al., 2002). Removal of H3S10ph by the Aurora B inhibitors such as Hesperadin and ZM447439 has also been shown to lead to erroneous spindle-kinetochore attachment in multiple species including mammalian and *xenopus* studies (e.g. Cimini et al., 2006; Gadea & Ruderman, 2005; Lane et al., 2010). This could indicate a role of H3S10 phosphorylation in chromosome segregation. However, the role of Aurora B in mitotic regulation is complex, with numerous proposed substrates. For example, a gradient of Aurora B centred at the spindle midzone has recently been proposed to phosphorylate multiple kinetochore substrates, including the KMN (Knl1 complex, Mis12 complex, Ndc80 complex) protein Dsn1, during anaphase (Fuller et al., 2008; Papini et al., 2021). Aurora B also has numerous downstream substrates important in mitosis; for example, Dsn1 is involved in building kinetochores, and Ndc80 required to detach incorrect microtubules. Aurora B inhibition, as used in the above-described studies, is therefore likely to have more widespread impact on mitotic regulation, leaving the importance of H3S10ph alone unknown. Studies mutating H3S10 to alanine led to chromosome mis-segregation in fission yeast and *Tetrahymena*, but did not produce any detectable phenotype in budding yeast (Mellone et al., 2003; Wei et al., 1999). However, homogenous mutant histones are almost impossible to generate in multicellular eukaryotes, due to multiple histone gene copies (often at multiple loci), and these studies are further complicated by histone variants such as H3.3 with overlapping function (Nevil & Duronio, 2021).

H3S10ph nicely exemplifies a large challenge in studying mammalian histone PTMs, where direct histone mutations are largely impossible, and indirect approaches knocking down their kinases/writers (when known) can impact the multiple targets (non-histone or other histone

residues) of that kinase, making it difficult to infer the importance of the one target histone PTM. These limitations make it difficult to identify the true functional gains or losses resulting from H3S10 phosphorylation.

Interestingly, H3S10ph also has proposed enrichment in chromatin "open" regions during interphase, with studies showing H3S10ph abundance in early DNA replication timing domains and negative correlation with repressive PTM H3K9me2 and lamina-associated domains (Chen et al., 2018; Komar & Juszczynski, 2020). Both H3S10ph and H3S28ph have reported involvement in transcription regulation in interphase, with multiple kinases reported to be able to phosphorylate these residues (Komar & Juszczynski, 2020; Wang & Higgins, 2013; Yung et al., 2015). A few example studies and proposed mechanisms linking H3S10ph and H3S28ph to transcription are discussed below.

In 1991,Mahadevan et al. saw that when fibroblast cells were stimulated with growth factors as well as phosphatase and protein synthesis inhibitors, the expression profile of immediate-early gene activation was closely mirrored by a rapid phosphorylation of histone H3S10 distinct from that seen in dividing cells. DeManno et al. (1999) also found increased H3S10ph in response to early differentiation signalling by FSH in ovarian follicles. Other evidence includes a study in the suprachiasmatic nucleus (SCN) of rats. Light stimulation induced H3 phosphorylation staining in SCN sections, with H3S10ph distribution coinciding with transcription profiles of the immediate-early gene *c-fos* and circadian gene *Per1 (*Crosio *et al.,* 2000), suggesting a role for H3S10ph in circadian regulation. In yeast where extensive transcriptome analysis has been performed, H3S10ph has been shown as essential for expression of only certain genes, indicating H3S10ph is not a global requirement but rather may play a distinct role at certain promoters. This evidence could be used to argue a potential role for H3S10ph in bookmarking genes for transcription regulatory pathways. The role of H3S10ph in transcriptional activation has also emerged as a potential player in carcinogenesis (for review see Komar & Juszczynski, 2020). It is possible that H3S10ph function is context dependent and may be affected by other histone modifications.

In the case of H3S28 phosphorylation, studies in mammalian cells have shown H3S28ph involvement in transcription regulation. Polycomb repressive complex 1 and 2 (PRC1 and PRC2) bind to repressive H3K27me2 and H3K27me3 to repress certain genes. It is possible

that H3S28ph acts in a methyl-phos switch similar to that of H3S10ph and H3K9me3, in which H3S28 phosphorylation can displace PRC complexes from their H3K27me2/3 binding sites. Mammalian studies have shown that in response to stress and developmental signalling, stress-activated kinase MSK phosphorylates H3S28, and that H3S28ph displaces polycomb group proteins PRC1 and PRC2, "de-repressing", or effectively activating, polycomb target genes (Gehani et al., 2010; Lau & Cheung, 2011). H3S28ph has also been seen to activate stress-response genes by displacing corepressor complex HDAC (Sawicka & Seiser, 2014). However, other biochemical studies have found that H3S28ph in combination with H3K27me3 (H3K27me3S28ph) prevents demethylation of H3K27me3 by UTX and JMJD3, thereby preserving repressive mark H3K27me3 (Kruidenier et al., 2012; Sengoku & Yokoyama, 2011).

Multiple mechanisms have been proposed as to how these phosphorylations may regulate transcription, such as the methyl-phos switch model described above. As another example, H3S10ph and H3S28ph peptides were found to bind 14-3-3, a protein recruited to *FOS* and *JUN* promoters upon gene activation. 14-3-3 also has demonstrated importance for transcriptional activation of genes *HDAC1* and *GAL1* in mammalian cells and budding yeast respectively (Macdonald et al., 2005; Walter et al., 2008; Winter et al., 2008). 14-3-3 binding could also act to recruit additional factors, such as histone remodelers, and thus stimulate transcription (Drobic et al., 2010).

Overall, it can be seen that H3S10ph and H3S28ph have numerous proposed functional roles, with evidence varying depending on experimental methods (e.g. kinase knockdowns), cell lines, and cell cycle stage. It could be argued that these histone phosphorylations may well contribute to multiple regulatory mechanisms, and that their roles and enrichments likely differ depending on cell contexts and in relation to other histone modifications and regulatory factors. We suggest that a broader analysis approach, taking into account both cell cycle stage and the context of other regulatory markers, will improve our understanding of the functional contributions of these histone phosphorylations in chromatin regulation.

1.2.3 How to elucidate functional roles of mitotic histone phosphorylations

Among the main obstacles in understanding histone phosphorylations during mitosis are mapping their precise genomic loci - particularly for abundant phosphorylations where any enrichments are likely to be subtle. Historically, techniques such as confocal microscopy imaging and mass spectrometry generated data on the broad distribution of mitotic histone phosphorylations at the chromosomal level, as demonstrated in **Figure 1.5**. This can suggest broadly the likely roles – for example H3T3ph localised distribution (see **Figure 1.5A**) suggests a role at centromeric regions is most likely. However, the roles of more widely enriched phosphorylations are harder to identify.

More recent approaches have still been unable to identify significant mitotic enrichments in the more abundant mitotic histone phosphorylations. The main example of this is mitotic H3S10ph. It is conceivable that even though H3S10ph is broadly enriched genome-wide, there could be more subtle but still significant true enrichment peaks at specific loci. However, chromatin immunoprecipitation (ChIP), arguably the gold-standard approach to histone modification mapping of the last decade, struggles to distinguish true but subtle enrichments compared to Input signal. In ChIP, an Input sample is generated containing the full genome fragmented. Ideally, this Input sequencing should produce a perfectly uniform read count across the genome. However, in real experiments, technical variation such as dimer artefacts or duplicates in repetitive regions, or biological variation such as increased fragmentation in more open chromatin regions, cause variation in read counts. The input is therefore sequenced in order to theoretically represent and control for these biological and technical read count variations. However, Input sequencing from traditional ChIP-seq often produces quite high, "noisy" sequencing. For ChIP-seq of abundant, broadly spread H3S10ph, it is therefore extremely difficult to distinguish true, significant H3S10ph enrichment from input biological or technical read count "noise".

To exemplify this problem, in **Figure 1.5B**, we process H3S10ph ChIP-seq raw sequencing data generated as part of (Javasky et al., 2018) study, and align to the mouse genome (mm9) alongside the corresponding Input sequencing. This Figure highlights the issue that it is very difficult to differentiate genuine enrichment for abundant histone phosphorylation compared to the input. These technique limitations are discussed in further detail in Section 1.3.

A

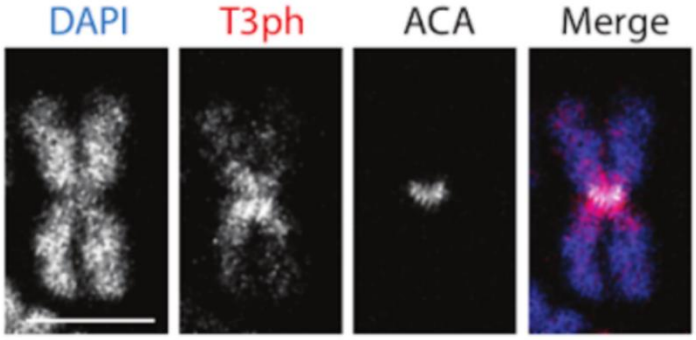
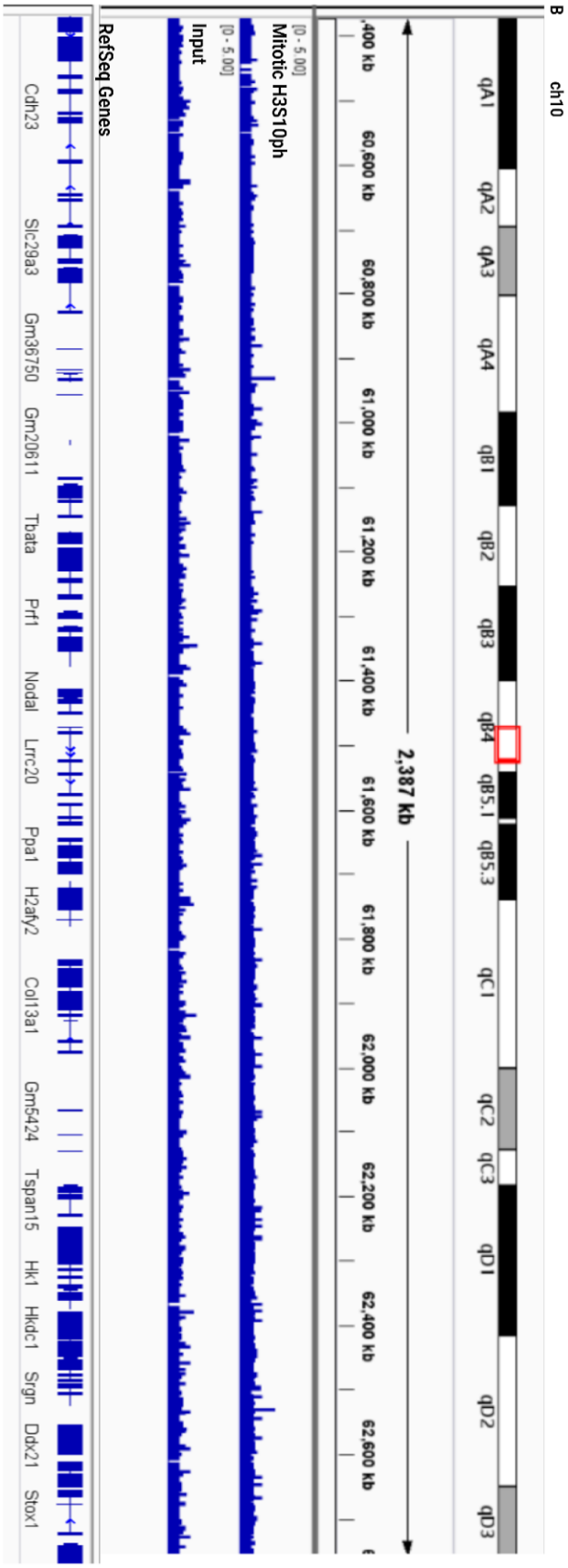


Figure 1. 5: Traditional techniques to map mitotic histone phosphorylation distribution.

A: H3T3ph immunostaining. *Mitotic chromosome spreads, performed by Rebecca Harris, immunostained against H3T3ph, DNA (DAPI) and centromeres (ACA), showing H3T3ph centromeric enrichment.*

B: H3S10ph ChIP-seq. *Javasky et al., 2018 performed ChIP-seq on H3S10ph (track above) as well as an Input sample (track below). Here, we aligned the raw sequencing generated by Javasky to the mm9 genome, and show an example section of chromosome 10. RefSeq gene annotations are shown below. Note both H3S10ph and Input alignments are shown with the same y axis scale. Track blue signal represents sequencing read count score.*



It is also important to consider the cell type utilised in any genomic mapping studies. The nature, distribution and potential functional roles of any given histone phosphorylation are likely to differ depending on the model organism and cell lines used, as described in the H3S10ph and H3S28ph sections above.

Another key challenge in mapping mitotic histone phosphorylations is isolating and optimising experimental techniques in mitotic cells. Issues can arise when trying to design experimental manipulations that only affect mitotic events and do not disrupt interphase cells. Therefore it is vital that we are able to isolate a cell population that is as purely mitotic as possible with minimal interphase contamination. Numerous techniques for isolating mitotic cells are described in the sections below.

Technique development is allowing rapid progress in identifying histone modifications and their enrichment patterns across the cell cycle; numerous techniques are described below in Section 1.3. It remains notoriously challenging to decipher the functions of individual histone modifications, as studies knocking down histone PTM writers or erasers will affect multiple co-regulated modification sites.

These challenges, and the known studies previously surrounding mitotic histone phosphorylations, allowed us to define an overarching goal of this project. We were able to describe our idealistic research goal: what would the "ideal" perfect experiment look like in order to accurately demonstrate the roles of histone phosphorylations in mitosis?

Our ideal experiment would:

- **isolate mitotic cells** with high purity while retaining the natural biology of real mitotic cells
- **Accurately map** the exact genomic loci of abundant histone phosphorylations
- Map in a **quantifiable** way, to allow us to distinguish subtle enrichments of abundant phosphorylations, and accurately control for biological and technical variation
- Be able to quantitatively **compare between different cell cycle stages**
- Integrate wider epigenetic data to place histone phosphorylation enrichments **in the context of other regulatory factors**

The following discussion and review of methodologies is therefore presented in the context of our overarching question for this project: how do we design experiments and analyses that are as close to our "ideal" perfect experiment as possible?

1.3 Methods of mapping mitotic histone phosphorylations

The first goal of our ideal experiment is to accurately map histone phosphorylations in isolated mitotic cells that retain natural mitotic biology. By identifying genomic loci, we can begin to hypothesise functional roles based on the function of enriched genes and/or regulatory elements where the phosphorylation is found. This section provides a brief overview of the key methods which have historically been used in the literature to map histone phosphorylation distribution.

1.3.1 Imaging

With the development of antibodies directly targeting and binding specific histone phosphorylations, visualisation of histone phosphorylations became possible through immunostaining. H3T3ph nicely demonstrates the contributions of imaging. Mitotic, condensed chromosomes can be isolated to be imaged via fluorescence confocal microscopy. After addition of primary antibody to the target histone modification, a secondary antibody can then be introduced to recognise and bind the primary antibody. This secondary antibody is manufactured to contain a fluorescent tag, which can then be visualised with imaging equipment. **Figure 1.5A** as discussed above demonstrates this for H3T3ph, which can be seen to enrich primarily at the centromeric region of chromosomes, with some fainter spread along the arms, in prometaphase.

Anti-phospho antibodies were first used by Hendzel and colleagues in 1997 to show H3S10ph initially is deposited in the centre of chromosomes in late G2, before strongly enriching all

along the chromosome arms by prometaphase. However, imaging techniques only allow identification of broad distribution patterns; the exact genomic loci of histone phosphorylation deposition cannot be seen through imaging. Thus, techniques have been developed over the past few decades to enable far more precise mapping of histone modifications.

1.3.2 Mass spectrometry

Among the most-utilised tools in identifying and quantifying histone modifications has been mass spectrometry (MS), an analytical technique that determines the mass/charge ratio of molecules including peptides and proteins (Noberini et al., 2022). MS provides a highly accurate quantitative approach that bypasses issues in immunoprecipitation techniques such as anti-PTM antibody availability, cross-reactivity or epitope masking(Janssen et al., 2017; Markaki et al., 2009; Önder et al., 2015; Schmitz et al., 2020; Stransky et al., 2020). Multiple PTMs can be measured in a single MS run, increasing throughput. In the most commonly-used "bottom-up" MS approach, histones are enzymatically digested into peptides roughly 5-20 amino acids long, for example using Arg-C protease (Noberini et al., 2022). In one example study, Dastidar et al. (2013) used MS in tandem with liquid chromatography to show a 14-3-3 protein binding to H3S10ph and in particular H3S28ph, in *P.falciparum*. Javasky and colleagues in 2018 used proteomics and MS to study global histone modification levels in human HeLa S3 cells, using synthetic peptide standards to reproducibly compare histone PTM levels between mitosis and interphase. They found that H3S10ph levels enriched 17-fold in mitotic cells compared to interphase. However, this only measures global histone PTM levels; mass-spectrometry requires extraction of histones, preventing genomic distribution information.

1.3.2 ChIP-seq

The most direct and commonly-used tool for identifying genomic distribution of modified histones, such as histone phosphorylations, is chromatin immunoprecipitation (ChIP). First,

the full genome is fragmented into mono-nucleosomal chromatin fragments of around 250 bp, either using sonication or digestion enzyme micrococcal nuclease (MNase). Then, an antibody specific to the target protein is used to immunoprecipitate the chromatin fragments containing the DNA-protein complex of interest (see **Figure 1.6**) (Gilmour & Lis, 1984; Small et al., 2021; Solomon et al., 1988). Microarrays can then be used to identify the DNA fragments by hybridisation and gain a genome-wide view of DNA-protein binding, in a technique known as ChIP-chip (Bernstein et al., 2002; Blat & Kleckner, 1999; Robyr et al., 2002; Robyr & Grunstein, 2003).

Development of next-generation sequencing (NGS) was then quickly applied to directly sequence ChIP DNA fragments (ChIP-seq), allowing precise genomic mapping with higher precision, greater coverage and fewer artefacts. Barski et al. performed the first large-scale profiling studies of chromatin marks in 2007, including 20 histone methylations as well as histone H2A.Z, RNA polymerase II, and CTCF in human T cells. ChIP-seq became a cornerstone for The Encyclopedia of DNA Elements (ENCODE) consortium, a collaborative, international public research project aiming to identify all functional elements of the human and mouse genomes (ENCODE Project Consortium, 2012).

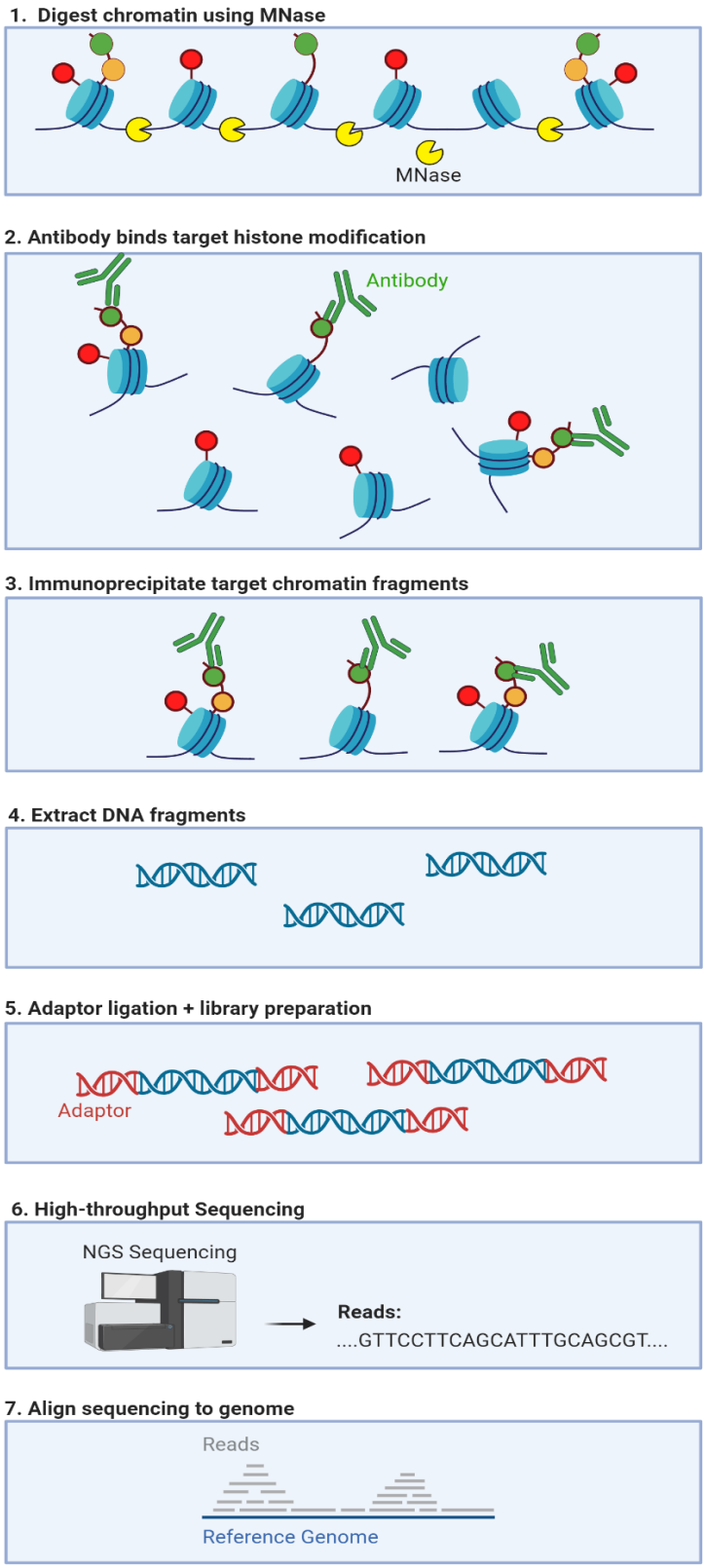


Figure 1. 6: ChIP-seq for histone modifications.

In X-ChIP, formaldehyde fixation cross-links DNA and modified histone protein Native ChIP (N-ChIP) contrastingly doesn't involve cross-linking. DNA can be fragmented using sonication or MNase digestion (top, yellow). DNA fragments bound to modified histone of interest are immunoprecipitated with an antibody specifically targeting the histone PTM of interest (green). DNA is purified from isolated fragments and can be amplified and sent for massively-parallel sequencing. Figure is an original illustration.

There are two major approaches to performing a ChIP, with different advantages and disadvantages depending on experimental aims. The first approach, pioneered by Varshavsky and colleagues, uses the fixative formaldehyde (or UV radiation, e.g. Gilmour & Lis, 1985) to covalently crosslink proximal (< 2 Å) protein-protein and protein-DNA interactions, termed crosslinked-ChIP/X-ChIP (Jackson, 1978; Solomon et al., 1988; Solomon & Varshavsky, 1985). X-ChIP is historically more common, as most nonhistone proteins (eg TFs) require crosslinking to stabilise weaker/transient DNA binding. Chromatin is then sheared, usually by sonication, although sonication produces variable fragment lengths of 200-600 bp (P. J. Park, 2009). Some studies have suggested crosslinking can promote epitope masking and generate false positive binding sites (Baranello et al., 2016; Jain et al., 2015; Meyer & Liu, 2014; D. Park et al., 2013; Teytelman et al., 2013). For example, Teytelman *et al.*, (2013) found that prolonged formaldehyde exposure time (20-60 min) enriched heterologous GFP (green fluorescent protein) at active, open chromatin regions, suggesting non-specific protein capture. The signal-to-noise ratio and false positive rate can be improved in X-ChIP by careful, precise timing of the fixation step (Baranello et al., 2016; Guéron et al., 1988). Cell fixation has historically been used to maintain the mitotic state of cells. However, formaldehyde fixation poses its own problems; many in the field now report that formaldehyde can strongly disrupt cellular components and displace dna-interacting factors such as TFs (for review see Baranello et al., 2016). These effects could potentially impact on observed histone modifications, calling into question whether histone modification sites observed represent true biological enrichment, or an artefact of fixation.

The alternative major approach, N-ChIP (or "native" ChIP), digests native chromatin from isolated cell nuclei using micrococcal nuclease (MNase) (Hebbes et al., 1988). Histone-DNA interactions, being more stable, can forgo fixation, allowing N-ChIP. By removing harsh fixation and sonication, N-ChIP is thought to reduce artefacts from non-specific crosslinking, increase signal-to-noise ratio, and improve precipitation efficiency, requiring lower starting cell culture numbers (for review see Small et al., 2021). It is worth noting that while N-ChIP is now the more commonly used approach to study modified histones, selective chromatin shearing by MNase in more open internucleosomal linker regions could have more pronounced

sequence bias compared to sonication (Park, 2009; Tolstorukov et al., 2009). Also, lack of fixation in N-ChIP means nucleosome rearrangements or histone PTM changes could occur during the experiment; this could especially cause issues in experiments investigating histone modification patterns at specific time points, such as distinct phases of mitosis.

Traditional chromatin immunoprecipitation (ChIP) involves digestion of the full genome, and as such some off-target chromatin fragments can be extracted unintentionally during immunoprecipitation. These off-target chromatin fragments then contribute "read count noise" to sequencing data, introducing artefact sequencing reads which are not true readings of the targeted histone modification's enrichment. Also, some regions of the genome may be more easily digested, for example more accessible chromatin regions may be more effectively bound and digested by MNase digestion enzyme. These regions may therefore be enriched in sequencing, creating more "read count noise". Read count noise in ChIP can be controlled for by including an Input sample, in which no immunoprecipitation is performed. The full genome is digested, and the fragments sequenced as "Input" to effectively sequence the genome background, accounting for variation in accessibility or off-target noise. Input sequencing can be subtracted from target sequencing, to in theory leave only "true" target enrichment remaining.

In ChIP experiments antibody specificity can affect background noise, reproducibility and can lead to misinterpretation of biological data (Baker, 2015; Nishikori et al., 2012; Weller, 2018). This is of particular concern when studying histone modifications, where knockdown studies testing antibody specificity are often not feasible due to either unknown or multiple histone PTM "writers/erasers". While ChIP-qPCR at known genomic loci is commonly used to validate antibodies, this relies on prior knowledge of specific loci often lacking for histone PTMs. The gold standard approach for histone PTM antibody validation is therefore testing against a library of modified histone peptides – particularly to test target recognition in presence of adjacent residue modifications (I. Bock et al., 2011; Fuchs et al., 2011; Rothbart et al., 2012). There is also an emerging approach using recombinant nucleosomes for antibody validation, thought to give more accurate predictions of antibody efficacy *in vivo* (Small et al., 2021). Antibody validation is discussed in greater detail in Chapter 2, and provides crucial context in Chapter 4.

ChIP-seq has effectively demonstrated clear patterns in multiple histone modifications and helped to elucidate functional roles for these modifications in mitotic studies. Perhaps the most well-known example is tri-methylation of lysine 4 of histone 3 (H3K4me3). When ChIP-seq is performed for H3K4me3 and the resulting sequenced reads mapped to the reference genome, H3K4me3 shows strong, sharp enrichment peaks at genomic regions associated with transcription activity – primarily, H3K4me3 peaks are seen at active promoters and active enhancer regions. A contrasting example is H3K27me3, a known repressive marker whose ChIP-seq data shows peaks at repressed genome regions across many species. ChIP-seq has produced strong evidence of functional roles for some histone phosphorylations. ChIP-seq was recently performed against H3T3ph in mitotic and asynchronous HeLa S3 cells by the Higgins lab, led by Rebecca Harris and Jonathan Higgins. These findings showed specific H3T3ph enrichment peaks in the centromeric regions, and interestingly reported an absence of H3T3ph in regions enriched for H3K4me3. This ChIP-seq data provided valuable evidence relating to the methyl-phos switch model; Harris and coworker's findings showed an anticorrelation of H3T3ph and H3K4me3 in HeLa cells, which would prevent a methyl-phos switch mechanism as that would require the two PTMs to be present together (Harris et al., 2023). This ChIP-seq study is discussed in further detail in Chapter 5.

1.3.3 CUT+RUN

In light of the potential limitations of ChIP-seq discussed above, a technique was more recently developed that could tackle ChIP-seq's issues with distinguishing true enrichment from "read count noise", and immunoprecipitation efficiency. CUT&RUN, "Cleavage Under Targets and Release Using Nuclease", was developed in 2017 by modifying the ChIC (chromatin immune-cleavage) method by (Skene & Henikoff, 2017a). Briefly, nuclei attached to magnetic beads are treated with antibody against the target protein (e.g. histone PTM), followed by a fusion protein of MNase and protein A, which specifically binds Immunoglobulin G, allowing specific DNA cleavage at the target site. After brief digestion kept on ice, the protein-DNA complex released into the supernatant is quickly recovered before MNase can diffuse along the genome and cleave at non-specific sites (see **Figure 1.7** for details). Initial studies suggest

CUT&RUN is able to extract chromatin markers even in less accessible chromatin regions, reporting reduced contamination from off-target fragments, reducing sequencing noise, and 10-fold lower required sequencing depth for quantitative chromatin mapping, likely thanks to digestion *in situ* without need for crosslinking (Skene and Henikoff, 2017).

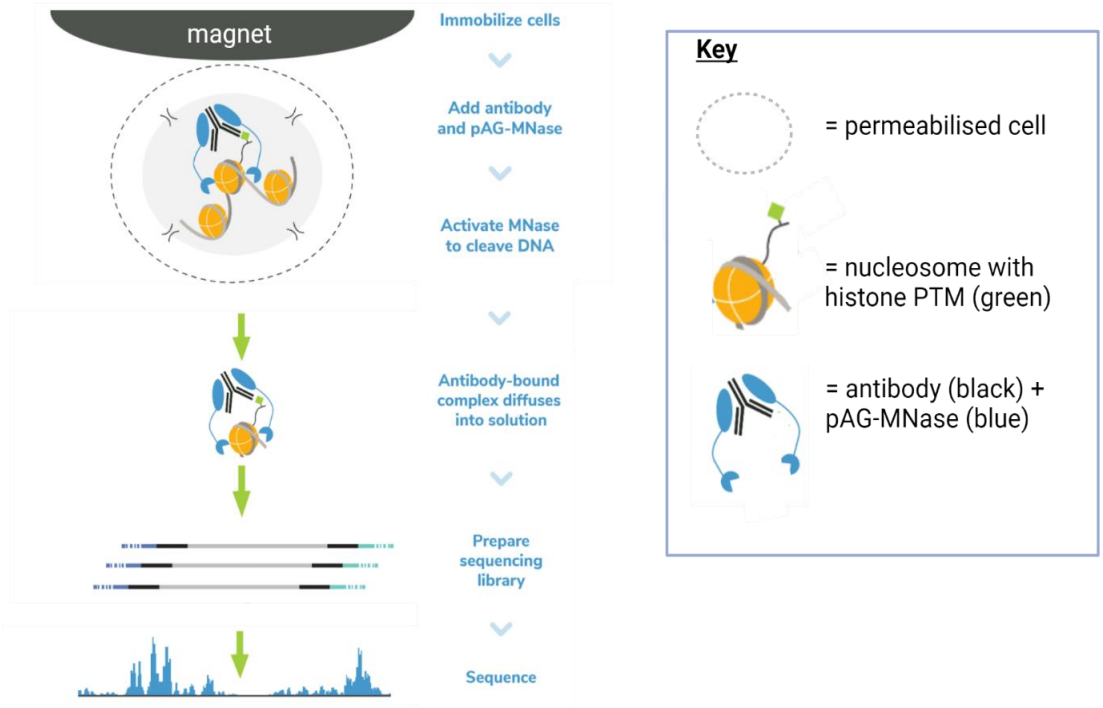


Figure 1. 7: CUT&RUN.

Cells attached to magnetic beads are permeabilised using digitonin. Cells are then treated with an antibody specific to the target histone PTM, and then Protein A-MNase (pA-MNase): pAG binds the antibody, and MNase once activated with Ca^{2+} *digests the chromatin at the target site. MN-pA-Antibody-Protein-DNA complexes diffuse out of the nuclei and cell, which is held immobilised by a magnet, and DNA is quickly extracted from supernatant before MNase-pA can diffuse along the genome and cleave at non-specific accessible sites. Extracted DNA fragments are used to prepare libraries for paired-end sequencing. Figure adapted from Epicypher.*

CUT+RUN was initially developed for use with isolated nuclei (Skene & Henikoff, 2017b), and later adapted to be performed using whole cells (Skene et al., 2018). Both methods have rarely been performed in mitotic cell samples; this could be due to difficulties in isolating pure and healthy mitotic cells, the lack of mitotic nuclear envelope preventing nuclei isolation, or simply that the technique is still relatively novel and therefore has not been optimised and standardised in a broad enough range of cell contexts yet. The optimisation of CUT+RUN for mitotic histone phosphorylation mapping is discussed in depth in Chapter 3.

1.3.4 Barcoding methods: MINUTE-ChIP-seq

Traditional ChIP-seq remains largely qualitative in terms of measuring target enrichments; partly because biological variations (e.g. varying chromatin accessibility) and technical variations (e.g. duplication artefacts) remain an issue in distinguishing true enrichment signals. Also, even if variation is perfectly controlled for, ChIP-seq does not allow the user to compare between experimental samples. It is not possible to quantitatively compare enrichment for example between different histone PTMs, or different cell cycle phase samples. A number of more quantitative approaches have therefore gained popularity - one example is barcoding-first ChIP techniques. Prior to DNA amplification, a barcode sequence is ligated to the extracted ChIPed fragments that allows pooling of barcoded samples. This greatly increases throughput, and allows *in silico* quantification of fragment counts, effectively removing technical variability by allowing precise identification of artefact contamination (Arrigoni et al., 2018; Chabbert et al., 2018; Lara-Astiaso et al., 2014; van Galen et al., 2016).

In 2016, the Bernstein lab developed a technique using barcoding to allow a more high-throughput ChIP approach (van Galen et al., 2016). Following MNase digestion of chromatin, DNA ends are repaired and ligated to a double-stranded adapter containing a T7 RNA polymerase promoter and, crucially, a sample barcode sequence. Each experimental sample is ligated to a different barcode (e.g. cell line 1 ligated to barcode 1, cell line 2 to barcode 2 etc). The samples can then be pooled and ChIP-seq experiments performed on this pool, allowing high-throughput immunoprecipitation, library amplification and sequencing while effectively removing technical variability between pooled samples. The samples can then be separated

back out *in silico* using the known barcode sequences. This method also allows post-ChIP linear amplification needing only one T7 adaptor per chromatin fragment, improving accuracy of quantification. **Figure 1.8** shows a diagram of this approach, called the MINT-ChIP technique.

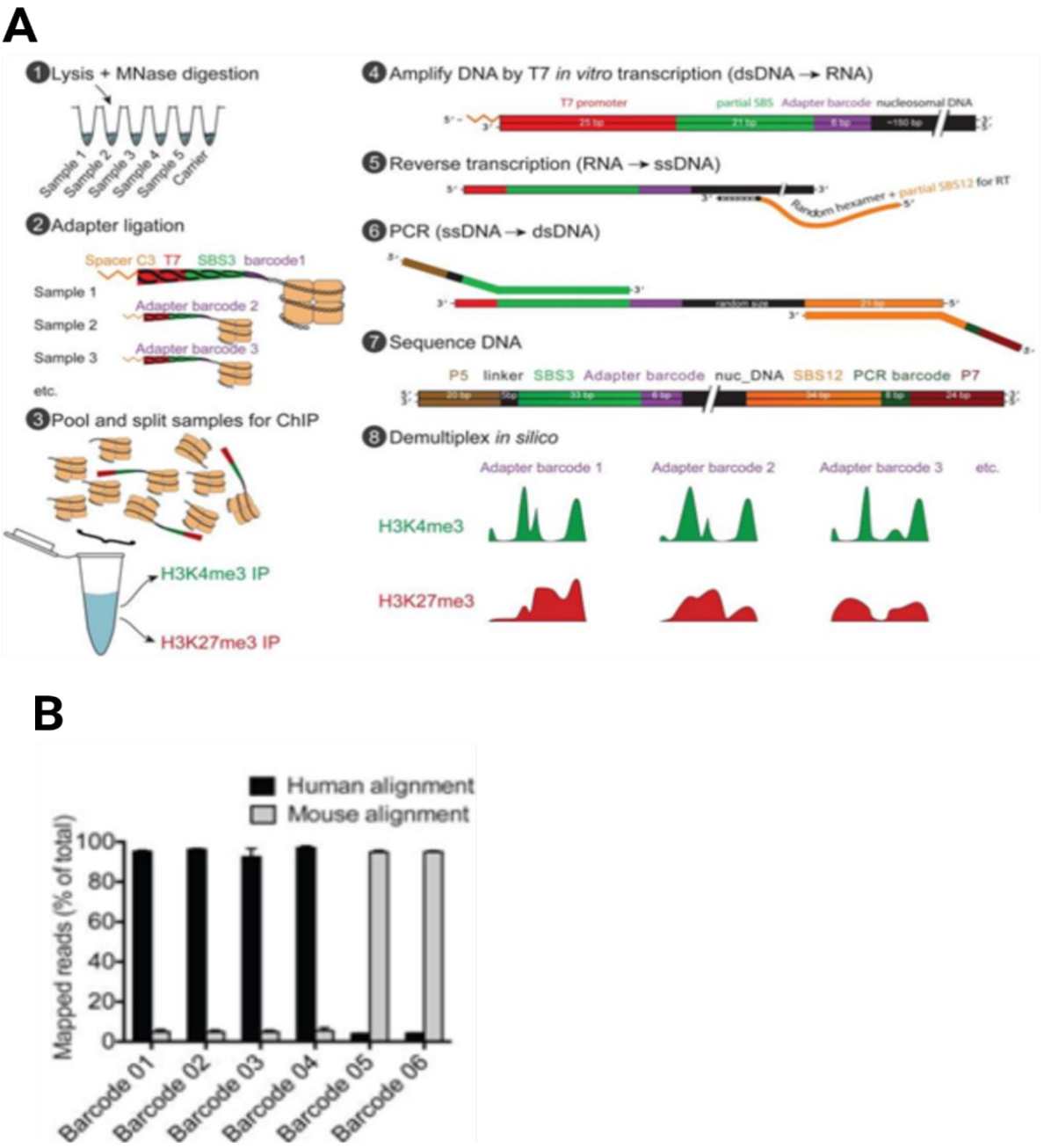


Figure 1. 8: Multiplexed, indexed T7 (MINT)-ChIP.

A) *In MINT-ChIP, samples are first lysed and chromatin fragmented by micrococcal nuclease enzyme (1). MNase activity is then stopped, and a double-stranded DNA adapter ligated to nucleosomes: this adapter contains a T7 RNA polymerase promoter (T7, red), an Illumina sequencing primer (SBS3, green), and a sample barcode sequence (purple) (2). Samples are then pooled and ChIP experiments performed on the pooled samples (3). T7 RNA polymerase*

is then used for in vitro transcription (4) to produce RNA, which is then reverse transcribed (5). Low-cycle library PCR amplification is then performed (6) to make an Illumina sequencing library (7). De-mulitplexing is then carried out in silico (8) to separate out samples based on sample barcodes. ***B)*** *Bernstein's study used four human samples (K562, T7-adapter barcode 1–4) and two mouse samples (YAC-1, T7-adapter barcode 5–6), which were then indexed/barcoded, pooled and split for three parallel Mint-ChIP assays. Plot depicts the proportions of reads for each barcode that align to the human or mouse genomes. Data is shown as mean ± SD of 3 ChIP assays. Figure adapted from van Galen et al 2016.*

The Elsasser research group then sought to further advance this quantitative approach, by creating "MINUTE-ChIP": Multiplexed indexed unique molecule T7 amplification end-to-end sequencing (Kumar and Elsasser, 2019). MINUTE-ChIP ligates to fragmented DNA ends an adapter, containing the same SBS3 Illumina sequencing primer, a T7 RNA polymerase adapter for linear amplification, and sample barcode as seen in MINT-ChIP. Novel to MINUTE-ChIP, between the SBS3 and barcode at the 3' end, a 6-base pair (bp) randomised sequence is introduced as a "unique molecular identifier" (UMI). This UMI is reported to contribute multiple advantages:

- Firstly, UMI sequence information greatly increases the confidence of calling amplification duplicates, and improves quantitative representation of repetitive sequencing (Kumar & Elsässer, 2019a). This deduplication improves the viability of the Input sample, where dimers and duplicates have often been seen to create artefacts. By creating a viable Input, MINUTE-ChIP experimental samples' sequencing can then be normalised against the Input to more accurately account for background sequencing "noise". MINUTE-ChIP is therefore able to provide a sequencing read count closer to the true biological enrichment of the target histone modification. By improving read count more accurate to true enrichment, the MINUTE-ChIP sequencing data can be used quantitatively.

- Secondly, Elsasser's research sought to calibrate MINUTE-ChIP sequencing to allow quantitative comparison of multiple ChIP-ed epigenetic features. H3K27me3 is a repressive histone modification known to be enriched at repressive polycomb regions of the genome throughout the cell cycle. Elsasser employed Ezh2-inhibitor (EPZ-6438) to deplete H3K27me3, or JMJD3/UTX inhibitor (GSK-J4) to cause very high H3K27me3. This generated H3K27me3-high and –low cell populations. Cell mixing was then used to create a % scale from 0% to 100% H3K27me3-enriched samples. These H3K27me3-scaled samples were then MINUTE-ChIPed and used to confirm that MINUTE-ChIP sequencing data can indeed provide a biologically accurate quantitative measure of H3K27me3 enrichment. **Figure 1.9** explains this calibration experiment in further detail. These experiments nicely demonstrated the ability of MINUTE-ChIP to produce biologically accurate measures of histone modification enrichments that can be directly compared between different cell samples. This ability to quantitatively analyse histone modification enrichment comparing samples provides an exciting opportunity to analyse dynamics of histone modifications in different cell cycle stages.

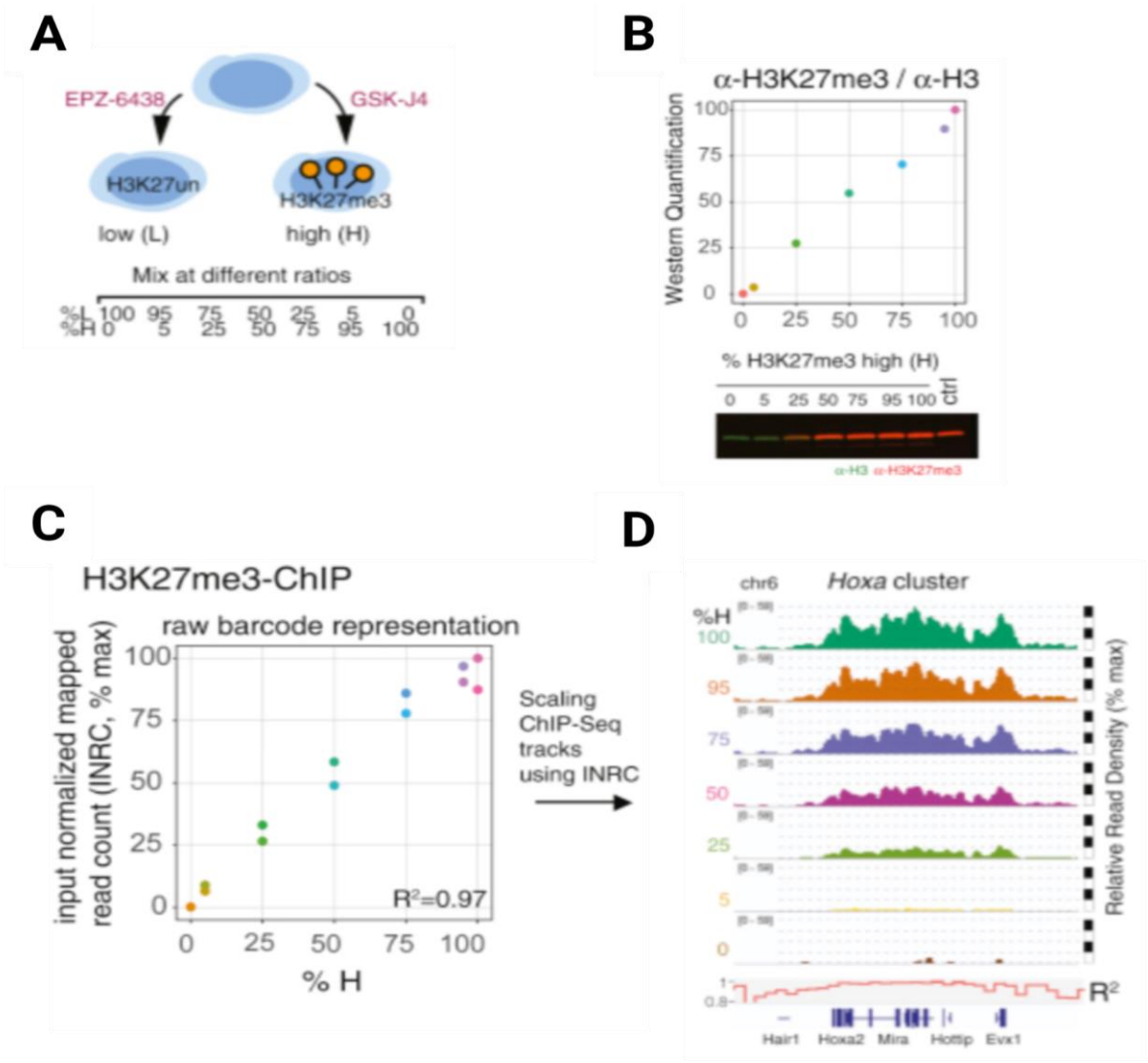


Figure 1. 9: Calibrating MINUTE-ChIP with H3K27me3 gradient.

A) Setting up calibration. *mESCs were treated with Ezh2 inhibitor EPZ-6438 to create a low H3K27me3 population (L), or with JMJD3/UTX inhibitor GSK-J4 creating a high H3K27me3 population (H). L and H populations were mixed at different ratios to create a gradient from 0% to 100% H3K27me3.* ***B) Quantitative infrared fluorescent western blot*** *confirmed the H3K27me3 gradient. Control sample (Ctrl, bottom right) = untreated mESC sample. Raw western blot H3K27me3/H3 signal was arbitrarily scaled from 0 to 100.* ***C) Input-normalised ChIP-seq readcount accurately replicated biological H3K27me3.*** *Input-normalised read count (INRC) of two technical replicates were plotted against the mixing ratio of each sample (H). R^2 = coefficient of determination for a linear regression.* ***D) Example INRC-scaled***

genomic tracks for each mixing ratio sample. *Hox cluster is shown as an example genomic region. ChIP-seq genomic tracks were scaled using Input-Normalised Read Count (INRC). Figure adapted from (Kumar & Elsässer, 2019b).*

While MINUTE-ChIP presents many potential advantages in advancing our knowledge of histone modification behaviours, it is still important to consider antibody efficacy. MINUTE-ChIP has been demonstrated to enable quantitative measurements of mark enrichment, and allow quantitative comparison between samples in a given pool. However, to date this method does not contain a solution to control for differences in efficacy of antibodies. Antibodies for different histone modification targets may well possess different binding affinities, different optimal experimental concentrations, different specificities and different possible inhibitory or cross-reacting adjacent histone modifications. The importance of assessing antibody efficacies is discussed in depth in Chapter 4. Therefore if two different histone modifications are MINUTE-ChIPed, it cannot be concluded whether any quantitative difference in enrichment measurements is due to a genuine biological difference, or due to difference in two antibodies' efficacies. We caveat that antibody efficacy must always be considered when interpreting enrichments, as it will still have impacted immunoprecipitation of the respective targets.

1.3.5 Generating mitotic data

Many of the laboratory techniques described above have rarely been applied in a mitotic cell context. For this project, we are particularly interested in understanding functional histone phosphorylation enrichments specifically in mitotic cells. In order to meet this objective, this project must first isolate pure mitotic samples that ideally retain natural biology of mitotic cells. There are multiple approaches which could be utilised to do this, described below.

Cell synchronisations

Synchronisation methods can be used to accumulate cells in a particular cell cycle phase, and have the advantage of potentially generating larger cell samples. Chemical synchronisation uses drug treatments to block and accumulate cells in a cell cycle phase; the gold standard drug used for mitotic blockage is nocodazole. Nocodazole acts by disrupting microtubule polymerisation, preventing separation of sister chromatids to opposite poles, blocking the cells before anaphase. The toxicity of nocodazole should be considered and varies depending on cell line; some will exit mitosis even in nocodazole. Some cell lines can be incubated in nocodazole for extended periods, but in other cell types nocodazole can quickly cause cell death. In most cases, arguably cell cycle arrest will disrupt the normal natural state of cells and could be cytotoxic to some degree (for review see Eastman & Guo, 2020). Alternatively, cells can be isolated in mitosis without chemicals using mitotic shake off. In adherent cell populations, mitotic cells will detach from the surface and can be collected through gentle shaking of the cell culture flask. However, this method generates a smaller yield of mitotic cells.

Cell cycle phase labelling and Cell Sorting

Instead of chemical synchronisation, cells can be labelled for cell cycle-phase specific markers, and sorted into cell cycle phases using fluorescence-assisted cell sorting: FACS. Cellular proteins, enzymes or factors that are enriched specifically in mitotic cells can be labelled and used to identify and isolate mitotic cells. Examples of mitotic probes include: phosphorylation histone H3, cyclin B1, and the most common, mitotic protein monoclonal #2 (MPM2). MPM2, for example, is fluorescently tagged, and flow cytometry performed. MPM2-enriched cells can then be isolated (Campbell & Turner, 2013). FACS for mitotic cell isolation is described in further detail in Chapters 2 and 4.

1.3.6 Bioinformatics: processing high-throughput sequencing data from histone phosphorylation mapping

Chromatin immunoprecipitation techniques, including CUT&RUN and MINUTE-ChIP, extract DNA from the site of the histone modification of interest. Massively parallel high-throughput sequencing then produces very large genomic data sets, containing 10s of millions of sequencing reads of roughly 35-150 bp length. A growing field of bioinformatics techniques and tools has developed in parallel to process these large epigenomic datasets.

The sequenced reads can be filtered to remove low-quality sequence reads, depending on sequencing depth. Reads are then mapped to the reference genome, commonly using bowtie, bowtie2 or BWA alignment tools. (Langmead et al., 2009; Langmead & Salzberg, 2012; Li & Durbin, 2009; Nakato & Sakata, 2021). Aligned reads can then be filtered to remove blacklisted regions (Amemiya et al., 2019), reduce redundancy and/or multi-alignment reads, although studies investigating highly repetitive regions, such as the centromeres, may keep multi-alignments to retain valuable sequencing data in repetitive regions that may represent true enrichment. Peak calling, often with the MACS2 tool, then identifies loci significantly enriched for the ChIP target (e.g. histone modification), using a corresponding input control sample to account for read count variation/noise, GC bias and chromatin structure (Nakato & Shirahige, 2017; Y. Zhang et al., 2008). Numerous other peak calling tools exist utilising different algorithms to model background noise (for review see Thomas et al., 2017). While no peak calling tool can identify true peaks with 100% accuracy, additional tools can be used to filter out false positives/noise, for example by peak calling with more strict thresholds and extracting peaks consistent across biological replicates, for example using the Irreproducible Discovery Rate (IDR) (Landt et al., 2012). Called peaks, representing likely histone modification sites, can then be integrated with binding motif, gene ontology, and other epigenomic data such as gene expression and chromatin conformation assays, to identify functional annotations associated with the histone modification of interest. See **Figure 1.10** for example downstream functional analysis workflows. In the context of this project, it is worth noting that peak calling is difficult for abundant histone phosphorylations, because their mitotic high enrichment is not likely to be seen in classic narrow (or even broad) peaks that peak calling tools are programmed to identify.

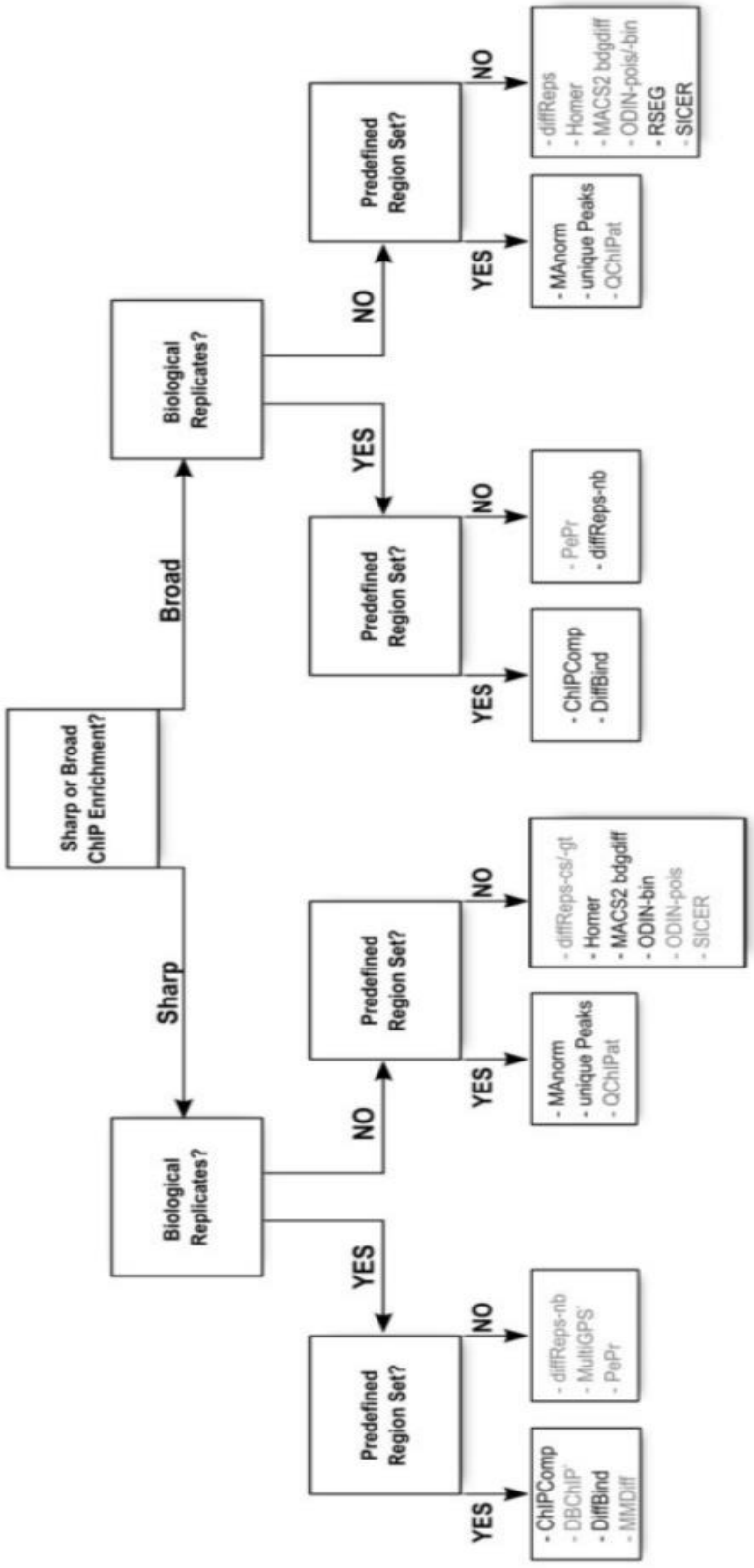


Figure 1. 10: Differential enrichment analysis decision tree.

Choice of tool depends on data set, signal shape, replicates, existence of predefined genomic regions of interest. Highlighted are tools which give good results with default parameters. In grey are tools requiring parameter tuning to achieve optimal results. Figure from (Steinhauser et al., 2016).

There are various factors affecting the quality of obtained sequence reads, including antibody quality, sequence depth, input sample quality, biological or technical variation affecting read counts, and read length. As such, despite efforts to streamline processing, there is no single workflow for ChIP-seq analysis that is optimal under all circumstances. Many experimental considerations are relevant to choice of tools, as well as the quality controls and checks which should be implemented.

1.4 Methods of analysing wider epigenetic regulatory landscape

1.4.1 Integrative approach to histone phosphorylation functional analysis

Laboratory techniques such as chromatin immunoprecipitation and CUT+RUN provide viable techniques for precise mapping of histone phosphorylations in mitosis. However, these wet-laboratory techniques alone still limit our understanding of the roles histone phosphorylations might be playing in chromatin regulation, and do not account for the possibility that histone phosphorylation function might depend on other interacting histone modifications and/or regulatory factors. Here, we argue that integration of wider epigenetic datasets will provide crucial contexts to hypothesise functional roles of mitotic H3S10ph and other histone phosphorylations with undiscovered functions. ChIP-seq genomic loci data can be co-analysed with other epigenomic datasets to paint a broader picture of chromatin interactions and mitotic regulation in relation to histone phosphorylation.

The benefits of integrative analysis can be seen for H3T3ph. H3T3ph shows centromeric enrichment in prometaphase of mitosis, and was found to be starkly absent at loci containing H3K4me3 (Dai & Higgins, 2005a; Harris et al., 2023). This anti-co-localisation of H3K4me3 and H3T3ph seriously calls into question the methyl-phos switch model for this residue, as discussed in earlier sections. Nonetheless, the functional role of H3T3ph enrichment, or absence, is not clear until integrated with wider epigenetic datasets. Studies depleting the H3T3ph kinase haspin by RNAinterference found disruption of CPC centromeric localisation, suggesting an interaction between H3T3 phosphorylation and the CPC. *In vitro* studies

including nuclear magnetic resonance spectroscopy then showed that H3T3ph specifically binds the BIR domain of the survivin subunit of the CPC (Kelly et al., 2010; Wang et al., 2010; Yamagishi et al., 2010). Together, this allowed researchers to elucidate the functional mechanism by which centromeric H3T3 phosphorylation by haspin recruits the CPC to centromeres via interactions between H3T3ph and the BIR domain of Survivin.

The field of epigenetic computational analysis is rapidly expanding and advancing, with a wide range of analytical techniques able to paint a detailed picture of the chromatin regulatory landscape. A few key techniques are detailed below.

1.4.2 Gene functional enrichment analyses

Genomic loci at which histone modifications of interest have been mapped can be interrogated against databases of known gene functions**.** By integrating gene set functional databases with histone modification enrichment, software can now analyse whether histone modifications correlate significantly with any particular cellular processes, pathways or regulatory networks.

For example, Over-Representation Analysis (ORA) software allows comparison of a defined set of genes – for example genes showing high histone phosphorylation – against the reference genome to identify any statistically significant functional enrichments (for review see Wieder et al., 2021). ORA software will then output any gene sets or gene families showing significant correlation with the high-phosphorylation gene list. This offers a statistical method of assessing and identifying possible functions correlating with histone phosphorylation.

An advance on ORA is Gene Set Enrichment Analysis (GSEA) (Mootha et al., 2003; Subramanian et al., 2005). Instead of a small gene list, GSEA takes all genes and ranks them by enrichment of the target, e.g. histone phosphorylation signal. This allows a more comprehensive and robust statistical analysis of histone phosphorylation functional enrichments, assessing both negative and positive correlations. GSEA and ORA are explored further in Chapter 4.

1.4.3 RNA-seq gene expression analyses

One method of introducing chromatin regulatory context to histone phosphorylation is to integrate gene expression data. Although transcription shows significant reduction during mitosis, low levels of transcription do remain (for review see Ito & Zaret, 2022). Transcription is then rapidly reactivated in a precisely regulated manner on exit from mitosis to begin G1 phase. How transcription is regulated and reactivated during and exiting mitosis is extensively studied with continually developing techniques (e.g. Kang et al., 2020; Schmitz et al., 2020).

The gold-standard approach to gene expression analysis is RNA-seq. Typically RNA-seq experiments consist of isolating RNA from the cells or tissues, converting it to complementary DNA (cDNA) by reverse transcription, preparing a library from the cDNA and performing Next-Generation Sequencing (for review, see Kukurba & Montgomery, 2015). Importantly, RNA-seq can be adapted depending on the specific type of RNA desired. Historically, RNA-seq and transcriptomics focussed on targeting messenger RNA (mRNA), generated by transcription of protein-coding genes. However, other RNA species exist, including ribosomal RNA (rRNA), and non-coding RNA (ncRNA). Example methods of subsetting RNA molecules include enriching for polyadenylated (polA) transcripts, or ribo-depletion to remove rRNA (Kukurba & Montgomery, 2015). Importantly, traditional RNA-seq does not account for the time at which isolated transcripts were generated; this limits its ability to measure gene expression over time. Techniques have therefore been developed to isolate only nascent RNA, enabling quantitative tracking of nascent gene expression (Wissink et al., 2019).

As one example of nascent RNA-seq, EU-RNA-seq was recently developed, utilising pulse-labelling of nascent RNA transcripts to measure mitotic and post-mitotic transcription activation temporally (Palozola et al., 2017, 2019). In one study, EU-RNA-seq data was examined comparatively with histone modification data exiting mitosis, finding some correlation between Histone H3 Lysine27 acetylation (H3K27ac) at enhancers and gene reactivation (Kang et al., 2020). Nascent RNAseq presents an interesting approach which could be effective in observing transcriptional activation during mitotic exit. However, functional mechanisms linking histone modifications to gene expression regulation during and exiting mitosis are not yet well understood.

1.4.4 Long-range chromatin interaction analyses

As another example of chromatin epigenetic analysis techniques, Chromatin conformation capture carbon copy (5C) combined with high-throughput sequencing (Hi-C) uses ligation to identify long-range chromatin interactions to reveal higher order 3D chromatin structure (see **Figure 1.11**) (Dekker et al., 2002; Dostie et al., 2006; Lieberman-Aiden et al., 2009). Chromatin interaction analysis with paired-end tag sequencing (ChIA-PET, see **Figure 1.11**) introduces a ChIP step to specifically identify chromatin interactions with a particular protein of interest, exemplifying co-analysis of histone modification and chromatin structural data.

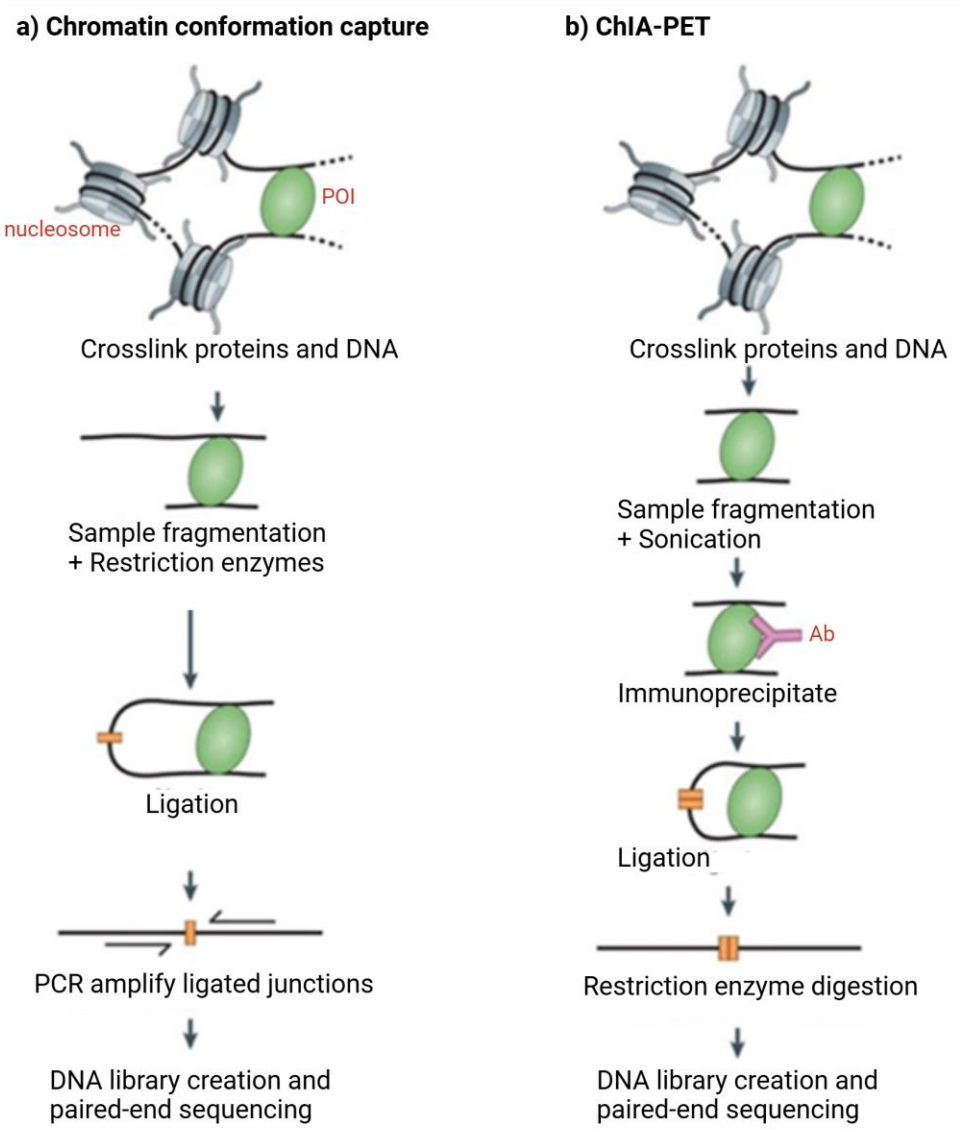


Figure 1. 11: Chromatin structural and interaction analyses

(a) Chromatin conformation capture (3C) *uses crosslinking and ligation to capture interactions between chromatin fragments and proteins of interest (POI, green). Ligation is used to join chromatin fragments that interact in 3D chromatin space.* ***(b) Chromatin interaction analysis with paired-end tag sequencing (ChIA-PET)*** *identifies interacting chromatin fragments at sites of binding of the protein of interest (POI, green) using POI-targeting antibodies (Ab, pink). Figure adapted from (Furey, 2012).*

1.4.5 Co-localisation analysis

Individual chromatin immunoprecipitation experiments focus on single histone modifications or epigenetic marks. One relatively simple way to greatly advance understanding of functional roles, is to assess the interactions of a histone modification of interest with other epigenetic markers. By analysing co-localisation of histone phosphorylations with other histone modifications with known functional and regulatory roles, we can generate an informative interaction network for histone phosphorylations. This could lead to improved hypotheses of likely functional or regulatory pathways that histone phosphorylations might contribute to. Co-localisation analyses use computationally intensive, statistics based methods and algorithms to assess the co-enrichment of two or more genomic features, be it histone modifications, transcription factors or other elements (Dozmorov, 2017; for review see Lawrence et al., 2013). A wide range of co-localisation tools are available which can: compare regions against public data (e.g. Bock et al., 2009; Halachev et al., 2012; Sheffield & Bock, 2016); visualise intersecting genomic regions of enrichment (e.g. Conway et al., 2017); and perform statistical tests for colocalisation between two or more tracks (e.g. Favorov et al., 2012; Layer et al., 2018). Co-localisation analysis provides a simple but effective and informative integrative approach to broaden our understanding of the regulatory network histone phosphorylations are associated with.

1.4.6 ChromHMM chromatin state analyses

The methods described above often study individual marks, such as single histone modifications, either looking at their genome-wide signal at predetermined annotations (transcription start sites), or by binarisation methods marking signal enrichment relative to other regions (e.g. MACS2). However, as epigenetic research has progressed, bioinformatics tools have been developed which use binarised feature enrichment data to study combinations of multiple epigenetic marks in their spatial context, to create patterns labelled as “epigenetic states” or “chromatin states”. Chromatin states often capture known regulatory elements, with ENCODE using epigenomic datasets across 127 cell types and tissues to accurately identify promoters and enhancers (Zacher et al., 2017). A range of tools exist for chromatin state modelling, such as Segway (Hoffman et al., 2012), chromstaR (Taudt et al., 2016) and

Spectacle (Song & Chen, 2015). The most common tool is arguably ChromHMM (Ernst & Kellis, 2012). Using the multivariate Hidden Markov Model (HMM), ChromHMM is trained with multiple datasets, such as ChIP-seq, of different epigenetic features with documented regulatory functions; typically six core regulatory histone modifications H3K4me1, H3K4me3, H3K6me3, H3K9me3, H3K27me3 and H3K27ac are used to train the model alongside corresponding input/control (Arcila-Galvis, 2024; Juan et al., 2016). ChromHMM generates a trained model of chromatin states, capturing the combinatorial interactions between these chromatin marks. The model is configured during training changing parameters such as the number of states, according to ChromHMM developer guidelines. The trained chromatin state model can then be used as a powerful tool for the integrative analysis of epigenomic data. Some examples of published ChromHMM models include 18- and 15-state models by the Roadmap Consortium (Roadmap Epigenomics Consortium et al., 2015) and the 11-state model by Carrillo-de-Santa-Pau et al., 2017. Chromatin state modelling could be a powerful method to elucidate chromatin regulation mechanisms across different cell types and potentially at different cell cycle stages. For example, ChromHMM could theoretically be used to combine mitotic histone modification data with EU-RNA-seq data to provide insight into post-mitotic transcriptional reactivation. Currently, research is lacking taking integrative approaches to analysing mitotic epigenomic data, largely due to technical challenges of isolating purely mitotic cells to extract epigenetic data from.

1.5 Aims of my project

1.5.1 General aim

The main aim of this project was to use integration of histone phosphorylation mapping techniques with comprehensive chromatin regulatory state data to understand the distribution and functional contributions of histone phosphorylation during mitosis.

With this overarching aim in mind, this project addresses multiple specific aims, detailed below:

1.5.2 Specific aims and objectives

1. **Our first specific aim was to accurately and precisely map mitotically abundant histone phosphorylations H3S10ph and H3S28ph**, and elucidate whether any subtle but significant enrichments are present during mitosis. H3S10ph and H3S28ph show broad, genome-wide strong mitotic enrichment. Technical limitations of traditional immunoprecipitation, such as sequencing noise and lack of quantification, have previously prevented investigation of any subtle enrichment patterns. We addressed this aim using two approaches:

    a) In **Chapter 3**, we describe optimisation of CUT+RUN methods in synchronised mitotic HeLa S3 cells. The objective was to utilise the reduced signal:noise ratio reported for CUT+RUN techniques to improve identification of more subtle H3S10ph enrichment peaks and distribution patterns.
    b) In **Chapter 4,** we employed data generated using MINUTE-ChIP-seq in collaboration with the Elsasser research group: this barcoding technique generated quantitative enrichment data for H3S10ph and H3S28ph in FACS-sorted mitotic mESCs. Through statistical, quantitative analysis we identified significant enrichment peaks in both H3S10ph and H3S28ph that occur specifically at a subset of promoters.

2. We then wanted to investigate the possible regulatory roles of these mitotic promoter peaks. **Our second specific aim was to integrate quantitative histone phosphorylation mapping with publicly available epigenetic datasets to assess mitotic phosphorylation in the context of chromatin regulation.** To address this, in **Chapter 4** we employed a comprehensive chromatin state model of 20 regulatory chromatin states with defined epigenetic characteristics, to map the mm9 genome by chromatin state regions**.** In a novel approach, mitotic histone phosphorylation was then mapped to chromatin state regions, allowing H3S10ph and H3S28ph enrichment to be

assessed in the context of the chromatin regulatory state, revealing significant histone phosphorylation peaks at promoters occurred in specific chromatin regulatory states.

3. **Our third specific aim was to identify potential functional roles of identified histone phosphorylation peaks at promoters in particular chromatin states.** In **Chapter 4** we identified genes with promoters showing significant phosphorylation peaks in chromatin states of interest. We then interrogated those isolated genes to perform Gene Set Enrichment Analysis, revealing functional enrichments for promoters with phosphorylation peaks. Additionally, partial correlation networks were generated from mitotic MINUTE-ChIPseq data to study co-localisation of H3S10ph and H3S28ph with a range of 16 epigenetic markers, many with known regulatory roles. Combined with the chromatin state regulatory context, these findings allowed us to hypothesise functional roles each histone phosphorylation might be playing in mitosis.

