INTERPHASE HISTONE H3 SERINE 10 PHOSPHORYLATION IN MOUSE EMBRYONIC STEM CELLS

by

Carol Chia-Lu Chen

B.Sc., The University of British Columbia, 2009

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Medical Genetics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2017

ABSTRACT

Phosphorylation of histone H3 at serine 10 (H3S10ph) has been observed in two paradoxical contexts, depending on the stage of the cell cycle – it is a hallmark of condensed mitotic chromatin, but is also found at specific mitogen-inducible genic promoters and distal enhancers. In the work presented in this thesis, I derived a mouse embryonic stem cell (ESC) line harbouring an endogenous fluorescent cell cycle reporter (FUCCI) in combination with next-generation sequencing to comprehensively map H3S10ph at distinct stages of the mammalian cell cycle and examined the crosstalk between this mark and the repressive mark, H3K9me. I found that H3S10ph broadly demarcates gene-rich, early-replicating euchromatic regions in G1, marking up to 30% of the ESC genome. Reminiscent of H3S10ph deposited by JIL-1 kinase in *Drosophila*, interphase H3S10ph pervasively marks gene-dense regions to prevent the accumulation of H3K9me2 at actively transcribed genes in ESCs. In contrast to H3S10ph at euchromatin, mitotic phosphorylation mediated by Aurora kinase is also detectable in ESCs in G1 at young endogenous retroviruses (ERVs), specifically in combination with H3K9me3. Finally, to identify the kinases responsible for interphase H3S10ph, I generated knockout (KO) ESC lines of serine kinases homologous to JIL-1, including *Msk1/2* and *Rsk1/2/3/4*. *Msk2*$^{-/-}$ ESCs showed the greatest loss of interphase H3S10ph at active promoter/enhancers. Furthermore, known H3K9me3 repressed targets, including imprinted genes, young ERVs and germline genes, were downregulated in MSK2-deficient cells. Taken together, this work revealed that H3S10ph plays a previously unappreciated role in interphase chromatin architecture and facilitates appropriate genic transcription by countering the repressive effects exerted by H3K9 methyltransferases.

LAY SUMMARY

In human cells, DNA strands are bound to structural proteins called histones to form chromatin. One mechanism for regulating the activity of genes is the chemical modification of such histones. Phosphorylation of the histone H3 is a hallmark of highly condensed chromosomes in actively dividing cells (mitosis), but has also been associated with loosely packaged chromatin at expressed genes in cells not undergoing division. To understand the role of this enigmatic mark in mammalian cells, I used a fluorescent cell cycle reporter coupled with a method called high-throughput sequencing to characterize phosphorylation of histone H3 during the cell cycle. I found that this phosphorylation is surprisingly pervasive in non-dividing cells, marking up to 30% of the genome, and implicate this conserved mark in maintaining chromatin structure at genomic regions that are rich in genes and are preferentially copied during DNA replication earlier than genomic regions that are gene poor.

PREFACE

I was the lead investigator responsible for designing and carrying out the experiments presented, and conducting the data analysis for all the research presented in this dissertation. Where highly specialized materials, equipment, or expertise was necessary, I collaborated with others to generate the relevant data. Dr. Julie Brind'Amour performed the Illumina sequencing library construction of H3K9me3 and H3K9me3S10ph ChIP-Seq in ESCs described in Chapter 3. Dr. Mohammad Karimi generated the bioinformatics analysis of transposable elements in Chapters 3 and 4. Dr. Yoichi Shinkai provided the *G9a*$^{-/-}$ and *Glp*$^{-/-}$ ESC lines used in Chapter 4. Dr. Hiroshi Kimura provided the monoclonal H3S10ph antibody used in Chapter 4, 5 and 6.

Collaborators:

Dr. Mohammad M. Karimi

Qatar Biomedical Research Institute,

Hamad Bin Khalifa University,

Doha, Qatar

Dr. Hiroshi Kimura

Cell Biology Center,

Institute of Innovative Research, Tokyo Institute of Technology

Yokohama, Japan

Chapter 3 is a submitted manuscript. Carol CL Chen, Preeti Goyal, Mohammad M. Karimi, Marie H. Abildgaard, Hiroshi Kimura, and Matthew C. Lorincz. **H3S10ph broadly marks early-replicating domains in interphase ESCs and shows reciprocal antagonism with H3K9me2**. I performed all major experiments described in this study, including generation of ESC transgenic lines (except *G9a*$^{-/-}$ and *Glp*$^{-/-}$), chromatin immunoprecipitation (ChIP), sequencing library construction and bioinformatics analyses and I wrote the paper with input from Dr. Matthew Lorincz. Preeti Goyal generated the initial H3K9me2 ChIP dataset, and Dr. Mohammad Karimi contributed to bioinformatics analysis.

Several of the datasets analyzed in Chapter 4 were published previously. For these datasets, I performed the ChIP and constructed the sequencing library. The H3K9me3 ChIP-seq dataset was used as the "gold-standard" control data set reported in: Julie Brind'Amour, Sheng Liu, Matthew Hudson, Carol Chen, Mohammad M Karimi, and Matthew C Lorincz (2015) **An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations.** *Nature Communications*, 6 p. 6033. The HP1β & H3K9me3 ChIP-seq data described was published in Kyoko Hiragami-Hamada, Szabolcs Soeroes, Miroslav Nikolov, Bryan Wilkins, Sarah Kreuz, Carol Chen, Inti A De La Rosa-Velázquez, Hans Michael Zenn, Nils Kost, Wiebke Pohl, Aleksandar Chernev, Dirk Schwarzer, Thomas Jenuwein, Matthew Lorincz, Bastian Zimmermann, Peter Jomo Walla, Heinz Neumann, Tuncay Baubec, Henning Urlaub, and Wolfgang Fischle (2016). **Dynamic and flexible H3K9me3 bridging via HP1β dimerization establishes a plastic state of condensed chromatin.** *Nature Communications*, 7 p. 11310.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

2i	GSK3β and MEK1/2 inhibitors
A	Adenine
Ab	Antibody
ADD	ATRX-DNMT3-DNMT3L
AGC	Protein Kinase A, G, and C
AP-1	Activator Protein 1
ATF1	Activation Transcription Factor 1
ATRX	α-thalassemia mental retardation X-linked
AURKB	Aurora B kinase
BLAT	Basic local alignment search tool
bp	Base pair
C	Cytosine
C-terminus	Carboxyl terminus
CaMK	Ca2+/calmodulin-dependent protein kinase
Cas9	CRISPR associated protein 9
CDYL1	Chromodomain Y-like 1
CHD4	Chromodomain helicase DNA binding protein 4
ChIP	Chromatin immunoprecipitation
Chr	Chromosome
CLS	Coffin-Lowry Syndrome
CMV	Cytomegalovirus

CPC	Chromosome Passenger Complex
CpG	Cytosine-phosphate-guanine
CREB	cAMP response element binding transcription factor
CRISPR	Clustered regularly interspaced short palindromic repeats
CTCF	CCCTC-binding factor
DAPI	4',6-diamidino-2-phenylindole
DHS	DNase hypersensitivity sites
DKO	Double knockout
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNMT1	DNA methyltransferase 1
DNMT3	DNA methyltransferase 3
Dox	Doxycycline
DTT	Ditiothreitol
E	Embryonic day post-coitum
E(var)	Enhancer of variegation
E2F	E2 Factor
E2F6	E2 Factor 6
EB	Embryoid body
EDTA	Ethylenediaminetetracetic acid
EGF	Epidermal growth factor
ERK	Extracellular signal-regulated kinase

ERV	Endogenous retroviral element
ESC	Embryonic stem cell
ETn	Early transposon
FACS	Fluorescent-activated cell sorting
FGF	Fibroblast growth factor
FGFR	Fibroblast growth factor receptor
FKBP6	FK506 binding protein 6
FUCCI	Fluorescent ubiquitin-mediated cell cycle indicator
G	Guanine
G1	Gap phase 1
G2	Gap phase 2
G4	Guanine quadruduplex
GATA6	GATA binding protein 6
GLP	G9a like protein 1
GO	Gene Ontology
GREAT	Genomic Regions Enrichment of Annotations Tool
h	Hour
H1	Histone H1
H2A	Histone H2A
H2A.X	Histone H2A variant X
H2B	Histone H2B
H3	Histone H3
H3.2	Histone H3 variant H3.2

H3.3	Histone H3 variant H3.3
H3K14ac	Histone H3 lysine 14 acetylation
H3K27ac	Histone H3 lysine 27 acetylation
H3K27me3	Histone H3 lysine 27 trimethylation
H3K36me3	Histone H3 lysine 36 trimethylation
H3K4me1	Histone H3 lysine 4 monomethylation
H3K4me3	Histone H3 lysine 4 trimethylation
H3K9	Histone H3 lysine 9
H3K9ac	Histone H3 lysine 9 acetylation
H3K9me	Histone H3 lysine 9 methylation
H3K9me2	Histone H3 lysine 9 dimethylation
H3K9me3	Histone H3 lysine 9 trimethylation
H3K9me3S10ph	Histone H3 lysine 9 trimethylation serine 10 phosphorylation
H3S10	Histone H3 serine 10
H3S10A	Histone H3 serine 10 to alanine mutation
H3S10ph	Histone H3 serine 10 phosphorylation
H3S28	Histone H3 serine 28
H4	Histone H4
H4K16	Histone H4 lysine 16
H4K20	Histone H4 lysine 20
HA	Hemagglutinin
HCl	Hydrochloric acid
HDAC	Histone deacetylase

HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
Hesp	Hesperadin
HMG-14	High mobility group protein 14
hnRNP K	Heterogeneous nuclear ribonucleoprotein K
HotSHOT	Hot sodium hydroxide and Tris
HP1	Heterochromatin Protein 1
IAP	Intracisternal A-type particle
ICM	Inner cell mass
IgG	Immunoglobulin G
INCENP	Inner Centromere Protein Antigens
IP	Immunoprecipitation
IPL1	Increase in ploidy protein 1
JMJD2A	Jumoji domain containing 2A
KAP1	KRAB Associated Protein 1
kb	kilobase
KDI	Kinase domain I
KDII	Kinase domain II
KMT	Lysine methyltransferase
KO	Knockout
KRAB-ZFPs	Krüppel-associated box zinc finger protein
LAD	Lamin associated domain
LIF	Leukemia inhibitory factor
LINE	Long interspersed nuclear element

LSD1	Lysine specific demethylase 1
LTR	Long terminal repeat
M	Mitosis
MAPK	Mitogen-activated protein kinases
Mb	Megabase
MEF	Mouse embryonic fibroblast
MEK1	Mitogen activated protein kinase kinase 1
MEME	Multiple em for motif ellicitation
mIgG	Mouse immunoglobulin G
mKO	Monomeric kusabira orange
MNase	Micrococcal nuclease
MOF	Male absent on the first
MPP8	M-phase phosphoprotein 8
MSK	Mitogen and stress-activated protein kinase
MT	Mouse transposon
N-terminus	Amino terminus
NChIP	Native chromatin immunoprecipitation
NFκB	Nuclear factor kappa B
NLS	Nuclear localization signal
NP-40	Nonyl phenoxypolyethoxylethanol
NRF1	Nuclear respiratory factor 1
P-TEFb	Positive transcription elongation factor b
PBS	Phosphate-buffered saline

PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PDK1	Phosphoinositide-dependent kinase 1
PEV	Position effect variegation
PGC	Primordial germ cell
PhosStop	Phosphatase inhibitor cocktail
PIC	Protease inhibitor cocktail
piRNA	Piwi-interacting RNA
PMSF	Phenylmethylsufonyl fluoride
Pol II	RNA polymerase II
PP1	Protein phosphatase 1
PrE	Primitive endoderm
PTM	Post-translational modification
PxVxL	Proline-x-Valine-x-Leucine
qPCR	Quantitative polymerase chain reaction
rDNA	Ribosomal DNA
rIgG	Rabbit immunoglobulin G
RLTR	Repeat long terminal repeat
RNA	Ribonucleotide acid
RPKM	Reads per kilobase mapped
RSK	Ribosomal s6 kinase
RT	Replication timing
S	Synthesis phase

S10A	Serine 10 to Alanine mutation
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Seq	Sequencing
SET	Suppressor of variegation 3-9, Enhancer of zeste, Trithorax
sgRNA	Short guide RNA
SRA	SET and RING finger associated domain
STAT	Signal transducer and activator of transcription
Su(var)	Suppressor of variegation
T	Thymine
TAD	Topologically associating domain
TE	Trophectoderm
TKO	Triple knockout
TSEA	Tissue specific expression analysis
TSS	Transcription start site
TTD	Tandem Tudor Domain
TTR	Replication Timing Transition Regions
TTS	Transcription termination site
UHRF1	Ubiquitin like with PHD and ring finger domains 1
UV	Ultraviolet ray
UV-B	Ultraviolet ray B
WT	Wildtype
XMChIP	Crosslink MNase chromatin immunoprecipitation

YFP	Yellow fluorescent protein

ACKNOWLEDGEMENTS

I would like to express my foremost gratitude to my supervisor, Dr. Matthew Lorincz, for his patience, guidance, and support. Thank you for entrusting a fascinating research project to me. Your eternal passion for science is truly unmatched, and was often the source of motivation.

Thanks to the past and present members of the Lorincz Lab for support and stimulating discussions, both in and out of the lab. Specifically, Julie Brind'Amour, Preeti Goyal, Peter Thompson, Irina Maksakova, Aaron Bogutz, Julien Richard Albert, and Kris Jensen, all of whom were instrumental in my research training and to this project. I would also like to extend my gratitude to my supervisory committee, Dr. LeAnn Howe, Dr. Louis Lefebvre, and Dr. Hugh Brock for their time, wisdom, and feedback over the years. Thank you for teaching me the importance of brevity. To Dr. Dixie Mager and Cheryl Bishop, who came to my rescue numerous times during this PhD, thank you for your dedication to trainees. I'd also like to mention the collaborators who provided their time and reagents that make this project possible, especially Dr. Hiroshi Kimura, who provided the antibodies used in this project. I truly could not have completed my thesis work without your generosity.

To my Vancouver "framily", including those in the Lorincz Lab, M & K Elliott, J. Wilson, A. Chang, N. Afacan, G. Tharmarajah, and many others - thank you for providing the necessary counterbalance to work, and for indulging me in all the occasional instances of gluttony and insobriety. To Carlin, for distractions and encouragements at the very end. Lastly, my family – thank you for your love and patience over the years.

DEDICATION

For dad.

1. INTRODUCTION

1.1 Eukaryotic genome packaging

The genome - the summation of an organism's genetic material - encompasses essential information necessary for development, maintenance, and reproduction of an organism. A massive amount of genetic information is encoded within one cell, ~3 billion base-pairs of DNA in the case of the human and mouse haploid genomes (Lander et al. 2001; Mouse Genome Sequencing Consortium et al. 2002). Stretched end-to-end, the human genome would theoretically reach a length of 2 m (Smith et al. 1996). To fit into the confines of a 10 μm nucleus, the eukaryotic genome is packaged as a chromatin fiber – an ordered macromolecule made up of predominantly DNA and histones. The basic building block of the chromatin fiber is the nucleosome core particle: 146 bp of DNA wrapped around an octamer of core histone proteins, 2 copies each of histones H2A, H2B, H3 and H4 (Luger et al. 1997). Nucleosome structure not only facilitates physical compaction of the genome, but is also dynamically regulated for essential cellular processes including transcription, replication, repair, and division.

Based upon the cytological appearance of DNA in the nucleus, regions of the genome are traditionally classified in one of two categories based on the compaction state of chromatin– loose and densely compacted regions of DNA were referred to as euchromatin and heterochromatin, respectively (Angell and Jacobs 1975). Visualized linearly on condensed mitotic chromosomes, euchromatin and heterochromatin compartments are classically also used to define cytogenetic bands. Intriguingly, cytogenetic banding roughly corresponds to gene

density and isochores – with gene-dense C:G rich regions and gene-poor A:T rich regions in euchromatin and heterochromatin bands, respectively. As such, euchromatin/heterochromatin have sometimes been used to broadly describe transcriptionally active vs. inactive chromatin compartments; however, Bickmore and colleagues have shown conclusively that euchromatin, as isolated by sucrose gradient fractionation, reflects gene density and not transcription activity (Gilbert et al. 2004)– suggesting the euchromatin state is a higher order structure maintained independently of transcription. Fundamentally, heterochromatin and euchromatin represent non-random organization of the genome in terms of both sequence and chromatin composition.

Pontecorvo was the first to observe that chromocentres of non-homologous polytene chromosomes coalesce together, reminiscent of somatic pairing between sister chromatids (Pontecorvo 1944). This observation led Pontecorvo to postulate that heterochromatin structure arises because of duplicated/repetitive sequences. Indeed, regions that are constitutively packaged as heterochromatin, such as centromeric and telomeric repeats, are composed of large tandem arrays of simple repeats. Repetitive sequences in total make up ~50% of mammalian genomes (Treangen and Salzberg 2012), consisting of both tandem satellite repeats and transposable element derived elements. Such repetitive DNA has been shown to cause genomic instability due to polymerase slippage, illegitimate recombination, and retrotransposition (Castel et al. 2010; O'Sullivan and Karlseder 2010; Slotkin and Martienssen 2007). The subsequent observation that heterochromatin can be ectopically induced upon insertion of tandem array transgenes (Henikoff 1998; Garrick et al. 1998) reinforces Pontecorvo's notion that heterochromatin serves as a defensive mechanism against aberrant copy number amplification in the genome.

1.2 Endogenous retroviral elements in the murine genome

Transposable elements, or transposons, describes mobile DNA segments within the genome. Compared to protein-coding regions, which account for 1.5% of the murine genome, transposons make up approximately half of the genome in humans and mice (Slotkin and Martienssen 2007). Depending on the mechanism of transposition, transposons are classified as DNA transposons or retrotransposons, with the latter requiring reverse transcription as an intermediate step in the viral life cycle. Retrotransposons are further subdivided based on the absence or presence of long terminal repeats (LTR). Non-LTR retrotransposons, such as the long and short interspersed nuclear elements (LINEs and SINEs), encompass 27.4% of the mouse genome and represent the single largest fraction of interspersed repeats in human and mice. LTR retrotransposons, also called endogenous retroviral elements (ERVs), arise as a result of exogenous retroviral integration in the germline (Stocking and Kozak 2008), and constitute over 10% of the mouse genome. In the mouse, specifically, a number of ERV families are transcriptionally active, with a small number being competent for retrotransposition (Stocking and Kozak 2008).

Structurally, full length ERVs possess identical long terminal repeats (LTRs) flanking open reading frames (ORFs) encoding viral proteins, including *gag* (group specific retroviral antigen), *pol* (reverse transcriptase), and *env* (envelope protein). ERVs are classified based on phylogenetic clustering of the reverse transcriptase and their homology to exogenous retroviruses (Stocking and Kozak 2008). Class I ERVs, which constitute 0.7% of the genome, are type-C viruses resembling exogenous gamma- and episilon-retroviruses. The class I ERV murine leukemia virus (MLV) entered the murine genome recently (<1.5 Mya) and has been extensively

studied due to its proto-oncogenic activation of *Wnt1* and *Notch1* leading to leukemia/lymphoma (Kassiotis 2014; Stoye and Coffin 1988). Class II ERVs, which comprise 3.14% of the mouse genome, are type-B and D retroviruses resembling exogenous beta- and alpha-retroviruses (Stocking and Kozak 2008). Compared to the other two classes, more class II subfamilies are capable of viral particle production (McCarthy and McDonald 2004). Well-studied members of class II includes the intracisternal A particles superfamily, which is responsible for variegating expression of the *Agouti* coat-colour gene (Morgan et al. 1999). Finally, Class III ERVs, which make up 5.4% of the mouse genome (Bénit et al. 1999), are the most ancient retroviruses, predating speciation of placental mammals (> 70 million mya). Class III ERVs include active subfamilies such as MaLR and MuERV-Ls, and a subset are co-opted as alternative promoters for host genes (Maksakova et al. 2006; Macfarlan et al. 2012; Peaston et al. 2004).

Active transposons are highly mutagenic, and are often referred to as "parasitic" elements due to the deleterious effect exerted on the host genome (Slotkin and Martienssen 2007). Retrotransposition of ERVs is responsible for up to 10% of spontaneous mutations in the mouse germline (Maksakova et al. 2006). Furthermore, infectious copies of MMTV and MLV have been shown to cause leukemia through activation of oncogenes, or disruption of tumour-suppressor genes (Kassiotis 2014). As such, the host employs various defence mechanisms against newly integrated ERVs, including epigenetic silencing and viral protein inhibition to repress proviral amplification (Mager and Stoye 2015). Over time, most ERV-derived sequences accumulate deleterious mutations under pressures of negative selection, and eventually become benign vestiges in the genome. Indeed, the vast majority of ERVs in the murine genome exist as

solo LTRs, resulting from recombination between the 5' and 3' LTRs of full-length elements (Mager and Stoye 2015).

Such solitary LTRs contains promoter sequences and various transcription factor binding sites, and have the potential to act as *cis* regulatory elements for host genes (Jacques et al. 2013). In a seminal study by Peaston et al., it was shown that expression of various ERV families are developmentally regulated in the mouse, and class III MT elements are used as alternative promoters in the oocyte to drive transcription of a number of host genes (Peaston et al. 2004). Since then, work by others also supports the idea that ERV LTRs can be co-opted by the host as *cis*-acting elements to deploy tissue-specific transcription programs (Thompson et al. 2016). The emerging paradigm reinforces the proposed role of transposons, first put forth by McClintock after their discovery (McClintock 1963), as a functional component of complex gene regulatory networks in multicellular organisms.

1.3 Position effect variegation

In 1930, harnessing the mutagenic potential of X-rays, Herman Muller described a new class of mutant flies that exhibited "mottled" eyes in which patches of red facets are interspersed with patches of white facets (Muller 1930). As opposed to the *white* mutants, these mottled mutants exhibit "metastable" expression of the fully functional *white* gene in patches of cells, suggesting that the cause for the variegating-pattern of expression cannot be explained by genic mutation in *white*, and therefore, was epigenetic in nature. Subsequently, Schultz discovered that the variegating nature of the mutation was due to inappropriate juxtaposition of the *white* gene near pericentromeric heterochromatin, "with the point of breakage acting as a centre from which the

[inactivation] spreads" (Schultz 1936). The general phenomenon, now termed "position effect variegation" (PEV), was the first evidence that the chromatin environment in which a gene resides can exert influence over transcription capacity and phenotypic outcome.

PEV has since been demonstrated in other eukaryotes, including *Schizosaccharomyces pombe* and mammals, through studying the variable transcriptional output from randomly integrated reporter constructs (Allshire & Ekwall, 2015; Blewitt & Whitelaw, 2013). From such studies (Sun et al. 2004a), it was shown that integration site strongly dictates transcriptional capacity across eukaryotes. In mouse embryonic stem cells (ESCs), individual integrated reporters can exhibit over 1000-fold difference in basal transcriptional output (Akhtar et al. 2013). The local chromatin environment, i.e. integration into constitutive restrictive "closed" heterochromatin vs. permissive "open" euchromatin (partitioning the genome into ~1 megabase domains in ESCs), accounts for a 100-fold difference. The additional 10-fold difference in expression was attributed to immediate adjacency (~ 20 kilobase in ESCs) to an active gene/enhancer.

The extent of transgene variegation is also correlated with proximity to a transposable element in *Drosophila* (Sun et al. 2004b), suggesting that, similar to constitutive heterochromatin, transposons-induced heterochromatin can also propagate beyond the repetitive sequence into flanking regions. In rare cases, these types of variegation induced by transposable elements are transgenerationally inherited as metastable epialleles. Such epialleles was first documented in mouse, whereby an integrated copy of class II IAP ERV 10 kb upstream of the *agouti* gene (A^{vy}) leads to stochastic expression of the agouti protein in isogenic animals,

resulting in a continuous spectrum in yellow (*Agouti*) to black (wildtype) coat-colours (Morgan et al. 1999).

Subsequent mutagenesis screens on PEV reporter lines in both flies and mice have been fruitful in identifying dominant modifier genes, and establishing the critical role of histone post-translation modifications (PTMs) in mediating transcriptional variability (**Table 1.1**). In *Drosophila,* moodifiers of PEV were categorized in two phenotypic classes - Suppressor, or *Su(var),* mutants showed more muted variegation and more closely resembled wildtype, while Enhancer, or *E(var)* mutants showed more drastic silencing of *white*. A total of ~150 genes have been identified as modifiers in *Drosophila*, and most have since been shown to be either *bona fide* chromatin modifiers or readers, or components of metabolic pathways for relevant histone post-translational modifications (Elgin and Reuter 2013).

Table 1.1 Dominant genetic modifiers of position effect variegation in *Drosophila.*
* indicates significant divergence between the orthologs. (Elgin and Reuter 2013).

PEV Modifier	Mouse Ortholog(s)	Description
Su(var)4-20	*Kmt5b*	Histone H4K20 methyl transferase
Sam-S	*Mata2a, Mata1a*	S-adenosylmethionine synthetase
chm	*Kat7*	Chameau, MYST histone acetyltransferase
Su(var)205	*Cbx3, Cbx1*	Heterchromatin Protein 1 alpha
Su(var)2-HP2	*Rp1*, Rif1**	Heterochromatin Protein 2
Su(var)2-10	*Pias1*	Protein inhibitor of activated STAT (PIAS)
HDAC1	*Hdac1*	Histone deacetylase 1
SuUR	-	Suppressor of Under-Replication
JIL-1	*Rps6ka5*, Rps6ka4**	JIL-1 kinase
Su(var)3-3	*Kdm1a*	LSD1, lysine speicifc demethylase 1
jumu	*Foxn4*	jumeau, Fork head winged-helix protein
Pp1-87B	*Ppp1cc, Ppp1ca*	Protein phosphatase 1 alpha
Su(var)3-7	-	Zinc-finger protein, heterochromatin-associated
Su(var)3-9	*Suv39h1/2*	Histone H3K9 methyl transferase
mod	-	Modulo
Usp47	*Usp47*	Ubiqutin specific protease 47
Trl	*Zbtb49**	Trithorax-like, GAGA sequence binding factor
mod(mdg4)	*Bach2**	Modifier of mdg4
E2F1	*E2f3*	E2F transcription factor 1

1.4 Histone post-translational modifications

The nucleosome is heavily modified with covalent PTMs concentrated at the N-terminal tails of histones, which extend out from the octamer structure of the core nucleosome. There is an increasing catalog of histone PTMs including, but not limited to, acetylation, methylation, phosphorylation, ubiquitination, sumoylation, citrullination, and biotinylation (Bannister and Kouzarides 2011). While histone PTMs may directly alter DNA-histone interactions by changing the electrostatic charge of the core nucleosome, it is now generally believed that histone PTMs act as a "histone code", in which specific combinations of PTMs are recognized by chromatin-binding proteins with differing affinities to these marks. Once bound these factors exert essential DNA-templated functions, including transcription, repair, replication, recombination and segregation (Strahl and Allis 2000).

Methylation of histones occurs on lysine and arginine residues on H2A, H3 and H4 (**Fig. 1.1**). Both arginine and lysine methyltransferases (KMTs) utilize the co-substrate *S*-adenosyl methionine to catalyze methyl group transfer, and this multivalent modification is further classified into mono- di- and tri- depending on the number of methyl groups transferred to the ε-amino group of lysine, or the guanidine group arginine. Nearly all KMTs possess a highly-conserved SET (Su(var)3-9, Enhancer-of-zeste, Trithorax) catalytic domain, which confers substrate specificity by recognition of 2-6 amino acids flanking the relevant lysine (reviewed in Dillon, 2005). In mammals, lysine methylation is associated with either transcriptional activation or repression depending on the specific lysine residue modified - residues associated with transcriptional activity include H3K4 and H3K36, whereas "repressive" methylation is found on H3K9, H3K27 and H4K20. As methylation does not notably impact the charge of the modified

histone, this PTM is thought to exert its function through recruitment of methyl-lysine binding proteins.

Histone phosphorylation is known to occur on serine, threonine and tyrosine residues of core nucleosomes (**Fig. 1.1**), some histone variants, and the linker histone H1. Histone phosphorylation, a transient mark lasting between 10 and 30 minutes, is associated with mitosis/meiosis, mitogenic signalling, transcription and DNA damage repair (Rossetto et al. 2012). Unlike histone methyltransferases, histone kinases do not share a conserved catalytic domain, and many are capable of phosphorylating multiple residues across different core histones. While histone phosphorylation is known to be coupled to specific nuclear events, the function of histone phosphorylation is not well understood compared to histone methylation. The addition of a phosphate group imparts a negative charge and could theoretically loosen DNA-histone interactions - indeed phosphorylation of sites in the nucleosome core have been associated with DNA entry site unwrapping and nucleosome destabilization (Bowman and Poirier 2015).

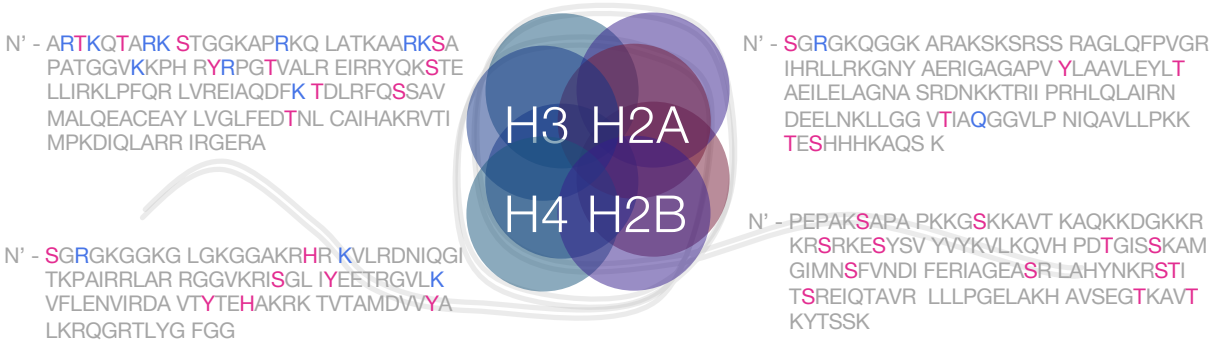


Figure 1.1 Methylation and phosphorylation sites on nucleosomal histones.
The core histones are heavily modified with post-translational modifications (PTMs), such as methylation and phosphorylation. Methylation of histones occur on lysine or arginine residues of histone H3, H4 and H2A and are

depicted in blue. While serine, threonine, tyrosine and histidine residues can be phosphorylated and are depicted in pink. In addition to the core nucleosome, phosphorylation of linker histone H1 and the H2A histone variant H2AX have also been reported.

The focus of this thesis work is on a pair of histone PTMs deposited on adjacent residues on the histone H3 N-terminus tail – namely methylation of H3K9 (H3K9me) and phosphorylation of H3S10 (H3S10ph). H3K9me2/3 is widely associated with transcriptional silencing, while H3S10ph is thought to promote transcription activity in interphase. The enzymes responsible for these marks in flies, *Su(var)3-9* and *Su(var)3-1/JIL-1*, were first identified from PEV screens; subsequent work has shown that together, H3K9 methylation and H3S10 phosphorylation influence heterochromatin and euchromatin structure in mammals as well, as described in detail below.

1.5 Histone H3 lysine 9 methylation

H3K9 can be mono-, di- (H3K9me2) or trimethylated (H3K9me3). Both of these marks are enriched at pericentromeric and telomeric heterochromatin in most eukaryotes, likely playing a role in suppressing recombination and silencing transcription, in addition to playing a structural role in the maintenance of genome integrity at centromeres and telomeres (Bickmore and van Steensel 2013). In mice, the five *bona fide* H3K9 methyltransferases: G9a, GLP, SUV39H1, SUV39H2, and SETDB1, all contain the conserved catalytic SET domain and the pre-/post-SET domains to facilitate efficient H3K9 methylation (Leung and Lorincz 2011).

G9a and GLP are paralogous proteins that form a functional heterodimeric complex to mono- or di-methylate H3K9, and also possesses Ankyrin repeats that act as reader modules for H3K9me1/2 to propagate repressive heterochromatin in neighboring nucleosomes (Tachibana et

al. 2002; 2008). G9a-dependent H3K9me2 is developmentally regulated, but is generally distributed as broad, megabase-sized domains in ESCs and somatic cells (Wen et al. 2009; Lienert et al. 2011). *G9a* or *Glp* KO mES cells are viable, but show severe loss of H3K9me2 in euchromatin regions, and derepression of late-replicating genes and class III endogenous retroviral elements (ERVs) (Tachibana et al. 2002; Yokochi et al. 2009; Maksakova et al. 2013). Deletion of either G9a or GLP in the mouse results in severe growth restrictions and early lethality at E9.5 (Tachibana et al. 2002; Tachibana 2005).