4. **Our fourth specific aim was to employ developed integrative techniques to analyse enrichment of another phosphorylation, H3T3ph, at promoters in a specific regulatory context.** H3T3ph shows overall depletion at promoters genome-wide, but we observe a small enrichment, suggesting a subset of promoters may contain H3T3ph peaks. In **Chapter 5**, we categorise promoters based on presence or absence of H3K4me3, an active chromatin mark adjacent to H3T3ph. This allowed us to identify H3T3ph enrichment peaks at promoters in the absence of H3K4me3, which we then expanded on to assess gene functional enrichments at these T3ph-enriched promoters.

The research discussed throughout this chapter identifies numerous demonstrable roles of histone phosphorylations in regulation of chromatin during mitosis. Correct regulation and maintained stability of chromatin during mitosis is vital; errors in these processes can lead to

numerous issues in daughter cells, including aneuploidy, DNA damage or erroneous gene expression. Thus, the regulatory functions of mitotic histone phosphorylations have far-reaching implications in numerous research fields, including cancer and developmental biology research where the molecular mechanisms underpinning erroneous cell division and chromatin stability are essential to improvement of our understanding and targeting/ treatment of disease. The research performed throughout this PhD project aims to improve our knowledge of how histone phosphorylations H3S10ph, H3S28ph and H3T3ph might function to regulate chromatin in multiple eukaryotic cell lines. Going forward, the implications of our findings could be explored across numerous research fields; for example, could these histone phosphorylations, or their perturbation, impact carcinogenesis?

Chapter 2: Materials and Methods

2.1 HeLa S3 Cell Culture

HeLa Kyoto, S3 (ATCC CCL-202) were cultured in high glucose Dulbecco's modified Eagle Medium (DMEM) (Sigma-Aldrich), supplemented with 10% foetal bovine serum (FBS), 100 µg/mL penicillin, 100 µg/mL streptomycin and 2 mM L-glutamine. Cells were grown in a humidified incubator at 37^0C, 5% CO_2, and were passaged every 2-3 days by washing in Dulbecco's phosphate buffered saline (DPBS, Sigma-Aldrich), incubating with TrypLE express (Gibco) for detachment, and then diluting with media to roughly 10% confluency in T75 flasks. Cell cultures were routinely tested for mycoplasma using Mycoplasma PCR Detection Kit (ThermoFisher Scientific) according to manufacturer guidelines.

Cell counting was performed where appropriate by adding detached cells at appropriate concentration to a haemocytometer (ThermoFisher Scientific). Manual counting of cells under a confocal microscope at 10 X objective allowed cells per mL to be calculated.

2.1.1 HeLa S3 Cell culture treatments: mitotic synchronisation

To synchronise HeLa S3 cell line in mitosis, cells were seeded at roughly 25% confluency, or 2.5 million cells per T75 flask and incubated until the following day. After removing any mitotic cells by shake-off, 2.5 mM thymidine was added and cells incubated for 18 hr under normal growth conditions detailed above. To release the cells, thymidine was removed by 3x washes in PBS followed by 3x washes in normal growth media, before incubating in fresh normal growth media for 8 hr. This thymidine 18 hr block and 8 hr release was repeated a second time. 0.5 µM nocodazole was then added and cells incubated for 5 hr. Mitotic cells were then collected by gentle shake-off.

2.2 Antibodies

Table 2. 1: List of antibodies.

Abbreviations: Mono = monoclonal; Poly = polyclonal; Rb = Rabbit; N/A = Not Applicable

Antibody, host, clone	**Source, catalogue #**	**ELISA dilution**	**CUT+RUN dilution**
CENP-A, Rb Mono, C51A7	Cell Signalling 2048S	-	-
H2A.Z, Rb Poly	Active Motif 39113	-	-
H2AK119Ub, Rb Mono	Cell Signalling #8240	-	-
H3, Rb Poly	Abcam Ab1791	-	-
H3K27ac, Rb Poly	Abcam ab4729	-	-
H3K27me3, Rb Poly	Millipore 07-449	-	-

H3K27me3, Rb Poly	Diagenode pAb-195-050	-	-
H3K4me1, Rb Mono, D1A9	Cell Signalling #5326S	-	-
H3K4me3, Rb Mono	Millipore 04-745	-	-
H3K4me3, Rb Mono, C42D8	Cell Signalling #9751	1:100,000	1:50
H3K9me3, Rb Poly	Abcam Ab8898	-	-
H3S10ph, Mouse Mono IgG1K, 3H10	Millipore 05-806	1:50,000	1:100
H3S10ph, Mouse Mono, RR002	Millipore 05-598	1:100,000	-
H3S10ph, Mouse Mono, IgG1 6G3	Cell Signalling #9706	1:10,000	-

H3S10ph, Rabbit Poly, R1 3.4	F Wang, Zhejiang A007024	1:5,000	-
H3S10ph, Rabbit Poly, R2 3.6	F Wang, Zhejiang A007712	1:5,000	-
H3S28ph, Rat Mono, 5D10	Millipore MABE941	1:2000	-
H3S28ph, Rat Mono, HTA28	Millipore MABE76	1:1000	-
H4K20me1, Rb Poly	Abcam Ab9051	-	-
IgG, Rb Poly	Diagenode #C15410206	-	1:200

2.3 Peptide ELISA

To investigate the binding properties of histone phosphorylation-targeting antibodies, enzyme-linked immunosorbent assays (ELISAs) were carried out using peptides composed of residues 1-21 of histone H3 with various modifications: peptides used are detailed in Appendix, Table A1. In ELISA, peptides (the antigens) are stuck to a streptavidin-coated plate (one peptide species per well). The primary antibody of interest is then added to each well and will bind any target peptide present. Any unbound or excess primary antibody is washed off, and then a

secondary antibody is added which is conjugated to Horseradish-peroxidase (HRP) enzyme. HRP-conjugated secondary antibody recognises and binds any primary antibody present. Addition of substrate in the presence of hydrogen peroxide then causes a reaction catalysed by HRP that produces a coloured substrate, causing a colour-change. The absorbance of this colour-change can be measured in a plate reader, acting as a direct measure of the amount of primary antibody that remained bound to the peptide. **Figure 2.1** illustrates the ELISA mechanism.

Figure 2. 1: Enzyme-linked immunosorbent assays. *Manufactured peptides representing the target protein(s) of interest (Target protein, blue) are bound to streptavidin-coated plate as antigens. Primary antibodies (orange) targeting the antigen are added which bind any target protein present on the plate. Unbound primary antibody is washed off, and then secondary antibody (green) is added which binds any bound primary antibody. Secondary antibody is conjugated to the HRP enzyme, which then reacts when hydrogen peroxide is added producing a coloured substrate (here shown as yellow). Spectrophotometer can then be used to measure the absorbance of this colour change as a linear measure of the amount of primary antibody-bound target antigen. Figure adapted from MBL Life Science.*

Prior to ELISAs, all peptides were first quantified with HABA-Avidin reagent according to manufacture guidelines (Sigma #H2153). High-Capacity Streptavidin-coated 96-well plates (Pierce) were washed X3 with TBS, 0.5% Tween20 (TBST). Then, 100 µL of peptide at 0.1 µM concentration in TBST was added and plates incubated at room temperature for 1-2 hr. Plates were washed with Dulbecco's phosphate buffered saline (PBS) before adding 100 µL

primary antibody in TBST with 0.1% bovine serum albumin (BSA) and incubated for a further 1-2 hr at room temperature. After washing with PBS, a 100 µL of HRP-conjugated secondary antibody was added at 1 in 2000 dilution (1/2000), incubated for 1 hr at room temperature. After washing with PBS again, signal was developed by addition of 100 µL TMB substrate (Pierce) and stopped using 50 µL of 2 M sulfuric acid once sufficient signal was detected. Using a Polarstar Omega microplate reader (BMG Labtech), absorbance (optical density) readings were taken at 450 nm to measure signal, and at 570 nm to measure background to subtract. All ELISAs were performed in duplicate, and each experiment performed independently three times. Peptide details including amino acid sequences, sources, and modifications can be found in the Appendix, Table A1.

2.4 CUT+RUN

2.4.1 CUT+RUN: Isolated nuclei chromatin digestion

Cleavage **u**nder **t**argets and **r**elease **u**sing **n**uclease (CUT+RUN) was performed according to the protocol for isolated nuclei published by (Skene & Henikoff, 2017a), with minor changes for mitotic HeLa S3 cells as seen in (Oomen et al., 2019a). For every CUT+RUN experiment, fresh asynchronous and mitotic synchronised (see Section 2.1.1) HeLa S3 cells were harvested by gentle scraping and gentle mitotic shake-off respectively. Cells were counted, and 500,000 cells used as input for every experiment. To perform CUT+RUN on isolated mitotic chromatin extracts, cells were spun at 600 xg for 3 min at 4C for every wash and buffer exchange, as recommended in (Oomen et al., 2019b; Skene & Henikoff, 2017a).

All CUT+RUN buffers were freshly made prior to each experiment, using the chemicals listed below in concentrations and quantities as described by Skene & Henikoff (2017a) and Oomen *et al.* (2019). Pre-manufactured CUT+RUN buffers were not chosen for use, due to inability to know or modify their exact chemical components.

Cells were washed in ice-cold PBS, followed by cell lysis by incubating cells for 10 min in nuclei extraction (NE) buffer (20 mM HEPES-KOH pH 7.9; 10 mM KCl; 0.5 mM Spermidine**;** 0.1% Triton-X-100; 20% Glycerol; 1X Roche cOmplete EDTA-free protease inhibitor (Sigma-Aldrich 11873580001); 1X phosphatase inhibitor (Merck 4906845001)). Cells were washed in Wash Buffer 1 (20 mM HEPES pH 7.5; 150 mM NaCl; 2 mM EDTA; 0.5 mM Spermidine; 0.1% BSA; Roche cOmplete EDTA-free protease inhibitor; 1X Roche PhosSTOP phosphatase inhibitor), then resuspended in Wash Buffer 2 (20 nM HEPES pH 7.5; 150 mM NaCl; 0.5 mM Spermidine; 0.1% BSA; Roche cOmplete EDTA_free protease inhibitor; Roche PhosSTOP phosphatase inhibitor). Nuclei were incubated with primary antibody at recommended concentrations (See Table 2.1) for 2 hr, rotating gently at 4C using an Eppendorf tube circular rotator on low speed; this included a negative control sample to which no primary antibody was added, and a negative control sample with non-specific rabbit IgG primary antibody. Nuclei were washed 3X in Wash Buffer 2, then incubated with protein-AG-conjugated-Micrococcal Nuclease (MNase) (Stratech, 15-1016-EPC) at 1:400 dilution for 1 hr gently rotating at 4C. Nuclei were washed X3 in Wash Buffer 2, resuspended in Wash Buffer 2 and samples transferred to an ice water bath equilibrated to 0C. $CaCl_2$ was added at 2 mM final concentration to samples, activating MNase digestion, while vortexing to allow rapid thorough mixing. Samples were incubated on ice for 30 min to allow digestion before addition of 2XSTOP buffer (200 mM NaCl; 20 mM EDTA; 4 mM EGTA; 50 µg/mL RNase A; 40 µg/mL Glycogen) to stop MNase activity. Samples were then immediately incubated at 37C for 20 min to allow RNA digestion, followed by adding 0.1% SDS and 0.25 mg/mL proteinase K and incubating at 65C for 10 min.

Digested DNA was extracted using 1:1 phenol chloroform extraction as described in (Oomen et al., 2019), and size selection was performed for <700 bp using room temperature AMPure XP beads (Beckmann Agencourt A63881). Purified DNA was quantified using the 1X dsDNA High-Specificity Assay Kit (ThermoFisher Scientific) following manufacturer guidelines.

2.4.2 CUT+RUN: digitonin optimisation

Whole cell CUT+RUN requires cell membrane permeabilization to allow introduction of reagents and diffusion of soluble digested chromatin fragments out of the nucleus and cell. Membrane permeabilization was done using digitonin. Prior to CUT+RUN, the optimal digitonin concentration was determined to allow permeabilization in >95% of cells while maintaining >95% viability. A digitonin dilution series was generated in CUT+RUN wash buffer (20 mM HEPES-KOH pH 7.5; 150 mM NaCl; 0.5 mM Spermidine; 1X Roche cOmplete EDTA-free Protease inhibitor) with 0%, 0.001%, 0.0025%, 0.005% and 0.01% digitonin. A 10 uL aliquot of cells was extracted after incubation in digitonin dilution series, and Trypan Blue dye added to assess cell permeability. A cell counter was used to count total cells in a chosen grid segment on a haemocytometer under confocal microscope. Lysed cells and permeabilized cells were then also counted in that chosen grid, and the % cell lysis and cell permeabilization calculated. This digitonin optimisation experiment was carried out in triplicate on three independent asynchronous HeLa S3 cell samples.

2.4.3 CUT+RUN: Whole cell chromatin digestion

CUT+RUN was performed for harvested whole cells as described originally by (Skene & Henikoff, 2017a). 500,000 asynchronous or synchronous mitotic cells were harvested for each sample, by gentle scraping or gentle mitotic shake-off respectively, at room temperature to minimise cell stress and DNA breakage. Cells were washed 3X in Wash Buffer (20 mM HEPES-KOH pH 7.5; 150 mM NaCl; 0.5 mM Spermidine; 1X Roche cOmplete EDTA-free Protease inhibitor). For whole cell CUT+RUN, concanavalin A-conjugated (ConA) paramagnetic beads (Stratech, 21-1401) were used for each reagent and buffer exchange. ConA beads were activated in Binding Buffer (20 mM HEPES pH7.5; 10 mM KCl; 1 mM $CaCl_2$; 1 mM $MnCl_2$). We added 10 µL of activated beads to each harvested washed cell sample, and samples were incubated with the beads at room temperature while rotating for 10 min. Extracts of 10 µL per cell sample were viewed under a confocal light microscope to observe ConA magnetic bead binding and cell viability (see Chapter 3 for further details). Cells were

permeabilised by addition of Wash Buffer containing 0.005% digitonin, while vortexing to mix thoroughly. Membrane permeabilization was checked by adding Trypan Blue to a 10 µL extract of sample and examining under microscope (see Chapter 3). EDTA 0.8 mM was then added to halt cell metabolism, and primary antibody added at appropriate concentration (see **Table 2.1**) while vortexing. Negative control samples using non-specific rabbit IgG, and no antibody, were both generated alongside experimental samples. Cells were incubated with primary antibody at 4C while rotating gently overnight (16 hr). Cells were then washed X3 in Wash Buffer containing 5% digitonin, before adding 700 ng/mL of pAG-MNase fusion protein while vortexing. Cells were incubated in pAG-MNase at 4C on a rotator for 1 hr. After again washing X3 in Wash Buffer containing 5% digitonin, cells were placed in wet ice equilibrated to 0C. While gently vortexing, 100 mM $CaCl_2$ was added to the samples to activate MNase, before immediately returning to 0C ice bath. Cells were incubated to allow digestion for 30 min before addition of 2XSTOP buffer (340 mM NaCl; 20 mM EDTA; 4 mM EGTA; 0.05% Digitonin; 100 µg/mL RNAse A; 50 µg/mL Glycogen) to halt MNase digestion. Samples were immediately incubated at 37C for 30 min to release digested fragments from insoluble nuclear chromatin, placed on a magnet, and supernatant retrieved. SDS 10% and 20 mg/mL Proteinase K were added to the retrieved supernatant and incubated at 50C for 1 hr. DNA was then extracted from digested chromatin samples using 1:1 phenol chloroform extraction as described in (Meers, Bryson and Henikoff, 2019; 'CUT&RUN: Targeted in situ genome-wide profiling with high efficiency for low cell numbers', no date).

2.4.4 CUT+RUN: library preparation and PCR amplification

Library preparation was performed on extracted DNA samples using Roche Kapa Hyperprep kit (KK8502)**,** according to the manufacturer instructions. End repair & A-tail buffer (Hyperprep), and end repair & A-tail enzyme (Hyperprep), were added to samples and incubated at 20C for 30 min followed by 65C for 30 min, then allowed to chill on ice. 600 nM Truseq adapters were then added to samples while on ice, followed by addition of ligation buffer and ligation enzyme (Hyperprep) and incubation at 20C for 15 min in a thermocycler. SPRIselect cleanup beads (Beckmann Coulter, B23318) were added at 0.95X volume, vortexed gently, and rotated at room temperature for 10 min. Samples were washed X3 with 85% EtOH,

then beads were airdried until shiny. Beads were then resuspended in TLE (10 mM Tris-HCl pH 8, 0.1 mM EDTA), vortexed to thoroughly mix, and incubated at 37C for 10 min. Supernatant was collected, completing cleanup. Truseq PCR primers (Hyperprep) and 2X KAPA HiFi Hotstart ReadyMix (Hyperprep) were added. PCR was then performed to the following settings: 45 sec @98C, 11X cycles of (15 sec @98C + 10 sec @62C), 1 min @72C, hold @12C. 0.95X vol SPRI bead cleanup was performed according to manufacturer guidelines. The retrieved samples were then assessed for sequencing viability as described below. Truseq Adapter sequences can be found in Appendix Table A2.

2.4.5 CUT+RUN: TapeStation

For CUT+RUN samples, Agilent High Sensitivity D5000 ScreenTape Assay kit was used following library preparation to assess fragment size distribution and concentration of the library, following manufacturer instructions.

2.5 MINUTE-ChIP-seq performed by Simon Elsasser and colleagues

The data analysed in Chapter 4 was generated externally by Simon Elsasser's research laboratory using their method: Multiplexed **in**dexed **U**nique molecule **T**7 amplification **e**nd-to-end **seq**uencing: MINUTE-ChIP-seq (Kumar and Elsässer, 2019). Simon Elsasser and colleagues generated G1, S1, S2, S3, S4, G2 and Mitotic cell cycle phase mESC samples using FACS-sorting, and then performed MINUTE-ChIP on FACS-sorted cells from each cell cycle phase, described below. Simon Elsasser has kindly provided unpublished MINUTE-ChIP sequencing data, which was then used for analysis in this project. The MINUTE-ChIP protocol used to generate this dataset is detailed below and was performed following instructions from the EpiFinder Genome Kit (Epigenica).

2.5.1 mESC cell culture

Rw4 murine (male, 129X1/SvJ) embryonic stem cells (mESCs) (ATCC SCRC-1018) were cultured in the Elsasser laboratory, in 0.1% gelatin-coated dishes, in a traditional serum condition to maintain stable pluripotent mESCs (knockout DMEM (Life Technologies); 2 mM Glutamax (Gibco); 0.1 mM non-essential amino acids (Gibco); 15% ESC-grade foetal bovine serum (FBS) (Gibco); 0.1 mM β-mercaptoethanol; leukaemia inhibitory factor (LIF) (Millipore)). Cells were passaged every 48 hr using accutase (Sigma) for detachment.

2.5.2 FACS-sorting mESCs

Prior to MINUTE-ChIP-seq, mESCs were separated into cell cycle phases by Elsasser *et al* using Fluorescence-Activated Cell Sorting (FACS). Around 1.5-2 x 10^9 mESCs were pulsed for 25 min with 10 μM EdU (5-ethynyl 2'-deoxyuridine) before CLICK labelling with Alexa Fluor®-488-Picolyl-Azide. mESCs were then stained first with primary antibody anti-MPM2, then secondary antibody anti-mouse Alexa Fluor®-647, and finally with DAPI. FACS sorting was performed on BD FACSAria™ Fusion and BD FACSAria™III. Cell cycle phases G1, S1, S2, S3, S4, G2 and M were FACS-sorted using example sorting gates shown in **Figure 2.2.**

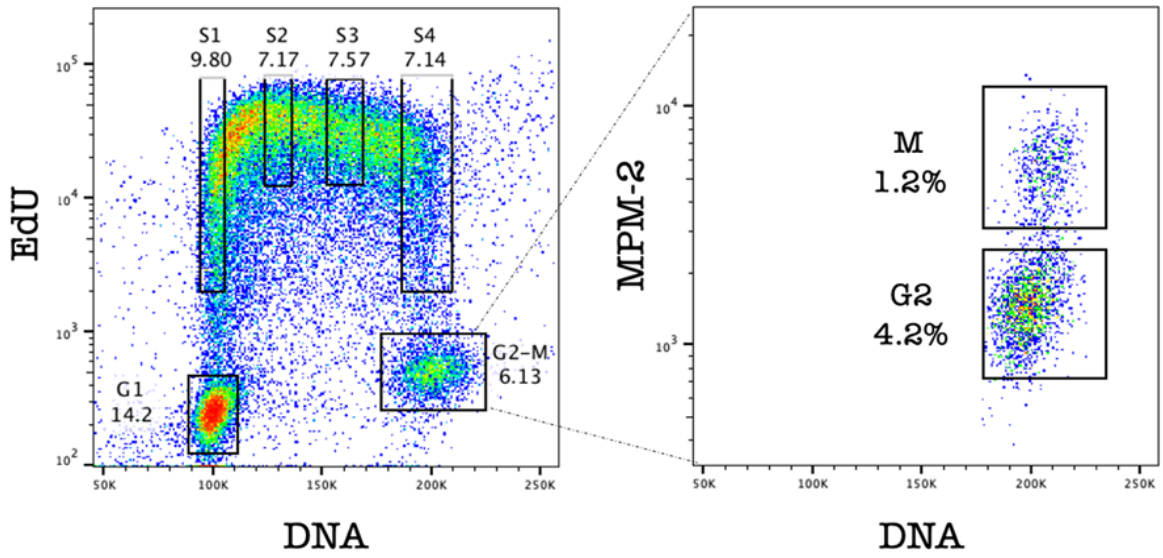


Figure 2. 2: Fluorescence-Activated Cell Sorting (FACS)

FACS-sorting was performed by Elsasser's research laboratory, sorting phases G1, S1-4, G2 and M using sorting gates exemplified here. 5-ethynyl-2'-deoxyuridine (EdU) is incorporated into newly synthesised DNA - cells with higher EdU are thus identified as those in Synthesis (S) phase of the cell cycle (S1-4). Cells with higher DNA levels have undergone DNA replication, and are therefore in the later G2 or Mitosis cell cycle stages (G2-M). Mitotic protein monoclonal 2 (MPM-2) antibody binds the heavily phosphorylated proteins occurring only in mitosis, thus allowing isolation of Mitotic cells (M). Figure by Elsasser et al., unpublished.

2.5.3 MINUTE-ChIP: Cell lysis

For Elsasser's group to perform MINUTE-ChIP, 2 x 10^6 million mESCs were harvested in formaldehyde-fixed pellets for each FACS-sorted cell cycle phase (see Section 2.1.1). Pellets were washed X2 in PBS and flash frozen at –80C prior to use. Cells were resuspended in PBS, and then lysed and digested simultaneously by adding 2X Lysis Buffer (100 mM Tris-HCl pH 8; 0.2% Triton-X-100; 0.1% sodium deoxycholate; 10 mM $CaCl_2$; 1X protease inhibitor cocktail) containing 2 µM/µL micrococcal nuclease (MNase) (New England BioLabs #M0247S). Lysed digested cells were incubated on ice for 20 min, then at 37C for 10 minutes.

2.5.4 MINUTE-ChIP: DNA adapter ligation

Double-stranded DNA adapters were generated by slow annealing of complementary single-stranded oligos (see Appendix Table A3 for adapter sequences). This annealed sequence contained a partial sequence compatible with Illumina Sequencing-bySynthesis (SBS) sequencing technology, in which fluorescently labelled nucleotides are used to enable massively-parallel sequencing. This SBS sequence is flanked by a T7 RNA Polymerase promoter sequence for linear amplification. At the 3' end of the SBS was a randomised 6 base-pair (bp) **u**nique **m**olecular **i**dentifier (UMIs); randomisation creates 4096 possible UMIs, providing sufficient diversity to identify and control for duplicates arising from PCR amplification artefact. UMI sequence is followed by an 8bp sample barcode at the 3' end of the adapter, allowing de-multiplexing *in silico* (see Chapter 4 for further details). These ds-DNA adapters were ligated to the 5' end of digested chromatin fragments by taking the whole cell lysates and performing blunt end ligation reaction using End-It DNA repair kit and Fast-Link DNA ligation kit (Epicenter) at 16C, with optimised adaptor concentration of 2.5 µM/reaction to reduce adapter dimers. Ligation reaction was then stopped using Lysis Dilution Buffer (50 mM Tris-HCl pH 8; 150 mM NaCl; 1% Triton-X-100; 50 mM EGTA; 50 mM EDTA; 0.1% sodium deoxycholate; 1X protease inhibitor cocktail). Barcoded samples were then combined into a single pool and spun down at 4C for 10 min. A typical 1-2 x 10^6 cell equivalent of the pool was used for each ChIP, with 10% of that used as Input.

2.5.5 MINUTE-ChIP: Chromatin Immunoprecipitation

SureBeads Protein A/G Magnetic Beads (BioRad #161-4013/4023) were washed X2 with PBS and 0.1% Tween20 and incubated with primary antibody at room temperature with rotation for 1 hr. Beads were quickly washed with RIPA (radio-immunoprecipitation assay) lysis buffer. The lysate pool (see Section 2.5.3 above) was combined with the antibody-coupled beads and incubated at 4C with rotation for 4 hr. Note, 5% of the lysate pool was kept for the input, and processed through the remaining IP protocol but without primary antibody addition. Beads were next washed in RIPA, RIPA high salt, LiCl and TE (Tris-EDTA) buffer, and resuspended

in ChIP Elution Buffer (10 mM Tris-HCl pH 8; 1 mM EDTA; 0.1% SDS; 300 mM NaCl) containing 0.25 mg/mL Proteinase K (Thermo Fisher Scientific), and eluted at 63C for 1 hr.

2.5.6 MINUTE-ChIP: library preparation, PCR amplification and Illumina sequencing

ChIP DNA fragments >100 bp were isolated using SPRI beads (Beckmann Coulter), and a linear amplification performed by transcription reaction *in vitro* using HiScrobe T7 Quick Yield RNA Synthesis kit (New England Biolabs) as per manufacturer instructions. Transcribed RNA was then purified using Silane beads (Life Technologies). RNA was primed by ligating to an RNA 3' adaptor using truncated T4 RNA Ligase 2 (New England Biolabs) at 63C for 1 hr, followed by addition of reverse transcription components (SuperScript III First-Strand Synthesis SuperMix, Life Technologies) to produce double-stranded cDNA. Purified cDNA was then put through 8 cycles of PCR amplification (High-Fidelity 2X master mix, New England BioLabs), using 0.2 µM PCR primers containing a second ChIP-specific 8 bp barcode sequence compatible with the Illumina sequencing platform. Primer details are provided in Appendix Table A4 and A5. Qubit (Life Technologies) and BioAnalyser (Agilent) were used to assess library quality and concentration. MINUTE-ChIP prepared libraries were sequenced in house using the Illumina NextSeq500 platform, using 50:8:34 cycles (Read1:Index1:Read2). This produced paired-end (PE) sequenced reads, at a minimum of 30 million reads per sample.

2.6 Bioinformatics

2.6.1 MINUTE-ChIPseq: Preliminary data processing, quality control and visualisation

MINUTE-ChIP raw sequencing data preliminary processing was **performed by the Elsasser lab,** as described in (Kumar & Elsässer, 2019a)**.** Software details are provided in **Table 2.2**. Briefly, sequencing reads were demultiplexed using Illumina bcl2fastq, based on the PCR primer barcode, the 6 nucleotide(nt) UMI and 8nt sample barcode to separate fastq files by

sample and replicate. Duplicate reads containing identical first 24nt (encompassing UMI, barcode and 10bp genomic sequence) were discarded, as well as reads matching parts of SBS or T7 promoter adapter sequence. This generated de-multiplexed, de-duplicated reads in fastq files for subsequent mapping. Quality assessments were performed using fastqc (Andrews, 2010).

Paired fastq files were mapped to the mm9 reference mouse genome using bowtie2 v2.3.4.3 (Langmead & Salzberg, 2012) with the –fast parameter, producing BAM files. BEDtools version 2.27.1 (Quinlan, 2014) was used to remove blacklisted regions, removing repetitive and other overrepresented sequences; blacklisted regions can be found at https://github.com/Boyle-Lab/Blacklist/blob/master/lists/Blacklist_v1/mm9-blacklist.bed.gz. The UMI-sensitive deduplication tool Je (Girardot et al., 2016) was used on the BAM files to move UMIs to the header, to estimate library diversity.

Input sample BAM files were converted to 1 bp-resolution BigWig format coverage tracks, using deepTools bamCoverage v3.1.0 (Ramírez et al., 2014) with scaling to 1X reads-per-genome-coverage (1XRPGC) and mm9 genome size 2,654,895,218. Input BigWig coverage tracks were also smoothened using a 10kb moving window and 500bp resolution. ChIP coverage tracks were also generated using deepTools bamCoverage, using only uniquely mappable reads (ie removing multi mapped reads).

Quantitative scaling of ChIP-seq coverage tracks among samples within each pool was performed based on their Input-Normalised Mapped Read Count (INRC). INRC was calculated as:

$$INRC[ChIPSample] = (\#unique\ mapped[ChIP]reads\ /\ \#unique\ mapped[Input]reads)$$

This INRC effectively measures the ratio of ChIP:Input and was used to correct for an uneven representation of barcodes in the Input.

One sample in each pool was then set as the "Reference Sample" and used to generate a scaling factor of 1XRPGC[ReferenceSample]. All other samples in that pool were then scaled to the reference sample using the following calculation:

*(INRC [Sample] / INRC [ReferenceSample]) * 1XRPGC[ReferenceSample]*

Together, this works to both normalise the ChIP sample to the input, and to scale all ChIP samples to the reference sample, allowing quantitative comparisons within the sample pool.

2.6.2 MINUTE-ChIPseq: Analysing data in collaboration with the Elsasser research group

As part of this project, we worked in collaboration with the Elsasser research group to perform downstream analyses on the normalised, scaled ChIP coverage data the Elsasser lab had generated. We utilised this MINUTE-ChIP dataset to address the project's key research questions, as detailed in the sections below.

Table 2.2 details the software, tools and packages utilised during this project for analysis of MINUTE-ChIP-seq data.

Table 2. 2: Software and packages used for analysis of MINUTE-ChIP-seq data.

Software used by the Elsasser group prior to this project are shaded in orange. Software used in our research lab as part of this project for downstream analyses are shaded in green.

Software or package (version)	**Source**	**Link**
Bcl2fastq (v2.20)	Illumina	https://emea.support.illumina.com/sequencing/sequencing_software/bcl2fastq-conversion-software.html
Bowtie 2 (v2.3.4.3)	(Langmead & Salzberg, 2012)	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
BEDtools (v2.27.1)	(Quinlan, 2014)	https://bedtools.readthedocs.io/en/latest/
Je	(Girardot et al., 2016)	https://gbcs.embl.de/portal/tiki-index.php?page=Je
deepTools (v3.10)	(Ramírez et al., 2014)	https://deeptools.readthedocs.io/en/develop/
R (v4.2.1 (2022-06-23))	(R Core Team, 2022)	https://www.R-project.org/

BioCManager (v1.30.23)	(Morgan and Ramos, 2024)	https://CRAN.R-project.org/package=BiocManager
GenomicRanges (v1.50.2)	(Lawrence et al., 2013)	http://www.ploscompbiol.org/article/info%3Adoi%2F10.1371%2Fjournal.pcbi.1003118
Ggplot2 (v3.5.1)	(Wickham, 2016)	https://ggplot2.tidyverse.org
Rtracklayer (v1.58.0)	(Lawrence et al., 2009)	http://bioinformatics.oxfordjournals.org/content/25/14/1841.abstract
Dplyr (v1.1.4)	(Wickham et al., 2023)	https://CRAN.R-project.org/package=dplyr
Tidyr (v1.3.1)	(Wickham et al., 2024)	https://CRAN.R-project.org/package=tidyr

Ggpubr (v0.6.0)	(Kassambara, 2023)	https://CRAN.R-project.org/package=ggpubr
Profileplyr (v1.14.1)	(Carroll et al., 2023)	https://mirror.sjtu.edu.cn/bioconductor/packages/3.16/bioc/manuals/profileplyr/man/profileplyr.pdf
Herper (v1.8.1)	(Paul et al., 2024)	https://github.com/RockefellerUniversity/Herper
Seqinr (v1.0-2)	(Charif & Lobry, 2007)	-
Qgraph (v1.9.8)	(Epskamp et al., 2012)	http://www.jstatsoft.org/v48/i04/
BiomaRt (v2.54.1)	(Durinck et al., 2005, 2009)	-
FSA (v0.9.5)	(Ogle et al., 2023)	https://cran.r-project.org/web/packages/FSA/index.html
Rcompanion (v2.4.36)	(Mangiafico, 2024)	https://CRAN.R-project.org/package=rcompanion

Macs2 (v2.2.9.1)	-	https://anaconda.org/bioconda/macs2

2.6.3 MINUTE-ChIPseq: genomic enrichment analyses

To assess genome-wide distribution of MINUTE-ChIP signal, BigWig files of normalised coverage tracks for each MINUTE-ChIP sample were visualised using Interactive Genome Browser: IGV (Robinson et al., 2011; Thorvaldsdóttir et al., 2013). Genomic distributions of samples were illustrated in R (version 4.2.1) through histograms, violin plots and boxplots, with descriptive statistics generated using FSA and rcompanion (see **Table 2.2**).

ChIP enrichment was assessed specifically at promoters. Genomic coordinates for principal isoform, protein-coding transcription start sites (TSSs) were downloaded from Ensembl from the GRCm38.p6 assembly, release number 102 (download date 16/04/2024), using Biomart (Yates et al., 2020). Window sizes ranging from 500 bp to 10 kb around the TSS were initially assessed by generating enrichment profile plots for a MINUTE-ChIP target with known promoter peak characteristics: H3K4me3. A 2 kb region (+/-1 kb) was chosen as the optimal window size to capture MINUTE-ChIP promoter enrichment peaks (see **Chapter 4 Section 4.2.4** for detailed rationale). Profile plots showing ChIP distribution centred at the TSSs were generated using the profileplyr R package. Enrichment profiles were also plotted across the full gene body from the TSS to the TES (transcription end site).

MACS2 software package was used to identify statistically significant peaks in MINUTE-ChIP enrichments, using addition or removal of the -broad parameter to identify broad or narrow peaks respectively. Default parameters for MACS2 were otherwise used. The proportion of called peaks occurring at TSSs (+/-2 kb) was assessed as is discussed in Chapter 4. The gene Ensemble IDs for TSSs found to overlap called ChIP peaks were also extracted and utilised for subsequent gene ontology analysis (see Section 2.6.6).

2.6.4 MINUTE-ChIPseq: Integrating epigenetic state data

We sought to further investigate MINUTE-ChIP enrichment patterns by segmenting the genome into chromatin states. (Juan et al., 2016) used ChromHMM software (Ernst & Kellis, 2012) to define 20 chromatin regulatory states. The model was trained through input of ChIP-seq mESC datasets targeting three cytosine modifications, 13 histone modifications and the CTCF insulator protein (details and sources for these datasets are listed in Appendix, Table A5). The mm9 mouse reference genome was divided into bins of 200 bp. Using the trained chromHMM model, each 200 bp bin was then labelled as one of the 20 epigenetic states. Confidence filtering was performed such that loci were only included where the epigenetic state could be called with >95% posterior probability. Adjacent genomic bins with the same epigenetic state were then combined to form epigenetic state regions genome-wide, as shown in the illustrative diagram below, **Figure 2.3**.

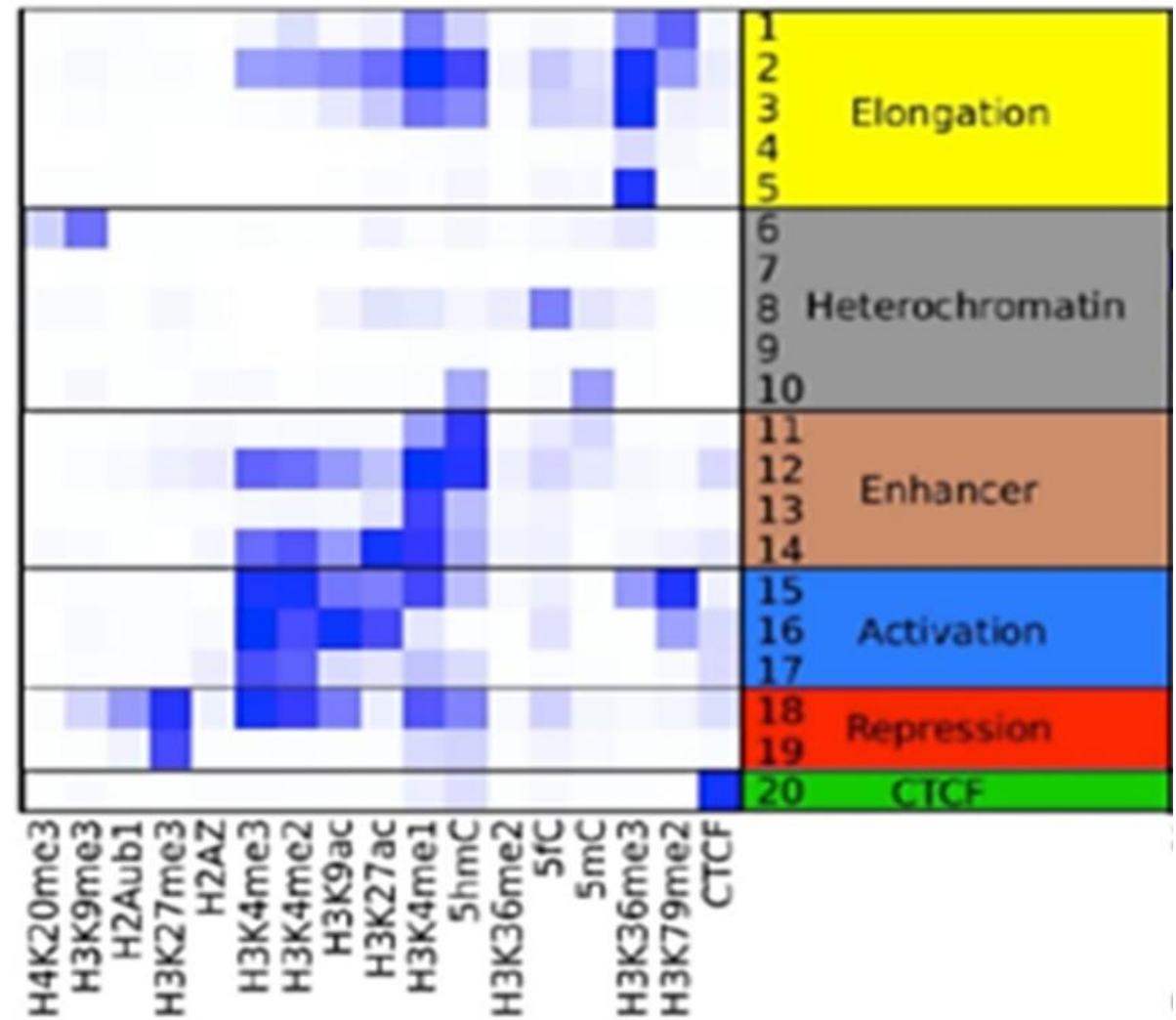


Figure 2. 3: Juan et al (2016) 20 chromatin state model, generated by chromHMM.

The figure shows enrichment heatmap of the epigenomic features (X axis, left) and genomic annotations (x axis, right) input by Juan et al across 20 defined chromatin states (y axis). The number of states was determined as the best biologically accurate representation, with some states grouped into categories: activation, elongation, repression, heterochromatin, and CTCF state. Figure adapted from Juan et al., 2016.

Our MINUTE-ChIP enrichment scores were segmented into chromatin state regions of the genome, and we performed statistical analyses to assess whether histone phosphorylation enrichment showed any significant variations between TSSs (+/-1 kb) in different chromatin regulatory states. Kruskal-Wallis non-parametric test was first performed with a Null Hypothesis that the MINUTE-ChIP-seq enrichment signal for a given histone phosphorylation shows no significant difference between TSSs in any of the 20 chromatin states. Having found significant variation we rejected the Null Hypothesis, and Dunn tests were then performed to assess significant differences between each of the 20 chromatin states, and the resulting p-values plotted to violin plots using ggstats software.

2.6.5 MINUTE-ChIPseq: Co-localisation analysis

To assess the co-localisation of epigenetic features mapped by MINUTE-ChIP-seq, the partial correlation coefficients between all ChIP samples in a given cell cycle phase were calculated, and significant ($p<0.05$) correlations, both positive and negative), were plotted in a correlation network, using qgraph software (see **Table 2.2** for details).

2.6.6 MINUTE-ChIPseq: Gene functional enrichment analyses

The Ensembl IDs were isolated for genes of interest (Yates *et al.*, 2020), and functional enrichment analyses performed using Webgestalt analysis software (Elizarraras et al., 2024; Zhang et al., 2005). Over-representation analysis (ORA) was performed via Webgestalt, analysing the isolated gene list against the mus musculus reference genome. Benjamini-Hochberg (BH) multiple test adjustment method was used (Benjamini and Hochberg, 1995). Gene set enrichment analysis (GSEA) via Webgestalt was then employed, ranking extracted genes of interest based on MINUTE-ChIP enrichment scores. Functional enrichment was assessed against all 15 databases available via Webgestalt, testing against a total 150,937 functional categories. Minimum genes in an enriched set were set to 10 as recommended, for both ORA and GSEA tests, and all other parameters were set as default. Enriched categories

were first ranked by their false discovery rate (FDR), then the top 50 most significant categories displayed based on p values. Categories with FDR<0.25 and p<0.05 were considered as significant enrichments. Example functional databases interrogated for GSEA are listed below, **Table 2.3**.

Table 2. 3: Functional databases interrogated for GSEA via Webgestalt.

Database "class"	**Database**
Gene ontology	Biological Process
	Cellular Component
	Molecular Function
Pathway	KEGG
	Panther
	Reactome
Network	TF target
	miRNA target
Phenotype	Mammalian Phenotype Ontology
Chromosomal Location	Cytogenetic Band

2.6.7 CIDOP-seq and ChIP-seq: Data processing, quality control and downstream analyses

Prior to this project, CIDOP-seq was performed by Rebecca Harris using chromatin from asynchronous and mitotic HeLa S3 cells. Harris also performed paired-end ChIP-seq targeting H3T3ph in asynchronous and mitotic HeLa S3s; both CIDOP-seq and ChIP-seq studies are described in detail in **Chapter 5.** The raw sequencing data produced from both CIDOP-seq and ChIP-seq were fully processed, quality controlled and checked, and analysed by myself during this project, as described below.

Fastq files generated by either CIDOP-seq or ChIP-seq were quality assessed using fastqc (Andrews, 2010) and sequence reads were trimmed using Trimmomatic software (Bolger *et al.,* 2014) to remove contaminating adaptor sequences. Repetitive sequence regions were retained to allow analyses in regions of interest such as centromeres. Trimmed fastq files were quality checked using fastqc to ensure high quality sequencing. Fastq files were then aligned to the GrCh38 human reference genome using bowtie2 v2.3.4.2 using the –very-sensitive option, and the resulting sam files were sorted and converted to Bam format using Samtools (Danecek *et al.,* 2021). Bam files were then converted to BigWig coverage tracks, normalising against Input samples and scaling as a ratio of Sample:Input, using Deeptools bamCompare (Ramírez et al., 2014). The resulting normalised BigWig files were then used for visualisation and subsequent analyses.

For CIDOP-seq analysis, described in **Chapter 5 Section 5.2.1,** Deeptools software (ramirez *et al.,* 2014) was used to generate enrichment profile plots and heatmaps centred at the Transcription Start Sites (TSSs). The centromere-proximal regions aligned to in **Figure 5.2** were defined previously by Rebecca Harris as detailed in Harris *et al.,* 2023; 200 bp bins across the genome were categorised as H3T3ph-positive or -negative based on binarised H3T3ph sequencing files. The largest contiguous block of H3T3ph-positive bins was then found, which Harris *et al.* found to correspond to the centromere as expected given H3T3ph's reported centromeric enrichment. This H3T3ph "domain" was then extended outwards until 3000 consecutive bins (i.e. 600 kb) showed no H3T3ph signal, and this region defined as the centromere-proximal region. In my analysis, **Figure 5.2** aligns CIDOP-seq data to TSSs within this centromere-proximal region.

For ChIP-seq analyses, Deeptools software was also used to generate all enrichment profiles and heatmaps presented in **Chapter 5.** MACS2 peakcalling software was also utilised to binarise single-end H3K4me3 ChIP-seq BigWig coverage tracks previously generated by Rebecca Harris. Binarisation allowed us to label bins genome-wide as H3K4me3-positive or H3K4me3-negative based on whether a MACS2 peak was called. GenomicRanges packages were then used in R to categorise TSS regions based on the presence or absence of H3K4me3 called peaks. The Ensembl IDs (Yates *et al.*, 2020) of genes identified as H3K4me3-negative were functionally assessed using Over-Representation Analysis (ORA) via Webgestalt software (Elizarraras et al., 2024; Zhang et al., 2005). Then, Gene Set Enrichment Analysis (GSEA) was performed, ranking all H3K4me3-negative gene IDs based on H3T3ph ChIP-seq signal. Functional analyses via Webgestalt used the following parameters: Benjamini-Hochberg (BH) multiple test adjustment method (Benjamini and Hochberg, 1995); minimum genes in an enriched set = 10; rank by False Dsicoverey Rate then display top 50 categories displayed based on p value (FDR<0.25 and p<0.05 considered as significant enrichments).

Chapter 3: CUT+RUN as a method to determine mitotic histone phosphorylation distribution

Summary

The first specific aim of the work in this chapter was to generate a high-resolution, genome-wide map of histone phosphorylations that are abundant during mitosis. One histone phosphorylation historically showing enrichment in mitosis relative to interphase is histone H3 serine 10 phosphorylation (H3S10ph). Arguably the most well-studied histone phosphorylation, H3S10ph appears faintly in the centromeric regions before spreading to be seen along the chromosome arms by prometaphase/metaphase of mitosis (Hendzel et al., 1997). Standard chromatin immunoprecipitation and sequencing (ChIP-seq) techniques arguably may have limited ability to interpret histone phosphorylations with widespread signal, due to difficulties distinguishing background sequencing "noise" from biologically relevant enrichment.

In this Chapter, we aimed to improve the precise mapping of mitotically abundant histone phosphorylations, with an initial focus on H3S10ph, to improve our understanding of any subtle variations in their enrichment. We describe development of the targeted immunoprecipitation approach CUT+RUN for use in mitotic HeLa S3 cells. We report successful isolation of mitotic HeLa S3 cells, and demonstrate that HeLa S3 cells can be successfully permeabilised and attached to magnetic beads as per the CUT+RUN method. Although we demonstrate promising results for extraction of H3K4me3 enriched chromatin fragments in positive control experiments, further optimisation is required to improve reproducibility of mitotic CUT+RUN to allow isolation of H3S10ph enriched chromatin in mitosis.

3.1 Introduction

Numerous histone phosphorylations historically show significant enrichment during mitosis, such as H3S10ph and H3S28ph (Hendzel et al., 1997; Gurley et al., 1978; Hsu et al., 2000; Prigent & Dimitrov, 2003; Wei et al., 1998; Dai & Higgins, 2005a; Goto et al., 2002; Kawashima et al., 2010b). It is possible that H3S10ph and other mitotically-abundant histone phosphorylations may indeed be uniformly enriched across all chromosomes *in vivo*, as many have reported: there may be no true functional enrichments. However, an alternative possibility is that limitations in techniques, both laboratory and bioinformatic, have restricted the accuracy of sequencing distribution data, which might have historically prevented us from detecting subtle but significant enrichments of these histone phosphorylations.

We described in Chapter 1 how standard ChIP-seq studies include an Input sample of genome-wide fragmented chromatin; sequencing of this input should represent any variation in read count due to biological variation. For example, increased fragmentation due to higher chromatin accessibility in certain regions of the genome should be represented in the input sample and can then be controlled for. This biological variation represented by the input is often referred to as "background noise". The input control should also enable detection of technical artefacts present in the input and corresponding experimental sample, such as PCR amplification dimers. However, this traditional input sample can produce quite high read counts which are termed "noisy". In the case of histone phosphorylations which are highly abundant genome-wide in mitosis, such as H3S10ph, ChIP sequencing often looks very similar to the background "noise" represented by the Input. Because these abundant histone phosphorylations characteristically do not show strong, sharp enrichment peaks, it is therefore difficult to understand where there is biologically relevant, true enrichment of phosphorylation, if any, rather than what is noise. It is possible that these histone phosphorylations are truly uniformly, broadly distributed genome-wide, as much of the literature has suggested.

However, we argue it is possible that there are subtle, but biologically significant enrichment variations in phosphorylation that are effectively "hidden" by high input sequencing noise. The

schematic below, **Figure 3.1,** illustrates this idea. In this theoretical schematic, it is possible that true histone phosphorylation shows: a) uniform enrichment; b) no enrichment, or: c) subtle enrichment. Normalising experiment sequencing by removing the noisy input sequencing, will result in all three scenarios in low level, uniform normalised "signal" that is not necessarily an accurate representation of the true biological distribution.

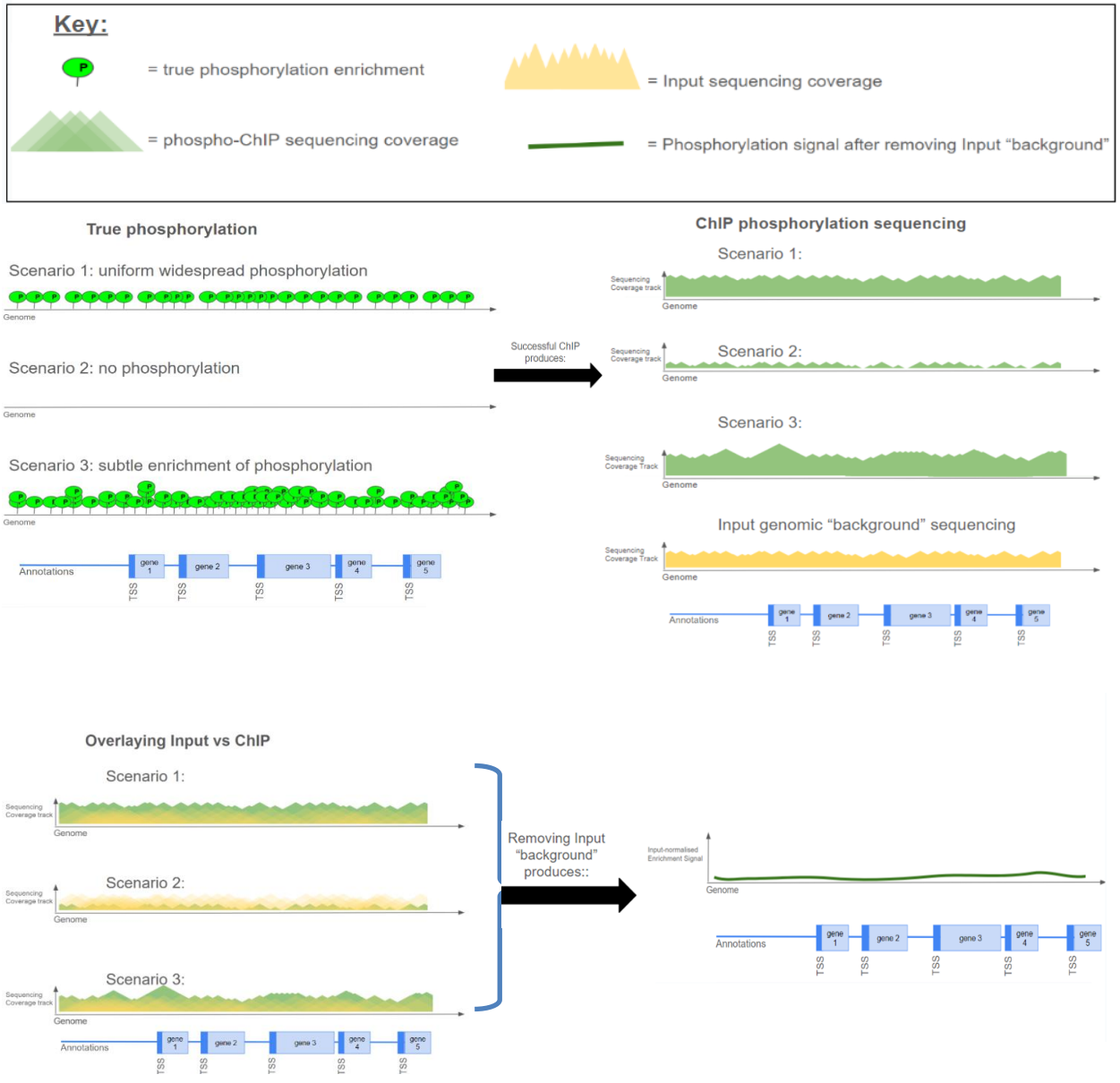


Figure 3. 1: Mitotically abundant histone phosphorylation could have biologically true enrichment peaks that are hidden due to high background "noise".

True in vivo histone phosphorylation (bright green, "P") could be enriched broadly across the genome (Scenario 1:), not enriched anywhere (Scenario 2:) or there could be true but subtle enrichments (Scenario 3:). Traditional ChIP uses an Input sample of the whole genome fragmented as a genome-wide sequence, as a way to measure genome "background" variation (orange). This background sequence signal is then used to normalise (ratio ChIP: Input) ChIP sequence signal to create a normalised measure of phosphorylation enrichment (dark green line, bottom right). However, it is difficult to determine from the resulting normalised signal which Scenario, 1, 2 or 3, is truly occurring in cells. All three scenarios here, once Input Normalisation is performed, produce low, uniform enrichment (bottom right). Example genome simple annotations are provided in blue. Figure is an original illustration.

Chromatin immunoprecipitation techniques continue to advance to reduce the impact of technical artefacts, background "noise" and technical variation. One major example of this is Cleavage Under Targets and Release Using Nuclease: CUT+RUN. An alternative to ChIP-seq, CUT+RUN uses targeted MNase activity to digest chromatin only at target genomic sites, rather than digesting the entire genome as in ChIP. As detailed in Chapter 1, CUT+RUN uses MNase digestion enzyme fused to protein-AG to guide MNase specifically to chromatin sites enriched for the target histone modification, thereby releasing only target chromatin fragments. By leaving off-target chromatin undigested, CUT+RUN has shown significant reductions in genomic "background" noise, and as such does not require an Input control (Skene et al., 2018; Skene & Henikoff, 2017a). CUT+RUN also provides increased throughput, experiment speed, and lower cell number requirements per experiment sample.

In this chapter, we aim to improve our understanding of whether histone phosphorylations with high mitotic abundance truly have broad genome-wide distribution, or whether specific enrichments are present. To do this, we address some of the limitations of traditional immunoprecipitation by instead using the CUT+RUN method, in order to reduce genomic "background" noise in sequencing. Here, we develop the CUT+RUN technique in synchronised mitotic HeLa S3 cells, for extraction of target histone modifications H3S10ph,

and H3K4me3 as a positive control experiment. We successfully permeabilised HeLa S3 cells synchronised in prometaphase, and present promising results for extraction of H3K4me3-enriched chromatin fragments. Further work is needed to improve reproducibility of this relatively new technique, with the hope that future work can be done to generate mitotic histone phosphorylation CUT+RUN data.

3.2 Results

3.2.1 CUT+RUN in HeLa cells shows promise in isolating H3K4 tri-methylated chromatin fragments, but requires further optimisation for histone phosphorylations in mitotic cell lines

HeLa S3 cells were successfully synchronised in mitosis

CUT+RUN first had to be optimised for use in mitotic cells. An asynchronous cell population in cell culture will contain a mixture of cells in different stages of the cell cycle - they are "out of sync". On average, only 2% of cells in the asynchronous population are undergoing mitosis at any given time. Therefore, methods are needed to isolate pure populations of mitotic cells. One such method is mitotic synchronisation. We selected this method to isolate mitotic HeLa S3 cells because the mitotic synchronisation protocol had already been optimised in HeLa S3 cells by Rebecca Harris in our laboratory, as detailed in Chapter 2 Section 2.1. Briefly, asynchronous cells are treated with two rounds of thymidine; thymidine disrupts the nucleotide metabolism pathway, preventing DNA synthesis which halts cells in early synthesis S phase of the cell cycle. By performing 2 rounds of thymidine incubation and release, the majority of cells become synchronised and are held in S phase. The cells are then released by washing off thymidine, allowing the

cells to progress in synchronicity into mitosis. After 8 hours, the cells are treated with nocodazole drug for 5 hours. Nocodazole binds β-tubulin which blocks assembly of spindle microtubules that form during metaphase of mitosis. This therefore blocks the synchronised cells in prometaphase of mitosis. These HeLa S3 cells, halted before chromosome segregation, have condensed chromosomes aligned at the centre of the cell, and the cell morphology appears round and slightly smaller than interphase cells when viewed under microscope. The HeLa S3 cells also become detached from the culture flask surface during mitosis. Because of this detachment, gentle shake-off is used to release these prometaphase cells, which are then collected as the mitotic synchronised cell population. These observable characteristics of prometaphase HeLa S3 cells can also be used to roughly assess the success of mitotic cell synchronisation. Previous optimisation work by Rebecca Harris also performed flow cytometry with propidium iodide and MPM2-Alexa488 staining, to quantify more precisely the proportion of cells in mitosis. It was found that this synchronisation protocol produced a HeLa S3 cell population with >90% mitotic purity (Harris et al., 2023).

This synchronisation protocol was carried out for this experiment and 0.5 million mitotic cells were collected for each CUT+RUN sample. Before beginning the CUT+RUN initial optimisation experiments, these synchronised mitotic cells were observed under confocal microscope to assess cell health and mitotic morphology, to roughly confirm the presence of mitotic cells and absence of interphase cell contamination for the purpose of these preliminary experiments. As shown in Figure 3.3 panel C, collected mitotic sample cells displayed bright, round morphology characteristic of mitotic cells. It was then anticipated that following successful CUT+RUN optimisation experiments, flow cytometry with propidium iodide and MPM2-Alexa488 staining would be performed on future mitotic synchronised cell samples to more accurately assess mitotic purity prior to sequencing of final CUT+RUN experiment samples. **Figure 3.2** shows the mitotic synchronisation timeline and the corresponding microscopy images of HeLa S3 cells at different stages of synchronisation.

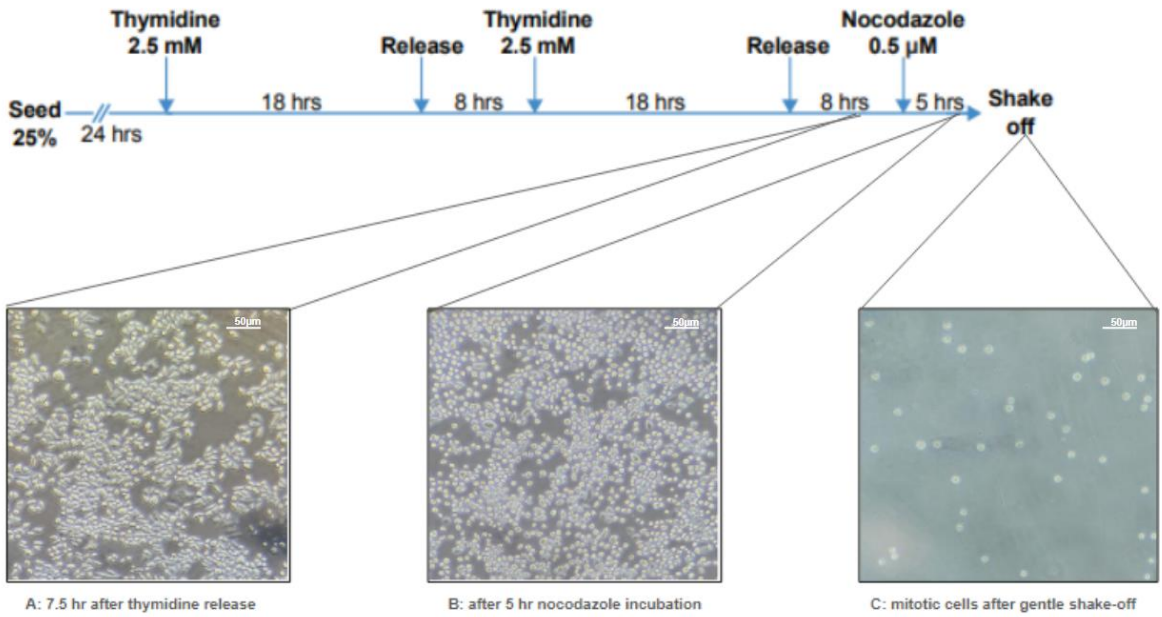


Figure 3. 2: Mitotic HeLa S3 cells following mitotic synchronisation process.

A: *HeLa cells as viewed under confocal microscope 10X objective/resolution, 7.5 hours after release from second thymidine block (See section 2.1.1 for details). Cells show minimal lysis and show healthy interphase morphology.* ***B:*** *The same HeLa S3 cell sample after 5 hours incubation with nocodazole. The majority of cells display bright, round morphology associated with mitosis.* ***C:*** *HeLa S3 mitotic cells are isolated by collecting after gentle shake-off show round morphology and appear bright under confocal microscope. Mitotic cells appear healthy with minimal lysis and minimal contamination with interphase cells.*

CUT+RUN with isolated mitotic chromatin did not yield digested target chromatin

Following mitotic synchronisation as described above, collected mitotic cells were carried forward into CUT+RUN, with asynchronous freshly grown HeLa S3 cells providing samples for comparison. To our knowledge, the only published study to date performing CUT+RUN on mitotic cells was by Job Dekker's group (Oomen et al., 2019a)**.** Their study also used the HeLa S3 cell line, therefore we chose to begin CUT+RUN optimisation by trying to replicate their published methods. Dekker and colleagues based their methods on the Henikoff 2017

protocol for isolated nuclei, rather than whole cells. As described in Chapter 1, original CUT+RUN involves binding magnetic beads to the cell membrane; in the isolated nuclei method, the magnetic beads bind the nuclear envelope membrane. Briefly, the nuclear membrane is then permeabilised to allow target antibodies and digestion enzymes to enter the nucleus enabling targeted chromatin digestion. Digested target chromatin 250 bp fragments are small and soluble, and are released from the larger insoluble nuclear chromosomes. The smaller, lighter soluble target fragments diffuse out of the permeabilised nucleus, which is held to a magnet via the bound magnetic beads, and the digested target fragments can thus be extracted. However, Dekker and colleagues made critical modifications to the Henikoff protocol. Dekker explained that mitotic cells in prometaphase do not possess a nuclear membrane, because the nuclear envelope breaks down at the beginning of mitosis. Therefore, there is no membrane to permeabilise or bind magnetic beads to. Instead, Dekker used nuclear extraction buffer to isolate mitotic chromosome extract (rather than intact nuclei). Permeabilisation was not required, because mitotic chromosomes were free in suspension. In order to perform each buffer exchange and addition of reagents (e.g. digestion enzyme, antibody), centrifugation was instead used. Henikoff's 2017 protocol provides centrifugation as a valid alternative to membrane permeabilisation and magnetic beads. Centrifugation separates cellular content based on weight; large mitotic chromosome clusters were pelleted and then resuspended in new buffer and/or reagent. After digestion of chromatin at target sites, digested 250 bp fragments are smaller than the large mitotic chromosomes; therefore, a final centrifugation was used to pellet the unwanted larger mitotic chromosomes, and the supernatant containing the smaller target chromatin fragments collected.

We chose to replicate this method from Dekker *et al.,* firstly because of their published success with CUT+RUN in HeLa S3 cells in mitosis. Secondly, the nuclear extraction approach means that samples do not have to be kept blocked in mitosis with nocodazole. While nocodazole blocking does produce high mitotic purity in samples, arguably nocodazole, as with any drug, will display some cytotoxicity and likely disrupt normal functioning of the cells.

We replicated the isolated nuclei protocol detailed by Dekker *et al.,* with both asynchronous and our synchronised mitotic HeLa S3 cell samples. For these preliminary experiments, antibody C42D8 against H3K4me3 was used, an antibody previously characterised and

validated by Rebecca Harris through peptide ELISAs. H3K4me3 is a modification with known enrichment at active promoter regions that shows consistent enrichment levels in mitosis and interphase. As negative controls, non-specific rabbit IgG antibody was used, which should not show enrichment for any particular genome sites. A no antibody sample was also included in both mitotic and asynchronous samples, to identify and control for any off-target MNase digestion activity.

Multiple independent replicate experiments were performed, however we were unable to extract any detectable DNA from these Dekker-protocol CUT+RUN experiments for either mitotic or asynchronous samples. We would expect CUT+RUN with the C42D8 antibody to yield at least 50ng of H3K4 tri-methylated chromatin fragments in both asynchronous and mitotic samples, according to original CUT+RUN protocol guidelines (Skene & Henikoff, 2017), and therefore we can assume that our CUT+RUN experiments were not successful using this protocol.

Whole-cell CUT+RUN successfully extracted chromatin fragments in asynchronous HeLa S3 and CA46 cells targeting H3K4me3

After no success replicating Dekker's isolated mitotic chromosome method, whole-cell CUT+RUN methods were also tested. To mediate the concerns that cells might exit mitosis during the CUT+RUN experiment, a low concentration of nocodazole was added to CUT+RUN buffers to hold the cells in prometaphase. As described in Chapter 2 section 2.4.2, this Henikoff approach (Skene et al., 2018)harvests whole cells and uses magnetic bead binding to lipids of the cell membrane and digitonin-mediated membrane permeabilization to allow introduction of target antibody and micrococcal nuclease (MNase) into the cell. Protein-A/G conjugated to MNase recognises and binds primary antibody, recruiting MNase to target chromatin sites. Upon MNase activation, this allows targeted digestion of chromatin to release soluble target site chromatin mono-nucleosomal fragments which then diffuse out of the magnet-bound cell and be collected, leaving larger insoluble nuclear chromatin behind.

For whole-cell CUT+RUN, the Henikoff whole-cell most recent optimised protocol was used (Skene et al., 2018)**.** Several independent replicate experiments were performed using whole-

cell asynchronous HeLa cells, and synchronised mitotic cells following the above mitotic synchronisation process. As with the previous experiments, we used H3K4me3 as a positive control antibody, non-specific IgG as a negative control, and no-antibody sample to measure off-target MNase activity.

Prior to CUT+RUN, digitonin concentration was optimised to maximise the % of cells permeabilised while keeping cell lysis below 5%: high concentrations of digitonin causes full cell membrane breakdown, lysing the cell. Cells were observed under light microscope to assess cell lysis, and Trypan Blue dye was used to assess cell permeablility; manual cell counting was performed for accuracy, and independent tests performed in triplicate. **Figure 3.3** shows the calculated percentage of cells lysed, and permeabilised, for a digitonin dilution series. From these results, 0.005% digitonin was selected as the optimal concentration, producing 87% permeabilisation with only 4% cell lysis observed. Higher digitonin concentration of 0.01%, while increasing permeability to 99%, saw higher lysis at 37.8% of cells. Reduced digitonin levels lessened cell lysis to <4%, but saw dramatic reduction in permeabilisation. 0.005% digitonin was used for all subsequent CUT+RUN preliminary experiments carried out.

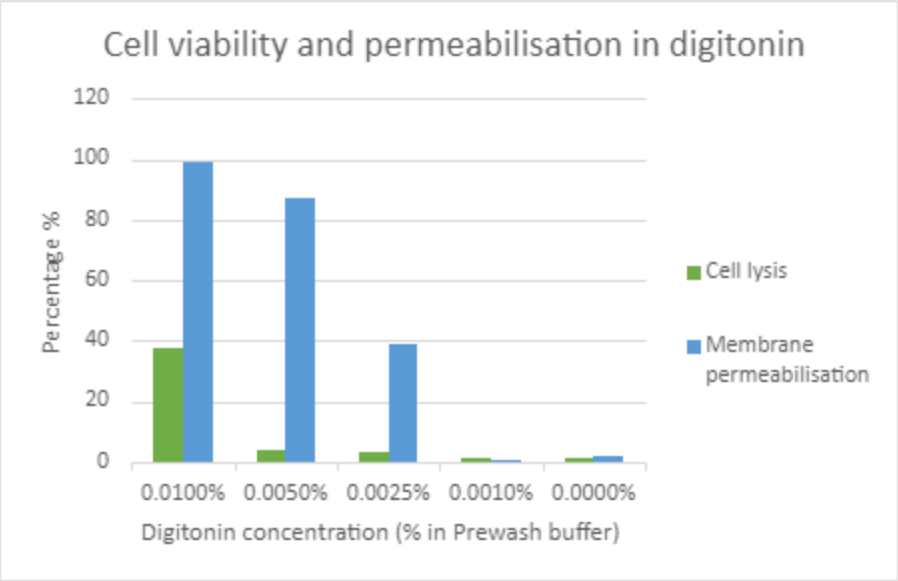


Figure 3. 3: Cell viability and permeabilisation following digitonin incubation.

A dilution series of digitonin (X axis) was added to asynchronous HeLa S3 cells, and the number of lysed cells counted to calculate % viability (green). Trypan Blue was added to cells and the number of cells showing blue staining, indicating membrane permeability, counted to calculate % permeabilisation (blue). Data represents a mean of triplicate independent tests.

During initial validation experiments, aliquots of experiment samples were extracted and examined under confocal microscope to assess correct binding of magnetic concanavalin A beads, as well as membrane permeabilisation by digitonin. Imaging is displayed in **Figure 3.4** below. After cell harvest, cells showed healthy round morphology with minimal cell lysis observed. Note that more rigorous assessment of cell viability would have been performed prior to sequencing of final optimised CUT+RUN samples. Magnetic concanavalin A beads can be seen at high levels evenly coating the outer cell membrane, which allows binding of the cells to magnets during the CUT+RUN experiment. Minimal clumping of cells was seen (which can occur if bead concentrations are too high), which would negatively impact even mixing of reagents during subsequent CUT+RUN steps. Trypan Blue dye diffused into the permeabilised cell membranes staining the cells blue, showing successful cell permeabilisation essential for CUT+RUN. Cell lysis was not seen, confirming optimal digitonin concentration was used allowing permeability without causing cell lysis.

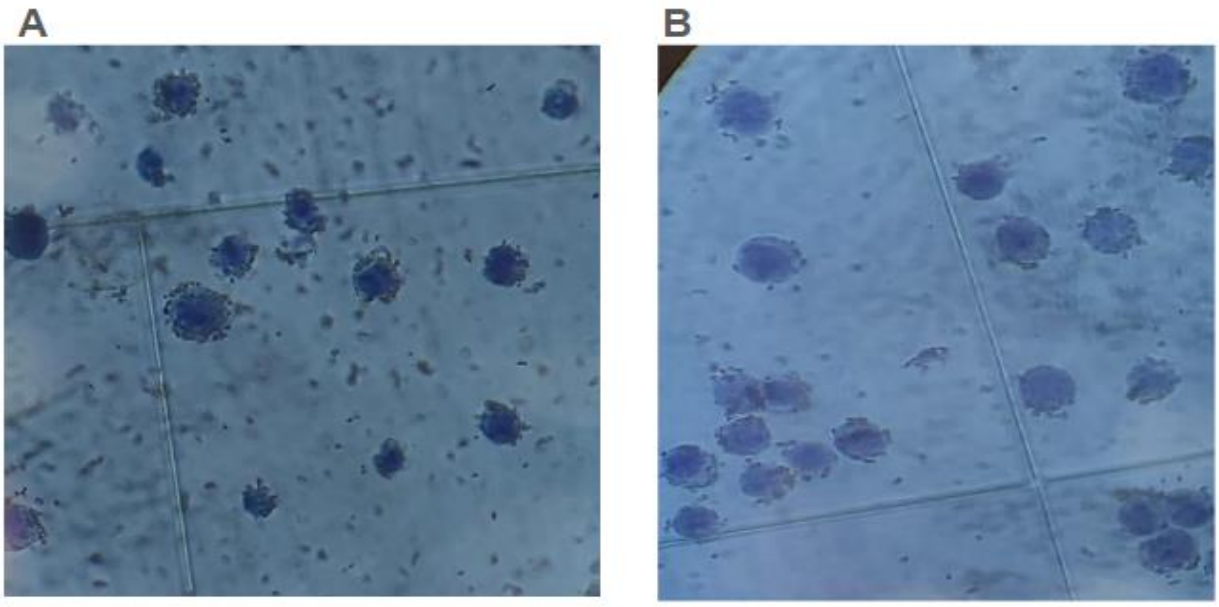


Figure 3. 4: Microscopy shows successful membrane permeabilisation and concanavalinA bead binding in asynchronous HeLa S3 cells.

A(left) *and* ***B(right)*** *display two cell samples tested independently during separate CUT+RUN experiments. CUT+RUN experimental cell samples were examined under confocal microscope 10X objective after addition of Trypan Blue dye. Cells show healthy round morphology and concanavalinA beads can be seen bound evenly and at high levels around all cells. Cells show minimal clumping. Cells appear blue due to Trypan Blue dye diffusion, confirming membrane permeabilisation.*

These validation steps to assess cell viability, bead binding and membrane permeability were performed and were successful for every preliminary CUT+RUN experiment performed. However, experiments failed to yield digested chromatin. We continued to try and optimise parameters according to published studies. A few example parameters modified include: rotation speed during enzyme and antibody incubations; precise vortexing timings to mix thoroughly while mitigating chromatin damage; ensuring minimal contact with sample tubes to reduce variation in MNase enzyme activity during digestion steps. However, QuBit analysis found extracted DNA concentrations to be below detectable range (<1 pg), indicating CUT+RUN was unsuccessful.

We then performed further optimisation experiments in the laboratory of Dr Lisa Russell, a research group we have ongoing collaborations with. Russell and colleagues were also performing CUT+RUN validation experiments in CA46, a Burkitt's lymphoma cell line. Russell *et al.* have successfully isolated and sequenced chromatin at target histone modification sites, using a protocol based on Janssen and Henikoff's whole-cell approach (Janssens et al., 2018).

We hypothesised that the success seen in Russell CUT+RUN experiments could be due to improved CUT+RUN performance in the CA46 cell line. We also postulated CUT+RUN standard conditions might not be optimal for the HeLa S3 cell line. We therefore performed the Russell protocol for CUT+RUN on H3K4me3 and IgG. Both CA46 and HeLa S3 cell line samples were included, and both Higgins and Russell laboratory reagents were used in parallel to test whether differing reagent quality was affecting CUT+RUN performance. 500,000 cells per sample were used, the maximum cell count recommended for CUT+RUN, in order to maximise yield, and optimised primary antibody working concentrations were used throughout.

This troubleshooting experiment was successful, with chromatin fragments successfully extracted for H3K4me3 and IgG samples in both HeLa S3 and CA46 cell lines. Tapestation was used to assess concentration of DNA yielded per sample, and the length of digested DNA fragments. The results of these Tapestation analyses are shown in **Figure 3.5** and **Table 3.1**. Use of Russel laboratory reagents did produce a slightly increased DNA yield in all samples, but more significantly the CA46 cell line generated much higher DNA yield of both IgG and H3K4me3 in all conditions. Nonetheless, all experiment samples exceeded the minimum DNA yield of 1 ng/mL recommended by Henikoff *et al.* as sufficient for successful sequencing**.** This yield of > 1 ng/mL had allowed Russell group to generate good quality sequencing datasets of > 10 million reads previously according to their fastqc reports (Andrews, 2010).

Table 3. 1: CUT+RUN troubleshooting samples and Tapestation-calculated DNA concentrations.

Target antibody	Cell Line	Experiment conditions (reagents used)	Tapestation DNA concentration (ng/mL)
IgG	HeLa S3	Higgins	**1.9**
		Russell	**4.48**
	CA46	Higgins	**14.48**
		Russell	**20.4**
H3K4me3	HeLa S3	Higgins	**3.708**
		Russell	**4.8**
	CA46	Higgins	**10**
		Russell	**14.24**

TapeStation analysis showed successful isolation of DNA fragments for both IgG and H3K4me3 CUT+RUNs, in both HeLa and CA46 cell lines. **Figure 3.5** focusses on Tapestation results for HeLa S3 samples using Higgins laboratory reagents; full Tapestation results can be seen in Appendix in Figure A2. Fragments were found to be around 250-1500 bp in length, representing mono, di and tri-nucleosome fragments of increasing length. Fragment length is seen to peak around 250-400 bp suggesting most isolated fragments were mono-nucleosomal, as desired and recommended by published protocol guidelines. Minimal primer-dimers were seen.

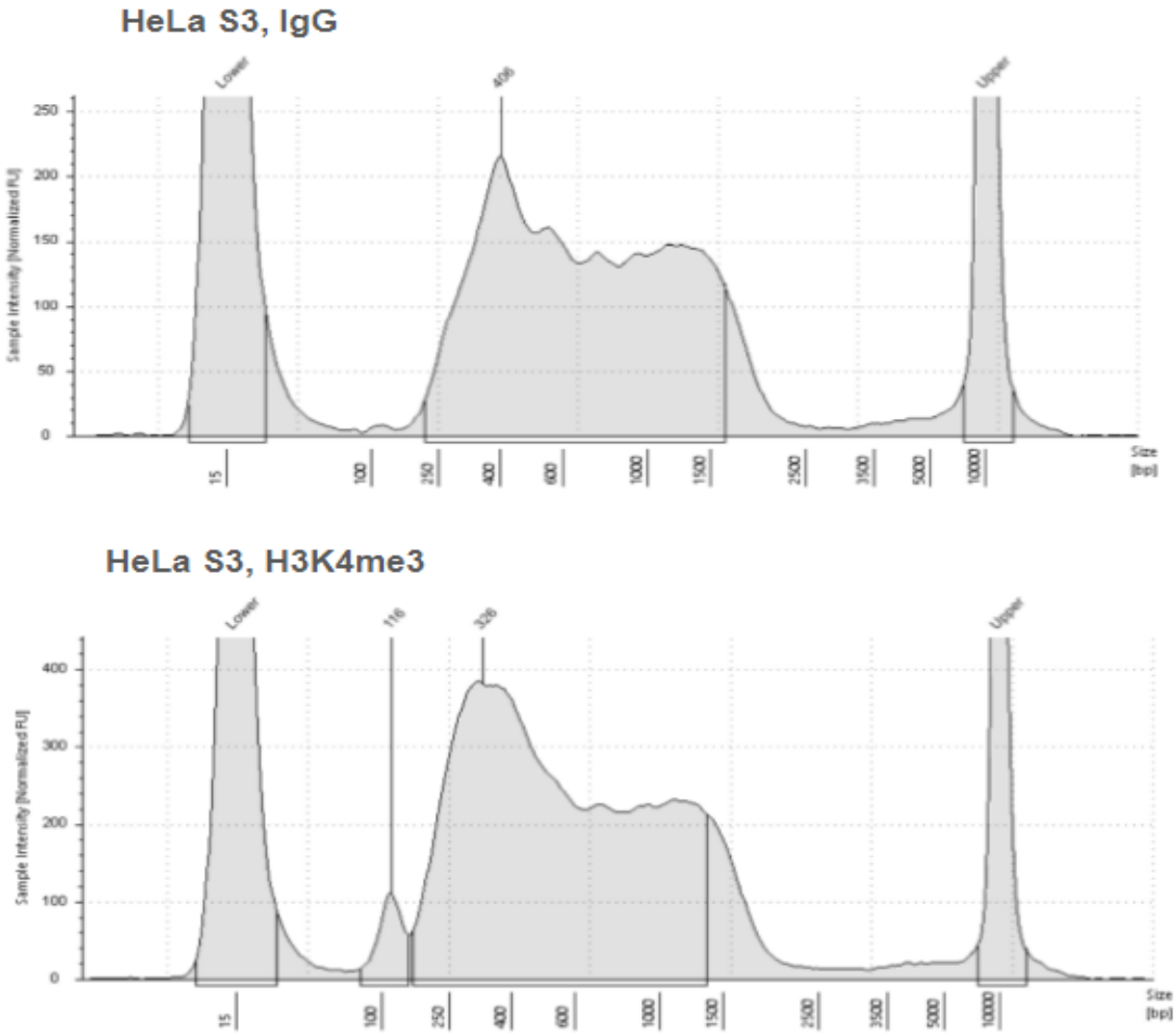


Figure 3. 5: Tapestation analysis shows successful isolation of chromatin fragments from IgG and H3K4me3 CUT+RUN experiments in all cell lines and conditions.

Sample intensity (y axes) was measured against known fragment lengths (x axis, bp) based on a DNA ladder (see Appendix Figure A1). CUT+RUN experiments were performed for non-specific IgG (top) and H3K4me3 (bottom), in HeLa S3 cell line, using Higgins laboratory reagents. A small peak likely resulting from dimer artefacts can be seen in H3K4me3 tapestation (116 bp). Both IgG and H3K4me3 fragments show peaks at a bp length likely to indicate mononucleosome fragments (406 and 326 respectively).

CUT+RUN shows reproducibility issues across multiple cell lines and was not able to extract H3S10ph chromatin sites

Following this successful validation experiment, CUT+RUN experiments were set up and performed under final desired conditions. HeLa S3 cells were synchronised in mitosis (see Section 2.1.1) and CUT+RUN was performed for IgG (negative control), no antibody (negative control), H3K4me3 (positive control), and H3S10ph in mitotic and asynchronous samples, according to the Janssen and Henikoff whole cell protocol as used for the successful preliminary experiments shown above. However, following library preparation and PCR amplification, Tapestation showed that these experiments did not successfully extract any digested DNA fragments. This experiment was replicated twice, but we were unable to yield DNA. We were unable to continue trying to reproduce successful CUT+RUN experiments due to time constraints of the project.

It is evident from the extensive troubleshooting performed during this project that CUT+RUN presents reproducibility issues, and requires further optimisation and validation experiments to develop its use in mitotic HeLa S3 cells. We recommend further validation work for CUT+RUN of histone phosphorylations and in mitotic cells. Our thoughts as to reasons for inconsistencies are discussed below.

3.3 Discussion

The histone phosphorylation H3S10ph has been extensively studied for decades since the phosphorylation was discovered to strongly enrich genome-wide during prometaphase of mitosis; this phosphorylation is highly conserved across eukaryotes, with enrichment spread roughly correlating with chromosome condensation (Hendzel et al., 1997; Gurely et al., 1978; Hsu et al., 2000; Prigent & Dimitrov, 2003; Wei et al., 1998). However, despite this broad enrichment being demonstrated in a range of eukaryotes, previous studies have reported that

the impact and functional role of mitotic H3S10ph varies greatly depending on species and experimental technique (de la Barre et al., 2001; S. I. Houston et al., 2008b; Neurohr et al., 2011; Tada et al., 2011; Wei et al., 1999). It is conceivable that H3S10ph may contribute different functions in different eukaryotes; alternatively, the varying evidence may perhaps be due to redundancy in mitotic regulatory pathways and factors. Further investigation is required in order to build evidence as to the functional contribution(s), if any, of H3S10ph in mitosis.

Here, we emphasise that functional hypotheses can be improved by improving the accuracy and precision of mapping this abundant histone phosphorylation. While it is possible that H3S10 phosphorylation is truly uniformly distributed along the chromosomes in mitosis, it is also possible that instead, there are subtle but significant and biologically relevant enrichments of H3S10ph that have historically gone undetected. We hypothesise that the high "background" noise of control input samples in traditional ChIP-seq method could "hide" subtle enrichments in H3S10ph. In this chapter we assess, validate and develop several steps of the CUT&RUN chromatin immunoprecipitation technique, an advancement of the ChIP-seq technique. CUT&RUN uses a protein-A/G-MNase fusion protein to bind target sites, enabling targeted chromatin digestion and digestion that does not require an input control. CUT&RUN has been reported to significantly reduce genomic "background" noise and generate less noisy sequencing. We argue that CUT&RUN is a strong candidate for mapping of H3S10ph, and would allow detection of any subtle H3S10ph enrichments that may be present.

In preparation for CUT&RUN experiments, we were able to employ the mitotic synchronisation protocol optimised by (Harris et al., 2023), which has previously been demonstrated to generate >90% mitotically pure cell populations. To address the research aims of this chapter, we were specifically interested in the H3S10ph distribution in mitosis; therefore it was vital we minimise contamination of interphase cells in our mitotic samples. Preliminary assessment of mitotic cell samples via microscope observations showed promising results, with high mitotic purity and low cell lysis. However we stress that much more vigorous assessment of mitotic purity would be needed prior to experimental CUT&RUN sequencing of samples. This assessment would be performed as in (Harris et al., 2023), using quantitative analysis of flow cytometry. We note that alternative methods of isolating mitotic cells could instead be adopted; for example, Fluorescence-Assisted Cell Sorting using fluorescent tagging of cell-

cycle-specific molecules can be used to identify and isolate mitotic cells (for review see Garg et al., 2024). This method has shown success with high mitotic purity in published studies (e.g. (Kumar & Elsässer, 2019a). Also, FACS may be advantageous over synchronisation because FACS does not require addition of nocodazole, which will arguably always cause some cytotoxicity that affects ”normal” cell biology in synchronisation approaches. Non-chemical synchronisations could be used alternatively, but these significantly reduced cell numbers per sample.

Whole cell CUT&RUN experiments showed success in extracting chromatin fragments of the correct ~250 bp length expected of the desired mononucleosomal fragments. Fragments were successfully extracted both for H3K4me3 and non-specific IgG, with sufficient yield to recommend sequencing according to original protocol guidelines (Janssens et al., 2018; Skene & Henikoff, 2017a). We do note that extraction of expected fragment lengths does not confirm whether targeted digestion was successful; until sequencing, we cannot conclude whether extracted fragments map to expected areas of H3K4me3 enrichment, for example. Interestingly during validation experiments, we found that yield of chromatin was considerably higher in CA46 cell line than in HeLa S3s; this could indicate that the default/standard experimental conditions of CUT&RUN are better suited to CA46 cell line, and may require further optimisation for HeLa S3 cells.

We were unable to extract chromatin fragments in either asynchronous or mitotically synchronised HeLa S3 cells when we replicated whole cell experiments. Numerous steps were taken to ensure consistency between experiments, referring to guidelines of original CUT&RUN protocols. We assessed magnetic bead binding for every experiment sample and replicate, to ensure oversaturation did not occur, as this would cause cell clumping and reduce chromatin digestion success. Fresh reagents were also used for every experiment replicate where appropriate. Precise protocol instructions for followed to ensure exact timings and temperatures were met for each step.

The inconsistencies seen in our CUT&RUN results may be due to a number of factors, and we encourage further work to optimise this relatively new method (Meers et al., 2019a). We draw attention to the original protocols which themselves reported issues with standardisation, including sensitivity to digestion time, with small variations of mere minutes in digestion

activation time producing dramatic changes from reduced yield up to severe over-digestion and off-target non-specific digestion of the genome (Meers et al., 2019b; Skene et al., 2018; Skene & Henikoff, 2017a). Other factors that could have affected consistency that I suggest include:

- minute differences in timings, e.g. time required to pipette during critical enzyme activation and reaction-stopping steps may have varied
- strength of vortex used to mix reagents during buffer exchange or reagent addition; too low a vortex may not mix reagents well enough to thoroughly saturate cells, whereas too strong a vortex may cause cell or chromatin damage
- temperature changes, for example a difference in room temperature on different days for different experiment repeats, or change in hand temperature when tubes held
- our rotator; some protocols recommend very specific types of rotator, for example Epicypher advise against sample tubes ever being entirely upside down. Our rotator did have tubes upside down, and was sometimes slightly faulty so could possibly have stopped rotating well during overnight incubations which may result in clumping or cells/beads drying out

Moreover, we emphasise that CUT&RUN, to our knowledge, has rarely been performed in mitotic cells, with only a few exceptions (e.g. Oomen et al., 2019b). It is conceivable that in mitotic cells, MNase may be less able to bind target sites due to the condensation of chromosomes; however, developers reported no change in performance in less accessible DNA regions in asynchronous cells (Skene et al., 2018).

We suggest that CUT&RUN experiment parameters may need further optimisation specifically for the HeLa S3 cell line. Furthermore, we also suggest that future work could look to further assess CUT&RUN success specifically in mitotic samples. For example, a target histone modification which is only enriched in mitosis (e.g. H3T3ph) could be used. A successful experiment would be expected to produce chromatin fragments in a mitotic cell sample, but no fragments should be seen in an asynchronous sample, consisting mainly of interphase cells where H3T3ph is not seen. This experiment could be used to further assess the viability of CUT&RUN in mitotic cells.

Chapter 4: Integrating MINUTE-ChIP-seq with epigenetic chromatin state modelling reveals significant enrichment of H3S10ph and H3S28ph at promoters in specific regulatory regions

Summary

The first main aim of the work described in this Chapter was to use a quantitative approach to improve our understanding of exact distributions for histone phosphorylations that are highly abundant during mitosis. The MINUTE-ChIP-seq method, described in Chapter 1, allows quantitative comparison between cell cycle phase samples, and potentially improves quality of within-sample read count normalisation through improved identification of technical artefacts such as PCR over-duplication. The quantitative nature of MINUTE-ChIP-seq provides a unique opportunity to better identify subtle but significant enrichments in abundant, widely-distributed mitotic histone phosphorylations. In this Chapter, we analyse unpublished, quantitative MINUTE-ChIP-seq data targeting the abundant mitotic histone phosphorylations H3S10ph and H3S28ph, generated and provided by the Elsasser laboratory. These MINUTE-ChIP-seq datasets were generated from isolated G1, S, G2, and mitotic mESC cell samples, allowing us to quantitatively analyse how these histone phosphorylation enrichments vary through the cell cycle.

The second main aim of this chapter was to integrate quantitative histone phosphorylation mapping with publicly available epigenetic and gene ontology datasets to assess mitotic phosphorylation in the context of chromatin regulation. In this Chapter, we demonstrate a novel integrative approach combining MINUTE-ChIP-seq H3S10ph and H3S28ph enrichment data with chromatin state modelling. This analysis reveals significant mitotic histone phosphorylation enrichment peaks specifically at promoter regions, and we demonstrate that

these phosphorylations only enrich at specific promoters depending on the chromatin regulatory state. These findings contribute insights into phosphorylation enrichments in the context of mitotic chromatin regulation and provide valuable information as to the possible functional roles these phosphorylations play during mitosis.

4.1 Introduction

This thesis has extensively discussed the difficulties faced in understanding the precise genome-wide distributions of histone phosphorylations that are highly enriched and abundant in mitosis genome-wide; examples include H3S10ph and H3S28ph phosphorylations, both deposited by Aurora B kinase. A large body of research, including microscopy and traditional chromatin immunoprecipitation studies, supports the widespread, likely broad, distribution of H3S10ph along the genome arms during prometaphase of mitosis. While interphase H3S10ph and H3S28ph both show distinct enrichments in regions associated with higher transcriptional activity, as discussed in Chapter 1, the functional roles, if any, of the broad, high abundance seen genome-wide for both H3S10ph and H3S28ph in mitosis are less clear. Studies have reported varying roles for high mitotic H3 phosphorylation, often differing depending on the cell type and/or techniques used in each study. For example, a ChIP-seq dataset generated by (Javasky et al., 2018) in human HeLa S3 cells observed broad, uniform distribution of H3S10ph with no significant enrichment peaks, as seen in the majority of H3S10ph immunoprecipitation studies historically. Contrastingly, a recent study by Meel et al. (2024) in human HeLa cells has observed enrichment of H3S10ph characterised by broad “islands” that they propose spread from specific genomic locations where H3S10ph enriches in interphase. They concluded that H3S10ph may play a role in bookmarking genes that are transcriptional active early after mitotic exit. The precise genome-wide distribution of H3S10ph and H3S28ph in mitosis continues to be studied, and further evidence will help to elucidate the likely functional roles, if any, of mitotic histone phosphorylation enrichments.

We aim to address the biological question: do abundant mitotic histone phosphorylations H3S10ph and H3S28ph show any significant enrichments at specific genome regions during mitosis, and can any functional pathways be identified showing a role of mitotic histone phosphorylation in gene regulation?

In this chapter we examine a likely limitation in experimental techniques to date; traditional ChIP normalisation does not allow quantitative comparison between different ChIP samples. Normalisation cannot be performed between samples because it cannot be assumed that biological and technical variation are constant between different samples and their different corresponding Inputs. There may well be global alterations, both biologically and technically, between samples, and we therefore cannot determine whether an increase in ChIP enrichment in Sample 1 compared to Sample 2 is due to biological true enrichment, or due to a technical difference between samples. **Figure 4.1** illustrates this issue:

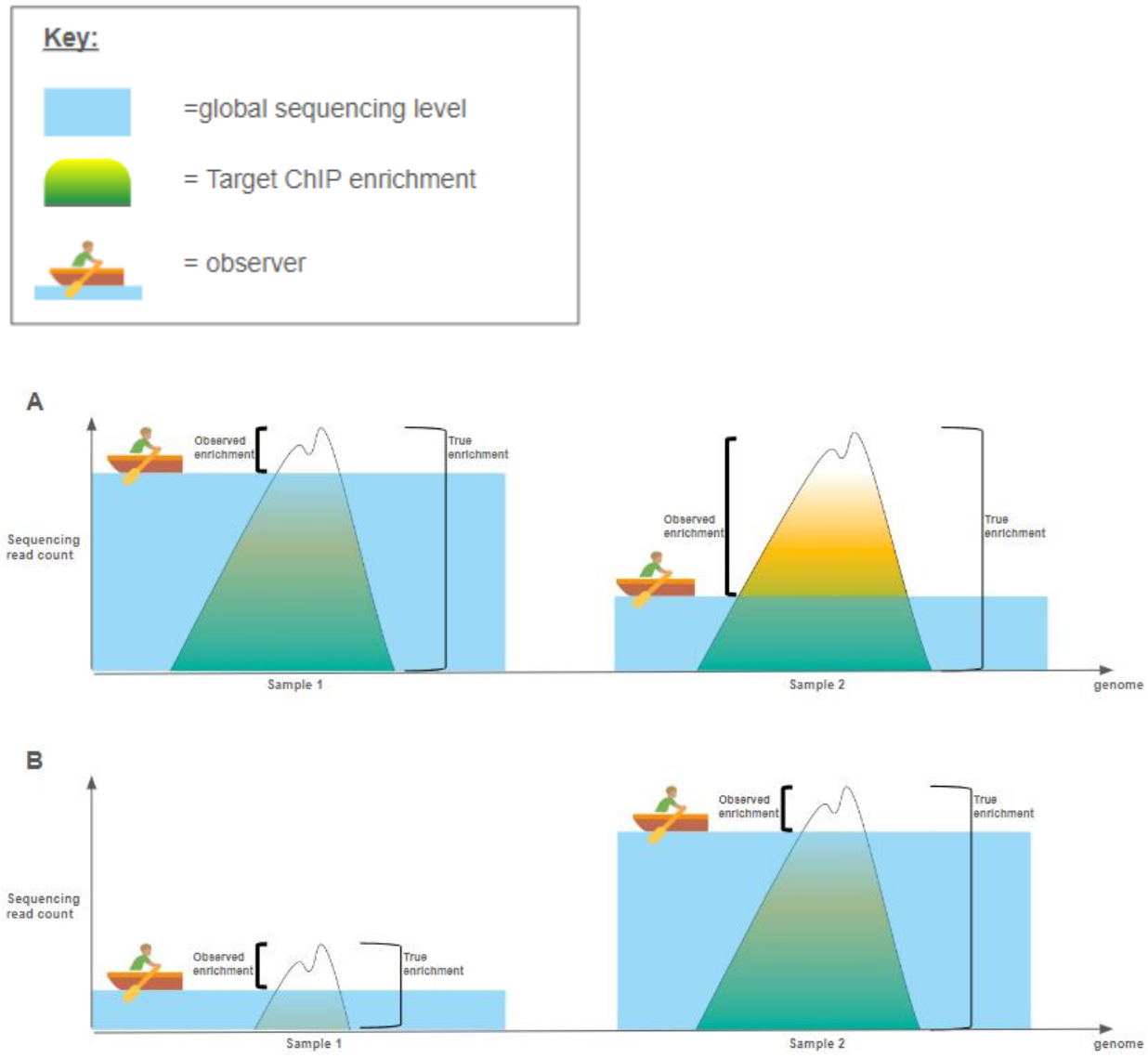


Figure 4. 1: Traditional ChIP-seq normalisation does not account for global changes in sequencing signal between different samples.

Panel A: *ChIP of the target produces the same enrichment (green) between experimental Sample 1 (left) and Sample 2 (right). However, there can be global alterations in sample sequencing (blue) that the observer cannot quantify with traditional ChIP techniques.* ***Panel B:*** *ChIP of target produces different enrichment (green) in Sample 1 (left) and Sample 2 (right). However, global increase in sequence read count in Sample 2 means the observer sees no enrichment change. Figure is an original illustration.*

The MINUTE-ChIP-seq method, using a barcoding approach detailed in Chapter 1, allows experiment samples to be pooled. Pooling increases experiment throughput, and should reduce technical variation between samples in a given pool. After sequencing, barcode sequences can be used *in silico* to separate samples and quantify, and thus control for, any global differences in barcode representation between samples in a given pool. This approach enables quantitative comparisons between pooled samples, generating a more accurate measurement of true biological enrichment of a ChIP target relative to other samples in a pool. Elsasser's lab validation studies using H3K27me3 calibration curves (Kumar and Elsasser, 2019), as detailed in Chapter 1, showed that MINUTE-ChIP and subsequent associated normalisation methods generate a quantitative measure of histone modification enrichment that is accurate to the true biological enrichment (Kumar & Elsässer, 2019). This quantitative method is crucial to our analysis in this chapter, allowing us to quantitatively compare histone phosphorylation enrichment between pooled cell cycle phases G1, S, G2 and Mitosis.

In this Chapter, we harness the quantitative MINUTE-ChIP technique to generate quantitative genomic maps of H3S10 and H3S28 phosphorylations in mouse embryonic stem cells (mESCs). We analyse an unpublished MINUTE-ChIP-seq dataset provided to us by the Elsasser research group, as part of an ongoing collaboration between Simon Elsasser and supervisors Jonathan Higgins and Daniel Rico. Elsasser and colleagues sorted mESCs into cell cycle phases G1, G2, S phase (separated into 4 sub-phases S1-S4), S4 and prometaphase of mitosis, using flow cytometry (see Chapter 2 for details). The samples in each cell cycle phase were then pooled, and MINUTE-ChIP-seq performed on the pool for each of 16 epigenetic features, including mitotically-abundant H3S10ph and H3S28ph, and multiple histone modifications with well-documented roles in chromatin regulation.

First, we use biochemical assays to show that antibodies used by Elsasser *et al.* for these MINUTE-ChIP experiments are able to specifically bind the targets H3S10ph and H3S28ph, and identify any cross-reactivity and impacts of adjacent histone modifications. Our antibody validation studies clearly demonstrate how vital vigorous antibody assessment is in interpreting immunoprecipitation sequencing data. We then show that this MINUTE-ChIP-seq dataset provides a quantitative, high-quality enrichment signal map for H3S10ph and H3S28ph genome-wide, with quantified comparisons between G1, G2 and Mitosis samples. We

observed significant enrichment, as expected, of all three histone phosphorylations in mitotic mESCs. Using bioinformatics tools including ProfilePlyr and MACS2, we reveal an enrichment of both H3S10ph and H3S28ph at promoters.

To deepen our analysis of histone phosphorylation enrichment, we develop an integrative approach combining MINUTE-ChIP-seq H3S10ph and H3S28ph data with publicly available chromatin state data. We present for the first time that H3S10ph and H3S28ph promoter enrichment is affected by the chromatin regulatory state of that promoter region, suggesting an interplay between these histone phosphorylations and mitotic chromatin regulation.

Having identified genes possessing significant histone phosphorylation peaks at promoters in chromatin regulatory states of interest, we then expanded this analysis to assess functional enrichments. By performing gene set enrichment analyses against publicly available gene ontology databases, we find that genes with high H3S10ph at promoters during mitosis are significantly enriched for functions associated with haematopoietic and immune cell phenotypes, and that these genes are in a bivalent regulatory state outside of mitosis. Moreover, we show genes with high H3S28ph at promoters are significantly enriched for negative cell cycle regulatory functions, and that these genes are actively transcribed during interphase. How the mitotic enrichment of H3S10ph and H3S28ph contribute to these functional pathways remains unclear, and we propose future studies to further assess the role of specifically mitotic histone phosphorylation in gene regulation.

4.2 Results

4.2.1 Thorough antibody characterisation is crucial to ensure binding patterns are biologically relevant

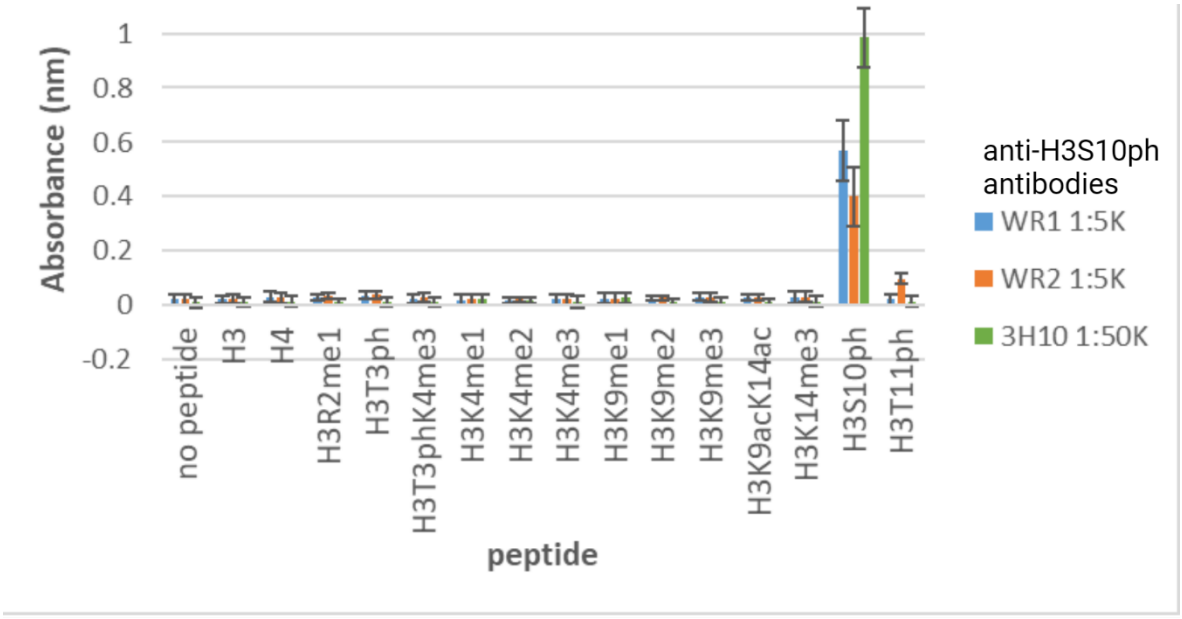
It is vital that the antibodies used in chromatin immunoprecipitation are characterised to assess any impact of adjacent histone modifications. As part of our collaboration with Dr Simon Elsasser and colleagues, during this project we performed peptide ELISAs (Enzyme-Linked Immunosorbent Assays) to characterise the antibodies used to perform MINUTE-ChIP by Elsasser's team. MINUTE-ChIP-seq was performed by Elsasser's team targeting H3S10ph, using the antibody referred to as "3H10", and two antibodies, clones HTA28 and 5D10, were used to target H3S28ph.

We describe thorough validation of numerous antibodies targeting H3S28ph and H3S10ph: namely, anti-H3S28ph antibodies 5D10 and HTA28, and anti-H3S10ph antibodies 3H10, 6G3, RR002, WR1 and WR2 (see Chapter 2 Table 2.1 for antibody details). We argue that in order to accurately interpret which ChIP-seq binding events are biologically relevant, antibodies must first be able to demonstrate ability to recognise and bind the target histone phosphorylation, to **not** bind phosphorylation at other neighbouring histone residues (eg H3T3ph, H3T11ph), and any cross-reactivity of antibodies with adjacent modifications needed to be identified. Any inhibition of antibody target binding by adjacent modifications needed to be known to accurately interpret sequencing enrichments. Optimal experimental concentrations allowing strong antibody binding were also assessed, because overloading of antibody increases costs and more importantly can disrupt normal cell behaviour and reduce ChIP efficacy.

Antibodies were validated using peptide ELISAs (enzyme-linked immunosorbent assays), as detailed in Chapter 2. Peptides consisting of residues of histone H3 were selected possessing an array of modifications (the amino acid residues contained in the peptide varied depending on the location of modifications present). These modified peptides were selected and designed to represent the target H3S10 and H3S28 phosphorylations, and the most commonly observed adjacent modifications of histone H3 that may lead to antibody cross-reactivity. A fairly large range of promising anti-H3S10ph antibodies were identified whose manufacturers reported

strong, specific binding of H3S10ph, and which were available to our laboratory. We tested a total of five candidate antibodies reported to target H3S10ph, hereby referred to by their clone names: 3H10, WR1 (Wang Rabbit 1), WR2, 6G3 and RR002. Antibodies targeting H3S28ph, clones 5D10 and HTA28, were also tested. Thorough validation was performed for each of these seven antibodies against a wide range of peptides, enabling us to confidently select the optimal antibody for subsequent immunoprecipitation experiments. The antibodies validated and their sources are detailed in Chapter 2. Note that 6G3 and RR002 ELISAs were performed prior to this project by Rebecca Harris within our Higgins laboratory.

The absorbance measurements of each antibody were plotted against H3(1-21)S10ph, H3(8-28/20-40)S28ph and an array of adjacent peptides. The absorbance findings (averaged across triple replicates) are shown below in **Figure 4.2**. Note that optimal antibody concentrations for strong, specific antibody binding were first determined through antibody dilution series where possible, else manufacturer guidelines were followed.



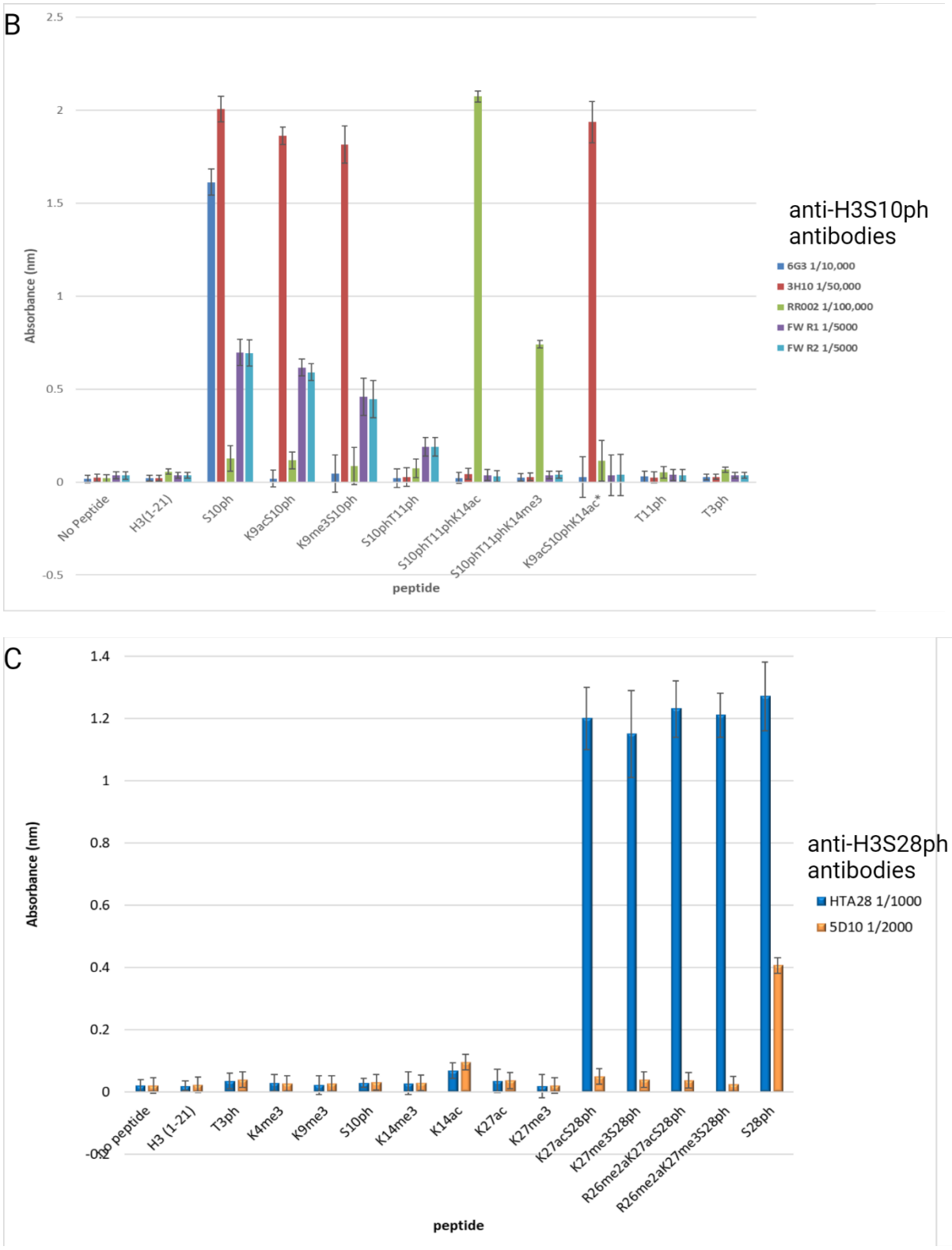


Figure 4. 2: ELISA absorbance readings showing binding affinities of anti-H3S10ph and -H3S28ph antibodies, against H3(1-21) S10ph, H3S28ph, and adjacent peptides.

ELISAs were performed in duplicate with 3 independent replicates; the mean absorbance readings are presented here. Peptides were composed of residues of the histone amino acid sequence (in this example amino acids 1-21), with an array of different post-translational modifications (PTMs). A range of peptides with a variety of possible PTMs neighbouring H3S10ph (see x axis key) were each added to a well in a streptavidin coated 96-well plate. The bar charts below show absorbance (nm, y axis), indicating antibody binding, for each peptide (x axis). Standard Deviation error bars shown. Panel A: Anti-H3S10ph antibodies Wang Rabbit clone 1 (WR1), Wang Rabbit 2 (WR2) and 3H10 ELISA against serine10-adjacent histone modifications lacking H3S10ph. Panel B: Anti-H3S10ph antibodies 6G3, 3H10, RR002, WR1 and WR1 ELISA against serine10-adjacent histone modifications including those with H3S10ph. Panel C: Anti-H3S28ph antibodies Millipore clone 5D10 and HTA28 against serine28-adjacent histone modifications.

These ELISAs clearly demonstrate the variation that exists in binding characteristics of antibodies reported to target the same histone modification. Antibody RR002 showed poor binding to the target S10ph in some peptides (see Panel B, green), which would limit interpretation of chromatin immunoprecipitation using RR002 and make true S10ph enrichment, or lack of enrichment, impossible to distinguish from antibody artefact. Antibody 6G3 showed binding specificity to H3S10ph peptide, with low absorbance readings for peptides absent of S10ph. However, use of peptides with multiple modifications revealed that 6G3 antibody S10ph binding may be inhibited by some adjacent modifications. 6G3 binding affinity appears to be inhibited by any of the common adjacent modifications: including acetylation of histone H3 lysine 9 or 14, H3T11ph, or H3 lysine 9 trimethylation. Therefore, presence of any of these adjacent histone modifications could inhibit antibody binding of H3S10ph.

The two rabbit antibodies from F Wang (Zhejiang University), WR1 and WR2, and 3H10 arguably showed the most promising results in efficacy for chromatin immunoprecipitation. Both 3H10 and WR1+2 showed strong S10ph binding, including to peptides with adjacent methylation and/or acetylations, and did not bind off-target modifications, as shown in **Figure**

4.2A. However, all three of these antibodies show some cross-reactivity with adjacent modifications. 3H10, WR1 and WR2 all showed reduced absorbance, and thus binding, to H3S10ph in peptides where T11 is also phosphorylated. WR1 and WR2 also show reduced H3S10ph binding when K9 and K14 were acetylated as well as S10 phosphorylation. These results suggest that T11ph may inhibit S10ph binding by 3H10, WR1 and WR2, and that adjacent lysine acetylation may inhibit WR1 and WR2 S10ph recognition. This binding inhibition in the presence of T11ph or lysine acetylation is of vital importance as context when interpreting any immunoprecipitation sequencing data generated using these antibodies; for example, a reduced enrichment in H3S10ph sequencing could, based on these antibody characteristics, be due to either T11ph or lysine9/14 acetylation rather than true absence of S10ph.

As shown in **Figure 4.2**, we found that 3H10 antibody has stronger binding ability relative to working antibody concentration for H3S10ph than either F Wang antibodies, with absorbance readings suggesting 2-fold antibody binding despite 3H10 being used at 10X weaker concentration. We caveat that other H3S10ph-targeting antibodies likely exist, however time and financial constraints on the project limited our pool of candidate antibodies to those shown above.

Focussing on H3S28ph target, antibody HTA28 showed strong specific binding to peptides containing the S28ph PTM, and HTA28 was able to bind S28ph when other adjacent PTMs were also present demonstrating minimal cross-reactivity. HTA28 did not bind peptides absent of S28ph and did not recognise other off-target phosphorylations S10ph or T3ph. These results validate HTA28 as a strong candidate for accurate binding of H3S28ph in chromatin immunoprecipitation studies.

Contrastingly, antibody 5D10 showed cross-reactivity and was unable to bind S28ph when adjacent PTMS K27me3, K27ac and/or R26me2a (asymmetrical di-methylation) were present. This antibody could provide an interesting strategy for analysis of H3S28ph interactions and co-localisations in future studies. However this was beyond the scope of this project, and for our aims, this inhibition would make it extremely difficult to interpret immunoprecipitation sequencing results - for example a sequencing depletion could be due to true biological absence of H3S28 phosphorylation, or due to inhibitory presence of any of these adjacent PTMs.

These validation experiments highlight the importance of thorough antibody characterisation prior to interpretation of immunoprecipitation sequencing, and suggest that 3H10 anti-H3S10ph and HTA28 anti-H3S28ph antibodies are able to strongly and specifically bind their targets with minimal cross-reactivity, validating their use in this MINUTE-ChIP-seq dataset.

4.2.2 MINUTE-ChIP performed by the Elsasser team produced high-quality sequenced samples of H3S10ph and H3S28ph in prometaphase mESCs

MINUTE-ChIP raw sequencing data was generated by Elsasser *et al.*, who then performed the initial data processing as described in detail in **Chapter 2.** Briefly, sequencing reads were aligned to the mm9 reference genome, excluding repetitive regions. The Unique Molecular Identifier (UMI) allowed more accurate identification and removal of dimers and duplicates than traditional ChIP-seq; these artefacts can contaminate samples following PCR amplification. As detailed in Chapter 2, input-based normalisations allowed effective scaling to read coverage and controlled for within-sample read count variation. As an additional step in Elsasser's normalisation method, the sample read counts in a given pool are scaled to a single selected "Reference Sample" within that pool - critically, this allows quantitative comparison of sequencing signals between different samples within a given pool.

This initial sequence processing produced normalised genome coverage files in Bigwig format, of MINUTE-ChIP-seq samples in G1, S1, S2, S3, S4, G2 and Mitotic (M) samples, using antibodies for H3S10ph and H3S28ph, along with 24 other histone modification and transcription factor targets:

- H3K4me1
- H3K4me3
- H3K9me3
- H3S10ph
- H3K27ac
- H3K27me3
- H3S28ph
- H3 (total)
- H2AK119Ub
- H3.3
- H4K20me1
- H4K20me2
- H4K20me3
- CENP-A

- H2A.Z
- H2Aub-1
- H2Aub-2
- CTCF
- Ezh2
- Ki67
- Nanog
- Ring1b
- RNApolII
- RNApolII-S2P
- RNApolII-S5P

The antibodies used for these targets are detailed in Chapter 2, Table 2.1. For this project, we used these normalised Bigwig files for all subsequent analyses, detailed below.

The quantitative nature of MINUTE-ChIP-seq data allowed us to analyse more subtle variations in H3S10ph and H3S28ph enrichment and distribution than traditional qualitative ChIP-seq. The Bigwig files contained a score, based on the INRC-normalised, scaled read count, at each genome coordinate. From here on, we refer to this score as a "normalised signal". As a preliminary assessment of the data quality, H3S10ph and H3S28ph enrichment signals genome-wide were plotted for each of the 7 cell cycle phases sampled. **Figure 4.3** shows the distribution of these normalised signals as whisker plots. Both H3S10ph and H3S28ph MINUTE-ChIP-seq show significant signal increase in mitotic samples compared to interphase. This finding aligns with previous literature demonstrating that both of these histone phosphorylations enrich markedly in mitosis. H3K4me3, reported widely to remain at a constant level through the cell cycle, showed a consistent similar global signal distribution in mitosis. Total histone H3, targeted using antibody anti-H3 Ab1791, also remains constant through the cell cycle, although we do note that overall H3 enrichment levels are very low which may suggest poor ChIP success for the H3 targeting experiment. For further comparison, MINUTE-ChIP of the repression-associated modification H3K27me3 was also assessed. As demonstrated in the whisker plots, H3S10ph and H3S28ph show a 7/8-fold increase in Mitosis compared to G1, whereas H3K4me3 and H3K27me3 do not increase from G1 to Mitosis. This preliminary visualisation supports that the MINUTE-ChIP-seq performed are of good quality, and support reported enrichments of these histone modifications through the cell cycle.

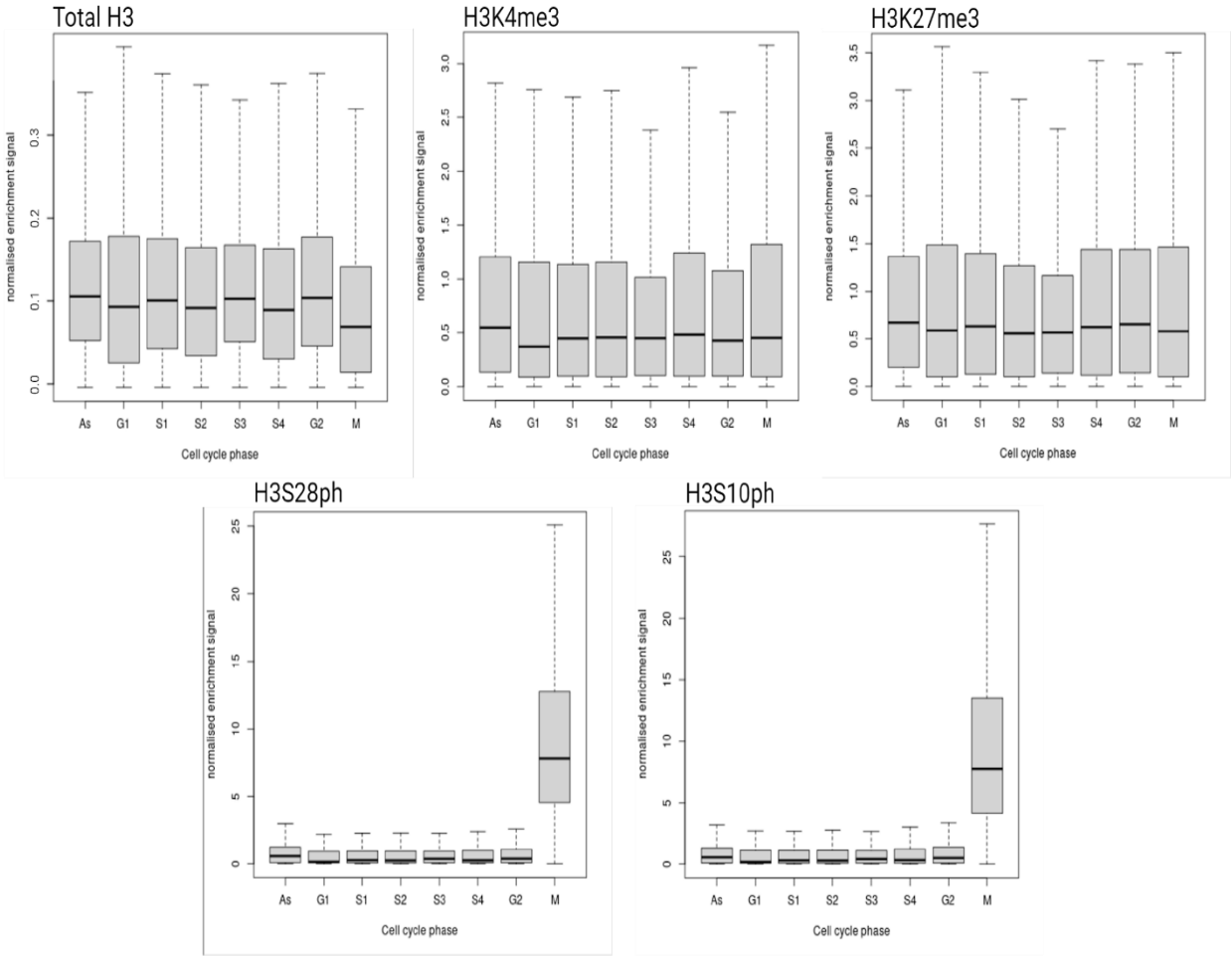


Figure 4. 3: Whisker plot of MINUTE-ChIP-seq genome-wide enrichment signal, per cell cycle phase.

Plot shows total H3 (top left), H3K4me3 (top centre), H3K27me3 (top right), H3S28ph (bottom left) and H3S10ph (bottom right) input-normalised MINUTE-ChIPseq enrichment signal separated into cell cycle phases G1, S1-4, G2, mitosis (M) and asynchronous sample (As). Whisker plots show the mean (central black lines), interquartile range (grey box) and the 5th and 95th percentile (upper and lower bars) of the normalised MINUTE-ChIP-seq signal for each cell cycle phase sample. Note the varying scale (y range).

We then sought to quantitatively analyse H3S10ph and H3S28ph MINUTE-ChIP-seq in mitosis to investigate whether enrichments occur in certain genome regions, or whether they are in fact uniformly distributed genome-wide.

4.2.3 H3S10ph and H3S28ph MINUTE-ChIP-seq shows enrichment at promoters

To investigate potential functional roles of abundant histone phosphorylations in mitotic chromatin regulation, we focussed analysis on regulatory regions and elements. Using Deeptools software, we examined the signal of H3S10ph, H3S28ph and other regulatory histone modifications to assess their enrichment at Transcription Start Sites (TSSs). Only the principal isoform TSS of protein-coding genes (according to APPRIS annotation (Rodriguez *et al.*, 2022)) were assessed, to ensure only one TSS per gene was analysed; this prevented any downstream gene enrichment analysis from being biassed by varying isoform numbers or pseudogenes. We do note that exclusion of these non-principal isoforms and alternative gene types could potentially exclude enrichments and putative functions of histone phosphorylation at these alternative regions. **Figure A1** in the Appendix is provided for details of the number of TSSs in each gene type and isoforms based on Ensembl Biomart downloaded TSS coordinates.

Figure 4.4 below shows the genome-wide mean enrichment profile of H3S10 and H3S28 phosphorylations centred at the TSS, in G1, G2 and Mitosis(M) to allow cell cycle phase comparison. For these preliminary analyses, 10 kilobase (kb) windows around the TSS coordinate were used to generate profile plots, to ensure even broad peaks would be observable. H3K4me3, a histone modification associated with actively expressed gene promoters, was used as a positive control with expected promoter enrichment. H3K27me3 MINUTE-ChIP profile was also plotted. H3K27me3 is a histone mark associated with repressed chromatin regions, and is deposited by polycomb group proteins in heterochromatin regions (Lau & Cheung, 2011). Therefore, we would not expect to see H3K27me3 enrichment at promoters in an active, open/accessible regulatory state. Total histone H3 was also profiled, which was not expected to be significantly enriched at promoters.

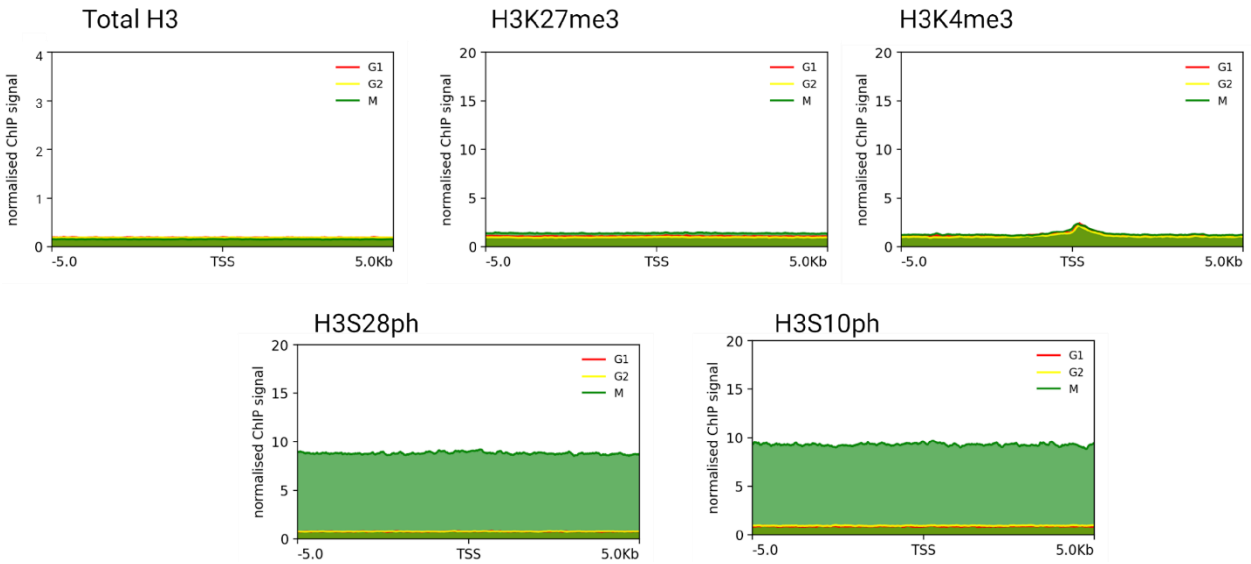


Figure 4. 4: Enrichment Profile Plots of H3S10ph (bottom right), H3S28ph (bottom left), H3K4me3 (top right), H3K27me3 (top centre) and total histone H3 (top left).

Profile plots of G1 (red), G2 (yellow) and Mitotic(M, green) samples are shown overlaid for each histone modification.

When plotting genome-wide MINUTE-ChIP-seq H3S10ph and H3S28ph signal, it can be seen that the expected mitotic enrichment was evident, but no TSS enrichment was observed. The H3K4me3 sequencing did show an enrichment at TSSs**.** Importantly, we propose that assessing genome-wide enrichment signal is too simplistic; the genome is composed of regions in many different regulatory states, ranging from highly accessible, "open" chromatin associated with active gene expression, to tightly condensed, repressed regions associated with reduced gene expression. These different regulatory "states" are characterised by enrichment of different histone modifications; for example, H3K4me3 is expected to enrich in sharp, narrow peaks specifically at promoters that are in active, euchromatin regions of the genome, whereas H3K27me3 is found in more repressive polycomb-associated regions**.** Therefore, enrichment of these regulatory modifications will only be observed in specific regions of the genome in certain regulatory states. This analysis highlighted to us that in order to confidently assess whether significant enrichment peaks occur for a given histone modification, the chromatin regulatory context must first be taken into account.

We therefore hypothesised that regions with genuine enrichment in mitotic H3S10ph and H3S28ph might be "hidden" in genome-wide distribution visualisations, and that separating the genome into different epigenetic regulatory categories might reveal genuine histone phosphorylation enrichments previously undetected.

4.2.4 Mapping the mouse genome based on chromatin regulatory state

We created an integrative approach to place MINUTE-ChIP-seq histone phosphorylation data within a chromatin regulatory landscape context. This aimed to test whether H3S10ph or H3S28ph may have significant enrichment peaks at specific promoters within specific chromatin regulatory states.

MINUTE-ChIP-seq data were integrated with a chromatin epigenetic state dataset for mESCs generated as part of an international collaborative study co-led by by Daniel Rico and Alfonso Valencia published in 2016 (Juan et al., 2016). To generate these data, David Juan and colleagues gathered a wide range of published epigenomic datasets for asynchronous mESCs, including 139 ChIP-seq, MEDIP (methylated DNA immunoprecipitation) and GLIB (glucosylation, periodate oxidation and biotinylation, used to biotinylate and elute 5hmc chromatin fragments) datasets containing 77 epigenomic features including 13 histone modifications. These cytosine and histone modification datasets were used to train a 20 chromatin regulatory state ChromHMM model, illustrated in the figure below (**Figure 4.5**). Each chromatin regulatory state is characterised by the relative enrichments of the epigenomic features displayed on the x axis (Figure 4.5). These 20 states were then labelled in categories based on their likely biological relevance: elongation, heterochromatic, enhancer, activation, repression, and CTCF (Juan *et al.*, 2016).

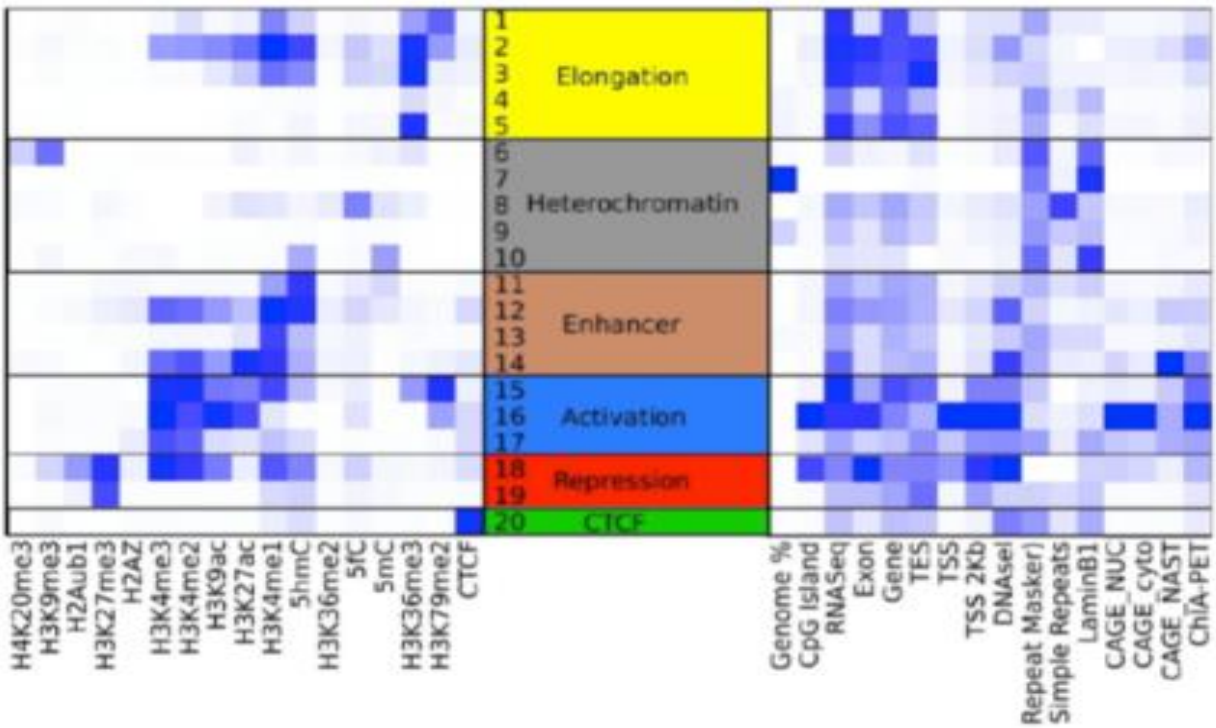


Figure 4. 5: Juan et al 2016 20 chromatin state model, generated by chromHMM.

The figure shows enrichment heatmap of the epigenomic features (X axis, left) and genomic annotations (x axis, right) input by Juan et al across 20 defined chromatin epigenetic states (y axis). The number of states was determined as the best biologically accurate representation, with some states grouped into categories: activation, elongation, repression, heterochromatin, and CTCF state. CAGE_NUC, CAGE_cyto and CAGE_NAST (Fort et al 2014) refer to CAGE in nuclear compartment, cytoplasmic compartment and non-annotated stem transcripts respectively. Figure from Juan et al, 2016 supplementary material.

The chromatin state model was then used by Juan et al to categorise the mm9 genome into these 20 epigenetic state regions, as described in Chapter 2. Briefly, the mm9 mouse reference genome was divided into 200 bp bins, and each bin categorised into these 20 chromatin states, only including intervals with > 0.95 probability of a confident chromatin state. 200 bp bin sizes were presumably chosen by Juan *et al.* to provide a high resolution of chromatin state mapping. Adjacent bins categorised as the same chromatin state were then pooled to form a chromatin state region, illustrated in schematic **Figure 4.6**. This provided an epigenetic state "map" of the mm9 genome against which we could align MINUTE-ChIP-seq H3S10ph and H3S28ph signal.

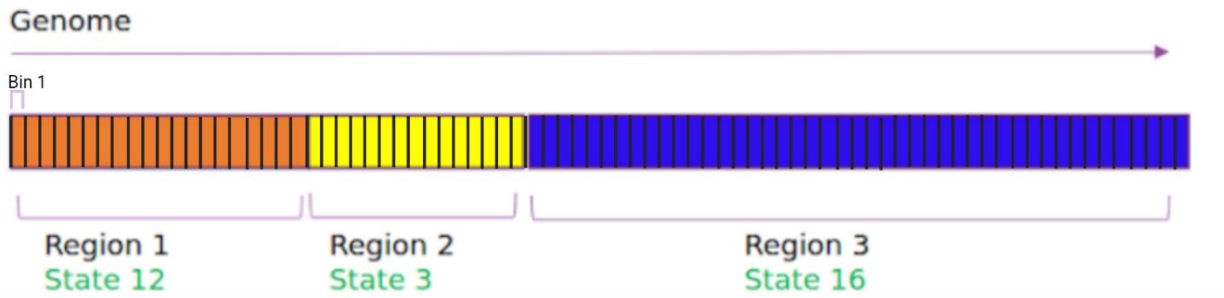


Figure 4. 6: Illustrative diagram of how allocation of epigenetic state across genomic loci was performed to generate chromatin state regions.

200 bp bins were classified as 1 of 20 chromatin states, and adjacent bins with the same state were grouped to form one chromatin state region. Only bins where state could be classified with >95% posterior probability were included. Figure is an original illustration.

We first chose to assess promoter regions. We defined our promoter regions as principal isoform, protein-coding TSSs +/- 1 kb; a 2kb window around the TSS was selected as optimal bin size based on observations that this 2kb window captured genome-wide H3K4me3 promoter enrichment peaks, which are commonly sharp but can also be more broadly enriched at some genes (e.g. Benayoun *et al.*, 2014; Chen *et al.*, 2015; Dahl *et al.*, 2016; Zhang *et al.*, 2016; Liu *et al.*, 2016; Lv and Chen, 2016). Using the R package GenomicRanges, I developed R scripts to categorise principal, protein-coding TSSs (+/- 1 kb) into one of the 20 epigenetic states. Explorative analyses were performed to assess the mapping of the mm9 genome by chromatin epigenetic state. **Figure 4.7** shows the number of protein-coding TSSs mapped as each of the 20 epigenetic states. The number of TSSs in each chromatin state varies considerably; for example, state 9, a heterochromatin-associated state accounting for a large percentage of the mm9 genome, encompasses the highest number of TSSs. These variations were considered when analysing differences in histone phosphorylation enrichments between states. It was also noted that the filtering of >95% confidence for chromatin state calling filtered out ~48% of the mm9 genome. Therefore, any subsequent chromatin state analysis was interpreted with the caveat of only referring to the remaining 52% of the genome where epigenetic state is confidently labelled.

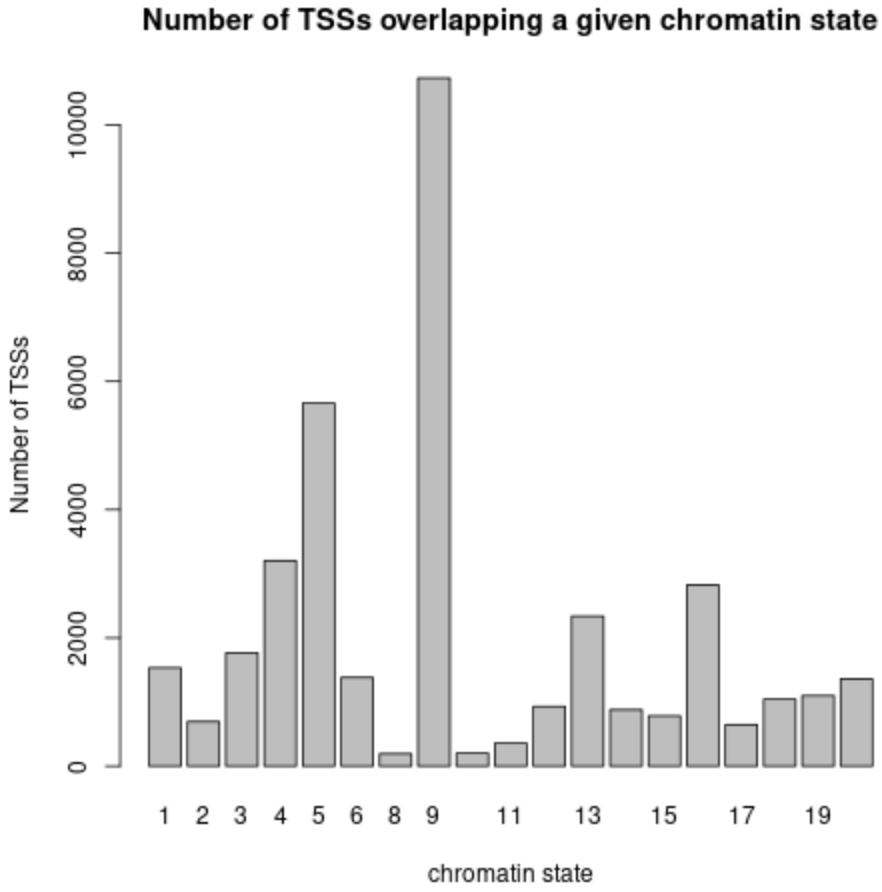


Figure 4. 7: Number of principal isoform, protein-coding TSSs called as each chromatin epigenetic state, based on Juan et al 2016 20-state ChromHMM model.

We then aligned the MINUTE-ChIP-seq normalised signal to the epigenetic state-labelled TSSs (+/-1kb), producing a data frame of mean normalised signal per state-labelled TSS, for each MINUTE-ChIP-seq epigenetic feature and each cell cycle phase. **Figure 4.8** describes this data frame. Mapping H3S10ph and H3S28ph MINUTE-ChIP-seq enrichment signal to the epigenetic state-mapped genome allowed us to subsequently quantitatively compare phosphorylation levels at promoters in different chromatin regulatory states.

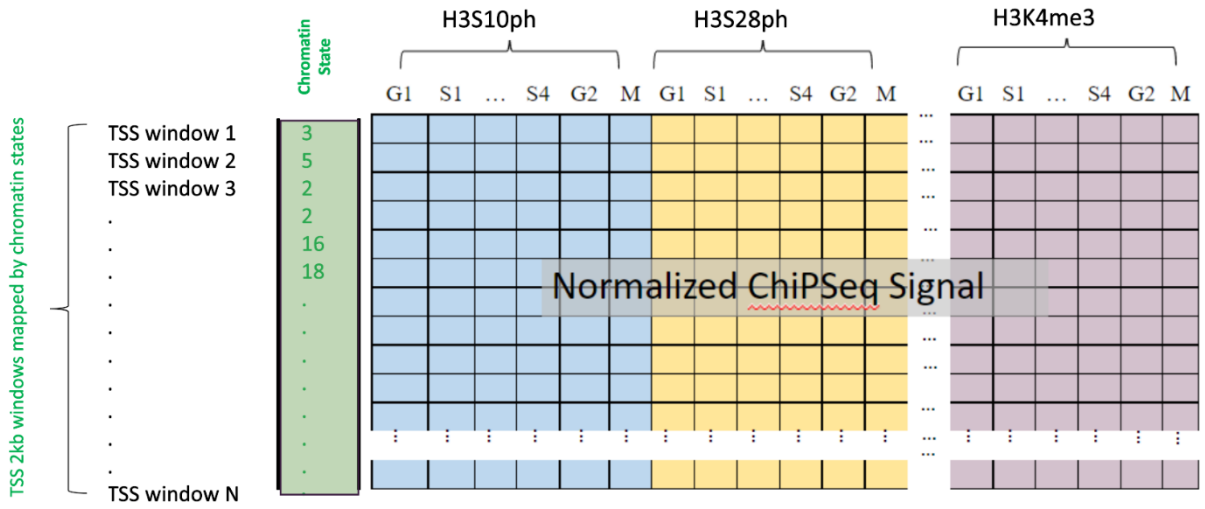


Figure 4. 8: The data frame generated of MINUTE-ChIP-signal for each chromatin state region.

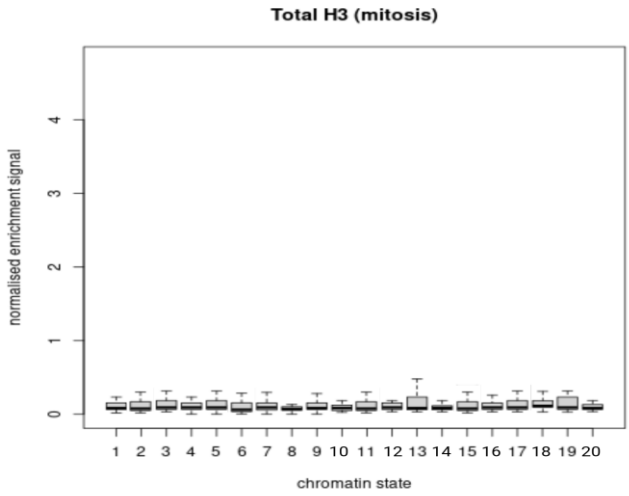
4.2.5 H3S10ph and H3S28ph show significant variation in enrichment depending on the chromatin regulatory state

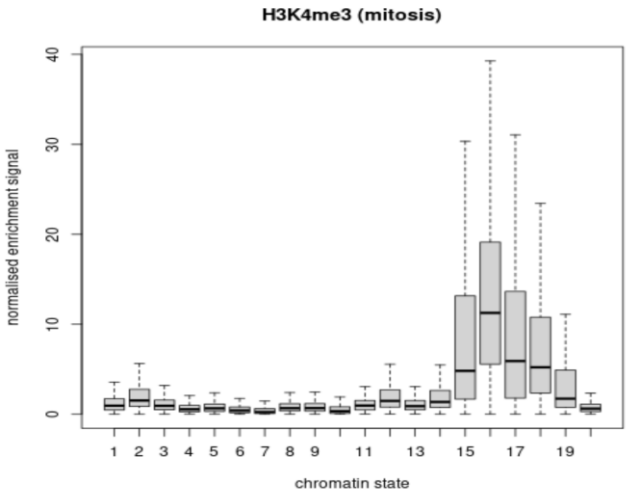
Once MINUTE-ChIP-seq signal had been quantified for each epigenetic state-labelled TSS (+/- 1 kb) (see **Figure 4.8**), the promoter-centred enrichment of a given histone modification could then be assessed in a specific chromatin regulatory state context. We were able to clearly observe using histone modifications with well-known regulatory localisations, enrichments that are specific to promoters in certain epigenetic states. **Figure 4.9** below shows how separating TSSs into epigenetic states reveals state-specific enrichments that were not observable in genome-wide analysis. H3K4me3 MINUTE-ChIP-seq, for example, shows dramatic enrichment increase in TSSs specifically in epigenetic states 15, 16, 17, 18 and 19; these are chromatin states associated with active or poised promoters, and were characterised by H3K4me3 enrichment during chromHMM model training by Juan *et al.*, 2016.