The paralogous SUV39H1 and SUV39H2 catalyze the di- and tri-methylation of H3K9 at constitutive heterochromatin and maintain transcriptional silencing of pericentromeric satellites (Aagaard et al. 2000). *Suv39h1/2* DKO mice are recovered below the expected Mendelian ratio and viable *Suv39h1/2* DKO mice show growth retardation at birth and are susceptible to tumour formation, likely due to chromosomal instability (Peters et al. 2001). *Suv39h1/2* DKO ES cells are viable, but show DNA hypomethylation and loss of H3K9me2/3 exclusively at pericentromeric regions (Lehnertz et al. 2003).

The final member of the Suv3-9 family, SETDB1, which possesses a signature bifurcated SET domain, is responsible for depositing H3K9me3 at euchromatin and confers transcriptional repression of some classes of young endogenous retrotransposons in the mouse (Matsui et al. 2010; Karimi et al. 2011; Liu et al. 2014; Dodge et al. 2004). In the context of both ERVs and imprinted genic loci, sequences are recognized by specific KRAB-ZFPs (Ecco et al. 2016), which in turn, recruit the KAP1/SETDB1 complex to deposit H3K9me3 and maintain transcriptional silencing (Rowe et al. 2010; Quenneville et al. 2011). SETDB1 ablation in ESCs

is cell lethal and results in derepression of ERVs, germline-specific genes and retroelement-driven chimeric transcripts prior to apoptosis (Matsui et al. 2010; Karimi et al. 2011). Deletion of SETDB1 *in vivo* results in peri-implantation lethality at E3.5 – 5.5, with surviving mutants showing a discernable ectoplacental cone but lacking the embryo proper at E6.5 (Dodge et al. 2004); these observations are in line with abnormal activation of extraembryonic trophoblast markers observed in SETDB1 knockdown ESCs (Yuan et al. 2009).

1.6 Heterochromatin Protein 1

Heterochromatin Protein 1 (HP1) was first identified as a protein associated with constitutive heterochromatin, and is best-known as a canonical reader of H3K9me2/3 (Lachner et al. 2001). In mammals, three distinct HP1 isoforms exist in the genome – HP1α (*Cbx5*), HP1β (*Cbx1*), and HP1γ (*Cbx3*). The isoforms are similar in domain structure and closely match the domain organization found in *Drosophila* HP1 paralogs – with a conserved N-terminal chromodomain and a C-terminal chromoshadow domain, which are separated by a variable hinge region. The chromodomain is homologous with other chromatin reader or writer proteins associated with transcription repression, including *Drosophila* Polycomb and SUV39H1 (Jacobs 2002).

In vitro, the chromodomains of HP1s show high affinity for a higher state of H3K9 methylation – predominantly di- or tri-methylation (Rothbart et al. 2012). In addition to reading out H3K9me2/3, HP1 is capable of oligomerization through the chromoshadow domain, and is proposed to bridge neighboring nucleosomes for chromatin compaction (Hiragami-Hamada et al. 2016). HP1 also complexes with a myriad of chromatin-modifying proteins via interaction in the PxVxL motif found in the chromoshadow domain – including SUV39H1, DNMT1/3a, and HP1

itself (Fritsch et al. 2010; Peters et al. 2001; Nozawa et al. 2010). The unique ability of HP1 to simultaneously recognize H3K9me3 heterochromatin and complex chromatin-modifying enzymes have led to the hypothesis that HP1 may function as a platform protein to assemble repressive machinery at H3K9me3-marked regions and to perpetuate linear spreading of heterochromatin. Indeed artificial tethering of an HP1α chromoshadow domain to a specific genomic region in ESCs is sufficient to establish a domain of H3K9me3, conclusively confirming that these readers actively reinforce heterochromatin structure (Hathaway et al. 2012). While mice lacking *HP1α* are viable with no overt phenotype, *HP1γ* KO mice show severe hypogonadism and *HP1β* KO mice die shortly after birth (Allan et al. 2012; Takada et al. 2011; Aucott et al. 2008).

1.7 Histone H3 serine 10 phosphorylation

1.7.1 Aurora B & mitotic H3S10ph

Phosphorylation of histone H3 is a classic hallmark of condensed mitotic and meiotic chromosomes in eukaryotes (**Fig. 1.2**) (Bradbury et al. 1973; Gurley et al. 1974). In mammals, Aurora B/C (homologous to *IPL1* in yeast) is the mitotic kinase responsible for phosphorylating Ser10 and Ser28 on histone H3 (Wei et al. 1998; Hsu et al. 2000). Disruption of Aurora kinase activity generally results in chromosome decondensation, segregation errors and cell cycle arrest (Van Hooser et al. 1998); however, Aurora B is an essential component of the Chromosome Passenger Complex (CPC) and its functions in mitosis extend far beyond phosphorylating H3S10, including facilitating chromosome-microtubule attachment, activation of the spindle check point, and regulation of cytokinesis (Carmena et al. 2012).

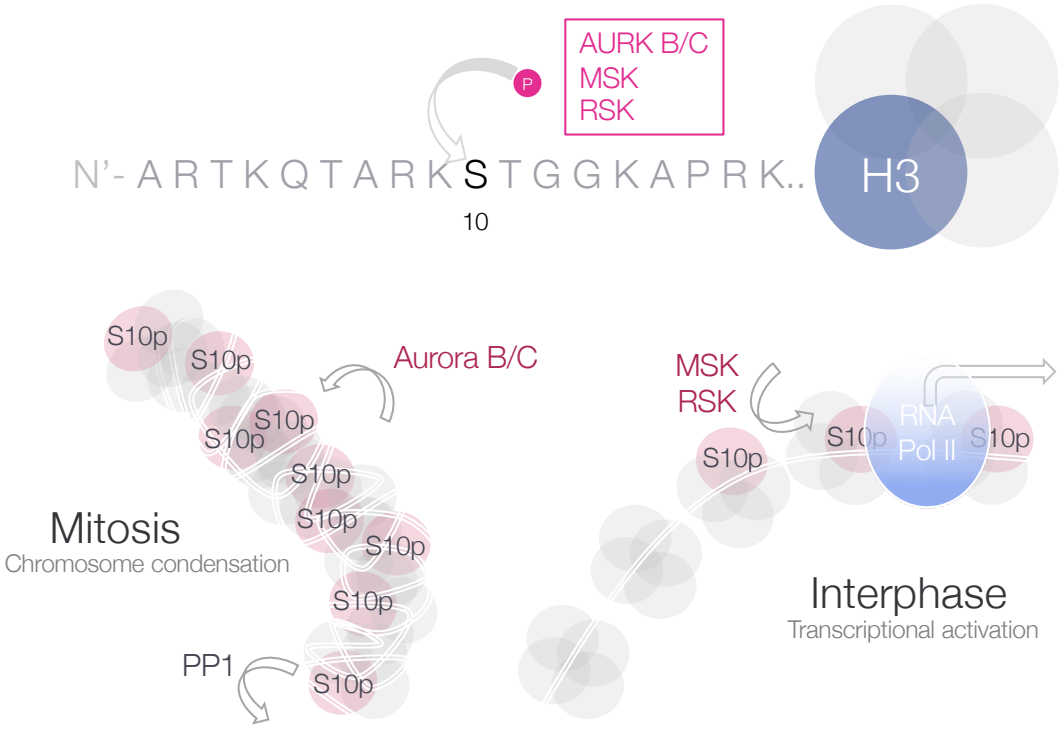


Figure 1.2 Paradoxical role of histone H3 serine 10 phosphorylation in mitosis and interphase.
In mitosis, Aurora B/C mediates genome-wide phosphorylation of histone H3 serine 10 during prophase, which is then dephosphorylated by the phosphatase PP1 at telophase. This process is essential for metaphase chromosome alignment and segregation. Conversely, H3S10ph has also been observed upon mitogenic or stress stimulation at promoter and enhancers of immediate early genes. This phosphorylation event, termed the "nucleosomal response", is mediated by MSKs & RSKs on a small fraction of nucleosomes to promote the release of paused RNA Pol II. Adapted from Bode et al. 2005.

The CPC complex is composed of a localization module (consisting of survivin and borealin) and a kinase module (Aurora B kinase), bridged by the protein INCENP on its N- and C-terminus, respectively. Aurora B kinase activity is initiated by INCENP binding and is fully potentiated by local density of CPC (Adams et al. 2001) – which may explain the tight temporal and spatial regulation of this highly promiscuous kinase. In interphase, the CPC is first recruited by Heterochromatin Protein 1, or HP1, via the PxVxL/I motif on INCENP to pericentromeric heterochromatin during late-S phase (Ainsztein et al. 1998). The temporal and spatial distribution of H3S10ph, as determined by microscopy, also begins with phosphorylation of pericentromeric

heterochromatin at late S, reaching completion to cover chromosome arms by prometaphase. This mark is removed by Protein Phosphatase 1 (PP1) at the end of telophase (Crosio et al. 2002). As with most cell cycle regulators, Aurora B/C is highly expressed in ESCs. This kinase is also amplified in human colorectal cancers (Bischoff et al. 1998), likely facilitating rapid cell cycling.

Despite its conservation, the function of H3S10ph during mitosis is controversial. Histone peptide injections into dividing cells indicates that H3S10ph is only necessary for the initiation of condensation, but not for maintenance of compact mitotic chromatin structure, *per se,* since dephosphorylated chromosomes remain condensed (Van Hooser et al. 1998). Conversely, inhibition of protein phosphatases induces abnormally condense chromosomes and premature mitotic exit into interphase (Guo et al. 1995). In yeast, the H3S10A mutant (which cannot be phosphorylated) shows neither mitotic nor meiotic abnormalities and exhibits identical generation time compared to wild-type, indicating that while the kinase is essential for proper cycling, H3S10ph is not essential in yeast (Hsu et al. 2000). It is worth noting, however, that the same mutation in *Tetrahymena* leads to a cell-lethal phenotypes indicative of mitotic defects, such as aneuploidy, chromosome lagging and mitotic arrest (Wei et al. 1999). In mammalian cells, H3S10ph has been reported to act as a molecular switch to eject HP1 proteins from the mitotic chromatin (Fischle et al. 2005) untethering the CPC for dynamic localization in mitosis (Nozawa et al. 2010).

1.7.2 Nucleosomal response: inducible H3S10ph

During interphase, H3S10ph is reportedly deposited on specific promoters as a part of downstream signal transduction following quiescent cell stimulation and is associated with activating transcription of inducible genes (**Fig. 1.2**) (Bode 2005). In contrast to the mitotic signal, inducible H3S10ph is thought to be limited to a small fraction of bulk nucleosomes (Barratt et al. 1994). Inducible H3S10ph is generally stimulated by extracellular ligands (e.g. phorbol esters, growth factors, UV), which initiate an intracellular kinase cascade in the cytoplasm, culminating in activation of immediate early genes in the nucleus (Lau and Nathans 1987). This induction of transcription is preceded by the rapid and transient phosphorylation of histone H3 at S10/S28 and transcription factors HMG-14 and CREB, described as the "nucleosomal response" (Mahadevan et al. 1991). Importantly, the nucleosomal response is not a consequence of transcription, as transcription inhibitors did not notably impact H3 phosphorylation following induction. Despite the fact that S10 and S28 share similar ARKS sequence motifs, inducible phosphorylation at S10 and S28 occurs on separate pools of nucleosomes (Khan et al. 2017), and shows distinct differences in kinetics following induction in activated macrophages (Josefowicz et al. 2016), suggesting the two phospho epitopes are not functionally equivalent.

In the context of mammalian transcriptional activation, numerous kinases involved in signal transduction have been shown to deposit H3S10ph at target promoters; including MSK1/2 (homologous to *JIL-1* in flies), IKKα and PIM1 (Bode 2005). At the human *FOSL1* promoter, for example, PIM1 kinase deposited H3S10ph upon serum stimulation triggers a cascade of chromatin modifications, including acetylation of H4K16 by MOF and recruitment of the

phospho-scaffolding protein 14-3-3, which together induce RNA polymerase II elongation (Zippo et al. 2009). Consistent with its role in promoting transcription initiation, mitogenic H3S10ph is only observed at promoter or enhancer regions of a handful of genes, but not within coding regions (Healy et al. 2012). Given that only a few genes are subjected to the nucleosomal response, it is unlikely that most transcription events require H3S10ph. Thus, inducible H3S10ph has been proposed to play a transient role in overriding existing repressive modifications at a subset of *cis* regulatory regions, thus permitting transient transcription of inducible genes (Sawicka and Seiser 2012).

1.7.3 JIL-1 and H3S10ph in Drosophila

Originally recovered from an expression screen as a candidate chromatin regulator in *Drosophila* (Jin et al. 1999), JIL-1 was later identified as an essential H3S10 kinase that associates with chromatin throughout the cell cycle (Wang et al. 2001). Indeed, JIL-1, which contains tandem kinase domains with intrinsic autophosphorylation activities, efficiently phosphorylates H3S10 *in vitro*. Using ChIP-chip, JIL-1 was subsequently found to localize to gene-dense regions forming blanket domains over gene bodies, regardless of transcriptional activity, but is especially enriched on the hypertranscribed X chromosome in male flies (Regnard et al. 2011).

Mutation studies indicate that in the absence of JIL-1 kinase activity, mitotic H3S10ph is unperturbed and heterochromatin marks, including H3K9me2 and HP1, spread ectopically (**Fig. 1.3**) (Zhang et al. 2006). Null mutations in *JIL-1* suppress or enhance the PEV phenotype, depending on the site of integration of the *white* reporter; this is probably due to uneven ectopic spreading of heterochromatin in the absence of H3S10ph (Wang et al. 2012). Intriguingly, the

lethality phenotype observed in *JIL-1* null mutants can be rescued via simultaneous loss-of-function in *Su(var)3-9*, suggesting that the aberrant gene expression observed in *JIL-1* hypomorphs is likely due to gross chromatin changes, rather than defects in transcription *per se* (Deng et al. 2008; Regnard et al. 2011). Furthermore, physical tethering of JIL-1 to a genomic location induces local chromosomal decondensation without enhancing transcriptional activity (Li et al. 2013). Taken together, these data suggest that JIL-1 acts upstream of Su(var)3-9 to reinforce euchromatin-heterochromatin boundaries by antagonizing heterochromatin propagation into gene-dense regions.

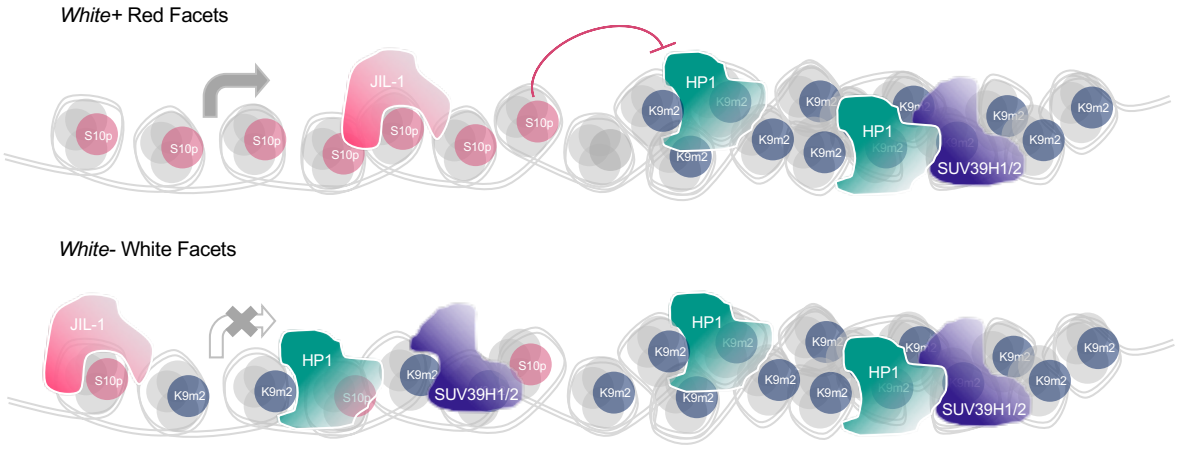


Figure 1.3 Model of methyl-lysine and phospho-serine conflict at the translocated *white* locus in *Drosophila*. Position effect variegation (PEV) in *Drosophila* describes the mosaic pattern of red and white patches of facets in the eye of the fly. PEV is caused by translocation or P element-mediated transposition of the euchromatic gene, *white,* with pericentromeric heterochromatin. The aberrant juxtaposition causes stochastic spreading of repressive heterochromatin modifications, H3K9me2 and HP1, into *white,* and causes transcriptional silencing resulting in white eye facets. This process is stochastic, with cell-to-cell variability in heterochromatin spreading and thus red patches of cells are also observed. JIL-1, a genetic modifier of the PEV phenotype, antagonizes heterochromatin spreading into *white* by depositing H3S10ph at euchromatin regions.

1.7.4 The p90 RSK family in mammals

MSK, or Mitogen and Stress active protein Kinase family, shares 63% and 32% similarity, respectively, over the two kinase domains (KDI and KDII) with JIL-1 to form a distinct monophyletic clade (**Fig. 1.4**), and is generally accepted as the direct ortholog of JIL-1 (Jin et al. 1999). RSK or Ribosomal S6 Kinases, is the closest kinase family to JIL-1 and MSK, sharing 40% identity with the latter. Together RSK and MSK kinases are classified as the p90 RSK family.

MSKs are constitutively localized to the nucleus and are thought to be the most potent kinases for H3S10 in interphase cells (Tomas-Zuber et al. 2001); MSK1/2, however, have additionally acquired the ability to phosphorylate H3S28 (Vermeulen et al. 2009). The orthologous kinases share similarities in the kinase activation sequences – both require the phosphorylation of an activation loop within KDII prior to autophosphorylation and subsequent activation of KDI, which acts as the effector kinase domain for target substrates (Reyskens and Arthur 2016). The most significant disparity between MSKs and JIL-1, however, is the intrinsic autophosphorylation activity of JIL-1 which does not show any evidence of requiring external stimuli for activation (Jin et al. 1999). MSK kinases, on the other hand, are the distal effectors of two separate signalling cascades, – ERK and p38 MAPK, which are mediated by two classes of stimuli: serum/EGF and stress/heat-shock/UV, respectively. This is likely due to the dissimilarity of the KDII sequence, which serves as a MAPK docking sites for MSKs but is missing in JIL-1 (Reyskens and Arthur 2016).

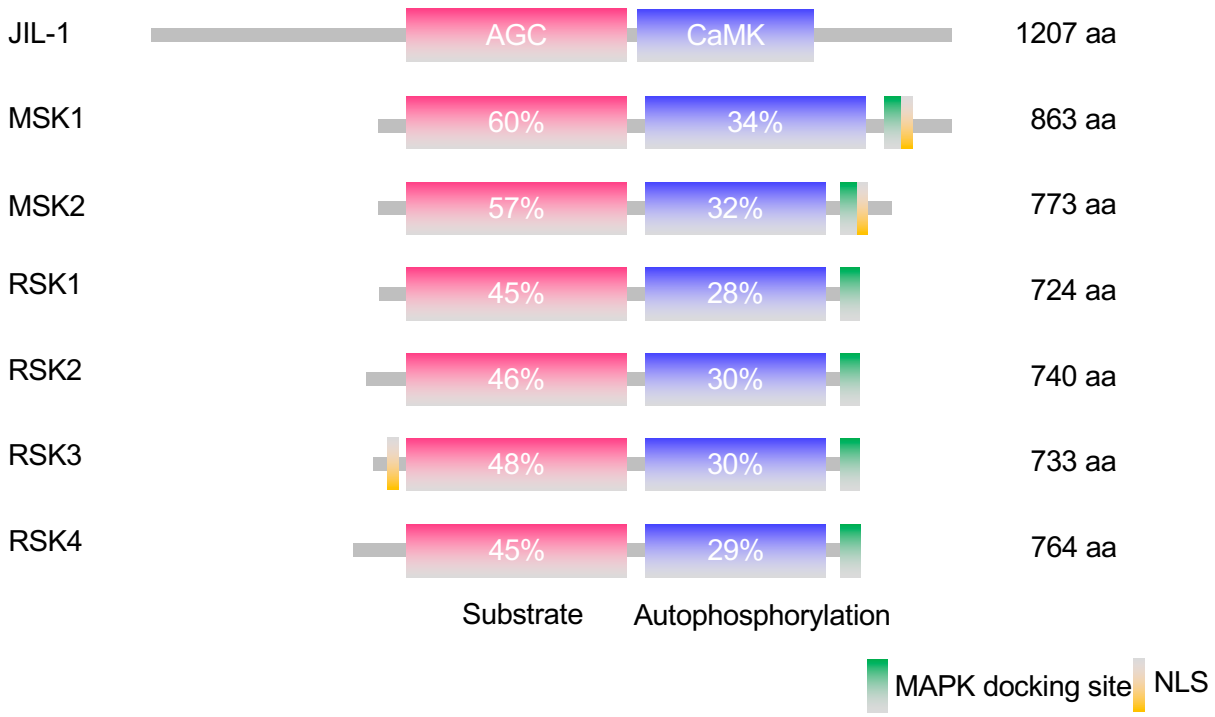


Figure 1.4 Protein structure of JIL-1 and MSKs & RSKs homologs in mouse.
Cartoon schematic depicting protein domain structures of MSK1/2 & RSK1/2/3/4 in relation to JIL-1. JIL-1 & MSK/RSK share significant homology (percentages depicting amino acid identity) over both kinase domains, named KDI and KDII. The N-terminal KDI domain is an AGC kinase domain, and is the effector kinase domain for histone H3 and other nuclear substrates. The C-terminal KDII is CaMK domain and is the regulatory site for kinase activity. During activation, both JIL-1 and MSK/RSK undergo sequential phosphorylation events initiating at the C-terminus, which then leads to activation of autophosphorylation of KDI by KDII. MSK/RSK require upstream MAPK kinases and possess docking sites in the C-terminus, whereas JIL-1 can be fully activated by autophosphorylation. In addition, MSKs possess a canonical nuclear localization signal (NLS) at the C-terminal tail region and RSK3 encodes a putative NLS signal at the N-terminus. RSK1/RSK2 are known to localize to the nucleus, likely transported via a chaperone.

In mouse, the MSK family includes MSK1 (*Rps6ka5*) and MSK2 (*Rps6ka4*), with MSK1 being predominantly expressed in the spleen, lymphocytes, central and peripheral nerves and oocytes, while MSK2 is expressed in bone, leukocytes and fibroblasts. Both are necessary for maximal induction of the nucleosomal response via phosphorylation of nuclear substrates such as CREB, ATF1 transcription factor and histone H3 upon stimulation of MAPK, which consequently results in the transcription of immediate early genes such as *Egr1*, *Junb*, *c-Fos*

(Soloaga et al. 2003). *Msk1*$^{-/-}$ ESCs showed no overt growth or morphological defects, likely owing to the low expression of MSK1 or compensation by the functionally redundant MSK2 (Arthur and Cohen 2000). Although *Msk1*$^{-/-}$ mice are born at the expected Mendelian ratio, and are overtly normal and fertile, aged *Msk1*$^{-/-}$ mice shows spontaneous striatal neuronal degeneration and increased ventricular volume at 9 months (Martin et al. 2011). *Msk1/2* double KO (DKO) mice were derived later and found to be overtly normal and fertile under pathogenic-free conditions. Phenotypic differences do not begin emerging until the KO animals are subjected to stress. *Msk1/2* DKO mice show hyperinflammation and increase susceptibility to endotoxin shock (Ananieva et al. 2008), reduced neuronal progenitor proliferation following induced seizure (Choi et al. 2012), and deficit in fear conditioning and spatial memory impairment (Chwang et al. 2007).

In mouse, the p90 RSK family includes four members with ubiquitous expression patterns, consistent with redundancy within the family. Structurally, the RSKs, including RSK1/2/3/4, are similar to MSKs, with two distinct kinase domains separated by a linker region (**Fig. 1.4**). MSKs and RSKs show considerable overlap in substrates, including transcription factors CREB, ATF1 and histone H3 (Chen et al. 1992; Anjum and Blenis 2008). In contrast to MSKs, which can respond to both mitogenic and stress stimuli, cytoplasmic RSK1/2/3 are activated by ERK as a consequence of serum/EGF stimulation and only translocate into the nucleus after PDK1-mediated activation (**Fig. 1.5**). RSK4, however, is the most functionally diverged isoform, showing constitutive activation independent of growth factor stimulation and is predominantly cytosolic (Dummler 2005). RSK1 (*Rps6ka1*) is expressed in highly proliferative tissues including epithelial cells, blastocysts and trophoblast stem cells, while RSK2

(*Rps6ka3*) is expressed in synaptic tissues, skeletal muscle, heart, pancreas and RSK3 (*Rps6ka2*) is expressed in developing neural and sensory tissues. Finally, RSK4 (*Rps6ka6*) is highly expressed in the oocyte, placenta and embryo.

Notably, null mutations of *RSK2* in humans are associated with Coffin-Lowry Syndrome (CLS), a rare X-linked dominant disorder characterized by small stature, hypertelorism, flat nose, large ears, tapered fingers, skeletal malformations, auditory/visual deficit, and cognitive impairment (Coffin et al. 1966; Lowry et al. 1971). Premature death is common amongst CLS patients, often due to congestive heart failure. Due to its relevance to CLS, *Rsk2* KO mice have been extensively studied. *Rsk2* null mice are 10-15% smaller than wildtype littermates and have a normal life span, but show impaired learning, cognition and coordination, and progressive skeletal disease due to impaired osteoblast differentiation (Yang et al. 2004). *Rsk1/2/3 TKO* mice are viable, and phenocopy the phenotype of RSK2-null mice in terms of abnormal craniofacial development, but were not extensively characterized (Laugel-Haushalter et al. 2014). Previously, RSK2 was proposed to be an H3S10 kinase *in vitro* and *in vivo*; *RSK2*-null fibroblast derived from CLS patient and *Rsk2* KO ESCs showed defects in inducible interphase H3S10ph in response to EGF (Sassone-Corsi et al. 1999). In later studies, however, the MSKs were shown to be the predominant kinases for induction of the nucleosomal response, and EGF-induced phosphorylation of H3S10 and CREB is unimpaired in CLS-derived fibroblasts (Wiggin et al. 2002); however the extent to which these kinases regulate global H3S10ph or transcription has never been examined in parallel.

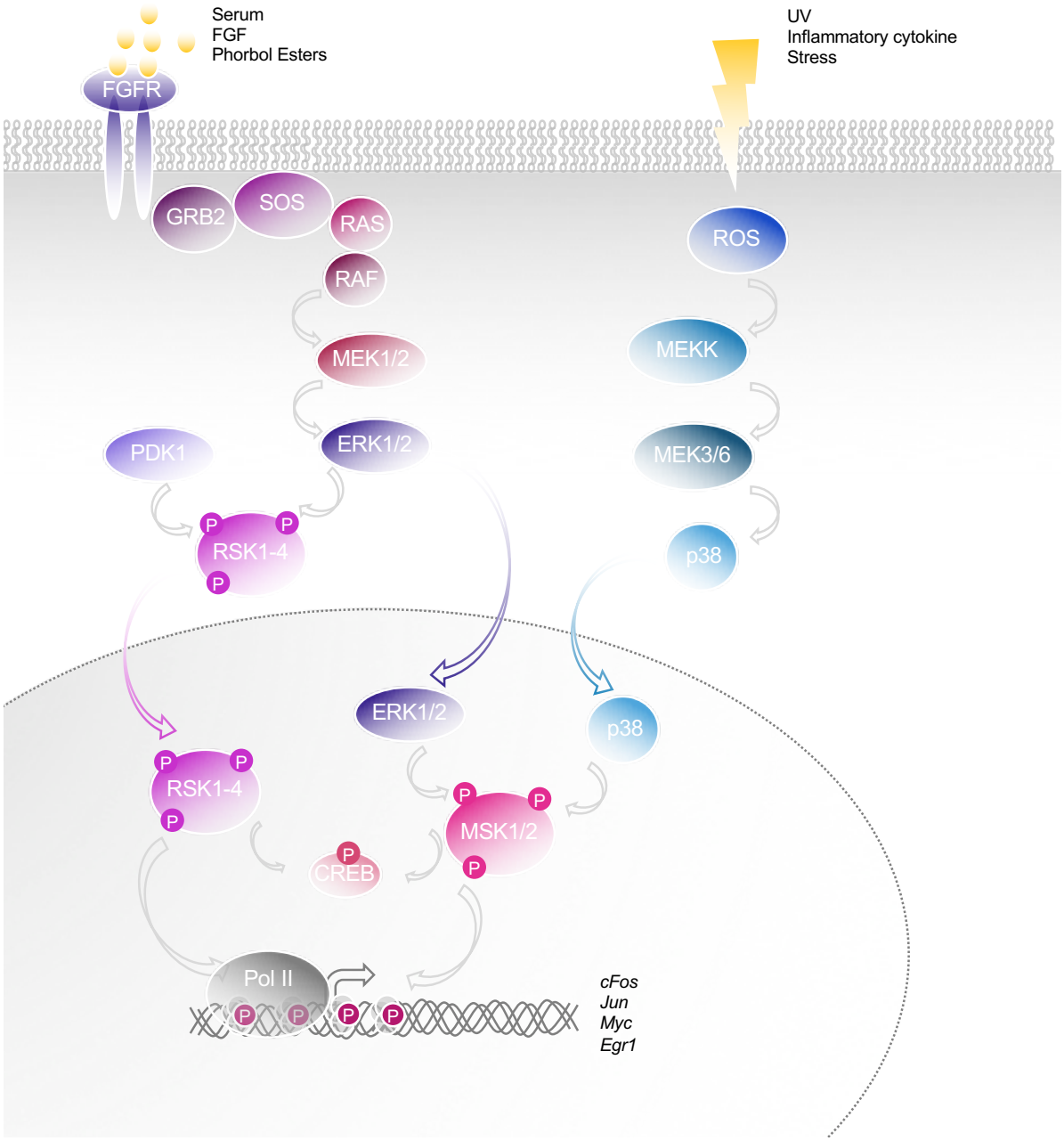


Figure 1.5 The MSK and RSK signalling pathways in mammalian cells.
In quiescent (i.e. serum starved) cells, mitogens bind to the extracellular domain of FGFR to initiate a signalling cascade to activate intracellular MAPKs, including MEK and ERK. Activated ERK then phosphorylates RSK1/2/3/4 in the cytoplasm, or translocates into the nucleus to activate MSK1/2. RSKs additionally require PDK1 for full activation prior to nuclear translocation. Once in the nucleus, both MSKs and RSKs phosphorylate transcription factors, such as CREB and ATFs, or the chromatin itself to induce transcription of immediate early genes. In addition to mitogen signalling, MSKs can also responds to p38-dependent signalling in response to UV-B/cytokines /stress.

1.8 The phospho-methyl switch and H3K9me "readers"

Biochemically, H3S10ph antagonizes the deposition and recognition of methylation on the adjacent H3K9. This occurs at the level of substrate incompatibility for H3K9 KMTs, where the H3S10ph peptide tail cannot be methylated by SUV39H1, G9a/GLP or SETDB1 (**Fig. 1.6**) (Rea et al. 2000; Duan et al. 2008; Schultz 2002). H3S10ph also sterically hinders recognition of H3K9me1/2/3 and precludes binding of many methyl lysine reader modules, in a conserved mechanism known as the "phospho-methyl" switch (Fischle et al. 2005).

The structural basis for methyl lysine recognition is conserved amongst eukaryotes, even between distinct reader modules. Crystal structures of methyl lysine reader modules that recognize different states of methylation (mono-, di-, tri-) on different residues reveal structural similarities in the domain relevant for histone tail recognition (Taverna et al. 2007; Musselman et al. 2012). The reader modules typically consist of 2-4 aromatic amino acids positioned perpendicularly, forming a hydrophobic pocket, or an "aromatic cage", around the methylated lysine residue. The specificity for mono-, di- or tri- methylation is determined by the composition and size of the pocket, while preference for a specific lysine residue is imparted by interactions with amino acids surrounding the cage. The H3K9-specific readers that have been extensively characterized include the chromodomain, ankryin repeats, tandem tudor and WD40 domains, all of which utilize a methyl lysine recognition mechanism similar to that described above (Taverna et al. 2007).

The phospho-sensitivity of methyl lysine readers has been characterized in greatest detail for the chromodomain of HP1, where Asn57 forms a hydrogen bond with unmodified H3Ser10

(Jacobs 2002). Phosphorylation of S10 not only breaks this stabilizing interaction, but creates steric hindrance with surrounding aromatic residues, decreasing binding affinity of histone:chromodomains by ~100 fold (Andrews et al. 2016). Similarly, hydrogen bonding of Glu87 within the ankryin repeat of G9a to mono- and di-methylated H3K9 is disrupted by phospho S10 (Collins et al. 2008).

Until recently, with the characterization of the binding characteristics of UHRF1/NP95 and ATRX, H3S10ph was thought to universally exclude all H3K9 readers. UHRF1 recognizes nascent hemi-methylated DNA via its SRA domain, and is an obligate partner of maintenance DNA methylation transferase 1, DNMT1 (Sharif et al. 2007). Aside from possessing an SRA domain, UHRF1 contains a tandem tudor domain (TTD) that not only recognizes H3K9me3, but also exhibits increased affinity for H3K9me3S10ph (Rothbart et al. 2012). While the TTD also possesses an aromatic cage specific for H3K9me3, the residue Asn147 just outside of the cage exhibits flexible conformation and allows for rotational freedom of phospho S10. The ADD (ATRX-DNMT3-DNMT3L) domain of ATRX is also a phospho-tolerant H3K9me3 reader, with S10ph positioned away from the pocket and stabilized internally via a salt bridge with H3R8 (Noh et al. 2014; Eustermann et al. 2011; Kunowska et al. 2015). The biochemical basis of H3K9 recognition is consistent with the observation that both UHRF1 and ATRX are reported to be retained on mitotic chromosomes, while HP1s are displaced, likely by H3S10ph (Rothbart et al. 2012; McDowell et al. 1999b; Fischle et al. 2005).

K9 Reader	Protein Module	K9m3	K9m2	K9m1	K9m3S10ph	K9m2S10ph	K9m1S10ph
HP1α/β	Chromodomain	++	+	-	-	-	-
MPP8	Chromodomain	++	+	-	-	-	-
CHD4	PHD2	+	-	-	-	-	-
ATRX	ADD	+	-	-	++	-	-
UHRF1	Tandem Tudor	+	+	-	++	++	-
GLP	Ankryn Repeat	-	-	+	-	-	-
JMJD2A	JmjC	+	+	-	-	-	-
CDYL1	Chromodomain	++	++	+	-		
LSD1	AOD		+	+		-	-

Figure 1.6 The phospho-methyl switch at H3K9/S10.
Phosphorylation of H3S10 hinders the catalytic activity of all known H3K9 KMTs. In addition, most known readers of H3K9 methylation have been shown to be displaced by H3S10ph, in a mechanism known as the "phospho-methyl switch". Reader modules that exhibit phospho-sensitivity includes the chromodomain in H3K9me readers HP1, MPP8, and CDYL. Similar intolerance to H3K9meS10ph has been reported for GLP, an obligate paralog of G9a, JMJD2A & LSD1, the H3K9 demethylases. Comparatively, few reader modules tolerate the phospho epitope. Thus far only the ADD domain of ATRX (a SWI/SNF helicase) and the TTD of UHRF1/NP95 (an essential cofactor to DNMT1), have been shown to be insensitive to the methyl-lysine phospho-serine bivalent histone tail. (Fischle et al. 2005; Rothbart et al. 2012; Musselman et al. 2009; Noh et al. 2014; Collins et al. 2008; Ng et al. 2007; Franz et al. 2010; Forneris et al. 2005).

1.9 Thesis objectives

In **Chapter 1**, I review the current state of knowledge on H3S10ph, showing that while H3S10ph was among the earliest histone PTMs described and is universally observed on mitotic chromosomes, surprisingly little is known about the role of this epigenetic mark in interphase. Biochemical and genetic evidence suggest that phosphorylation of the adjacent serine 10 antagonizes both the methylation of the adjacent H3K9 by KMTs and the recognition of H3K9me2/3 by readers of this mark, in a conserved mechanism known as the "phospho-methyl switch". We and others have previously demonstrated that di- and tri-methylation of lysine 9 on histone H3 by GLP/G9a and SETDB1, respectively, is instrumental in transcriptional repression of ERVs and a subset of genes (Maksakova et al. 2013; Yokochi et al. 2009; Matsui et al. 2010; Karimi et al. 2011). I hypothesize that H3S10ph may function to maintain euchromatin and transcription potential by barring certain genomic regions from the H3K9 KMT machinery. The general objective of my thesis project was to comprehensively map H3S10ph in interphase mammalian cells, and to understand the relationship between H3S10ph and H3K9me2/3 in the context of transcriptional regulation in ESCs using next-generation sequencing methods.

In **Chapter 2**, I describe the materials and methods relevant to data presented in **Chapters 3, 4, and 5**. I established several ESC lines used in this thesis work, including the reporter ESC line harbouring the FUCCI (fluorescent ubiquitin-mediated cell cycle indicator) reporter, an ESC line overexpressing YFP-tagged H3.3S10A, and a panel of CRISPR-mediated KO ESC lines. I also optimized and adapted an existing native chromatin immunoprecipitation (N-ChIP) protocol to capture the labile H3S10ph epitope, using crosslinking followed by MNase digestion (XM-ChIP). Details of library preparation and sequencing on the Illumina platform are

also presented. Finally, I describe the bioinformatics methods used to analyze the relevant sequencing data.

In **Chapter 3**, I analyzed the distribution of H3S10ph in specific cell cycle stages in ESCs and MEFs, and found that this mark is enriched in interphase in broad gene-dense euchromatin domains. Reminiscent of JIL-1 dependent H3S10ph in *Drosophila,* interphase H3S10ph is strongly anti-correlated with H3K9me2 domains in both cell types. In ESCs, H3S10ph and H3K9me2 show high concordance with the replication timing (RT) program, marking early- and late-replicating fractions of the genome, respectively. I further explored the reciprocal antagonism between H3K9me2 and H3S10ph, using histone mutants and ESCs deficient in the H3K9 KMT *GLP*. This work uncovered interphase H3S10ph as a pervasive PTM in euchromatin in ESCs and specifically identified replication timing transition regions (TTRs) as hotspots for aberrant transcription in the absence of H3K9me2.

In **Chapter 4,** I discovered that a large fraction of H3K9me3 is converted into H3K9me3S10ph by Aurora B/C at mitosis, and this bivalent mark persists into the next G1. To probe the function of H3K9me3S10ph in interphase, I also mapped HP1β localization and found that HP1β localized robustly to H3K9me3S10ph-marked regions. These results suggest that H3K9me3S10ph and H3K9me3 co-exist heterogeneously in G1, with HP1β binding to the latter. Finally, I discuss the implications of the co-existence of H3K9me3- and H3K9me3S10ph-marked nucleosomes in recruiting separate classes of readers with differential H3S10ph tolerance at ERVs and germline-specific genes.

In **Chapter 5,** using the panel of *Msk1/2 & Rsk1/2/3/4* KO ESCs that I generated using the CRISPR/Cas9 system, I demonstrate that MSK and RSK kinases, homologs of the JIL-1 kinase in *Drosophila*, are dispensable for the broad interphase H3S10ph domains present in ESCs. Among the KO ESC lines, *Msk2*$^{-/-}$ showed the greatest loss of H3S10ph at promoters and distal enhancers, with many genes deregulated. Notably, genes down-regulated in *Msk2*$^{-/-}$ ESCs are enriched for testis-specific expression, and many are *bona fide* targets of the repressor SETDB1. These results conclusively show that MSK and RSK are not required for the majority of H3S10ph present in interphase ESCs, and potentially uncover a role for MSK2 in mediating expression of specific germline genes in ESCs.

Finally, in **Chapter 6**, I summarize experimental results from **Chapters 3-5**, and discuss potential future directions for this project. Given that the JIL-1 homologs were dispensable for most of the interphase H3S10ph observed in ESCs, I specifically describe a novel screening strategy to identify interphase H3S10ph kinases. Furthermore, the reiterative observation that germline-specific genes are deregulated upon disruption of H3K9me or H3S10ph suggest these genes act as the nexus for the phospho-methyl interplay, and thus, are obvious pursuits for future work. Finally, I discuss the merits of characterizing the distribution of interphase H3S10ph *in vivo* to advance our understanding of this prevalent mark in chromatin biology.

2. MATERIAL & METHODS

2.1 Cell lines and culturing conditions

Mouse embryonic stem cells, including TT2 (WT), *G9a*$^{-/-}$, *Glp*$^{-/-}$, were cultured in DMEM supplemented with 15% fetal bovine serum (HyClone Laboratories, Logan, UT, USA), 20 mM HEPES, 1 mM L-glutamine, 1 mM non-essential amino acids, 1mM sodium pyruvate, 0.1 mM betamercaptoethnol, 100 U/mL penicillin-streptomycin, and recombinant leukemia inhibitory factor (LIF) on gelatinized plates. ESCs were passaged every 2-3 days. Hesperadin (Millipore) was diluted to working stock of 1 mM in DMSO, and diluted in complete ES media to a final concentration of 200 nM for 3 hours prior to harvesting. Mouse embryonic fibroblasts were cultured in DMEM supplemented with 10% fetal bovine serum, 1 mM L-glutamine, and 100 U/mL penicillin-streptomycin.

2.2 Generation of cell lines

2.2.1 FUCCI

To derive an ESC line stably expressing FUCCI, 1.5 ug of CAG-hCdt1:mKO2-T2A-hGeminin:Venus plasmid (kind gift from U. Naumann) was lipotransfected into WT TT2 ESCs using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's recommendation. Stable transfectants were obtained through 2 successive rounds of FACS for mKO2^{+} or Venus^{+} at day 4 and day 11 post-transfection.

2.2.2 H3.3-YFP

CMV-H3.3-YFP vector was obtained from Addgene. Ser10Ala point mutation was introduced using the QuikChange Lightning site-directed mutagenesis kit (Agilent, Cat#210518), and sequenced using T7 primer to verify the mutation. H3.3 and H3.3S10A were then cloned into a TetOn Puromycin resistance plasmid using EcoRI and SpeI; vector was then lipotransfected using Lipofectamine 2000 into rTTa2-M2-expressing TT2 ESCs and subjected to puromycin selection for 24 hours. Clones were then isolated and characterized using flow cytometry 24 hours after Dox induction.

2.2.3 CRISPR-mediated deletion of Msk1/2 & Rsk1/2/3/4 KO

Short guide RNA, or sgRNA, were designed *in silico* using the sgRNA Designer webtool from the Broad Institute against RefSeq annotated transcripts of *Rps6ka5* (MSK1), *Rps6ka4* (MSK2), *Rps6ka1* (RSK1), *Rps6ka3* (RSK2), *Rps6ka2* (RSK3) and *Rps6ka6* (RSK4). Two top ranking sgRNA were chosen and cloned into a multicistronic vector encoding the Cas9 endonuclease and puromycin resistance. Inserts encoding sgRNA were then verified using Sanger sequencing. WT Fu.7 were then transfected using Lipofectamine 3000 (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's recommendation, and subjected to puromycin selection for 48 hours. Cells were then seeded at low density to obtain clonal colonies, and 10-20 clones per each kinase were genotyped at the targeted exon using Sanger sequencing.

2.3 Immunofluorescence, flow cytometry and cell cycle analysis

Indirect immunofluorescence staining was performed using standard methods. Cells were harvested using trypsinization were then crosslinked with 1% formaldehyde, permeabilized with

0.25% Triton-X-100 and blocked with 1% bovine serum albumin (Sigma- Aldrich). Cells were then incubated with anti-H3S10ph (CMA311, 1:100), anti-PCNA (Abcam ab29) at 37 °C for 1 hour and subsequently incubated with Alexa Fluor 488 and 594-labeled secondary antibodies (Life Technologies). DNA was counter- stained with 500 ng/mL Hoescht 33342 (Sigma-Aldrich). Flow cytometry analysis of FUCCI cells was performed as previously described. Briefly, cells were resuspended in phosphate buffered saline and analyzed on a BD LSRII561 flow cytometer using BD FACS Diva software. Cells were successively gated on forward and side scatter, then singlet cells, and lastly mKO^+ or $Venus^+$ cells, using the WT TT2 ESC line as mKO^-Venus^- to set the gates.

2.4 Native chromatin immunoprecipitation (NChIP)

NChIP-seq experiments were performed as previously described (Karimi et al. 2011), with minor modifications. To generate chromatin, 10^7 cells were trypsinized, washed in PBS and flash frozen. Cells were resuspended in 250 μL of douncing buffer (10 mM Tris-HCl pH 7.5, 4 mM $MgCl_2$, 1 mM $CaCl_2$, 1 x protease inhibitory cocktail or PIC) and homogenized through a 25-gauge needle syringe for 25 repetitions. Subsequently, 1.25 μL of 50 U/mL of MNase was added to the nuclei and incubated at 37 °C for 7 minutes. The reaction was quenched by the addition of 0.5M EDTA and incubation on ice for 5 minutes. 1 mL of hypotonic lysis buffer (0.2 mM EDTA pH 8.0, 0.1 mM benzamidine, 0.1 mM phenylmethylsufonyl fluoride, 1.5 mM dithiothreitol, 1 x PIC) was then added and the mix incubated with rotation for 1 hour at 4 °C. Protein A/G Dynabeads (Life Technologies) were mixed and washed twice with 1 mL of IP buffer (10 mM Tris-HCl pH 8.0, 1% Triton X-100, 0.1% Deoxycholate, 0.1% SDS, 90 mM NaCl, 2 mM EDTA, 1 x PIC), and then resuspended in the original volume with IP buffer. Cellular debris was

pelleted and the soluble chromatin fractions were then pre-cleared by rotation with Protein A/G Dynabeads for 2 hours at 4 °C. 100 μL of the pre-cleared chromatin was purified using phenol:chloroform extraction and fragment sizes were analyzed on a 1.5% agarose gel.

Antibody-bead complexes were prepared using antibodies specific for H3K9me2 (Abcam ab1220, 5 μL), H3K4me3 (Abcam ab1012, 5 μL), H3K36me3 (Abcam ab9050, 5 μL), and YFP (Abcam ab290, 5 μL) incubated with Protein A/G dynabeads at an antibody:bead ratio of 1:4, in IP buffer at 4 °C, rotating for 2 hours. 10^6 cell equivalents of chromatin were then added to each antibody-bead complex aliquot and rotated overnight at 4 °C. The immunoprecipitated complexes were then washed twice with 400 μL of ChIP wash buffer (20 mM Tris-HCl pH 8.0, 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 1 x PIC), followed by a single wash with ChIP final wash buffer (20 mM Tris-HCl pH 8.0, 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 500 mM NaCl). The protein-DNA complexes was eluted by incubating the beads in 200 μL of elution buffer (100 mM $NaHCO_3$, 1% SDS) and RNAse at 68 °C for 2 h. Beads were then washed and eluted in the presence of RNase A. DNA was then purified with 2 x volume of phenol:choloroform:isoamyl (25:24:1) in MaXtract High Density phase-locked tubes (Qiagen) and ethanol precipitated at -20 °C overnight. DNA pellets were washed twice with 70% ethanol, and resuspended in 60 μL ddH_2O.

2.5 Crosslinked MNase ChIP (XMChIP)

For crosslinked MNase ChIP of H3S10ph and HP1β, 10^7 ES cells were harvested and crosslinked with 1% formaldehyde in 10 mL of PBS for 10 minutes at room temperature then

quenched with 0.125 M glycine for 5 minutes. Cells were washed 1x with PBS, and resuspended in 1 mL of EZ Nuclei Isolation Buffer (Sigma). The cytoplasmic supernatant was discarded and the pellets containing nuclei were flash frozen with liquid N_2. Nuclei were washed once with MNase Wash buffer (50 mM Tris pH 8.0, 1.5 mM DTT, 1 mM PMSF, and 1 x Protease Inhibitor Cocktail (Roche)) and resuspended in MNase Digestion Buffer (10 mM Tris-Cl pH 7.5, 4 mM $MgCl_2$, 1 mM $CaCl_2$, 1 mM PMSF, 1 x PIC, and 1 x PhosStop). Chromatin were then digested with micrococcal nuclease (NEB) and incubated at 37 °C for 7 minutes, yielding predominantly mononucleosomes. The MNase digestion was quenched with the addition of EDTA to a final concentration of 10 mM. Cells were then lysed with 1 mL of IP Buffer (0.5 % NP-40, 0.1% NaDOC, 0.1% SDS, 150 mM NaCl, 10 mM EDTA, 1 mM PMSF, 1 x PIC, and 1 x PhosStop) at 4 °C on a rotator for 1 hour. Cell debris were then pelleted, and the soluble chromatin fraction was precleared with Protein A/G Dynabeads (Life Technologies) at 4 °C, rotating for 2 hours. During preclearing, antibody-bead complexes were prepared using antibodies specific for H3S10ph (clone CMA311) and incubated with sheep anti-mouse IgG dynabeads, in IP buffer at 4 °C, on a rotator for 2 hours. 10^6 cell equivalents of chromatin was then added to each antibody-bead complex, and rotated overnight. Beads were then washed and eluted in the presence of RNAse A. To reverse crosslinking, eluted DNA was incubated with Proteinase K and high salt overnight at 65 °C. DNA was then purified with 2x volume of phenol:choloroform:isoamyl (25:24:1) in MaXtract High Density phase-locked tubes (Qiagen) and ethanol precipitated at -20°C overnight. DNA pellets were washed two times with 70% ethanol, and resuspended in 60 µL ddH_2O.

2.6 Genomic DNA isolation and genotyping

Genomic DNA was isolated using the HotSHOT protocol. A small number of trypsinized cells (~1,000) were lysed and boiled in 50 μL HotShot I alkaline buffer (25 mM NaOH, 0.2 mM EDTA) for 30 minutes, cooled at 4 °C for 10 minutes and neutralized immediately with HotShot II (40 mM Tris-HCl pH 5.0). 2 μL of the DNA prep is then used for PCR in genotyping reactions. Genotyping primers were designed to flank the sgRNA-target sequence, and validated in WT parent line to verify the size of amplicons. Clones were first screened for overt deletions as evident by a band shift – positive clones were then sequenced using the Sanger method.

2.7 RNA extraction and RT-PCR

RNA was isolated using the RNeasy kit (Qiagen) according to the manufacturer's protocol. DNaseI treated RNA was subjected to first-strand cDNA synthesis using the RevertAid H Minus kit (Fermentas) in the presence or absence of reverse transcriptase. qRT-PCR using primers of interest, or β-actin specific primers as an internal control, was conducted with EvaGreen dye (Biotium) on an Opticon 2 thermal cycler (Bio-Rad). Relative expression levels were determined by normalizing to the β-actin gene.

2.8 Whole-cell extraction and western blot

For preparing cell extracts for western blotting, cells were lysed in RIPA buffer (50 mM Tris pH 8.0, 150 mM NaCl, 1% NP-40, 0.25% deoxycholate, 0.1% SDS). Histones were isolated from ESCs for westerns by boiling cells in SDS-PAGE loading buffer. Western blotting was performed as previously described using anti-H3S10ph (H. Kimura, CMA311) and anti-pan H3

(Sigma H0164). Primary antibodies were detected using IRDYE-conjugated secondary antibodies and scanning on the Odyssey imager (LiCOR Biosciences).

2.9 Sequencing library construction and bioinformatic analysis

Construction of Illumina sequencing libraries was as described previously (Brind'Amour et al. 2015). Briefly, ChIP and input DNA (10-20 ng total) was subjected to end repair with T4 and Klenow, A-tailing, and adaptor ligation. Following ligation, libraries were amplified (8-10 PCR cycles) using primers complementary to the adaptors. PCR products in the range of 200-700 bp was then purified on a 2% EX Gel (Life Technologies) and quantified on a Qubit and Agilent Tape station. Paired-end sequencing was performed on an Illumina Hi-Seq 2000 following the recommended protocol. Reads were aligned to mm9 using Burrows-Wheeler Aligner (BWA) using default parameters. PCR duplicates and multi-match reads (Q=0) were removed from all subsequent statistical analyses. Wiggle plots were generated from genome coverage counts using reads of MAPQ $\geq$ 5. RPKM (Reads Per Kilobase per Million mapped reads) and z-score values were calculated using the equations below as per Karimi et al. (Karimi et al. 2011) with SeqMonk and VisRSeq (Younesy et al. 2015) software.

Strand-specific mRNA-seq libraries were constructed from purified polyA RNA, as described in Morrissy et al. (Morrissy et al. 2016), from 6 ug of DNAse1 treated total RNA and sequenced on an Illumina Genome Analyzer$_{iix}$ following the manufa cturer's recommended protocol. The resulting sequence reads were aligned to the mouse reference genome (mm9) using MAQ v0.7.1 (Li et al. 2009), with Smith-Waterman alignment disabled and annotated exon-exon junctions compiled from Ensembl v54 (Flicek et al. 2009), RefSeq (Pruitt and Maglott 2001) and UCSC

(Rhead et al. 2009) gene annotation sources (downloaded from http://genome.ucsc.edu on March 17, 2009), as described (Karimi et al. 2011). Sequence reads that could be uniquely assigned a position in the transcript resource (exon-exon junctions) were computationally repositioned to the genomic mm9 coordinates and a single merged bam file generated for downstream analyses. The Samtools pileup utility was used to generate data tracks for visualization (wig and bigWig (Kent et al. 2010)) in the UCSC browser. Repbase (Jurka et al. 2005), a comprehensive database of repetitive elements, was used for alignment of reads to specific repetitive elements.

To quantify expression levels and histone marks we calculated RPKM values. For each genomic region of interest, RPKM was calculated using the following formula: $RPKM_x = (n/L) * (1000000/N_x)$, where n is number or reads aligned to the region, L is the length in kilobases, and N_x is the total number of aligned reads used for normalization. For pair-wise sample comparisons, an empirical Z-score was calculated assuming the distribution of RPKMs for each sample followed a Poisson model: $Z = (RPKM_A - RPKM_B)/\sqrt{(RPKM_A + RPKM_B)}$ where $\mathrm{RPKM_A}$ and $\mathrm{RPKM_B}$ are RPKMs in the region of interest of A and B samples respectively. H3S10ph ChIP-seq data was normalized by subtracting input RPKM from ChIP RPKM values, z-scores were calculated from the normalized score. Gene Ontology enrichment was performed using InnateDB (Breuer et al. 2012) or PantherDB (Mi et al. 2017). Kendall correlation was performed by ranking all genomic 5 kb bins by RPKM values and comparing the ranks of chromatin features, followed by unsupervised clustering to generate the correlation matrix. Data visualization and plots were generated in VisRSeq (Younesy et al. 2015), Excel or RStudio.

2.10 Deposited sequencing datasets accession

Dataset	Accession	Reference
ESC_Fucci_G1_H3S10ph	GSE97947	Chapter 3
ESC_Fucci_G1_Input	GSE97947	Chapter 3
ESC_Fucci_S_H3S10ph	GSE97947	Chapter 3
ESC_Fucci_S_Input	GSE97947	Chapter 3
ESC_Fucci_SG2M_H3S10ph	GSE97947	Chapter 3
ESC_Fucci_SG2M_Input	GSE97947	Chapter 3
ESC_TT2asynch_H3S10ph	GSE97947	Chapter 3
ESC_TT2_Ctrl_H3S10ph	GSE97947	Chapter 3
ESC_TT2_Ctrl_Input	GSE97947	Chapter 3
ESC_TT2_Hesp_H3S10ph	GSE97947	Chapter 3
ESC_TT2_Hesp_Input	GSE97947	Chapter 3
ESC_TT2_H3K9me2	GSE97947	Chapter 3
ESC_H3.3WT_YFP	GSE97947	Chapter 3
ESC_H3.3S10A.1_YFP	GSE97947	Chapter 3
ESC_H3.3S10A.2_YFP	GSE97947	Chapter 3
ESC_H3.3WT_H3K9me2	GSE97947	Chapter 3
ESC_H3.3S10A.1_H3K9me2	GSE97947	Chapter 3
ESC_H3.3S10A.2_H3K9me2	GSE97947	Chapter 3
ESC_GlpKO_H3K9me2	GSE97947	Chapter 3
ESC_GlpKO_Hesp_H3S10ph	GSE97947	Chapter 3
ESC_GlpKO_Hesp_Input	GSE97947	Chapter 3
ESC_TT2_H3K4me3	GSE97947	Chapter 3
ESC_TT2_H3K36me3	GSE97947	Chapter 3
ESC_TT2_ssRNAseq	GSE97947	Chapter 3
ESC_GlpKO_H3K4me3	GSE97947	Chapter 3
ESC_GlpKO_H3K36me3	GSE97947	Chapter 3
ESC_GlpKO_ssRNAseq	GSE97947	Chapter 3
ESC_G9aKO_H3K4me3	GSE97947	Chapter 3
ESC_G9aKO_H3K36me3	GSE97947	Chapter 3
ESC_G9aKO_ssRNAseq	GSE97947	Chapter 3
MEF_Ctrl_H3S10ph	GSE97947	Chapter 3
MEF_Ctrl_Input	GSE97947	Chapter 3
MEF_Hesp_H3S10ph	GSE97947	Chapter 3
MEF_Hesp_Input	GSE97947	Chapter 3
TT2_H3K9me3_1million	GSE63523	Chapter 4
TT2_HP1b	GSE71114	Chapter 4

2.11 Primer sequences

Primer	Sequences
***Sox2* promoter ChIP F**	GAAAGCCTGGCCTTTTGCAGAATC
***Sox2* promoter ChIP R**	CTGTGTTTCCTACGTCGACCATCA
IAP LTR F	GCTCCTGAAGATGTAAGCAATAAA
IAP LTR R	CTTCCTTGCGCCAGTCCCGAG
Chr2 IAP Proximal F	AGAAGATTCTGGTCTGTGGTGT
Chr2 IAP Proximal R	TCTCAATTGGCTATAGTGTC
Chr2 IAP Distal F	TCACCAGGTCCAGAGTCTAAC
Chr2 IAP Distal R	TCGCCAGGGCTCAGAGTCAC
***Msk1* ex5 sgRNA**	TGTGGAAAGGGCTCTTCACCGCCTGGATTCTAACGGCCATGGTTTTAGAGCTAGAAATA GCAAGTTAAAATAAGGCTAG
***Msk2* ex3 sgRNA**	TGTGGAAAGGGCTCTTCACCGCCAGCACCGAGCGTTCGGTGGTTTTAGAGCTAGAAATA GCAAGTTAAAATAAGGCTAG
***Rsk1* ex3 sgRNA**	TGTGGAAAGGGCTCTTCACCGCTCCATCACACACCACGTCAGTTTTAGAGCTAGAAATA GCAAGTTAAAATAAGGCTAG
***Rsk2* ex9 sgRNA**	TGTGGAAAGGGCTCTTCACCGCCAGAAGTAGTTAACCGCAGGTTTTAGAGCTAGAAATA GCAAGTTAAAATAAGGCTAG
***Rsk3* ex3 sgRNA**	TGTGGAAAGGGCTCTTCACCGAAGGTTACAGGATCAGACGCGTTTTAGAGCTAGAAATA GCAAGTTAAAATAAGGCTAG
***Rsk4* ex12 sgRNA**	TGTGGAAAGGGCTCTTCACCGCCTGAAGTAGTAAACAGACGGTTTTAGAGCTAGAAATA GCAAGTTAAAATAAGGCTAG
***Msk1* int5 F**	CTTTTGTCTCCGAGGGACCTTGAA
***Msk1* int4 R**	ATTTGCTCGCAGAGTTGCCA
***Msk2* int3 F**	TGGGAGTGAGGAGACGTTGAAATG
***Msk2* int2 R**	CCCAGGCTAGTTTGCACTGCAAAT
***Rsk1* int3 F**	CGTCCTCTGATCCTTTGGTGCATT
***Rsk1* int2 R**	TCTCCTGGACCTGTTTCCATCT
***Rsk2* int10 F**	TATGGTCAACTGACGGTTGGGT
***Rsk2* int12 R**	CCCTAGTCCTACACTTCAGGCAAT
***Rsk3* int2 F**	GTCACCCGTAATGGACTCGTTCTT
***Rsk3* int3 R**	GCCACAGTGGAACAAATGGCTT
***Rsk4* int12 F**	ACAGCTAGGGAAGGTGTGAAAC
***Rsk4* int11 R**	ACTGAAGCCAGACAAGGAAGCA

3. H3S10PH BROADLY MARKS EARLY-REPLICATING DOMAINS IN INTERPHASE

3.1 Introduction

H3S10ph is a highly-conserved histone PTM; however, the influence of this mark on chromatin structure remains elusive, partly due to H3S10ph being observed in two seemingly paradoxical contexts – broadly marking condensed chromosomes during mitosis and transiently marking the promoters of inducible genes in interphase. By early metaphase, histone H3 is extensively phosphorylated by Aurora kinase B (AURKB) at Ser10 and S28 (Wei et al. 1998), a process essential for the initiation of chromosome condensation and transition through mitosis (Van Hooser et al. 1998). Conversely, during interphase, activation of the MAPK pathway by mitogens or stress promotes rapid H3S10 phosphorylation at the promoter and enhancer regions of immediate early genes (Mahadevan et al. 1991; Cheung et al. 2000), stimulating the release of paused RNA Pol II and P-TEFb dependent transcription elongation (Zippo et al. 2009). While several nuclear kinases have been implicated as the downstream effectors of such inducible H3S10 phosphorylation at specific loci, including RPS6KA5/MSK1 RPS6KA4/MSK2 (the homologs of JIL-1 in flies), RPS6KA3/RSK2, CHUK/IKKα and PIM1 (Thomson et al. 1999b; Sassone-Corsi et al. 1999; Anest et al. 2003; Zippo et al. 2009), a genome-wide view of the dynamics of H3S10ph during the mammalian cell cycle has not been reported.

H3S10ph inhibits the lysine methyltransferase (KMT) activity of mammalian Su(var)3-9 family members *in vitro*, including SUV39H1, SETDB1 and G9a and promotes displacement of Chromobox proteins, also called Heterochromatin Protein 1 (HP1) from H3K9 methylated regions (Rea et al. 2000; Schultz 2002; Rathert et al. 2008; Fischle et al. 2005). While histone peptide binding studies have shown that H3K9me3 binding modules vary in their sensitivity to the presence of H3S10ph, the chromodomain of HP1 and ankyrin repeats of G9a /GLP are inhibited by H3S10ph-marked peptides (Fischle et al. 2005; Rathert et al. 2008). Such antagonism between H3S10ph and methylation of the adjacent H3K9, termed the "phospho-methyl switch" (Fischle et al. 2005), has been best characterized in the context of mitosis; however, the extent of crosstalk between H3S10ph and H3K9me2/3 in interphase mammalian cells remains unexplored.

Interplay between H3S10ph and H3K9 methylation is well documented in *Drosophila*. The essential H3S10 kinase JIL-1, isolated originally as an antimorph of position effect variegation (PEV), associates with chromatin throughout the cell cycle (Wang et al. 2001) and localizes to gene-dense regions, forming "blanket" domains over gene bodies (Regnard et al. 2011). Notably, chromosome morphology is severely impacted in a *JIL-1* hypomorph, with widespread loss of euchromatin interbands, despite normal mitotic H3S10 phosphorylation (Jin et al. 1999). Furthermore, H3K9me2 and HP1 spread ectopically into euchromatin polytene bands, suggesting that H3S10ph functions to reinforce euchromatin boundaries in flies by antagonizing heterochromatin propagation (Zhang et al. 2006). Intriguingly, the lethality observed in *JIL-1* null mutants is rescued in Su(var)3-9 double mutants, indicating that the aberrant gene expression observed in *JIL-1* hypomorphs may be a consequence of ectopic H3K9

methylation. Indeed, depletion of interphase H3S10ph in the *JIL-1* hypomorph leads to global redistribution of H3K9me2, with gain or loss of heterochromatin corresponding with down- and up-regulation of associated genes, respectively (Cai et al. 2014). Taken together, these observations reveal that crosstalk between H3K9me2 and H3S10ph plays a pivotal role in determining large-scale chromatin structure and in turn transcription potential in interphase *Drosophila* cells.

In mammals, H3K9me2 is deposited in euchromatic regions by G9a and its obligate paralog, GLP/Ehmt1 (Shinkai and Tachibana 2011). In mouse ESCs, H3K9me2 is enriched in megabase (Mb)-scale LADs that are gene-poor and late-replicating (Guelen et al. 2008; Yokochi et al. 2009; Lienert et al. 2011). Furthermore, G9a and GLP are required for repression of a number of late-replicating genes (Yokochi et al. 2009), including several gene clusters on the X-chromosome (Shinkai and Tachibana 2011), as well as specific families of retroelements (Maksakova et al. 2013). Interestingly, a recent report examining randomly integrated reporter constructs in ESCs revealed that transcription permissive integration sites cluster in large Mb-scale domains that are anti-correlated with H3K9me2 and LADs, indicating that the transcription potential of transgenes is dependent on chromatin features of the integration site (Akhtar et al. 2013). However, whether H3S10ph and H3K9me2 are mutually antagonistic in interphase mammalian cells has not been addressed, due largely to contaminating AURKB-dependent H3S10ph in mitotic cells.