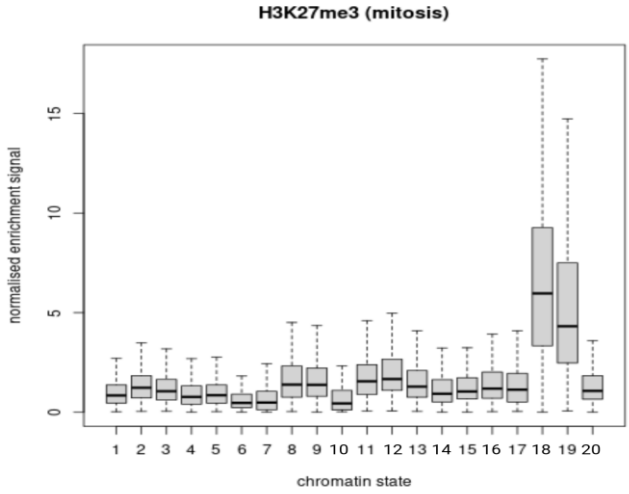
However, it is possible that all histone modifications might show increased enrichment in epigenetic states that are more open and accessible. To assess this, the repressive marker

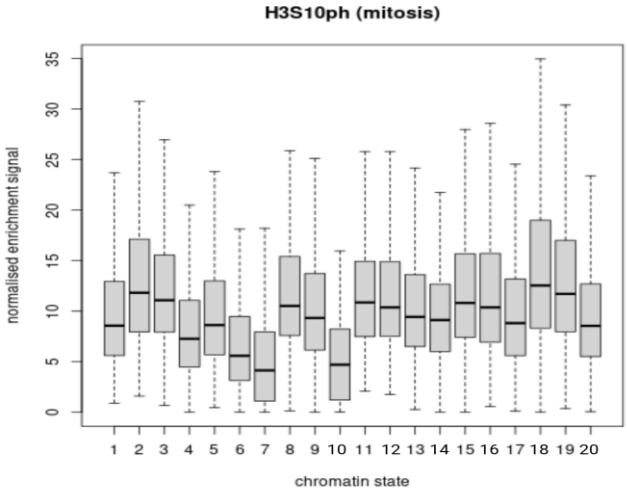
H3K27me3 was analysed as a second control. **Figure 4.9** illustrated H3K27me3 enrichment specifically at promoters in chromatin states 18 and 19, both of which are repressive polycomb-associated promoters, while H3K27me3 was not enriched in more open accessible states 15-17. Furthermore, total H3 showed uniform enrichment across the epigenetic states, shown in **Figure 4.9**, although it is noted that overall enrichment is lower for H3 than other ChIP experiments (note Figure 4.9 y-axis ranges).

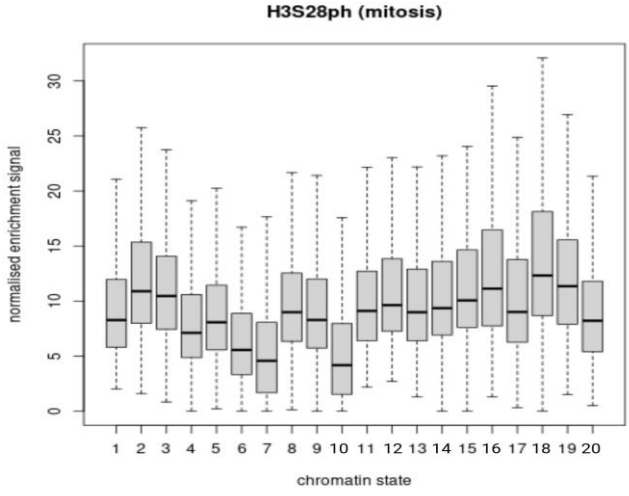
We can see from **Figure 4.9** that H3S10ph and H3S28ph signal at TSSs (+/- 1 kb) shows variation depending on the chromatin state.


Total H3 (mitosis)
normalised enrichment signal
chromatin state


H3K4me3 (mitosis)
normalised enrichment signal
chromatin state


H3K27me3 (mitosis)
normalised enrichment signal
chromatin state


H3S10ph (mitosis)
normalised enrichment signal
chromatin state


H3S28ph (mitosis)
normalised enrichment signal
chromatin state

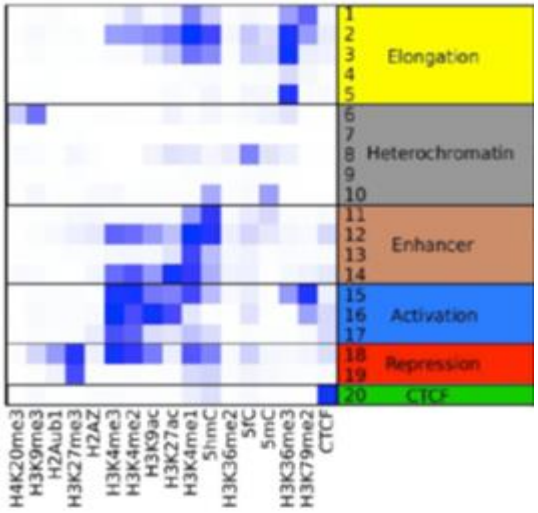

Elongation
Heterochromatin
Enhancer
Activation
Repression
CTCF
H4K20me3
H3K9me3
H2Aub1
H3K27me3
H2AZ
H3K4me3
H3K4me2
H3K9ac
H3K27ac
H3K4me1
5hmC
H3K36me2
5fC
5mC
H3K36me3
H3K79me2
CTCF

Figure 4. 9: MINUTE-ChIP-seq mitotic enrichment signal varies significantly between different chromatin state regions.

Enrichment signals when TSSs (+/-1 kb) are separated into different epigenetic states 1-20 (x axis). Note the differences in y axis range. For reference, bottom right shows the ChromHMM 20-state model and the relative enrichments of regulatory markers used to train the model, from Juan et al 2016.

Having established this method of integrating MINUTE-ChIP-seq quantitative signal with chromatin regulatory state labelling, we next analysed whether variation in H3S10ph and H3S28ph between promoters in different chromatin states was statistically significant. Non-parametric Kruskal-Wallis test found that there was significant difference in both mitotic H3S10ph and H3S28ph between at least two of the 20 chromatin states. Post-hoc Dunn tests were then performed to identify which epigenetic states showed significant difference in promoter H3S10ph and H3S28ph enrichment. It was determined that almost all 20 epigenetic states showed significant difference in mitotic H3S10ph and mitotic H3S28ph enrichment at promoters. The results of Kruskal-Wallis tests are shown below, **Table 4.1,** with tests for H3, H3K4me3 and H3K27me3 also shown for comparison**.** Post-hoc Dunn test results can be seen in Appendix Figure A4.

Table 4. 1: Kruskal-Wallis test for significant variation in normalised signal between TSSs (+/- 1 kb) in different chromatin states. Note that higher chi-squared values indicate a larger difference

ChIP	chi-squared	df	p
H3S10ph (M)	29901	19	$2.2e^{-16}$
H3S28ph (M)	27153	19	$2.2e^{-16}$
H3 (M)	1875	19	$2.2e^{-16}$
H3K4me3 (M)	9695	19	$2.2e^{-16}$
H3K27me3 (M)	6139	19	$2.2e^{-16}$

These findings clearly demonstrate significant variation in mitotic H3S10 and H3S28 phosphorylation enrichment at promoters depending on the TSS's chromatin regulatory state, which to the best of our knowledge has not previously been seen in genome-wide distribution studies. We therefore argue against H3S10ph or H3S28ph being distributed uniformly genome-wide, but instead suggest significant patterns in H3S10ph and H3S28ph correlating with chromatin regulatory landscape.

4.2.6 Co-localisation networks reveal significant interplay between histone phosphorylations and regulatory markers that vary depending on both cell cycle phase and chromatin regulatory state.

MINUTE-ChIP-seq experiments were performed by Simon Elsasser's research group across a range of histone modifications and regulatory markers as well as the histone phosphorylations analysed above. Thanks to the quantitative nature of MINUTE-ChIP-seq, we were able to

statistically assess co-localisation of regulatory markers with histone phosphorylations H3S10ph and H3S28ph across the cell cycle. This provided valuable insight into potential regulatory interactions of mitotic histone phosphorylations, and provides further support that H3S10ph and H3S28ph regulatory interactions and enrichments may be revealed by examining the chromatin regulatory context.

Partial correlation was chosen to assess co-localisation as opposed to pairwise correlation. Pairwise correlation measures the correlation of two factors, but is unable to account for other indirect correlations in calculations; if A correlates with B, and B correlates with C, then A will likely correlate with C, but the importance of each correlation is more difficult to interpret. Therefore, pairwise correlations between chromatin regulatory markers could be strongly impacted by the myriad of other markers present. This is illustrated in **Figure 4.10** (left) using an example selection of histone modifications, transcription factors and regulatory proteins, all with well-documented regulatory roles and interactions, in G1 cell samples. A high number of correlations is seen, but it is near-impossible to interpret how each correlation might be indirectly affecting the others. Contrastingly, partial correlations calculate the correlation between two factors *while accounting for indirect correlations.* Partial correlations were therefore calculated between each two regulatory markers, while factoring in the indirect impacts of other regulatory markers. Partial correlation networks greatly reduce the "noise" of a correlation network, only showing significant, direct correlations after the impact of indirect correlations has been controlled for.

As shown in the right panel of **Figure 4.10**, this allowed us to more confidently isolate significant direct correlations between MINUTE-ChIP targets. For example, a significant strong positive partial correlation is seen between genome-wide Ezh2, a subunit of the polycomb complex, and H3K27me3, a histone modification deposited by polycomb in repressive chromatin regions. H3K9me3, a heterochromatin-enriched histone modification, is seen to negatively correlate with H3K4me3, a euchromatin-enriched histone modification associated with active gene expression. These partial correlations which support the documented functional relationships between these markers, are not seen in the pairwise correlation network shown on the right.

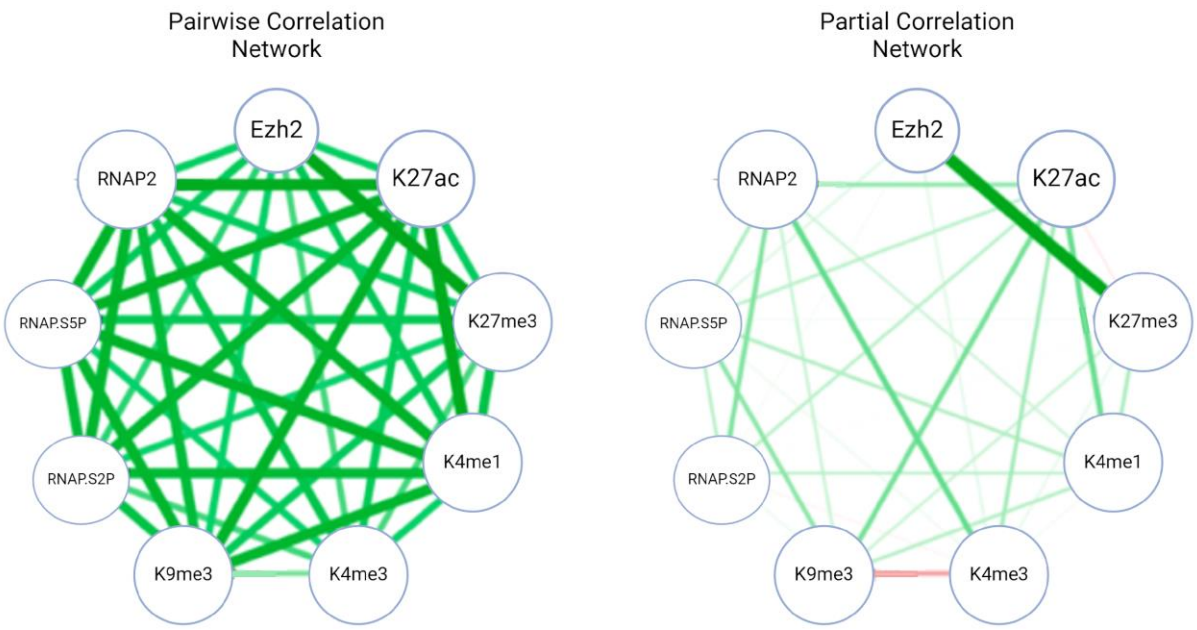


Figure 4. 10: Pairwise (left) vs Partial (right) correlation networks for a set of example regulatory markers in G1 cells.

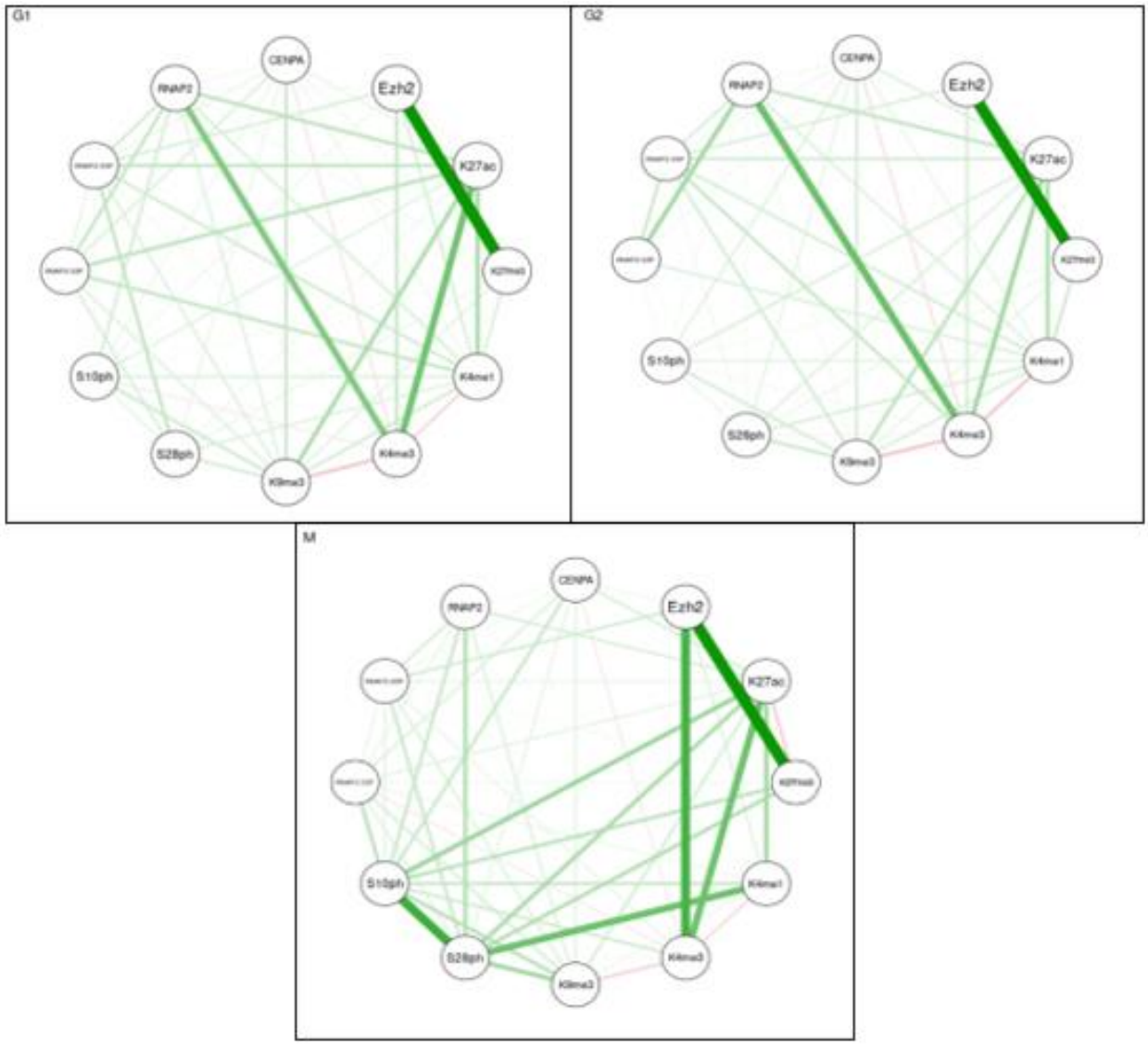
MINUTE-ChIP-seq genome-wide normalised signal at TSSs(+/-1kb) in G1 cell samples for a set of histone modifications, transcription factors and regulatory proteins were used to calculate the pairwise correlation coefficient between each two markers, plotted in the network below, left. Partial correlations were also calculated between each marker and are plotted in a network below, right. Plotted correlations are filtered to only significant correlations ($p<0.05$). Green = positive correlation, Red = negative correlation. Edge width represents correlation coefficient value, indicating correlation strength. Abbreviations: K27ac = histone H3 Lysine 27 acetylation; K27me3 = histone H3 lysine 27 trimethylation; K4me1/3 = histone H3 lysine 4 mono/trimethylation; K9me3 = histone H3 lysine 9 trimethylation; RNAP.S2P/S5P = RNA polymerase 2 phosphorylated on serine 2/5; RNAP2 = RNA polymerase 2; Ezh2 = polycomb complex subunit Ezh2.

MINUTE-ChIP-seq normalised signal at TSS (+/- 1 kb) windows were used to calculate partial correlations. We continued to focus on TSSs in order to further elucidate the regulatory landscape surrounding H3S10ph and H3S28ph in these promoter regions. We note that other co-localisations

may well occur with potential significance elsewhere along the genome, and these could be studied in future work.

Partial correlations were measured across a range of MINUTE-ChIP-ed targets with known regulatory roles, and H3S10ph and H3S28ph. The partial correlation coefficients, a measure of correlation strength, were plotted in co-localisation networks, filtering to show only those correlations that were significant ($p<0.05$). First, the correlations were plotted using MINUTE-ChIP enrichment signal genome-wide in arbitrary 2kb windows; resulting genome-wide co-localisation networks are depicted below, **Figure 4.11**. Expected co-localisations between regulatory markers can be seen in each cell cycle phase; for example, repressive polycomb subunit Ezh2 shows positive correlation with repressive marker H3K27me3. H3K4me3 shows co-localisation with transcription enzyme RNA polymerase 2 (RNAP2) during interphase, and with H3K27ac across the cell cycle; this makes sense given both H3K4me3 are known to enrich at promoter and enhancer regions. This genome-wide network demonstrates variation in the regulatory landscape between G1, G2 and Mitosis. For example, the H3K4me3-RNApolII co-localisation is lost in Mitosis, likely because RNApolII largely dissociates from mitotic chromosomes (Contreras and Perea-Resa, 2024).

Focussing on histone phosphorylations, numerous co-localisations were identified at all protein-coding TSSs genome-wide, that were specific to certain cell cycle phases. In mitosis, where H3S10ph and H3S28ph are by far the most enriched, H3S28ph and H3S10ph show strong positive correlation. This might be expected given both are deposited by the mitotic kinase Aurora B. Mitotic H3S10ph shows significant positive correlation ($P<0.05$) with marks associated with both repression and activation regulatory pathways, including H3K27ac, H3K4me1, H3K4me3, H3K27me3 and H3K9me3. Mitotic H3S28ph shows the strongest positive correlation with H3K4me1 followed by H3K27ac, both of which are typically enriched at enhancers. H3S28ph also co-localises with H3K9me3 and H3K27me3 during mitosis. The majority of the significant TSS co-localisations seen for histone phosphorylations in mitosis are weaker in G2 samples; however, significant interactions are still evident, and appear slightly stronger in G1 – it is conceivable some of these co-localisations could remain on mitotic exit into G1.

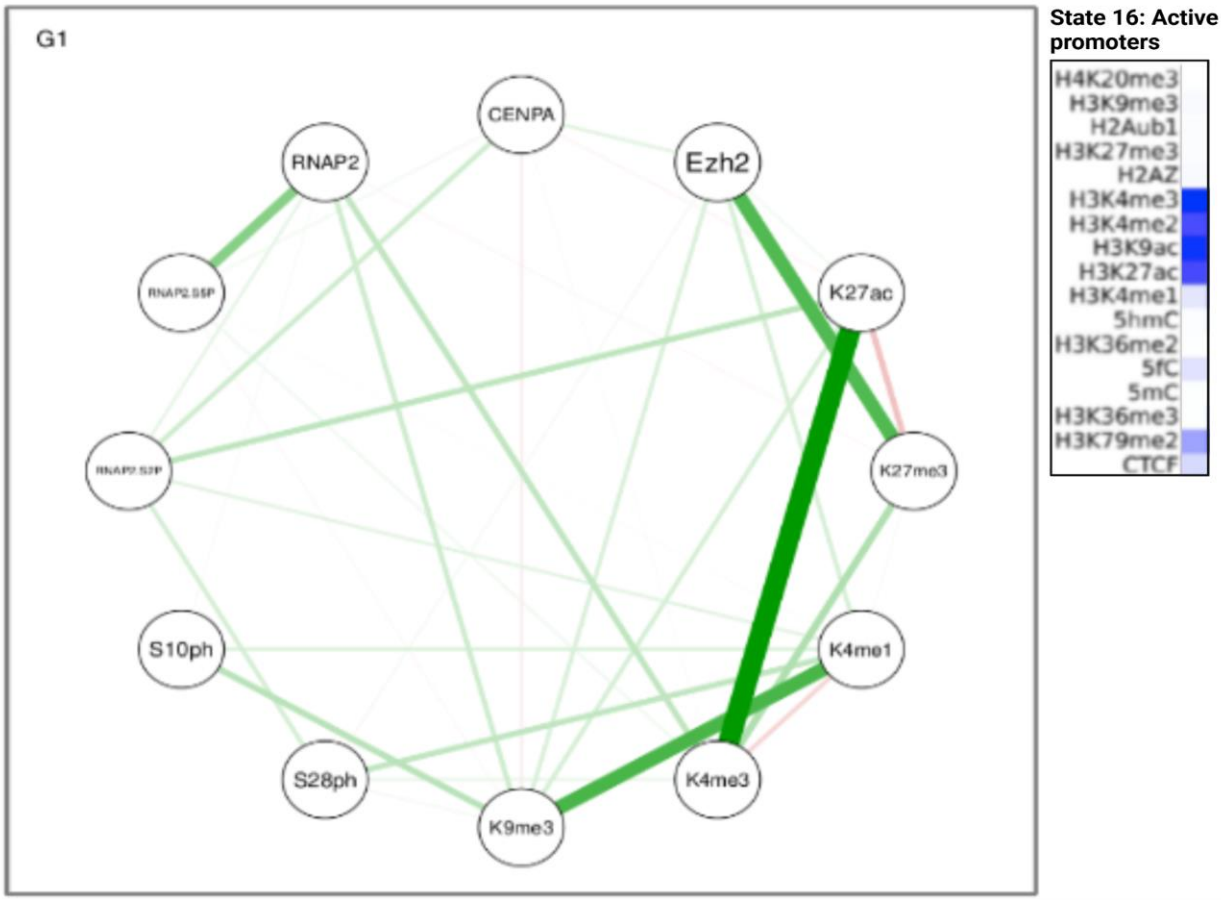


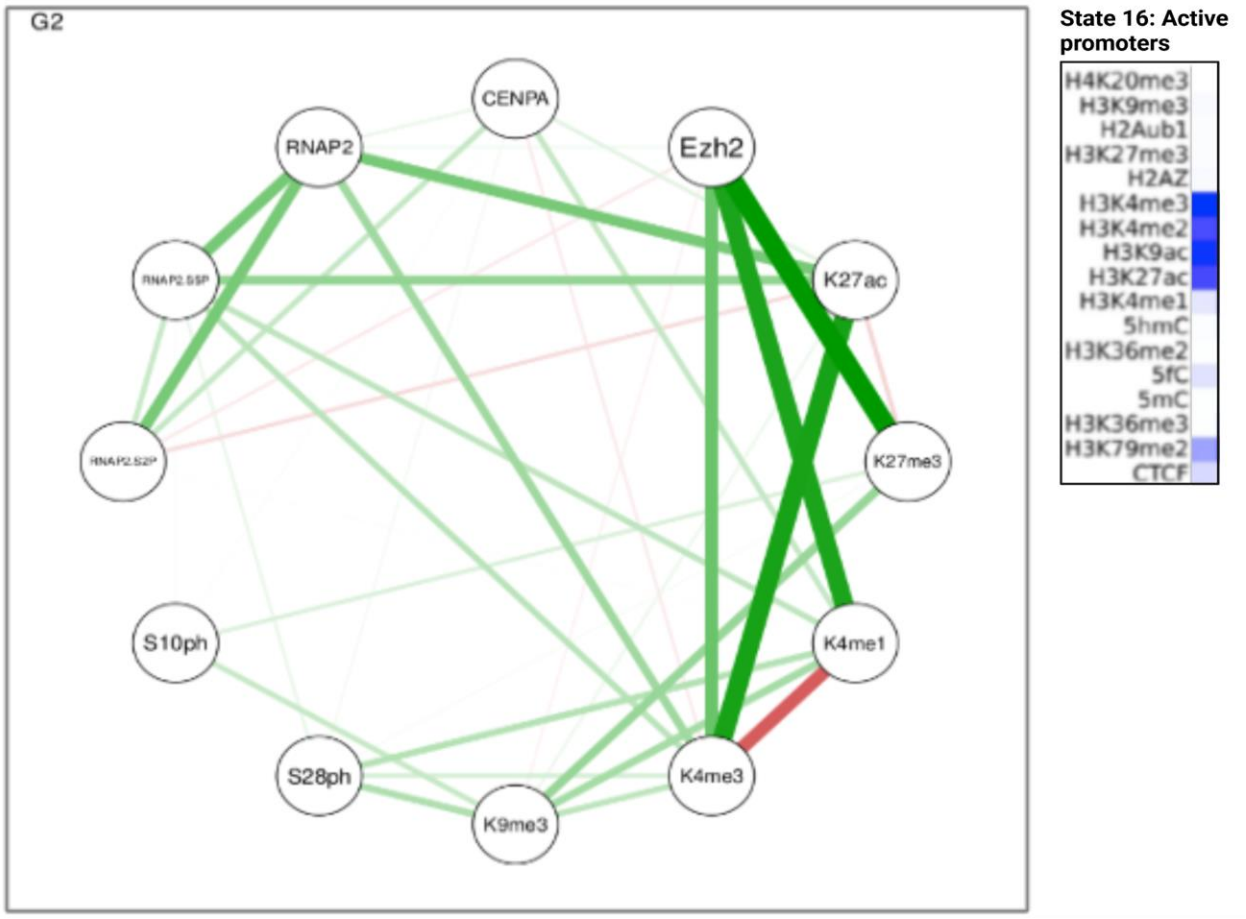
Figre 4. 11: Partial correlation networks of histone phosphorylations H3S10ph, H3S28ph and regulatory chromatin markers at all TSSs (+/- 1 kb), across the cell cycle.

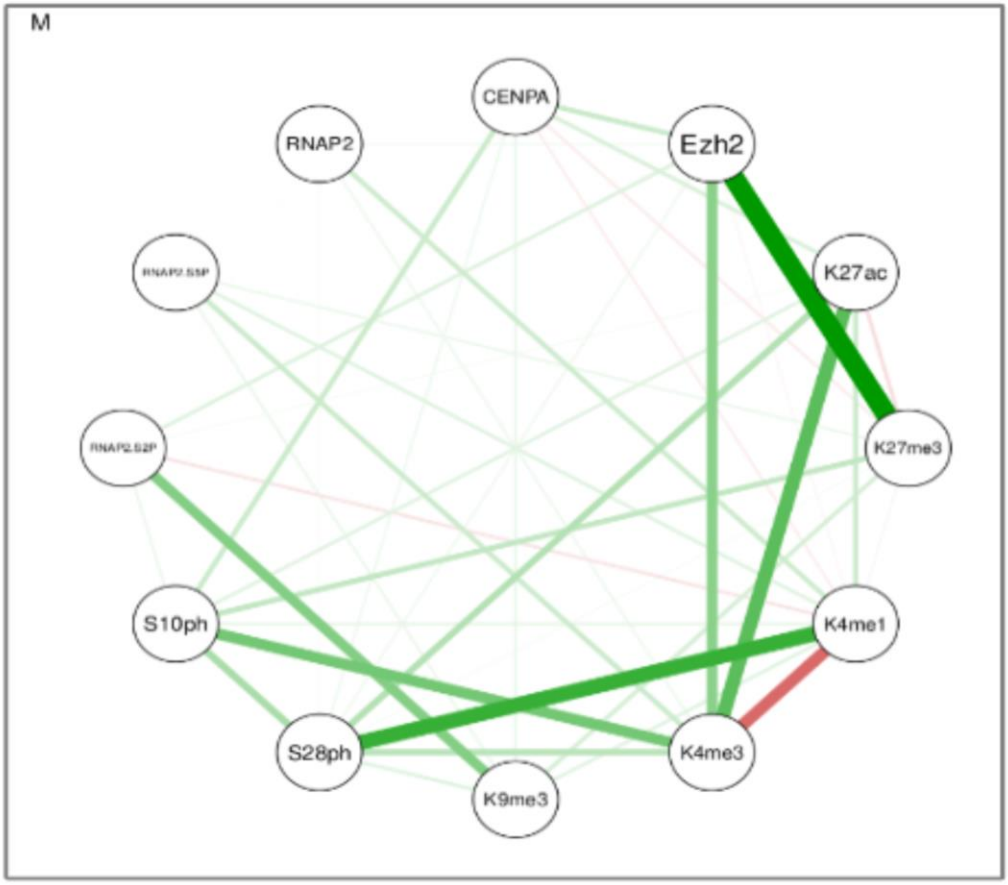
A set of histone modifications, transcription factors and regulatory proteins were selected and the partial correlation coefficient between each two markers was calculated, using MINUTE-ChIP-seq normalised enrichment signal at all protein-coding TSSs (+/- 1 kb windows). Plotted correlations are filtered to only significant correlations ($p<0.05$). Green = positive correlation, Red = negative correlation. Edge width indicates correlation coefficient value, indicating correlation strength. Networks were generated across the cell cycle: Left = G1, Centre= G2, Right= Mitosis(M). Nodes, clockwise from top: CENPA, Ezh2, H3K27ac (K27ac), H3K27me3 (K27me3), H3K4me1 (K4me1), H3K4me3 (K4me3), H3K9me3 (K9me3), H3S10ph (S10ph),

H3S28ph (S28ph), RNApolII-phosphorylated at serine 2 (RNAP2.S2P), RNApolII-phosphorylated at serine 5 (RNAP2.S5P, RNApolII (RNAP2)

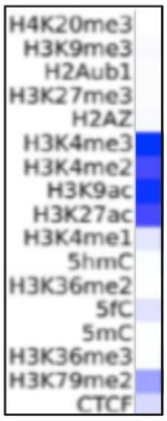
We then wanted to assess how co-localisations of histone phosphorylation vs regulatory markers changed depending on the chromatin state of promoters. Partial correlation coefficients were calculated for subsets of MINUTE-ChIP signal at TSS (+/- 1 kb) in each chromatin state. Our aim was to determine whether histone phosphorylations co-localise with different regulatory markers depending on the chromatin regulatory state. The resulting partial correlation networks are presented below in **Figure 4.12,** showing co-localisations in two example states: active promoter chromatin state 16, and bivalent promoter chromatin state 18. G1, G2 and Mitosis networks are presented for comparison. Again, networks were filtered so that only statistically significant ($p<0.05$) correlations are plotted, with edge weight indicating strength of correlation.

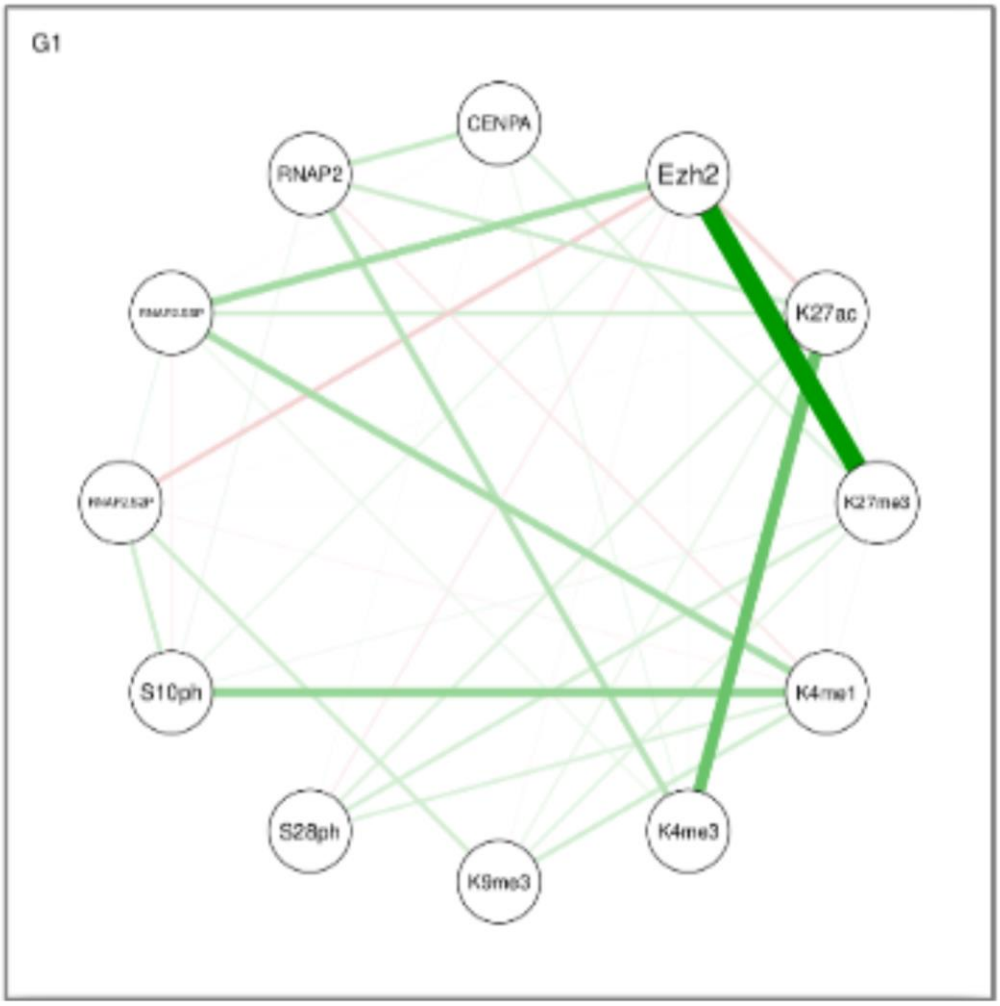



G2
CENPA
RNAP2
Ezh2
RNAP2.S5P
K27ac
RNAP2.S2P
K27me3
S10ph
K4me1
S28ph
K4me3
K9me3
State 16: Active promoters
H4K20me3
H3K9me3
H2Aub1
H3K27me3
H2AZ
H3K4me3
H3K4me2
H3K9ac
H3K27ac
H3K4me1
5hmC
H3K36me2
5fC
5mC
H3K36me3
H3K79me2
CTCF

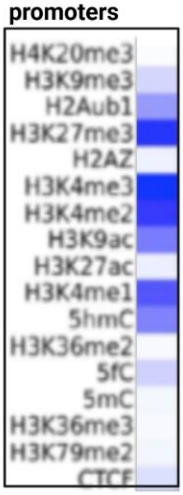


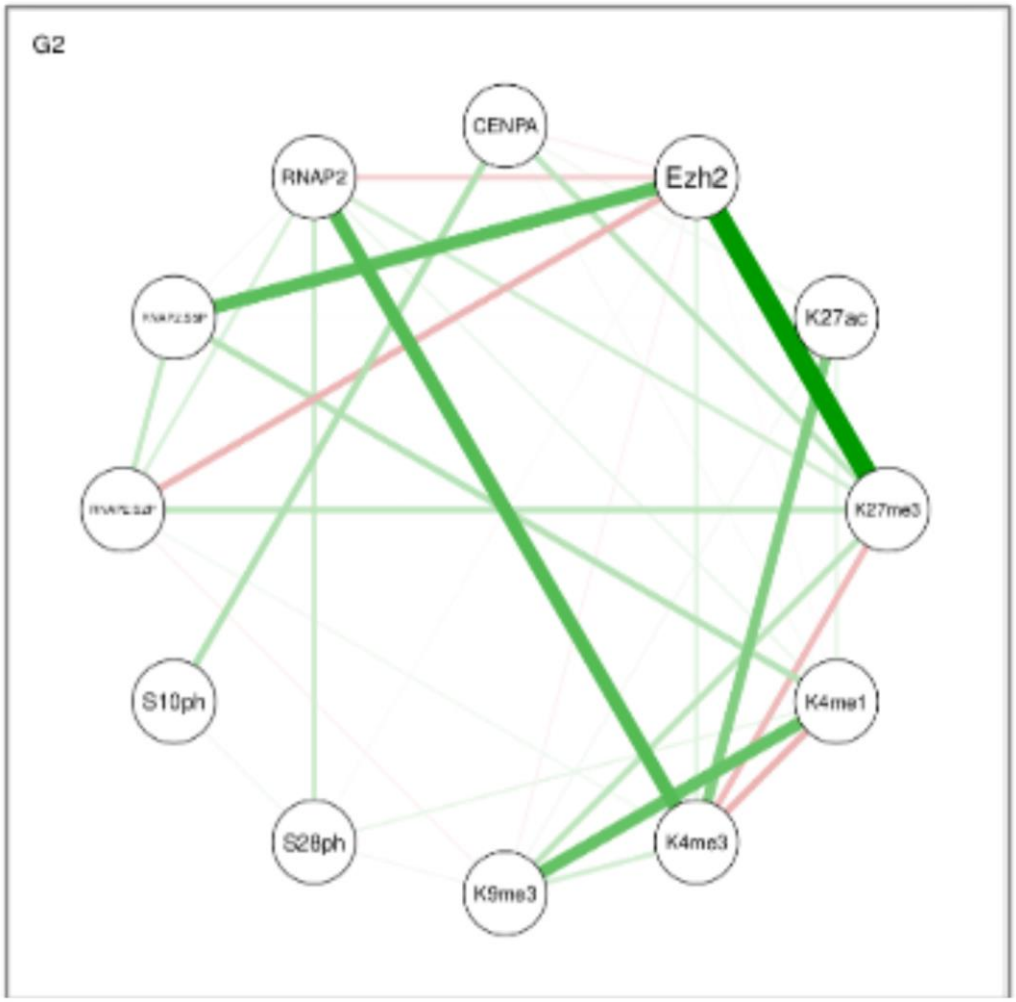
State 16: Active promoters



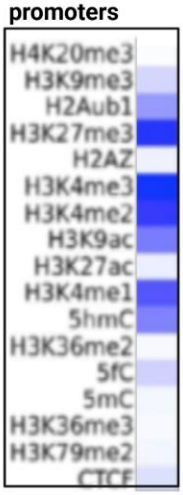


State 18: Bivalent promoters





State 18: Bivalent promoters



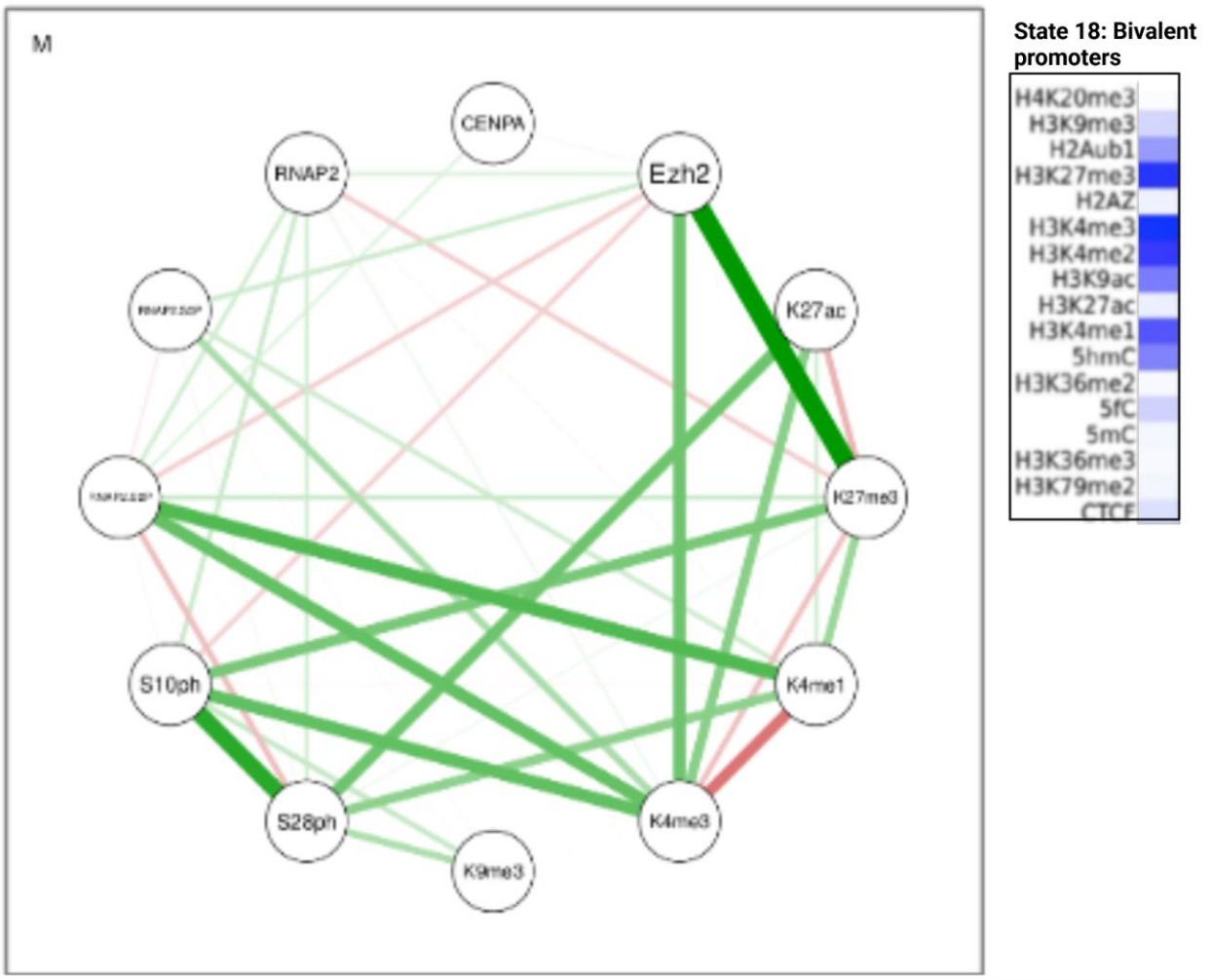


Figure 4. 12: Partial correlation networks of histone phosphorylations H3S10ph and H3S28ph and regulatory chromatin markers, across the cell cycle, focussing on transcription start sites (+/- 1 kb) in chromatin epigenetic states 16 and 18.

A set of histone modifications, transcription factors and regulatory proteins were selected and the MINUTE-ChIP-seq enrichment signal for transcription start sites (TSSs) +/-1 kb measured. From top moving clockwise, nodes are: CENPA, Ezh2, H3K27ac, H3K27me3, H3K4me1, H3K4me3, H3K9me3, H3S28ph, H3S10ph, RNApolII-S2P, RNApolII-S5P, RNApolII. A subset of TSSs identified as chromatin regulatory state 16 (top three networks) and 18 (bottom three networks) were isolated, and the partial correlations between all markers calculated for G1, G2 and Mitosis (M) samples. Only significant correlations ($p<0.05$) are plotted. Green = positive correlation, Red = negative correlation. Edge width indicates correlation coefficient value, indicating correlation strength. Heatmaps showing relative enrichment of regulatory markers used to train the chromatin state model are shown for state 16/18 to the right of each network, for reference.

These partial correlation networks illustrate that histone phosphorylations co-localise with different regulatory markers depending on a) the epigenetic state of promoters, and b) the cell cycle phase. Interpreting all of these co-localisations was not within the scope of this thesis, and so a few co-localisations of interest were selected for each histone phosphorylation, discussed below:

H3S10ph co-localisations

H3S10 phosphorylation was found to show different co-localisations depending on both chromatin state and cell cycle phase. At TSSs in epigenetic state 16, the state representing actively expressed promoter regions, in G1, H3S10ph showed significant ($p<0.05$) positive correlation with H3K4me1 and H3K9me3. In G2, H3K4me1 co-localisation is lost, and a new weak but significant co-localisation with H3K27me3 is seen. In mitosis, H3S10ph shows strong positive co-localisation with H3K4me3, but also co-localises with polycomb-associated H3K27me3 and, interestingly, the centromeric protein CENPA.

Focussing then on epigenetic state 18 TSSs, representing bivalent promoters, H3S10ph shows significant ($p<0.05$) positive correlation with H3K4me1 in G1, and some correlation with RNApolII-S2P. In G2, these co-localisations are lost, and H3S10ph instead co-localises with CENPA. Numerous strong correlations are then seen in Mitosis; H3S10ph positively correlates with H3S28ph, as well as H3K4me3 and H3K27me3 - two modifications important for the bivalent state of these promoters. H3S10ph also correlates more weakly but significantly ($p<0.05$) with RNA polymerase II in mitosis. Finally, mitotic H3S10ph anti-correlates with polycomb complex subunit Ezh2.

H3S28ph co-localisations

In epigenetic state 16, H3S28p co-localised with H3K4me1 and RNApolII-S2P in G1, although relatively weakly. In G2, H3S28ph continues to co-localise with H3K4me1, and also correlates with H3K9me3. Mitotic state 16 TSSs then show relatively strong positive correlation between H3S28ph and H3K4me1, as well as new mitosis-specific correlations with H3K4me3 and H3K27ac, modifications associated with transcriptional activation.

Bivalent promoters in state 18 show weak co-localisations for H3S28ph in G1 and G2; for example, H3K4me1, H3K27me3 and H3K27ac in G1, and RNApolII in G2. In mitosis, H3S28ph then strongly co-localises with H3K27ac and H3K4me1, and anti-correlates with RNApolII-S2P. **Table 4.2** below is provided to aid visualisation of these histone phosphorylation co-localisations with regulatory markers through the cell cycle. Markers associated with activation or repression are highlighted in orange and blue respectively to aid understanding.

Table 4. 2: Co-localisations of histone phosphorylations with regulatory markers through the cell cycle.

H3S10ph (left) and H3S28ph (right) co-localisations at TSSs in epigenetic state 16 (top) and state 18 (bottom) were identified using partial correlation coefficients. Table shows regulatory markers found to have significant ($p<0.05$) positive or negative correlations with either histone phosphorylation in G1 (top), G2 (centre) or Mitotic (bottom) cells. Marks associated with activation or repression are highlighted as orange or blue respectively

State 16				
Phase	**S10ph co-localisations**		**S28ph co-localisations**	
	Positive	**Negative**	**Positive**	**Negative**
G1	**H3K4me1** **H3K9me3**	-	**H3K4me1** **RNAP2.S2P**	-
G2	**H3K9me3** **H3K27me3**	-	**H3K9me3** **H3K4me1**	-
M	**H3K4me3** H3S28ph **H3K27ac** **H3K27me3** CENPA	-	**H3K4me3** **H3K4me1** **H3K27ac** H3S10ph	-

State 18				
Phase	**S10ph co-localisations**		**S28ph co-localisations**	
	Positive	**Negative**	**Positive**	**Negative**
G1	**H3K4me1** **RNAP2.S2P**		**H3K4me1** **H3K27me3** **H3K27ac**	**Ezh2**
G2	CENPA		**RNAP2** **H3K4me1**	
M	**H3K9me3** H3S28ph **H3K4me3** **H3K27me3** **RNAP2**	**Ezh2**	**H3K9me3** H3S10ph **H3K27ac** **H3K4me1**	**RNAP.S2P**

It was noted that the partial correlation networks separated by epigenetic state, generated to address our research questions, also revealed numerous significant ($p<0.05$) co-localisations between other regulatory markers that are only observed once the genome is separated by epigenetic state. For example, our correlation networks identify significant negative correlation between H3K4me3 and H3K4me1, which occurs more strongly in certain epigenetic states, such as state 16. Also, this negative correlation is strongest in G2 and Mitosis but not present in G1, suggesting that while these marks remain at consistent levels genome-wide through cell cycle, their interactions and specific co-localisations vary depending on cell cycle phase.

These partial correlation networks categorised by epigenetic state and cell cycle phase likely offer many contributions to knowledge of regulatory mark distributions and interactions, and how these change with cell cycle. As such, we generated partial correlation maps as seen above for all cell cycle phases G1, G2 and Mitosis, for each of the 20 chromatin states defined by Juan *et al.*, 2016. These were generated both using MINUTE-ChIP signal at TSSs +/- 1 kb to visualise promoter co-localisations, and using MINUTE-ChIP signal genome-wide separated into windows based on

chromatin state regions. These partial correlation networks are outside of the focus of this thesis but will contribute to future analysis and publication in collaboration with Dr Simon Elsasser and colleagues.

We concluded from these analyses that at mitotic promoter regions, both H3S10ph and H3S28ph show significant variation in both enrichment and co-localisation with other regulatory markers, depending on the chromatin regulatory state and the cell cycle phase. We highlight the importance of taking into account the regulatory landscape as a crucial context when assessing the enrichment patterns of these abundant mitotic histone phosphorylations.

4.2.7 Mitotic H3S10ph and H3S28ph show significant enrichment at promoters in specific chromatin regulatory states

As mentioned previously, both H3S10ph and H3S28ph have been reported to enrich at promoters during interphase, and we have demonstrated above that both phosphorylations significantly vary at promoters in mitosis depending on the chromatin regulatory state. We have also shown that it is important to take into account the chromatin regulatory state in order to identify genuine enrichments in histone PTMs. We therefore hypothesised that H3S10ph and H3S28ph may be significantly enriched, and/or have "peaks" in distribution, at promoters in specific epigenetic states, as was observed for H3K4me3 and H3K27me3.

We first sought to test whether H3S10ph or H3S28ph show higher enrichment at promoters compared to other genomic regions during mitosis. Although significant variation in mitotic H3S10ph and H3S28ph is seen between promoters in different regulatory states, this does not confirm whether mitotic H3S10ph and H3S28ph is higher at TSSs (+/- 1 kb) than elsewhere on the genome, such as the gene body or non-genic regions.

To address this question, we used MACS2 peak calling software to transform the continuous MINUTE-ChIP-seq signal into a binary absence/presence classification. MACS2 software scans along the genome using a 200 bp sliding window/bin, and using statistical algorithms classifies each bin as either "present" or "absent" of sequencing signal, **relative to neighbouring sequence.** MACS2 binarisation therefore allowed us to statistically isolate genomic loci (i.e. bins) with highest H3S10ph/H3S28ph signal relative to other regions: these high H3S10ph or H3S28ph bins

were defined as “MACS2 called peaks” for subsequent analyses. Note that MACS2 was performed without providing an optional Input sample, as Elsasser’s team had already performed Input normalisation as described above. MACS2 thus allowed us to classify loci with highest H3S10ph and H3S28ph in a systematic and statistically sound way.

We then analysed whether either H3S10ph or H3S28ph MACS2-called peaks occurred significantly more in promoter regions compared to the rest of the genome. The mouse mm9 genome, taken to be 2,725,521,370 bp long, was divided arbitrarily into 2 kb windows, giving 1,362,760 windows in total. Based on Ensembl Biomart data (see Chapter 2 for version details) we estimated there are 23,010 (principal isoform, protein-coding) TSSs in the mm9 genome. We therefore calculated that these TSSs are present in 1.68% of these arbitrary 2 kb windows across the genome.

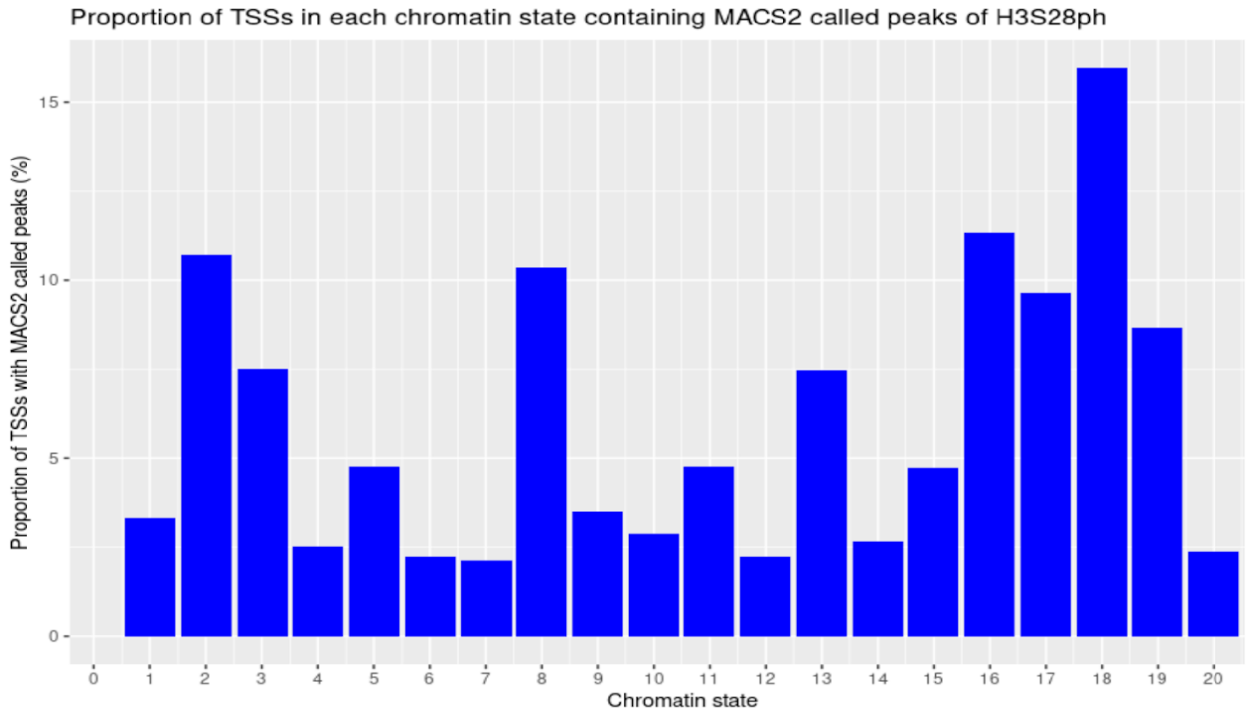
In mitosis, MACS2 identified a total 192,611 significant H3S10ph “called peaks” (i.e. highest H3S10ph bins relative to other genomic loci) genome-wide. Based on the above assumptions, we would therefore expect, if H3S10ph were randomly distributed, that 1.68% of those peaks (i.e. 3,252) would occur at TSSs (+/- 1 kb). Our actual findings calculated that 10,747 H3S10ph peaks occurred within TSS 2 kb windows, meaning 5% of H3S10ph peaks occurred at TSSs - more than would be expected for a random peak distribution. For comparison, 3117 (1.62%) H3S10ph called peaks occurred at TESs (+/- 1 kb), similar to the 1.68% expected at random.

Following the same calculations for H3S28ph, MACS2 identified 198,843 significant H3S28ph peaks genome-wide in mitosis, meaning we would expect if H3S28ph were randomly distributed that 1.68% of those peaks, i.e. 3,341 would occur at a TSS (+/- 1 kb). Our actual findings showed that 8,617 H3S28ph called peaks occurred within TSS 2kb windows, 4.33% of all H3S28ph peaks. In contrast, 2,150 H3S28ph peaks (1.18% of all peaks) occurred at TESs.

We therefore argue that H3S10ph and H3S28ph enrich significantly at promoters compared to other genomic regions, more than would be expected given a random uniform distribution.

In the previous Section, we described how separating TSSs based on epigenetic state revealed significant variation in both H3S10ph and H3S28ph enrichment. Therefore, we hypothesised that MACS2-called H3S10ph and H3S28ph peaks may be more prevalent in specific epigenetic states. To address this hypothesis, TSSs were categorised into epigenetic states, as detailed previously in Section 4.2.4. The proportion of TSSs in each chromatin state containing MACS2 called peaks was then calculated, depicted below in **Figure 4.13**.

Variation was seen in the proportion of TSSs containing MACS2-called peaks per epigenetic state. Epigenetic state 18, proposed to represent bivalent promoters, contained the highest proportion of both H3S10ph and H3S28ph called peaks. A higher proportion of TSSs with H3S10ph and H3S28ph called peaks were also seen in state 2, associated with elongation. A high proportion of TSSs in active promoter state 16 contained H3S28ph peaks, while H3S10ph peaks were seen more frequently in states 15 and 17, thought to represent poised active promoters. State 6 had the lowest proportion of H3S10ph peaks; this state is characterised by the serine10-adjacent modification H3K9me3.



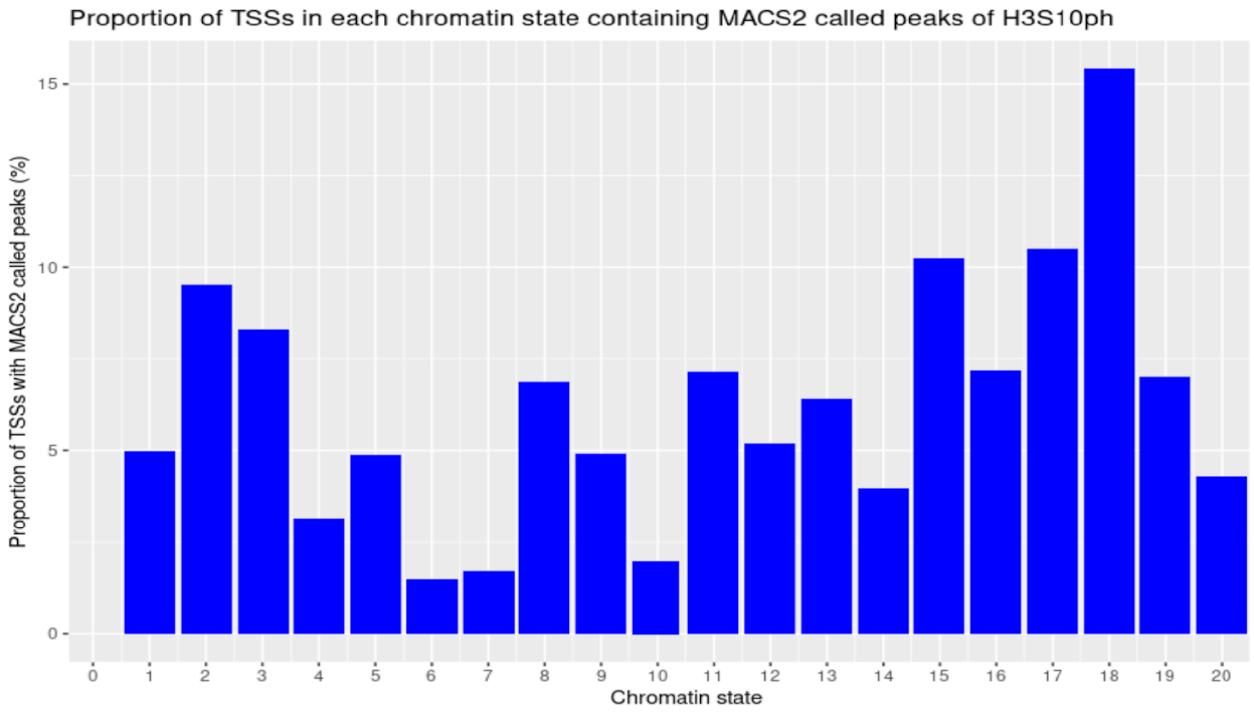


Figure 4. 13: The proportion (%) of TSSs in each defined chromatin epigenetic state 1-20 which contain MACS2-called mitotic H3S28ph(top***) and H3S10ph(***bottom***) peaks.***

TSSs (+/- 1 kb) were labelled as chromatin epigenetic states 1-20 based on the trained ChromHMM 20 state model generated in Juan et al., (2016). MACS2 was used to call significant peaks for mitotic H3S10ph and H3S28ph, shown in separate plots. The proportion of TSSs in a given state that overlap with at least one H3S10ph or H3S28ph called peak were plotted.

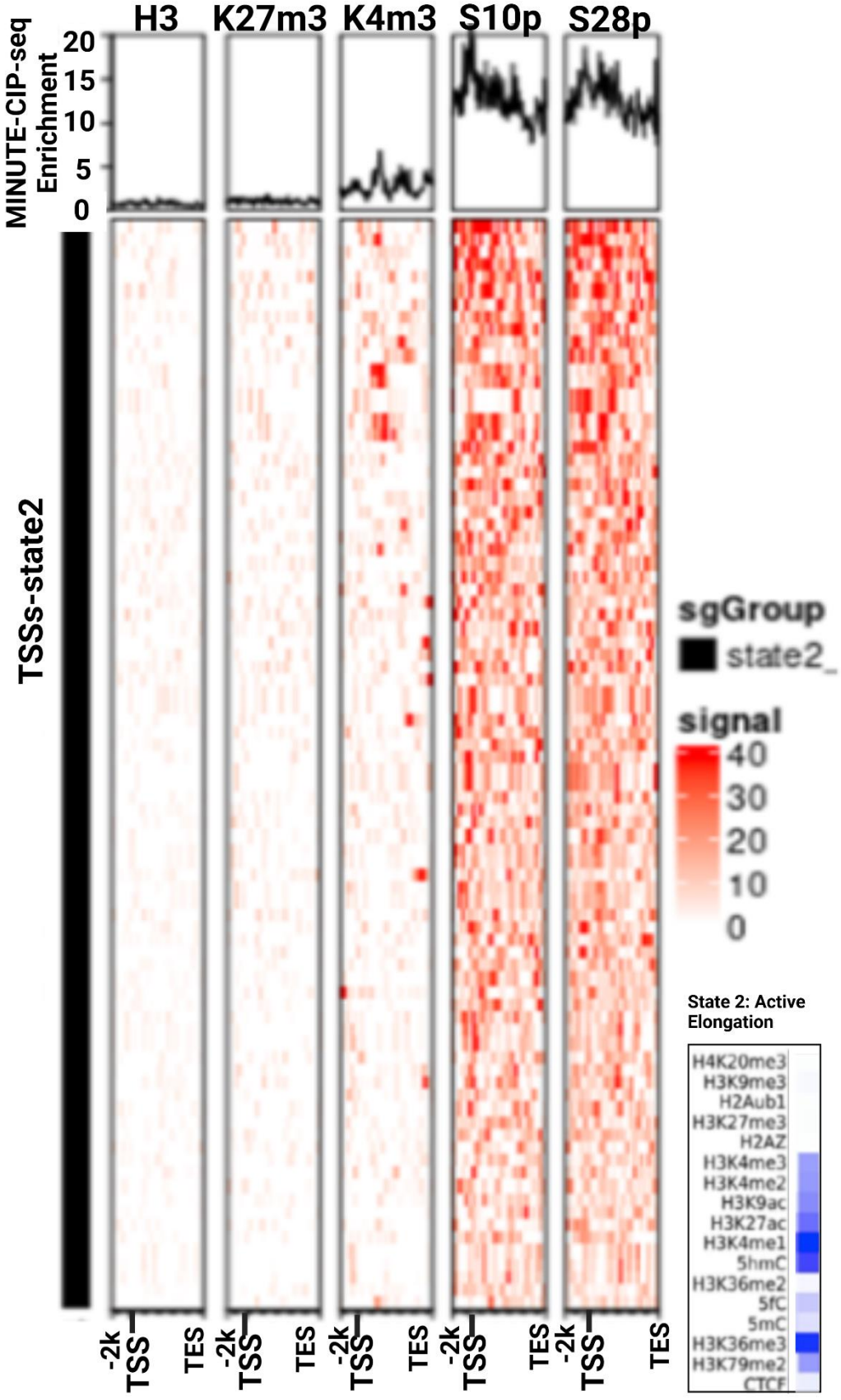
To consolidate our findings of H3S10ph and H3S28ph mitotic enrichment at promoters in specific epigenetic states, profile plots and corresponding enrichment heatmaps were generated for TSSs in each individual epigenetic state. Here, we focus on/highlight:

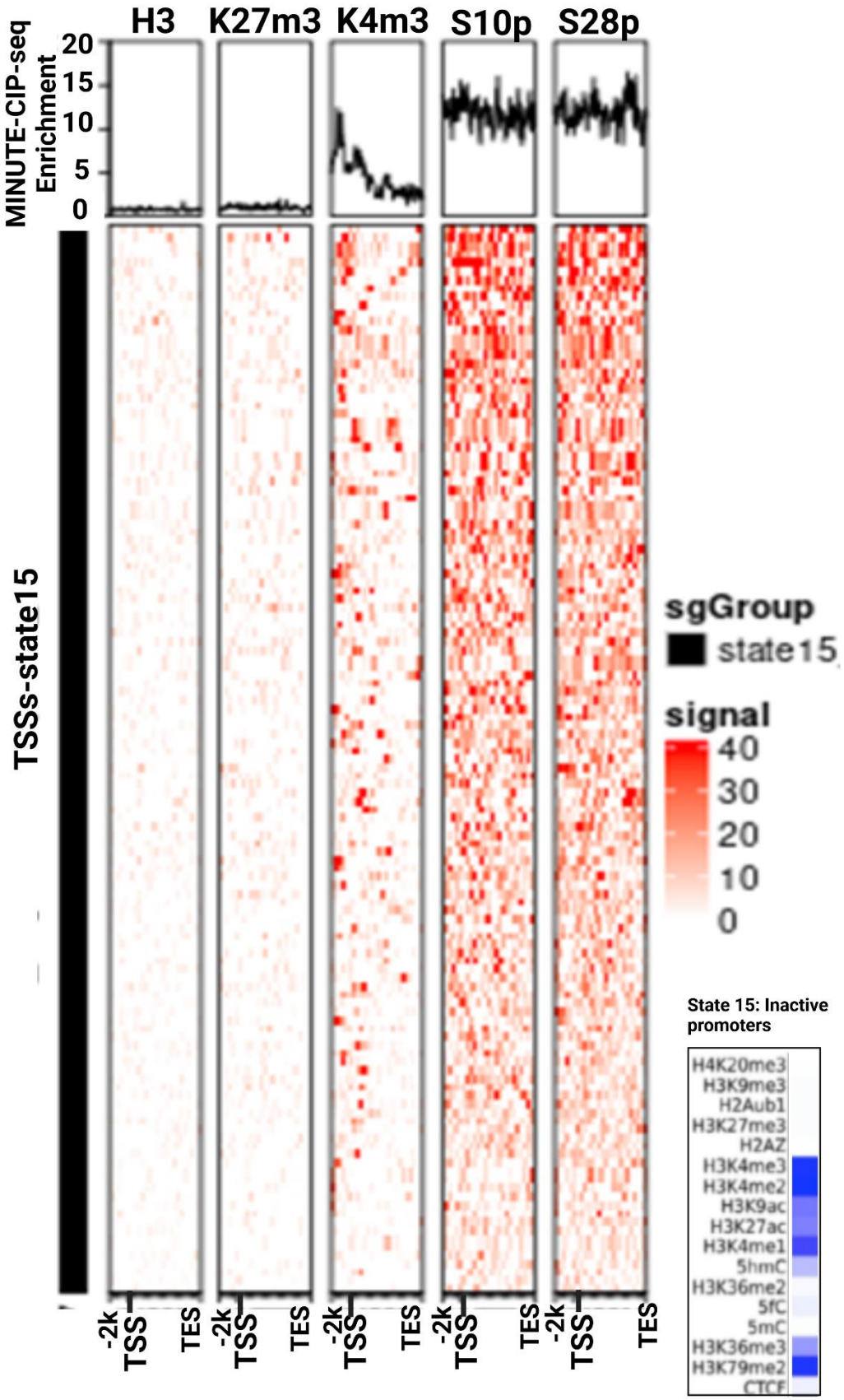
- states of most interest in the context of gene expression regulation
- states where high proportion of TSSs in that state were found to contain H3S10ph or H3S28ph called peaks

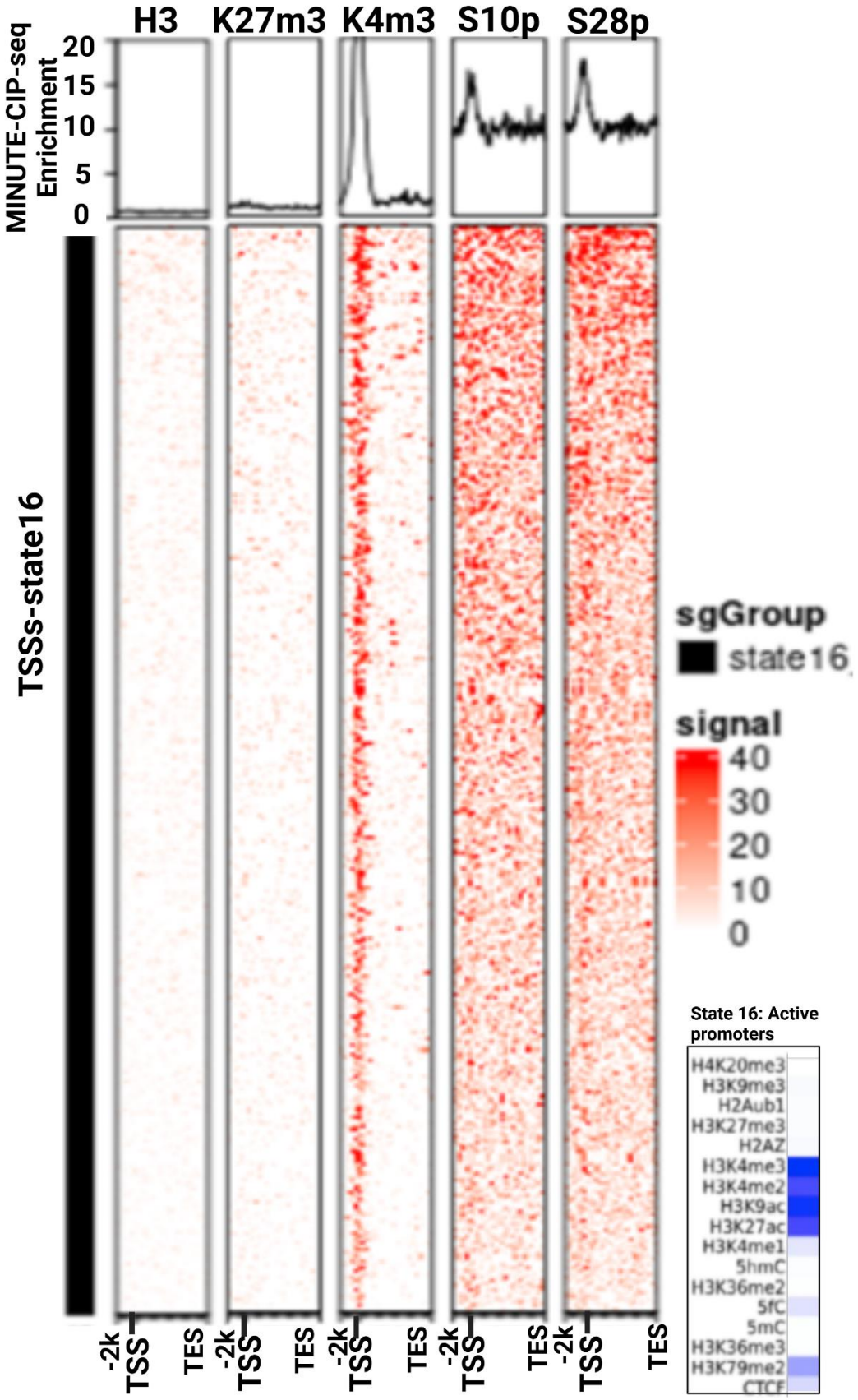
This allowed us to focus on the following chromatin states of interest:

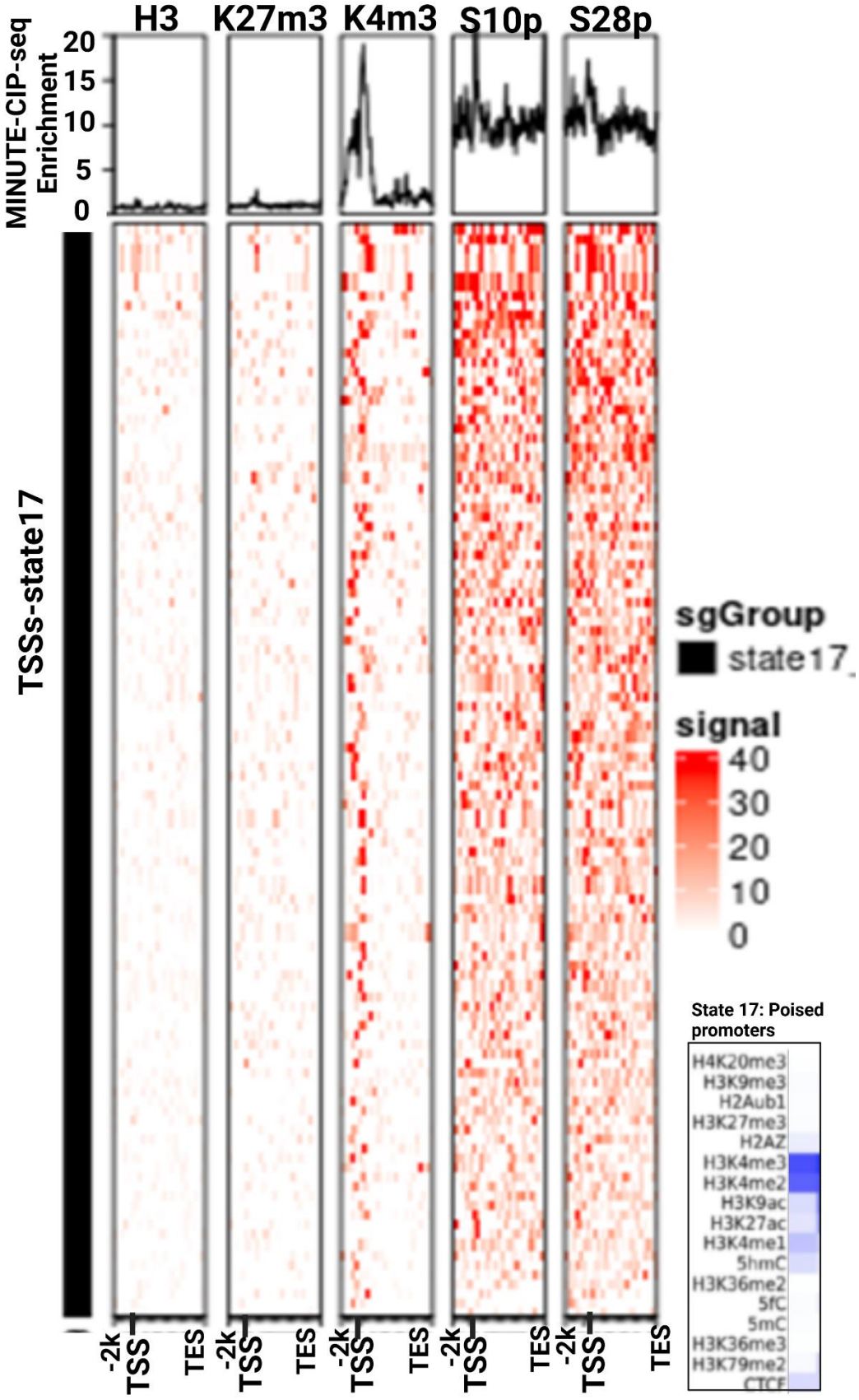
- “activation” chromatin states 15, 16 and 17
- “repression” chromatin states 18 and 19
- “elongation” state 2

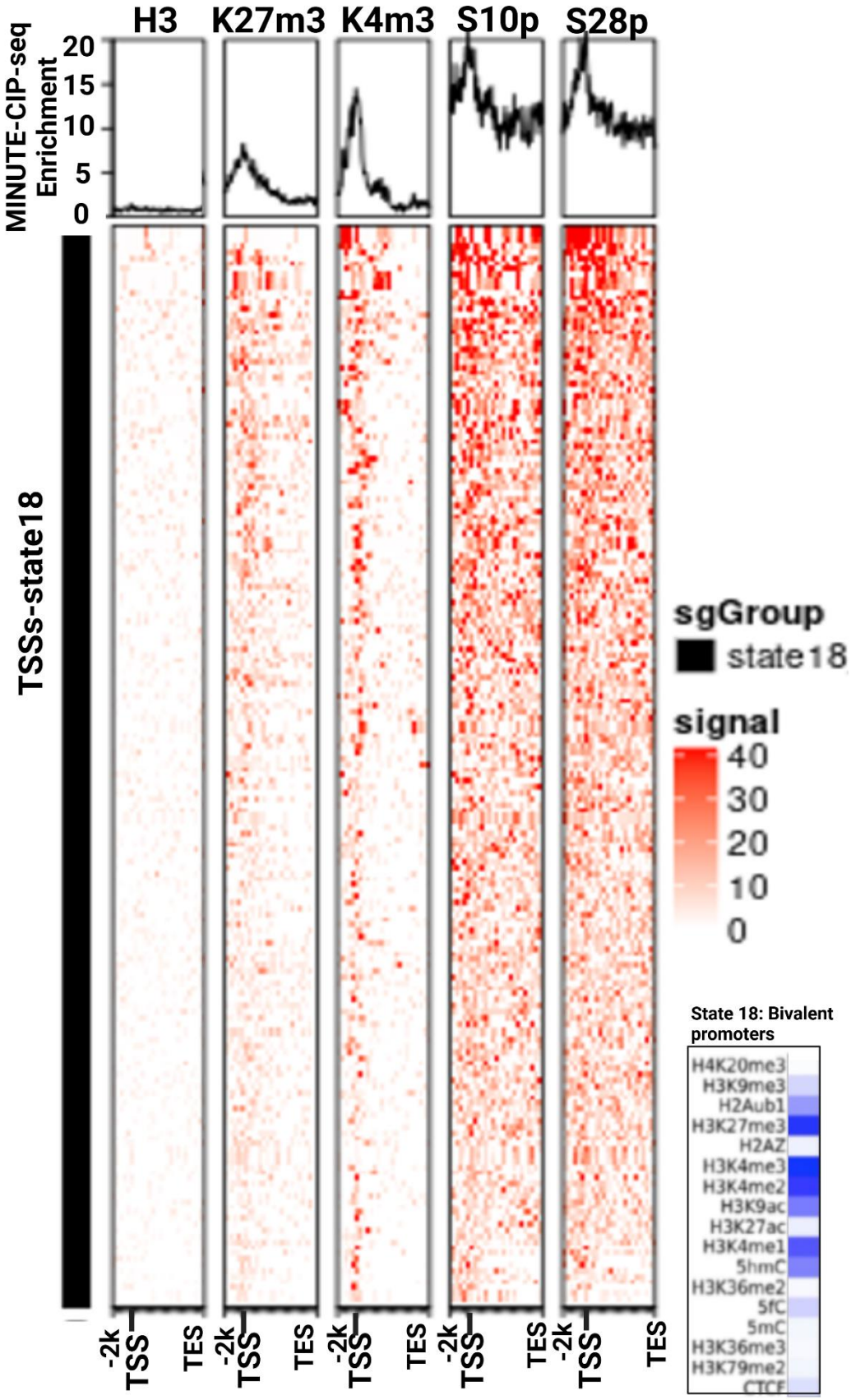
Profile plots and corresponding heatmaps allowed us to visualise promoter enrichment of H3S10ph and H3S28ph in specific epigenetic states. Plots were generated using profileplyr and are illustrated in **Figure 4.14**. Depicted are enrichment profiles and heatmaps of both H3S10ph and H3S28ph, in mitotic samples. Total histone H3, H3K4me3 and H3K27me3 are also displayed for comparison. The whole gene body is depicted to confirm whether H3S10ph or H3S28ph peaks can be seen at TSSs in certain epigenetic states.