To examine the crosstalk between these PTMs in mammalian cells, I first sought to comprehensively map H3S10ph in interphase ESCs. Using the FUCCI system for purification of

live cells based on cell cycle stage (Sakaue-Sawano et al. 2008) and chromatin immunoprecipitation sequencing (ChIP-seq), I found that H3S10ph marks Mb-scale domains encompassing gene-dense regions in interphase. Surprisingly, while actively transcribed genes are uniformly marked with H3S10ph in G1, enrichment of this mark is generally correlated with domains of early replication timing (RT). In contrast, genomic regions depleted of H3S10ph are generally gene poor, late replicating and enriched in LADs and H3K9me2. Consistent with the phospho/methyl switch model, expression of a non-phosphorylatable S10A H3.3 variant resulted in aberrant accumulation of H3K9me2 in gene-dense regions. Conversely, in the absence of GLP, H3S10ph domains expand into late-replicating regions marked with H3K9me2 in WT ESCs. The expansion of H3S10ph is accompanied by aberrant transcription of cryptic promoters near the boundaries of early and late replicating regions, defined previously as timing transition regions (TTRs) (Pope et al. 2014), with a clear bias towards transcription of the template strand used during leading strand DNA synthesis. H3S10ph and H3K9me2 are also anti-correlated in MEFs; however, H3S10ph domains cover only transcribed gene bodies in interphase MEFs, while H3K9me2 covers intergenic regions within gene-rich domains. Taken together, these observations reveal that pervasive marking of early replicating regions by H3S10ph may be a distinguishing feature of ESCs, and that H3S10ph and H3K9me2 likely have opposing functions in transcriptional regulation in interphase mammalian cells.

3.2 Results

3.2.1 H3S10ph in interphase ESCs

AURKB activity is recognized as a canonical marker of mitotic chromosomes, initiating phosphorylation of H3S10ph in G2/M at pericentromeric heterochromatin and gradually

modifying entire chromosomes as mitosis progresses. We utilized a previously characterized H3S10ph-specific antibody (CMA311) that is insensitive to the presence of H3K9 mono- or di-methylation (Hayashi-Takanaka et al. 2009), but is occluded by H3K9 acetylation or trimethylation, to profile the distribution of H3S10ph in ESCs. Staining of asynchronous ESCs cultures with CMA311 reveals intense signal at chromocenters in G2 and condensed chromosomes in mitotic cells (**Fig. 3.1A**). However, abundant but faintly stained small H3S10ph foci are also clearly present in euchromatic and nucleolar compartments of interphase ESCs (**Fig. 3.1A**), consistent with previous studies using polyclonal or monoclonal H3S10ph antibodies (Fazzio and Panning 2010; Sassone-Corsi et al. 1999; Hayashi-Takanaka et al. 2009). While considerable heterogeneity in interphase H3S10ph signal is apparent across ESC nuclei, presumably reflecting dynamic regulation of this labile PTM, these observations indicate that H3S10ph likely marks numerous genomic regions in interphase ESCs.

To generate high-quality maps of H3S10ph in interphase cells in the absence of any confounding effects associated with chemical synchronization, I employed the dual-colour FUCCI cell cycle reporter system (**Fig. 3.1B**), which allows for cell sorting based on fluorescent tagging of the labile cell cycle regulated proteins Cdt1 and Geminin (Sakaue-Sawano et al. 2008). I derived a stable polyclonal ESC line expressing hCdt1-monomeric Kusabira Orange (mKO) and hGeminin-Venus from a multi-cistronic construct separated by T2A cleavage tag (Falconer et al. 2012) and validated the fidelity of these FUCCI fluorescent tags as cell cycle stage reporters in ESCs using Hoechst 33342 DNA counterstain (**Fig. 3.2A, 2B**). As expected, the $mKO^{+}Venus^{-}$ subpopulation of these "FUCCI ESCs" is enriched for cells in G1, whereas the $mKO^{-}Venus^{+}$ subpopulation includes cells in S, G2 and M phases. I then sorted mKO^{+},

mKO^{+}Venus^{+}, and Venus^{+} cells to > 90% purity and 10^{6} nuclei, representing G1, S, S/G2/M stage fractions, respectively, were crosslinked and MNase-digested followed by ChIP with the aforementioned H3S10ph antibody. Quantification of unamplified nucleosomal DNA recovered by ChIP revealed a ~6-fold higher level of H3S10ph in the S/G2/M fraction (**Fig. 3.2C**), reflecting the widespread deposition of this mark in mitotic cells. However, consistent with fluorescence microscopy and western analysis of synchronized HeLa cells (Rothbart et al. 2015), the presence of nucleosomal DNA in chromatin isolated from G1 and S-phase cells indicates that H3S10ph is also present in interphase cells, albeit at lower levels. Indeed, ChIP-qPCR of several randomly chosen loci representing both genic and intergenic regions reveals that these regions are enriched for H3S10ph in G1 sorted cells (**Fig. 3.2D**). Depending on the region analyzed, the percentage of input DNA immunoprecipitated ranged from ~8- 50%, which is well above the mitotic index of ESCs, ruling out the possibility that such enrichment is simply a result of contamination of the sorted G1 or S-phase populations with mitotic cells. G1-sorted cells showed higher levels of enrichment of H3S10ph than S-phase sorted cells, consistent with nascent histone deposition during DNA replication. While H3S10ph enrichment is also detected in the S/G2/M-sorted cells by ChIP-qPCR at all tested loci, the absolute levels of enrichment are significantly lower than observed in G1, likely due to saturation of the antibody by the pervasive phosphorylation of H3S10 during mitosis.

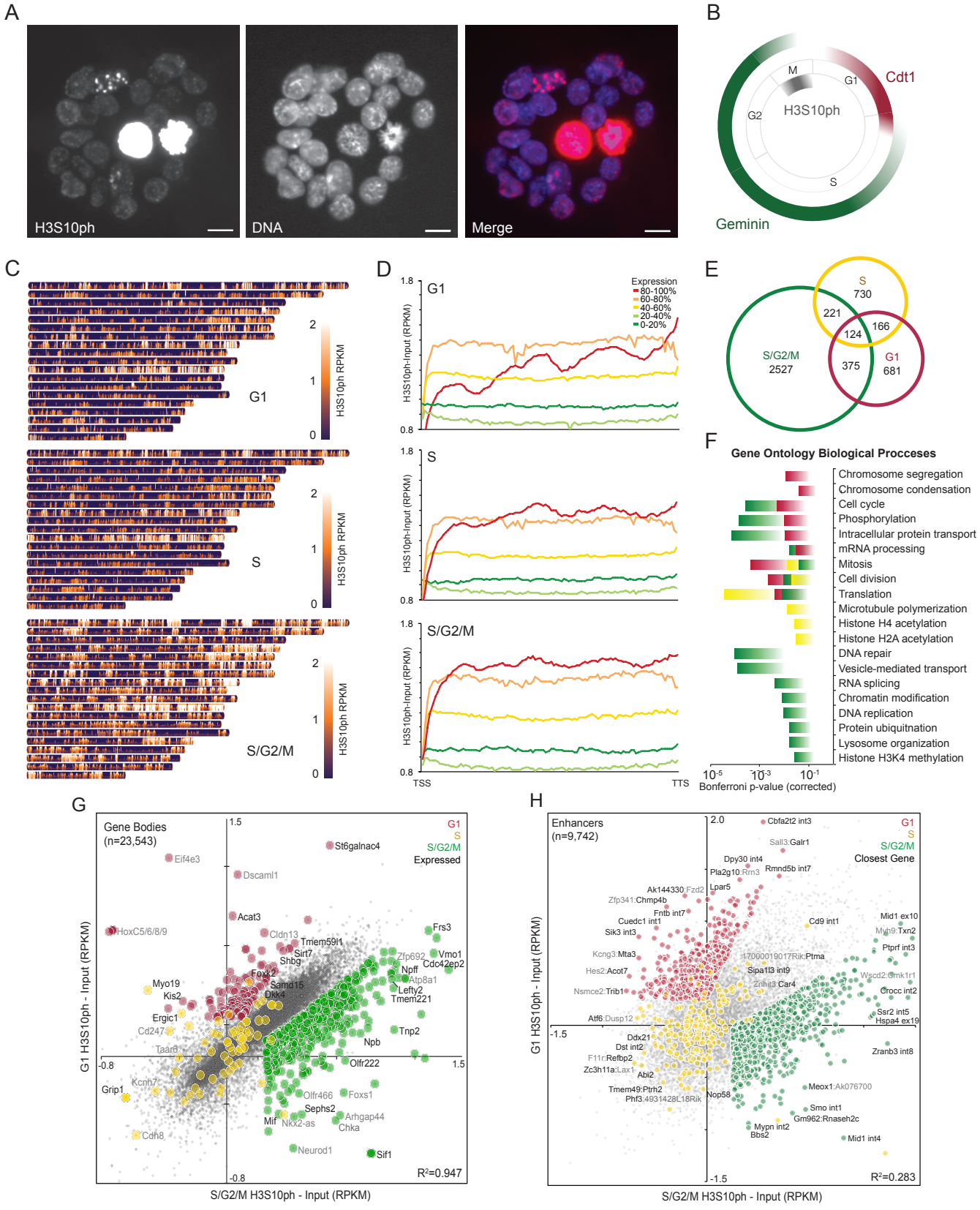

A
H3S10ph
DNA
Merge
B
Cdt1
M
G1
G2
H3S10ph
S
Geminin
C
G1
S
S/G2/M
H3S10ph RPKM
D
G1
S
S/G2/M
H3S10ph-Input (RPKM)
Expression
80-100%
60-80%
40-60%
20-40%
0-20%
TSS
TTS
E
S
730
221
124
166
S/G2/M
2527
375
G1
681
F
Gene Ontology Biological Procceses
Chromosome segregation
Chromosome condensation
Cell cycle
Phosphorylation
Intracellular protein transport
mRNA processing
Mitosis
Cell division
Translation
Microtubule polymerization
Histone H4 acetylation
Histone H2A acetylation
DNA repair
Vesicle-mediated transport
RNA splicing
Chromatin modification
DNA replication
Protein ubiquitnation
Lysosome organization
Histone H3K4 methylation
Bonferroni p-value (corrected)
G
Gene Bodies
(n=23,543)
G1
S
S/G2/M
Expressed
G1 H3S10ph - Input (RPKM)
S/G2/M H3S10ph - Input (RPKM)
R²=0.947
H
Enhancers
(n=9,742)
Closest Gene
G1 H3S10ph - Input (RPKM)
S/G2/M H3S10ph - Input (RPKM)
R²=0.283

Figure 3.1 Genome-wide characterization of cell-cycle-specific H3S10ph.
A. H3S10ph immunostaining of asynchronous WT TT2 ESCs, counterstained with Hoechst 33342. 10 μm scale bar. **B**. Schematic of the FUCCI reporter system for dissection of cell cycle specific H3S10ph. Cdt1-mKO and Geminin-Venus are fluorescent markers of G1 and S/G2/M, respectively. **C**. Heatmap depicting H3S10ph enrichment across all murine autosomes in G1, S and S/G2/M sorted fractions. **D**. Metagene analysis of normalized H3S10ph enrichment in ESCs (input-subtracted RPKM) over all ENSEMBL annotated gene bodies, clustered by expression levels (in quintiles). **E**. Venn diagram of genes enriched for H3S10ph (> 2.0-fold over input in gene bodies) in G1, S and S/G2/M sorted fractions. **F**. Gene Ontology of Biological Processes from genes identified as in Fig. 3.1E (corrected $p < 0.05$). **G**. Pair-wise comparison of H3S10ph enrichment (input-subtracted RPKM) over all gene bodies in G1 vs. S/G2/M sorted fractions. Colored dots represent genes with cell cycle skewing ($z > 0.5$). R^2 = Pearson correlation. **H**. Pair-wise comparison of H3S10ph enrichment (input-subtracted RPKM) over annotated enhancer regions in G1 vs. S/G2/M sorted fractions. Colored dots represent sites with cell cycle skewing as in panel G. A subset of enhancers is labelled with the overlapping gene(s), ex: exon, int: intron, or flanking gene(s) with the closest gene in black.

3.2.2 Cell cycle dynamics of interphase H3S10ph

Evidence for the presence of H3S10ph in interphase prompted us to survey the landscape of H3S10ph genome-wide via ChIP-seq in G1, S and S/G2/M purified ESCs. Consistent with bulk chromatin quantification, S/G2/M cells show widespread H3S10ph enrichment relative to input (**Fig. 3.1C**), reflecting the global accumulation of H3S10ph during mitosis. Notably, a significant fraction of the genome was also enriched with H3S10ph relative to input in G1 and S phases (**Fig. 3.2E**), indicating that this mark may play a broader role in transcriptional regulation and/or other extra-mitotic chromatin-related pathways than previously recognized. Input samples showed consistent recovery across the cell cycle, indicating that the dynamic profile of H3S10ph is unlikely to be an artifact of differential efficiency of nucleosome recovery in these cell cycle fractions.

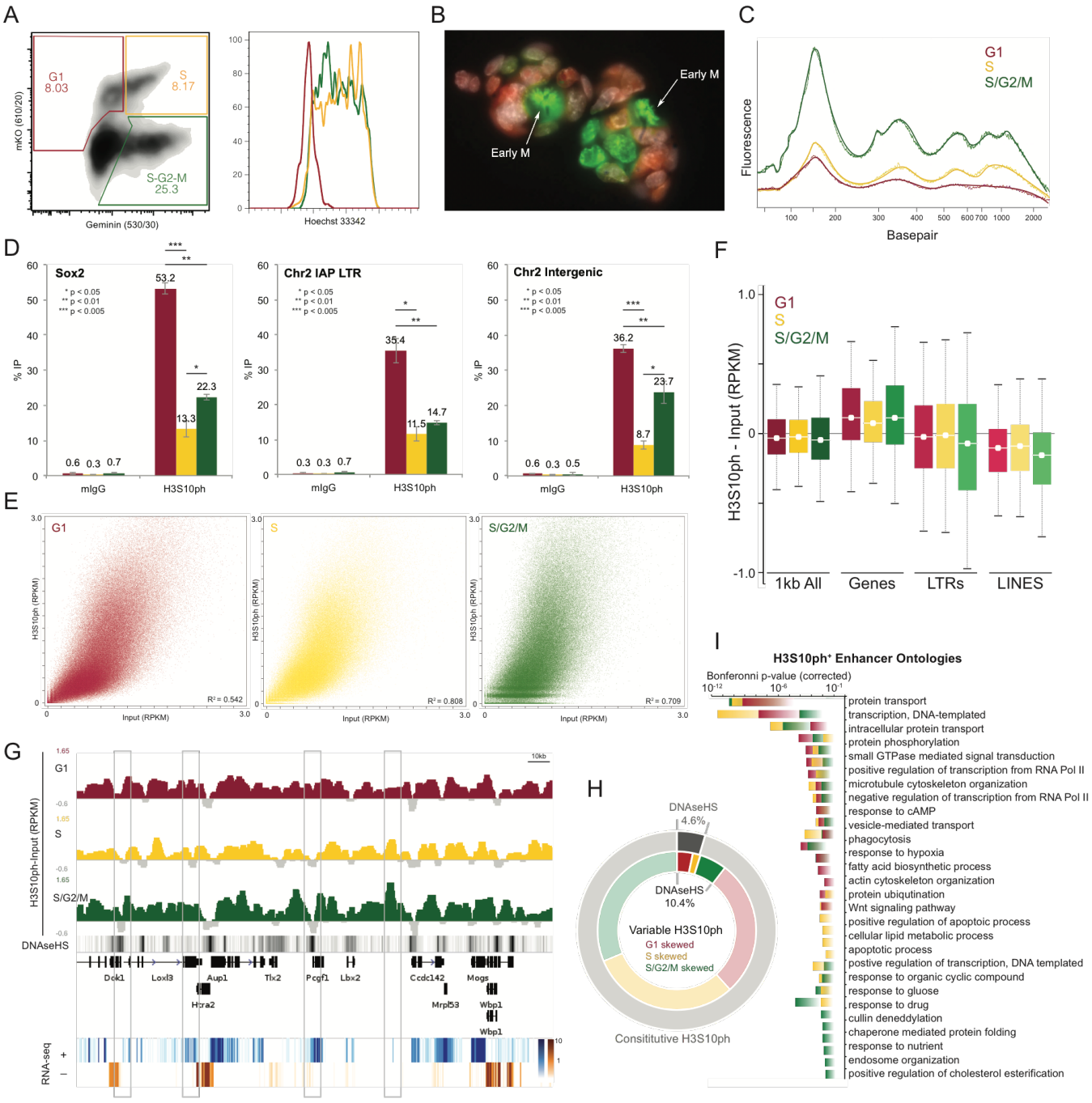


Figure 3.2 Genome-wide characterization of cell cycle specific H3S10ph (supplemental to Fig. 3.1). **A.** Flow cytometry analysis of ESCs stably transfected with hCdt1-mKO and hGeminin-Venus FUCCI constructs. Gates for G1, S and S-G2-M fractions are shown, along with the percentage of fluorescent cells in each fraction. Right panel histograms show Hoescht 33342 staining profiles for each gated fraction. **B.** Fluorescence microscopy of FUCCI ESCs showing Cdt1-mKO2 (red) and Geminin-Venus (green) fluorescence, counterstained with Hoescht 33342 (white). Arrows highlight mitotic cells with condensed chromosomes expressing Geminin-Venus. **C.** Analysis of the size distribution of unamplified DNA recovered from G1, S and S-G2-M fractions following H3S10ph ChIP quantitated using capillary electrophoresis. **D.** H3S10ph ChIP-qPCR of amplicons at the *Sox2*

promoter, a specific IAP LTR on Chromosome 2 (chr2) and an intergenic region on chr2. Enrichment levels are normalized to total input DNA. ChIP with mIgG was conducted in parallel as a negative control. **E.** Pairwise comparison of H3S10ph enrichment versus matched input DNA genome-wide using 5kb tiling bins. R^2: Pearson correlation coefficient. **F.** H3S10ph preferentially marks gene bodies at all cell-cycle stages, including in G1. Box and whisker plot of cell cycle specific H3S10ph enrichment (RPKM, normalized to input) genome-wide (1kb bins), at all gene bodies, at non-genic Long Terminal Repeats (LTRs), and at non-genic LINE elements. **G.** Genome-browser screenshot of the *Pcgf1* locus and flanking genes. Input corrected H3S10ph tracks for G1, S and S-G2-M fractions are shown, as well as strand-specific RNA-seq, revealing persistent H3S10ph over gene-bodies of actively transcribing genes as well as in intergenic regions. Boxed regions represent Dnase I hypersensitive sites (DNaseHS) (Yue et al. 2014) that exhibit dynamic H3S10ph enrichment during the cell cycle. **H.** Total number of sites overlapping DNAseHS sites in cell cycle skewed H3S10ph enrichment, and regions showing constitutive/invariant H3S10ph enrichment. Regions that show cell cycle-variable H3S10ph are overrepresented for DNAseHS compared to constitutively marked regions. **I.** Biological Process ontology analysis of genes physically interacting with an enhancer that is skewed for cell cycle specific H3S10ph identified as in Fig. 3.1H.

To characterize the genomic distribution of H3S10ph in greater detail, I determined whether H3S10ph marks distinct DNA sequence features across cell cycle fractions by comparing input- normalized H3S10ph at genes and repetitive sequences. H3S10ph was relatively more enriched at genes than repetitive sequences across all cell cycle stages, as measured in 1kb bins (**Fig. 3.2F**). Metagene analysis of H3S10ph enrichment in FUCCI-sorted ESCs revealed that H3S10ph broadly marks gene-bodies in interphase and is generally positively correlated with mRNA levels (**Fig. 3.1D, Fig. 3.2G**). However, the most highly expressed quintile of genes shows moderately lower levels of enrichment than the 2nd quintile; this pattern is most evident towards the 5' end of genes, likely due to the high level of nucleosomal turnover at such highly expressed genes. The apparent depletion of H3S10ph at TSSs may reflect epitope masking by H3K9ac (Hayashi-Takanaka et al. 2009), a mark generally detected in the promoter regions of active genes. In addition, H3S10ph levels in the intragenic regions of expressed genes were modestly lower in S-phase relative to G1 and G2/M cells (**Fig. 3.2D-E**), presumably due to replication-coupled deposition of nascent unmodified H3 during early S phase and re-establishment of H3S10ph, in G2/M.

To identify genes that may be differentially marked at distinct cell-cycle stages, I analyzed enrichment levels across ENSEMBL annotated genes in the G1 versus S/G2/M datasets. Notably, 77% of all annotated genes showed H3S10ph enrichment relative to the input control in G1. While H3S10ph levels at most gene bodies were highly correlated across the cell cycle (R^2= 0.947), a minority of genes show clear cell cycle-specific skewing of H3S10ph enrichment. To ascertain the function of those genes showing a cell cycle enrichment bias, genes with the highest levels of H3S10ph enrichment were identified (≥ 2-fold enriched over matched input RPKM) and Gene Ontology (GO) analysis was conducted. S/G2/M cells include a much larger subset of differentially marked genes (n = 3,247) than G1 (n = 1,346) or S (n = 1,241) cells (**Fig. 3.1E**), as expected given that H3S10ph levels increase dramatically with the onset of mitosis. Across all cycle stages, heavily phosphorylated genes are enriched (Bonferroni-corrected p-value < 0.05) in housekeeping functions, specifically cell cycle-regulated biological processes (**Fig. 3.1F**). In contrast, genes enriched with H3S10ph in G1 were significantly over-represented in mitosis-related GO terms, such as chromosome condensation and segregation, whereas genes with functions in DNA repair as well as DNA replication were prominent in S/G2/M. Furthermore, several GO terms were enriched in S compared to G1 and S/G2/M, including translation, microtubule polymerization and histone acetylation. Taken together, these observations suggest a link between enrichment of H3S10ph in interphase and transcriptional activity, which for at least a subset of genes is apparently dynamically regulated during the cell cycle (Sasagawa et al. 2013).

To further characterize regions showing cell cycle-variable H3S10ph enrichment in ESCs, I first scored genomic bins that exhibit skewed H3S10 phosphorylation and found that 10.4% of these differentially phosphorylated regions are within 200 bp of a DNase I hypersensitive site (DHS), compared to 4.6% overlapping DHS in constitutively H3S10 phosphorylated regions (**Fig. 3.2H**). To determine whether sites exhibiting dynamic H3S10ph include putative enhancers, ENCODE DHS coordinates were filtered for regions enriched for H3K4me1 and H3K27ac, a signature of active enhancers. Subsequent pair-wise comparisons of normalized H3S10ph in G1 vs. S/G2/M populations at the 9,742 active enhancers identified in ESCs yielded much greater cell cycle variability in H3S10ph enrichment (R^2=0.250) than observed in gene-bodies (**Fig. 3.1G-H**), with 1,044 sites showing strong stage-specific H3S10ph enrichment (normalized RPKM > 0.5, z-score > 0.5) across G1, S and S/G2/M, and 595 enhancer regions that were consistently marked in all 3 fractions. To annotate gene targets of cell cycle-biased distal enhancers in ESCs, I parsed enhancers that physically contact promoters, as determined by Chromatin Interaction Analysis by Paired-End Sequencing (Sahlén et al. 2015). Gene Ontology analysis of the putative gene targets revealed enrichment for housekeeping functions, as observed for GO analysis of H3S10ph-marked gene bodies (**Fig. 3.2I**), with the distinct emergence of fatty acid biosynthesis in G1, DNA replication and protein ubiquitination in S, and response to glucose, nutrient and drugs in S/G2/M. Taken together, these observations reveal that across the cell cycle, H3S10ph deposition and/or removal is dynamic in gene bodies as well as distal regulatory regions, perhaps reflecting transcriptional regulation at these stages.

3.2.3 Interphase H3S10ph is not dependent on AURKB

Curiously, as opposed to other PTMs positively associated with transcriptional activity, inspection of individual loci reveals that interphase H3S10ph is not limited to gene bodies or enhancers, but also broadly marks intergenic regions (**Fig. 3.2G**). Indeed, H3S10ph at gene-dense regions encompasses Mb-scale domains in interphase, covering approximately 30% of the genome (**Fig. 3.4A**). To determine whether these broad domains reflect incomplete dephosphorylation of H3S10ph deposited by AURKB during mitosis, I treated ESCs with Hesperadin (Hesp), an inhibitor of AURKB/c kinases (Fischle et al. 2005). As expected, a 3h treatment with 200nM Hesp effectively eliminated > 98% of mitotic cells, as measured by DNA content and H3S10ph signal (**Fig. 3.3A**). Nevertheless, quantitative Western blotting revealed that Hesp treated cells retain ~60% of the level of H3S10ph in control asynchronous ESCs, likely reflecting AURKB independent activity (**Fig. 3.3B**). Furthermore, H3S10ph ChIP-seq on control and Hesp treated ESCs revealed reduced global enrichment of H3S10ph following AURKB inhibition, as expected, but increased prominence of putative interphase H3S10ph domains, recapitulating the profile observed in G1-sorted cells (**Fig. 3.4B**). Based on these observations, I conclude that the euchromatic H3S10ph domains observed in interphase ESCs are deposited by a serine kinase or kinases other than AURKB.

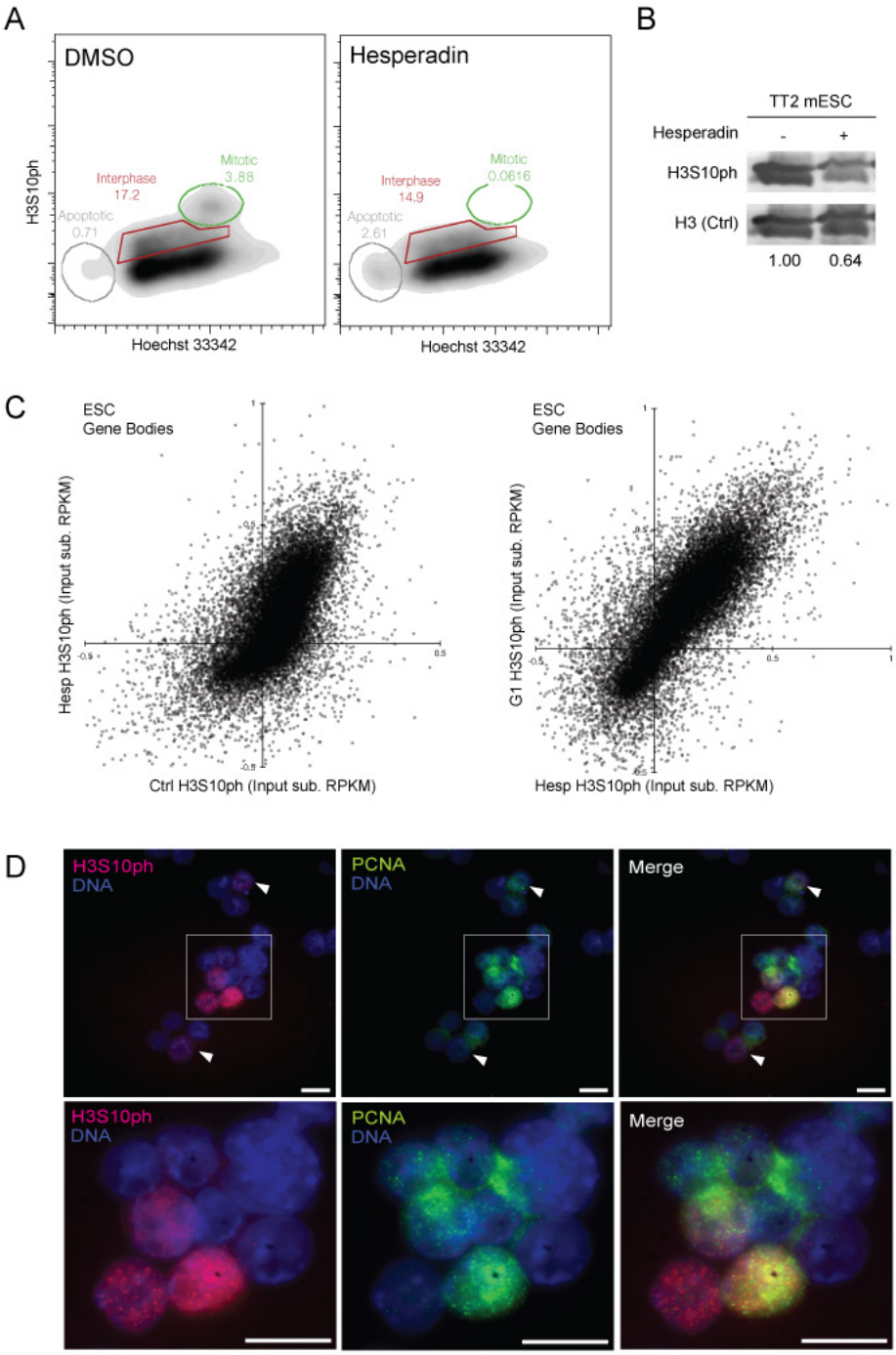


Figure 3.3 ***Interphase H3S10ph is not dependent on Aurora kinase B (supplemental to Fig. 3.4).***
A. Flow cytometry analysis of control (DMSO) and Hesp-treated ESCs stained with H3S10ph and Hoechst 33342. The percentage of interphase and mitotic cells, as determined by H3S10ph staining and DNA content, is also shown. Note the absence of H3S10ph+ mitotic cells in the Hesp-treated sample, as expected following culture with this AURKB/c-specific inhibitor. **B.** Western analysis of H3S10ph and H3 (loading control) in ESCs treated with 200nm Hesperadin or DMSO (vehicle control). Levels of each were quantitated using the Odyssey (LICOR) platform and

the ratio of H3S10ph/H3 signal is presented. **C.** Comparison of enrichment of H3S10ph (Input subtracted RPKM) over all gene bodies in Hesp-treated ESCs versus asynchronous (Ctrl) or G1-sorted FUCCI ESCs. **D.** Immunofluorescence analysis of WT TT2 ESCs stained with H3S10ph (red), PCNA (green), and Hoechst 33342 (blue). Note that a subset of interphase cells negative for the S-phase marker PCNA show H3S10ph staining.

3.2.4 H3S10ph and H3K9me2 partition the genome into early- and late-replicating domains

To further characterize the chromatin state of regions marked with H3S10ph in interphase, I intersected the genome-wide distribution of this mark with additional chromatin features in ESCs, including those generated by the ENCODE consortium (Yue et al. 2014) (**Table 3.1**). Given that H3S10ph domains in interphase ESCs were relatively broad, I used a tiling bin size of 5kb. As expected, H3S10ph shows a positive correlation with RNA Pol II, transcript level (RNA-seq) and transcription associated PTMs (H3K4me3, H3K9ac, H3K36me3) as well as enhancer PTMs (H3K4me1, H3K27ac) (**Fig. 3.4C**). However, consistent with the observation that H3S10ph is also enriched in intergenic regions of gene-rich genomic domains, 42% of interphase H3S10ph marked bins do not coincide with annotated genes. Notably, these intergenic H3S10ph-marked regions are almost universally replicated relatively early in S-phase, with boundaries of interphase H3S10ph domains precisely aligning with TTRs (**Fig. 3.4B**). Indeed, RT was concordant with H3S10ph enrichment (r2= 0.462) genome-wide (**Fig. 3.4D**), and H3S10ph in G1 cells scales quantitatively with RT progression (**Fig. 3.4E**). To determine whether H3S10ph in interphase is specific to cells undergoing DNA replication, I co-stained proliferating ESCs with H3S10ph and the S-phase marker PCNA (**Fig. 3.3D**). While replicating/PCNA+ nuclei were generally marked with punctate H3S10ph foci, PCNA- nuclei also harboured punctate H3S10ph foci, consistent with the observation that H3S10ph is present in G1 cells, i.e. prior to the onset of DNA replication.

Table 3.1 Published datasets used for meta-epigenomic analysis used in Kendall correlation matrix.

Cell Type	ChIP / RNA	GEO Accession#	Citation
ESC	CTCF	GSM918748	(Yue et al. 2014)
ESC	DNase HS	GSM1014154	(Yue et al. 2014)
ESC	H3.3-HA	GSM1429928	(Elsässer et al. 2015)
ESC	H3K27ac	GSM1000126	(Yue et al. 2014)
ESC	H3K27me3	GSM307619	(Mikkelsen et al. 2007)
ESC	H3K36me3	GSM1000125	(Yue et al. 2014)
ESC	H3K4me1	GSM1000121	(Yue et al. 2014)
ESC	H3K4me3	GSM1000124	(Yue et al. 2014)
ESC	H3K9ac	GSM1000123	(Yue et al. 2014)
ESC	H3K9me3	GSM1551524	(Brind'Amour et al. 2015)
ESC	Lamin B1	GSM426758	(Peric-Hupkes et al. 2010)
ESC	mRNA	GSM929718	(Yue et al. 2014)
ESC	p300	GSM918750	(Yue et al. 2014)
ESC	Pol II	GSM918749	(Yue et al. 2014)
ESC	RT	GSM450274	(Hiratani et al. 2010)
MEF	CTCF	GSM918743	(Yue et al. 2014)
MEF	H3K27ac	GSM1000139	(Yue et al. 2014)
MEF	H3K27me3	GSM307609	(Mikkelsen et al. 2007)
MEF	H3K36me3	GSM307610	(Mikkelsen et al. 2007)
MEF	H3K4me3	GSM307608	(Mikkelsen et al. 2007)
MEF	H3K9me2	GSM887877	(Fang et al. 2012a)
MEF	H3K9me3	GSM307611	(Mikkelsen et al. 2007)
MEF	H3Kme1	GSM769028	(Yue et al. 2014)
MEF	Lamin B1	GSM426764	(Peric-Hupkes et al. 2010)
MEF	mRNA	GSM929719	(Yue et al. 2014)
MEF	Pol II	GSM918761	(Yue et al. 2014)
MEF	RT	GSM450291	(Hiratani et al. 2010)

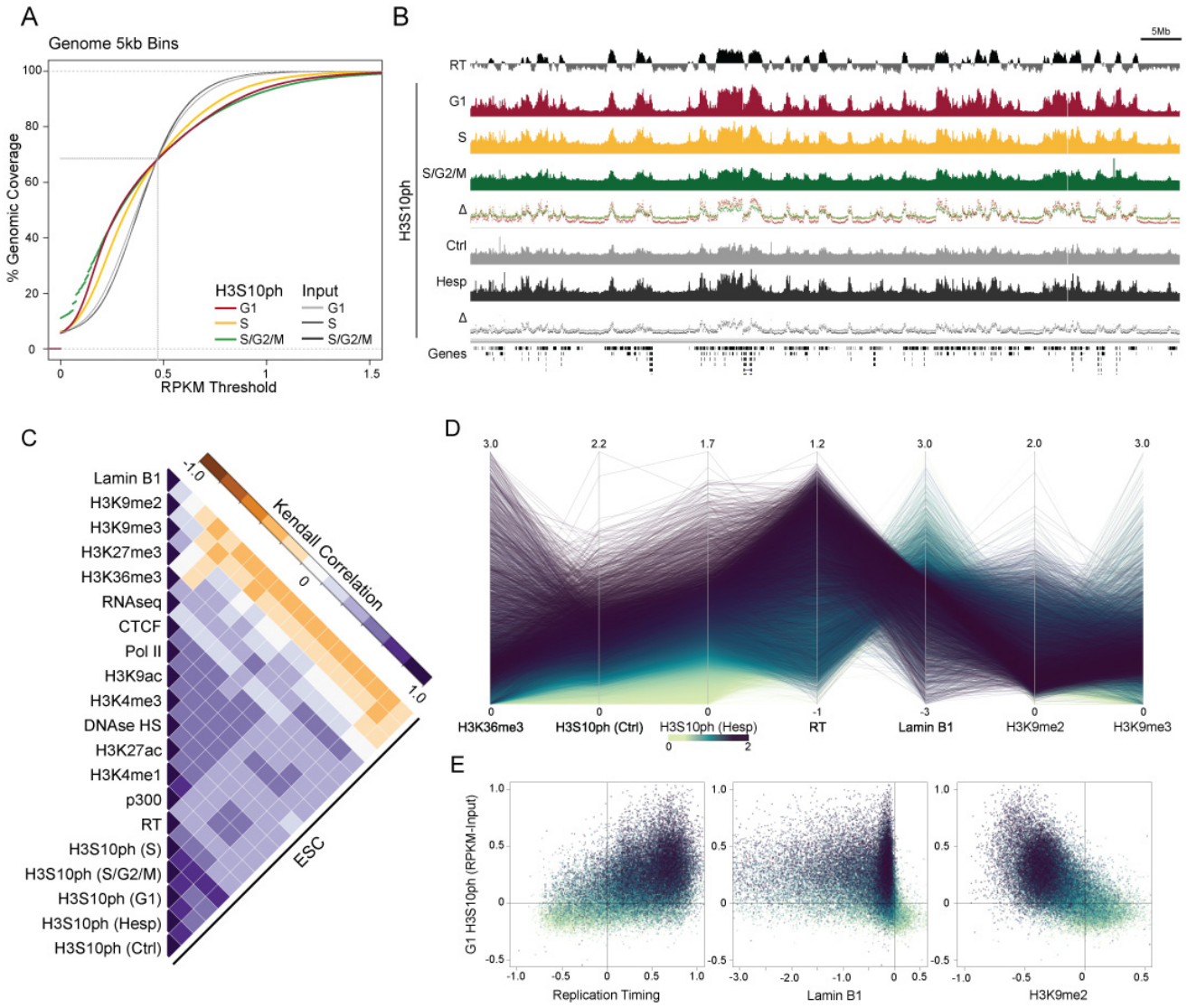


Figure 3.4 Interphase H3S10ph forms domains highly concordant with replication timing.
A. Cumulative distribution frequency plot of ESC H3S10ph ChIP vs. Input libraries in 5kb bins. **B**. Genome browser screenshot of H3S10ph enrichment and replication timing (RT) (Hiratani et al. 2010) in TT2 ESCs across chr18. **C**. Kendall correlation matrix of H3S10ph with chromatin features (Table 3.1) across 5kb genomic bins in ESCs, following unsupervised hierarchical clustering. H3S10ph ESC datasets include sorted FUCCI subpopulations (S, S/G2/M and G1), Hesp treated and vehicle control (Ctrl). **D**. Parallel coordinate comparison of H3S10ph enrichment (RPKM) with histone marks H3K36me3, H3K9me2 and H3K9me3 as well as Lamin B1 and RT in randomly sampled 5kb bins. Heat map colour-coding is based on H3S10ph enrichment in Hesp treated ESCs. Note the strong positive correlation between H3S10ph in interphase and RT and the negative correlation with H3K9me2. **E**. Pair-wise comparisons of G1 H3S10ph enrichment with RT, LADs and H3K9me2 (input subtracted RPKM) enrichment. Randomly sampled 5kb bins are coloured as in D.

In contrast, regions enriched for H3S10ph in G1 show a clear discordance with lamin-associated domains (LADs) (**Fig. 3.4C, D and E**), shown previously to replicate relatively late in S-phase (Kind et al. 2015). Regions enriched for H3S10ph are also anti-correlated with H3K9me2, a mark deposited by the KMTs G9a and GLP at late-replicating genomic regions (Yokochi et al. 2009; Kind and van Steensel 2010) and to a lesser extent H3K9me3, which marks constitutive heterochromatin and specific retrotransposons in ESCs (Karimi et al. 2011). As H3K9me2 is specifically associated with LADs (Kind et al. 2013), I focused on the relationship between H3S10ph and H3K9me2. Consistent with the Mb-scale domains of H3K9me2 observed in ESCs via ChIP-chip (Lienert et al. 2011), ChIP-seq revealed that H3K9me2 demarcates large Mb domains, covering ~40% of the mappable genome. As predicted, this mark is clearly anti-correlated with interphase H3S10ph (r^2=0.787, 5kb bins) (**Fig. 3.4E**), indicating that H3S10ph may antagonize H3K9me2 deposition, and vice versa, in interphase ESCs. While the apparent anti-correlation may reflect epitope masking of the H3K9me2 antibody by H3S10ph (Duan et al. 2008; Rothbart et al. 2015), such a technical artifact is unlikely, as H3S10ph covers ~30% of the genome and quantitative mass spectrometry analysis of WT ESCs revealed that under 20% of bulk H3 is dimethylated at H3K9 (Kubicek et al. 2007).

3.2.5 Expression of non-phosphorylatable H3.3 promotes aberrant accumulation of H3K9me2

To determine whether H3S10ph influences H3K9me2 deposition, I generated ESC lines stably expressing tagged H3.3 and H3.3S10A, which mimics the loss of phosphoserine (**Fig. 3.5A**). I reasoned that overexpression of the non-phosphorylatable H3.3 variant would mitigate the likelihood of mitotic defects, as H3.3 is deposited in a replication-independent pathway at

specific genomic regions, including transcribed gene bodies (Ahmad and Henikoff 2002; Chen et al. 2013) - regions which are enriched for interphase H3S10ph. Indeed, two independent clonal lines, "S10A.1" and "S10A.2", which stably express the mutant H3.3-YFP transgene, showed mitotic progression similar to that of the line expressing the wild type (WT) H3.3-YFP transgene (**Fig. 3.5B, 3.5C**). To determine whether WT H3.3-YFP and H3.3S10A-YFP are targeted to intragenic regions of active genes, I performed ChIP-seq using a YFP-specific antibody. Metagene analysis revealed that the level of exogenous H3.3 enrichment in gene bodies increases with increasing transcription level (**Fig. 3.6A**), as expected. Importantly, WT and S10A mutant H3.3-YFP show similar enrichment over active gene-bodies and genome-wide (**Fig. 3.6A**), revealing that the incorporation of exogenous H3.3 is not dependent upon H3S10 *per se*. Furthermore, the genomic distribution of YFP-tagged H3.3 is similar to that of HA-tagged H3.3 (Elsässer et al. 2015), indicating that YFP-tagging does not perturb intragenic deposition of H3.3 (**Fig. 3.3B**).

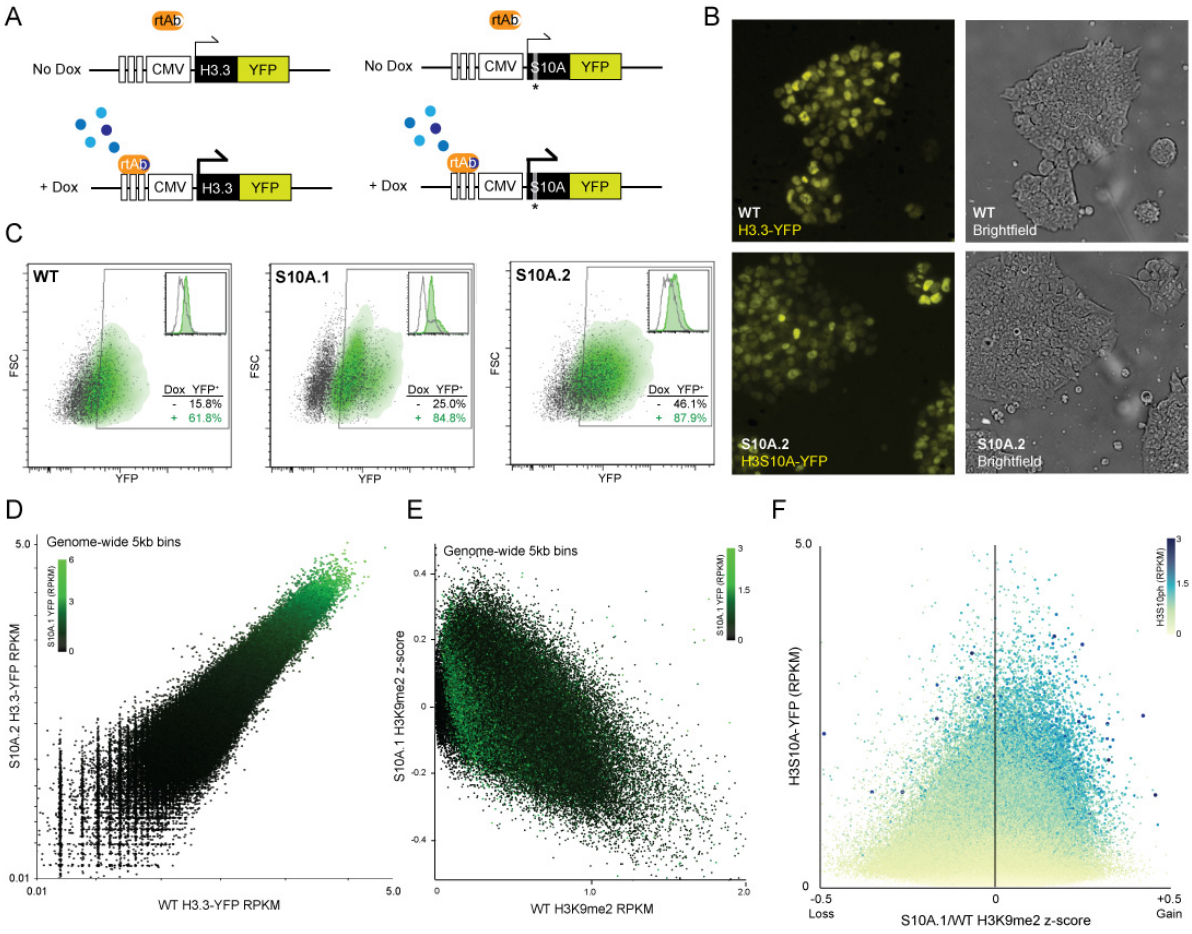


Figure 3.5 Overexpression of H3S10A in ESCs causes H3K9me2 accumulation in early-replicating regions (supplemental to Fig. 3.6).
A. Schematic of Dox inducible system for induction of H3.3-YFP (WT) and H3.3S10A-YFP. **B.** Live fluorescence imaging of ESC clones expressing H3.3-YFP and H3.3(S10A)-YFP. **C.** Flow cytometry analysis of YFP expression in WT, S10A.1 and S10A.2 lines with or without Dox induction. The percentage of GFP^+ cells is shown for each condition. H3.3S10A is incorporated at higher levels than WT H3.3. Greater heterogeneity in expression was observed in the clonal S10A.1 mutant line than the S10A.2 or WT H3.3 lines. **D.** Scatterplot of ChIP-seq data (5kb bins) showing enrichment of H3.3-YFP in WT versus S10A.2 expressing lines, with enrichment of H3.3-YFP in the clonal S10A.1 line overlaid in heat map format. **E.** Scatterplot comparing the genome-wide change in H3K9me2 (z-score) in the S10A.1 line versus H3K9me2 levels (RPKM) in WT cells. Both S10A clones show increased H3K9me2 at regions that are relatively H3K9me2-poor in WT ESCs. **F.** Scatterplot of YFP incorporation levels versus the change in H3K9me2 enrichment in the S10A.1 line (z-score). Note that accumulation of H3K9me2 as a direct result of incorporating H3S10A occurs primarily at regions that are normally H3S10ph-rich.

To determine whether incorporation of H3.3S10A influences the genomic distribution of H3K9me2, I performed H3K9me2 ChIP-seq, as above, on each H3.3-YFP expressing line. Strikingly, differences in the distribution of H3K9me2 are clearly apparent in the H3.3S10A-

YFP lines relative to the H3.3-YFP WT line (**Fig. 3.6B**). Analysis of the genome in 5 kb bins reveals a clear increase in the fraction of bins with intermediate levels of H3K9me2 in the mutant expressing lines (**Fig. 3.6C**), with 3.8% and 7.7% genome-wide bins showing increased H3K9me2 ($z > 0.2$) in S10A.1 and S10A.2 clones, respectively. In both H3.3S10A clonal lines, the ectopic gain of H3K9me2 at regions normally harboring low levels of H3K9me2 is accompanied by loss of this mark at regions normally showing high enrichment (**Fig. 3.6C-D, Fig. 3.5E**). The majority of regions showing such a gain in H3K9me2 enrichment may reflect indirect effects, since H3.3S10A-YFP signal was not detected over most of these regions. To identify regions that may accumulate H3K9me2 as a direct consequence of S10A-YFP incorporation, I filtered genomic bins showing a gain in this mark based on YFP enrichment level (>1 RPKM), which yielded 1,846 and 1,007 regions in S10A.1 and S10A.2 lines, respectively, with 40.2% overlap between these biological replicates. Importantly, in the regions identified as "direct targets", the majority (92%) reside within H3S10ph-rich domains in WT cells (**Fig. 3.6E, Fig. 3.5F**). Thus, the disruption of H3S10 phosphorylation via incorporation of ectopic H3.3S10A leads to accumulation of H3K9me2 at regions enriched for H3S10ph in WT cells, indicating that interphase H3S10ph may insulate gene-dense regions from G9a/GLP-mediated H3K9me2 deposition.

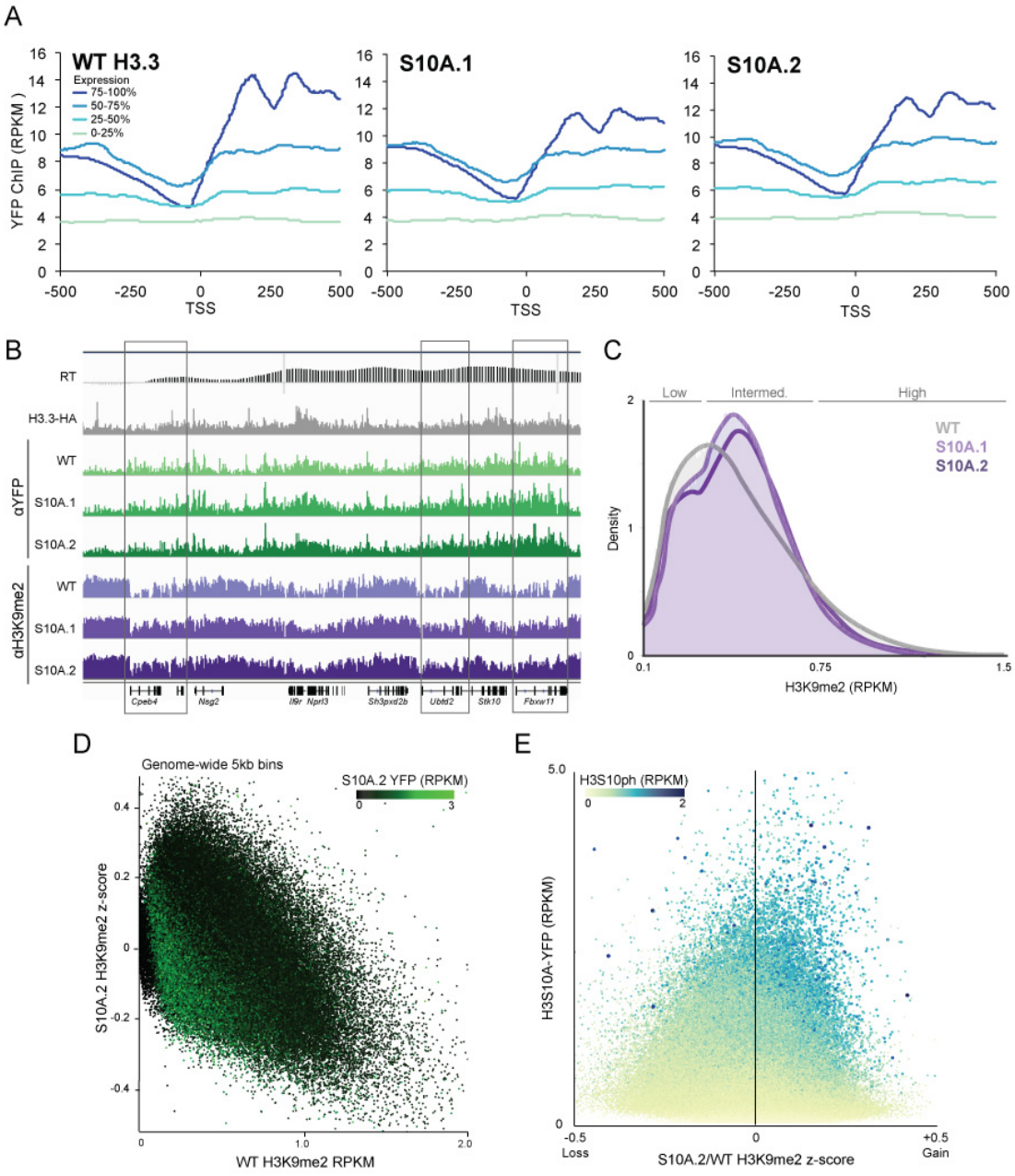


Figure 3.6 Overexpression of H3S10A in ESCs promotes H3K9me2 accumulation in early-replicating regions. **A**. ChIP enrichment of YFP (RPKM) around transcription start sites (TSS +/- 500bp) in WT and S10A H3.3-YFP expressing lines, clustered by genic expression quartiles. **B**. Genome browser screen shot spanning *Cpeb4* to *Fbxw11* loci on chromosome 11 (31,696,800 -32,668,200) with YFP and H3K9me2 ChIP enrichment in H3.3-YFP WT and S10A mutant lines presented, as well as previously published RT (Yokochi et al. 2009) and H3.3-HA (Elsässer et al. 2015) tracks. **C**. Histogram showing the global density distribution of H3K9me2 in H3.3-YFP WT and S10A lines. **D**. 2D scatterplot comparing the gain or loss (z-score) of H3K9me2 in the S10A.2 mutant line with WT H3K9me2 enrichment levels (RPKM). **E**. Comparison of YFP incorporation levels with the change in H3K9me2 enrichment in H3.3-YFP S10A.2 relative to WT cells (z-score). Heat map and data-point size represent the range of WT H3S10ph enrichment levels. Note that gain of H3K9me2 occurs predominantly within regions that are enriched for H3S10ph in WT ESCs.

3.2.6 Interphase H3S10ph domains expand upon loss of H3K9me2

To investigate whether H3K9me2 restricts the expansion of H3S10ph domains at TTRs, I performed ChIP-seq on previously characterized *Glp*$^{-/-}$ ESCs (Maksakova et al. 2013; Shinkai and Tachibana 2011). As expected, genome-wide depletion of H3K9me2 was observed (**Fig. 3.7A**), with 40% of genomic bins, generally overlapping with late-replicating domains, showing reduced enrichment ($z < -0.2$). Regions that retain H3K9me2 in *Glp*$^{-/-}$ ESCs are generally enriched for H3K9me3 in WT ESCs (**Fig. 3.7B**), likely reflecting the independent deposition of this mark by SETDB1 at retrotransposons (Karimi et al. 2011). Remarkably, 70% of genomic bins that lose H3K9me2 in Hesp-treated Glp$^{-/-}$ cells show a gain of H3S10ph (**Fig. 3.7A, 3.7C**). Notably, ectopic H3S10ph is clearly observed in regions adjacent to previously identified TTRs (Pope et al. 2014) (**Fig. 3.7D**), consistent with spreading of H3S10ph into gene-poor, relatively late-replicating regions normally marked by H3K9me2 in interphase.

3.2.7 Asymmetrical transcriptional initiation at TTRs in G9a$^{-/-}$ *and Glp*$^{-/-}$ *ESCs*

To determine whether ectopic H3S10ph at TTRs in H3K9me2 depleted ESCs is associated with changes in transcription, I conducted strand specific mRNA-seq in *Glp*$^{-/-}$ and *G9a*$^{-/-}$ ESCs as well as the control WT parent line. Transcripts observed exclusively in *G9a*$^{-/-}$ and *Glp*$^{-/-}$ ESCs are apparent in regions extending up to ~500 kb beyond the boundaries of the H3S10ph/gene-rich domains observed in WT cells (**Fig. 3.7A**). Strikingly, these aberrant transcripts show strong strand bias depending on the orientation of the replication fork (**Fig. 3.7A**), with reads mapping to the minus strand in late to early TTRs and to the plus strand in early to late TTRs. While meta-analysis of TTRs reveals that this asymmetry is present in WT ESCs, consistent with the

previously described co-orientation of genes with replication forks in human cells (Huvet et al. 2007), this phenomenon is greatly enhanced in the *Glp*$^{-/-}$ and *G9a*$^{-/-}$ lines (**Fig. 3.7E**). Transcripts upregulated in the vicinity of TTRs (+/- 250kb from the inflection point) include both repetitive sequences and genes, such as *MageA* and *Rhox* gene clusters (**Fig. 3.7F**), shown previously to be upregulated in *G9a*$^{-/-}$ and *Glp*$^{-/-}$ ESCs and *in vivo* (Shinkai and Tachibana 2011; Auclair et al. 2016). Taken together, these results indicate that G9a/Glp play a particularly important role in regulating transcription units co-oriented with the replication fork in TTRs and adjacent late-replicating regions.

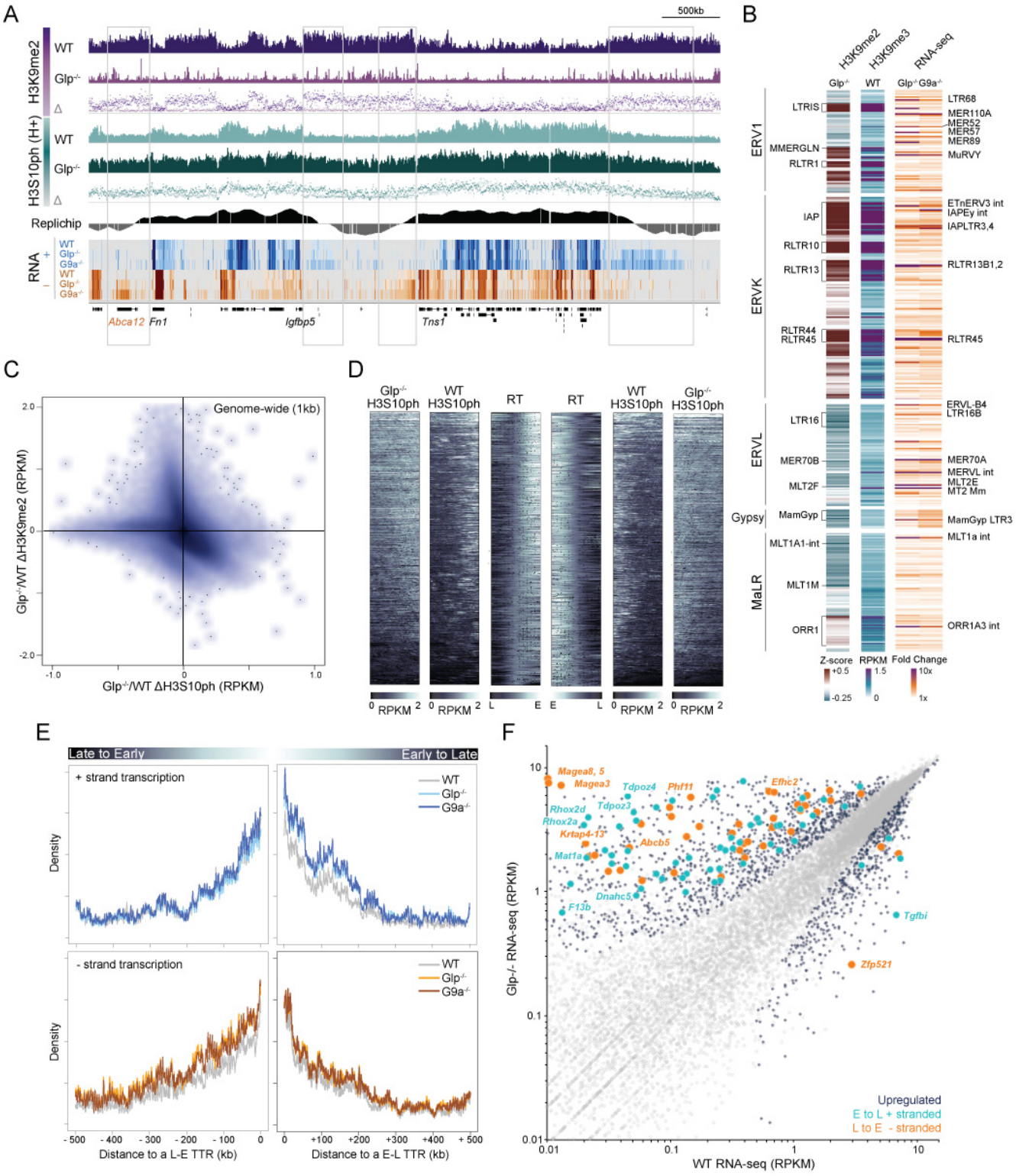


Figure 3.7 H3K9me2 restricts spreading of H3S10ph and inhibits aberrant transcription near TTRs.
A. Genome browser screenshot of a region on chromosome 1 showing H3K9me2, Hesp-treated (H+) H3S10ph and strand-specific RNA-seq in WT and *Glp*$^{-/-}$ ESCs. **B**. Heat maps of the change in H3K9me2 enrichment in *Glp*$^{-/-}$ vs.

WT (Z-score), H3K9me3 enrichment (RPKM) in WT and differential expression in *Glp*$^{-/-}$ vs. WT ESCs for all LTR transposable element subfamilies present at > 100 copies in the reference mouse genome. **C**. Scatterplot showing the changes in H3K9me2 and H3S10ph enrichment in Hesp-treated *Glp*$^{-/-}$ ESCs relative to WT. **D**. Heatmap of H3S10ph enrichment in Hesp-treated WT and *Glp*$^{-/-}$ ESCs over 500kb centred on all RT transition boundaries in WT ESCs (RT data from (Yokochi et al. 2009)). **E**. Cumulative density plot of RNA-seq coverage on the plus (top plot) and minus (bottom plot) strands at TTR boundaries (0-500kb) in WT, *Glp*$^{-/-}$ and *G9a*$^{-/-}$ ESCs. Note the increase in RNA coverage in the mutant lines on the plus strand in early to late transition regions and the minus strand in late to early transition regions. **F**. 2D scatterplot comparing genic transcription in *Glp*$^{-/-}$ and WT ESCs. Note that many of the genes up-regulated in *Glp*$^{-/-}$ ESCs reside within RT transition boundaries (0-500kb) and show a RT orientation strand-bias consistent with that described in panel E.

As G9a-dependent H3K9me2 was recently reported to mediate Pol II transcriptional termination via R-loop formation (Skourti-Stathaki et al. 2014), I considered the possibility that regions showing aberrant transcription may also be the result of inefficient termination of the last transcribed gene oriented towards the replication timing boundary within the early replicating domain. Analysis of H3K36me3 ChIP-seq data generated from *G9a*$^{-/-}$, *Glp*$^{-/-}$ and WT ESCs however, revealed similar profiles of reduced H3K36me3 downstream of annotated genic transcription termination sites (TTSs) (**Fig. 3.8A, B**), indicating that aberrant Pol II processivity beyond the TTS is unlikely to account for such ectopic transcription. On the other hand, ChIP-seq analysis of H3K4me3 in *G9a*$^{-/-}$ and *Glp*$^{-/-}$ ESCs revealed numerous ectopic peaks of H3K4me3 at TTRs in the mutant lines, likely reflecting aberrant transcription initiation (**Fig. 3.8A, C**). To identify candidate "*de novo"* initiation sites, I called all H3K4me3 peaks in WT and mutant lines (n = 55,000). 7,878 sites gained H3K4me3 in both *G9a*$^{-/-}$ and *Glp*$^{-/-}$ ESCs (R^2=0.992 between *G9a*$^{-/-}$ and *Glp*$^{-/-}$), with 1,081 common regions showing no enrichment in the WT control (**Fig. 3.8C**). While most of the shared and gained H3K4me3 peaks were found at annotated promoters*, de novo* sites were also frequently found within LTR and LINE sequences, and were significantly closer to TTRs than constitutive H3K4me3 peaks (**Fig. 3.8D-E**). As expected, *de novo* H3K4me3 peaks are frequently found adjacent to regions of increased RNA-seq coverage on the plus or minus strand (**Fig. 3.8F**), indicative of the presence of ectopic

promoters in *G9a*$^{-/-}$ and *Glp*$^{-/-}$ ESCs. Taken together, these observations indicate that transcription units near TTRs are particularly prone to aberrant initiation following loss of H3K9me2.