H3
K27m3
K4m3
S10p
S28p
MINUTE-CIP-seq Enrichment
20
15
10
5
0
TSSs-state2
-2k
TSS
TES
sgGroup
state2_
signal
40
30
20
10
0
State 2: Active Elongation
H4K20me3
H3K9me3
H2Aub1
H3K27me3
H2AZ
H3K4me3
H3K4me2
H3K9ac
H3K27ac
H3K4me1
5hmC
H3K36me2
5fC
5mC
H3K36me3
H3K79me2
CTCF


H3
K27m3
K4m3
S10p
S28p
MINUTE-CIP-seq Enrichment
20
15
10
5
0
TSSs-state15
-2k
TSS
TES
sgGroup
state15
signal
40
30
20
10
0
State 15: Inactive promoters
H4K20me3
H3K9me3
H2Aub1
H3K27me3
H2AZ
H3K4me3
H3K4me2
H3K9ac
H3K27ac
H3K4me1
5hmC
H3K36me2
5fC
5mC
H3K36me3
H3K79me2
CTCF


H3
K27m3
K4m3
S10p
S28p
MINUTE-CIP-seq Enrichment
20
15
10
5
0
TSSs-state16
-2k
TSS
TES
sgGroup
state16
signal
40
30
20
10
0
State 16: Active promoters
H4K20me3
H3K9me3
H2Aub1
H3K27me3
H2AZ
H3K4me3
H3K4me2
H3K9ac
H3K27ac
H3K4me1
5hmC
H3K36me2
5fC
5mC
H3K36me3
H3K79me2
CTCF


H3
K27m3
K4m3
S10p
S28p
MINUTE-CIP-seq Enrichment
20
15
10
5
0
TSSs-state17
sgGroup
state17
signal
40
30
20
10
0
State 17: Poised promoters
H4K20me3
H3K9me3
H2Aub1
H3K27me3
H2AZ
H3K4me3
H3K4me2
H3K9ac
H3K27ac
H3K4me1
5hmC
H3K36me2
5fC
5mC
H3K36me3
H3K79me2
CTCF
-2k
TSS
TES


H3
K27m3
K4m3
S10p
S28p
MINUTE-CIP-seq Enrichment
20
15
10
5
0
TSSs-state18
-2k
TSS
TES
sgGroup
state18
signal
40
30
20
10
0
State 18: Bivalent promoters
H4K20me3
H3K9me3
H2Aub1
H3K27me3
H2AZ
H3K4me3
H3K4me2
H3K9ac
H3K27ac
H3K4me1
5hmC
H3K36me2
5fC
5mC
H3K36me3
H3K79me2
CTCF

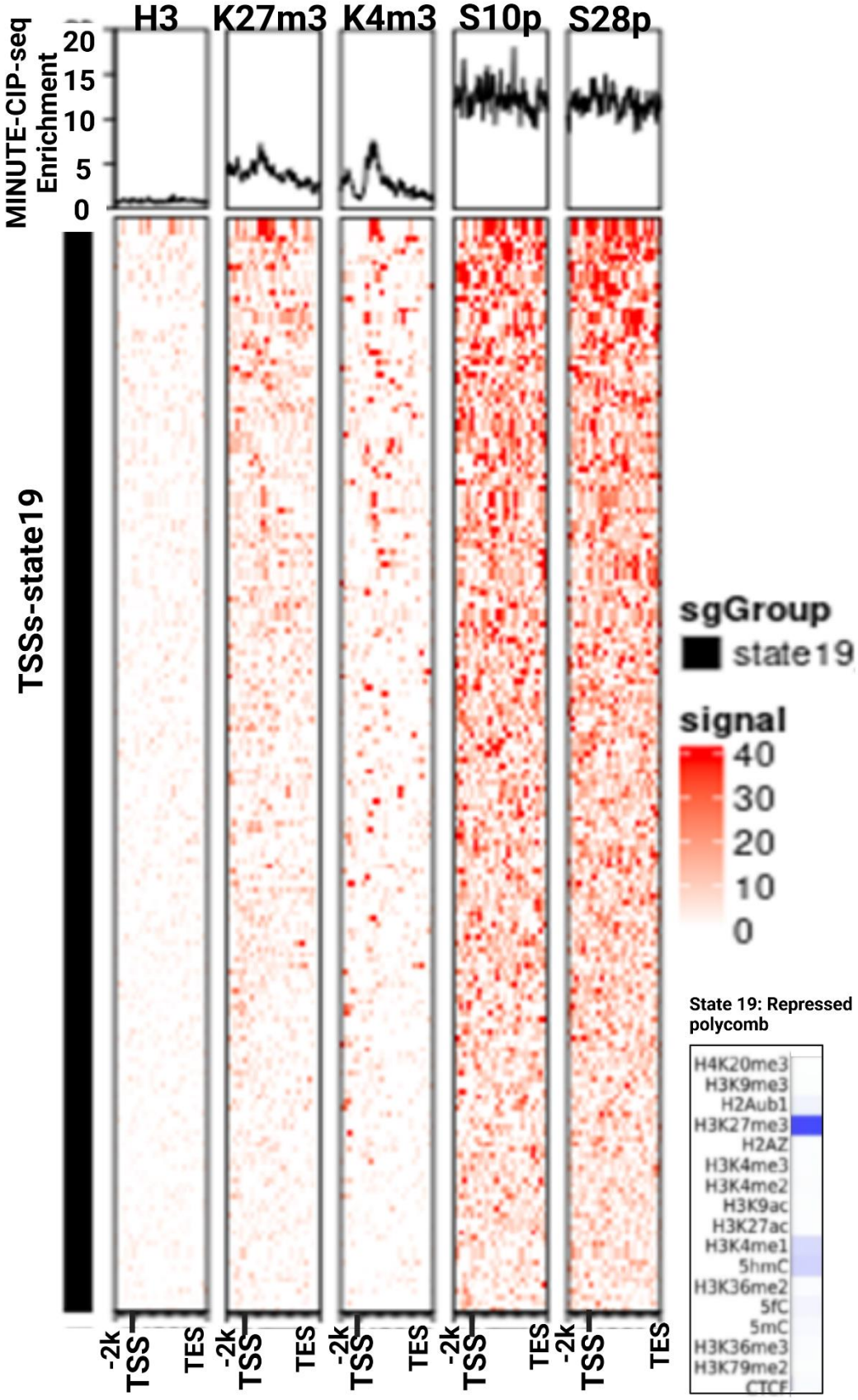

H3
K27m3
K4m3
S10p
S28p
MINUTE-CIP-seq Enrichment
20
15
10
5
0
TSSs-state19
-2k
TSS
TES
sgGroup
state19
signal
40
30
20
10
0
State 19: Repressed polycomb
H4K20me3
H3K9me3
H2Aub1
H3K27me3
H2AZ
H3K4me3
H3K4me2
H3K9ac
H3K27ac
H3K4me1
5hmC
H3K36me2
5fC
5mC
H3K36me3
H3K79me2
CTCF

Figure 4. 14: Enrichment profiles and heatmaps for MINUTE-ChIP-seq in mitotic samples.

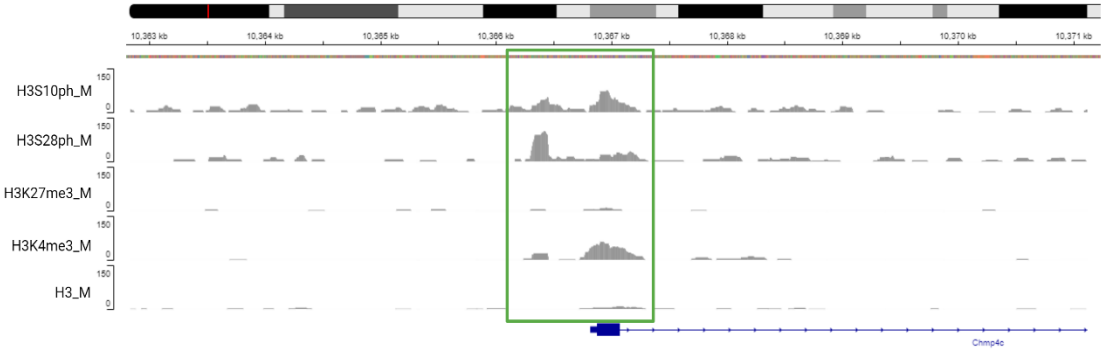
Figure shows enrichment of (left to right): total histone H3, H3K27me3 (K27m3), H3K4me3 (K4m3), H3S10ph (S10p) and H3S28ph (S28p). Enrichment is plotted to the gene body, with labelled TSS, -2kb upstream of TSS, and TES. Chromatin states of interest 2, 15, 16, 17, 18 and 19 are shown. For reference, a heatmap is provided (bottom right) for each chromatin state showing the relative enrichment of regulatory markers used to train the chromatin state model (Juan et al, 2016).

Both H3S10ph and H3S28ph enrichment peaks were clearly observed at TSSs in epigenetic states 16 and 18. Some enrichment peaks were seen for both H3S10ph and H3S28ph at TSSs in states 2 and 17. No enrichment peaks were seen for either histone phosphorylation at TSSs in states 15 or 19. We therefore suggest that H3S10ph and H3S28ph peak at promoters specifically in active promoter-associated state 16 and in bivalent promoter-associated state 18.

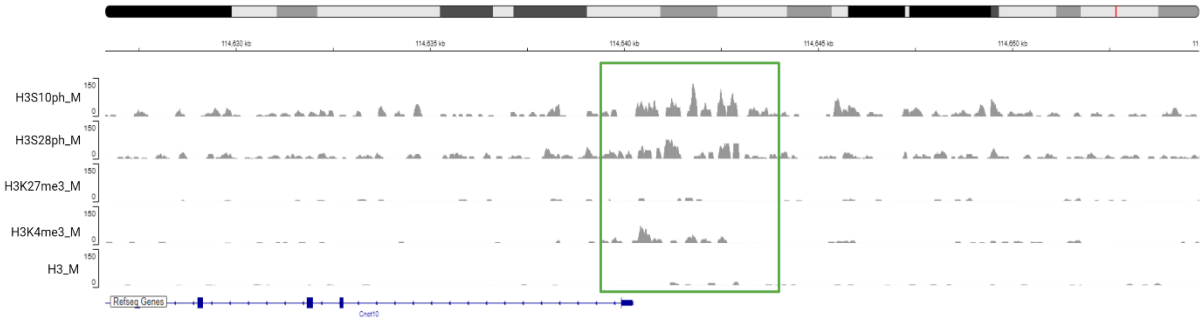
Interestingly, it must be noted that not all TSSs within a chromatin state can be said to have enrichment peaks. Heatmaps generated by Profileplyr, also displayed above (**Figure 4.14**), demonstrate that in state 16 and 18, where average profile plots show clear peaks, not all TSSs contain enrichment peaks. **Figure 4.13** also supports this, showing that only a proportion of TSSs contain MACS2-called significant peaks, including in states 16 and 18. In order to isolate TSSs within chromatin states 16 and 18 which did contain H3S10ph and H3S28ph peaks, we referred to our MACS2 binary classification of loci with statistically highest H3S10ph and H3S28ph. A list of TSSs was compiled containing only promoters in states 16 and 18 where MACS2 called peaks in H3S10ph or H3S28ph were present. Integrative Genomics Viewer (IGV) was used to confirm that these isolated TSSs corresponded to observable H3S10ph or H3S28ph peaks. **Figure 4.15** below shows a selection of example genes identified as having either H3S10ph or H3S28ph called peaks in states 16 and 18, alongside other histone modification tracks for comparison. It can be seen that state 16 TSSs show H3K4me3 enrichment, while state 18 TSSs show both H3K4me3 and H3K27me3 enrichment. As a negative control, a state 16 and state 18 gene examples Gtf2b and Zfp472 are shown where H3K4me3 enriches but H3S10ph and H3S28ph enrichment was **not** observed (i.e. no called peaks).

Chromatin state 16 (active promoter state)

E.g. 1: Chmp4c (charged multivesicular body protein 4C, associated with negative cell cycle regulation)



E.g. 2: Cnot10 (CCR4-NOT transcription complex, subunit 10, associated with gene expression)



E.g. 3: Zfp472 (Zinc finger protein 472)

Chromatin state 18 (bivalent promoter state)

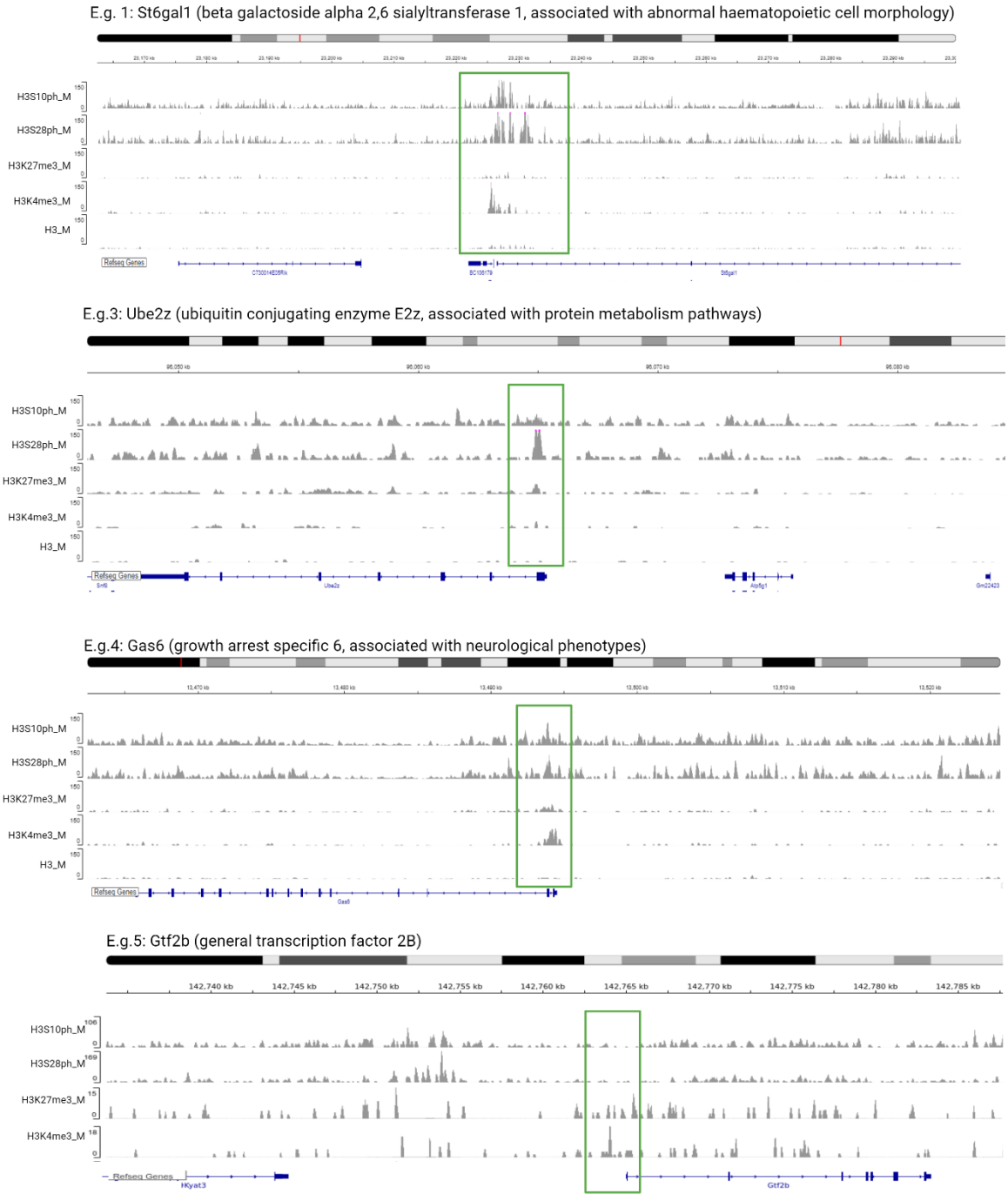


Figure 4. 15: IGV coverage tracks of H3S10ph, H3S28ph, H3K27me3, H3K4me3 and H3, at example gene coordinates.

Input-normalised, read count scaled MINUTE-ChIP-seq scores (y axis) of mitotic mESC samples are shown as coverage tracks, top to bottom: H3S10ph, H3S28ph, H3K27me3, H3K4me3 and total H3. Displayed are example regions containing TSS enrichment peaks. TSSs in chromatin state 16, representing active promoters, and chromatin state 18, representing bivalent promoters, are shown.

This list of genes in epigenetic state 16 and 18, containing mitotic H3S10ph or H3S28ph called peaks at TSSs, was carried forward into subsequent gene functional enrichment analyses, detailed in Section 4.2.8.

4.2.8 H3S10ph and H3S28ph peaks enrich at promoters in different chromatin regulatory states depending on cell cycle phase

The above analyses focussed on mitotic samples, where histone phosphorylations H3S10ph and H3S28ph are seen to strongly enrich. The MINUTE-ChIP-seq method allowed us to quantitatively compare phosphorylation enrichment between different cell cycle phases which were also sampled by Elsasser et al. Therefore, importantly, we also assessed histone phosphorylation enrichment peaks in G1 and G2 samples, in comparison to our mitotic observations. We assessed whether the H3S10ph and H3S28ph peaks identified in mitotic samples were also observed in G1 or G2. The plots below are presented for comparison: (**Figure 4.16, 4.17**).

Firstly, the proportion of TSSs in each chromatin state containing MACS2 called H3S10ph or H3S28ph peaks were compared across G1, G2 and M samples, shown below (**Figure 4.16**). It is immediately clear that the chromatin states in which the most H3S10ph and H3S28ph promoter peaks occur vary strongly depending on the cell cycle phase. The high number of H3S10ph and H3S28ph peaks seen at state 18 promoters in mitosis is not seen in G2 and is greatly reduced in G1. The H3S10ph peaks seen in mitosis in poised active promoters in state 17 is also high in G2, but much lower in G1; similarly H3S28ph peaks at state 17 promoters are only seen in mitosis; this could suggest that H3S10ph deposition at these poised promoters begins in G2 phase and continues into mitosis, while H3S28ph deposition at state 17 promoters begins in mitosis. Contrastingly, promoters in some chromatin states show high levels of H3S10ph and H3S28ph peaks in interphase that are not seen in mitotic samples; for example, H3S28ph peaks are high in

state 11 promoters in G1, and state 10 and 12 promoters in G2, none of which are seen in mitosis. H3S10ph peak are high in state 8 promoters in G2, and state 10 and 11 promoters in G1, and again this is lost in mitosis. This analysis presents numerous interesting implications as to possible deposition timings of H3S10 and H3S28 phosphorylation through the cell cycle at promoters in different regulatory regions. These interphase analyses could be expanded in future work and would likely provide valuable evidence as to potential roles of H3S10ph and H3S28ph in interphase chromatin regulation.

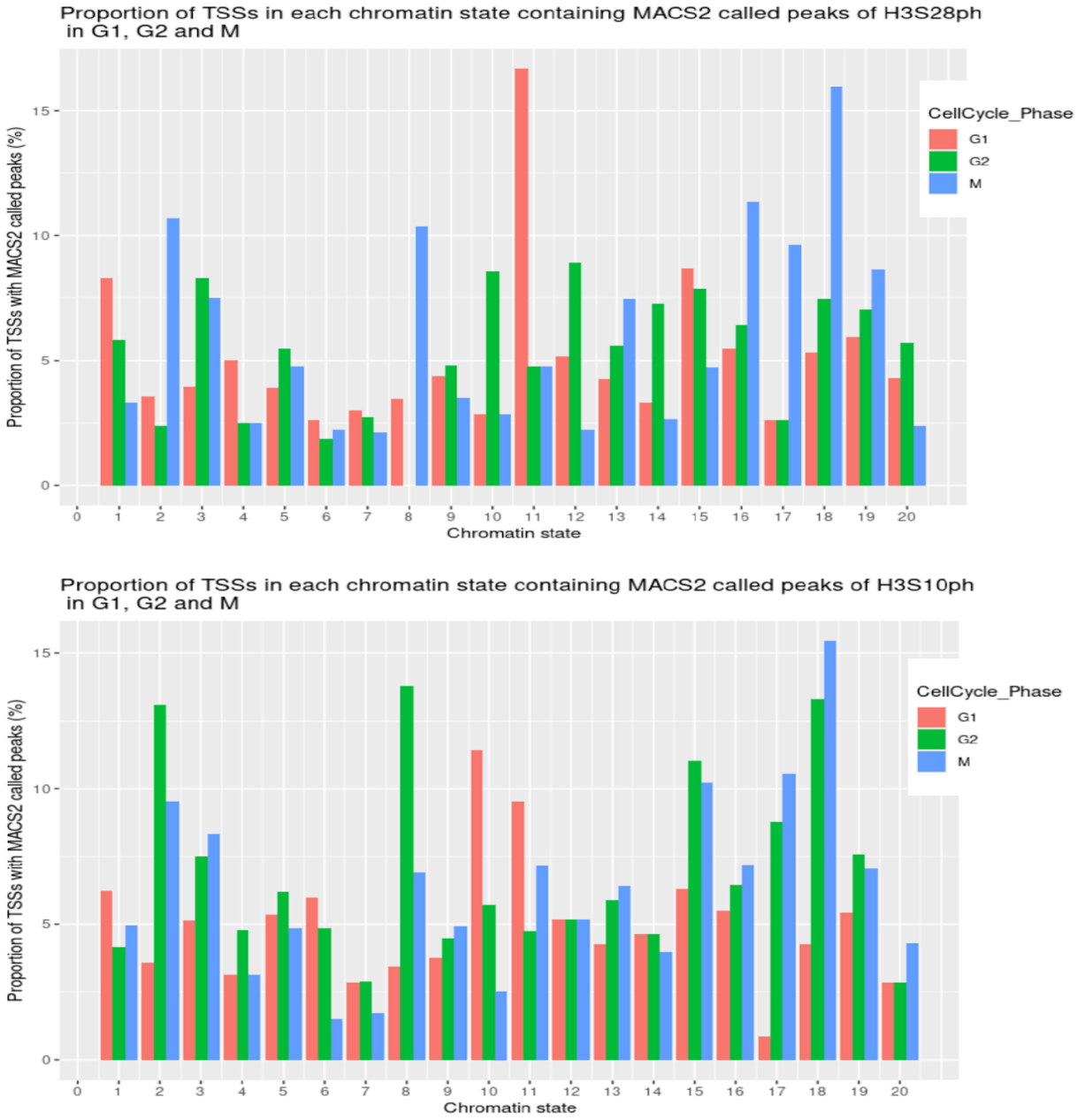


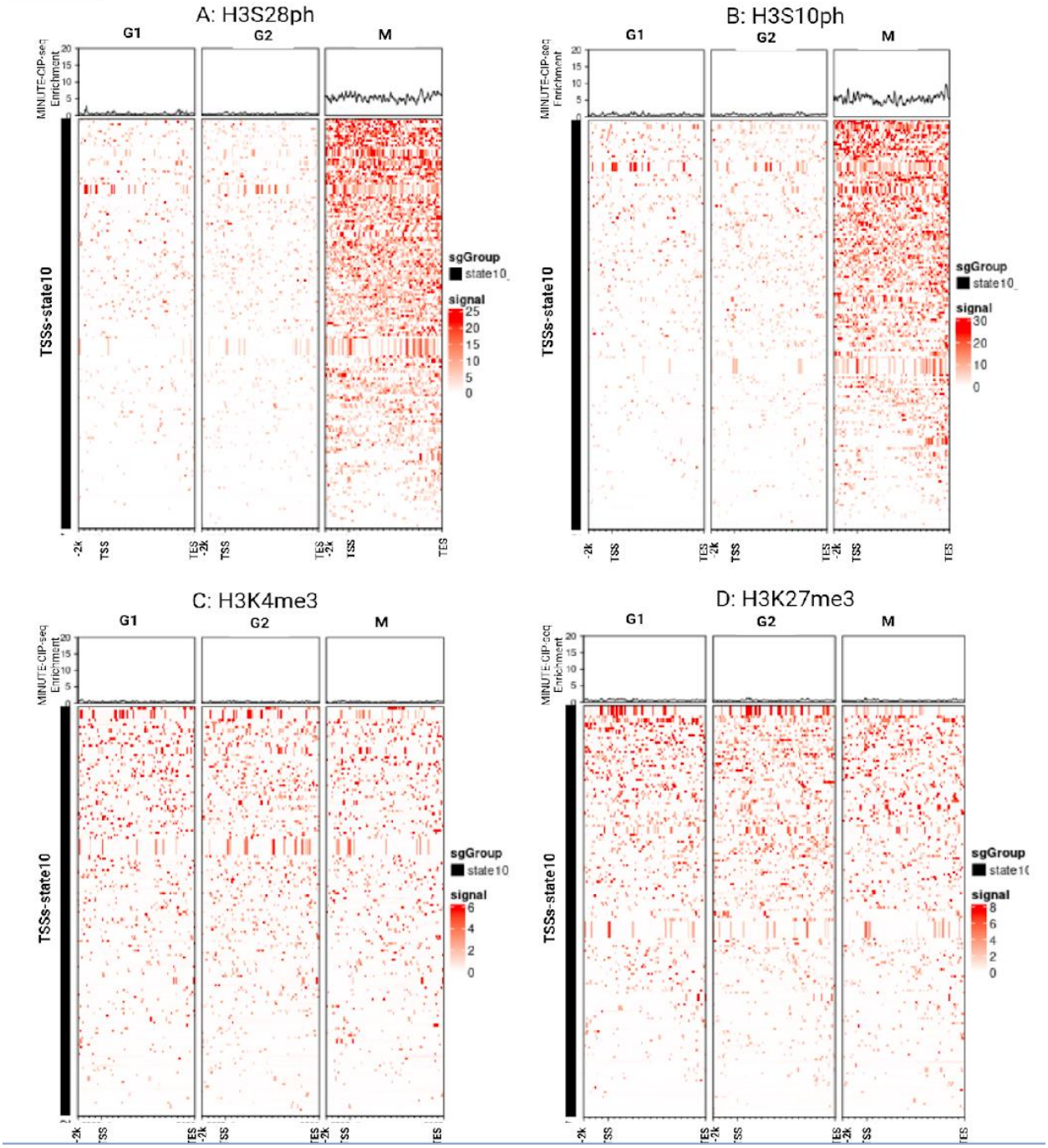
Figure 4. 16: The proportion (%) of TSSs in each defined chromatin epigenetic state 1-20 which contain MACS2-called mitotic H3S28ph(left) and H3S10ph(right) peaks, comparing G1, G2 and Mitotic (M) samples.

TSSs (+/- 1 kb) were labelled as chromatin epigenetic states 1-20 based on the trained ChromHMM 20 state model generated in Juan et al., (2016). MACS2 was used to call significant peaks for mitotic H3S10ph and H3S28ph, shown in separate plots. The proportion of TSSs in a

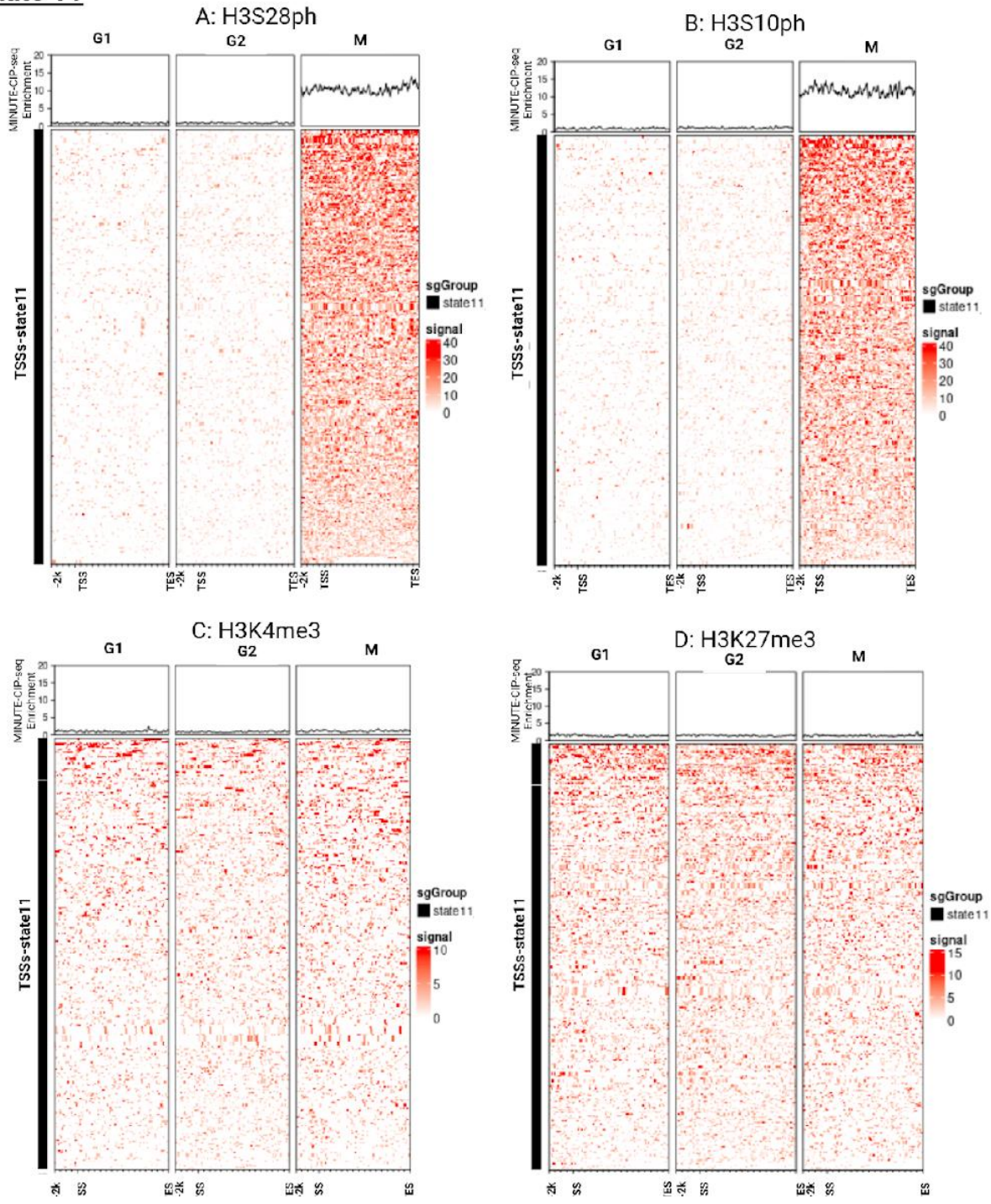
given state that overlap with at least one H3S10ph or H3S28ph called peak were plotted. The plot shows G1 (red), G2 (green) and M (blue) values clustered for comparison in each chromatin state.

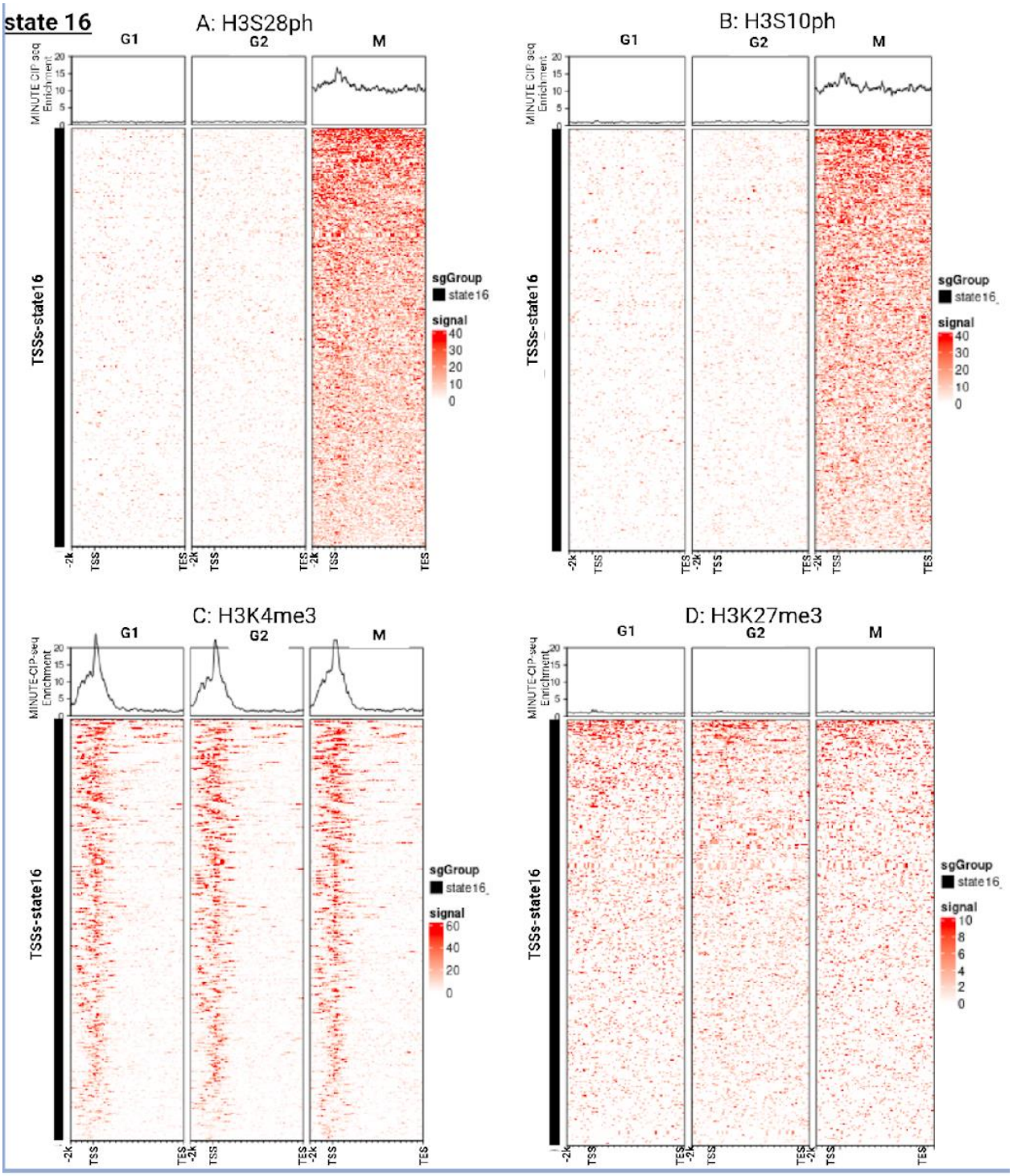
We then assessed whether the H3S10ph and H3S28ph enrichment peaks we observed in mitosis at TSSs in chromatin states 16 and 18 were present in G1 or G2 samples, and whether enrichment peaks could be observed in G1 or G2 samples that weren't present in Mitosis in states 10 and 11. As can be seen in **Figure 4.17**, the promoter H3S10ph and H3S28ph enrichment seen in mitosis in state 16 and state 18 is not seen in G1 or G2. Enrichment peaks are not observed in interphase or mitosis at promoters in state 10 or 11, despite high proportion of TSSs containing MACS2 called peaks in these states. It is possible that the overall lower enrichment signal seen in G1 and G2 samples restricts ability to identify peaks. It is noted that while phosphorylation enrichments may be seen in other chromatin regulatory state regions; further analysis could be done in future to investigate interphase enrichment distributions in greater depth.

state 10


A: H3S28ph
B: H3S10ph
C: H3K4me3
D: H3K27me3
G1
G2
M
MINUTE-ChIP-seq Enrichment
TSSs-state10
sgGroup
state10
signal
-2k
TSS
TES

state 11


A: H3S28ph
B: H3S10ph
C: H3K4me3
D: H3K27me3
G1
G2
M
MINUTE-ChIP-seq Enrichment
TSSs-state11
sgGroup
state11
signal
-2k
TSS
TES


state 16
A: H3S28ph
B: H3S10ph
C: H3K4me3
D: H3K27me3
G1
G2
M
MINUTE-ChIP-seq Enrichment
TSSs-state16
-2k
TSS
TES
sgGroup
state16
signal

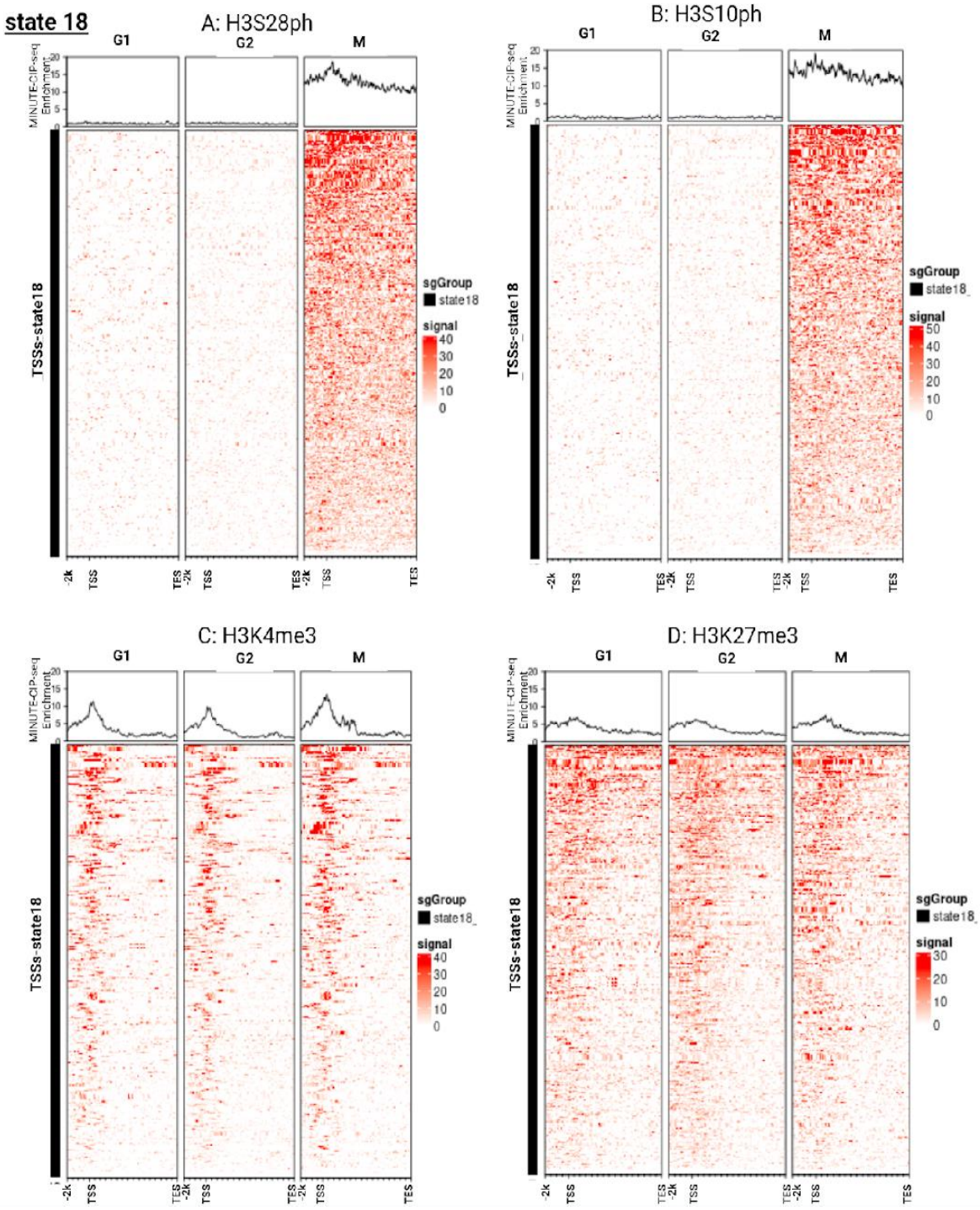


Figure 4. 17: Enrichment profiles and heatmaps for MINUTE-ChIP-seq comparing G1, G2 and Mitosis (M).

Figure shows enrichment of H3S28ph (A), H3S10ph (B), H3K4me3(C) and H3K27me3 (D) in chromatin states 10 (top left box), 11 (top right), 16 (bottom left) and 18 (bottom right). Each cluster of plots compares cell cycle phases G1 (left), G2 (centre) and M (right). Enrichment is plotted to the gene body, with labelled TSS, -2 kb upstream of TSS, and TES.

4.2.9 Gene Set Enrichment Analysis reveals promoters with H3S10ph and H3S28ph peaks significantly enrich for multiple functional gene sets

Our analysis of mitotic H3S10ph and H3S28ph in the context of chromatin regulatory states allowed us to isolate for the first time mitotic promoter peaks in actively transcribed promoter regions (state 16) and in bivalent repressed promoter regions (state 18). Our next aim was to identify possible functional roles of these H3S10ph or H3S28ph promoter peaks.

Chromatin epigenetic states 16 and 18, as defined by Juan et al 2016, where H3S10ph and H3S28ph enrichment peaks were seen at TSSs, were selected as the focus of our Gene Ontology analysis. TSSs were selected that contained MACS2-called peaks in H3S10ph or H3S28ph, in order to filter to only loci classified as "high" H3S10ph or H3S28ph (see Section 4.2.7). This provided a list of genes for TSSs in state 16 and 18, with likely H3S10ph or H3S28ph peaks. These genes were then analysed against publicly available functional databases.

We utilised the online platform Webgestalt, a commonly-used software tool for mining functional sets of genes (Zhang *et al.*, 2005). Webgestalt interrogates 15 databases in 7 classes, illustrated in **Figure 4.18.** Gene sets derived from these central and public databases are then subjected to Boolean statistical operations to generate unions, intersections and/or differences between different gene sets (Zhang *et al.*, 2005). Webgestalt then organises gene sets for visualisation in various biological context, presenting biological areas including Gene Ontology, chromosome distribution, metabolic pathways and signalling pathways that are important for that gene set and warrant further investigation. Users can input the ID of their genes of interest, and Webgestalt can identify which of its functional gene sets are over-represented in those input genes of interest. This method is called **Gene Set Enrichment Analysis (GSEA).**

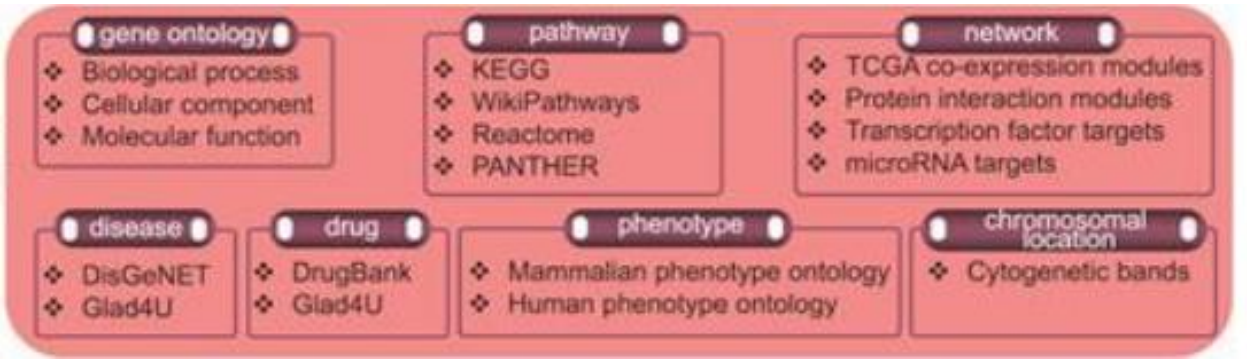


Figure 4. 18: Webgestalt used 15 databases in 7 classes to generate gene sets across 150,937 functional categories.

Figure from Wang et al., 2017

For our analysis, the genes isolated containing H3S10ph or H3S28ph MACS2-called peaks at promoters were isolated, and our gene list was ranked based on H3S10ph or H3S28ph MINUTE-ChIP-seq normalised enrichment values. GSEA was then performed through Webgestalt, using recommended statistical parameters, comparing these phosphorylation enrichment-ranked gene lists against Webgestalt functional gene sets. The aim was to identify any functional enrichments of gene families, biological pathways, cell types or disease phenotypes that are over-represented/enriched in our genes with histone phosphorylation peaks. Gene sets >= 10 were considered, and False Discovery Rate <0.25 was considered as potentially significant (provided $p<0.05$), as per Webgestalt recommendations. Numerous functional gene sets were found to significantly enrich in our high-H3S10ph and high-H3S28ph genes, described below:

H3S10ph

Genes in active promoter chromatin state 16 that contained H3S10ph MACS2-called peaks showed significant functional enrichment for protein metabolism pathways ($p = 0.037$, FDR = 0.212); this enriched gene set is listed below, **Table 4.3**.

Table 4. 3: GSEA found H3S10ph peaks at state 16 active promoters show functional enrichment for the below genes sets Webgestalt reports are associated with protein metabolism.

H3S10ph MINUTE-ChIP normalised Enrichment Signal	**Gene Symbol**	**Gene Name**	**Gene Known Function(s)**
65	*St6gal1*	Beta galactoside alpha 2;6 sialyltransferase 1	Generation of cell-surface carbohydrate determinants and differentiation antigens
39	*Dcun1d2*	DCN1, defective in culin neddylation 1, somain containing 2	(predicted) enables ubiquitin conjugating enzyme activity, (predicted) involved in protein neddylation, (predicted) part of ubiquitin ligase complex
37	*Mrpl40*	Mitochondrial ribosomal protein L40	(predicted) involved in mitochondrial translation
34	*Commd2*	COMM domain containing 2	(predicted) modulate cullin-RING E3 ubiquitin ligase activity, (predicted) down-regulate NF-kappa-B activation
33	*Mrpl28*	Mitochondrial ribosomal protein L28	Involved in mitochondrial protein synthesis
31	*Pex2*	Peroxisomal biogenesis factor 2	Involved in peroxisome organisation. (Predicted) enables ubiquitin ligase activity
21	*Mepl43*	Mitochondrial ribosomal protein L43	Involved in mitochondrial protein synthesis

20	*Rce1*	Ras converting CAAX endopeptidase 1	Enables endopeptidase activity
20	*Cul4a*	Cullin 4A	Part of Cul4a-RING E3 ubiquitin ligase complex. Involved in several processes, including cellular response to UV; in utero embryonic development; and ribosome biogenesis. Acts upstream of or within several processes, including negative regulation of granulocyte differentiation; regulation of cell cycle phase transition; and regulation of nucleotide-excision repair
19	*Marchf6*	Membrane-associated ring finger (C3HC4) 6	(Predicted) involved in ubiquitin protein- ligase, -protease and -conjugating enzyme activity.

Interestingly, H3S10ph MACS2-called peaks in chromatin state 18, representing bivalent promoters, showed many significant enrichments for genes associated with abnormal haematopoietic cell morphologies and abnormal immune cell morphologies ($p<0.05$, FDR<0.25), which are presented in **Table 4.4**. H3S10ph at state 18 promoters also showed significant enrichment for protein metabolism genes; namely *Gas6* (growth arrest specific 6), *Klhl22* (kelch-like 22), *Ube2z* (ubiquitin conjugating enzyme E2Z) and *Gcg* (Glucagon).

Table 4. 4: GSEA found H3S10ph peaks at state 18 bivalent promoters show significant enrichment for gene sets Webgestalt reports are associated with haematopoietic and immune cell morphologies.

Gene Set Description	Gene Set Size	GSEA Enrichment Score	P value	FDR
abnormal hematopoietic system morphology/development	15	0.59799	0.016860	0.089038
abnormal immune system cell morphology	11	0.65282	0.016234	0.097748
abnormal leukocyte morphology	11	0.65282	0.016234	0.097748
immune system phenotype	19	0.54376	0.030114	0.10130
abnormal hematopoietic cell morphology	12	0.63192	0.014100	0.10220
hematopoietic system phenotype	19	0.54577	0.018634	0.10960
abnormal mononuclear cell morphology	10	0.63037	0.035398	0.12056
abnormal immune system morphology	14	0.63444	0.010638	0.12969
nervous system phenotype	14	-0.38298	0.010870	0.16731
abnormal myeloid cell morphology	10	0.69411	0.0032931	0.18485
abnormal nervous system morphology	11	-0.37113	0.12500	0.21680

H3S28ph

H3S28ph was also analysed for functional gene enrichments. For genes in active promoter state 16, H3S28ph MACS2 peaks showed significant enrichment for negative regulation of the cell cycle; the genes found to be enriched are shown below, **Table 4.5**. In bivalent promoter state 18 genes, H3S28ph MACS2 peaks show significant enrichment for the same genes associated with protein metabolism as were enriched for H3S10ph state 18 peaks.

Table 4. 5: GSEA for S28ph peaks at state 16 active promoters show functional enrichment for the below genes associated with negative cell cycle regulation.

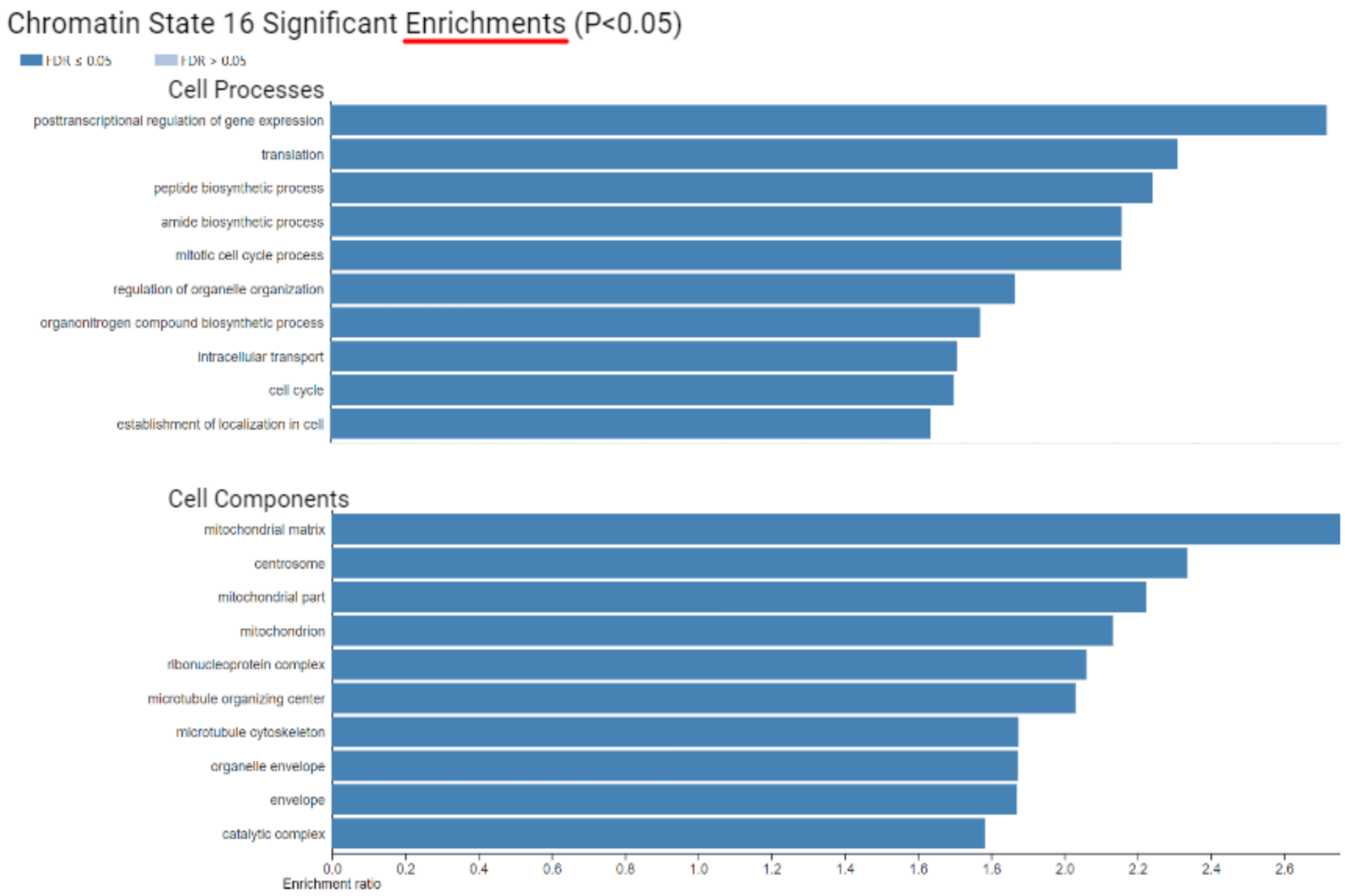
H3S28ph MINUTE-ChIP normalised Enrichment Signal	**Gene symbol**	**Gene name**	**Gene Known Function(s)**
40.2302	*Cul4a*	Cullin 4A	Part of Cul4a-RING E3 ubiquitin ligase complex. Involved in several processes, including cellular response to UV; in utero embryonic development; and ribosome biogenesis. Acts upstream of or within several processes, including negative regulation of granulocyte differentiation; regulation of cell cycle phase transition; and regulation of nucleotide-excision repair
39.4559	*Pcid2*	PCI domain containing 2	Involved in several processes, including positive regulation of mitotic cell cycle spindle assembly checkpoint; regulation of DNA-templated transcription; and spleen development.
31.6955	*Mcph1*	Microcephaly, primary autosomal recessive 1	Acts upstream of or within several processes, including

			establishment of mitotic spindle orientation; neuronal stem cell population maintenance; and protein localization to centrosome. (Predicted) located in centrosome and cytoplasm. Is expressed in brain. Used to study microcephaly.
20.2727	*Btg3*	B cell translocation gene 3	Thought to have anti-proliferative properties, and may be involved in regulating the G1-S transition to suppress cell cycle progression.
17.4134	*Cep192*	Centrosomal protein 192	Predicted to enable phosphatase binding activity. Acts upstream of or within response to bacterium.
17.346	*Chmp4c*	Charged multivesicular body protein 4C	(Predicted) enables protein homodimerization activity. (Predicted) involved in several processes, including proteolysis involved in protein catabolic process; regulation of cell cycle process; and vacuolar transport

We noted that these functional enrichments could be biassed by any functional enrichments that all genes in a given chromatin state have regardless of phosphorylation signal. To address this, we listed all protein-coding genes with promoters classified as chromatin state 16 or 18. We then performed an **Over-Representation Analysis (ORA).** Unlike GSEA, ORA does not use ChIP signal ranking. Instead, ORA compares the gene list in question against the full reference genome, and identifies any functional pathways which the gene list is significantly enriched for compared to the reference genome. Therefore, ORA allowed us to show what gene sets are functionally enriched in chromatin state 16 and 18 compared to the mouse reference (mm9) genome, **regardless of H3S10ph or H3S28ph levels**.

We found that chromatin state 16 genes significantly enriched for cell processes and components such as cell cycle regulation, peptide biosynthesis, organelle organisation, centrosomes and microtubule organisation, shown in **Figure 4.19A**; these are arguably important "housekeeping" gene functions that would be expected to be in an active regulatory state such as state 16. Chromatin state 18, thought to represent bivalent promoters, showed different ontology enrichments, including DNA binding functions and multiple neural cell components (see **Figure 4.19B**). By performing ORA in the context of specific epigenetic states, we were able to interpret findings as to histone phosphorylation functional enrichments while accounting for variation between different states across the genome.findings as to histone phosphorylation functional enrichments while accounting for variation between different states across the genome.

A


Chromatin State 16 Significant Enrichments (P<0.05)
FDR ≤ 0.05
FDR > 0.05
Cell Processes
posttranscriptional regulation of gene expression
translation
peptide biosynthetic process
amide biosynthetic process
mitotic cell cycle process
regulation of organelle organization
organonitrogen compound biosynthetic process
intracellular transport
cell cycle
establishment of localization in cell
Cell Components
mitochondrial matrix
centrosome
mitochondrial part
mitochondrion
ribonucleoprotein complex
microtubule organizing center
microtubule cytoskeleton
organelle envelope
envelope
catalytic complex
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
Enrichment ratio

B

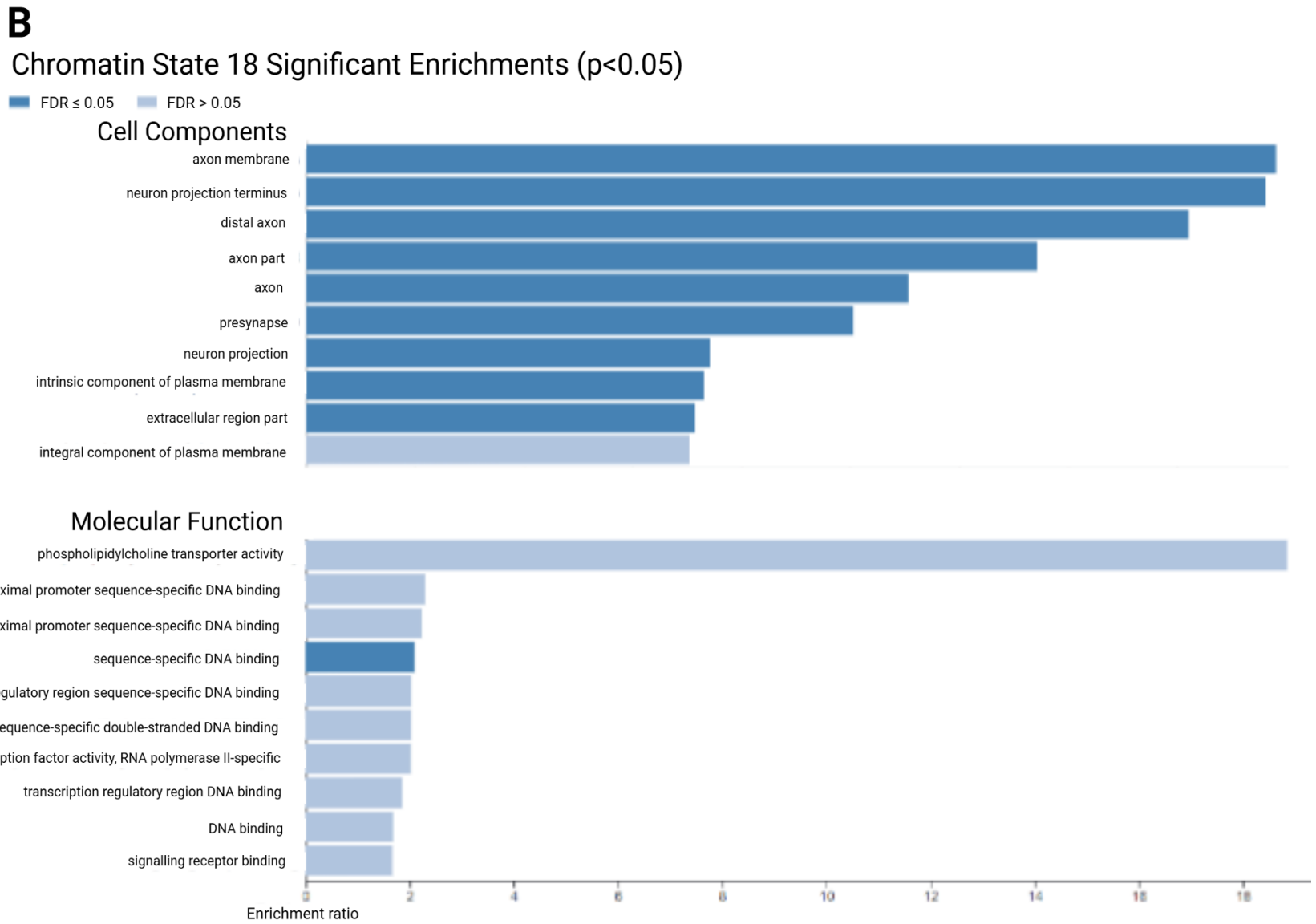

Chromatin State 18 Significant Enrichments (p<0.05)
FDR ≤ 0.05
FDR > 0.05
Cell Components
axon membrane
neuron projection terminus
distal axon
axon part
axon
presynapse
neuron projection
intrinsic component of plasma membrane
extracellular region part
integral component of plasma membrane
Molecular Function
phospholipidylcholine transporter activity
proximal promoter sequence-specific DNA binding
proximal promoter sequence-specific DNA binding
sequence-specific DNA binding
cis regulatory region sequence-specific DNA binding
sequence-specific double-stranded DNA binding
transcription factor activity, RNA polymerase II-specific
transcription regulatory region DNA binding
DNA binding
signalling receptor binding
0
2
4
6
8
10
12
14
16
18
Enrichment ratio

Figure 4. 19: Over-representation analysis of genes in chromatin state 16 and 18, compared to reference genome.

Webgestalt ORA analysis was performed on genes whose TSSs were categorised as chromatin state 16 (A) or chromatin state 18 (B), using the mm9 reference genome as reference. Webgestalt showed significant enrichment ($p<0.05$, $FDR<0.05$) in genes in chromatin state 16 for the biological processes and cell components shown below.

Importantly, the gene functions which enriched at promoters with high H3S10ph and H3S28ph, identified through GSEA, were not functions seen to enrich across all state 16 or state 18 genes.

Overall, we can conclude that genes containing H3S10ph and H3S28ph MACS2-called peaks at promoters are significantly enriched in functional pathways including protein metabolism, negative cell cycle regulation, and haematopoietic and immune cell morphology. However, the mechanisms by which these histone phosphorylations may contribute to these functional pathways remain unclear, and require further validation studies.

We emphasise that these new insights into H3S10ph and H3S28ph promoter enrichment were only identified by improving the "resolution" of our H3S10ph and H3S28ph distribution maps; isolating specific chromatin regulatory regions allowed us to uncover subtle but significant H3S10ph and H3S28ph enrichment peaks which were previously undiscovered amidst genome-wide high H3S10ph and H3S28ph sequencing signal levels.

4.3 Discussion

4.3.1 The importance of thorough antibody binding characterisation

We performed rigorous characterisation of antibodies reported to target our histone phosphorylations of interest H3S10ph and H3S28ph. By thoroughly validating a range of anti-H3S10ph and H3S28ph antibodies we hoped to improve accuracy of interpretation of MINUTE-ChIP-seq results. Using peptide ELISAs we tested the ability of antibodies to recognise and bind H3S10ph or H3S28ph, with and without adjacent additional histone modifications, allowing us to assess cross-reactivity as well as any blocking adjacent modifications. Our results clearly highlight the importance of rigorous antibody validation prior to chromatin immunoprecipitation studies. Antibodies RR002, 6G3, reported by their respective manufacturers to bind H3S10ph, showed poor binding of H3S10ph in the presence of many of the common neighbouring histone modifications, including H3K9me3, H3K9ac, H3T11ph and H3K14ac. RR002 in particular showed poor binding to H3S10ph peptide. Antibodies 3H10 and WR1+2 showed good recognition of H3S10ph; however, both showed binding inhibition when adjacent phosphorylation H3T11ph was present. It is crucial that these characteristics are accounted for when interpreting immunoprecipitation sequencing data; for example, any genomic depletions of H3S10ph sequencing could in fact be due to inhibition by H3T11ph, rather than a true biologically relevant H3S10ph depletion.

Regarding H3S28ph antibodies, HTA28 showed strong and specific binding to all peptides containing S28ph, with minimal inhibition by adjacent modifications. Contrastingly, antibody 5D10 was unable to bind S28ph when adjacent methylations or acetylations were present, such as K27me3 and K27ac. We chose not to perform our subsequent analyses on 5D10 sequencing, as these blocking/inhibitory modifications make interpreting signal much more difficult. However, we do note that 5D10 could serve as a useful antibody for future experiments, for example to specifically target and identify regions where S28ph is present alone without adjacent modifications.