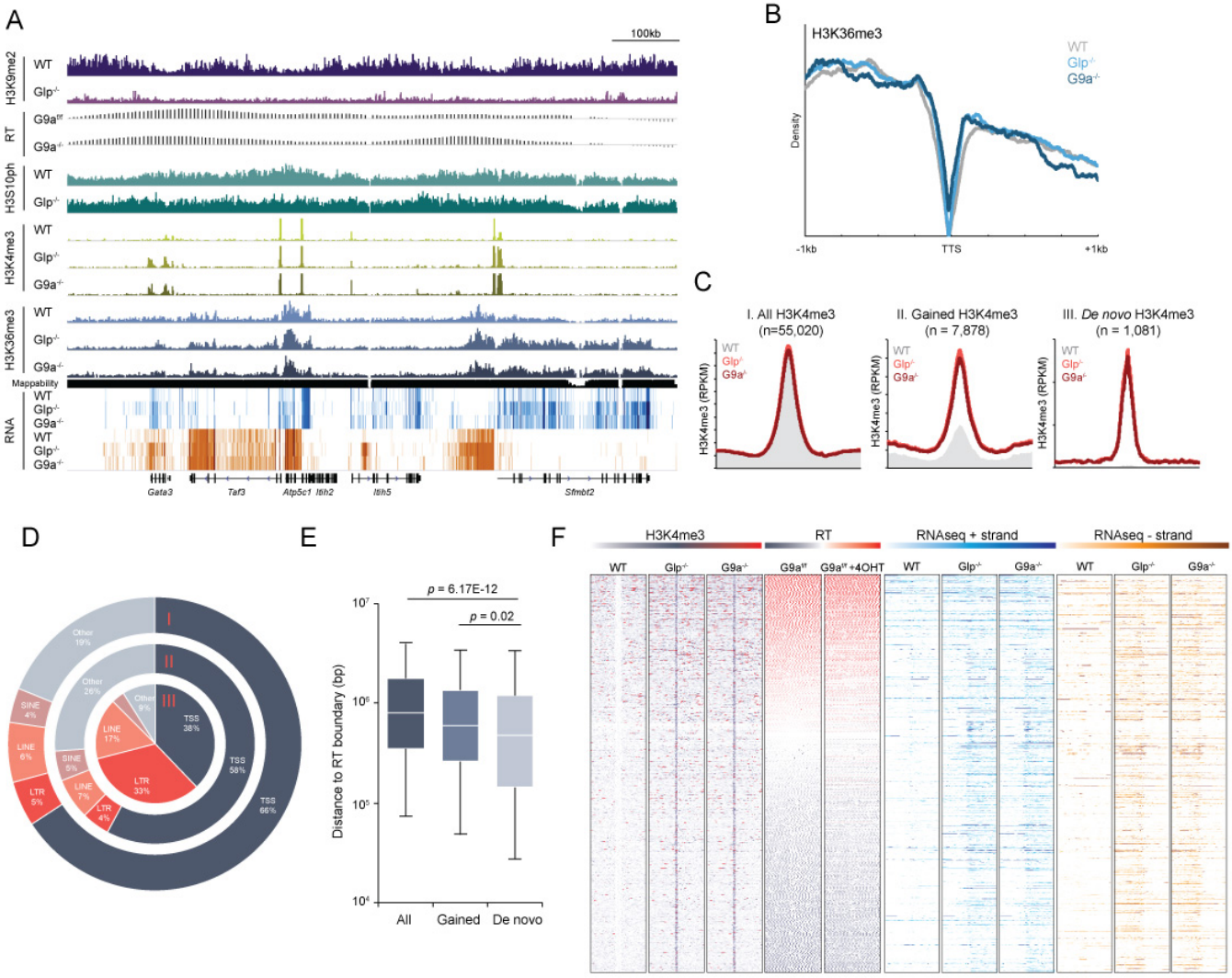


Figure 3.8 Glp-mediated H3K9me2 restricts H3K4me3 deposition and prevents aberrant activation of promoters proximal to TTRs (supplemental to Fig. 3.7).
A. Genome browser screenshot of the *Gata3 – Sfmbt2 locus* showing tracks for H3S10ph (Hesp-treated), H3K4me3, H3K36me3and strand-specific RNA-seq in WT, *Glp*$^{-/-}$ and *G9a*$^{-/-}$ ESCs. **B.** Meta-profile of the distribution of H3K36me3 around transcription termination sites (TTS +/- 1kb) in WT, *Glp*$^{-/-}$ and *G9a*$^{-/-}$ ESCs. Note that H3K36me3 levels are not increased downstream of the TTS, indicating that transcription run-through events are not prevalent in the mutant lines. **C.** Profiles of H3K4me3 enrichment around transcription start sites in WT, *Glp*$^{-/-}$ and *G9a*$^{-/-}$ ESCs. Class I includes H3K4me3 peaks that are common in all three lines (All), Class II includes sites that show higher enrichment for H3K4me3 in *Glp*$^{-/-}$ and *G9a*$^{-/-}$ than in WT cells (Gained) and Class III includes H3K4me3 peaks that are present in *Glp*$^{-/-}$ and *G9a*$^{-/-}$ cells, but are absent in WT ESCs (De novo). **D.** Genomic features overlapping with the 3 classes of H3K4me3 peaks. Note that *de novo* H3K4me3 peaks predominantly fall

within LTR and LINE repeats. **E.** Box and whisker plot showing distance, in base pairs (bp) to the closest RT boundary for each of the three H3K4me3 classes. D*e novo* H3K4me3 sites are found closer to RT boundaries than All and Gained sites. **F.** Heat maps of H3K4me3, replication timing (RT) (Yokochi et al. 2009) and strand-specific RNA-seq centered on the 1,081 *de novo* H3K4me3 sites (+/- 5 kb) in each subpanel. Note the clear increase in RNA-seq coverage downstream or upstream of the *de novo* H3K4me3 peak in the + strand and - strand heat maps, respectively.

3.2.8 H3S10ph domains are restricted to gene bodies in interphase MEFs

As somatic cells exhibit chromatin marks and replication timing profiles distinct from ESCs (Core et al. 2008; Mikkelsen et al. 2007), I next determined whether early replicating regions are also enriched in H3S10ph in interphase MEFs. As observed in asynchronous ESCs, MEFs show heterogeneity in H3S10ph staining in interphase, with a subset of cells showing small but abundant foci exclusive of DAPI-dense regions (**Fig. 3.9A**). However, ChIP-seq of Hesp-arrested MEFs revealed distinct differences in H3S10ph distribution in early replicating regions relative to ESCs, with extended domains in ESCs resolved into smaller domains in MEFs (**Fig. 3.9B)**. Metagene analysis reveals prominent enrichment of H3S10ph in gene bodies, with a bias towards the 3'end of transcribed genes (**Fig. 3.9C**). As in ESCs, gene body H3S10ph enrichment was generally higher in Hesp-treated than control MEFs (**Fig. 3.10A**), indicating that such regions are maintained by an AURKB-independent H3S10 kinase in interphase. However, consistent with ChIP-seq on asynchronous MEFs, H3S10ph in Hesp-treated MEFs does not mark broad domains. Indeed, while H3S10ph covers ~30% of the genome in interphase ESCs, H3S10ph is enriched at only 11% of all genomic bins (H3S10ph RPKM – input > 0.1) in Hesp-arrested MEFs.

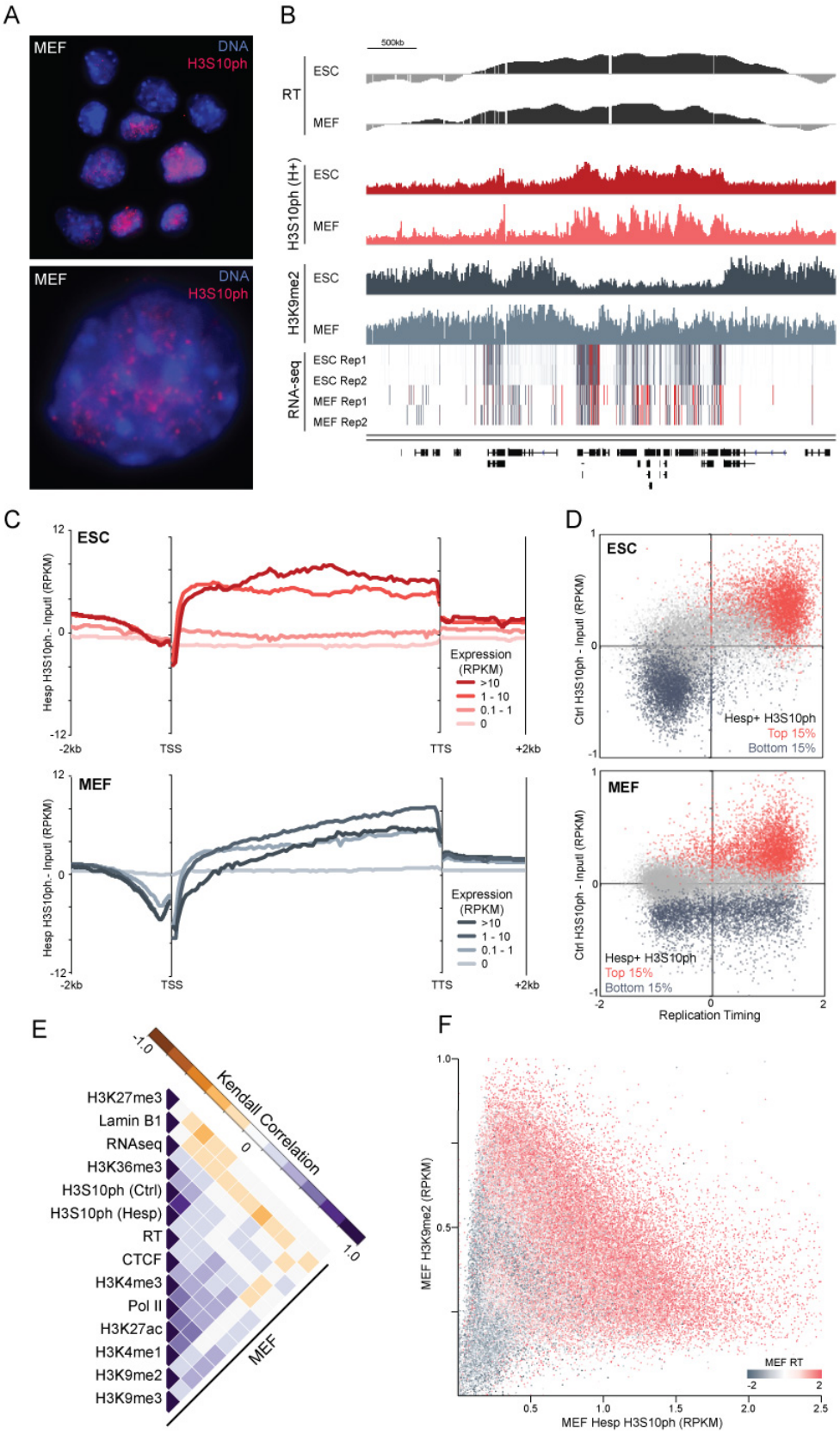

A
MEF
DNA
H3S10ph
MEF
DNA
H3S10ph
B
500kb
RT
ESC
MEF
H3S10ph (H+)
ESC
MEF
H3K9me2
ESC
MEF
RNA-seq
ESC Rep1
ESC Rep2
MEF Rep1
MEF Rep2
C
ESC
Hesp H3S10ph - Input (RPKM)
12
0
-12
-2kb
TSS
TTS
+2kb
Expression (RPKM)
>10
1 - 10
0.1 - 1
0
MEF
Hesp H3S10ph - Input (RPKM)
12
0
-12
-2kb
TSS
TTS
+2kb
Expression (RPKM)
>10
1 - 10
0.1 - 1
0
D
ESC
Ctrl H3S10ph - Input (RPKM)
1
0
-1
Hesp+ H3S10ph
Top 15%
Bottom 15%
MEF
Ctrl H3S10ph - Input (RPKM)
1
0
-1
Hesp+ H3S10ph
Top 15%
Bottom 15%
-2
0
2
Replication Timing
E
-1.0
Kendall Correlation
0
1.0
H3K27me3
Lamin B1
RNAseq
H3K36me3
H3S10ph (Ctrl)
H3S10ph (Hesp)
RT
CTCF
H3K4me3
Pol II
H3K27ac
H3K4me1
H3K9me2
H3K9me3
MEF
F
1.0
0.5
MEF H3K9me2 (RPKM)
MEF RT
-2
2
0.5
1.0
1.5
2.0
2.5
MEF Hesp H3S10ph (RPKM)

Figure 3.9 H3S10ph is restricted to genic regions in interphase MEFs.
A. H3S10ph immunostaining of interphase WT MEFs, counterstained with Hoechst 33342 for DNA, as in Figure 3.1A. **B**. Genome browser screenshot showing RT, H3K9me2 and biological replicates of RNA-seq coverage in ESCs and MEFs as well as H3S10ph in Hesp-treated cells. **C**. Metagene analysis of normalized H3S10ph enrichment (input subtracted RPKM) over all murine gene bodies (-2kb-TSS-TTS+2kb), clustered by RNA-seq expression levels into quartiles in the respective cell type. **D**. Pair-wise comparison of H3S10ph with replication timing in asynchronous ESCs and MEFs using genome-wide 100kb tiling bins. Red and blue data points correspond to bins that were in the top and bottom 15% of H3S10ph enrichment, respectively, in Hesp-treated cells. **E**. Kendall correlation matrix of H3S10ph with chromatin features across 5kb genomic bins in MEFs, following unsupervised hierarchical clustering. **F**. Pair-wise comparison of H3S10ph (Hesp-treated) with H3K9me2 genome-wide (5kb bins) in MEFs, overlaid with a heatmap of MEF RT data with red and grey data points representing early- and late-replicating bins, respectively. Note that regions showing high H3K9me2 and low H3S10ph can still be early replicating in MEFs.

Inspection of the distribution of H3S10ph in relation to replication timing reveals that unlike in ESCs, this mark is not uniformly distributed across early replicating regions in MEFs (**Fig. 3.9B**). While H3S10ph-enriched regions are typically replicated early in MEFs, H3S10ph-depleted regions can be early- or late-replicating (**Fig. 3.9D**). Using 10kb tiled genomic-bins, H3S10ph is 94.0% sensitive and 97.9% specific as a predictor of early RT in ESCs; in stark contrast, in MEFs, sensitivity of H3S10ph to predict early RT is 91.7%, but specificity is only 46.7%. Notably, the subset of genomic regions that are late replicating in ESCs but early replicating in MEFs show relatively low levels of H3S10ph in the latter (**Fig. 3.10B**), indicating that H3S10ph is not a universal mark of early RT.

Meta-epigenomic comparisons in MEFs reveals many of the same trends observed in ESCs, with regions marked with H3S10ph in Hesp-treated cells showing a positive correlation with RT and other features of transcriptionally active regions and a negative correlation with H3K9me2, H3K27me3 and LADs (**Fig. 3.9E**, **Fig. 3.10C**). As H3K9me2 and H3S10ph are anti-correlated in ESCs as well as in *Drosophila*, I surmised that the absence of H3S10ph in intergenic early-replicating regions in MEFs could be coincident with a reciprocal enrichment of

H3K9me2 in these regions. Indeed, early replicating regions depleted of H3S10ph in MEFs are generally enriched for H3K9me2 (**Fig. 3.9F**) but depleted of H3K9me2 in ESCs (**Fig. 3.10D**). The trend towards increased abundance of H3K9me2 in MEFs relative to ESCs is consistent with previous mass spectrometry data, which revealed a 10% greater abundance of H3K9me2 in MEFs (Kubicek et al. 2007). Taken together, these results indicate that H3K9me2 accumulates in MEFs in intergenic regions that are not marked with H3S10ph, leading to an apparent "fragmentation" of the broad interphase H3S10ph domains observed in ESCs.

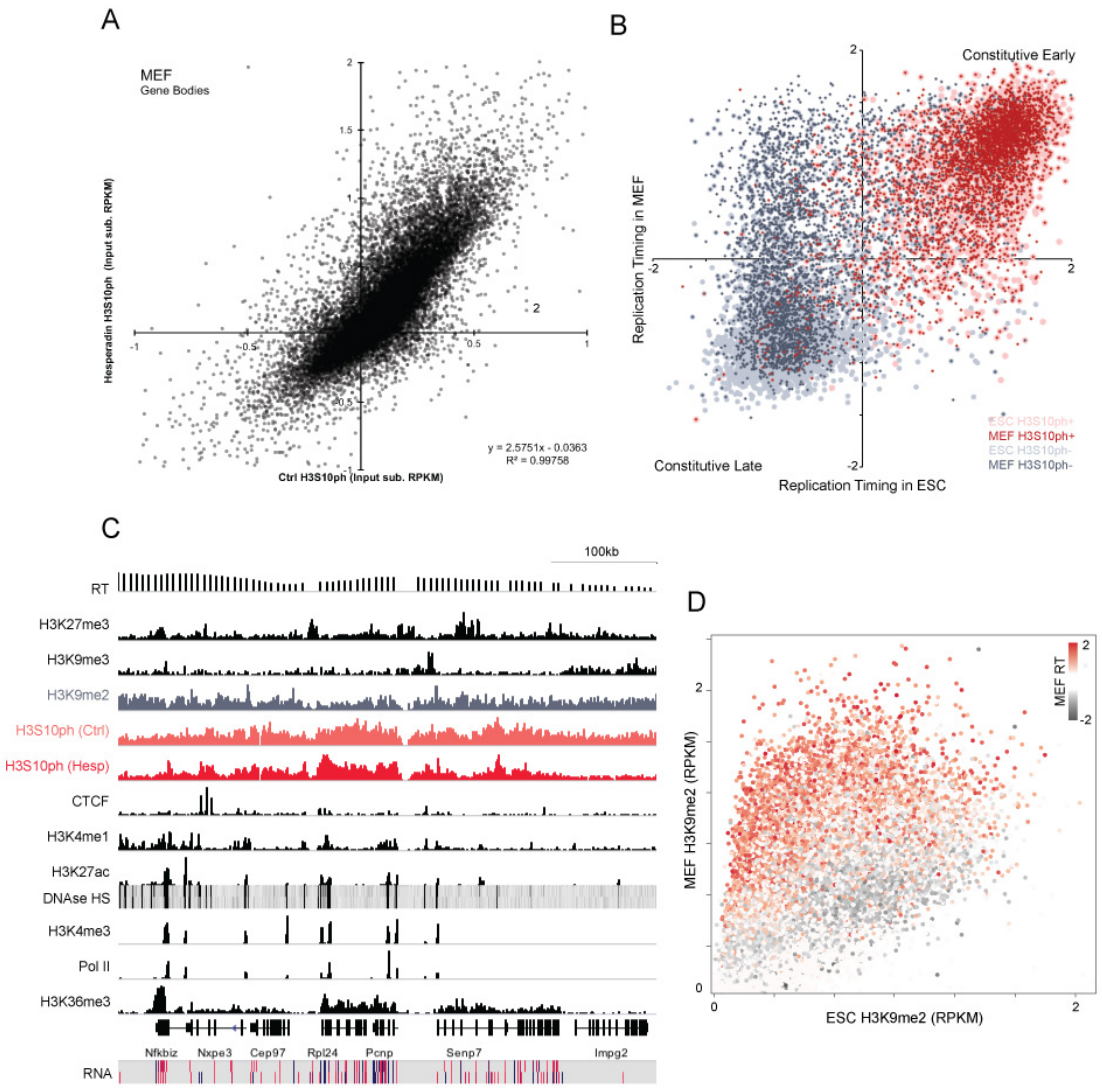


Figure 3.10 H3S10ph is enriched in gene-rich regions in interphase MEFs, but not as broad domains (supplemental to Fig. 3.9).
A. H3S10ph enrichment (input subtracted RPKM) in gene bodies in Hesp-treated MEFs versus asynchronous (Ctrl) MEFs. Note that Hesp-treatment increases interphase H3S10ph ChIP sensitivity by over 2-fold (linear regression slope). **B.** Genome-wide scatterplot of RT (100kb tiled bins) in MEFs versus ESCs (Hiratani et al. 2010), overlaid with a heatmap showing Hesp-treated H3S10ph enrichment in both cell types. Regions replicating early in both cell types are also marked with H3S10ph in interphase in both cell types, while regions that are early-replicating exclusively in MEFs are not marked with H3S10ph in interphase in either cell type. **C.** Genome browser screenshot of a gene-rich, early replicating region showing H3S10ph tracks (control and Hesp-treated) as well as other chromatin features in MEFs. **D.** Genome-wide (5kb tiled bins) pairwise comparison of H3K9me2 enrichment (RPKM) in MEFs (Fang et al. 2012b) versus ESCs, overlaid with a heat map of replication timing (RT) in MEFs. Note that regions showing a bias towards H3K9me2 enrichment in MEFs are generally early-replicating in MEFs.

3.3 Discussion

3.3.1 H3S10ph marks megabase-scale domains in interphase ESCs

H3S10ph has generally been recognized as a transient modification in interphase mammalian cells, marking specific gene promoter and enhancer regions following induction with external growth factors or stress (Wang and Higgins 2012; Bode 2005). Our genome-wide analysis of H3S10ph reveals that this PTM is not limited to transcriptional regulatory regions in interphase ESCs but rather broadly marks gene-rich regions, encompassing ~30% of the genome. The majority of gene bodies are marked with H3S10ph in interphase ESCs, a pattern similar to that observed for *Drosophila* JIL-1 occupancy (Regnard et al. 2011). The broad distribution of interphase H3S10ph in euchromatic regions is consistent with the "hyperdynamic plasticity" of chromatin reported in ESCs (Meshorer et al. 2006), and may play a role in the relatively short chromatin residency time of chromatin-associated proteins observed in ESCs in interphase and/or the exclusion of these regions from LADs. Regardless, a subset of intragenic regions as well as enhancers embedded in these broad euchromatic domains show varying levels of H3S10ph enrichment in G1 vs S vs S/G2/M datasets, perhaps reflecting the dynamics of transcriptional regulation during the cell cycle. Indeed, genes enriched preferentially with H3S10ph in specific cell stages are overrepresented in distinct GO biological processes associated with housekeeping functions.

While our ChIP-seq analyses are consistent with the model that AURKB promiscuously phosphorylates all nucleosomes in mitotic cells; AURKB is unlikely to be responsible for euchromatic H3S10ph in G1 ESCs, as treatment with the specific inhibitor Hesperadin did not result in a loss of H3S10ph in gene-dense regions. MSK1/Rps6ka5 and MSK2/Rps6ka4, the two

orthologs of JIL-1 in mammalian genomes, phosphorylate H3S10 upon ERK or p38-mediated signal transduction (Thomson et al. 1999a); however, both MSKs are expressed at low levels in ESCs relative to MEFs, and *Msk1/2* DKO MEF showed no appreciable loss of basal H3S10ph (Soloaga et al. 2003). RSK2, a member of the related kinase family, p90 RSKs, was implicated in EGF-induced H3S10ph in interphase ESCs (Sassone-Corsi et al. 1999), but like MSKs, RSK2 acts only at a subset of inducible promoters/enhancers. Indeed, I have found that H3S10ph in gene-rich regions is maintained in *Msk* and *Rsk* knock-out ESCs (**Chapter 5**). Identifying the kinase(s) responsible for H3S10ph deposition in interphase will require systematic depletion of individual candidates and perhaps combinatorial knock-outs.

3.3.2 Reciprocal H3K9me2/H3S10ph antagonism

H3K9me2 was previously recognized as a broadly distributed modification (Lienert et al. 2011). My analyses reveal that H3S10ph and H3K9me2 are anti-correlated in both MEFs and ESCs, with the majority of the mappable genome covered by one or the other mark in ESCs. H3K9me2 accumulates over transcribed genic regions in ESCs overexpressing H3.3S10A, indicating that H3S10ph may directly inhibit deposition of H3K9me2 in mammalian cells, consistent with previous reports of heterochromatin spreading in JIL-1 hypomorphs in *Drosophila* (Zhang et al. 2006; Cai et al. 2014). Notably, antagonism between H3K9me2 and H3S10ph in interphase ESCs seems to be reciprocal, as G9a/Glp-dependent H3K9me2 prevents the spreading of H3S10ph into gene-poor regions at most TTRs (**Fig. 3.11**). Such expansion of H3S10ph domains is unlikely to be a consequence of transcription, as this mark spreads over ~100kb beyond TTRs in $Glp^{-/-}$ ESCs regardless of the presence of aberrant transcripts. As few new H3S10ph domains appear in $Glp^{-/-}$ ESCs, ectopic H3S10ph at TTRs may depend upon spreading of this mark via

mass action into regions previously marked with H3K9me2. Whether H3K9me2 and/or readers of this mark, directly inhibit the catalytic activity of the relevant H3S10 kinase(s) remains to be determined.

3.3.3 Asymmetrical transcription at TTRs in the absence of H3K9me2

I observed a strong directional bias in *de novo* initiated transcripts in *G9a*$^{-/-}$ and *Glp*$^{-/-}$ ESCs, specifically in TTRs. Minus stranded transcription was favored at late-early TTRs, whereas plus stranded transcription was favored at early-late TTRs (**Fig. 3.7E, 3.11B**), consistent with co-orientation of transcription with the replication fork. Notably, many of the genes previously reported to be upregulated in the absence of G9a or GLP, including the *MageA* and *Rhox* gene clusters (Tachibana, 2002), are within TTRs and are co-oriented with the replication fork (**Fig. 3.7F**), revealing that this phenomenon may be extended to endogenous genes. Intriguingly, a reporter construct integrated in mammalian cells in the same locus in both orientations showed strong orientation-dependent transcription (Feng et al. 2001; Strathdee et al. 2006), perhaps reflecting the same phenomenon. Regardless, as replication timing in G9a deficient ESCs is largely unchanged (Yokochi et al. 2009), such aberrant transcription is unlikely to be a consequence of a shift in replication timing.

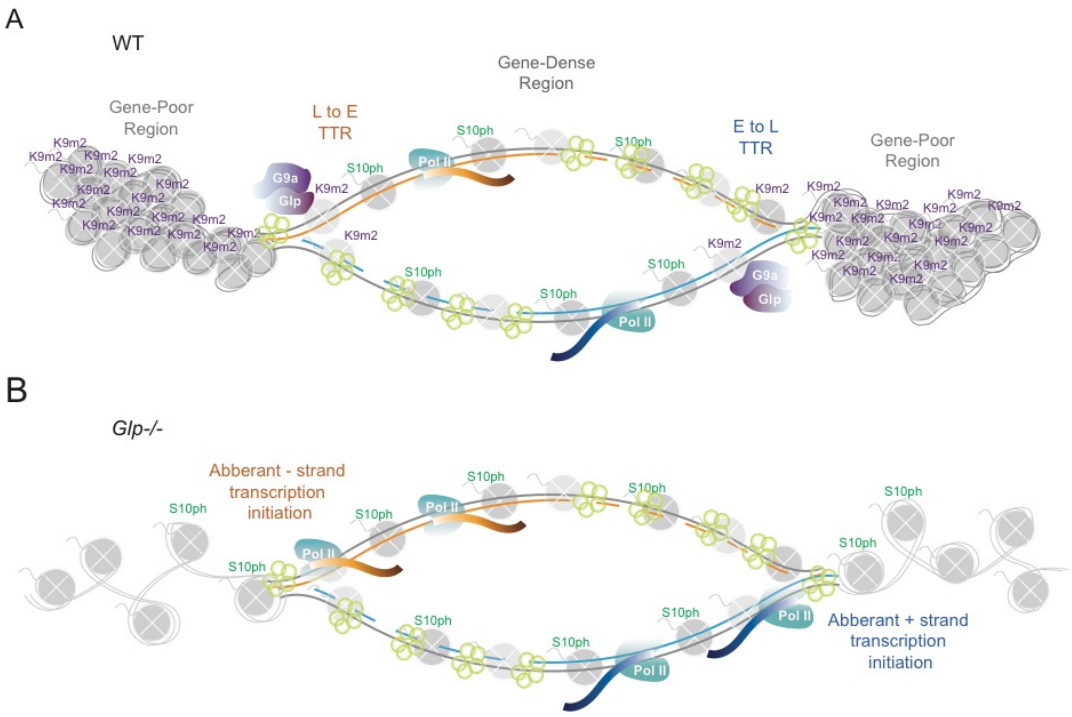


Figure 3.11 Model of the interplay between H3S10ph and H3K9me2 during DNA synthesis in ESCs, and the influence of perturbing H3K9me2 deposition on strand-biased transcription.
A. Schematic of a hypothetical replicon in WT ESCs, with AURKB-independent H3S10ph marking early-replicating, gene-dense regions and G9a/GLP-dependent H3K9me2 marking late-replicating, gene-poor regions. Replication timing transition regions (TTRs), present at the edges of the replicon, coincide with the interface of H3S10ph and H3K9me2 domains in WT ESCs. **B**. Schematic of the same replicon in *Glp*$^{-/-}$ ESCs showing that in the absence of H3K9me2, H3S10ph spreads into late replicating regions, coincident with aberrant transcription initiating asymmetrically at TTRs; showing a plus-strand bias at early-to-late (E to L) and a minus-strand bias at late-to-early (L to E) TTRs. Thus, in both cases, transcription is co-oriented with leading strand synthesis.

An alternative explanation for the strong strand bias in transcriptional upregulation observed at the boundaries of H3K9me2 and H3S10ph marked domains is an intrinsic asymmetry in the "transcription-competency" of leading versus lagging strand templates in the wake of the replication fork. A difference in the rate of chromatin maturation of the leading versus lagging strand following replication could render the former more sensitive to the establishment of an open chromatin state, potentiating transcription behind the replication fork.

3.3.4 H3S10ph domains in ESCs versus MEFs

While replication timing profiles are generally conserved in ESCs and MEFs, the broad uninterrupted domains of H3S10ph observed in early replicating, gene-dense regions in ESCs are frequently interspersed with smaller domains of H3K9me2 in MEFs. Recession of H3S10ph in these intergenic and 5' genic regions may reflect a decrease in the level or activity of the kinase responsible, or an increase in the level or activity of the relevant phosphatase, likely PP1 (Hayashi-Takanaka, 2007), in MEFs. PP1 is inactivated during mitotic exit as well as the G1/S transition (Liu et al. 1999). A long G1 phase may therefore allow for more robust phosphatase activity in MEFs, leading to punctuated H3S10ph profiles. The concomitant gain of H3K9me2 in intergenic early-replicating regions is not accompanied by establishment of new LADs – suggesting that H3K9me2 deposition is not sufficient for relocalization to the nuclear periphery. Regardless, the accumulation of H3K9me2 may be a general feature of somatic cells, as a global increase of G9a/GLP-dependent H3K9me2 was also observed in the E5.5 epiblast after implantation, specifically within gene bodies (Zylicz et al. 2015). G9a is essential for mid-gestation development with protracted expression of pluripotency markers such as *Nanog* and *Oct3/4* observed at E7.5 and lethality observed at E10.5 in *G9a*$^{-/-}$ embryos (Shinkai and Tachibana 2011; Yamamizu et al. 2012). Whether the broader domains of H3S10ph in ESCs are required for pluripotency or simply a consequence of differential cell cycle kinetics and/or kinase/phosphatase activities remains to be determined.

3.3.5 Reconciling the role of mitotic & interphase H3S10ph

To reconcile the paradoxical association of H3S10ph with condensed chromatin in mitosis and de-condensed chromatin in interphase, Johansen and Johansen proposed that this mark promotes

the detachment of chromatin from the nuclear scaffold (Johansen and Johansen 2006), facilitating chromosome segregation and transcription. Consistent with this model, following stimulation in B lymphocytes, immediate early genes physically shift from the periphery of the nucleus to constitutively active transcription factories prior to the onset of transcription (Osborne et al. 2007). Furthermore, the stable radial positioning of chromosomes is established *de novo* within the first 2h of G1 (Walter et al. 2003), coinciding with establishment of self-associating topological associating domains (TADs) and RT (Naumova et al. 2013; Pope et al. 2014). I show that interphase H3S10ph domains are most prominent in G1, are not associated with the nuclear lamina, and are universally early replicating (**Fig. 3.4E**). Thus, H3S10ph-mediated nuclear positioning may enhance both transcription and access to the replication machinery during interphase.

As H3S10ph exerts little influence on physical packing of nucleosome arrays *in vitro* (Fry and Shogren-Knaak 2004; Shogren-Knaak 2003), this histone mark may promote untethering through repulsion of lamin-associated proteins. Consistent with this model, disruption of mitotic H3S10ph through AURKB inhibition results in retention of HP1 and lamin resident protein PRR14 on metaphase chromosomes (Fischle et al. 2005; Poleshko and Katz 2014). Intriguingly, G9a itself is physically tethered to the nuclear lamina (Kind et al. 2013), possibly via interaction with the lamin-associated protein Barrier to Auto-integration Factor (Montes de Oca et al. 2014), and genes repressed by G9a are localized to the nuclear periphery in ESCs (Yokochi et al. 2009). Determining whether interphase H3S10ph is a driver or bystander in nuclear compartmentalization of marked euchromatic regions in mammals will depend upon identification of the kinase(s) involved.

4. PERSISTENCE OF AURORA B/C-DEPENDENT H3K9ME3S10PH IN INTERPHASE

4.1 Introduction

Most repetitive sequences are maintained in a transcriptionally silent state by multiple epigenetic mechanisms acting in concert, such as DNA methylation and H3K9 methylation. Repression of endogenous retroviruses (ERVs), specifically those that have recently integrated into the murine genome, requires SETDB1-dependent H3K9me3 in ESCs and primordial germ cells (PGCs) (Matsui et al. 2010; Liu et al. 2014). Establishment of SETDB1-mediated H3K9me3 is dependent upon KRAB-ZFPs, which encode an array of zinc-fingers that bind to specific DNA sequences, and the repressor scaffolding protein KAP1, which interacts directly with SETDB1 (Wolf and Goff 2009; Rowe et al. 2010; Schultz 2002). H3K9me3 readers, such as Heterochromatin Protein 1 (HP1), are also recruited to such H3K9me3-marked sites (Vogel et al. 2006), and HP1 is proposed to propagate H3K9me3 heterochromatin into flanking genomic regions. However, unlike *KAP1* or *Setdb1* KO ESCs which shows loss of DNA methylation and concurrent transcriptional de-repression of ERVs (Rowe et al. 2010; Leung et al. 2011), depletion of *HP1α* or *HP1β* have little effect on the transcription of satellite repeats or ERVs (Maksakova et al. 2013; 2011).

As discussed in **Chapter 1,** the phospho-methyl switch is a biological phenomenon in which the phosphorylation of H3 at serine 10 inhibits the recognition of the adjacent H3K9me3

by HP1 (Fischle et al. 2005). During mitosis, Aurora B broadly phosphorylates H3S10 and discharge the majority of chromatin-bound HP1 to the nucleoplasm, except for a small pool of HP1 at the centromeres (Hayakawa et al. 2003). The functional relevance of this mitotic-specific displacement of HP1 is unclear, but has been speculated to liberate chromatin-bound HP1 in order to potentiate the kinase activity of Aurora B towards kinetochore substrates (Abe et al. 2016). Alternatively, in fission yeast, mitotic displacement of *Swi6*, the HP1 homolog, was proposed to facilitate transcription of centromeric repeats for the generation of small RNAs required for the establishment of heterochromatin (Allshire and Ekwall 2015; Chen et al. 2008). Whether H3S10ph is relevant for HP1 exclusion from H3K9me3-marked regions in interphase mammalian cells has not been investigated.

A separate class of H3K9me3 readers are also reported to interact with H3K9me3S10ph-marked tails (Kunowska et al. 2015). NP95/UHRF1 and ATRX, for example, are tolerant of H3K9me3S10ph (Rothbart et al. 2012; Noh et al. 2014), and ATRX even shows a preference for the H3K9me3S10ph over H3K9me3 *in vitro*. Furthermore, ATRX is recruited to pericentromeric heterochromatin in forskolin-stimulated neurons in a H3S10ph-dependent manner (Noh et al. 2014). A mutation in NP95 that specifically disrupt its interaction with H3K9me3S10ph, but not H3K9me3, results in impaired mitotic retention of NP95/DNMT1 and loss of DNA methylation at ribosomal DNA in HeLa cells (Rothbart et al. 2012). Unlike the HP1s, both NP95 and ATRX are recruited to H3K9me3-marked ERVs, and are functionally relevant for SETDB1-dependent silencing of ERVs in ESCs (Sharif et al. 2016; Elsässer et al. 2015; Sadic et al. 2015). Taken together, these results suggest that the presence of H3K9me3S10ph may influence both the DNA methylation maintenance machinery and ATRX-dependent nucleosome remodeling at ERVs.

To understand the relationship between H3K9me3S10ph and HP1 in interphase, I took advantage of the FUCCI ESCs that I established previously in **Chapter 3** for cell cycle synchronization, and performed H3K9me3 and H3K9me3S10ph ChIP-seq in G1-sorted cells. I found that essentially all H3K9me3-marked regions are concurrently marked with H3K9me3S10ph in G1. Unlike the H3S10ph mark, described in **Chapter 3**, which is maintained independent of Aurora B/C kinase activity, H3K9me3S10ph is apparently deposited by Aurora B/C kinase at G2/M and persists in G1. To determine whether HP1 is ejected from H3K9me3S10ph-marked chromatin, as previously reported (Fischle et al. 2005), I also performed HP1β ChIP in G1-, S- and S/G2/M-sorted cells. HP1β is a faithful H3K9me3 reader, showing binding at ~80% of H3K9me3 peaks but is notably depleted from the promoter regions of SETDB1/H3K9me3-marked genes that are expressed in the ovaries/testis. Surprisingly, HP1β shows robust localization at H3K9me3S10ph-enriched class I & II ERVs, suggesting that the transition between H3K9me3S10ph and H3K9me3 may be dynamic in interphase. Taken together, the data presented here indicate that Aurora-dependent H3S10ph is not fully dephosphorylated at the end of mitosis in repetitive regions, generating the bivalent H3K9me3S10ph mark in interphase. I present a model for the deposition and persistence of H3K9me3S10ph during the cell cycle, and discuss implications of this mark in directing H3K9me3 readers at retroviral elements in ESCs.

4.2 Results

4.2.1 Validation of H3K9me3 and H3K9me3S10ph-specific antibodies

All commercial antibodies raised against H3K9me3 are excluded by the presence of H3S10ph (Bock et al. 2011; Duan et al. 2008), including those used previously by us to generate ChIP-seq datasets (Karimi et al. 2011). To investigate how such "phospho-epitope masking" confounds the interpretation of results obtained with a H3K9me3-specific antibody, I co-stained asynchronous ESC nuclei with the antibody used previously for H3K9me3 ChIP-seq, and an antibody raised specifically against the bivalent H3K9me3S10ph mark. While the H3K9me3 polyclonal antibody stains the condensed pericentromeric heterochromatin in most interphase cells, it is specifically excluded from heavily phosphorylated mitotic chromatin (**Fig. 4.1**, arrowhead), consistent with the reported histone peptide binding profile of this antibody (Bock et al. 2011). Conversely, the H3K9me3S10ph antibody stained interphase chromocenters and mitotic chromosomes, as expected. Interestingly, labelling of chromocenters with either H3K9me3 or H3K9me3S10ph is heterogeneous even amongst non-mitotic cells, with some cells staining heavily for H3K9me3 or H3K9me3S10ph, but not others – suggesting the antibody masking effect is also occurring in interphase ESCs.

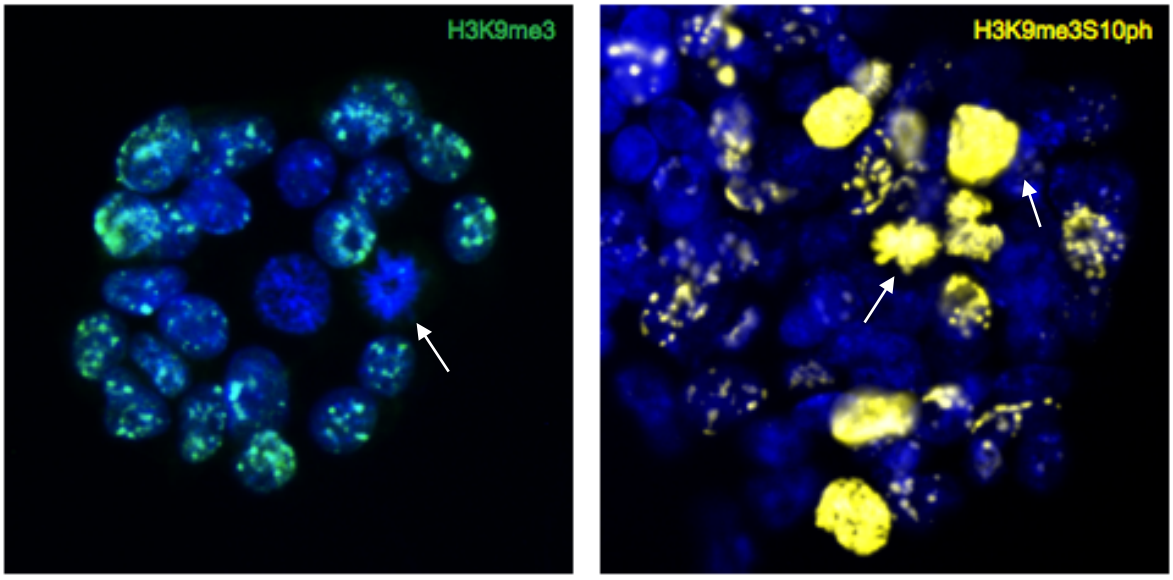


Figure 4.1 H3K9me3 antibody is excluded from mitotic chromatin.
Immunofluorescence of asynchronous ESCs stained with antibody specific for H3K9me3 (green) or H3K9me3S10ph (yellow), counterstained with Hoechst 33342 (blue). White arrows denote mitotic cells. While the H3K9me3S10ph-antibody stains mitotic nuclei (white arrowheads), H3K9me3-antibody is occluded.

To analyze the regions enriched for H3K9me3 and H3K9me3S10ph in greater detail, I conducted ChIP, followed by qPCR. Consistent with our previous studies, Intracisternal A Particle (IAP) long terminal repeats (LTRs) in general, as well as a specific IAP LTR on Chr2, were enriched for H3K9me3 (Karimi, 2011), while the IgG control IP showed very low enrichment (**Fig. 4.2)**. Surprisingly, very high levels of enrichment at the same regions were also detected with the H3K9me3S10ph antibody (**Fig. 4.2)**. Furthermore, the high level of enrichment in asynchronous cells suggests that H3K9me3S10ph may not be limited to mitotic chromatin and the high yield indicates that this dual mark may be more prevalent than H3K9me3 in asynchronous ESCs. I further confirmed that the H3K9me3 and H3K9me3S10ph antibodies were not cross-reactive in the ChIP reaction by pre-blocking the antibody using histone peptides. H3K9me3 ChIP signal was effectively abolished when the H3K9me3 antibody was pre-blocked

with H3K9me3 peptide. Conversely, signal for H3K9me3S10ph at the same region did not change upon addition of H3K9me3 or H3S10ph peptide, but was abolished upon addition of the H3K9me3S10ph peptide, suggesting that the H3K9me3S10ph-specific antibody is only sensitive to inhibition by its cognate substrate.

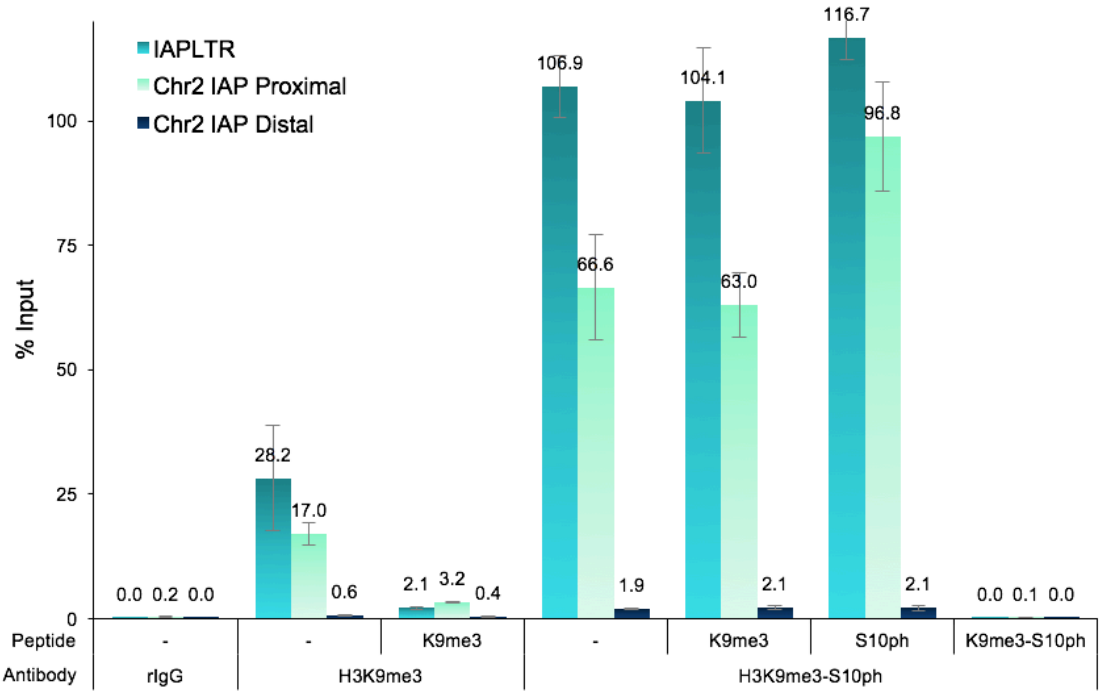


Figure 4.2 H3K9me3S10ph antibody specifically recognizes the bivalent H3K9me3S10ph tail.
ChIP-qPCR of chromatin IPed with the H3K9me3 and H3K9me3S10ph specific antibodies, blocked with various histone peptide tails at IAP ERVs. The "IAPLTR" amplicon recognizes multiple IAP elements in the genome, while the "Chr2 IAP Proximal" amplicon recognizes a specific IAP LTR on Chromosome 2 and the "Chr2 IAP Distal" amplicon recognizes a region on Chr2 distal from the LTR region shown previously to lack H3K9me3. Note that ChIP signal for the H3K9me3S10ph-specific antibody is abolished only upon addition of the cognate H3K9me3S10ph peptide, but not H3K9me3 or H3S10ph peptides. IgG was included as a negative control.

4.2.2 H3K9me3S10ph is dependent on Aurora B/C kinase

Given that SETDB1, the KMT responsible for deposition of H3K9me3 at IAP elements, is unable to methylate pre-phosphorylated H3 substrates (Schultz 2002), I postulated that phosphorylation of H3S10 occurs after methylation of H3K9. To examine whether Aurora B is

the kinase depositing H3S10ph in regions of the genome showing enrichment for H3K9me3S10ph, I utilized a potent Aurora B/C kinase inhibitor, hesperidin and analyzed the cell cycle distribution of H3K9me3S10ph (**Fig. 4.3A)**. In DMSO vehicle-treated ESCs, ~3% of the cell population stained intensely for the H3K9me3S10ph mark, as determined by flow cytometry. This population was exclusively of 4n DNA content, indicating that these cells are likely in M-phase of the cell cycle. A sub-population of cells in late-S to G2 was also observed with intermediate levels of H3K9me3S10ph. Following 3h treatment with 200nM hesperidin, both the mitotic and sub-mitotic fraction were significantly depleted of H3K9me3S10ph signal, suggesting that Aurora B/C phosphorylates H3K9me3-marked H3 tails, generating H3K9me3S10ph in G2. The staining as measured by flow cytometry likely reflects the phosphorylation of pericentromeric heterochromatin (**Fig 4.1),** which is the most abundant class of H3K9me3-marked repetitive DNA in the genome. To determine whether Aurora B/C was also phosphorylating interspersed H3K9me3-marked transposable elements, I performed ChIP-qPCR after hesperidin treatment (**Fig. 4.3B)**. Indeed, H3K9me3S10ph enrichment was significantly depleted following hesperidin treatment, indicating that Aurora B also phosphorylates H3S10 on H3K9me3-marked ERV LTRs.

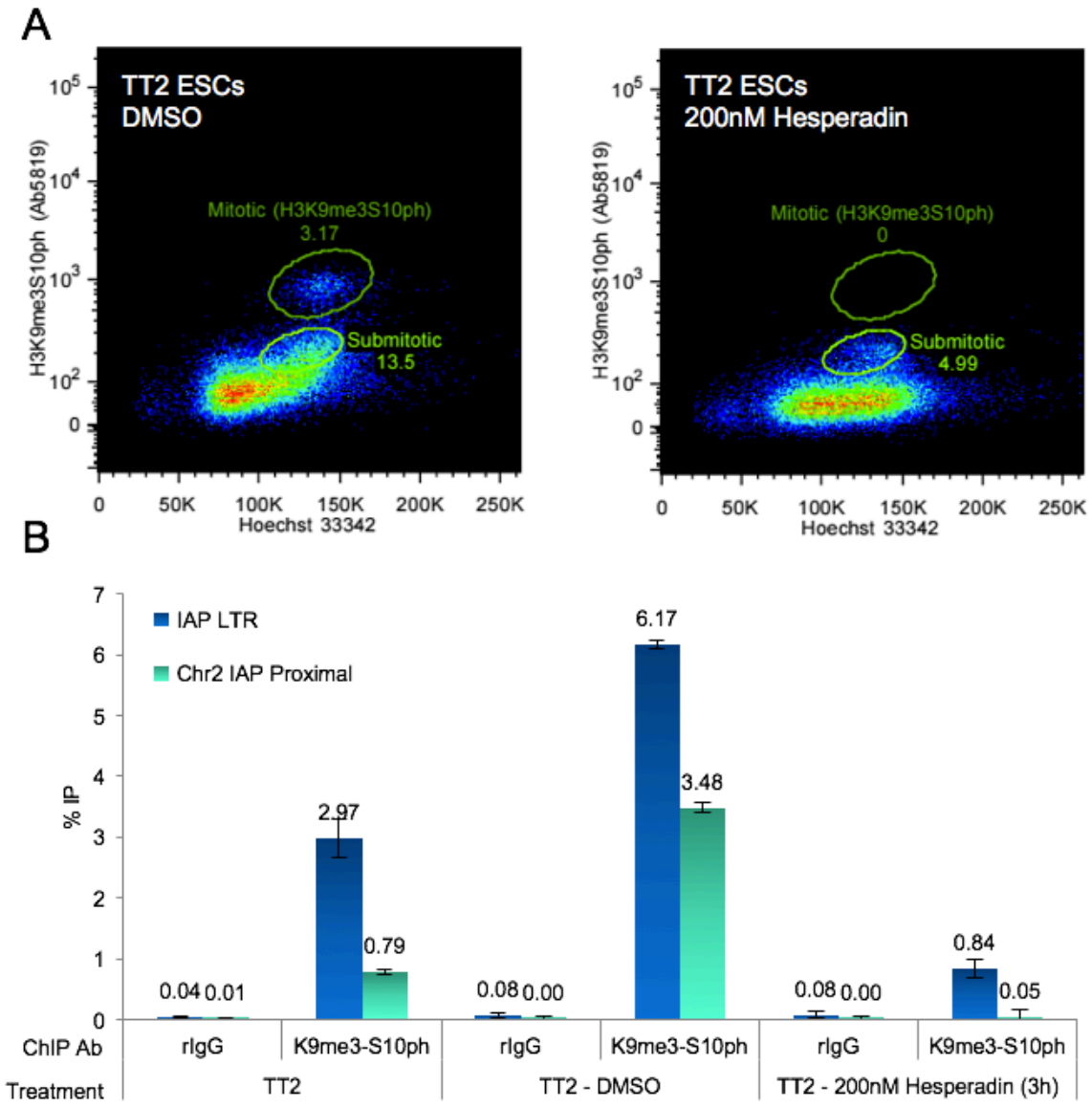


Figure 4.3 H3K9me3S10ph is Aurora B-dependent.
A. Flow cytometry cell cycle analysis of H3K9me3S10ph and Hoechst 33342-stained ESCs treated with hesperidin vs. vehicle (DMSO) control. Note the significant depletion of H3K9me3S10ph in the "mitotic" and "submitotic" fraction with 4n DNA. **B.** ChIP-qPCR of H3K9me3S10ph in ESCs treated with hesperadin vs. vehicle (DMSO) control. Note that hesperidin-treatment abolished H3K9me3S10ph signal at IAP LTRs using a multi-copy amplicon (IAP LTR) and a single copy of IAP on Chr2. rIgG: negative control.

4.2.3 H3K9me3S10ph co-localizes genome-wide with H3K9me3

Given the high level of H3K9me3S10ph enrichment at IAP elements, I wondered whether this dual mark was generally enriched at the same genomic regions marked with H3K9me3. I, therefore, performed parallel H3K9me3 and H3K9me3S10ph ChIP-seq in G1-sorted FUCCI ES cells (**Fig. 4.4A**). Unexpectedly, H3K9me3S10ph enrichment was strongly correlated with H3K9me3 enrichment in G1 cells ($R^2 = 0.99$), with 96% of the genomic bins showing no significant difference ($-0.25 < z < 0.25$), and both showing enrichment at repetitive regions with low mappability. I found slightly better ChIP recovery of H3K9me3S10ph vs. H3K9me3 at uniquely aligned regions, predominantly at rDNA clusters, but examination of the H3K9me3 ChIP track revealed similar peak profiles genome-wide, i.e. H3K9me3S10ph does not mark regions that are not marked with H3K9me3 (**Fig. 4.4B**).

To determine whether regions showing high levels of H3K9me3S10ph are depleted of HP1β enrichment in interphase ESCs, as would be expected if this H3K9me2/3 reader is sensitive to the presence of S10ph, I performed HP1β ChIP-seq in G1-sorted FUCCI ESCs. Surprisingly, HP1β was enriched at regions that are heavily enriched for H3K9me3S10ph. Genomic regions enriched for H3K9me3, H3K9me3S10ph and HP1β were predominantly class I & II ERV families (**Fig. 4.4C**), which we have shown previously to be targets of SETDB1 (Karimi et al. 2011). These results reveal that enrichment of H3K9me3S10ph does not necessarily predict HP1β exclusion, indicating that at such regions, dynamic turnover of phosphorylation at S10 is likely taking place in interphase, with HP1β bound to nucleosomes marked only by H3K9me3.

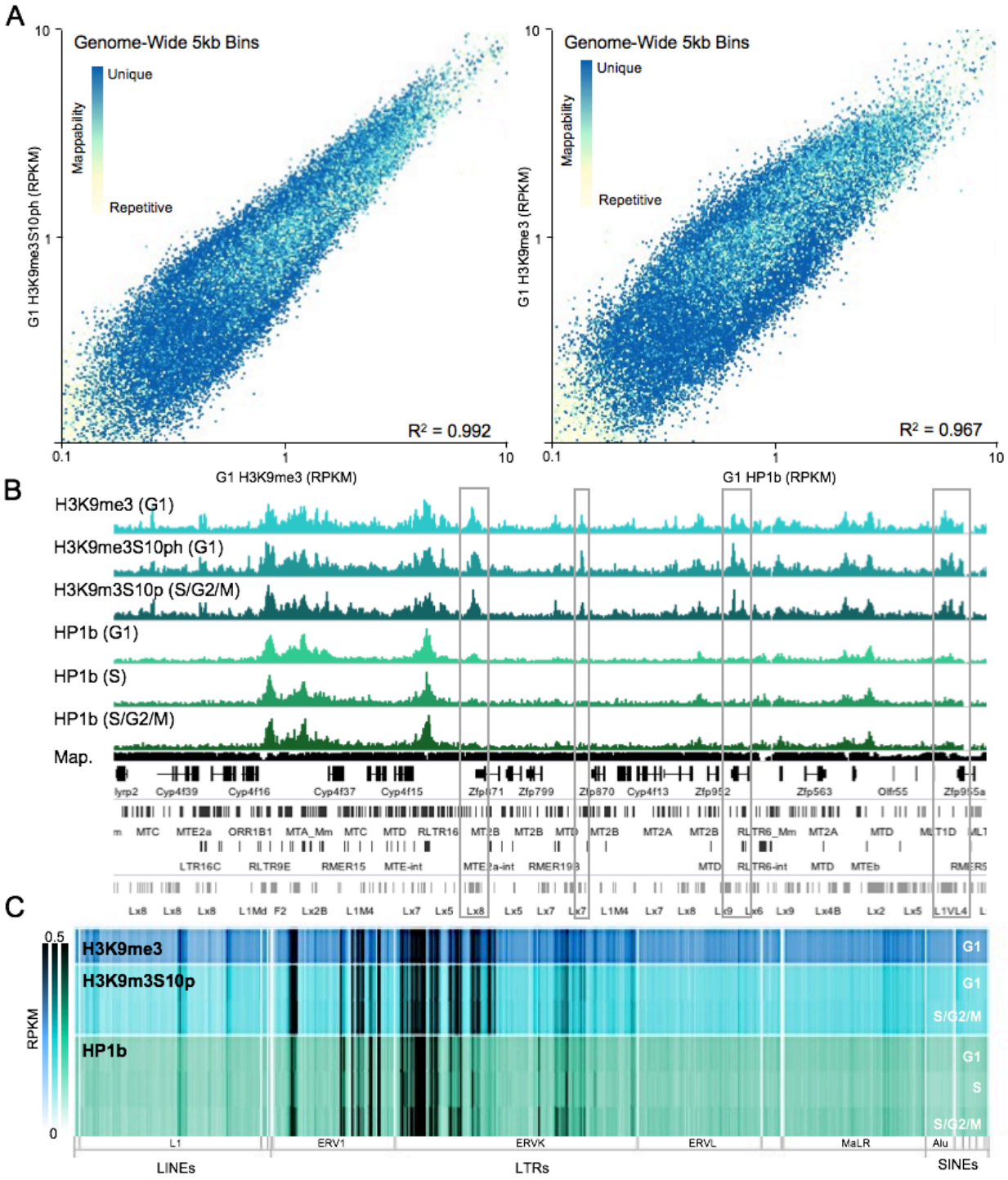


Figure 4.4 H3K9me3S10ph persists in G1 ESCs, and does not predict HP1β exclusion.
A. Genome-wide correlation of H3K9me3 with H3K9me3S10ph and HP1β in G1-sorted ESCs over 5kb bins. Bins are coloured according to mappability score. R^2 = Pearson correlation. H3K9me3, H3K9me3S10ph and HP1β are enriched at repetitive regions in the genome, and are highly correlated to each other. **B.** Browser screenshot of

H3K9me3, H3K9me3S10ph and HP1β tracks over a cluster of *Zfp* genes showing largely cell cycle invariant occupancy at repeat-rich regions. While all HP1β-bound genomic regions were highly enriched for both H3K9me3 and H3K9me3S10ph, the converse was not true, i.e. a small subset of regions with H3K9me3 and H3K9me3S10ph were not bound by HP1β (grey windowed). **C.** Heatmap of H3K9me3, H3K9me3S10ph and HP1β enrichment at repeat subfamilies in ESCs, sorted alphabetically by family name. H3K9me3-marked repeats are also marked with H3K9me3S10ph and HP1β throughout the cell cycle, likely reflecting dynamic turnover of H3S10ph and HP1β binding when H3K9me3 is present.

4.2.4 HP1β is bound at a subset of regions enriched for H3K9me3 and H3K9me3S10ph

While H3K9me3 and H3K9me3S10ph enrichment were correlated throughout the genome, HP1β showed preference for a subset of H3K9me3-marked regions (**Fig. 4.4A**). Consistent with the observation that all H3K9me3-marked regions can be phosphorylated at S10, analysis of H3K9me3 vs. H3K9me3S10ph at all ERV subfamilies reveal similar trends in enrichment (**Fig. 4.5A, B**). Indeed, comparison of H3K9me3 and HP1β binding reveals that for the majority of ERV subfamilies, the binding of HP1β scales quantitatively with H3K9me3 enrichment using linear regression. There were some exceptions to the near 1:1 correlation of H3K9me3/HP1β. For example, a subset of IAP subfamilies were disproportionately more enriched for HP1β. Class I and class II LTRs are also disproportionately enriched for H3K9me3/HP1β. Depending on the specific subfamily; some are enriched for H3K9me3 but low HP1β (e.g. RLTR1B), while others are enriched with HP1β but show low H3K9me3 (RLTR4, the LTR for MLV). Taken together, these results confirm that HP1β co-localizes with H3K9me3 at most transposable elements, with some exceptional ERV family showing biased HP1β enrichment that cannot be explained by the presence of H3K9me3S10ph.

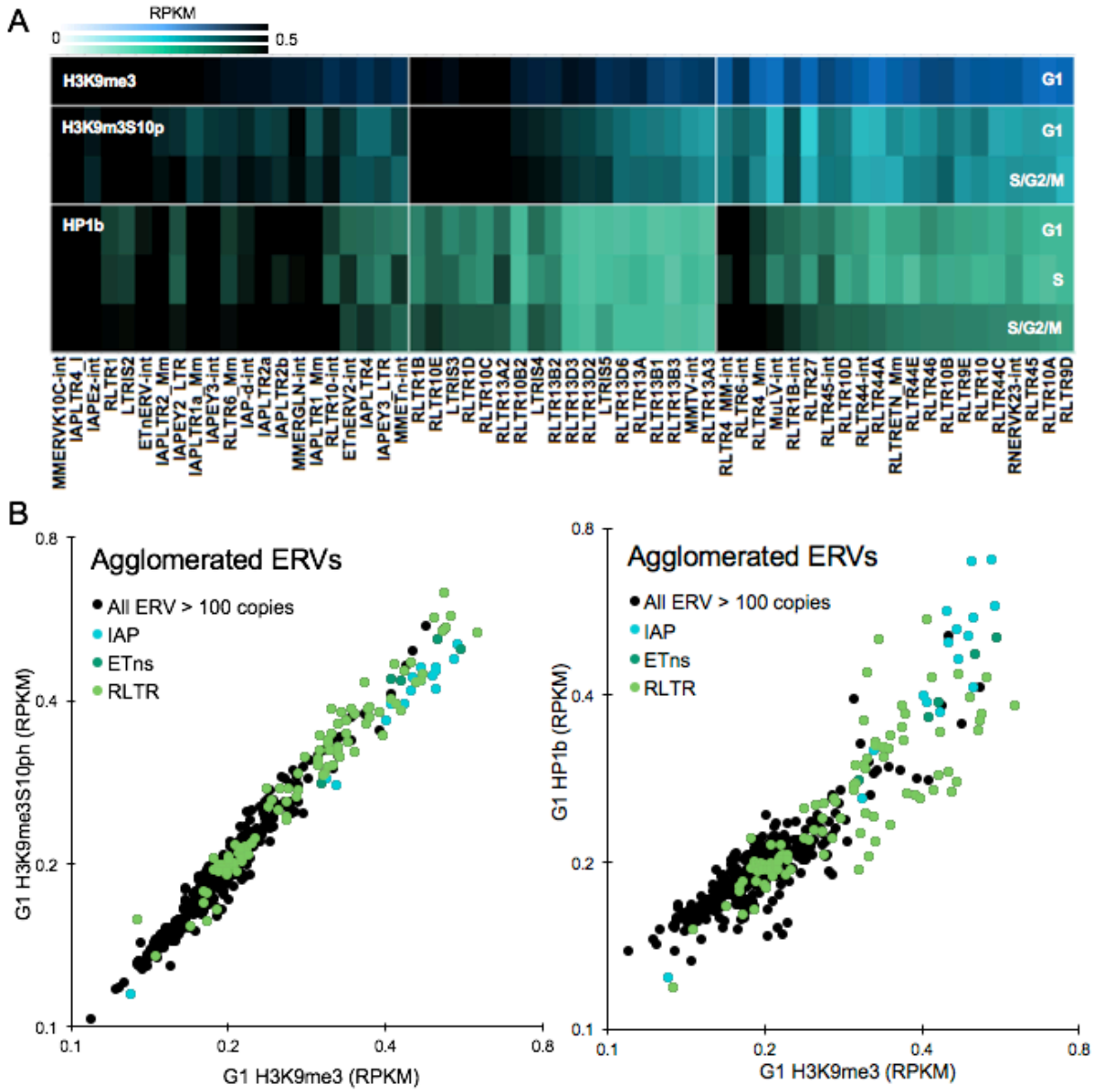


Figure 4.5 Enrichment of H3K9me3, H3K9me3S10ph and HP1β at ERV subfamilies.
A. Heatmap of a subset of class I and II ERV subfamilies with enrichment of H3K9me3, H3K9me3S10ph, and HP1β, clustered by enrichment patterns. IAP, MMERVK10C and ETns are uniformly enriched for all three marks, while RLTRs show differential H3K9me3 vs. HP1β enrichment. **B.** Pairwise comparison of H3K9me3 vs. H3K9me3S10ph and HP1β enrichment at all ERV subfamilies with over 100 copies, showing near identical enrichment pattern between H3K9me3 and H3K9me3S10ph, and modest difference between H3K9me3 and HP1β.

4.2.5 HP1β is not enriched at H3K9me3-marked germline genes

We have previously documented that in addition to ERVs, SETDB1-deposited H3K9me3 mediates transcriptional repression of a subset of germline-specific genes in ESCs (Karimi, 2011). Indeed, I confirmed robust H3K9me3 signal in G1 cells at promoters of several silent piRNA biogenesis genes, including *Ddx4, Dazl,* and *Mael* (**Fig. 4.6A**). Like the repeats, these germline promoters are also robustly marked with H3K9me3S10ph, indicating that phosphorylation of H3 tails by Aurora B also occurs at genic regions marked with H3K9me3.

HP1β enrichment at such germline promoters however, is surprisingly low in all phases of the ESC cell cycle. Indeed, a 2D scatterplot of the relationship between of HP1β and H3K9me3 at TSSs revealed that a subset of promoter regions show relatively low levels of HP1β enrichment, $R^2 = 0.755$ (**Fig. 4.6B**). Our analysis revealed that ~1,235 gene promoters show skewed H3K9me3 or HP1β enrichment ($z > 0.25$). Amongst the 619 genes that were enriched for H3K9me3 but not HP1β, piRNA metabolic process were strongly overrepresented in Gene Ontology analysis, with 8 out of 15 annotated piRNA genes showing this pattern (**Fig. 4.6C**). Other highly significant ontologies were all related to gametogenesis, including DNA methylation in gametogenesis, synapsis and spermatogenesis. On the other hand, promoter regions that were not biased for H3K9me3 or HP1β enrichment (n= 2,064) include large repetitive gene families, such as olfactory (*Olfr*), vomeronasal (*Vmn1r*), taste receptors (*Tas2r)*, and zinc finger proteins (*Zfp)*. Finally, HP1β-biased (n = 616) genes were weakly enriched for developmental pathways – specifically renal and cardiovascular related ontologies.