We also note that further characterisation could be performed, for example using whole nucleosome ELISAs. It cannot be confidently known whether *in vitro* biochemical assays can accurately predict the performance of antibodies in MINUTE-ChIP-seq experiments. Indeed,

Henikoff *et al* and many standardised CUT+RUN protocols (e.g. Epicypher) stress that ELISA validation does not guarantee suitability of an antibody for immunoprecipitation *in vivo*. Whole nucleosome ELISAs could be used to improve this, as they would arguably better represent the true biological behaviour of nucleosomes *in vivo* than shortened peptides.

4.3.2 Quantitative MINUTE-ChIP-seq analysis reveals H3S10ph and H3S28ph promoter enrichments that are specific to mitosis

The initial aim of this Chapter was to employ a quantitative method, MINUTE-ChIP-seq, to improve our understanding of the exact enrichment distribution of two highly abundant mitotic histone phosphorylations: H3S10ph and H3S28ph. Our collaborators Kumar and Elsasser (2019) performed extensive validation experiments using H3K27me3 as an example histone modification to assess the quantitative ability of MINUTE-ChIP-seq. By generating cell populations with known H3K27me3 levels from 0% to 100%, they were able to show that their MINUTE-ChIP-seq normalisation method, used for this chapter's mESC dataset, produces sequencing read count scores that are highly accurate to the true biological enrichment levels of the targeted histone modification.

Elsasser and colleagues isolated cell populations in cell cycle phases G1, G2, S phase and Mitosis (M) using flow cytometry. As detailed in Chapter 2, fluorescent cell-cycle probes were used to generate these populations, with mitotic cells isolated using MPM-2 predominantly in prometaphase, likely with some prophase and metaphase cells. Importantly, this dataset presented a unique opportunity to quantitatively compare histone phosphorylation signal between interphase (G1 or G2) and Mitosis - this would not be possible in the more qualitative traditional ChIP-seq, and is pivotal in this project in allowing us to identify regulatory enrichments in histone H3S10ph and H3S28ph that are specifically seen in mitosis.

In this Chapter, through MACS2 peak calling and metagene analysis of MINUTE-ChIP-seq data, we demonstrate that statistically significant enrichment peaks in H3S10ph and H3S28ph occur at transcription start sites (+/- 1kb) in mitotic mESC cells. Previous literature has interpreted histone phosphorylation at serine 10 and 28 primarily in the context of interphase gene regulation, rather than in mitosis. For example, H3S10 phosphorylation has been observed to enrich in specific

regions in asynchronous cells, coinciding with immediate-early gene transcriptional activation, when cells were stimulated with light, growth factors or early differentiation signalling in rat suprachiasmatic nuclei, mouse fibroblast cells or rat ovarian follicle cells respectively (Mahadevan *et al.*, 1991; DeManno *et al.*, 1999; Crosie *et al.*, 2000). Mammalian studies have also shown H3S28 phosphorylation in interphase displacing polycomb group proteins, effectively activating polycomb target genes, in response to stress and developmental signalling (Gehani *et al.*, 2010; Lau and Cheung, 2011; Josefowicz *et al.*, 2016). The majority of these studies use asynchronous cell populations to study interphase. However, we note that asynchronous cell populations are composed of multiple cell cycle phases, and a small proportion of any given asynchronous population are mitotic. Given the strong increase in levels of histone H3 phosphorylation during mitosis, it is conceivable that the phosphorylation enrichments interpreted as interphase-specific are actually due to contaminating mitotic cells. We were able to utilise the quantitative nature of MINUTE-ChIP-seq combined with FACS cell sorting to assess differences in H3S10ph and H3S28ph enrichment between mitotic samples and G1 and G2 interphase samples in mouse embryonic stem cells. We found that both H3S10ph and H3S28ph enrich at promoters in mitosis, and that this enrichment is lost in G1 and G2. Although, many of these previous studies only observed H3S28ph enrichment at early-transcribed genes following stimulation of cells, for example using external growth or stress stimuli. Our MINUTE-ChIP-seq experiments did not involve any stimulation of cells, and therefore it is important to caveat that we may not have seen interphase-specific promoter enrichments that might only occur following cell stimulation.

Metagene analysis did not observe any promoter peaks in FACS-sorted G1 or G2 cell samples, and overall phosphorylation levels are dramatically reduced in both G1 and G2 (Figure 4.14). We do note that the reduced levels of G1 and G2 histone phosphorylation could have prevented detection of more subtle enrichment peaks, although we argue that the quantitative power of MINUTE-ChIP-seq should account for this. Using alternative approaches, we also found co-localisations between histone phosphorylations and other regulatory histone modifications and TFs changed significantly genome-wide depending on cell cycle phase (Figure 4.17), suggesting that mitotic regulatory networks of histone phosphorylations could affect regulatory functions.

4.3.3 H3S10ph and H3S28ph co-localisation networks can be used to assess the methyl-phos switch model

Through our partial correlation networks, we found that H3S10ph and H3S28ph show significant co-localisations with numerous epigenetic markers with known regulatory roles, which again are specific to mitosis and change dramatically compared to G1 and G2. A myriad of interesting research avenues could be expanded from these co-localisation networks, depending on the regulatory markers and cell cycle phases of interest. Here we highlight one key regulatory mechanism that both H3S10ph and H3S28ph have often been implicated in through previous studies across multiple cell types: the methyl-phos switch model. In this model, histone phosphorylation displaces reader proteins and/or prevents their binding to adjacent histone methylations (H3K9me3 and H3K27me3 respectively). H3S28ph has been shown to prevent binding of proteins such as demethylase UTX, corepressor complex HDAC, and polycomb group proteins, from adjacent H3K27me3 in both interphase and mitotic studies (e.g. Sawick and Seiser, 2014; Kruidenier *et al.*, 2012; Sengoku and Yokoyama, 2011; Lau and Cheung, 2011; Gehani *et al.*, 2010), while H3S10 phosphorylation has been proposed to displace heterochromatin protein HP1 from H3K9me3 (for review see Kouzarides, 2007; Wang and Higgins, 2013). Our MINUTE-ChIP-seq dataset included samples targeting H3K9me3 and H3K27me3, and therefore we are able to contribute evidence towards the methyl-phos switch discussion. Our co-localisation analyses show that H3S10ph correlated with H3K9me3 in G1 and G2 cells at active promoters, and in mitosis at bivalent promoters, which could support a methyl-phos switch mechanism. H3S28ph also positively correlates significantly with H3K27me3 a bivalent promoters in G1. These positive correlations for H3S10ph and H3S28ph with their adjacent methyl modifications interestingly contrasts with previous findings by Harris *et al.*, (2023) of an anti-correlation between another histone phosphorylation, H3T3ph, and its adjacent mark H3K4me3 in HeLa cells, arguing against the methyl-phos model for this H3T3 phosphorylation. However, we cannot conclude whether a methyl-phos switch mechanism occurs in other cell cycle phases, or at promoters other than those in states 16 and 18. Previous studies assessed global H3 phosphorylation, often using biochemical assays of cell lysates, as opposed to our closer focus on promoter regions in specific regulatory states. Our co-localisation analysis can easily be expanded to investigate promoters in all 20 chromatin states, in order to further assess the methyl-phos switch question. Based on our findings, we suggest that a switch could occur at regions in specific regulatory states.

4.3.4 Hypothesising roles for mitotic H3S10ph and H3S28ph peaks at bivalent and active promoters

The second main aim of the chapter was to integrate H3S10ph and H3S28ph MINUTE-ChIP-seq data with chromatin regulatory state data to place histone phosphorylation enrichment within a regulatory context. By isolating subsets of promoters in each of the 20 chromatin states (as defined by Juan *et al.,* 2016), we revealed significant enrichment peaks in both H3S10ph and H3S28ph specifically at promoters in active promoter state 16, and in bivalent promoter state 18. These promoters showed clear histone phosphorylation profile peaks in metagene enrichment profiles and heatmaps (Figure 4.11), and particularly bivalent promoters contained a high proportion of MACS2-called peaks in both H3S10ph and H3S28ph (Figure 4.10). Furthermore, other chromatin states characterised by active markers and higher accessibility, such as state 15 and 17, did not show histone phosphorylation peaks, confirming that the enrichment peaks observed were not likely to be an artefact of more active/accessible regions. These findings contrast the evidence presented by Meel *et al.,* in their recent 2024 study, where they reported that H3S10ph enriched in broader "islands" specifically at genes in more open "euchromatic" regions in prometaphase, and that these genes were then transcription activated upon mitotic exit. In our studies, we do not find that H3S10ph enriches in more active states compared to more repressive states; on the contrary, H3S10ph did not show enrichment at gene promoters in chromatin states 15 or 17, associated with active transcription, and did show enrichment in state 18 regions, where chromatin is characteristically more compact and repressed and is enriched for polycomb-associated H3K27me3. Moreover, while Meel found that the spatial distribution of H3S10ph in mitosis aligned with asynchronous H3S10ph distribution, we found that H3S10 phosphorylation distribution varied significantly between Mitosis and Interphase samples, as discussed above. From our findings, H3S10ph may play a more dynamic role in mitotic chromatin regulation, that may contribute towards both activation and repression of transcription, which may vary between cell cycle phases. However it is important to consider that studies' findings may differ depending on numerous factors, such as cell type, antibodies used, method of cell cycle phase isolation.

We revealed significant promoter enrichment of H3S10ph and H3S28ph in mitosis that were dependent on the regulatory state of the chromatin (Figure 4.11). While we do highlight some significant functional enrichment pathways identified as over-represented in these genes with

H3S10ph and H3S28ph peaks, no clear single functional pathway became apparent. The molecular mechanisms by which H3S10ph and H3S28ph might contribute to regulation of these genes requires careful interpretation and we encourage further investigation. One hypothesis is that H3S10ph and H3S28ph may promote transcription of these genes during mitosis, contributing to mitotic gene regulation. However, transcription is greatly reduced globally during mitosis, and critically, our analysis does not assess nascent transcription during mitosis. The chromatin states we used to classify promoters are trained using asynchronous ChIP-seq data, and so the spatial distribution of chromatin states could look different in mitosis. On the other hand, there is little to no evidence to our knowledge that methylations change during mitosis compared to their interphase distribution, and furthermore our MINUTE-ChIP-seq data do not show significant changes in key markers, including H3K4me3 and H3K27me3, during mitosis. It is therefore reasonable to assume that chromatin states regions remain the same during mitosis. Nevertheless, future validation studies should focus on performing nascent RNA-seq experiments in order to determine the transcriptional activity of promoters found to contain H3S10ph or H3S28ph peaks, in order to measure gene expression through mitosis and interphase, as will be discussed in detail in Chapter 6.

We can however suggest an alternative hypothesis. From this analysis, we can confidently conclude that H3S10ph and H3S28ph are significantly enriched during mitosis, at promoters which are subsequently in active state 16 and bivalent state 18 during interphase. Therefore, we suggest that mitotic histone phosphorylation might serve to *bookmark* these specific promoters to regulate the timing of their transcriptional activation upon mitotic exit. Again, nascent RNAseq experiments would be crucial further studies in order to examine transcriptional re-activation timings at these promoters of interest during mitotic exit.

Our co-localisation findings also create a further hypothesis; given the significant positive correlation seen between H3S10ph/H3S28ph and H3K9me3/H3K27me3 respectively, these histone phosphorylations could serve to “flip” a methyl-phos switch specifically at these promoters of enrichment, during mitosis.

4.3.5 Future perspectives

While we have primarily assessed protein-coding gene functions here, due to our focus on TSS enrichments, it is also likely that both H3S10ph and H3S28ph play a more broad and dynamic role in mitotic chromatin regulation. Mitotic histone phosphorylation could contribute to other processes such as mitotic exit transcription timing of certain genes, rather than regulating individual gene families. As another example, both H3S10ph and H3S28ph have previously been implicated in more structural regulatory roles during mitosis (e.g. Neurohr *et al.,* 2011; Gehani *et al.,* 2010; Hirota *et al.,* 2005; Wei *et al.,* 1999). To expand analysis along this line, future work could focus instead on 3D genome analyses, such as Hi-C, to investigate whether H3S10ph or H3S28ph might appear in clusters. A simple example would be to analyse the loci of genes found here to have H3S10ph and H3S28ph enrichment: do these genes cluster? Alternatively, future analysis could assess the interaction of H3S10ph and H3S28ph with other landmarks more associated with structural regulation in mitosis, such as cohesin or Heterochromatin Protein 1.

We do caveat that some of these analyses might prove difficult currently in direct relation to this project, because there is limited data available for this mESC cell line in mitosis, compared to more commonly used mitotic human cell lines such as HeLa. It is also important to note that different studies have produced contrasting findings as to the distributions and functions of both H3S10ph and H3S28ph, in different cell lines. It is conceivable, for example, that the phosphorylation enrichment seen at bivalent promoters is seen here because embryonic stem cells are likely to have more bivalency. Other differentiated cell lines might lose this enrichment. Whether our findings are conserved across eukaryotes remains unknown, and we encourage future research performing similar integrative quantitative analysis of histone phosphorylations in other cell lines and species.

4.3.6 Conclusion

To conclude, we present a quantitative dataset for H3S10ph and H3S28ph across mitosis, G1 and G2 integrated with chromatin regulatory state data, and reveal for the first time mitotic histone phosphorylations peak at active and bivalent promoters. We suggest that H3S10ph and H3S28ph may contribute to the dynamic regulation of bivalent promoters during mitosis, and propose they may serve to bookmark genes through mitosis that are actively transcribed in interphase. In order

to consolidate our findings and test hypothesised bookmarking roles, we suggest future experimental validation using nascent RNA-seq techniques to analyse the timings of transcription and transcriptional activation during and exiting mitosis. We also encourage analyses in alternative cell lines and species to expand this investigation. We emphasise that only by annotating the genome by regulatory chromatin states, were we able to isolate and reveal histone phosphorylation peaks at promoter regions that have not been detected before in mitotic studies. We encourage future analyses of abundant histone phosphorylations, and indeed other histone modifications, to include assessment of the epigenetic landscape as crucial context when interpreting histone modification distribution enrichments and functional analyses.

Chapter 5: H3T3ph, a phosphorylation required for correct mitotic chromosome segregation, shows promoter enrichment in the absence of H3K4me3 in HeLa S3 cells.

Summary

Recently, we published findings on the interplay between H3K4me3 and another histone phosphorylation, histone H3 threonine 3 (H3T3ph). H3T3ph is deposited by the kinase haspin, enriching at the inner centromeric region of the chromosomes during prometaphase and metaphase of mitosis, with some distribution spread along the chromosome arms. H3T3ph has been previously reported to play vital roles in recruitment/localisation of the chromosome passenger complex (CPC) to the centromeric regions, contributing to CPC-mediated correction of erroneous kinetochore-microtubule binding. This ensures the spindle correctly attaches to the kinetochores, allowing correct separation of sister chromatids as mitosis progresses. Absence of haspin, the mitotic kinase of H3T3ph, resulted in chromosome mis-segregation (Kelly et al., 2010; Wang et al., 2010; Yamagishi et al., 2010). H3T3ph has also been proposed to have a role in a phospho-methyl switch mechanism with its adjacent modification, H3K4me3; in vitro peptide studies have found H3T3 phosphorylation prevented binding of H3K4me3-reader proteins (e.g.Flanagan *et al.*, 2005; Varier *et al.*, 2010). However, research in our group used single-end ChIP-seq in HeLa S3 cell line to show that H3K4me3 and H3T3ph show strong anti-correlation, and that H3K4me3-binding transcription factors were able to recognise H3K4me3 in the presence of H3T3 phosphorylation (Harris et al., 2023). In this chapter, we expand upon these H3T3ph studies to investigate H3T3ph enrichment in promoter regions. In an unpublished finding, we report that H3T3ph enrichment peaks are seen specifically at promoters where H3K4me3 is absent. We describe the processing and analysis of unpublished paired-end X-ChIP-seq in HeLa S3 cells, and perform functional enrichment analyses of these H3T3ph-enriched, H3K4me3-negative promoters. The findings of this chapter nicely complement those of Chapter 4, demonstrating another histone phosphorylation enriching at a subset of promoters depending on regulatory state, this time in human cells.

5.1 Introduction

In 2005 it was discovered that histone H3 is phosphorylated at threonine 3 (H3T3ph), deposited by the kinase haspin, and that this phosphorylation was specific to mitosis (Dai *et al.*, 2005). A methyl-phos switch mechanism has been proposed, largely based on *in vitro* observations, whereby H3K4me3-reading proteins are displaced by adjacent H3T3 phosphorylation (e.g. Flanagan *et al.*, 2005).

Our lab recently published our findings investigating the *in vivo* distributions of both H3T3ph and H3K4me3, using single-end ChIP-seq (Harris *et al.*, 2023). The results of this ChIP-seq are illustrated in **Figure 5.1**. H3T3ph showed centromeric enrichment (**Figure 5.1A),** consistent with previous immunofluorescence microscopy results. Metagene analysis, performed by Rebecca Harris, showed the expected strong enrichment of H3K4me3 centred at Transcription Start Sites (TSSs), as shown in **Figure 5.1B.** Interestingly, Harris *et al.*, then showed a strong anti-correlation between H3T3ph and H3K4me3, as illustrated in **Figure 5.1C and D.** Based on these findings, they argued against a methyl-phos switch *in vivo* in HeLa cells which would require the co-localisation of the two modifications.

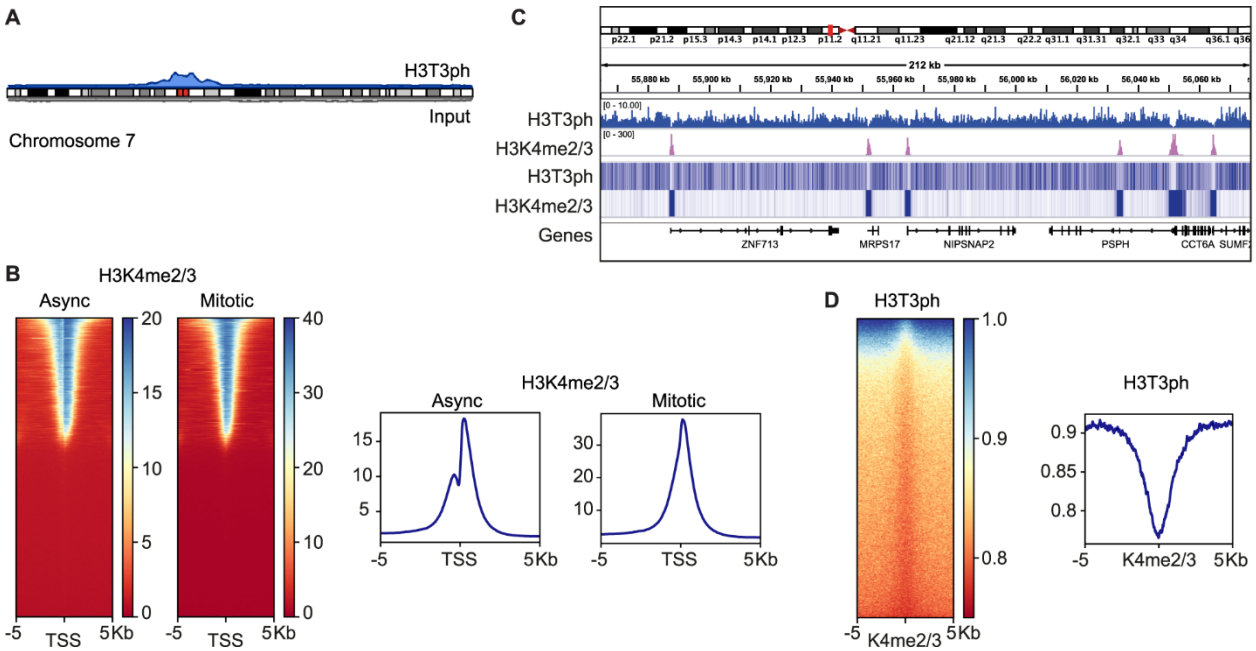


Figure 5. 1: Metagene analysis of single-end ChIP-seq, as seen in Harris et al., 2023.

A) Single-end ChIP-seq of H3T3ph performed by Rebecca Harris showed centromeric enrichment: chromosome 7 is shown as an example. B) Single-end H3K4me2/3 ChIP-seq shows enrichment at global Transcription Start Sites (TSSs) in both asynchronous and mitotic HeLa S3 cells. Mitotic loss of the nucleosome-depleted regions can be seen. C) Integrative Genomics Viewer (IGV) of a region of chromosome 7: single-end H3T3ph and H3K4me2/3 ChIP-seq, shown as coverage tracks (input-normalised) (top) and heatmaps (below). D) H3T3ph shows anti-correlation with H3K4me2/3. Normalised ChIP signal indicated on vertical axes.

In this Chapter, we further analyse the exact distribution of H3T3ph. Firstly, we describe the use of CIDOP-seq technique to assess the ability of transcription factors to bind H3K4me3 in the presence of H3T3ph, further contributing to our evidence against the methyl-phos switch. We also detail the processing, quality control, alignment and analysis of paired-end ChIP-seq of H3T3ph. This paired-end dataset improves our ability to align to repetitive centromeric regions where H3T3ph is known to enrich.

We also utilised this dataset to expand on our promoter-focussed analysis from Chapter 4, to see if H3T3ph shows any enrichment at promoters in human HeLa cells, when chromatin regulatory

state is taken into account. In an unpublished finding, our analysis reveals enrichment of H3T3ph at specific promoters in the absence of H3K4me3.

5.2 Results

5.2.1 CIDOP-seq shows TAF3 PHD finger protein is able to bind H3K4me3 in the presence of H3T3ph, arguing against a methyl-phos switch

To further assess the likelihood of a methyl-phos switch mechanism, the ability of transcription factors to recognise and bind H3K4me3 on mitotic chromatin in the presence of H3T3ph was investigated. Previous *in vitro* studies using synthetic peptides found that H3T3ph prevented binding of H3K4me3 reader proteins, supporting a methyl-phos switch (Ali et al., 2013; Flanagan et al., 2005; Garske et al., 2010; Gatchalian et al., 2016; Southall et al., 2009; Varier et al., 2010b). One such H3K4me3-recognising protein is the TAF3 PHD finger domain of TFIID transcription factor. Similarly to previous *in vitro* observations, we found that GST-TAF3 PHD finger fusion protein binding of H3K4me2/3 was strongly inhibited by adjacent H3T3ph in synthetic peptides (Harris *et al.*, 2023).

We then used a technique called chromatin-interacting domain precipitation and sequencing: CIDOP-seq. Instead of an antibody probe as in ChIP-seq, CIDOP uses a recombinant GST-TAF3 PHD finger fusion protein to bind sheared chromatin fragments *in vitro*. The chromatin-bound GST-TAF3 fusion protein can then be precipitated, and bound chromatin fragments sequenced to determine genome-wide binding sites. This allowed them to determine the binding sites of TAF3 in the absence of other TFIID components, and critically this technique allowed us to bypass the need for antibodies that can be impacted by cross-reactivity or blocking by adjacent modifications. CIDOP-seq was performed by Rebecca Harris on chromatin from mitotic HeLa S3 cells, to investigate whether this H3T3ph inhibition of H3K4me3 binding is also seen *in vivo*.

I then processed raw sequencing data from these CIDOP experiments, including quality control, alignment to genome, normalisation to input samples and scaling read coverage. Aligned, normalised data was then analysed for TSS enrichment, using Deeptools to generate enrichment

profiles and heatmaps across 10 kb windows centred at TSSs, both genome-wide and at centromere-proximal regions. The enrichment plots we generated in contribution to this publication are displayed in **Figure 5.2** below, as seen in Harris *et al.,* 2023.

As can be seen in **Figure 5.2,** CIDOP-seq from HeLa S3 chromatin found that GST-TAF3 PHD finger binding to H3K4me2/3 remained consistent across the genome, including at centromere-proximal regions where H3T3ph is known to be enriched *in vivo,* showing similar enrichment profile to H3K4me2/3 antibody ChIP-seq (Figure 5.1B). The binding enrichment seen here This research demonstrated that H3K4me3 reader protein TAF3, of TFIID, is able to bind H3K4me3 in mitosis in HeLa S3 cells, and this binding is not affected by H3T3ph. CIDOP-seq showed that enrichment of TAF3-PHD finger binding of H3K4me3 was not reduced in centromere-proximal regions, and displays mitotic spread into nucleosome-depleted regions as would be expected with mitotic nucleosome incursion.

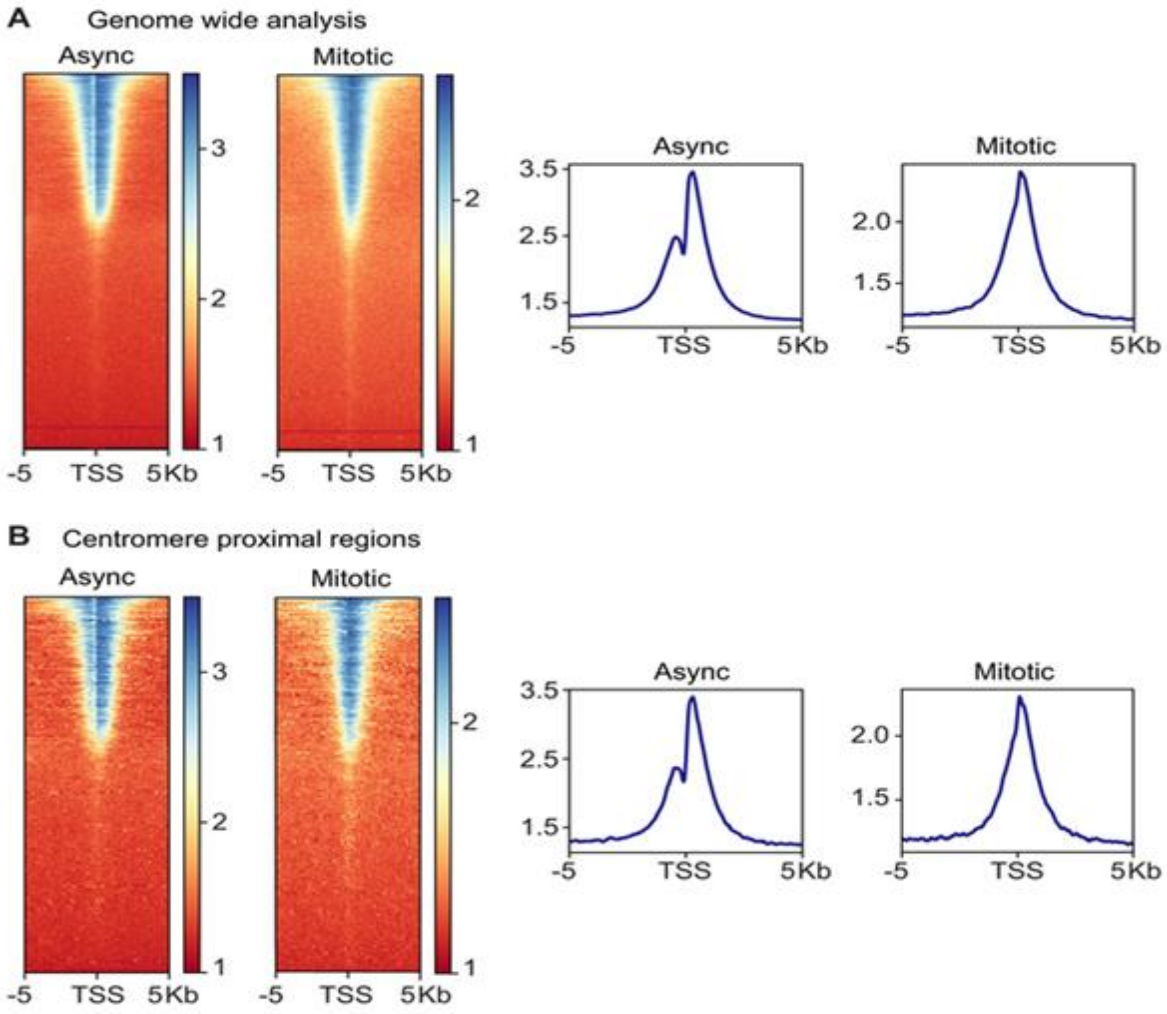


Figure 5. 2: CIDOP-seq analysis of TAF3 PHD enrichment at TSSs.

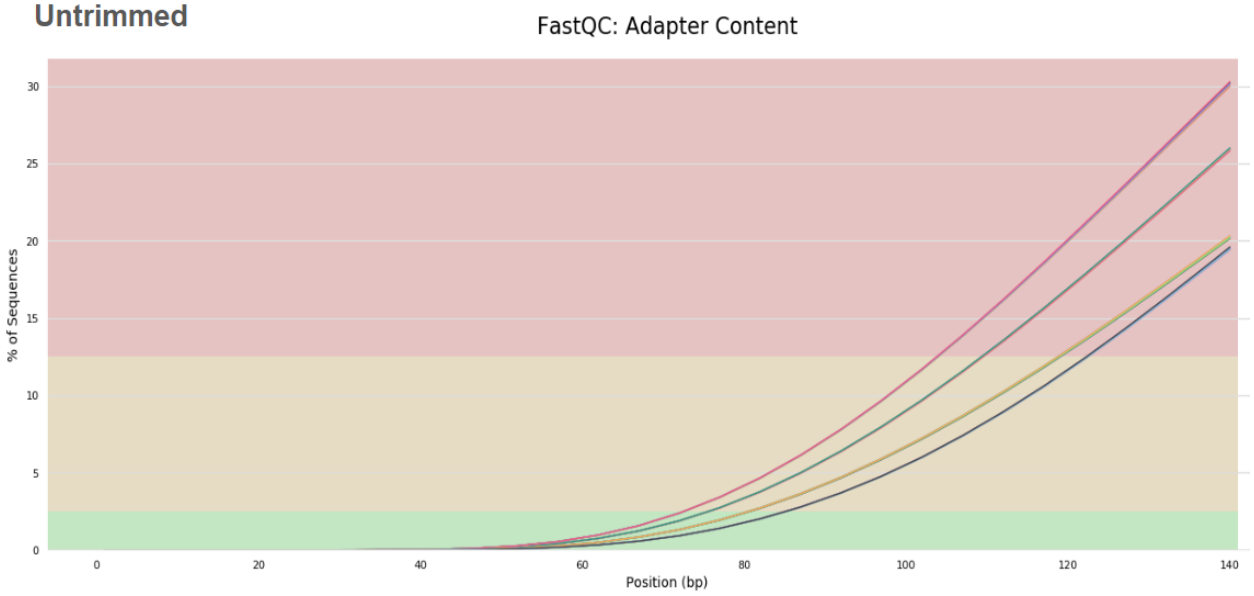
A: Analysis of TAF3 PHD enrichment genome-wide, centred at the TSSs. B TAF3 PHD enrichment at centromere proximal TSSs. *Heatmaps (left) and metagene plots (right) show TAF3-PHD binding (y axis shows enrichment scores) across 10 kb regions centred at TSSs for both asynchronous and mitotic HeLa S3 cell chromatin.*

5.2.2 Paired-end ChIP-seq produced high-quality sequencing of mitotic H3T3ph in HeLa S3 cells

For this project, we then aimed to expand upon H3T3ph analysis through study of an unpublished, paired-end H3T3ph ChIP-seq dataset in HeLa S3s. This paired-end dataset allowed us to replicate and consolidate on published findings, and will enable improved alignment of H3T3ph to repetitive centromeric regions where H3T3ph is known to enrich.

Paired-end ChIP raw sequencing data in fastq file format for H3T3ph in HeLa S3 cells, with two independent replicates, was produced by Rebecca Harris in the Higgins lab. As part of this project, I then quality assessed this raw sequencing using fastqc software, and sequencing was trimmed using Trimmomatic software to remove contaminating adapter sequences, as shown in **Figure 5.3**. This trimming retained >50 million high-quality sequence reads per sample.

Untrimmed


FastQC: Adapter Content
% of Sequences
Position (bp)
0
5
10
15
20
25
30
0
20
40
60
80
100
120
140


FastQC: Mean Quality Scores
Phred Score
Position (bp)
0
5
10
15
20
25
30
35
0
20
40
60
80
100
120
140

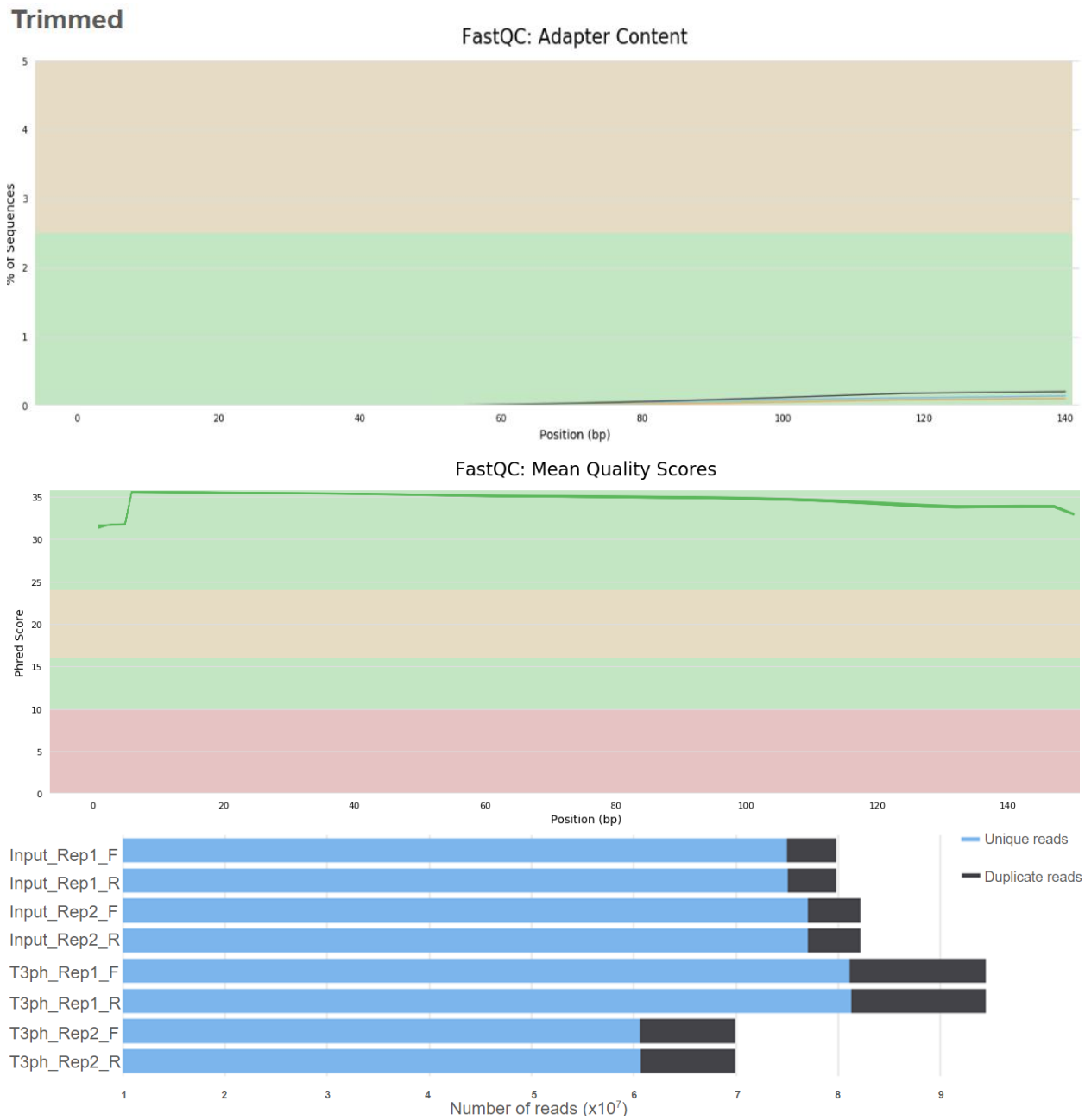


Figure 5. 3: Trimming of raw sequencing produced high-quality sequencing with >50 million sequence reads per sample.

Shown are the raw sequencing quality scores and adapter content per bp (see plot titles) of sequencing samples before trimming (top) and after trimming (bottom). Below, barchart (blue) shows number of sequencing reads for each sample replicate (Rep1 and Rep2), forward and reverse strand sequencing (F and R). Removal of adaptor contamination through trimming can be seen.

We then aligned H3T3ph paired-end sequencing reads to the hg38 human reference genome. Aligned H3T3ph sequencing was normalised against the corresponding Input sample for each independent replicate, and scaled to read count. The fingerprint enrichment shown in **Figure 5.4** shows enrichment of H3T3ph sequence compared to input samples, indicating targeted immunoprecipitation.

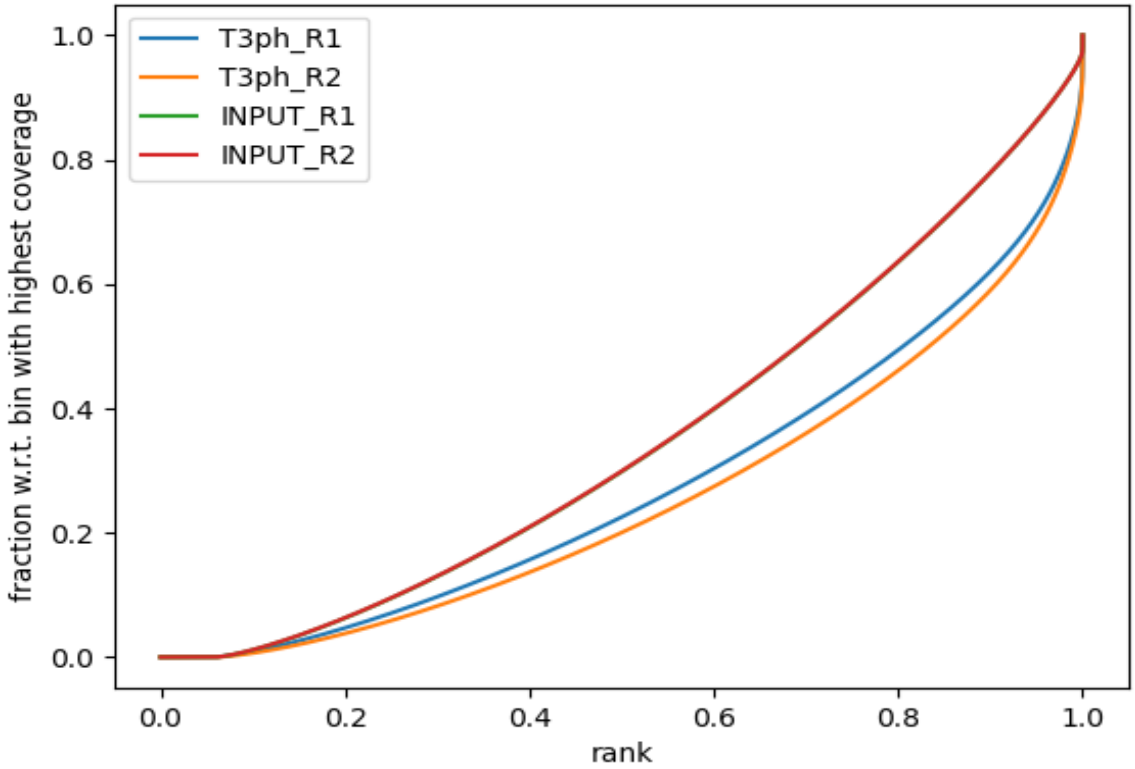


Figure 5. 4: Fingerprint enrichment plot of H3T3ph replicates 1 and 2 and corresponding Input sequencing.

Plot indicates the number of bins along the genome (x axis, "rank") and the percentage of sequencing reads contained within those bins (y axis, "fraction wrt bin with highest coverage"). An "elbow" to the right of the plot indicates a high proportion of sequencing is aligned to a small number of genome bins, indicating a small subset of genome regions contain high, narrow enrichment.

Following alignment, Integrative Genomics Viewer (IGV) was used to visualise the distribution of sequencing reads for normalised H3T3ph, displayed below (**Figure 5.5**). Sequencing showed

enrichment at the centromeric regions with some spread along the chromosome arms in all chromosomes (example chromosomes selected at random shown below), consistent with previous single-end sequencing findings of H3T3ph distribution (see Harris *et al.*, 2023, and Figure 5.1A). Together, these findings suggest that this paired-end ChIP-seq is of high quality and was successful in precipitating H3T3ph sites.

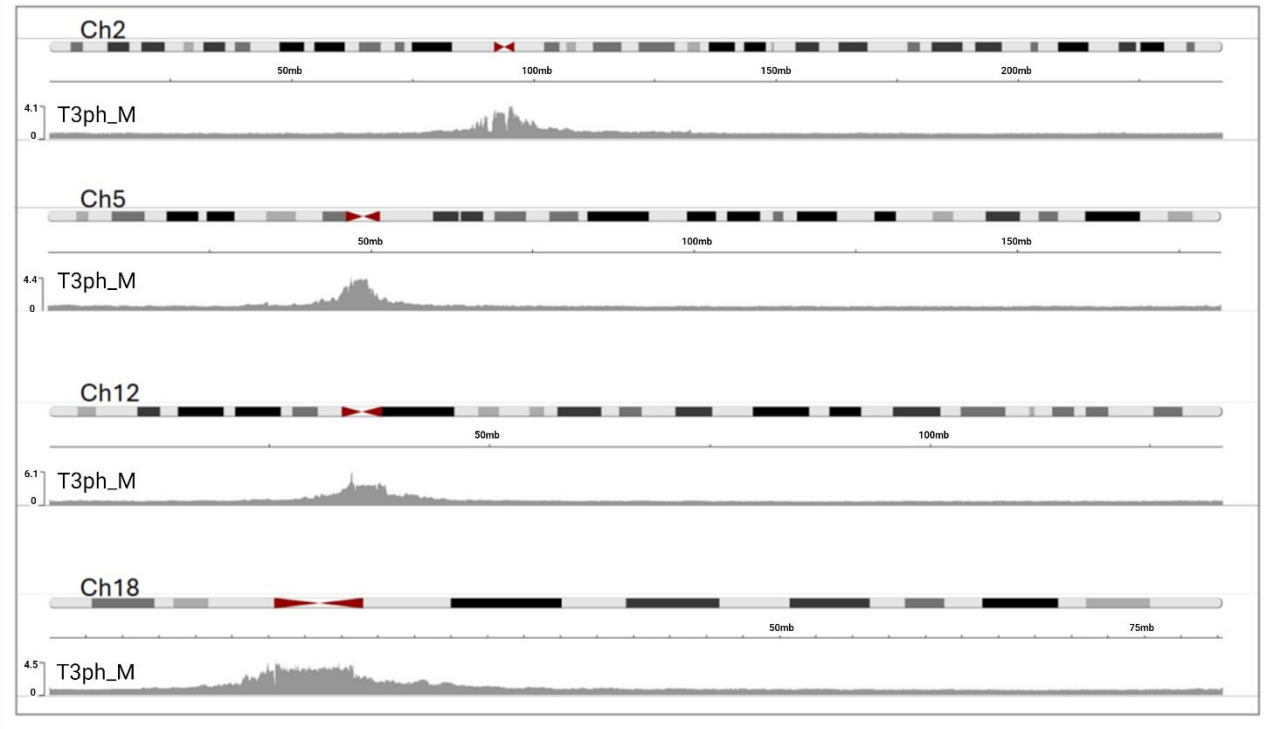


Figure 5. 5: IGV visualisation of sequencing coverage for H3T3ph paired-end ChIP-seq.

H3T3ph paired-end ChIP-seq (T3ph_M) was aligned to the hg38 reference genome and normalised to the corresponding input sample sequence, and scaled to read count. Example chromosomes 1, 5, 12 and 18 are displayed. H3T3ph sequencing coverage track shows enrichment at centromeric regions(red arrows).

These quality assessments found that the paired-end sequencing data looks as expected and similar to previous single-end sequencing findings. This paired-end sequencing will be of particular use compared to single-end data for future work investigating centromeric enrichments, a topic of

particular interest for the centromeric mark H3T3ph. However, for this project we focussed on assessing the distribution of H3T3ph at promoters, discussed below.

5.2.3 Metagene analysis of paired-end ChIP-seq reproduces findings of subtle H3T3ph enrichment at transcription start sites and anti-correlation with H3K4me3

Our previous analysis detailed in Chapter 4 found that H3S10ph and H3S28ph both enriched at a subset of promoters during mitosis in mESCs, and that it was crucial to take into account chromatin regulatory state in order to identify these enrichments. The paired-end H3T3ph ChIP-seq dataset analysed in this chapter therefore provided an interesting opportunity; we sought to test whether this other phosphorylation, H3T3ph, might also show promoter enrichments at a subset of promoters – this time in human cells.

To begin this analysis, normalised, paired-end H3T3ph alignments were used to generate metagene enrichment plots and heatmaps of H3T3ph, centred both at genome-wide transcription start sites (TSSs), and centred at H3K4me3. For H3K4me3-centred plots, MACS2 was used to call H3K4me3 peaks from single-end H3K4me3 ChIP-seq previously generated by Rebecca Harris and included in publication (Harris *et al.,* 2023). H3T3ph was then aligned to these H3K4me3 MACS2 peaks. The resulting heatmaps and enrichment profiles are shown in **Figure 5.6,** with H3T3ph showing consistent enrichment when aligned to H3K4me3 as was seen in single-end analysis (Figure 5.1D). For all subsequent plots, both replicates were visualised to confirm reproducibility, before merging replicates to produce the displayed plots.

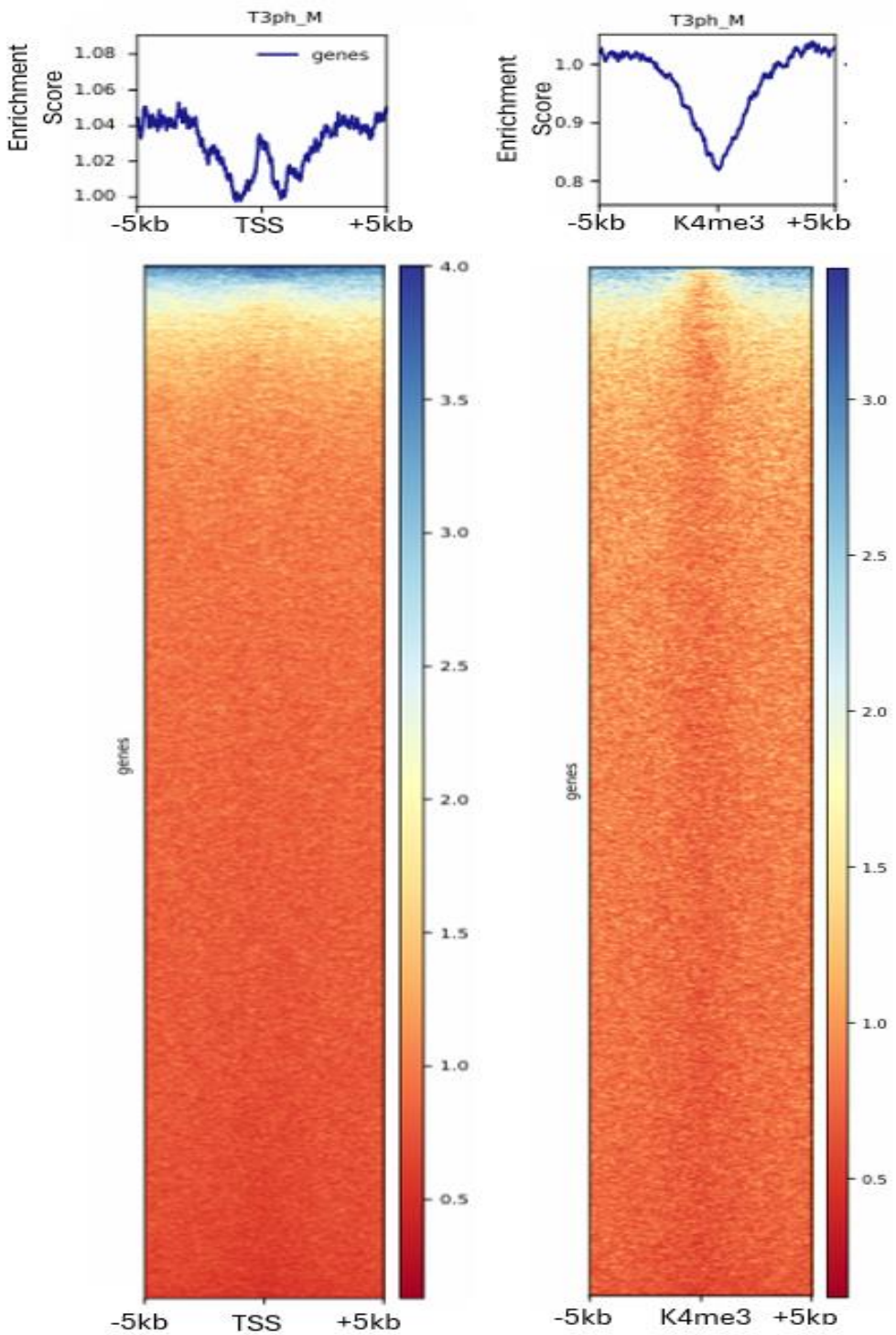


Figure 5. 6: Enrichment heatmaps and profile plots of paired-end H3T3ph ChIP-seq centred at genome-wide TSSs (left), and centred at H3K4me3 (right).

Enrichment scores provided as vertical scales.

This analysis reproduces and confirms the findings that H3T3ph anti-correlates with H3K4me3, as described in Harris *et al.*, 2023. Moreover, we also found that H3T3ph shows overall depletion at genome-wide TSSs, likely as a result of H3K4me3 anti-correlation. However, interestingly, an enrichment of H3T3ph can be seen within the overall depletion at TSSs (Figure 5.6, left). We hypothesised that there may be a subset of TSSs that lack H3K4me3 at which H3T3ph is enriched.

5.2.4 H3T3ph shows enrichment at promoters in the absence of H3K4me3

We hypothesised that the small enrichment of H3T3ph enrichment seen at genome-wide TSSs might be as a result of an H3T3ph enrichment at a subset of TSSs lacking H3K4me3. To test this, we separated TSSs into those which are H3K4me3-"positive" and H3K4me3-"negative". MACS2 was used to call peaks in H3K4me3 using single-end H3K4me3 ChIP-seq, and TSSs categorised as H3K4me3-positive or –negative based on the presence or absence of H3K4me3 called peaks respectively. H3T3ph paired-end ChIP-seq enrichment heatmaps and profile plots were then generated centred at H3K4me3-positive, and H3K4me3-negative, TSSs, displayed in **Figure 5.7**. This analysis shows a subtle but clearly observable enrichment peak in H3T3ph specifically at TSSs that do not contain H3K4me3 enrichment peaks.

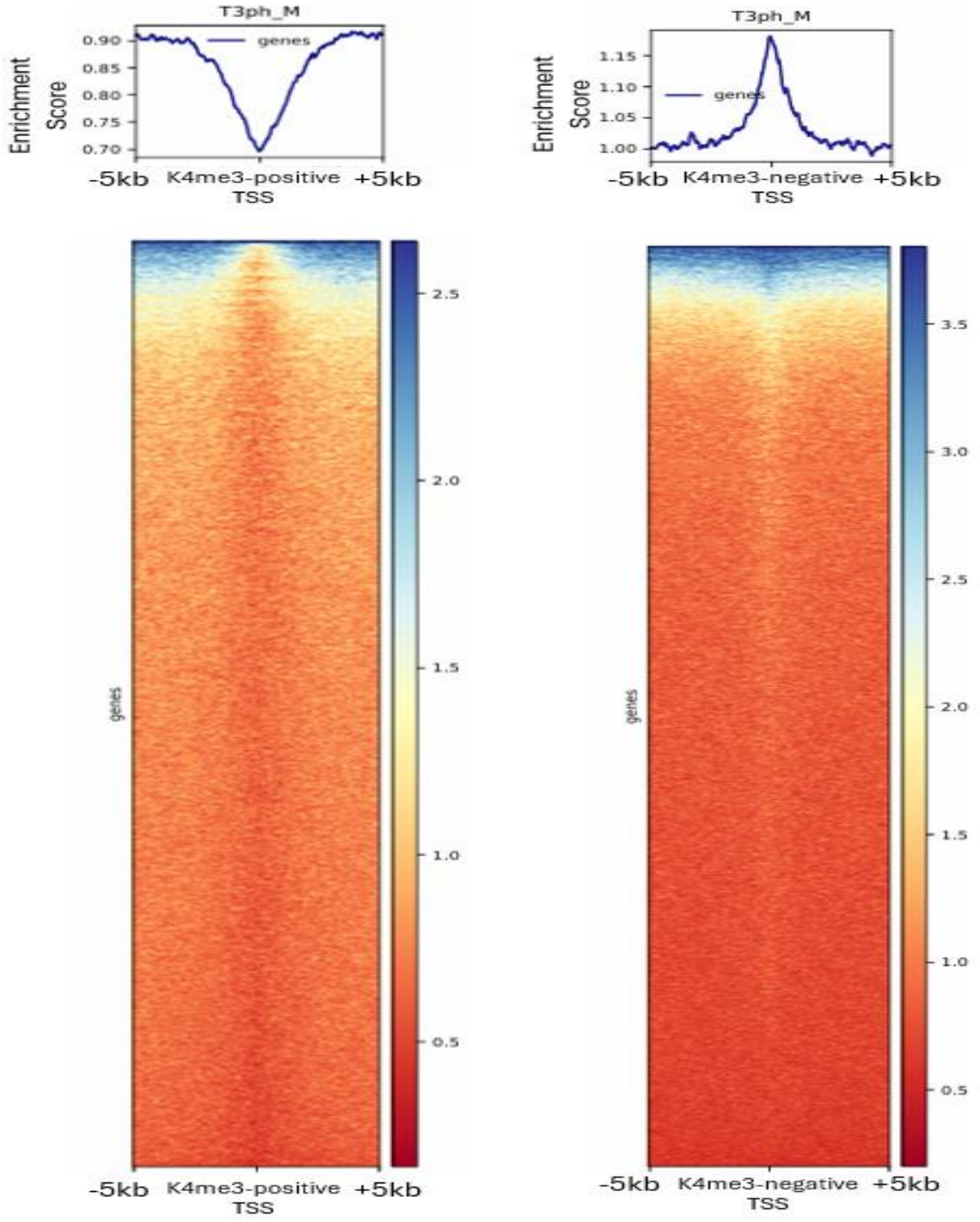
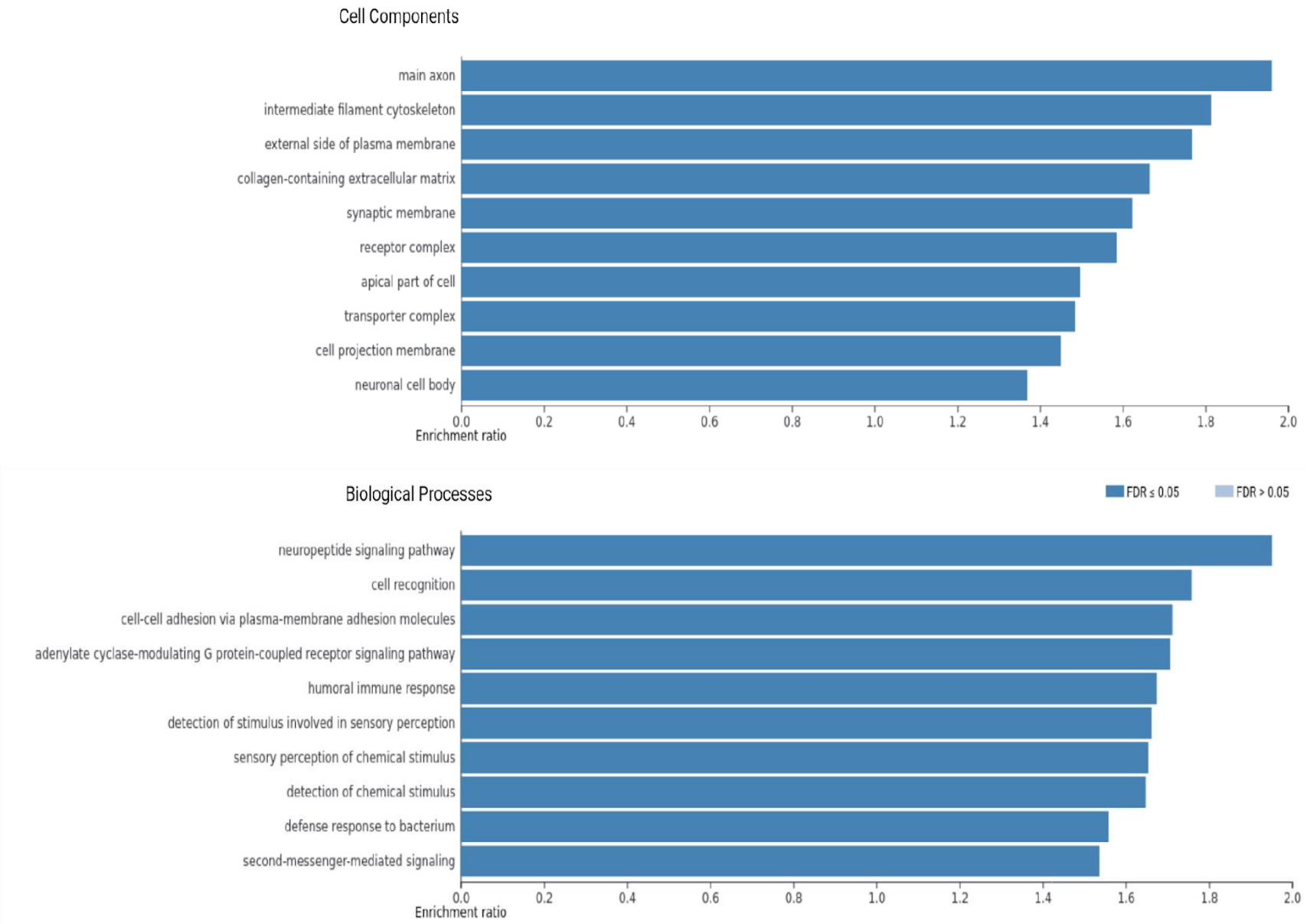


Figure 5. 7: Enrichment heatmaps and profile plots of paired-end H3T3ph ChIP-seq centred at H3K4me3-positive (left) and H3K4me3-negative (right) TSSs.

This analysis clearly demonstrates that mitotic H3T3ph shows enrichment at promoters that is only revealed when TSSs are subset based on the context of other regulatory markers.

5.2.5 Functional gene set enrichment analysis for H3T3ph promoter peaks

We then sought to investigate the putative function of these subtle H3T3ph peaks at H3K4me3-negative promoters. Firstly, we performed over-representation analyses for all TSSs absent of H3K4me3 MACS2 peaks, against the reference hg38 genome, to provide context of the overall functional enrichments for H3K4me3-negative TSSs regardless of H3T3 phosphorylation. The functional enrichments found based on Webgestalt functional databases are shown below, **Figure 5.8**. These findings do not necessarily highlight a single clear functional enrichment or role for H3K4me3-negative promoters, but provide important context for interpreting functional enrichments for H3T3ph performed subsequently.


Cell Components
main axon
intermediate filament cytoskeleton
external side of plasma membrane
collagen-containing extracellular matrix
synaptic membrane
receptor complex
apical part of cell
transporter complex
cell projection membrane
neuronal cell body
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
Enrichment ratio
Biological Processes
FDR ≤ 0.05
FDR > 0.05
neuropeptide signaling pathway
cell recognition
cell-cell adhesion via plasma-membrane adhesion molecules
adenylate cyclase-modulating G protein-coupled receptor signaling pathway
humoral immune response
detection of stimulus involved in sensory perception
sensory perception of chemical stimulus
detection of chemical stimulus
defense response to bacterium
second-messenger-mediated signaling
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
Enrichment ratio

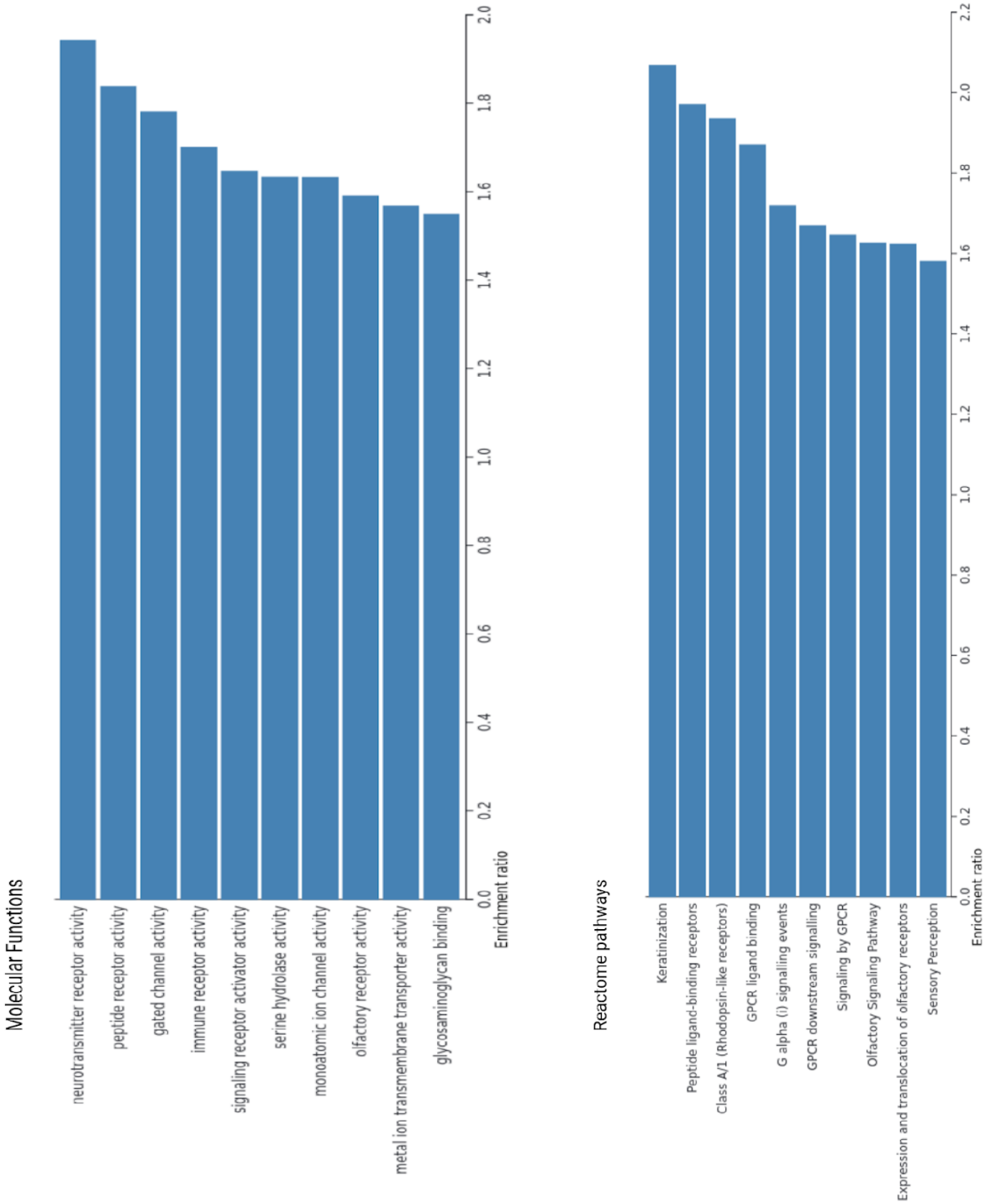


Figure 5. 8: Over-Representation Analysis (ORA) for H3K4me3-negative promoters found functional enrichments for numerous gene ontology sets.

Promoters (TSS +/- 1 kb) that do not contain H3K4me3 MACS2 called peaks were input into Webgestalt against the reference hg38 genome and ORA performed against all available gene ontology databases. Significant functional enrichments ($p<0.05$, FDR <0.05) were found for the biological components shown.

Gene Set Enrichment Analysis (GSEA) was then performed to identify any significant functional enrichments for H3T3 phosphorylation in this subset of H3K4me3-negative promoters where H3T3ph was shown to enrich. Parameters were set identically to analysis performed in Chapter 4, setting minimum gene set size to 10 as recommended. No functional enrichments were identified as significant, suggesting these H3T3ph promoter "peaks" are not functioning to regulate any one specific gene family or biological pathway.

5.3 Discussion

In this Chapter we describe the processing and metagene analysis of CIDOP-seq data, contributing to a study (Harris *et al.*, 2023) showing the TAF3 PHD finger domain is able to bind H3K4me2/3 sites independent of H3T3 phosphorylation in mitotic HeLa S3 cell chromatin. We subsequently expanded on this research, processing paired-end ChIP sequencing for H3T3ph in synchronised prometaphase HeLa S3 cells. Through alignment to the hg38 reference genome and metagene analysis, we were able to demonstrate that H3T3ph anti-correlates with H3K4me3, as seen in single-end sequencing in Harris *et al.*, 2023.

As well as supporting previous results, this paired-end sequencing will provide invaluable data for future analysis focussing on the centromeric regions. Our findings confirm this dataset as good quality, and centromeric enrichment seen on IGV indicates successful targeting of H3T3ph. H3T3ph shows enrichment at the centromeric regions with some spread along chromosome arms; these centromeres are composed of highly repetitive tandem repeat sequences, making accurate alignment of sequencing to the centromeric regions of the reference genome very difficult

(Altemose *et al.,* 2022). Depending on chosen alignment software parameters, sequenced reads are likely to either align many times to repetitive regions, or be filtered out if repetitive regions or multi-aligning reads are excluded. With the recent assembly of the full human centromere (Altemose *et al.,* 2022), bioinformatics techniques, software and algorithms are continually being developed to address the challenges of centromere alignment. Paired-end sequence provides advantages for aligning to repetitive regions, because sequencing of both forward and reverse DNA strands enables a more accurate alignment of each read to the genome; this is particularly helpful in aligning to repetitive regions like the centromere. Combined with advancing techniques such as long-read sequencing (e.g. Logsdon *et al.,* 2024) and use of unique *k-mer* algorithms designed for mapping in repetitive regions, the paired-end sequencing data presented in this Chapter can be used in future work to produce a more accurate map of mitotic H3T3ph distribution at the centromeric region.

Likely due to anti-correlation with H3K4me3, H3T3ph shows overall depletion at transcription start sites (TSSs). However, our metagene analysis shows a small enrichment within this; we hypothesised that this may be due to an enrichment of H3T3ph at a subset of TSSs which are absent of H3K4me3. By separating TSSs into those containing, or absent, H3K4me3 MACS-called significant peaks, we revealed a subtle but clear H3T3ph enrichment peak at H3K4me3-negative TSSs. To our knowledge, this promoter enrichment of H3T3ph has not been previously reported in human systems. We also conclude that this H3T3ph promoter enrichment is not likely to be due to chromatin accessibility artefacts, because enrichment is only seen at H3K4me3-negative promoters which are likely less accessible.

This analysis presents a nice complementary study to those in Chapter 4; whereas Chapter 4 focussed on H3S10ph and H3S28ph promoter enrichment in mouse cell lines, here we demonstrate promoter enrichment of another phosphorylation, H3T3ph, again dependent on regulatory state, this time in human cells. It is interesting that the H3T3ph promoter enrichment was detectable even in genome-wide analysis, seen in Figure 5.6, whereas H3S10ph and H3S28ph promoter enrichments were not seen genome-wide. This could indicate that the H3S10ph and H3S28ph promoter enrichments are more subtle; or perhaps could be a product of the high, genome-wide abundance of those H3S10ph and H3S28ph compared to H3T3ph centromeric enrichment.

Gene set enrichment analysis did not identify any significant functional enrichments in relation to these H3K4me3-negative promoter H3T3ph peaks, suggesting that H3T3ph is not functioning to regulate any specific gene family or pathway. H3T3ph could nevertheless play a broader regulatory role at these promoters during mitosis. Further analysis could be performed to assess the nature of these H3K4me3-negative TSSs; for example, are the TSSs containing H3T3ph peaks protein-coding, and/or principal isoforms? Gene-centred analysis may not be the most informative approach for H3T3ph, given H3T3ph enriches primarily at repetitive centromeric regions where there may be more non-protein-coding genes. Arguably, our findings align with previous literature where H3T3ph is far more clearly linked to structural, rather than genic, regulation. For example, we described in Chapter 1 the role of H3T3ph in recruiting CPC to the centromeres, contributing to correct chromosome segregation. Therefore, we encourage future analysis focussing on aligning this paired-end dataset to the centromere sequence; improving our knowledge of the exact distribution of H3T3ph may help improve our understanding of the structural roles and exact localisations of H3T3ph.

Additionally, it would be interesting to study the H3T3ph-enriched promoters in a more 3D approach; for example, Hi-C data, readily available in HeLa cell lines, could be integrated to assess the 3D arrangement and potential interactions of these genes. Do they show any clustering or long-range interactions? It would also be beneficial to look at the chromosomal distribution of these genes, to assess whether they are more centromeric, or further along chromosome arms.
Experiments could also be designed to assess whether these H3T3ph enrichments might be involved in transcriptional timing; nascent RNA-seq experiments could measure *de novo* transcription in different cell cycle phases, including mitosis and mitotic exit, to identify any patterns in transcriptional activation timing for these H3T3ph-enriched promoters. This could be used to test the hypothesis that H3T3ph may act to bookmark these promoters through mitosis.

We also recommend that further paired-end ChIP-seq be performed in support of this study. For example, generating paired-end ChIP-seq on an asynchronous, or isolated interphase HeLa S3 cell sample would allow us to further confirm successful H3T3ph targeting, as this phosphorylation has only been reported in mitotic cells. Also, H3K4me3 ChIP-seq data utilised in this analysis to generate MACS2 called peaks and perform metagene analyses was single-end data. H3K4me3

paired-end ChIP-seq would be more directly comparable with H3T3ph data, and single-end sequencing could potentially skew alignments.

We also propose that chromatin state analysis could be an interesting avenue to expand this research. The H3T3ph enrichment at H3K4me3-negative promoters demonstrates the value of placing histone phosphorylation distribution in the context of chromatin regulatory landscape, as was also demonstrated in Chapter 4. Chromatin state analysis describes the regulatory state of regions of the genome based on the relative enrichments of known regulatory markers. Assessing how H3T3ph distribution changes at promoters in different chromatin states - and particularly those states characterised by absence of H3K4me3 - may further our understanding of this phosphorylation in the context of gene regulation.

Chapter 6: Project Contributions and Future Perspectives

6.1 A synopsis: project outcomes and scientific contributions

The central aim of this project has been to improve our understanding of the exact distributions of abundant histone phosphorylations, and to investigate their putative functions in chromatin regulation in mitosis. Development of the CUT+RUN targeted immunoprecipitation method initially showed promise in mitotic HeLa S3 cells using anti-H3K4me3 antibodies; however, this relatively new technique requires further optimisation to successfully isolate histone phosphorylation chromatin sites in mitotic HeLa cells. Instead, we turned to the MINUTE-ChIP-seq approach. In collaboration with Simon Elsasser's group, we were able to utilise mitotic mESC sequencing data to quantitatively compare H3S10ph and H3S28ph levels in mitotic vs G1 and G2 cells. We report that these MINUTE-ChIP-seq data are high quality, and based on previous validation experiments we argue that MINUTE-ChIP-seq offers good quantitative accuracy in reflecting biological enrichment of these histone phosphorylations.

We found significant evidence that mitotic H3S10ph and H3S28ph are **not** uniformly distributed genome-wide**.** Focussing our analysis at promoter regions, we demonstrate that both H3S10ph and H3S28ph signal significantly varies depending on the regulatory state of the promoter. Critically, this enrichment pattern was only identified through our novel integrative approach combining MINUTE-ChIP-seq with chromatin regulatory state modelling data. ChromHMM chromatin state data previously generated by Juan *et al.,* 2016 allowed us to classify TSSs (+/- 1 kb) into one of 20 characterised chromatin states. We demonstrated that H3S10ph and H3S28ph, as well as other modifications with known regulatory roles (H3K4me3 and H3K27me3), significantly vary in almost all 20 chromatin states.

Moreover, we used partial correlation analysis to generate networks of significant co-localisations between H3S10ph, H3S28ph and 10 other regulatory markers; critically, partial correlations allowed us to take into account third or more other markers that might impact each correlation. We present a wealth of significant co-localisations between histone phosphorylations and regulatory markers that contribute to understanding of the regulatory networks which H3S10ph and H3S28ph

may play roles in specifically in mitosis. These co-localisation networks showed dramatic changes depending on both the chromatin state and the cell cycle phase. This further supports our argument that a) chromatin regulatory context is crucial when interpreting histone phosphorylation distributions, and b) histone phosphorylations H3S10ph and H3S28ph are involved in specific interactions with known regulatory markers that are specific to mitosis. These mitosis-specific findings potentially add a new avenue of histone phosphorylation research, which historically has largely studied gene regulation by histone phosphorylation only in the context of interphase (e.g. DeManno *et al.*, 1999; Gehani *et al.*, 2010; Komar & Juszczynski, 2020; Lau & Cheung, 2011; Mahadevan *et al.*, 1991). Importantly, these prior studies largely used asynchronous cell populations, but interpreted promoter histone phosphorylations as interphase events. We argue that asynchronous cell populations will contain mitotic cells, and that the proportion of mitotic cells in an asynchronous population will likely vary depending on numerous factors including cell type, experimental conditions and treatments used (e.g. stimulations). This arguably makes it difficult to interpret whether the histone phosphorylation enrichments seen in previous studies were actually due to contamination from mitotic cells with high histone phosphorylation levels.

We found that both H3S10ph and H3S28ph significantly are enriched at promoters, and specifically promoters which are in active and bivalent regulatory states in interphase. This was demonstrated through MACS2 binarisation, enrichment metagene profiling and heatmaps. To our knowledge, this is the first time that these abundant histone phosphorylations have been shown to enrich at promoters in mitosis; previous studies have only observed promoter enrichments in asynchronous, largely interphase cells (e.g. Mahadevan *et al.*, 1991; Lau & Cheung, 2011). In Chapter 5, we also demonstrate in human HeLa cells an enrichment of H3T3ph at promoters in the absence of H3K4me3, further emphasising the importance of considering the context of other regulatory markers when analysing phosphorylation enrichments across cell lines and species. We highly recommend that research is expanded to consider the role of mitotically-enriched histone phosphorylations in gene regulation.

While Gene Set Enrichment Analysis did show some significant enrichment of H3S10ph and H3S28ph in functional pathways such as protein metabolism and negative cell cycle regulation, whether this reflects a real functional association of histone phosphorylation with certain

molecular pathways remains unclear. Instead, we consider that H3S10ph and H3S28ph, and indeed H3T3ph, may play broader, more structural regulatory roles during mitosis.

We also suggest that these abundant histone phosphorylations could perform a bookmarking role in mitosis. “Bookmarking” theory has been hypothesised for many years (Hsiung et al., 2016; Kang et al., 2020a; Liang et al., 2015; Oomen et al., 2019a; Pelham-Webb et al., 2021; Valls et al., 2005), in which mitotically-enriched markers remain at specific important genomic loci in order to regulate the timing and/or activation of their transcription upon mitotic exit. We hypothesise that H3S10ph and H3S28ph may enrich at our isolated subsets of promoters in order to bookmark these genes which are then active or bivalent following exit from mitosis. Similarly, H3T3ph might bookmark specifically promoters which are absent of H3K4me3; the transcriptional activity of these genes in interphase would need to be further analysed. We recommend numerous further studies and validation experiments to test these hypotheses, discussed below.

Our findings strongly suggest that H3S10ph is not uniformly enriched across the genome, contrary to what has often been assumed in previous H3S10ph studies. This work therefore improves our understanding not only of the differential enrichment patterns of H3S10ph and H3S28ph during mitosis, but of how histone phosphorylation may well play important roles in mitotic chromatin and transcription regulation. Our research particularly contributes to biomedical research fields concerning dividing cells. Cancer research and developmental research investigating mechanisms of disease, for example, are progressed by improving our understanding of the base molecular mechanisms driving healthy versus dysregulated cell division. Our contribution highlights possible regulatory roles of H3S10ph and H3S28ph in correct gene expression specifically during mitosis; previous studies have only identified similar links between histone phosphorylation and transcription regulation in interphase cells. We therefore present here an exciting new avenue of research into the regulation of chromatin during mitosis and histone phosphorylation in contribution particularly to the cancer and developmental biology research fields.

6.2 Limitations and Future perspectives

6.2.1 Expanding analyses with different cell lines

The body of literature surrounding H3S10ph and H3S28ph often reports contrasting evidence depending on the cell line and organism used in each study. This is exemplified by evidence regarding H3S10ph involvement in mitotic chromosome condensation. While studies in *Tetrahymena* and budding yeast found H3S10ph depletion (via S10A mutation or Aurora B inhibition respectively) caused disrupted chromosome condensation, other *in vitro* studies found no evidence of H3S10A mutants impacting condensation, and no evidence of a role in condensation has been seen in human cell lines (e.g. Wei *et al.*, 1999; Hirota *et al,.*, 2005; Ditchfield *et al.*, 2003).

We must therefore caveat that our findings in mouse embryonic stem cells may well not be conserved across other eukaryotic cell types. In particular we note our evidence of H3S10ph and H3S28ph enrichment at promoters which are in a bivalent state requires careful interpretation. Embryonic stems cells are characterised by their ability to differentiate down many pathways into numerous tissue and cell types (for review see Zakrzewski *et al.*, 2019). As such, they are likely to possess higher levels of bivalency than differentiated cell types, in order to be poised to activate or repress many genes depending on the differentiation route taken. It is possible that the enrichment we found at bivalent promoters may be a biproduct of this overall enrichment of bivalent promoters in our chosen cell type. Therefore, we encourage future studies to perform MINUTE-ChIP-seq in other, perhaps differentiated cell types, to test whether stem cell bivalency affects the promoter enrichments we have observed. It would also be interesting to replicate our experiments and subsequent analyses in human cell lines, such as HeLa. On the other hand, we have shown that histone phosphorylations are enriched at subsets of promoters in two different mammalian cell lines, through our analysis of both H3S10ph/H3S28ph in mouse ESCs (Chapter 4) and H3T3ph in human HeLa cells (Chapter 5).

6.2.2 Enriching functional analyses with other genome regulatory regions

Our analysis of sequencing both from MINUTE-ChIP (Chapter 4) and paired-end traditional ChIP (Chapter 5) focussed primarily on promoter regions (TSS coordinates +/- 1 kb). Our MINUTE-ChIP data in particular provided an exciting opportunity to directly and quantitatively compare Mitotic vs G1 and G2 promoter enrichments, contributing interesting evidence towards previous research where histone phosphorylation has mainly only been seen at promoters in interphase. Following our findings that H3S10ph and H3S28ph both significantly varied between promoters depending on chromatin states, and that this was mitosis-specific, we chose to continue our focus on the promoter regions.

However, there may well be other genomic regions that could also show significant enrichments in mitotic H3S10ph/H3S28ph, including other regulatory regions, which we have not assessed in this project. For example, enhancer regions provide an interesting avenue of future research. Our co-localisation networks suggest a significant correlation between H3S28ph and H3K4me1, a characteristic marker of enhancers, across the cell cycle. This highlights a limitation of our study that these co-localisation networks only assess promoter regions; in order to interpret whether H3S28ph might interact with enhancer regions, co-localisation analysis would need to be expanded to enhancer regions**.**

Given the centromeric localisation of H3T3ph, we recommend that future analysis is performed to assess the chromosomal distribution of the H3K4me3-negative promoters we have isolated here. It would be interesting to determine whether these promoters found to have H3T3ph enrichment are in the regions farther from the centromeres, or whether they cluster at all. Moreover, our analysis focussed on protein-coding TSSs to avoid bias in enrichment analysis from genes with multiple TSSs. However, the centromeric and peri-centromeric regions might conceivably contain more non-protein coding genes and TSSs; analysis of these genes might generate additional insight into H3T3ph enrichments.

6.2.3 Elucidating the contribution of mitotic versus interphase phosphorylation in gene regulation

We have extensively discussed how previous studies have predominantly interpreted H3S10ph and H3S28ph enrichment at promoters to be an interphase-phenomenon (e.g. (e.g. DeManno *et al.*, 1999; Gehani *et al.*, 2010; Komar & Juszczynski, 2020; Lau & Cheung, 2011; Mahadevan *et al.*, 1991). The majority of these studies used asynchronous cell populations. Contrastingly, our analysis found H3S10ph and H3S28ph promoter enrichment was only seen in mitotic samples, and not in G1 or G2. We suggest a possibility that the lower-level enrichments seen at specific loci in asynchronous samples of previous studies might not be representing interphase cells, but instead be a product of contaminating signal from mitotic cells. Asynchronous populations contain some cells in mitosis, and we suggest that the abundance of histone phosphorylation events in mitotic cells might produce the low-level signal seen in asynchronous ChIP-seq studies. However, we do also caveat that our study did not involve any stimulation of interphase cell samples; which was required in the majority of previous studies to produce phosphorylation enrichments at promoters. Future studies could be performed introducing stimuli, such as growth or stress stimuli, to G1, G2 and Mitotic cell samples and compare the resulting MINUTE-ChIP sequencing to this non-stimulated dataset.

Our analysis contributes interesting evidence towards specific histone phosphorylation distributions and co-localisations in different cell cycle phases. In particular we draw attention to Figure 4.16; this demonstrates that the chromatin state regions where higher H3S10ph and H3S28ph are seen in mitosis are often very different from the states with highest H3S10ph/H3S28ph in G1 or G2 cells. In future, experiments could be performed to generate cell populations in clearly defined cell cycle stages; for example, inhibitor molecules such as CDK1 inhibitors (Vassilev, 2006) could be used to block and isolate cells in G2 phase, or cells could be released from mitosis and timed to collect cells at specific time points after mitotic exit. By performing ChIP or MINUTE-ChIP with these cell samples in specific cell cycle phases, as well as an asynchronous cell population, future studies could identify whether any promoter phosphorylation enrichment seen in asynchronous cells is also seen in purely interphase, G1 or G2 cells. This would test the hypothesis that previous phosphorylation enrichments seen in asynchronous cells could be due to mitotic contamination. We do caveat however that such

experiments, by introducing drug treatments, impact and potentially disrupt the "normal" biology of cells, as was discussed in Chapter 1. The choice between blocking cells in cell cycle phases through drug treatment, versus less invasive cell sorting as used by Elsasser *et al.,* for this MINUTE-ChIP dataset, requires careful consideration. Drug methods might be used to meet specific experiment objectives; for example, CDK1i would allow testing of whether high CDK1 is necessary for the mitotic promoter histone phosphorylation enrichments we have observed here.

6.2.4 Investigating non-genic roles of H3S10ph and H3S28ph in mitosis

Having assessed protein-coding promoter regions and found significant enrichment patterns in H3S10ph, H3S28ph and H3T3ph, we subsequently focussed our analysis for this project on genic regions. Our aim was to elucidate the possible roles of these histone phosphorylations in gene regulation specifically in mitotic cells.

However, the majority of literature surrounding these histone phosphorylations focusses instead on regulation of chromosome functions not directly related to gene expression (e.g. Neurohr *et al.*, 2011; Gehani *et al.*, 2010; Hirota *et al.*, 2005; Wei *et al.*, 1999). H3T3ph demonstrates clear examples of this, with the phosphorylation shown to bind and recruit the chromosomal passenger complex to the centromeric regions, enabling correct chromosome segregation (Wang *et al.*, 2010; Kelly *et al.*, 2010; Yamagishi *et al.*, 2010). When functional roles have been suggested for mitotic H3S10ph, they again have predominantly centred on structural: for example, numerous studies have hypothesised a role for H3S10ph in displacing heterochromatin protein 1 (HP1) from its H3K9me3 binding site, in a methyl-phos switch mechanism (e.g. Hirota *et al.*, 2005; Fischle *et al.*, 2005)**.** Extensive studies across a range of vertebrates have reported HP1 to be involved in promoter regulation, with HP1 enrichment or depletion at specific promoters reported to be dependent on contexts including the regulatory state of that chromatin region (for review see (Schoelz and Riddle, 2022). Combining this literature with our findings, we suggest that a methyl-phos switch may be important at these specific promoters depending on the chromatin regulatory state. Validation studies would be very valuable here to assess the co-localisation of H3S10ph and H3K9me3 with HP1 at this subset of promoters in mitotic cells.

Although we observed promoter enrichment of mitotic histone phosphorylation at protein-coding genes, we did not isolate any individual clear functional pathway or mechanism through which these histone phosphorylations might act. Therefore, we suggest that the histone phosphorylation promoter enrichments we saw in mitosis might instead be involved in structural regulation.

This presents an exciting new avenue of research into functional roles of histone phosphorylations in mitosis. For example, one hypothesis we made is that H3S10ph and H3S28ph might act to "bookmark" specific promoters which are then in an active or bivalent state in interphase. Validation experiments could be designed to test this. In particular, we encourage nascent RNA-seq experiments, such as the pulse-labelling RNA-seq previously performed by Palazola *et al.,* (Palozola *et al.,* 2017, 2019; Kang *et al.,* 2020)**.** Nascent RNA-seq in tandem with MINUTE-ChIP-seq would allow researchers to measure and track the transcription activity of genes during mitosis, and exiting mitosis. It would be interesting to then analyse the timing of transcriptional activation for those genes where H3S10ph or H3S28ph are seen at promoters during mitosis: are these genes activated early upon mitotic exit for example? This experiment would also allow us to determine whether genes with H3S10ph/H3S28ph promoter enrichment are active during mitosis, suggesting a direct role of H3S10ph/H3S28ph in regulating transcription during mitosis, or whether H3S10ph/H3S28ph may instead play a bookmarking role. Knockdown studies could also potentially be designed to assess whether reduction/silencing of H3S10ph/H3S28ph impacted transcription activation and/or timing; however, we stress that histone mutation and kinase knockdown/inhibition studies are notoriously difficult to interpret due to the complexity of their downstream substrates and pathways.

An alternative, perhaps simpler validation experiment would be to analyse H3S10ph/H3S28ph interaction with known structural regulatory markers. While we were able to co-analyse multiple genic regulatory histone modifications and proteins, we did not have MINUTE-ChIP-seq data for other potentially interesting markers. For example, sequencing data for cohesin protein, or heterochromatin protein 1 (HP1), could provide valuable insights into whether H3S10ph or H3S28ph might contribute to functions of chromosome regulation during mitosis that are not directly related to gene expression. We do caveat that limitations in terms of data availability in our mESC cell line might currently restrict our ability to perform such analyses; for example, to

our knowledge there is limited Hi-C data or analysis of structural regulatory proteins in the Rw4 murine ESC cell line (see Chapter 2) that are specifically mitotic cells.

Similarly, our MINUTE-ChIP-seq phosphorylation data could be co-analysed/integrated with structural genomic data; in particular, we suggest integration of Hi-C data. Hi-C experiments use ligation to assess 3D long-range interactions of chromatin. If Hi-C data were available in our mESC cell line, this could be integrated to assess whether H3S10ph or H3S28ph enrichments correlate with any long-range/3D chromatin interactions that would not be detectable through our linear genic analysis in this project. Many exciting research questions that have not been included in this project could be analysed in future through this type of integrative analysis. For example, we noted when looking at all TSSs (+/- 1 kb), H3S10ph was lower at chromatin state 10 promoters (see Figure 4.9). This state is characterised by enrichment of laminB1 see Figure 4.5) and is likely to represent lamina-associated domains (LADs) of repressive heterochromatin formed at the nuclear periphery in interphase through interaction of chromatin with the nuclear lamina. Prior studies have reported that H3K9me2/3 is involved in formation of LADs, possibly through binding of HP1 (e.g. Poleshko *et al.*, 2017; Kind *et al.*, 2013). In mitosis, it has been hypothesised that H3S10ph may regulate LADs by displacing HP1 from H3K9me2/3, preventing chromatin from binding to lamina (e.g. Poleshko *et al.*, 2019). These studies notably contrast with our findings here, where we observed lower H3S10ph at state 10 (lamina-associated) promoter regions, which would argue against a methyl-phos switch operating at these locations. Further studies focussing on the co-localisation (or lack) of H3S10ph and laminB1 in this mESC cell line could be performed to further investigate this, contributing to the methyl-phos switch debate and knowledge of LAD regulation through mitosis.

Finally, we hypothesise that H3S10ph and H3S28ph might be involved in a methyl-phos switch at specific promoters in active and bivalent states, given significant co-localisation with their adjacent methylations H3K9me3 and H3K27me3 was seen. We suggest future research consider that methyl-phos switches rather than being globally active/present or inactive/absent, might instead be important and/or strongly "flipped" at certain important genes in certain regulatory states. Again, RNA-seq validation experiments could be performed alongside further analysis of these adjacent methylation marks, to assess whether co-localisation is associated with changes in transcription levels or activation timing during mitotic exit.

6.3 Conclusion

Many histone phosphorylations are found at much higher levels in mitosis compared to interphase, a phenomenon observed across eukaryotes. However, the functional roles, if any, of this mitotic abundance are often unclear. In this project, we aimed to use advancements in chromatin immunoprecipitation techniques combined with integrative analysis to accurately assess the precise distributions of abundant mitotic histone phosphorylations H3S10ph, H3S28ph and H3T3ph. Crucially, combining quantitative histone phosphorylation sequencing with chromatin regulatory state data allowed us to quantitatively compare histone phosphorylation levels in mitosis vs G1 and G2, and for the first time assess histone phosphorylation in the context of chromatin regulatory landscape. We reveal significant enrichment of mitotic H3S10ph and H3S28ph specifically at promoters which are active and bivalent in interphase in mESCs, and show enrichment of H3T3ph at promoters in the absence of H3K4me3 in mitotic human HeLa cells. Previous studies have only interpreted promoter enrichment of H3S10ph and H3S28ph in asynchronous cells as occurring solely in interphase; from our analysis, we encourage future research to expand to consider a role of mitotic histone phosphorylation in gene regulation.

While we do present some significant functional enrichments associated with these mitotic H3S10ph and H3S28ph promoter peaks, we do not find any clear molecular mechanisms for mitotic gene regulation. Instead, we suggest a few functional roles that H3S10ph and H3S28ph might serve in mitotic chromosome regulation that are not directly regulating gene expression. Namely:

- H3S10ph and H3S28ph might act as mitotic “bookmarks”, enriching at specific promoters which are then active and bivalent in interphase. Nascent RNA-seq would provide crucial validation experiments to assess the timing and activation of transcription during mitosis and following mitotic exit, at these promoters of interest.
- H3S10ph and H3S28ph might be involved in methyl-phos switches in mitosis at promoters in specific chromatin regulatory states. This could be further investigated for example through co-localisation studies of regulatory proteins that H3S10ph or H3S28ph might displace (e.g. HP1, polycomb group proteins respectively)
- H3S10ph and H3S28ph might be involved in more 3D regulatory interactions that are not captured through our analysis here. Integration of long-range chromatin interaction

methods such as Hi-C would allow analysis of whether these H3S10ph and H3S28ph promoter enrichments correlate with any long-range interactions in mitosis.

We conclude by emphasising the importance of taking into account the wider chromatin regulatory context when interpreting the distribution of abundant mitotic histone phosphorylations.

Appendix

Table A 1: N-terminal histone H3 peptides.

Column 1(left) shows the histone residue, amino acid numbers of the peptide, and the modifications of the given peptide. Abbreviations and symbols: -Ahx = aminohexanoic acid; Ac- = N-terminal acetylation; Kac = acetyl-Lysine; Sph = phospho-Serine; Kme1 = monomethyl-Lysine; Tph = phospho-Threonine; Kme2 = dimethyl-Lysine; Rme1 = monomethyl-Argenine; Kme3 = trimethyl-Lysine; Rme2a = asym dimethyl-Argenine.

Peptide	**Sequence**	**Source**
H3 1-21	ARTKQTARKSTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: A2**
H3 1-21 R2me	ARme1TKQTARKSTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: A3**
H3 1-21 R2me2a	ARme2aTKQTARKSTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: A4**
H3 1-21 R2me2aT3phos	ARme2aTphKQTARKSTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: A5**
H3 1-21 R2me2aT3phosK4me3	ARme2aTphKme3QTARKSTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: A6**
H3 1-21 R2me2aK4me3	ARme2aTKme3QTARKSTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: A7**
H3 1-21 T3phos	ARTphKQTARKSTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: A8**

H3 1-21 T3phosK4me3	ARTphKme3QTARKSTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: A9**
H3 1-21 K4me	ARTKme1QTARKSTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: A10**
H3 1-21 K4me2	ARTKme2QTARKSTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: A11**
H3 1-21 K4me3	ARTKme3QTARKSTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: A12**
H3 1-21 K4me3K9ac	ARTKme3QTARKacSTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: B1**
H3 1-21 K4me3K9me3	ARTKme3QTARKme3STGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: B2**
H3 1-21 K9ac	ARTKQTARKacSTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: B3**
H3 1-21 K9acS10phos	ARTKQTARKacSphTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: B4**
H3 1-21 K9acT11phos	ARTKQTARKacSTphGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: B5**
H3 1-21 K9acS10phosT11phos	ARTKQTARKacSphTphGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: B6**
H3 1-21 K9me	ARTKQTARKme1STGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: B7**

H3 1-21 K9me2	ARTKQTARKme2STGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: B8**
H3 1-21 K9me3	ARTKQTARKme3STGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: B9**
H3 1-21 K9me3T11phos	ARTKQTARKme3STphGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: B10**
H3 1-21 K9me3S10phos	ARTKQTARKme3SphTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: B11**
H3 1-21 K9me3S10phosT11phos	ARTKQTARKme3SphTphGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: B12**
H3 1-21 S10phos	ARTKQTARKSphTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: C1**
H3 1-21 T11phos	ARTKQTARKSTphGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: C2**
H3 1-21 S10phosT11phos	ARTKQTARKSphTphGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: C3**
H3 4-24 S10phosT11phosK14ac	Ac-KQTARKSphTphGG/APRKQLATKA-Ahx-biotin	**Alta Histone Peptide Library Set 3: C4**
H3 4-24 S10phosT11phosK14me3	Ac-KQTARKSphTphGGKme3APRKQLATKA-Ahx-biotin	**Alta Histone Peptide Library Set 3: C5**
H3 4-24 T11phosK14ac	Ac-KQTARKSTphGG/APRKQLATKA-Ahx-biotin	**Alta Histone Peptide Library Set 3: C6**

H3 4-24 T11phosK14me3	Ac-KQTARKSTphGGKme3APRKQLATKA-Ahx-biotin	**Alta Histone Peptide Library Set 3: C7**
H3 4-24 K9acK14ac	Ac-KQTARKacSTGGKacAPRKQLATKA-Ahx-biotin	**Alta Histone Peptide Library Set 3: C8**
H3 8-28 K14ac	Ac-RKSTGGKacAPRKQLATKAARKS-Ahx-biotin	**Alta Histone Peptide Library Set 3: C9**
H3 8-28 K14acR17me2a	ARTKQTARKSTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: C10**
H3 8-28 K14me	ARme1TKQTARKSTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: C11**
H3 8-28 K14me2	ARme2aTKQTARKSTGGKAPRKQLA-Ahx-biotin	**Alta Histone Peptide Library Set 3: C12**
H3 8-28 K14me3	Ac-RKSTGGKacAPRme2aKQLATKAARKS-Ahx-biotin	**Alta Histone Peptide Library Set 3: D1**
H3 8-28 K14me3R17me2a	Ac-RKSTGGKme1APRKQLATKAARKS-Ahx-biotin	**Alta Histone Peptide Library Set 3: D2**
H3 8-28 R17me	Ac-RKSTGGKme2APRKQLATKAARKS-Ahx-biotin	**Alta Histone Peptide Library Set 3: D3**
H3 8-28 R17me2a	Ac-RKSTGGKme3APRKQLATKAARKS-Ahx-biotin	**Alta Histone Peptide Library Set 3: D4**
H3 8-28 K14acK18ac	Ac-RKSTGGKme3APRme2aKQLATKAARKS-Ahx-biotin	**Alta Histone Peptide Library Set 3: D5**

H3 12-32 R17me2aK18ac	Ac-RKSTGGKAPRme1KQLATKAARKS-Ahx-biotin	**Alta Histone Peptide Library Set 3: D6**
H3 12-32 K18ac	Ac-RKSTGGKAPRme2aKQLATKAARKS-Ahx-biotin	**Alta Histone Peptide Library Set 3: D7**
H3 12-32 K18acK23me3	Ac-RKSTGGKacAPRKacQLATKAARKS-Ahx-biotin	**Alta Histone Peptide Library Set 3: D8**
H3 12-32 K18acK23ac	Ac-GGKAPRme2aKacQLATKAARKSAPAT-Ahx-biotin	**Alta Histone Peptide Library Set 3: D9**
H3 18-38 K23ac	Ac-GGKAPRKacQLATKAARKSAPAT-Ahx-biotin	**Alta Histone Peptide Library Set 3: D10**
H3 18-38 K23acR26me2a	Ac-GGKAPRKacQLATKme3AARKSAPAT-Ahx-biotin	**Alta Histone Peptide Library Set 3: D11**
H3 18-38 K23acR26me2aK27ac	Ac-GGKAPRKacQLATKacAARKSAPAT-Ahx-biotin	**Alta Histone Peptide Library Set 3: D12**
H3 18-38 K23me	Ac-KQLATKme1AARKSAPATGGVKKP-Ahx-biotin	**Alta Histone Peptide Library Set 3: E1**
H3 18-38 K23me2	Ac-KQLATKme2AARKSAPATGGVKKP-Ahx-biotin	**Alta Histone Peptide Library Set 3: E2**
H3 18-38 K23me3	Ac-KQLATKme3AARKSAPATGGVKKP-Ahx-biotin	**Alta Histone Peptide Library Set 3: E3**
H3 18-38 K23me3R26me2a	Ac-KQLATKme3AARme2aKSAPATGGVKKP-Ahx-biotin	**Alta Histone Peptide Library Set 3: E4**

H3 18-38 K23me3R26me2aK2ac	Ac-KQLATKme3AARme2aKacSAPATGGVKKP-Ahx-biotin	**Alta Histone Peptide Library Set 3: E5**
H3 20-40	Ac-LATKAARKSAPATGGVKKPHR-Ahx-biotin	**Alta Histone Peptide Library Set 3: E6**
H3 20-40 R26me	Ac-LATKAARme1KSAPATGGVKKPHR-Ahx-biotin	**Alta Histone Peptide Library Set 3: E7**
H3 20-40 R26me2a	Ac-LATKAARme2aKSAPATGGVKKPHR-Ahx-biotin	**Alta Histone Peptide Library Set 3: E8**
H3 20-40 R26me2aK27ac	Ac-LATKAARme2aKacSAPATGGVKKPHR-Ahx-biotin	**Alta Histone Peptide Library Set 3: E9**
H3 20-40 R26me2aK27acS28phos	Ac-LATKAARme2aKacSphAPATGGVKKPHR-Ahx-biotin	**Alta Histone Peptide Library Set 3: E10**
H3 20-40 R26me2aS28phos	Ac-LATKAARme2aKSphAPATGGVKKPHR-Ahx-biotin	**Alta Histone Peptide Library Set 3: E11**
H3 20-40 K27ac	Ac-LATKAARKacSAPATGGVKKPHR-Ahx-biotin	**Alta Histone Peptide Library Set 3: E12**
H3 20-40 K27acS28phos	Ac-LATKAARKacSphAPATGGVKKPHR-Ahx-biotin	**Alta Histone Peptide Library Set 3: F1**
H3 20-40 K27me	Ac-LATKAARKme1SAPATGGVKKPHR-Ahx-biotin	**Alta Histone Peptide Library Set 3: F2**

H3 20-40 K27me2	Ac-LATKAARKme2SAPATGGVKKPHR-Ahx-biotin	**Alta Histone Peptide Library Set 3: F3**
H3 20-40 K27me3	Ac-LATKAARKme3SAPATGGVKKPHR-Ahx-biotin	**Alta Histone Peptide Library Set 3: F4**
H3 20-40 R26me2aK27me3	Ac-LATKAARme2aKme3SAPATGGVKKPHR-Ahx-biotin	**Alta Histone Peptide Library Set 3: F5**
H3 20-40 K27me3S28phos	Ac-LATKAARKme3SphAPATGGVKKPHR-Ahx-biotin	**Alta Histone Peptide Library Set 3: F6**
H3 20-40 R26me2aK27me3S28phos	Ac-LATKAARme2aKme3SphAPATGGVKKPHR-Ahx-biotin	**Alta Histone Peptide Library Set 3: F7**
H3 20-40 S28phos	Ac-LATKAARKSphAPATGGVKKPHR-Ahx-biotin	**Alta Histone Peptide Library Set 3: F8**

Table A 2: CUT+RUN library preparation Truseq Adapter sequences

Adapter	**Adapter Sequence**
1	ATCACG
2	CGATGT
3	TTAGGC
4	TGACCA

5	ACAGTG
6	GCCAAT
7	CAGATC
8	ACTTGA
9	GATCAG
10	TAGCTT
11	GGCTAC
12	CTTGTA
13	AGTCAA
14	AGTTCC
15	ATGTCA
16	CCGTCC
18	GTCCGC
19	GTGAAA
20	GTGGCC
21	GTTTCG
22	CGTACG
23	GAGTGG
25	ACTGAT

27	ATTCCT

Table A 3: MINUTE-ChIP T7 adapter sequences.

*Includes T7 promoter sequence, partial Illumina primer SBS3, Unique Molecular Identifer (UMI) and PCR barcode sequences. Used in MINUTE-ChIP, designed by Elsasser and Kumar, 2019, based on Mint-ChIP protocol from the Bernstein lab***.**

FORWARD	
AdRan_BC01_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NCT ACC AGG
AdRan_BC02_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NCA TGC TTA
AdRan_BC03_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NGC ACA TCT
AdRan_BC04_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NTG CTC GAC
AdRan_BC05_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NAG CAA TTC

AdRan_BC06_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NAG TTG CTT
AdRan_BC07_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NCC AGT TAG
AdRan_BC08_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NTT GAG CCT
AdRan_BC09_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NAC CAA CTG
AdRan_BC10_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NGG TCC AGA
AdRan_BC11_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NGT ATA ACA
AdRan_BC12_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NTT CGC TGA
AdRan_BC13_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NAG GGT ACT

AdRan_BC14_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NCG GAC TAT
AdRan_BC15_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NCT GCA CAA
AdRan_BC16_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NGA ATC GGT
AdRan_BC17_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NGT TGA GTC
AdRan_BC18_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NAA AGG GAC
AdRan_BC19_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NCA CAG GTT
AdRan_BC20_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NCC TAT GCA
AdRan_BC21_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NCT GTT GTG

AdRan_BC22_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NGA CCT ATC
AdRan_BC23_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NGA TAG TGC
AdRan_BC24_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NTA GCT GGA
AdRan_BC25_s	/5SpC3/GAA TTT AAT ACG ACT CAC TAT AGG GTA CAC GAC GCT CTT CCG ATC TNN NNN NTG ACA GTG
REVERSE	
AdRan_BC01_as	CCT GGT AGN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC02_as	TAA GCA TGN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC03_as	AGA TGT GCN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC

AdRan_BC04_as	GTC GAG CAN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC05_as	GAA TTG CTN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC06_as	AAG CAA CTN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC07_as	CTA ACT GGN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC08_as	AGG CTC AAN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC09_as	CAG TTG GTN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC10_as	TCT GGA CCN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC11_as	TGT TAT ACN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC

AdRan_BC12_as	TCA GCG AAN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC13_as	AGT ACC CTN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC14_as	ATA GTC CGN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC15_as	TTG TGC AGN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC16_as	ACC GAT TCN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC17_as	GAC TCA ACN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC18_as	AAA GGG ACN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC19_as	AAC CTG TGN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC20_as	TGC ATA GGN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC

AdRan_BC21_as	CAC AAC AGN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC22_as	GAT AGG TCN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC23_as	GCA CTA TCN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC24_as	TCC AGC TAN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC
AdRan_BC25_as	CAC TGT CAN NNN NNA GAT CGG AAG AGC GTC GTG TAC CCT ATA GTG AGT CGT ATT AAA TTC

Table A 4: MINUTE-ChIP-seq RNA 3' adapter and reverse transcription (RT) primer

RA3 adapter	*/5rApp/AGA TCG GAA GAG CAC ACG TCT /3SpC3/*
RT_primer	*AGA CGT GTG CTC TTC CGA TCT*

Table A 5: MINUTE-ChIP-seq PCR primers.

Designed in house by Elsasser et al., includes Illumina primer P7 and SBS12 sequences

PCR_forward	*AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTC TTCCGATCT*
PCR_BC1	*CAA GCA GAA GAC GGC ATA CGA GAT CCT GGT AGG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT*
PCR_BC2	*CAA GCA GAA GAC GGC ATA CGA GAT TAA GCA TGG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT*
PCR_BC3	*CAA GCA GAA GAC GGC ATA CGA GAT AGA TGT GCG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT*
PCR_BC4	*CAA GCA GAA GAC GGC ATA CGA GAT GTC GAG CAG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT*
PCR_BC5	*CAA GCA GAA GAC GGC ATA CGA GAT GAA TTG CTG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT*
PCR_BC6	*CAA GCA GAA GAC GGC ATA CGA GAT TCG CAC CTG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT*

PCR_BC7	*CAA GCA GAA GAC GGC ATA CGA GAT CTA ACT GGG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT*
PCR_BC8	*CAA GCA GAA GAC GGC ATA CGA GAT AGG CTC AAG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT*
PCR_BC9	*CAA GCA GAA GAC GGC ATA CGA GAT CAG TTG GTG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT*
PCR_BC10	*CAA GCA GAA GAC GGC ATA CGA GAT TCT GGA CCG TGA CTG GAG TTC AGA CGT GTG CTC TTC CGA TCT*

L1: Ladder

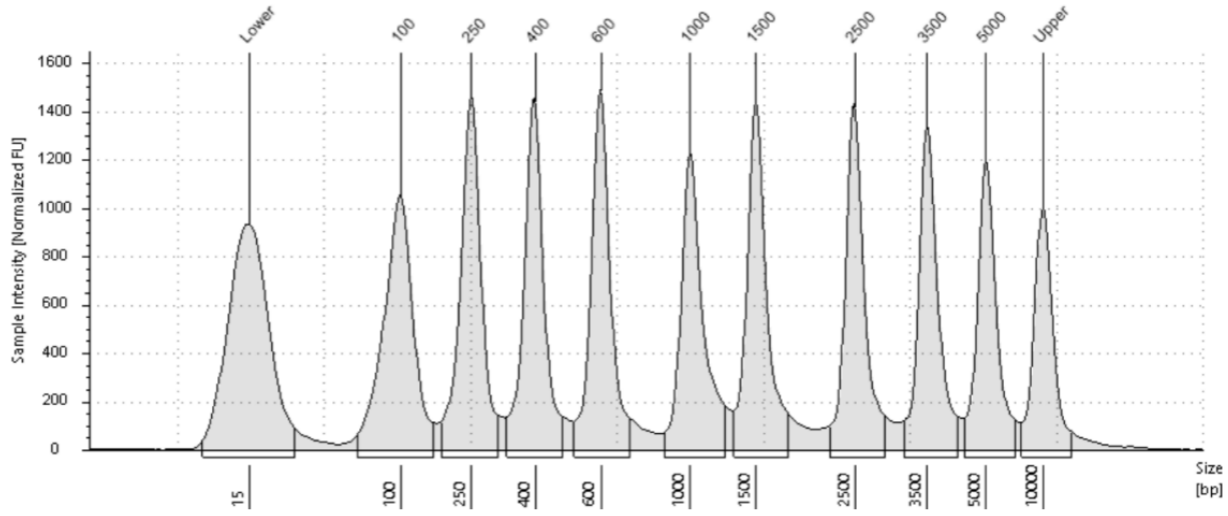
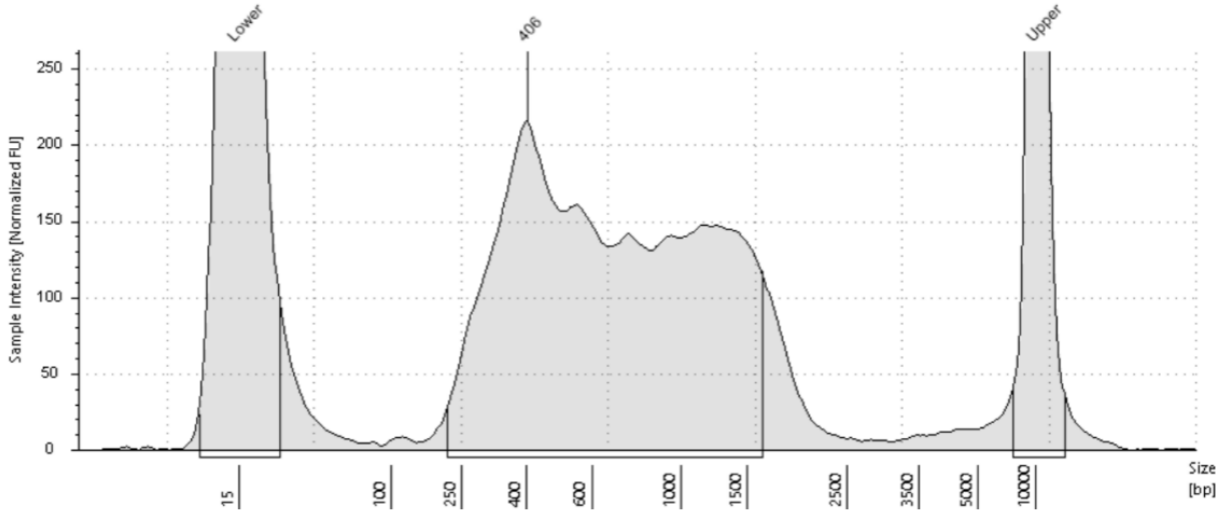
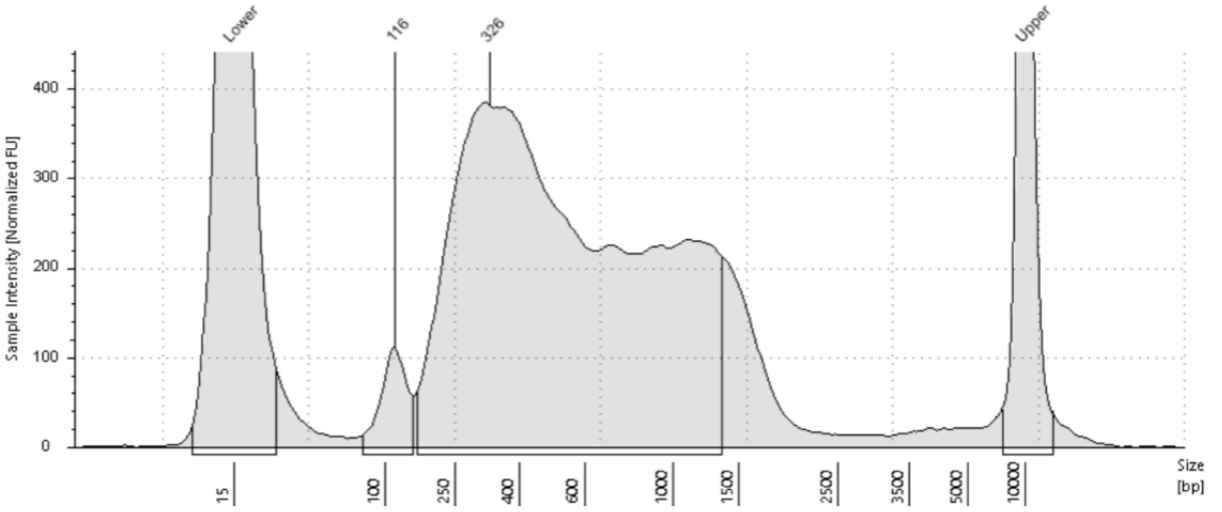


Figure A 1: Tapestation DNA ladder of known lengths (x axis, bp)

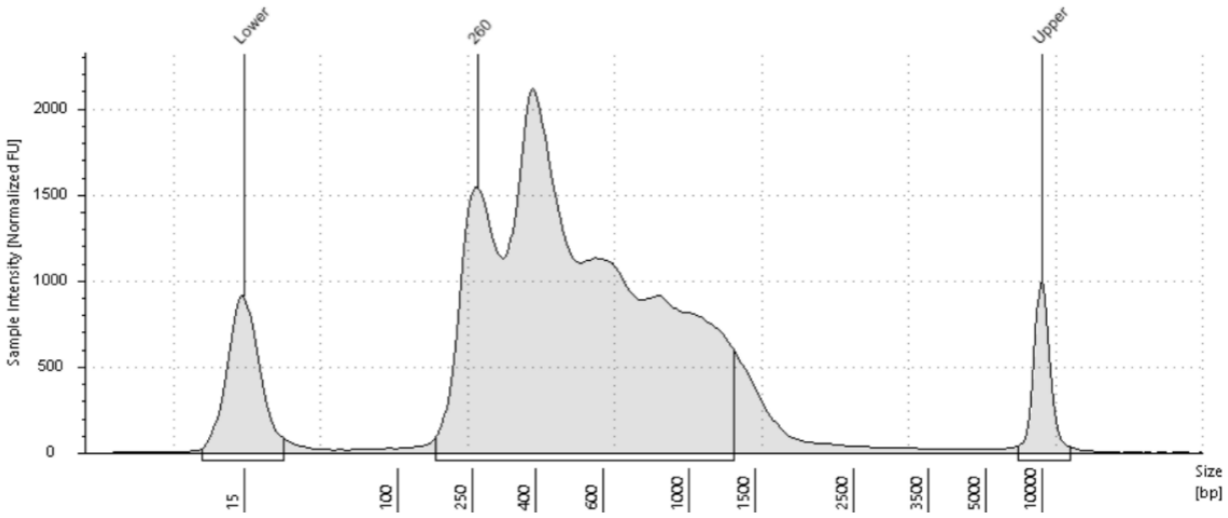
B1: JM HeLa IgG 1/4



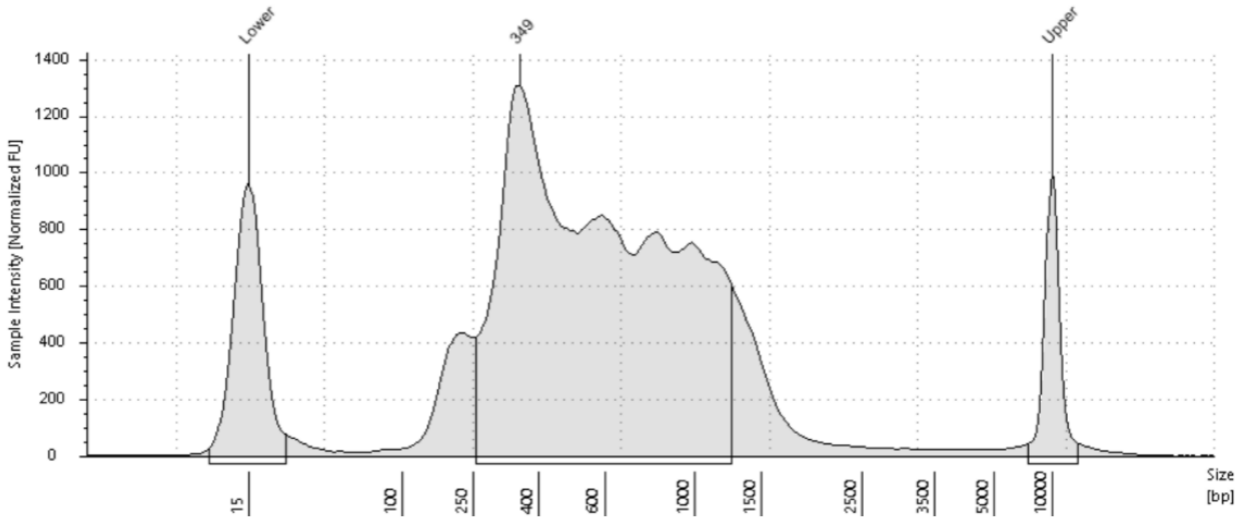
C1: JM HeLa K4me3 1/4



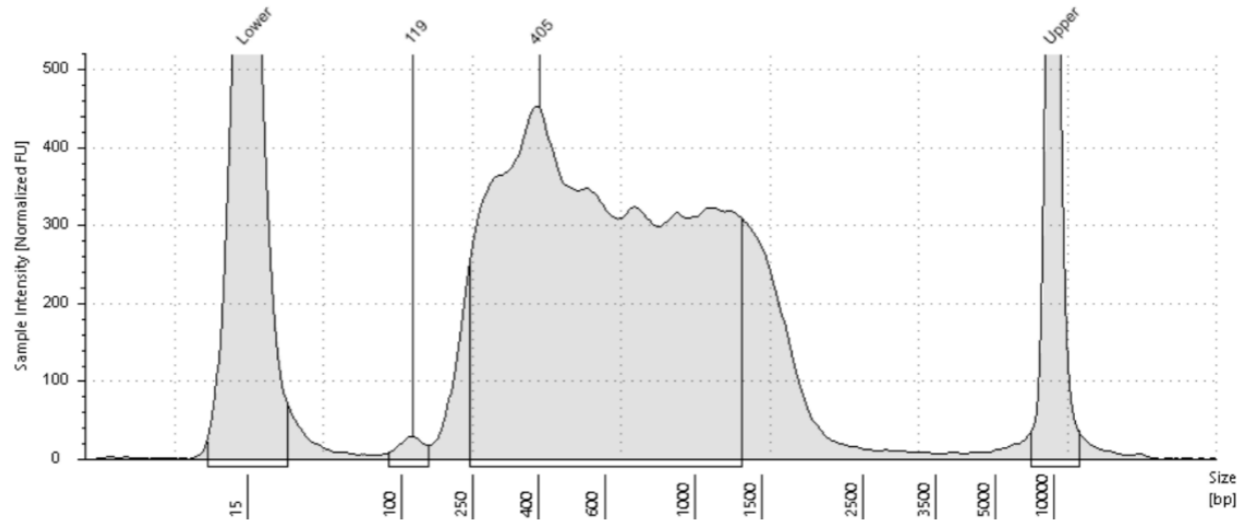
A2: PS CA46 K4me3 1/4



E1: JM CA46 K4me3 1/4



F1: PS HeLa IgG 1/4



H1: PS CA46 IgG 1/4

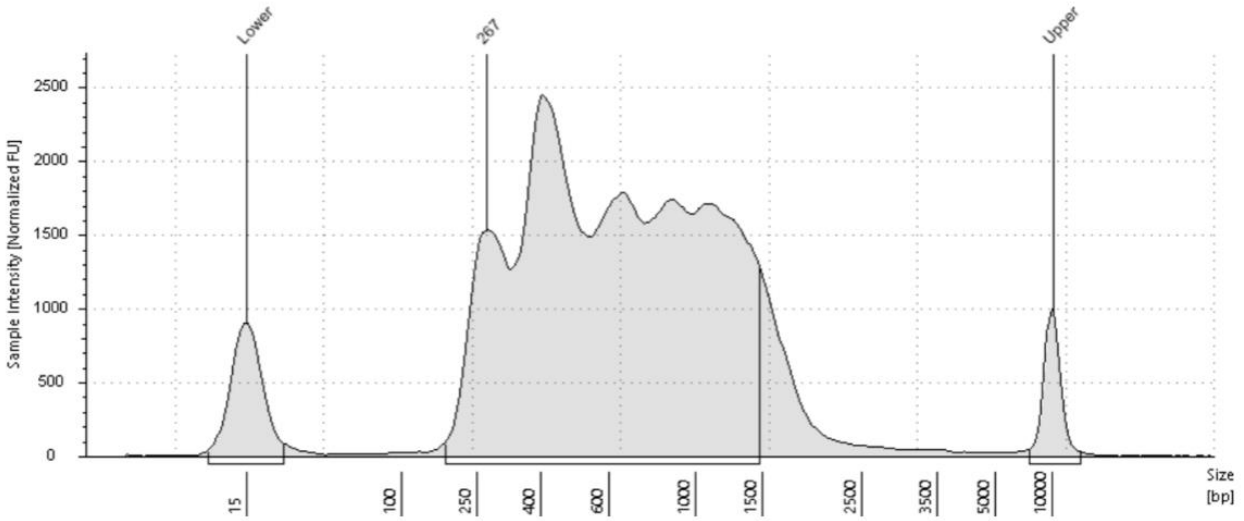


Figure A 2: Tapestation following CUT+RUN targeting IgG or H3K4me3, in HeLa S3 or CA46 cell lines, using Higgins (JM) or Russell (PS) laboratory reagents.

See title above each plot for sample info. ¼ indicated 1 in 4 dilution used for TapeStation samples.

.

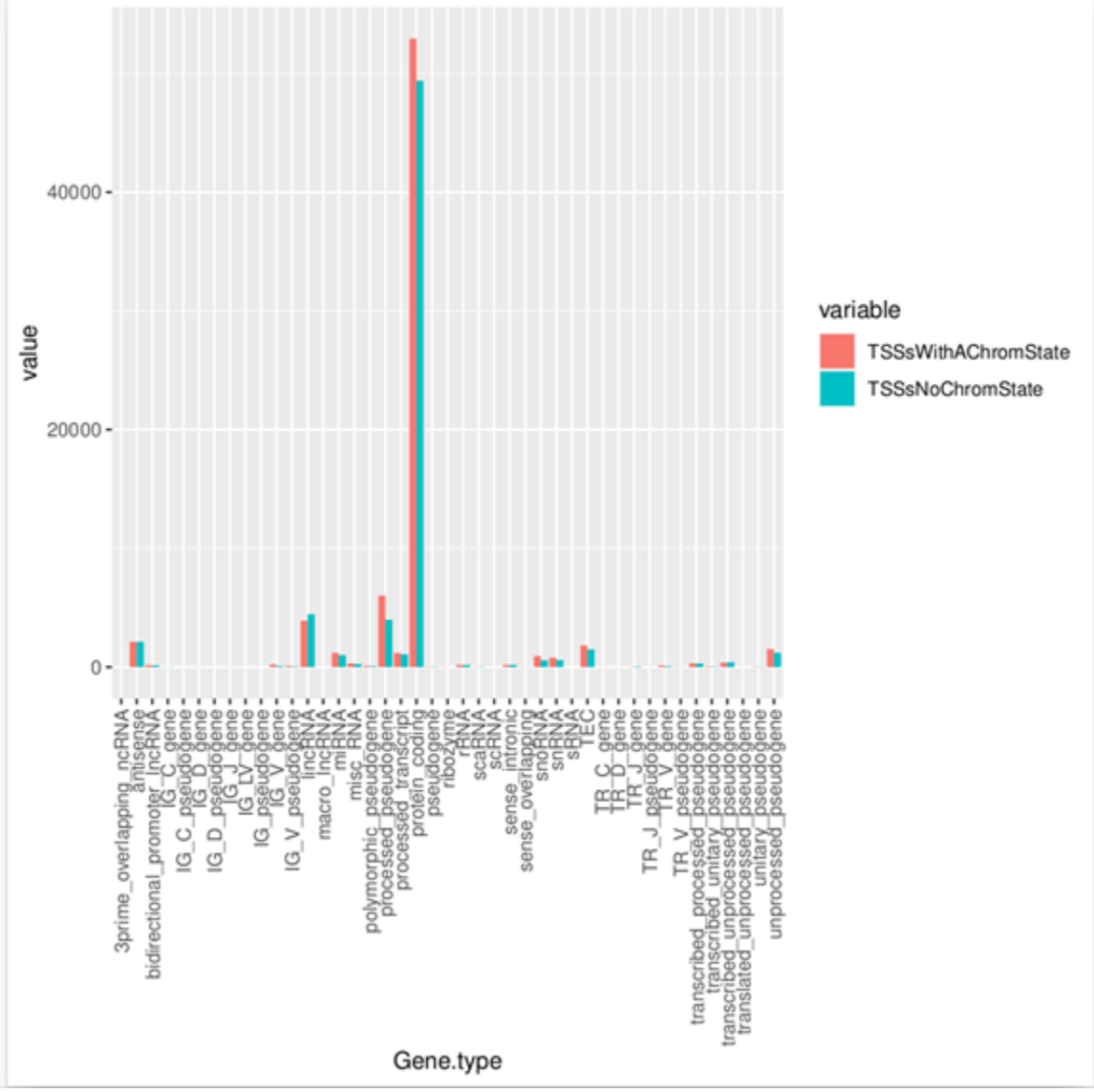
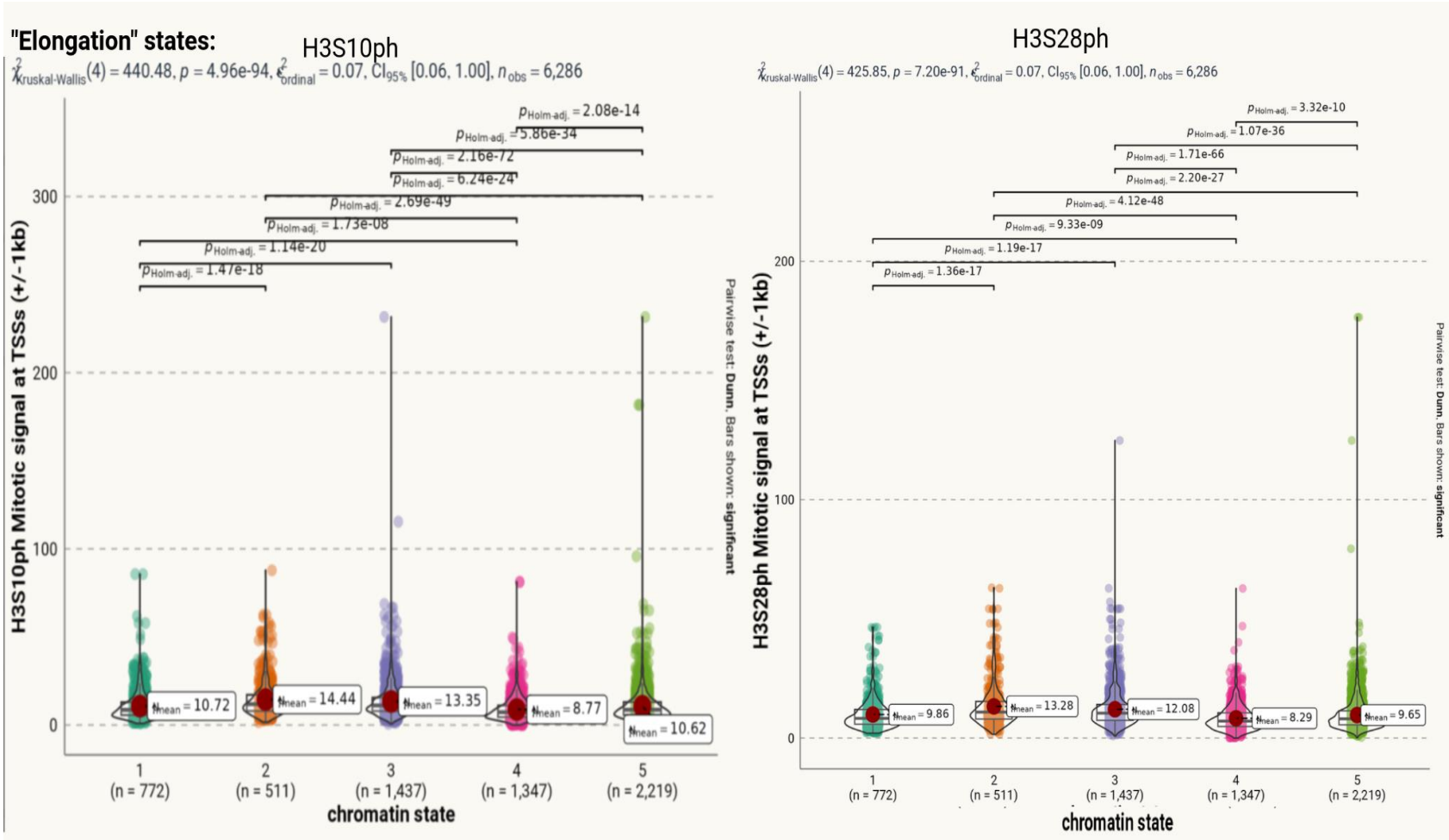
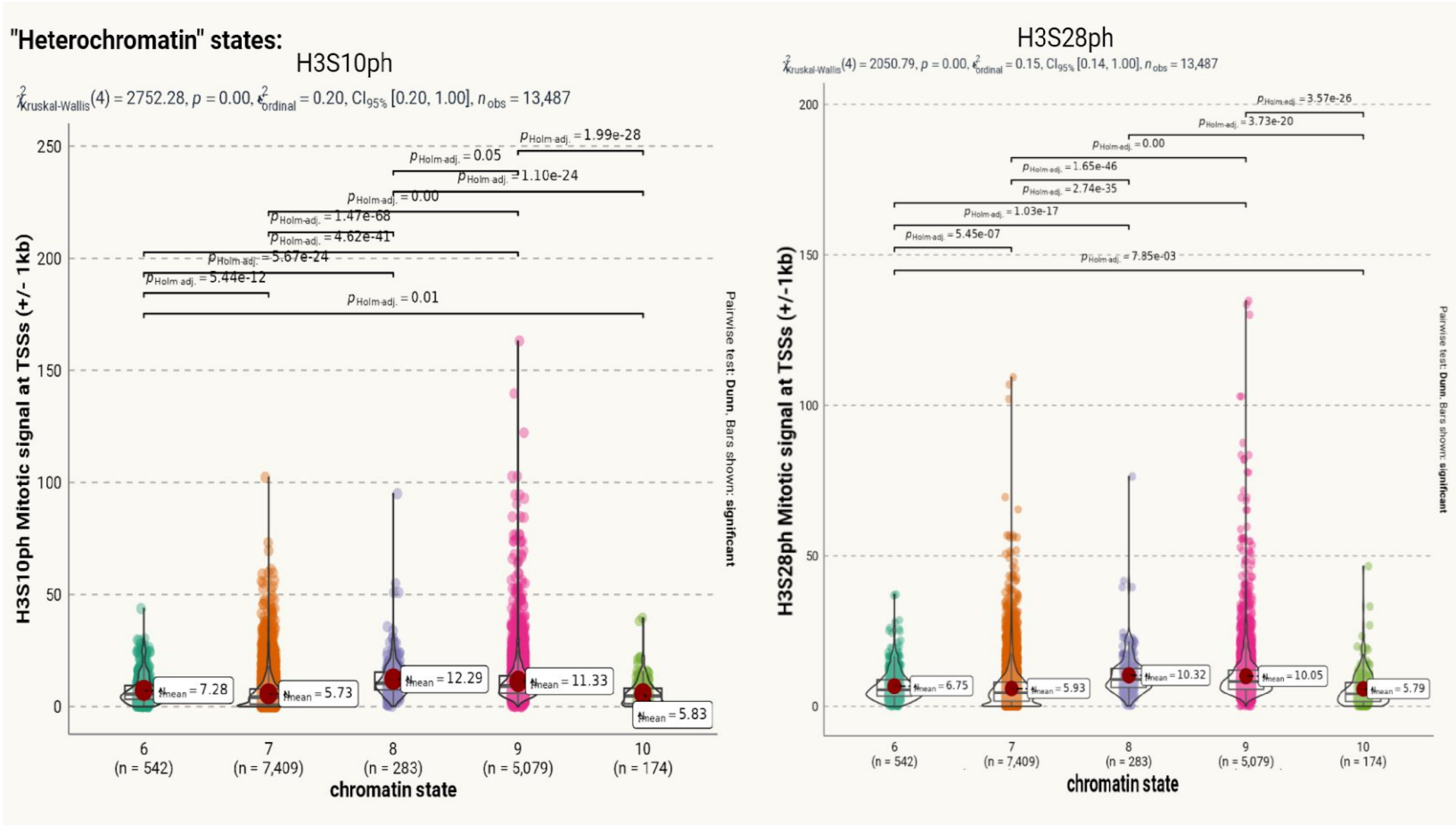
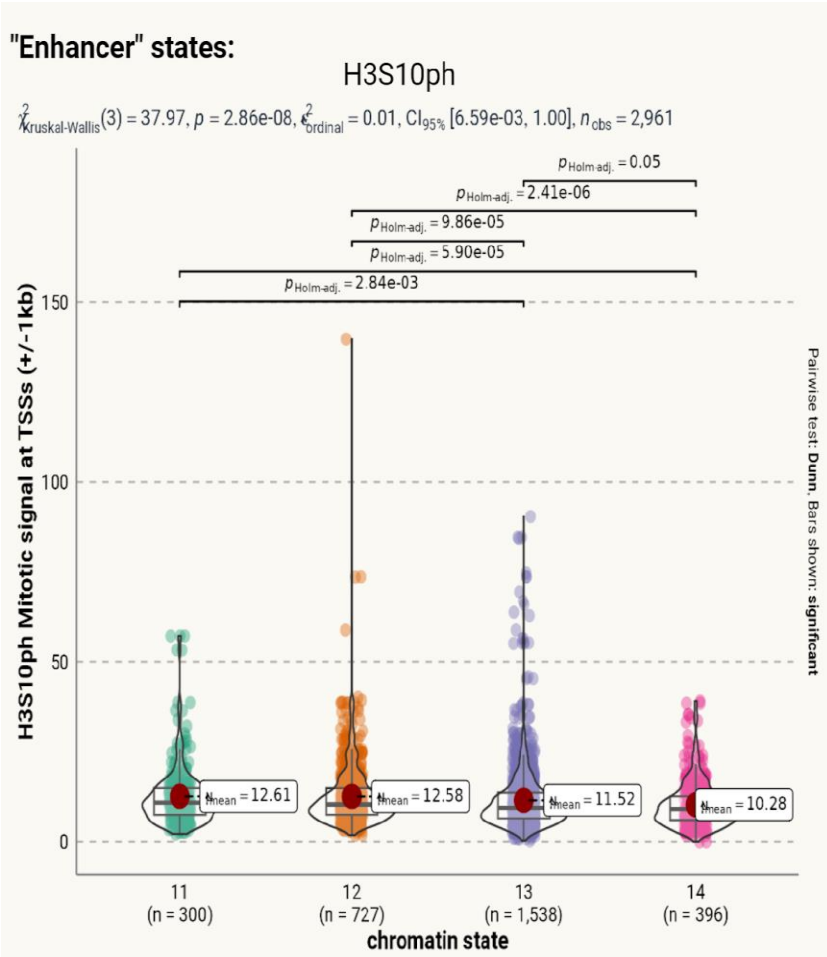


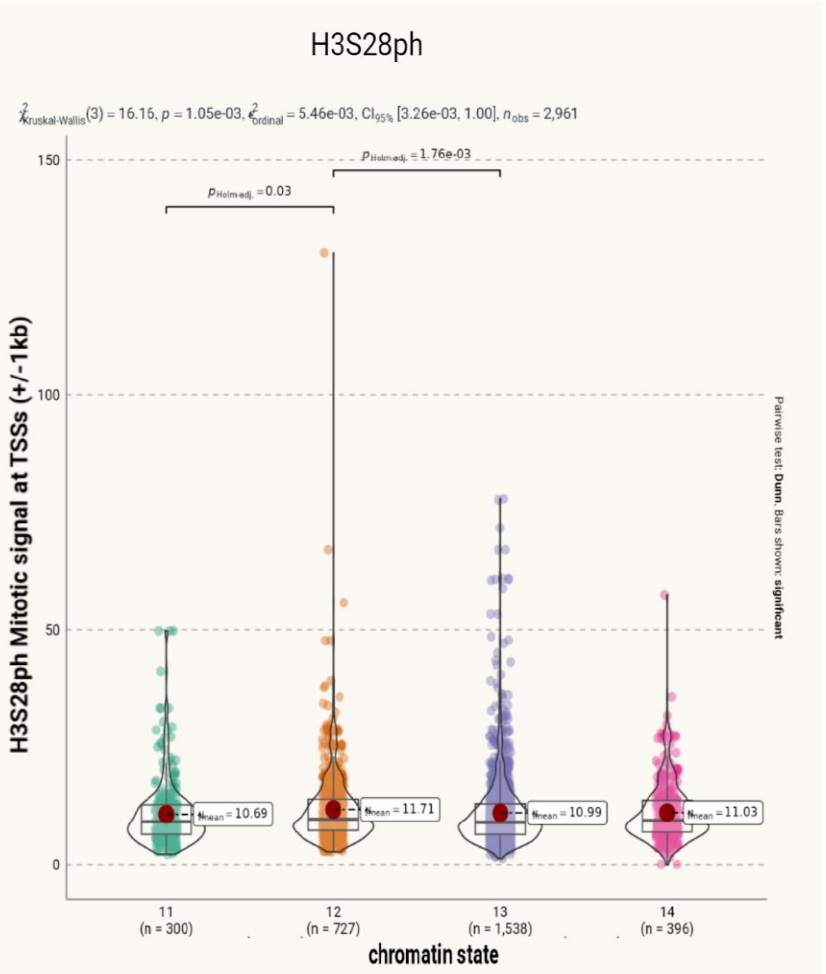
Figure A 3: Number of TSSs, both with and without a confident chromatin state called, with each gene type, according to biomart data


"Elongation" states:
H3S10ph
$\chi^2_{Kruskal\text{-}Wallis}(4) = 440.48, p = 4.96e\text{-}94, \epsilon^2_{ordinal} = 0.07, CI_{95\%}\ [0.06, 1.00], n_{obs} = 6,286$
$p_{Holm\text{-}adj.} = 2.08e\text{-}14$
$p_{Holm\text{-}adj.} = 5.86e\text{-}34$
$p_{Holm\text{-}adj.} = 2.16e\text{-}72$
$p_{Holm\text{-}adj.} = 6.24e\text{-}24$
$p_{Holm\text{-}adj.} = 2.69e\text{-}49$
$p_{Holm\text{-}adj.} = 1.73e\text{-}08$
$p_{Holm\text{-}adj.} = 1.14e\text{-}20$
$p_{Holm\text{-}adj.} = 1.47e\text{-}18$
H3S10ph Mitotic signal at TSSs (+/-1kb)
300
200
100
0
$\bar{x}_{mean} = 10.72$
$\bar{x}_{mean} = 14.44$
$\bar{x}_{mean} = 13.35$
$\bar{x}_{mean} = 8.77$
$\bar{x}_{mean} = 10.62$
1 (n = 772)
2 (n = 511)
3 (n = 1,437)
4 (n = 1,347)
5 (n = 2,219)
chromatin state
Pairwise test: Dunn. Bars shown: significant
H3S28ph
$\chi^2_{Kruskal\text{-}Wallis}(4) = 425.85, p = 7.20e\text{-}91, \epsilon^2_{ordinal} = 0.07, CI_{95\%}\ [0.06, 1.00], n_{obs} = 6,286$
$p_{Holm\text{-}adj.} = 3.32e\text{-}10$
$p_{Holm\text{-}adj.} = 1.07e\text{-}36$
$p_{Holm\text{-}adj.} = 1.71e\text{-}66$
$p_{Holm\text{-}adj.} = 2.20e\text{-}27$
$p_{Holm\text{-}adj.} = 4.12e\text{-}48$
$p_{Holm\text{-}adj.} = 9.33e\text{-}09$
$p_{Holm\text{-}adj.} = 1.19e\text{-}17$
$p_{Holm\text{-}adj.} = 1.36e\text{-}17$
H3S28ph Mitotic signal at TSSs (+/-1kb)
200
100
0
$\bar{x}_{mean} = 9.86$
$\bar{x}_{mean} = 13.28$
$\bar{x}_{mean} = 12.08$
$\bar{x}_{mean} = 8.29$
$\bar{x}_{mean} = 9.65$
1 (n = 772)
2 (n = 511)
3 (n = 1,437)
4 (n = 1,347)
5 (n = 2,219)
chromatin state
Pairwise test: Dunn. Bars shown: significant


"Heterochromatin" states:
H3S10ph
$\chi^2_{Kruskal-Wallis}(4) = 2752.28, p = 0.00, \epsilon^2_{ordinal} = 0.20, CI_{95\%} [0.20, 1.00], n_{obs} = 13,487$
$p_{Holm-adj.} = 1.99e-28$
$p_{Holm-adj.} = 0.05$
$p_{Holm-adj.} = 1.10e-24$
$p_{Holm-adj.} = 0.00$
$p_{Holm-adj.} = 1.47e-68$
$p_{Holm-adj.} = 4.62e-41$
$p_{Holm-adj.} = 5.67e-24$
$p_{Holm-adj.} = 5.44e-12$
$p_{Holm-adj.} = 0.01$
H3S10ph Mitotic signal at TSSs (+/- 1kb)
Pairwise test: Dunn. Bars shown: significant
$\bar{x}_{mean} = 7.28$
$\bar{x}_{mean} = 5.73$
$\bar{x}_{mean} = 12.29$
$\bar{x}_{mean} = 11.33$
$\bar{x}_{mean} = 5.83$
6 (n = 542)
7 (n = 7,409)
8 (n = 283)
9 (n = 5,079)
10 (n = 174)
chromatin state
H3S28ph
$\chi^2_{Kruskal-Wallis}(4) = 2050.79, p = 0.00, \epsilon^2_{ordinal} = 0.15, CI_{95\%} [0.14, 1.00], n_{obs} = 13,487$
$p_{Holm-adj.} = 3.57e-26$
$p_{Holm-adj.} = 3.73e-20$
$p_{Holm-adj.} = 0.00$
$p_{Holm-adj.} = 1.65e-46$
$p_{Holm-adj.} = 2.74e-35$
$p_{Holm-adj.} = 1.03e-17$
$p_{Holm-adj.} = 5.45e-07$
$p_{Holm-adj.} = 7.85e-03$
H3S28ph Mitotic signal at TSSs (+/-1kb)
Pairwise test: Dunn. Bars shown: significant
$\bar{x}_{mean} = 6.75$
$\bar{x}_{mean} = 5.93$
$\bar{x}_{mean} = 10.32$
$\bar{x}_{mean} = 10.05$
$\bar{x}_{mean} = 5.79$
6 (n = 542)
7 (n = 7,409)
8 (n = 283)
9 (n = 5,079)
10 (n = 174)
chromatin state

"Enhancer" states:





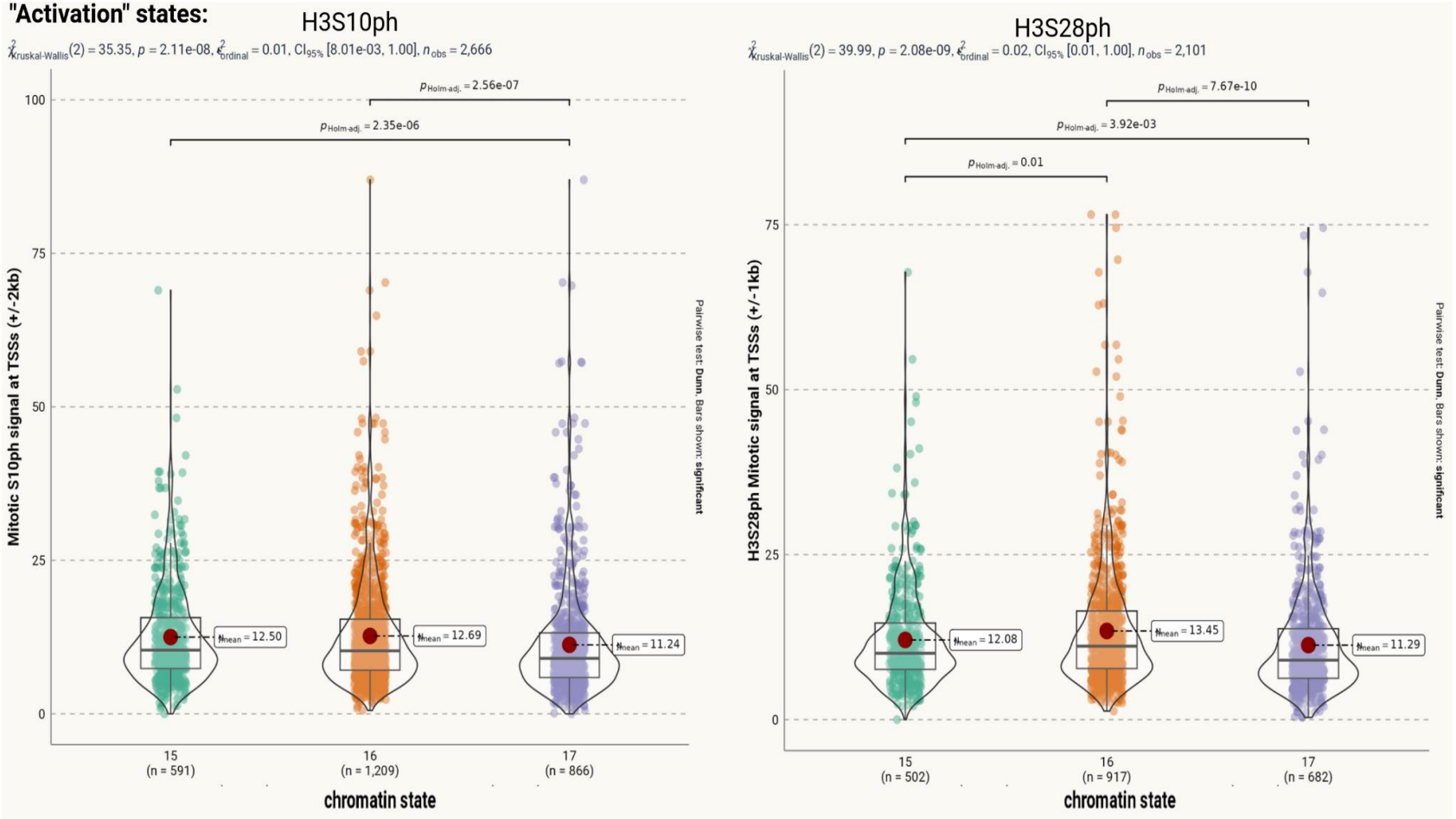

"Activation" states:
H3S10ph
$\chi^2_{Kruskal\text{-}Wallis}(2) = 35.35$, $p = 2.11e\text{-}08$, $\epsilon^2_{ordinal} = 0.01$, $CI_{95\%}$ [8.01e-03, 1.00], $n_{obs} = 2,666$
$p_{Holm\text{-}adj.} = 2.56e\text{-}07$
$p_{Holm\text{-}adj.} = 2.35e\text{-}06$
Mitotic S10ph signal at TSSs (+/-2kb)
100
75
50
25
0
$\bar{x}_{mean} = 12.50$
$\bar{x}_{mean} = 12.69$
$\bar{x}_{mean} = 11.24$
Pairwise test: Dunn. Bars shown: significant
15
(n = 591)
16
(n = 1,209)
17
(n = 866)
chromatin state
H3S28ph
$\chi^2_{Kruskal\text{-}Wallis}(2) = 39.99$, $p = 2.08e\text{-}09$, $\epsilon^2_{ordinal} = 0.02$, $CI_{95\%}$ [0.01, 1.00], $n_{obs} = 2,101$
$p_{Holm\text{-}adj.} = 7.67e\text{-}10$
$p_{Holm\text{-}adj.} = 3.92e\text{-}03$
$p_{Holm\text{-}adj.} = 0.01$
H3S28ph Mitotic signal at TSSs (+/-1kb)
75
50
25
0
$\bar{x}_{mean} = 12.08$
$\bar{x}_{mean} = 13.45$
$\bar{x}_{mean} = 11.29$
Pairwise test: Dunn. Bars shown: significant
15
(n = 502)
16
(n = 917)
17
(n = 682)
chromatin state

Figure A 4: Violin-box-scatter plots of H3S10ph and H3S28ph at TSSs +/- 1 kb in each of 20 chromatin states.

Kruskal-Wallis test results shown above each plot. Post-hoc Dunn test values are shown where significant.

References

A Grammar of Data Manipulation [R package dplyr version 1.1.4]. (2023). https://doi.org/10.32614/CRAN.PACKAGE.DPLYR

Access the Bioconductor Project Package Repository [R package BiocManager version 1.30.23]. (2024). https://doi.org/10.32614/CRAN.PACKAGE.BIOCMANAGER

Aihara, H., Nakagawa, T., Yasui, K., Ohta, T., Hirose, S., Dhomae, N., Takio, K., Kaneko, M., Takeshima, Y., Muramatsu, M., & Ito, T. (2004). Nucleosomal histone kinase-1 phosphorylates H2A Thr 119 during mitosis in the early Drosophila embryo. *Genes & Development*, *18*(8), 877–888. https://doi.org/10.1101/gad.1184604

Ali, M., Rincón-Arano, H., Zhao, W., Rothbart, S. B., Tong, Q., Parkhurst, S. M., Strahl, B. D., Deng, L. W., Groudine, M., & Kutateladzea, T. G. (2013). Molecular basis for chromatin binding and regulation of MLL5. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(28), 11296–11301. https://doi.org/10.1073/PNAS.1310156110/-/DCSUPPLEMENTAL/SD01.XLSX

Altemose, N., Logsdon, G. A., Bzikadze, A. v., Sidhwani, P., Langley, S. A., Caldas, G. v., Hoyt, S. J., Uralsky, L., Ryabov, F. D., Shew, C. J., Sauria, M. E. G., Borchers, M., Gershman, A., Mikheenko, A., Shepelev, V. A., Dvorkina, T., Kunyavskaya, O., Vollger, M. R., Rhie, A., … Miga, K. H. (2022). *Complete genomic and epigenetic maps of human centromeres*. *376*(6588). https://www.science.org/doi/10.1126/science.abl4178

Amanda J. Broad, Keith F. DeLuca, Jennifer G. DeLuca; Aurora B kinase is recruited to multiple discrete kinetochore and centromere regions in human cells. *J Cell Biol* 2 March 2020; 219 (3): e201905144. doi: https://doi.org/10.1083/jcb.201905144

Amemiya, H. M., Kundaje, A., Boyle, A. P. (2019). The ENCODE Blacklist: Identification of Problematic Regions of the Genome. *Scientific Reports*, 9, 9354. https://www.nature.com/articles/s41598-019-45839-z

Anna De Antoni, Stefano Maffini, Stefan Knapp, Andrea Musacchio, Stefano Santaguida; A small-molecule inhibitor of Haspin alters the kinetochore functions of Aurora B. *J Cell Biol* 15 October 2012; 199 (2): 269–284. doi: https://doi.org/10.1083/jcb.201205119

Arrigoni, L., Al-Hasani, H., Ramírez, F., Panzeri, I., Ryan, D. P., Santacruz, D., Kress, N., Pospisilik, J. A., Bönisch, U., & Manke, T. (2018). RELACS nuclei barcoding enables high-throughput ChIP-seq. *Communications Biology*, *1*, 214. https://doi.org/10.1038/s42003-018-0219-z

Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. Retrieved September 1, 2024, from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Baker, M. Reproducibility crisis: Blame it on the antibodies. *Nature* **521**, 274–276 (2015). https://doi.org/10.1038/521274a

Baranello, L., Kouzine, F., Sanford, S., & Levens, D. (2016). ChIP bias as a function of cross-linking time. *Chromosome Research: An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology*, *24*(2), 175–181. https://doi.org/10.1007/s10577-015-9509-1

Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., & Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell*, *129*(4), 823–837. https://doi.org/10.1016/j.cell.2007.05.009

bcl2fastq Conversion Software. (n.d.). https://emea.support.illumina.com/sequencing/sequencing_software/bcl2fastq-conversion-software.html

Behera, V., Stonestrom, A. J., Hamagami, N., Hsiung, C. C., Keller, C. A., Giardine, B., Sidoli, S., Yuan, Z-F., Bhanu, N. V., Werner, M. T., Wang, H., Hardison, R. C., Blobel, G. A. (2019). Interrogating histone acetylation and BRD4 as mitotic bookmarks of transcription. *Cell Reports*, 27(2), 400-415. https://doi.org/10.1016/j.celrep.2019.03.057.

Bellec, M., Dufourt, J., Hunt, G. *et al.* The control of transcriptional memory by stable mitotic bookmarking. *Nat Commun* **13**, 1176 (2022). https://doi.org/10.1038/s41467-022-28855-y

Benayoun BA, Pollina EA, Ucar D, Mahmoudi S, Karra K, Wong ED, Devarajan K, Daugherty AC, Kundaje AB, Mancini E, Hitz BC, Gupta R, Rando TA, Baker JC, Snyder MP, Cherry JM, Brunet A. H3K4me3 breadth is linked to cell identity and transcriptional consistency. Cell. 2014 Jul 31;158(3):673-88. doi: 10.1016/j.cell.2014.06.027.

Benjamini, Y. and Hochberg, Y. (1995), Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological), 57: 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bergmann JH, Rodríguez MG, Martins NM, Kimura H, Kelly DA, Masumoto H, Larionov V, Jansen LE, Earnshaw WC. Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. EMBO J. 2011 Jan 19;30(2):328-40. doi: 10.1038/emboj.2010.329.

Bernstein, B. E., Humphrey, E. L., Erlich, R. L., Schneider, R., Bouman, P., Liu, J. S., Kouzarides, T., & Schreiber, S. L. (2002). Methylation of histone H3 Lys 4 in coding regions of active genes. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(13), 8695–8700. https://doi.org/10.1073/pnas.082249499

Blat, Y., & Kleckner, N. (1999). Cohesins bind to preferential sites along yeast chromosome III, with differential regulation along arms versus the centric region. *Cell*, *98*(2), 249–259. https://doi.org/10.1016/s0092-8674(00)81019-3

Bock, C., Halachev, K., Büch, J., & Lengauer, T. (2009). EpiGRAPH: user-friendly software for statistical analysis and prediction of (epi)genomic data. *Genome Biology*, *10*(2), R14. https://doi.org/10.1186/gb-2009-10-2-r14

Bock, I., Dhayalan, A., Kudithipudi, S., Brandt, O., Rathert, P., & Jeltsch, A. (2011). Detailed specificity analysis of antibodies binding to modified histone tails with peptide arrays. *Epigenetics*, *6*(2), 256–263. https://doi.org/10.4161/epi.6.2.13837

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014 Aug 1;30(15):2114-20. doi: 10.1093/bioinformatics/btu170.

Bowtie 2: fast and sensitive read alignment. (n.d.). https://bowtie-bio.sourceforge.net/bowtie2/index.shtml

Campbell, A. E., Hsiung, C. C. S., & Blobel, G. A. (2014). Comparative Analysis of Mitosis-Specific Antibodies For Bulk Purification of Mitotic Populations By Fluorescence-Activated Cell Sorting. *BioTechniques*, *56*(2), 90–94. https://doi.org/10.2144/000114137

Caravaca JM, Donahue G, Becker JS, He X, Vinson C, Zaret KS. Bookmarking by specific and nonspecific binding of FoxA1 pioneer factor to mitotic chromosomes. Genes Dev. 2013 Feb 1;27(3):251-60. doi: 10.1101/gad.206458.112.

Carroll, T., Barrows, D., & Carroll, M. T. (2023). *Package "profileplyr" Type Package Title Visualization and annotation of read signal over genomic ranges with profileplyr*. https://git.bioconductor.org/packages/profileplyr

Chabbert, C. D., Adjalley, S. H., Steinmetz, L. M., & Pelechano, V. (2018). Multiplexed ChIP-Seq Using Direct Nucleosome Barcoding: A Tool for High-Throughput Chromatin Analysis. *Methods in Molecular Biology (Clifton, N.J.)*, *1689*, 177–194. https://doi.org/10.1007/978-1-4939-7380-4_16

Charif, D., & Lobry, J. R. (2007). Seqin\{{R}} 1.0-2: a contributed package to the \{{R}} project for statistical computing devoted to biological sequences retrieval and analysis. In U. Bastolla, M. Porto, H. E. Roman, & M. Vendruscolo (Eds.), *Structural approaches to sequence evolution: Molecules, networks, populations* (pp. 207–232). Springer Verlag.

Chen K, Chen Z, Wu D, Zhang L, Lin X, Su J, Rodriguez B, Xi Y, Xia Z, Chen X, Shi X, Wang Q, Li W. Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. Nat Genet. 2015 Oct;47(10):1149-57. doi: 10.1038/ng.3385.

Chen, C. C. L., Goyal, P., Karimi, M. M., Abildgaard, M. H., Kimura, H., & Lorincz, M. C. (2018). H3S10ph broadly marks early-replicating domains in interphase ESCs and shows reciprocal antagonism with H3K9me2. *Genome Research*, *28*(1), 37–51. https://doi.org/10.1101/gr.224717.117

Cimini, D., Wan, X., Hirel, C. B., & Salmon, E. D. (2006). Aurora Kinase Promotes Turnover of Kinetochore Microtubules to Reduce Chromosome Segregation Errors. *Current Biology*, *16*(17), 1711–1718. https://doi.org/10.1016/j.cub.2006.07.022

Complete Step-by-Step Guide to CUT\&{RUN}. (n.d.). In *EpiCypher*. https://www.epicypher.com/resources/blogcomplete-stepbystep-guide-to-cutrun/

Contreras A, Perea-Resa C. Transcriptional repression across mitosis: mechanisms and functions. Biochem Soc Trans. 2024 Feb 28;52(1):455-464. doi: 10.1042/BST20231071.

Conway, J. R., Lex, A., & Gehlenborg, N. (2017). UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics*, *33*(18), 2938–2940. https://doi.org/10.1093/BIOINFORMATICS/BTX364

D'Anna, J. A., Tobey, R. A., Barham, S. S., & Gurley, L. R. (1977). A reduction in the degree of H4 acetylation during mitosis in Chinese hamster cells. *Biochemical and Biophysical Research Communications*, *77*(1), 187–194. https://doi.org/10.1016/s0006-291x(77)80181-2

Dahl JA, Jung I, Aanes H, Greggains GD, Manaf A, Lerdrup M, Li G, Kuan S, Li B, Lee AY, Preissl S, Jermstad I, Haugen MH, Suganthan R, Bjørås M, Hansen K, Dalen KT, Fedorcsak P, Ren B, Klungland A. Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. Nature. 2016 Sep 22;537(7621):548-552. doi: 10.1038/nature19360.

Dai, J., Sullivan, B. A., Higgins, J. M. G (2006). Regulation of Mitotic Chromosome Cohesion by Haspin and Aurora B. *Dev. Cell, 11(5),* 741-750. https://doi.org/10.1016/j.devcel.2006.09.018

Dai, J., Sultan, S., Taylor, S. S., & Higgins, J. M. G. (2005). The kinase haspin is required for mitotic histone H3 Thr 3 phosphorylation and normal metaphase chromosome alignment. *Genes & Development*, *19*(4), 472–488. https://doi.org/10.1101/gad.1267105

Petr Danecek, James K Bonfield, Jennifer Liddle, John Marshall, Valeriu Ohan, Martin O Pollard, Andrew Whitwham, Thomas Keane, Shane A McCarthy, Robert M Davies, Heng Li, Twelve years of SAMtools and BCFtools, *GigaScience*, Volume 10, Issue 2, February 2021, giab008, https://doi.org/10.1093/gigascience/giab008

Dastidar, E. G., Dzeyk, K., Krijgsveld, J., Malmquist, N. A., Doerig, C., Scherf, A., & Lopez-Rubio, J.-J. (2013). Comprehensive Histone Phosphorylation Analysis and Identification of Pf14-3-3 Protein as a Histone H3 Phosphorylation Reader in Malaria Parasites. *PLOS ONE*, *8*(1), e53179. https://doi.org/10.1371/journal.pone.0053179

de la Barre, A.-E., Angelov, D., Molla, A., & Dimitrov, S. (2001). The N-terminus of histone H2B, but not that of histone H3 or its phosphorylation, is essential for chromosome condensation. *The EMBO Journal*, *20*(22), 6383–6393. https://doi.org/10.1093/emboj/20.22.6383

Dekker, J., Rippe, K., Dekker, M., & Kleckner, N. (2002). Capturing chromosome conformation. *Science (New York, N.Y.)*, *295*(5558), 1306–1311. https://doi.org/10.1126/science.1067799

Deluz C, Friman ET, Strebinger D, Benke A, Raccaud M, Callegari A, Leleu M, Manley S, Suter DM. A role for mitotic bookmarking of SOX2 in pluripotency and differentiation. Genes Dev. 2016 Nov 15;30(22):2538-2550. doi: 10.1101/gad.289256.116. Epub 2016 Dec 5.

Dhawal Jain, Sandro Baldi, Angelika Zabel, Tobias Straub, Peter B. Becker, Active promoters give rise to false positive 'Phantom Peaks' in ChIP-seq experiments, *Nucleic Acids Research*, Volume 43, Issue 14, 18 August 2015, Pages 6959–6968, https://doi.org/10.1093/nar/gkv637

Dileep, V., Ay, F., Sima, J., Vera, D. L., Noble, W. S., Gilbert, D. M. (2015). Topologically associating domains and their long-range contacts are established during early G1 coincident with the establishment of the replication-timing program. *Genome Res, 25(8),* 1104-13. https://pubmed.ncbi.nlm.nih.gov/25995270/

Ditchfield, C., Johnson, V. L., Tighe, A., Ellston, R., Haworth, C., Johnson, T., Mortlock, A., Keen, N., & Taylor, S. S. (2003). Aurora B couples chromosome alignment with anaphase by targeting BubR1, Mad2, and Cenp-E to kinetochores. *The Journal of Cell Biology*, *161*(2), 267–280. https://doi.org/10.1083/jcb.200208091

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S., Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, *485*, *276-280*. https://www.nature.com/articles/nature11082

Dostie, J., Richmond, T. A., Arnaout, R. A., Selzer, R. R., Lee, W. L., Honan, T. A., Rubio, E. D., Krumm, A., Lamb, J., Nusbaum, C., Green, R. D., & Dekker, J. (2006). Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Research*, *16*(10), 1299–1309. https://doi.org/10.1101/gr.5571506

Dozmorov, M. G. (2017). Epigenomic annotation-based interpretation of genomic data: from enrichment analysis to machine learning. *Bioinformatics*, *33*(20), 3323–3330. https://doi.org/10.1093/BIOINFORMATICS/BTX414

Drobic, B., Pérez-Cadahía, B., Yu, J., Kung, S. K.-P., & Davie, J. R. (2010). Promoter chromatin remodeling of immediate-early genes is mediated through H3 phosphorylation at either serine 28 or 10 by the MSK1 multi-protein complex. *Nucleic Acids Research*, *38*(10), 3196–3208. https://doi.org/10.1093/nar/gkq030

Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., Moor, B. de, Brazma, A., & Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics (Oxford, England)*, *21*(16), 3439–3440. https://doi.org/10.1093/BIOINFORMATICS/BTI525

Durinck, S., Spellman, P. T., Birney, E., & Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nature Protocols 2009 4:8*, *4*(8), 1184–1191. https://doi.org/10.1038/nprot.2009.97

Eastman, A. E., & Guo, S. (2020). The palette of techniques for cell cycle analysis. *FEBS Letters*, *594*(13), 2084–2098. https://doi.org/10.1002/1873-3468.13842

Egli, D., Birkhoff, G., & Eggan, K. (2008). Mediators of reprogramming: transcription factors and transitions through mitosis. *Nature Reviews. Molecular Cell Biology*, *9*(7), 505–516. https://doi.org/10.1038/nrm2439

Elizarraras, J. M., Liao, Y., Shi, Z., Zhu, Q., Pico, A. R., & Zhang, B. (2024). WebGestalt 2024: faster gene set analysis and new support for metabolomics and multi-omics. *Nucleic Acids Research*, *52*(W1), W415–W421. https://doi.org/10.1093/NAR/GKAE456

ENCODE Project Consortium. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*(7414), 57–74. https://doi.org/10.1038/nature11247

Enrique Carrillo-de-Santa-Pau, David Juan, Vera Pancaldi, Felipe Were, Ignacio Martin-Subero, Daniel Rico, Alfonso Valencia, on behalf of The BLUEPRINT Consortium, Automatic identification of informative regions with epigenomic changes associated to hematopoiesis, *Nucleic Acids Research*, Volume 45, Issue 16, 19 September 2017, Pages 9244–9259, https://doi.org/10.1093/nar/gkx618

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network Visualizations of Relationships in Psychometric Data. *Journal of Statistical Software*, *48*, 1–18. https://doi.org/10.18637/JSS.V048.I04

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

Favorov, A., Mularoni, L., Cope, L. M., Medvedeva, Y., Mironov, A. A., Makeev, V. J., & Wheelan, S. J. (2012). Exploring massive, genome scale datasets with the GenometriCorr package. *PLoS Computational Biology*, *8*(5), e1002529. https://doi.org/10.1371/journal.pcbi.1002529

Festuccia, N., Dubois, A., Vandormael-Pournin, S., Gallego Tejeda, E., Mouren, A., Bessonnard, S., Mueller, F., Proux, C., Cohen-Tannoudji, M., & Navarro, P. (2016). Mitotic binding of Esrrb marks key regulatory regions of the pluripotency network. *Nature Cell Biology*, *18*(11), 1139–1148. https://doi.org/10.1038/ncb3418

Flanagan, J. F., Mi, L. Z., Chruszcz, M., Cymborowski, M., Clines, K. L., Kim, Y., Minor, W., Rastinejad, F., & Khorasanizadeh, S. (2005). Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature*, *438*(7071), 1181–1185. https://doi.org/10.1038/NATURE04290

Fuller, B. G., Lampson, M. A., Foley, E. A., Rosasco-Nitcher, S., Le, K. v, Tobelmann, P., Brautigan, D. L., Stukenberg, P. T., & Kapoor, T. M. (2008). Midzone activation of aurora B in anaphase produces an intracellular phosphorylation gradient. *Nature*, *453*(7198), 1132–1136. https://doi.org/10.1038/nature06923

Gadea, B. B., & Ruderman, J. v. (2005). Aurora Kinase Inhibitor ZM447439 Blocks Chromosome-induced Spindle Assembly, the Completion of Chromosome Condensation, and the Establishment of the Spindle Integrity Checkpoint in Xenopus Egg Extracts. *Molecular Biology of the Cell*, *16*(3), 1305–1318. https://doi.org/10.1091/mbc.E04-10-0891

Garg, G., Patel, P., Gupta, G. das, & Kurmi, B. das. (2024). A Review on Working Principle and Advanced Applications of Fluorescence activated Cell Sorting Machine (FACS). *Current Pharmaceutical Analysis*, *20*(2), 85–97. https://doi.org/10.2174/0115734129301463240306064929

Garske, A. L., Oliver, S. S., Wagner, E. K., Musselman, C. A., Leroy, G., Garcia, B. A., Kutateladze, T. G., & Denu, J. M. (2010). Combinatorial profiling of chromatin binding modules reveals multisite discrimination. *Nature Chemical Biology*, *6*(4), 283–290. https://doi.org/10.1038/NCHEMBIO.319

Gatchalian, J., Gallardo, C. M., Shinsky, S. A., Ospina, R. R., Liendo, A. M., Krajewski, K., Klein, B. J., Andrews, F. H., Strahl, B. D., van Wely, K. H. M., & Kutateladze, T. G. (2016). Chromatin condensation and recruitment of PHD finger proteins to histone H3K4me3 are mutually exclusive. *Nucleic Acids Research*, *44*(13), 6102–6112. https://doi.org/10.1093/NAR/GKW193

Gehani, S. S., Agrawal-Singh, S., Dietrich, N., Christophersen, N. S., Helin, K., Hansen, K. (2010) Polycomb Group Protein Displacement and Gene Activation through MSK-Dependent H3K27me3S28 Phosphorylation. *Molecular Cell, 39(6),* 886-900. https://doi.org/10.1016/j.molcel.2010.08.020

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

Gilmour, D. S., & Lis, J. T. (1984). Detecting protein-DNA interactions in vivo: distribution of RNA polymerase on specific bacterial genes. *Proceedings of the National Academy of Sciences of the United States of America*, *81*(14), 4275–4279. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC345570/

Gilmour, D. S., & Lis, J. T. (1985). In vivo interactions of RNA polymerase II with genes of Drosophila melanogaster. *Molecular and Cellular Biology*, *5*(8), 2009–2018. https://doi.org/10.1128/mcb.5.8.2009-2018.1985

Girardot, C., Scholtalbers, J., Sauer, S., Su, S. Y., & Furlong, E. E. M. (2016). Je, a versatile suite to handle multiplexed NGS libraries with unique molecular identifiers. *BMC Bioinformatics*, *17*(1), 1–6. https://doi.org/10.1186/S12859-016-1284-2/FIGURES/2

Gonzalez, I., Molliex, A., & Navarro, P. (2021). Mitotic memories of gene activity. *Current Opinion in Cell Biology*, *69*, 41–47. https://doi.org/10.1016/j.ceb.2020.12.009

Goto, H., Tomono, Y., Ajiro, K., Kosako, H., Fujita, M., Sakurai, M., Okawa, K., Iwamatsu, A., Okigaki, T., Takahashi, T., & Inagaki, M. (1999). Identification of a novel phosphorylation site on histone H3 coupled with mitotic chromosome condensation. *The Journal of Biological Chemistry*, *274*(36), 25543–25549. https://doi.org/10.1074/jbc.274.36.25543

Goto, H., Yasui, Y., Nigg, E. A., & Inagaki, M. (2002). Aurora-B phosphorylates Histone H3 at serine28 with regard to the mitotic chromosome condensation. *Genes to Cells: Devoted to Molecular & Cellular Mechanisms*, *7*(1), 11–17. https://doi.org/10.1046/j.1356-9597.2001.00498.x

Gottesfeld, J. M., & Forbes, D. J. (1997). Mitotic repression of the transcriptional machinery. *Trends in Biochemical Sciences*, *22*(6), 197–202. https://doi.org/10.1016/s0968-0004(97)01045-1

Guéron, M., Kochoyan, M., & Leroy, J.-L. (1987). A single mode of DNA base-pair opening drives imino proton exchange. *Nature*, *328*(6125), 89–92. https://doi.org/10.1038/328089a0

Gurley, L. R., D'Anna, J. A., Barham, S. S., Deaven, L. L., & Tobey, R. A. (1978). Histone phosphorylation and chromatin structure during mitosis in Chinese hamster cells. *European Journal of Biochemistry*, *84*(1), 1–15. https://doi.org/10.1111/j.1432-1033.1978.tb12135.x

Gurley, L. R., Walters, R. A., Tobey, R. A. (1973). Histone phosphorylation in late interphase and mitosis. (1973). *Biochemical and Biophysical Research Communications*, *50*(3), 744–750. https://doi.org/10.1016/0006-291X(73)91307-7

Halachev, K., Bast, H., Albrecht, F., Lengauer, T., & Bock, C. (2012). EpiExplorer: live exploration and global analysis of large epigenomic datasets. *Genome Biology*, *13*(10), R96. https://doi.org/10.1186/gb-2012-13-10-r96

Harlen, K. M., & Churchman, L. S. (2017). The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nature Reviews Molecular Cell Biology 2017 18:4*, *18*(4), 263–273. https://doi.org/10.1038/nrm.2017.10

Harris, R.J., Heer, M., Levasseur, M.D. *et al*. Release of Histone H3K4-reading transcription factors from chromosomes in mitosis is independent of adjacent H3 phosphorylation. *Nat Commun* **14**, 7243 (2023). https://doi.org/10.1038/s41467-023-43115-3

Hauf, S., Cole, R. W., LaTerra, S., Zimmer, C., Schnapp, G., Walter, R., Heckel, A., van Meel, J., Rieder, C. L., & Peters, J.-M. (2003). The small molecule Hesperadin reveals a role for Aurora B in correcting kinetochore-microtubule attachment and in maintaining the spindle assembly checkpoint. *The Journal of Cell Biology*, *161*(2), 281–294. https://doi.org/10.1083/jcb.200208092

Hebbes, T. R., Thorne, A. W., & Crane-Robinson, C. (1988). A direct link between core histone acetylation and transcriptionally active chromatin. *The EMBO Journal*, *7*(5), 1395–1402. https://doi.org/10.1002/j.1460-2075.1988.tb02956.x

Hendzel, M. J., Wei, Y., Mancini, M. A., van Hooser, A., Ranalli, T., Brinkley, B. R., Bazett-Jones, D. P., & Allis, C. D. (1997). Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation. *Chromosoma*, *106*(6), 348–360. https://doi.org/10.1007/s004120050256

Hengeveld, R. C. C., Vromans, M. J. M., Vleugel, M., Hadders, M. A., & Lens, S. M. A. (2017). Inner centromere localization of the CPC maintains centromere cohesion and allows mitotic checkpoint silencing. *Nature Communications*, *8*. https://doi.org/10.1038/NCOMMS15542

Hirota, T., Lipp, J. J., Toh, B.-H., & Peters, J.-M. (2005). Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature*, *438*(7071), 1176–1180. https://doi.org/10.1038/nature04254

Hoffman, M. M., Buske, O. J., Wang, J., Weng, Z., Bilmes, J. A., & Noble, W. S. (2012). Unsupervised pattern discovery in human chromatin structure through genomic segmentation. *Nature Methods*, *9*(5). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3340533/

Houston, S. I., McManus, K. J., Adams, M. M., Sims, J. K., Carpenter, P. B., Hendzel, M. J., & Rice, J. C. (2008). Catalytic function of the PR-Set7 histone H4 lysine 20 monomethyltransferase is essential for mitotic entry and genomic stability. *The Journal of Biological Chemistry*, *283*(28), 19478–19488. https://doi.org/10.1074/jbc.M710579200

Hsiung, C. C.-S., Bartman, C. R., Huang, P., Ginart, P., Stonestrom, A. J., Keller, C. A., Face, C., Jahn, K. S., Evans, P., Sankaranarayanan, L., Giardine, B., Hardison, R. C., Raj, A., & Blobel, G. A. (2016). A hyperactive transcriptional state marks genome reactivation at the mitosis-G1 transition. *Genes & Development*, *30*(12), 1423–1439. https://doi.org/10.1101/gad.280859.116

Hsu, J. Y., Sun, Z. W., Li, X., Reuben, M., Tatchell, K., Bishop, D. K., Grushcow, J. M., Brame, C. J., Caldwell, J. A., Hunt, D. F., Lin, R., Smith, M. M., & Allis, C. D. (2000). Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. *Cell*, *102*(3), 279–291. https://doi.org/10.1016/s0092-8674(00)00034-9

Ito, K., & Zaret, K. S. (2022). Maintaining Transcriptional Specificity through Mitosis. *Annual Review of Genomics and Human Genetics*, *23*, 53–71. https://doi.org/10.1146/annurev-genom-121321-094603

Jackson, V. (1978). Studies on histone organization in the nucleosome using formaldehyde as a reversible cross-linking agent. *Cell*, *15*(3), 945–954. https://doi.org/10.1016/0092-8674(78)90278-7

Janssen, K. A., Sidoli, S., & Garcia, B. A. (2017). Recent Achievements in Characterizing the Histone Code and Approaches to Integrating Epigenomics and Systems Biology. *Methods in Enzymology*, *586*, 359–378. https://doi.org/10.1016/bs.mie.2016.10.021

Javasky, E., Shamir, I., Gandhi, S., Egri, S., Sandler, O., Rothbart, S. B., Kaplan, N., Jaffe, J. D., Goren, A., Simon, I. (2018). Study of mitotic chromatin supports a model of bookmarking by histone modifications and reveals nucleosome deposition patterns. *Genome Res, 28(10,* 1455-66. https://pubmed.ncbi.nlm.nih.gov/30166406/

John, S., & Workman, J. L. (1998). Bookmarking genes for activation in condensed mitotic chromosomes. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *20*(4), 275–279. https://doi.org/10.1002/(SICI)1521-1878(199804)20:4<275::AID-BIES1>3.0.CO;2-P

Joo, HY., Zhai, L., Yang, C. *et al.* Regulation of cell cycle progression and gene expression by H2A deubiquitination. *Nature* **449**, 1068–1072 (2007). https://doi.org/10.1038/nature06256

Josefowicz, S. Z., Shimada, M., Armache, A., li, C. H., Miller, R. M., Lin, S., Yang, A., Dill, B. D., Molina, H., Park, H. S, Garcia, B. A., Taunton, J., Roeder, R. G., Allis, C. D. (2016). Chromatin

Kinases Act on Transcription Factors and Histone Tails in Regulation of Inducible Transcription. *Molecular Cell, 64(2),* 347-61.

Juan, D., Perner, J., Carrillo, E., Pau, S., Vingron, M., Rico, D., & Valencia, A. (2016). Epigenomic Co-localization and Co-evolution Reveal a Key Role for 5hmC as a Communication Hub in the Chromatin Network of ESCs. *Cell Reports*, *14*. https://doi.org/10.1016/j.celrep.2016.01.008

Kadauke, S., Udugama, M. I., Pawlicki, J. M., Achtam, J. C., Jain, D. P., Cheng, Y., Hardison, R. C., Blobel, G. A. (2012) Tissue-specific mitotic bookmarking by hematopoietic transcription factor GATA1. *Cell, 150(4),* 725-37. https://pubmed.ncbi.nlm.nih.gov/22901805/

Kanduri, Christoph Bock, Sveinung Gundersen, Eivind Hovig, Geir Kjetil Sandve, Colocalization analyses of genomic elements: approaches, recommendations and challenges, *Bioinformatics*, Volume 35, Issue 9, May 2019, Pages 1615–1624, https://doi.org/10.1093/bioinformatics/bty835

Kang, H., Shokhirev, M. N., Xu, Z., Chandran, S., Dixon, J. R., & Hetzer, M. W. (2020). Dynamic regulation of histone modifications and long-range chromosomal interactions during postmitotic transcriptional reactivation. *Genes & Development*, *34*(13–14), 913–930. https://doi.org/10.1101/gad.335794.119

Kassambara, A. (2023). *"ggplot2" Based Publication Ready Plots [R package ggpubr version 0.6.0]*. https://doi.org/10.32614/CRAN.PACKAGE.GGPUBR

Kawashima, S. A., Tsukahara, T., Langegger, M., Hauf, S., Kitajima, T. S., & Watanabe, Y. (2007). *Shugoshin enables tension-generating attachment of kinetochores by loading Aurora to centromeres*. *21*(4), 420–435. https://pubmed.ncbi.nlm.nih.gov/17322402/

Kawashima, S. A., Yamagishi, Y., Honda, T., Ishiguro, K., & Watanabe, Y. (2010). Phosphorylation of H2A by Bub1 prevents chromosomal instability through localizing shugoshin. *Science (New York, N.Y.)*, *327*(5962), 172–177. https://doi.org/10.1126/science.1180189

Kelly., A, Ghenoiu, C., Xue, J. Z., Zierhhut, C., Kimura, H., Funabiki, H. (2010). Survivin Reads Phosphorylated Histone H3 Threonine 3 to Activate the Mitotic Kinase Aurora B.*Science***330**,235-239.DOI:10.1126/science.1189505

Kind, J., Pagie, L., Ortabozkoyun, H., Boyle, S., de Vries, S. S., Janssen, H., Amendola, M., Nolen, L. D., Bickmore, W. A., van Steensel, B. (2013). Single-Cell Dynamics of Genome Nuclear Lamina Interactions. *Cell, 153(1),* 178-192.

Kitajima, T. S., Kawashima, S. A., & Watanabe, Y. (2004). The conserved kinetochore protein shugoshin protects centromeric cohesion during meiosis. *Nature*, *427*(6974), 510–517. https://doi.org/10.1038/nature02312

Komar, D., & Juszczynski, P. (2020). Rebelled epigenome: histone H3S10 phosphorylation and H3S10 kinases in cancer biology and therapy. *Clinical Epigenetics*, *12*, 147. https://doi.org/10.1186/s13148-020-00941-2

Kouzarides, T. (2007) Chromatin Modifications and Their Function. *Cell, 128(4),* 693-705. https://doi.org/10.1016/j.cell.2007.02.005

Kruhlak MJ, Hendzel MJ, Fischle W, Bertos NR, Hameed S, Yang XJ, Verdin E, Bazett-Jones DP. Regulation of global acetylation in mitosis through loss of histone acetyltransferases and deacetylases from chromatin. J Biol Chem. 2001 Oct 12;276(41):38307-19. doi: 10.1074/jbc.M100290200.

Kruidenier, L., Chung, C. W., Cheng, Z., Liddle, J., Che, K., Joberty, G., Bantscheff, M., Bountra, C., Bridges, A., Diallo, H., Eberhard, D., Hutchinson, S., Jones, E., Katso, R., Leveridge, M., Mander, P. K., Mosley, J., Ramirez-Molina, C., Rowland, P., … Wilson, D. M. (2012). A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. *Nature*, *488*(7411), 404–408. https://doi.org/10.1038/NATURE11262

Kukurba, K. R., & Montgomery, S. B. (2015). RNA Sequencing and Analysis. *Cold Spring Harbor Protocols*, *2015*(11), 951–969. https://doi.org/10.1101/pdb.top084970

Kumar, B., Elsasser, S. J., (2019). Quantitative Multiplexed ChIP Reveals Global Alterations that Shape Promoter Bivalency in Ground State Embryonic Stem Cells. *Cell Reports, 28(12),* 3274-84. https://www.cell.com/cell-reports/abstract/S2211-1247(19)31094-0

Landt, S. G., Marinov, G. K., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S., Bernstein, B. E., Bickel, P., Brown, J. B., Cayting, P., Chen, Y., DeSalvo, G., Epstein, C., Fisher-Aylor, K. I., Euskirchen, G., Gerstein, M., Gertz, J., Hartemink, A. J., Hoffman, M. M., … Snyder, M. (2012). ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Research*, *22*(9), 1813–1831. https://doi.org/10.1101/gr.136184.111

Lane, S. I. R., Chang, H.-Y., Jennings, P. C., & Jones, K. T. (2010). The Aurora kinase inhibitor ZM447439 accelerates first meiosis in mouse oocytes by overriding the spindle assembly checkpoint. *Reproduction (Cambridge, England)*, *140*(4), 521–530. https://doi.org/10.1530/REP-10-0223

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods 2012 9:4*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

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

Lara-Astiaso, D., Weiner, A., Lorenzo-Vivas, E., Zaretsky, I., Jaitin, D. A., David, E., Keren-Shaul, H., Mildner, A., Winter, D., Jung, S., Friedman, N., & Amit, I. (2014). Immunogenetics. Chromatin state dynamics during blood formation. *Science (New York, N.Y.)*, *345*(6199), 943–949. https://doi.org/10.1126/science.1256271

Lau, P. N. L., Cheung, P. (2011). Histone code pathway involving H3 S28 phosphorylation and K27 acetylation activates transcription and antagonizes polycomb silencing. *PNAS, 108(7),* 2801-2806. https://doi.org/10.1073/pnas.1012798108

Lawrence, M., Gentleman, R., & Carey, V. (2009). rtracklayer: an R package for interfacing with genome browsers. *Bioinformatics*, *25*(14), 1841–1842. https://doi.org/10.1093/BIOINFORMATICS/BTP328

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M. T., & Carey, V. J. (2013). *No Title*. *9*(8), e1003118. https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003118

Layer, R. M., Pedersen, B. S., DiSera, T., Marth, G. T., Gertz, J., & Quinlan, A. R. (2018). GIGGLE: a search engine for large-scale integrated genome analysis. *Nature Methods*, *15*(2), 123–126. https://doi.org/10.1038/nmeth.4556

Lee, N. R., Kim, H. S., Kim, Y. S., Kwon, M. H., Choi, K. S., & Lee, C. W. (2014). Regulation of the subcellular shuttling of Sgo1 between centromeres and chromosome arms by Aurora B-mediated phosphorylation. *Biochemical and Biophysical Research Communications*, *454*(3), 429–435. https://doi.org/10.1016/J.BBRC.2014.10.103

Li, F., & Ding, J. (n.d.). *REVIEW Sialylation is involved in cell fate decision during development, reprogramming and cancer progression*. https://doi.org/10.1007/s13238-018-0597-5

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Liang, K., Woodfin, A. R., Slaughter, B. D., Unruh, J. R., Box, A. C., Rickels, R. A., Gao, X., Haug, J. S., Jaspersen, S. L., & Shilatifard, A. (2015). Mitotic Transcriptional Activation: Clearance of Actively Engaged Pol II via Transcriptional Elongation Control in Mitosis. *Molecular Cell*, *60*(3), 435–445. https://doi.org/10.1016/j.molcel.2015.09.021

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

Liu X, Wang C, Liu W, Li J, Li C, Kou X, Chen J, Zhao Y, Gao H, Wang H, Zhang Y, Gao Y, Gao S. Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. Nature. 2016 Sep 22;537(7621):558-562. doi: 10.1038/nature19362.

Liu, Y., Pelham-Webb, B., Campigli Di Giammartino, D., Li, J., Kim, D., Kita, K., Saiz, N., Garg, V., Doane, A., Giannakakou, P., Hadjantonakis, A. K., Elemento, O., Apostolou, E. (2017). Widespread Mitotic Bookmarking by Histone Marks and Transcription Factors in Pluripotent Stem Cells. *Cell Rep., 19(7),* 1283-1293. https://pubmed.ncbi.nlm.nih.gov/28514649/

Luger, K., Mäder, A., Richmond, R. *et al.* Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997). https://doi.org/10.1038/38444

Lv J, Chen K. Broad H3K4me3 as A Novel Epigenetic Signature for Normal Development and Disease. Genomics Proteomics Bioinformatics. 2016 Oct;14(5):262-264. doi: 10.1016/j.gpb.2016.09.001.

Macdonald, N., Welburn, J. P. I., Noble, M. E. M., Nguyen, A., Yaffe, M. B., Clynes, D., Moggs, J. G., Orphanides, G., Thomson, S., Edmunds, J. W., Clayton, A. L., Endicott, J. A., & Mahadevan, L. C. (2005). No Title. *Molecular Cell*, *20*(2), 199–211. https://doi.org/10.1016/j.molcel.2005.08.032

Macs2 Anaconda.org. (n.d.). https://anaconda.org/bioconda/macs2

Mahadevan LC, Willis AC, Barratt MJ. Rapid histone H3 phosphorylation in response to growth factors, phorbol esters, okadaic acid, and protein synthesis inhibitors. Cell. 1991 May 31;65(5):775-83. doi: 10.1016/0092-8674(91)90385-c. PMID: 2040014.

Mangiafico, S. (2024). *Functions to Support Extension Education Program Evaluation [R package rcompanion version 2.4.36]*. https://doi.org/10.32614/CRAN.PACKAGE.RCOMPANION

Markaki, Y., Christogianni, A., Politou, A. S., & Georgatos, S. D. (2009). *Phosphorylation of histone H3 at Thr3 is part of a combinatorial pattern that marks and configures mitotic chromatin*. *122*(Pt 16), 2809–2819. https://doi.org/10.1242/jcs.043810

Martínez-Balbás, M. A., Dey, A., Rabindran, S. K., Ozato, K., & Wu, C. (1995). Displacement of sequence-specific transcription factors from mitotic chromatin. *Cell*, *83*(1), 29–38. https://doi.org/10.1016/0092-8674(95)90231-7

Meel, S., Mishra, S., Mann, R., Hussain Najar, A., Singh, A. K., 1#, D. N., Najar, A. H., Singh, A. K., & Notani, D. (2024). Compaction of Active Chromatin Domains and Promoters by H3S10 Phosphorylation during Mitosis Preserves Interphase-Specific Chromatin Structure and Function. *BioRxiv*, 2024.05.20.595059. https://doi.org/10.1101/2024.05.20.595059

Meers, M. P., Bryson, T., & Henikoff, S. (2019). Improved CUT&RUN chromatin profiling and analysis tools. *BioRxiv*, 569129. https://doi.org/10.1101/569129

Meyer, C. A., & Liu, X. S. (2014). Identifying and mitigating bias in next-generation sequencing methods for chromatin biology. *Nature Reviews. Genetics*, *15*(11), 709–721. https://doi.org/10.1038/nrg3788

Moazed, D. (2011). Mechanisms for the inheritance of chromatin states. *Cell*, *146*(4), 510–518. https://doi.org/10.1016/j.cell.2011.07.013

Mootha, V. K., Lindgren, C. M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., Houstis, N., Daly, M. J., Patterson, N., Mesirov, J. P., Golub, T. R., Tamayo, P., Spiegelman, B., Lander, E. S., Hirschhorn, J. N., … Groop, L. C. (2003). PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genetics*, *34*(3), 267–273. https://doi.org/10.1038/ng1180

Mora-Bermúdez, F., Gerlich, D. & Ellenberg, J. Maximal chromosome compaction occurs by axial shortening in anaphase and depends on Aurora kinase. *Nat Cell Biol* **9**, 822–831 (2007). https://doi.org/10.1038/ncb1606

Mueller, R. D., Yasuda, H., Hatch, C. L., Bonner, W. M., & Bradbury, E. M. (1985). Identification of ubiquitinated histones 2A and 2B in Physarum polycephalum. Disappearance of these proteins at metaphase and reappearance at anaphase. *The Journal of Biological Chemistry*, *260*(8), 5147–5153.

Muramoto T, Müller I, Thomas G, Melvin A, Chubb JR. Methylation of H3K4 Is required for inheritance of active transcriptional states. Curr Biol. 2010 Mar 9;20(5):397-406. doi: 10.1016/j.cub.2010.01.017

Nakato R, Sakata T. Methods for ChIP-seq analysis: A practical workflow and advanced applications. Methods. 2021 Mar;187:44-53. doi: 10.1016/j.ymeth.2020.03.005.

Nakato, R., & Shirahige, K. (2017). Recent advances in ChIP-seq analysis: from quality management to whole-genome annotation. *Briefings in Bioinformatics*, *18*(2), 279–290. https://doi.org/10.1093/bib/bbw023

Narendra V, Rocha PP, An D, Raviram R, Skok JA, Mazzoni EO, Reinberg D. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. Science. 2015 Feb 27;347(6225):1017-21. doi: 10.1126/science.1262088

Naumova, N., Imakaev, M., Fudenberg, G., Zhan, Y., Lajoie, B. R., Mirny, L. A., & Dekker, J. (2013). Organization of the mitotic chromosome. *Science (New York, N.Y.)*, *342*(6161), 948–953. https://doi.org/10.1126/science.1236083

Neurohr, G., Naegeli, A., Titos, I., Theler, D., Greber, B., Díez, J., Gabaldón, T., Mendoza, M., & Barral, Y. (2011). A midzone-based ruler adjusts chromosome compaction to anaphase spindle length. *Science (New York, N.Y.)*, *332*(6028), 465–468. https://doi.org/10.1126/science.1201578

Nishikori, S., Hattori, T., Fuchs, S. M., Yasui, N., Wojcik, J., Koide, A., Strahl, B. D., & Koide, S. (2012). Broad ranges of affinity and specificity of anti-histone antibodies revealed by a quantitative peptide immunoprecipitation assay. *Journal of Molecular Biology*, *424*(5), 391–399. https://doi.org/10.1016/j.jmb.2012.09.022

Noberini, R., Robusti, G., Bonaldi, T., Noberini, C. R., & Bonaldi, T. (2022). Mass spectrometry-based characterization of histones in clinical samples: applications, progress, and challenges. *The FEBS Journal*, *289*(5), 1191–1213. https://doi.org/10.1111/FEBS.15707

Nora, E. P., Lajoie, B. R., Schulz, E. G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N. L., Meisig, J., Sedat, J., Gribnau, J., Barillot, E., Blüthgen, N., Dekker, J., & Heard, E. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*, *485*(7398), 381–385. https://doi.org/10.1038/nature11049

Ogle, D. H., Doll, J. C., Wheeler, A. P., & Dinno, A. (2023). *Simple Fisheries Stock Assessment Methods [R package FSA version 0.9.5]*. https://doi.org/10.32614/CRAN.PACKAGE.FSA

Önder, Ö., Sidoli, S., Carroll, M., & Garcia, B. A. (2015). Progress in epigenetic histone modification analysis by mass spectrometry for clinical investigations. *Expert Review of Proteomics*, *12*(5), 499–517. https://doi.org/10.1586/14789450.2015.1084231

Oomen, M. E., Hansen, A. S., Liu, Y., Darzacq, X., & Dekker, J. (2019). CTCF sites display cell cycle-dependent dynamics in factor binding and nucleosome positioning. *Genome Research*, *29*(2), 236–249. https://doi.org/10.1101/gr.241547.118

Palozola KC, Donahue G, Liu H, Grant GR, Becker JS, Cote A, Yu H, Raj A, Zaret KS. Mitotic transcription and waves of gene reactivation during mitotic exit. Science. 2017 Oct 6;358(6359):119-122. doi: 10.1126/science.aal4671.

Palozola, K. C., Lerner, J., & Zaret, K. S. (2019). A changing paradigm of transcriptional memory propagation through mitosis. *Nature Reviews. Molecular Cell Biology*, *20*(1), 55–64. https://doi.org/10.1038/s41580-018-0077-z

Papini, D., Levasseu, M. D., Higgins, J. M. G. (2021). The Aurora B gradient sustains kinetochore stability in anaphase. *Cell Rep, 37(6),* 109818. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8595645/

Park, D., Lee, Y., Bhupindersingh, G., Iyer, V. R. (2013). Widespread Misinterpretable ChIP-seq Bias in Yeast. *PLOS ONE*. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0083506

Park, P. J. (2009). ChIP-seq: advantages and challenges of a maturing technology. *Nature Reviews. Genetics*, *10*(10), 669–680. https://doi.org/10.1038/nrg2641

Parsons, G. G., & Spencer, C. A. (1997). Mitotic repression of RNA polymerase II transcription is accompanied by release of transcription elongation complexes. *Molecular and Cellular Biology*, *17*(10), 5791–5802. https://doi.org/10.1128/MCB.17.10.5791

Paul, M., Barrows, D., & Carrol, T. (n.d.). *GitHub - RockefellerUniversity/Herper: The Herper package is a simple toolset to install and manage Conda packages and environments from R.* https://github.com/RockefellerUniversity/Herper

Pawaletz, N. (2001) Walter Flemming: pioneer of mitosis research. *Perspectives, 2. https://www.nature.com/scitable/content/Walther-Flemming-pioneer-of-mitosis-research-12650/*

Pelham-Webb, B., Polyzos, A., Wojenski, L., Kloetgen, A., Li, J., di Giammartino, D. C., Sakellaropoulos, T., Tsirigos, A., Core, L., & Apostolou, E. (2021). H3K27ac bookmarking promotes rapid post-mitotic activation of the pluripotent stem cell program without impacting 3D chromatin reorganization. *Molecular Cell*, *81*(8), 1732-1748.e8. https://doi.org/10.1016/j.molcel.2021.02.032

Petsalaki, E., & Zachos, G. (2021). The Abscission Checkpoint: A Guardian of Chromosomal Stability. *Cells*, *10*(12). https://doi.org/10.3390/CELLS10123350

Petsalaki, E., Dandoulaki, M., & Zachos, G. (2018). Chmp4c is required for stable kinetochore-microtubule attachments. *Chromosoma*, *127*(4), 461–473. https://doi.org/10.1007/S00412-018-0675-8/METRICS

Poleshko, A., Shah, P. P., Gupta, M., Babu, A., Morley, M. P., Manderfield, L. J., Ifkovits, J. L., Calderon, D., Aghajanian, H., Sierra-Pagan, J. E., Sun, Z., Wang, Q., Li, L., Dubois, N. C., Morrisey, E. E., Lazar, M. A., Smith, C. L., Epstein, J. A., Jain, R. (2017). Genome-Nuclear Lamina Interactions Regulate Cardiac Stem Cell Lineage Restriction. *Cell, 171(3),* 573-587.

Poleshko, A., Smith, C. L., Nguyen, S. C., Sivaramakrishnan, P., Wong, K. G., Murray, J. I., Lakadamyali, M/. Joyce, E. F., Jain, R., Epstein, J. A. (2019). H3K9me2 orchestrates inheritance of spatial positioning of peripheral heterochromatin through mitosis. *Cell Biology, Chromosomes and Gene Expression, 8,* 49278.

Prakash, K., & Fournier, D. (2017). Histone Code and Higher-Order Chromatin Folding: A Hypothesis. *Genomics and Computational Biology*, *3*(2), e41. https://doi.org/10.18547/gcb.2017.vol3.iss2.e41

Prasanth, K. v, Sacco-Bubulya, P. A., Prasanth, S. G., & Spector, D. L. (2003). Sequential entry of components of the gene expression machinery into daughter nuclei. *Molecular Biology of the Cell*, *14*(3), 1043–1057. https://doi.org/10.1091/mbc.e02-10-0669

Prescott, D. M., & Bender, M. A. (1962). Synthesis of RNA and protein during mitosis in mammalian tissue culture cells. *Experimental Cell Research*, *26*, 260–268. https://doi.org/10.1016/0014-4827(62)90176-3

Prigent, C., & Dimitrov, S. (2003). Phosphorylation of serine 10 in histone H3, what for? *Journal of Cell Science*, *116*(Pt 18), 3677–3685. https://doi.org/10.1242/jcs.00735

Quinlan, A. R. (2014). BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Current Protocols in Bioinformatics*, *47*(1), 11.12.1-11.12.34. https://doi.org/10.1002/0471250953.BI1112S47

R: The R Project for Statistical Computing. (2024). https://www.r-project.org/

Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A., & Manke, T. (2014). DeepTools: A flexible platform for exploring deep-sequencing data. *Nucleic Acids Research*, *42*(W1), W187–W191. https://doi.org/10.1093/nar/gku365

Reuben Thomas, Sean Thomas, Alisha K Holloway, Katherine S Pollard, Features that define the best ChIP-seq peak calling algorithms, *Briefings in Bioinformatics*, Volume 18, Issue 3, May 2017, Pages 441–450, https://doi.org/10.1093/bib/bbw035

Ribeiro, S. A., Vagnarelli, P., Dong, Y., Hori, T., McEwen, B. F., Fukagawa, T., Flors, C., & Earnshaw, W. C. (2010). A super-resolution map of the vertebrate kinetochore. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(23), 10484–10489. https://doi.org/10.1073/pnas.1002325107

Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M. J., Amin, V., Whitaker, J. W., Schultz, M. D., Ward, L. D., Sarkar, A., Quon, G., Sandstrom, R. S., Eaton, M. L., … Kellis, M. (2015). Integrative analysis of 111 reference human epigenomes. *Nature 2015 518:7539*, *518*(7539), 317–330. https://doi.org/10.1038/nature14248

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature Biotechnology 2011 29:1*, *29*(1), 24–26. https://doi.org/10.1038/nbt.1754

Robyr, D., & Grunstein, M. (2003). Genomewide histone acetylation microarrays. *Methods (San Diego, Calif.)*, *31*(1), 83–89. https://doi.org/10.1016/s1046-2023(03)00091-4

Robyr, D., Suka, Y., Xenarios, I., Kurdistani, S. K., Wang, A., Suka, N., & Grunstein, M. (2002). Microarray deacetylation maps determine genome-wide functions for yeast histone deacetylases. *Cell*, *109*(4), 437–446. https://doi.org/10.1016/s0092-8674(02)00746-8

Rodriguez, J. M., Pozo, F., Cerdán-Velez, D., di Domenico, T., Vázquez, J., & Tress, M. L. (2022). APPRIS: selecting functionally important isoforms. *Nucleic Acids Research*, *50*(D1), D54–D59. https://doi.org/10.1093/NAR/GKAB1058

Rothbart, S. B., Dickson, B. M., Raab, J. R., Grzybowski, A. T., Krajewski, K., Guo, A. H., Shanle, E. K., Josefowicz, S. Z., Fuchs, S. M., Allis, C. D., Magnuson, T. R., Ruthenburg, A. J., & Strahl, B. D. (2015). An Interactive Database for the Assessment of Histone Antibody Specificity. *Molecular Cell*, *59*(3), 502–511. https://doi.org/10.1016/j.molcel.2015.06.022

Santaguida, S., & Amon, A. (2015). Short- and long-term effects of chromosome mis-segregation and aneuploidy. *Nature Reviews Molecular Cell Biology*, *16*(8), 473–485. https://doi.org/10.1038/nrm4025

Sarkies, P., & Sale, J. E. (2012). Cellular epigenetic stability and cancer. *Trends in Genetics: TIG*, *28*(3), 118–127. https://doi.org/10.1016/j.tig.2011.11.005

Schoelz, J.M., Riddle, N.C. Functions of HP1 proteins in transcriptional regulation. *Epigenetics & Chromatin* **15**, 14 (2022). https://doi.org/10.1186/s13072-022-00453-8

Schmitz, M. L., Higgins, J. M. G., & Seibert, M. (2020). *Priming chromatin for segregation: functional roles of mitotic histone modifications*. *19*(6), 625–641. https://doi.org/10.1080/15384101.2020.1719585

Schwarzkopf, M., Knobeloch, K. P., Rohde, E., Hinderlich, S., Wiechens, N., Lucka, L., Horak, I., Reutter, W., & Horstkorte, R. (2002). Sialylation is essential for early development in mice. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(8), 5267–5270. https://doi.org/10.1073/PNAS.072066199/ASSET/B74529C9-C381-4A9F-876A-A91E89433CCE/ASSETS/GRAPHIC/PQ0720661005.JPEG

Sebastian Steinhauser, Nils Kurzawa, Roland Eils, Carl Herrmann, A comprehensive comparison of tools for differential ChIP-seq analysis, *Briefings in Bioinformatics*, Volume 17, Issue 6, November 2016, Pages 953–966, https://doi.org/10.1093/bib/bbv110

Sengoku, T., Yokoyama, S. (2011) Structural basis for histone H3 Lys 27 demethylation by UTX/KDM6A. https://doi.org/10.1101/gad.172296.111

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., & Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. *Cell*, *148*(3), 458–472. https://doi.org/10.1016/j.cell.2012.01.010

Sheffield, N. C., Bock, C. (2016). LOLA: enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor, *Bioinformatics*, Volume 32, Issue 4, Pages 587–589, https://doi.org/10.1093/bioinformatics/btv612

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

Skene, P., Henikoff, J. & Henikoff, S. Targeted *in situ* genome-wide profiling with high efficiency for low cell numbers. *Nat Protoc* **13**, 1006–1019 (2018). https://doi.org/10.1038/nprot.2018.015

Small, E. C., Maryanski, D. N., Rodriguez, K. L., Harvey, K. J., Keogh, M.-C., & Johnstone, A. L. (2021). Chromatin Immunoprecipitation (ChIP) to Study DNA-Protein Interactions. *Methods in Molecular Biology (Clifton, N.J.)*, *2261*, 323–343. https://doi.org/10.1007/978-1-0716-1186-9_20

Solomon, M. J., & Varshavsky, A. (1985). Formaldehyde-mediated DNA-protein crosslinking: a probe for in vivo chromatin structures. *Proceedings of the National Academy of Sciences of the United States of America*, *82*(19), 6470–6474. https://doi.org/10.1073/pnas.82.19.6470

Song, J., & Chen, K. C. (2015). Spectacle: fast chromatin state annotation using spectral learning. *Genome Biology*, *16*(1), 33. https://doi.org/10.1186/s13059-015-0598-0

Southall, S. M., Wong, P. S., Odho, Z., Roe, S. M., & Wilson, J. R. (2009). Structural basis for the requirement of additional factors for MLL1 SET domain activity and recognition of epigenetic marks. *Molecular Cell*, *33*(2), 181–191. https://doi.org/10.1016/J.MOLCEL.2008.12.029

Stragl, B. D., Allis, C. D. (2000) The language of covalent histone modifications. *Nature*, *403*, 41-45. https://www.nature.com/articles/47412

Stransky, S., Aguilan, J., Lachowicz, J., Madrid-Aliste, C., Nieves, E., & Sidoli, S. (2020). Mass Spectrometry to Study Chromatin Compaction. *Biology*, *9*(6), 140. https://doi.org/10.3390/biology9060140

Su, X., Ren, C., & Freitas, M. A. (2007). Mass spectrometry-based strategies for characterization of histones and their post-translational modifications. *Expert Review of Proteomics*, *4*(2), 211–225. https://doi.org/10.1586/14789450.4.2.211

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, *102*(43), 15545–15550. https://doi.org/10.1073/pnas.0506580102

Sullivan, B. A., & Karpen, G. H. (2004). Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nature Structural & Molecular Biology*, *11*(11), 1076–1083. https://doi.org/10.1038/nsmb845

Symmons, O., Uslu, V. V., Tsujimura, T., Ruf, S., Nassari, S., Schwarzer, W., Ettwiller, L., & Spitz, F. (2014). Functional and topological characteristics of mammalian regulatory domains. *Genome Research*, *24*(3), 390–400. https://doi.org/10.1101/gr.163519.113

Tada, K., Susumu, H., Sakuno, T., & Watanabe, Y. (2011). Condensin association with histone H2A shapes mitotic chromosomes. *Nature*, *474*(7352), 477–483. https://doi.org/10.1038/nature10179

Taudt, A., Nguyen, M. A., Heinig, M., Johannes, F., & Colomé-Tatché, M. (2016). *chromstaR: Tracking combinatorial chromatin state dynamics in space and time*. bioRxiv. https://doi.org/10.1101/038612

Taylor, J. H. (1960). Nucleic acid synthesis in relation to the cell division cycle. *Annals of the New York Academy of Sciences*, *90*, 409–421. https://doi.org/10.1111/j.1749-6632.1960.tb23259.x

Teves SS, An L, Hansen AS, Xie L, Darzacq X, Tjian R. A dynamic mode of mitotic bookmarking by transcription factors. Elife. 2016 Nov 19;5:e22280. doi: 10.7554/eLife.22280.

Teytelman, L., Thurtle, D. M., Rine, J., & van Oudenaarden, A. (2013). Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(46), 18602–18607. https://doi.org/10.1073/pnas.1316064110

Thorvaldsdóttir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, *14*(2), 178–192. https://doi.org/10.1093/BIB/BBS017

Tolstorukov, M. Y., Kharchenko, P. v, Goldman, J. A., Kingston, R. E., & Park, P. J. (2009). Comparative analysis of H2A.Z nucleosome organization in the human and yeast genomes. *Genome Research*, *19*(6), 967–977. https://doi.org/10.1101/gr.084830.108

Turner BM. Acetylation and deacetylation of histone H4 continue through metaphase with depletion of more-acetylated isoforms and altered site usage. Exp Cell Res. 1989 May;182(1):206-14. doi: 10.1016/0014-4827(89)90292-9.

Valls, E., Sánchez-Molina, S., & Martínez-Balbás, M. A. (2005). Role of histone modifications in marking and activating genes through mitosis. *The Journal of Biological Chemistry*, *280*(52), 42592–42600. https://doi.org/10.1074/jbc.M507407200

Van Galen, P., Viny, A. D., Ram, O., Cross, M. B., Levine, R. L., Bernstein, B. E. (2016) A Multiplexed System for Quantitative Comparisons of Chromatin Landscapes. *Molecular Cell: Technology, 61(1),* 170-180. https://doi.org/10.1016/j.molcel.2015.11.003

van Hooser, A., Goodrich, D. W., Allis, C. D., Brinkley, B. R., & Mancini, M. A. (1998). Histone H3 phosphorylation is required for the initiation, but not maintenance, of mammalian chromosome condensation. *Journal of Cell Science*, *111 (Pt 23)*, 3497–3506. https://doi.org/10.1242/jcs.111.23.3497

Varier, R. A., Outchkourov, N. S., de Graaf, P., van Schaik, F. M. A., Ensing, H. J. L., Wang, F., Higgins, J. M. G., Kops, G. J. P. L., & Timmers, H. M. (2010). A phospho/methyl switch at histone H3 regulates TFIID association with mitotic chromosomes. *The EMBO Journal*, *29*(23), 3967–3978. https://doi.org/10.1038/EMBOJ.2010.261

Vassilev, L. T. (2006). Cell cycle synchronization at the G2/M phase border by reversible inhibition of CDK1. *Cell Cycle (Georgetown, Tex.)*, *5*(22), 2555–2556. https://doi.org/10.4161/CC.5.22.3463

Walczak, C. E., Cai, S., & Khodjakov, A. (2010). Mechanisms of chromosome behaviour during mitosis. *Nature Reviews. Molecular Cell Biology*, *11*(2), 91–102. https://doi.org/10.1038/nrm2832

Walter, W., Clynes, D., Tang, Y., Marmorstein, R., Mellor, J., & Berger, S. L. (2008). 14-3-3 Interaction with Histone H3 Involves a Dual Modification Pattern of Phosphoacetylation. *Molecular and Cellular Biology*, *28*(8), 2840–2849. https://doi.org/10.1128/MCB.01457-07

Wang F, Ulyanova NP, Daum JR, Patnaik D, Kateneva AV, Gorbsky GJ, Higgins JM. Haspin inhibitors reveal centromeric functions of Aurora B in chromosome segregation. J Cell Biol. 2012 Oct 15;199(2):251-68. doi: 10.1083/jcb.201205106.

Wang, F., & Higgins, J. M. G. (2013). Histone modifications and mitosis: countermarks, landmarks, and bookmarks. *Trends in Cell Biology*, *23*(4), 175–184. https://doi.org/10.1016/j.tcb.2012.11.005

Wang, F., Dai, J., Daum, J. R., Niedzialkowska, E., Banerjee, B., Stukenberg, P. T., Gorbsky, G. J., & Higgins, J. M. G. (2010). Histone H3 Thr-3 phosphorylation by Haspin positions Aurora B at centromeres in mitosis. *Science (New York, N.Y.)*, *330*(6001), 231–235. https://doi.org/10.1126/science.1189435

Wang, F., Ulyanova, N. P., van der Waal, M. S., Patnaik, D., Lens, S. M. A., & Higgins, J. M. G. (2011). A positive feedback loop involving Haspin and Aurora B promotes CPC accumulation at centromeres in mitosis. *Current Biology: CB*, *21*(12), 1061–1069. https://doi.org/10.1016/j.cub.2011.05.016

Wang, J., Vasaikar, S., Shi, Z., Greer, M., & Zhang, B. (2017). WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Research*, *45*(W1), W130–W137. https://doi.org/10.1093/NAR/GKX356

Wang, Y. C., Stein, J. W., Lynch, C. L., Tran, H. T., Lee, C. Y., Coleman, R., Hatch, A., Antontsev, V. G., Chy, H. S., O'Brien, C. M., Murthy, S. K., Laslett, A. L., Peterson, S. E., & Loring, J. F. (2015). Glycosyltransferase ST6GAL1 contributes to the regulation of pluripotency in human pluripotent stem cells. *Scientific Reports 2015 5:1*, *5*(1), 1–13. https://doi.org/10.1038/srep13317

Wei, Y., Mizzen, C. A., Cook, R. G., Gorovsky, M. A., & Allis, C. D. (1998). Phosphorylation of histone H3 at serine 10 is correlated with chromosome condensation during mitosis and meiosis in Tetrahymena. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(13), 7480–7484. https://doi.org/10.1073/pnas.95.13.7480

Wei, Y., Yu, L., Bowen, J., Gorovsky, M. A., & Allis, C. D. (1999). Phosphorylation of histone H3 is required for proper chromosome condensation and segregation. *Cell*, *97*(1), 99–109. https://doi.org/10.1016/s0092-8674(00)80718-7

Weller, M. G. (2018). Ten Basic Rules of Antibody Validation. *Analytical Chemistry Insights*. https://journals.sagepub.com/doi/10.1177/1177390118757462

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. *Springer Cham*, *2nd Edition.* https://doi.org/10.1007/978-3-319-24277-4

Wickham, H., Vaughan, D., Girlich, M., Ushey, K., Posit Software. *Tidy Messy Data [R package tidyr version 1.3.1]*. (2024). https://doi.org/10.32614/CRAN.PACKAGE.TIDYR

Wieder, C., Frainay, C., Poupin, N., Rodríguez-Mier, P., Vinson, F., Cooke, J., Lai, R. P. J. J., Bundy, J. G., Jourdan, F., & Ebbels, T. (2021). *Pathway analysis in metabolomics: Recommendations for the use of over-representation analysis*. *17*(9), e1009105. https://pubmed.ncbi.nlm.nih.gov/34492007/

Winter, S., Simboeck, E., Fischle, W., Zupkovitz, G., Dohnal, I., Mechtler, K., Ammerer, G., & Seiser, C. (2008). 14-3-3 Proteins recognize a histone code at histone H3 and are required for transcriptional activation. *The EMBO Journal*, *27*(1), 88–99. https://doi.org/10.1038/sj.emboj.7601954

Wissink, E. M., Vihervaara, A., Tippens, N. D., & Lis, J. T. (2019). Nascent RNA Analyses: Tracking Transcription and Its Regulation. *Nature Reviews. Genetics*, *20*(12), 705–723. https://doi.org/10.1038/s41576-019-0159-6

Woodcock CL, Dimitrov S. Higher-order structure of chromatin and chromosomes. Curr Opin Genet Dev. 2001 Apr;11(2):130-5. doi: 10.1016/s0959-437x(00)00169-6. PMID: 11250134.

Yamagishi, Y., Honda, T., Tanno, Y., & Watanabe, Y. (2010). Two histone marks establish the inner centromere and chromosome bi-orientation. *Science (New York, N.Y.)*, *330*(6001). https://doi.org/10.1126/science.1194498

Yates, A. D., Achuthan, P., Akanni, W., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Azov, A. G., Bennett, R., Bhai, J., Billis, K., Boddu, S., Marugán, J. C., Cummins, C., Davidson, C., Dodiya, K., Fatima, R., Gall, A., … Flicek, P. (2020). Ensembl 2020. *Nucleic Acids Research*, *48*(D1), D682–D688. https://doi.org/10.1093/NAR/GKZ966

Yung, P. Y. K., Stuetzer, A., Fischle, W., Martinez, A.-M., & Cavalli, G. (2015). Histone H3 Serine 28 Is Essential for Efficient Polycomb-Mediated Gene Repression in Drosophila. *Cell Reports*, *11*(9), 1437–1445. https://doi.org/10.1016/j.celrep.2015.04.055

Zacher, B., Michel, M., Schwalb, B., Cramer, P., Tresch, A., & Gagneur, J. (2017). Accurate Promoter and Enhancer Identification in 127 ENCODE and Roadmap Epigenomics Cell Types and Tissues by GenoSTAN. *PLOS ONE,* https://doi.org/10.1371/JOURNAL.PONE.0169249

Zakrzewski, W., Dobrzyński, M., Szymonowicz, M., & Rybak, Z. (2019). Stem cells: Past, present, and future. *Stem Cell Research and Therapy*, *10*(1), 1–22. https://doi.org/10.1186/S13287-019-1165-5/FIGURES/8

Zhang, B., Kirov, S., & Snoddy, J. (2005). WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Research*, *33*(Web Server issue). https://doi.org/10.1093/NAR/GKI475

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

Zhang, B., Zheng, H., Huang, B., Li, W., Xiang, Y., Peng, X., Ming, J., Wu, X., Zhang, Y., Xu, Q. and Liu, W., 2016. Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature*, *537*(7621), pp.553-557.

Zhiteneva, A., Bonfiglio, J. J., Makarov, A., Colby, T., Vagnarelli, P., Schirmer, E. C., Matic, I., & Earnshaw, W. C. (2017). Mitotic post-translational modifications of histones promote chromatin compaction in vitro. *Open Biology*, *7*(9). https://doi.org/10.1098/RSOB.170076

Zhou, Y., Sun, Y., Huang, D., & Li, M. J. (2020). epiCOLOC: Integrating Large-Scale and Context-Dependent Epigenomics Features for Comprehensive Colocalization Analysis. *Frontiers in Genetics*, *11*, 53. https://doi.org/10.3389/fgene.2020.00053