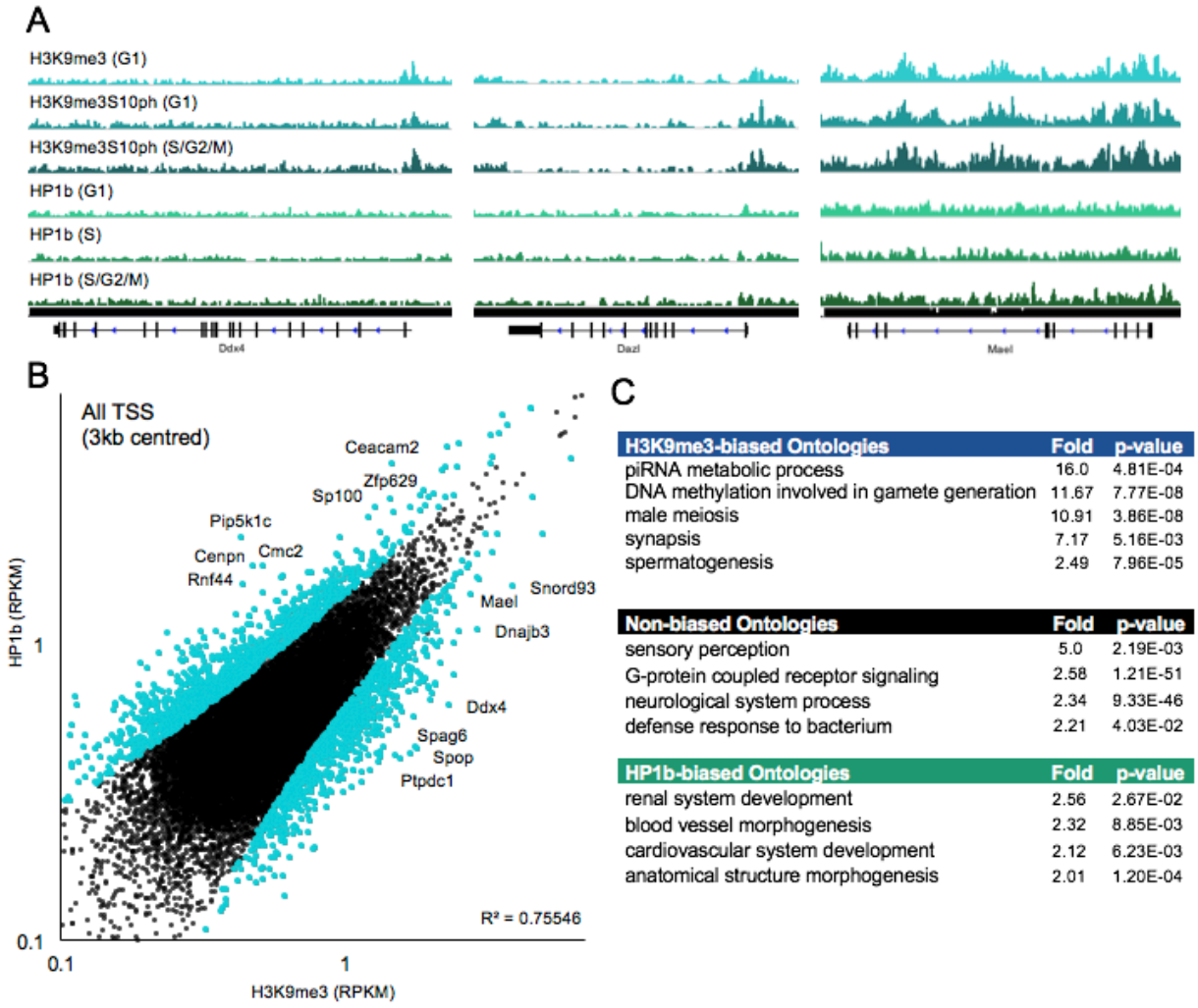


Figure 4.6 HP1β is excluded from H3K9me3-marked piRNA and germline-specific promoters.
A. Genome browser snapshot of H3K9me3, H3K9me3S10ph and HP1β enrichment at testis-specific piRNA biogenesis genes *Ddx4, Dazl,* and *Mael.* **B.** Pair-wise comparison of H3K9me3 and HP1β occupancy at all annotated TSS (3kb centred), coloured points denote TSS that were significantly biased towards H3K9me3 or HP1β. R^2: Pearson correlation. **C.** Ontologies of genic promoters that showed H3K9me3- or HP1β-skewed enrichment versus non-biased promoters. Note that H3K9me3-biased genes showed greater fold enrichmen and lower p-value scores for specific ontologies compared to non-biased or HP1β genes.

4.2.6 ATRX preferentially localizes to H3K9me3-marked regions not bound by HP1β

In addition to HP1 proteins, a number of "readers" have recently been identified that bind to H3 peptides with H3K9me3. Intriguingly, ATRX was recently shown to bind with high affinity to H3K9me3S10ph, as well as H3K9me3-marked peptides (Noh et al. 2014). Binding of such

alternative H3K9me readers may explain why a subset of H3K9me3-marked regions are relatively depleted of HP1β enrichment. To test this hypothesis I examined published ATRX ChIP-seq data generated from ESCs (Law et al. 2010). Consistent with Law et al., ATRX shows robust localization to centromeric and telomeric repeats, and was weakly-enriched at H3K9me3-marked tandem repeats within gene-dense regions. Interestingly, when H3K9me3 or H3K9me3S10ph-marked regions were stratified into three categories of HP1β occupancy – high, low, and absent, I found that the percentage of regions occupied by ATRX was highest in HP1β-depleted sites (**Fig. 4.7**). These results suggest that H3K9me3 readers may compete for their nucleosomal substrates in a sequence-dependent manner, with HP1β being preferentially enriched at ERVs and ATRX preferentially enriched at tandem repeats.

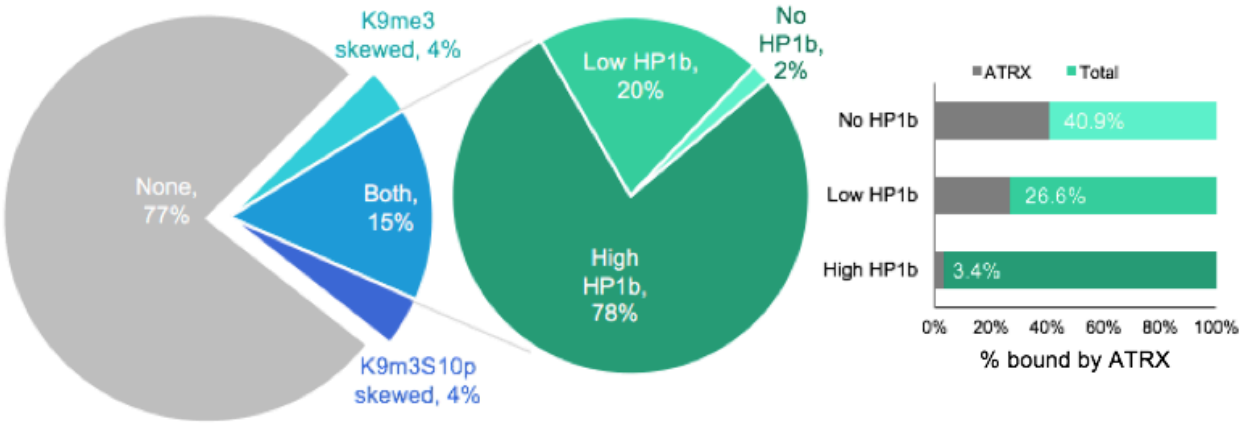


Figure 4.7 Occupancy of HP1β & ATRX at H3K9me3-marked regions.
H3K9me3 or H3K9me3S10ph marks ~23% of the mouse genome in ESCs. The H3K9me3-marked fraction was further stratified into regions with high, low and no HP1β - showing that HP1β localizes to ~80% of H3K9me3-marked regions. These categories of H3K9me3-marked regions were then intersected with ATRX occupancy, and percentage of $ATRX^{+}$ sites are represented in the bar graph. Notice ATRX predominantly occupy H3K9me3 regions that have low or no HP1β.

4.2.7 Analysis of H3K9me3-marked germline promoters

As H3K9me3-marked genes involved in piRNA biogenesis and spermatogenesis showed relatively low levels of HP1β in ES cells, I analyzed other chromatin features at these specific genes (**Fig. 4.8A**). While *Mov10l1*, *Fkbp6* and *Dpep3* are enriched for both KAP1 and SETDB1, the majority of piRNA/spermatogenesis genes are bound by SETDB1 but not KAP1. This is consistent with the observation that specific germline genes are de-repressed in *Setdb1*$^{-/-}$ but not *Kap1*$^{-/-}$ ESCs (Karimi et al. 2011; Rowe et al. 2010). I then determined whether HP1α, HP1γ or ATRX are bound to these promoters and found that these H3K9me3 readers were not enriched at these regions. As the phospho-methyl switch was previously reported to additionally require H3K14ac to completely displace HP1 from mitotic chromatin (Mateescu et al. 2004), I also analyzed a recently published H3K14ac ChIP-seq dataset (Karmodiya et al. 2012). Many germline promoters were indeed strongly enriched for H3K14ac. Interestingly, a subset of these genes, including *Gpa33*, *Tex19.1* and *Dppa3*, also show relatively high levels of expression despite bearing repressive H3K9me3, perhaps reflecting heterogeneity in the population in ESCs, with a subset harboring H3K9me3 at these genes and a distinct subset harboring H3K14ac. Taken together, the analysis of chromatin features at H3K9me3-marked germline genes revealed that these genes may not be stably repressed in WT ESCs by SETDB1, given that they lack KAP1 and all examined H3K9me3 readers, consistent with the observation that these genes also bear considerable amount of H3K14ac and are lowly transcribed in WT ESCs.

4.2.8 Identification of cis motifs from H3K9me3-marked germline-specific promoters

The KRAB-ZFPs/KAP1 complex is essential for the recruitment of SETDB1 to specific genomic loci, including imprinting control centres, zinc finger clusters, and retroviral-derived sequences

(Quenneville et al. 2011; Frietze et al. 2010; Wolf and Goff 2009; Rowe et al. 2010; Matsui et al. 2010). Given that these germline genes were repressed in a KAP1-independent manner, it is unlikely that SETDB1 is recruited by a KRAB-Zfp to these genic promoters. Thus, I sought to identify a common sequence motif that may indicate binding of a sequence-specific DNA binding factor responsible for recruiting SETDB1. I first parsed all H3K9me3-marked genes and intersected those that were specifically expressed in the testis or ovaries, according to Tissue Specific Expression Analysis (TSEA) (Dougherty et al. 2010), yielding 331 H3K9me3-marked germline-specific genes. Using RefSeq annotation, I parsed 1kb upstream of the TSS, masked known repeats and queried common sequence motifs (50bp in length) using MEME (Bailey et al. 2009). While 10 total motifs were generated, 240 out the 330 queried genes contained motif 1, 2 or 3 (**Fig. 4.8B-C**). Motif 1 and 2 were especially prevalent, but did not show strong sequence specificity – examination of individual promoter sequence alignment revealed that motif 1 and 2 are tandem monomeric repeats, being C-rich and T-rich respectively. Interestingly, promoter sequences that aligned to motif 1 tend to yield multiple imperfect matching genomic coordinates using BLAT, while motif 2 sequences tend to be unique and did not yield multiple coordinates - suggesting motif 1 may represent a repetitive sequence interspersed throughout the genome.

The top 3 motifs were further analyzed for transcription factor (TF) enrichment using Tomtom motif comparison tool (Bailey et al. 2009). TFs identified by such analysis on the three motifs yield several known regulators of gametogenesis. Strikingly, for a subset of the candidate TFs identified, including ZNF148, CPEB1, FOXO1, SRY, ZFP105, SOX17 , and ZFX, abnormal spermatogenesis phenotypes have already been reported (Takeuchi et al. 2003; Sousa Martins et al. 2016; Goertz et al. 2011; Zhou et al. 2010; Ng et al. 2013). E2F6 was previously

shown to recruit DNMT3b to germline promoters including *Mael* (Velasco et al. 2010), suggesting that motif 1 may be necessary for recruitment of DNMT3 in somatic cells. TFs enriched for motif 2, conversely, included master regulators of sex differentiation – such as SRY and SOX17, and may act at tissue-specific enhancer elements. Finally, motif 3 contained several CpG dinucleotides, likely representing the CpG island, given its immediate proximity to the TSS. We have previously shown that the transcriptional repression of these germline genes uniquely depends on both SETDB1 and DNA methylation (Karimi et al. 2011). Notably, several of the top TFs identified in motif 3, including NRF1 and CTCF, have previously been shown to only bind to unmethylated motifs to promote transcription (Domcke et al. 2015; Filippova et al. 2001). Taken together, the male germline genes show atypical chromatin signatures of SETDB1-dependent silencing, and *in silico* motif analysis yield two potential consensus sequences, consisting of simple mononucleotide tracts, that could act as repressive *cis* regulatory regions in ESCs.

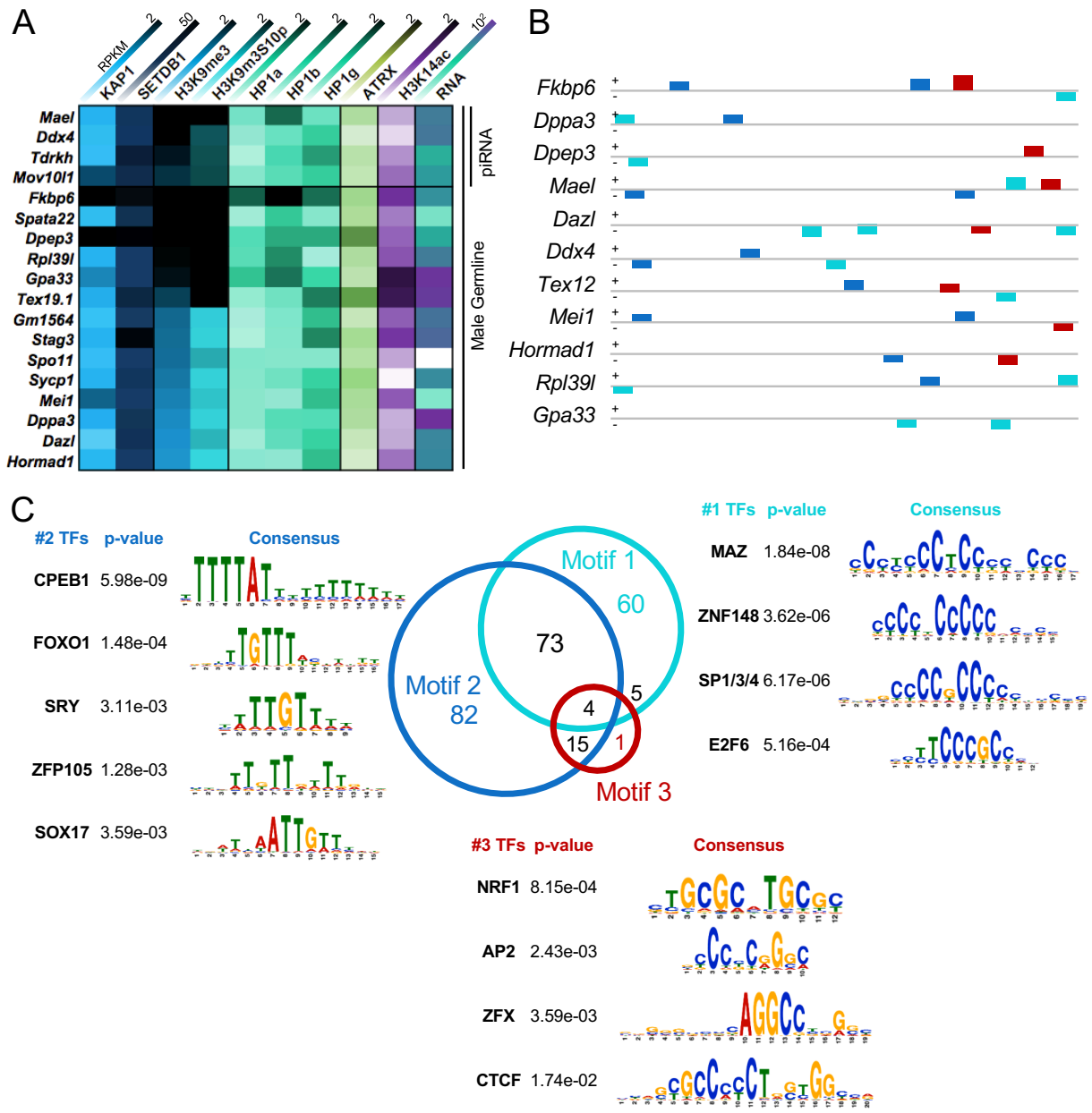


Figure 4.8 Chromatin and sequence features of H3K9me3-marked testis specific genes.
A. Heatmap of various chromatin features at SETDB1-targeted germline promoters, with expression in WT ESCs. Note that while SETDB1 localized to these genic promoters, KAP1 and HP1s are not. **B.** Presence, orientation, and locations of the top 3 consensus motifs for a subset of H3K9me3-marked germline promoters as identified by MEME. Height of the box denotes alignment score to identified motif. Grey line represents 1kb upstream of TSS, with + and – denoting plus- and minus- stranded orientation, respectively. Motif 1 is represented by teal boxes. Motif 2 is represented by blue boxes. Motif 3 is represented by red boxes. **C.** TFs that were significantly enriched for the Top 3 consensus sequences as identified by Tomtom. Venn diagram depicts the number of genes that possess the Top 3 consensus sequence out of the 331 promoters queried. Motif are as coloured in subpanel B. p-values represents the probability that the TF match to the input motif occurred randomly by chance.

4.3 Discussion

4.3.1 Aurora-dependent phosphorylation of H3K9me3S10 persists in G1

The spatiotemporal pattern of mitotic H3S10ph is tightly regulated by the coordinate reciprocal activities of Aurora B kinase and Protein Phosphatase 1. Thus, the finding that Aurora B-dependent H3K9me3S10ph persists in G1 is somewhat surprising. Regardless, the observation that H3K9me3 and H3K9me3S10ph show nearly identical profiles genome-wide reveals that this is a global phenomenon. Intriguingly, in asynchronous cells, >60% of nucleosomes at IAP ERVs are marked with H3K9me3S10ph, exceeding H3K9me3 in interphase ESCs. This is independently corroborated by mass spectrometry data, where twice the level of H3K9me3S10ph peptide was recovered compared to H3K9me3 peptide in ESCs (Jung et al. 2010).

There are two possible explanations for the observation that Aurora B-dependent H3S10ph persists in G1 – first, Aurora B may have activity towards H3K9me3-marked nucleosomes in interphase, and second, dephosphorylation by PP1 may be inefficient specifically on H3K9me3S10ph-marked H3 tails. Aurora B is highly expressed in ESCs and interphase activity of Aurora B was previously shown in highly proliferative cells (Hayashi-Takanaka et al. 2009). Based on the recruitment pattern of Aurora B and emergence of H3S10ph in G2 on pericentromeric heterochromatin (Crosio et al. 2002), it is possible that Aurora B preferentially phosphorylates H3K9me3-marked histones. Given that the kinase activity of Aurora B is highly promiscuous in G2, forming a spatial phosphorylation gradient radiating from the midzone (Fuller et al. 2008), the temporal pattern of Aurora B-dependent H3S10ph in anticipation of mitosis is more likely due to activity focused at centromeric spindle attachment sites rather than H3K9me3-marked regions *per se*. Consistent with this hypothesis, ablation of H3K9me3 in

Suv39h1/2 DKO MEF results in only modest dispersion of H3S10ph at DAPI-dense heterochromatin (Peters et al. 2001).

The abundance of Aurora-mediated H3K9me3S10ph in interphase, specifically in G1, is therefore likely not due to preferential targeting, but rather incomplete dephosphorylation by PP1 on exit from mitosis. Interestingly, such histone PTM antagonism has been documented for the H3K9me2/3 demethylase JMJD2A, which cannot demethylate H3K9me3S10ph (Ng et al. 2007), however, it is not known whether PP1 is inhibited by neighbouring methyl lysines. It is also possible that H3K9me3S10ph is masked by H3K9me readers bound during mitosis (**Fig. 4.9**). Intriguingly, both NP95 and ATRX are stably retained on mitotic chromosomes (Rothbart et al. 2012; McDowell et al. 1999a), and are potential candidates for future investigation into the mechanism of H3S10ph retention in interphase.

4.3.2 Model: Aurora B-dependent H3K9me3S10ph during the ESC cell cycle

Curiously, H3K9me3 and H3K9me3S10ph antibodies bind to mutually exclusive pools of histone tails, as evident by peptide blocking assays. Therefore, the near-perfect correlation between the ChIP-seq datasets in G1 ESCs would suggest ongoing turnover of either modification (**Fig. 4.9**). Mass spectrometry data in ESCs additionally revealed that H3K9me3S10ph is only found on H3.2 and not on H3.3, whereas H3K9me3 is present on both H3.3 and H3.2 (Jung et al. 2010). This is consistent with our model, describing the sequence of events during the cell cycle. I postulate that after deposition of nascent H3.2 during S-phase, KRAB-ZFPs target specific sequences and tether the HP1/KAP1/SETDB1 complex to deposit H3K9me3. At late G2/M, Aurora B is activated and phosphorylates all nucleosomes, converting

H3K9me3 to H3K9me3S10ph. This coincides with the release of HP1 proteins and concomitantly promotes binding of phospho-tolerant H3K9me3 readers, such as ATRX/DAXX and NP95, to mitotic chromatin. At telophase, most of the Aurora B-dependent H3S10ph is dephosphorylated by PP1, except for sites where the epitope is masked by the presence of such readers. ATRX/DAXX dependent deposition of H3.3 at IAP ERVs is subsequently followed by SETDB1 dependent H3K9me3 in the absence of H3S10ph on the same tail (Elsässer et al. 2015) , thus leading to co-occurrence of both H3K9me3 and H3K9me3S10ph at the same genomic location in G1 cells. Given that H3K9me3S10ph is incompatible with HP1β recognition of H3K9me3, the finding that HP1β localizes robustly to H3K9me3S10ph-marked regions is probably reflective of the presence of H3K9me3 alone at the same regions in a minority of cells in the population.

4.3.3 H3K9me3S10ph influences the binding of H3K9me3 readers

Genome-wide, HP1β localizes to 80% of H3K9me3-marked regions, suggesting it is a faithful H3K9me3 reader. Pairwise comparison of H3K9me3 and HP1β enrichment revealed preferential targeting of HP1β to a subset of ERVs, i.e. IAPs & ETns, somewhat irrespective of the H3K9me3 enrichment. The bias amongst H3K9me3-marked ERVs indicates chromatin stability of HP1β may additionally depend on other factors. One possibility is that HP1 proteins, recruited by H3K9me3 recognition, are additionally stabilized by interactions with KAP1 via the PxVxL motif (Cammas et al. 2007).

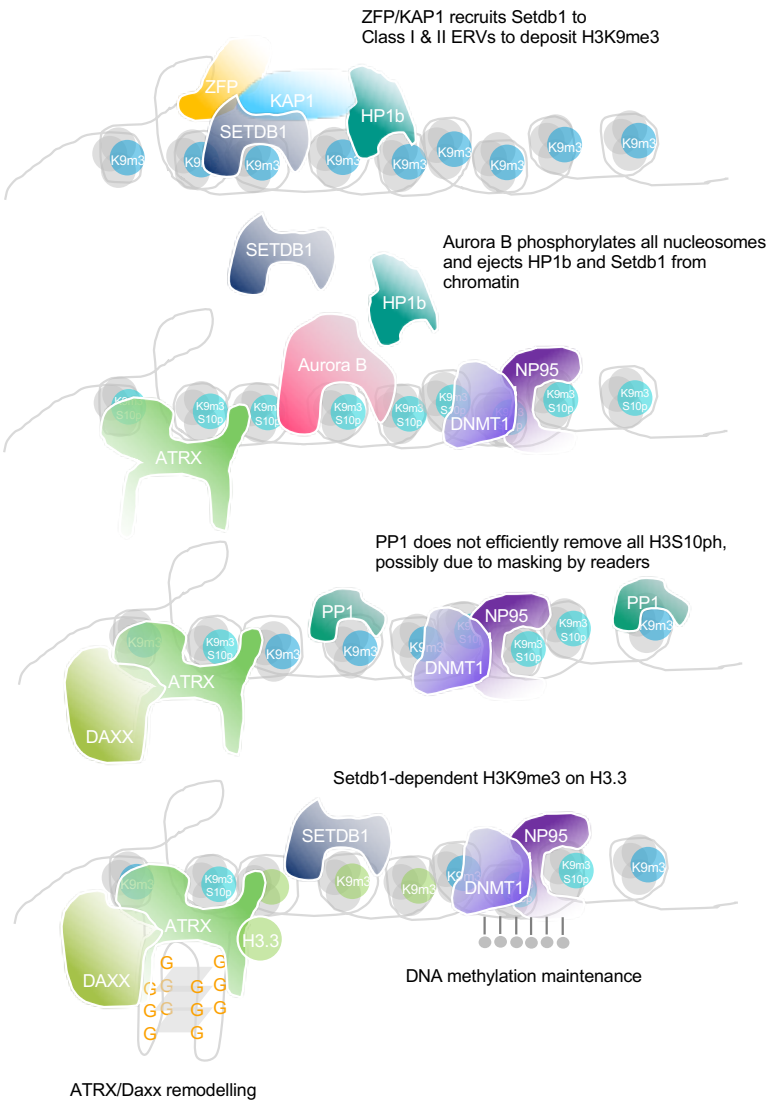


Figure 4.9 Model for co-existence of H3K9me3 and H3K9me3S10ph at young endogenous retroviruses in ESCs.

Sequences within class I & II ERVs, such as IAP or ETn elements, are recognized by specific KRAB-ZFPs, which then recruit the repressive KAP1/SETDB1 complex for H3K9me3 deposition, and in turn HP1 recruitment. During mitosis, Aurora B phosphorylates all H3 nucleosomes, converting H3K9me3 to H3K9me3S10ph. This phosphorylation then displaces HP1 from mitotic chromatin, and allows the loading and retention of phospho-tolerant H3K9me3 readers, such as NP95 or ATRX/DAXX, during chromosome segregation. As the cell enters telophase, most nucleosomes are dephosphorylated by PP1, except those that are bound by H3K9me3 reader proteins that are tolerant of H3S10ph – leading to persistence of H3K9me3S10ph in the following G1. As such, a heterogeneous population of H3K9me3- and H3K9me3S10ph-marked nucleosomes may facilitate a diversity of chromatin readers for heterochromatin maintenance at ERVs. These functions include DNA methylation maintenance by NP95, which is a necessary co-factor of DNMT1, and nucleosome remodeling at non-B-form DNA (e.g. G quadruduplexes) by the ATRX/DAXX complex. As a consequence of such remodelling, nascent H3.3 is then deposited by ATRX/DAXX at the ERVs and targeted by SETDB1 for H3K9me3.

Conversely, NP95 and ATRX are known readers of SETDB1-dependent H3K9me3 at ERVs that are tolerant of H3K9me3S10ph (Rothbart et al. 2012; Elsässer et al. 2015; Sharif et al. 2016). In addition, H3K9me3-binding of ATRX and HP1 has been shown to occur in a mutually exclusive manner *in vitro*, with excess HP1 displacing ATRX from H3K9me3 peptide tails (Eustermann et al. 2011). Indeed, in primary neurons, localization of ATRX to pericentromeric heterochromatin occurs only after H3K9me3S10 tails are converted to H3K9me3S10ph following forskolin stimulation (Noh et al. 2014). Importantly, this phosphorylation event coincides with ejection of HP1β, suggesting that ATRX cannot access H3K9me3 sites that are HP1β-bound. In ES cells, ATRX and HP1β show preference for distinct H3K9me3-marked genomic locations, with HP1β occupying ~80% of H3K9me3-marked regions. Thus, the presence of H3K9me3S10ph nucleosomes may counteract the "monopolizing" effect of HP1β, and promote loading of other heterochromatin factors (**Fig. 4.9**). Consistent with this model, introduction of a mutant of UHRF1/NP95 that abolishes H3S10ph tolerance (i.e. has a binding preference for H3K9me3, like HP1 proteins) results in loss of NP95 and DNMT1 from mitotic chromatin and defects in DNA methylation at rDNA (Rothbart et al. 2012); whether H3S10ph is similarly relevant for ATRX will be an interesting avenue of future research.

4.3.4 Atypical signatures of SETDB1-dependent silencing on male germline promoters

SETDB1-dependent H3K9me3 at genes shows two distinct patterns. At *Zfn* and *Olfr* genes, H3K9me3 is enriched over the entire length of the gene body (Frietze et al. 2010; Vogel et al. 2006); such intragenic SETDB1-dependent H3K9me3 does not seem to negatively impact expression. In contrast, a small subset of germline-specific genes essential for gametogenesis, including genes involved in piRNA biogenesis (*Mael, Ddx4)* and meiotic synapsis (*Sycp1,*

Mei1), show localized/punctate peaks of H3K9me3 over the promoter regions. Intriguingly, such genes are upregulated upon depletion of SETDB1 (Yuan et al. 2009; Karimi et al. 2011).

Most of these germline-specific promoters fall within the exceptional 2% of H3K9me3-marked regions that are devoid of HP1β at the promoter, and are not transcriptionally induced upon genetic ablation of HP1s (Rowe et al. 2013; Maksakova et al. 2013). It has been suggested that a third PTM, H3K14ac, in conjunction with H3S10ph, is required for efficient release of HP1 from mitotic chromatin (Mateescu et al. 2004). Indeed, I found that germline promoters that bear H3K9me3 and H3K9me3S10ph, but not HP1β, are generally marked with H3K14ac. Interestingly, transcription is also relatively high in WT ESCs for a subset of these genes, such as *Gpa33*, *Tex19.1* and *Dppa3*, despite bearing repressive H3K9me3 in the promoter region. The simultaneous detection of repressive and activating marks is likely due to cell-to-cell heterogeneity in serum-grown ESC cultures. That said, significant levels of H3K9me3S10phK14ac have been observed using mass spectrometry (Jung et al. 2010), suggesting that the trivalent histone PTM can occur on the same histone tail in ESCs. H3K9me3S10ph ChIP, followed by H3K14ac reChIP should address whether the trivalent mark is found at these specific genic promoters.

Curiously, these germline gene promoters are also apparently not bound by the co-repressor KAP1, and are not dependent on KAP1 for repression (Rowe et al. 2010), suggesting that SETDB1 is recruited to genic promoters independently of KRAB-ZFPs. To investigate this further, I reasoned that as nearly all SETDB1-dependent H3K9me3 is found over LTR elements, it is possible that the promoter sequences of these genes bear a consensus motif that is also

present in such retrotransposons. The top two consensus sequence from H3K9me3-marked germline promoters consist of simple nucleotide repeat. Interestingly, motif 1 is strongly predicted to form G-quadruduplex (G4) structures on the reverse complement orientation, which has interesting implications for recruiting hnRNP K, which can bind to G4 DNA (Law et al. 2010; Zamiri et al. 2015) and is functionally relevant for SETDB1-dependent silencing of germline genes (Thompson et al. 2015). Interestingly, the 3rd ranking motif identified is a consensus sequence for ubiquitous transcription factors, e.g. NRF1 and CTCF, that have previously been shown to exhibit sensitivity to the presence of DNA methylation *in vitro* (Domcke et al. 2015; Filippova et al. 2001). In support of the hypothesis that DNA methylation at these promoters is necessary to prevent aberrant activation by suppressing TF-binding, a subset of these germline genes are also de-repressed in ESCs deficient in all three DNMTs (Karimi et al. 2011). Future studies aimed at characterizing the relevance of these motifs for SETDB1-dependent silencing are clearly of interest. For example, integration of a stable reporter with and without the relevant *cis* motif elements in *Setdb1* cKO ESCs should yield functional insights to whether these motifs are required for SETDB1-dependent repression of these gametogenesis genes.

5. ROLE OF MSKS & RSKS IN ESCS

5.1 Introduction

Dynamic changes in chromatin structure are required for the orchestration of gene expression in response to developmental or environmental cues. ESCs, which are derived from the inner cell mass (ICM) of the E3.5 blastocyst, possess the capacity to differentiate into every cell type in the embryo proper, including germ cells. The exit from pluripotency can be induced by the mitogen Fibroblast Growth Factor 4 (FGF4), which triggers downstream ERK1/2 to promote differentiation and lineage commitment (Kunath et al. 2007; Stavridis et al. 2007). Genetic ablation of FGF or ERK in ESCs notably hinders their differentiation capacity towards all three germ layers, which underscores the requirement of mitogen-stimulated signalling in directing cell fate (Hamilton and Brickman 2014; Qi et al. 2004). The reliance on mitogenic signalling was recently exploited in the development of the "2i media", which inhibits differentiation signals exerted by MEK/ ERK to maintain ESCs in a naive "ground state" (Ying et al. 2008). Despite mounting evidence reinforcing the importance of ERK signalling in lineage commitment, the role and necessity of downstream MAPK kinases, such as MSKs and RSKs, remains poorly defined in this process.

The eventual endpoint of mitogen signalling is thought to be phosphorylation of transcription and chromatin factors in the nucleus, which then releases paused Pol II at the serum-response genes (Anjum and Blenis 2008). This rapid remodeling of chromatin, termed "the nucleosomal response", specifically involves the phosphorylation of H3 at S10 and S28 at

the promoters of immediate-early genes, such as *c-fos*, *jun, c-myc,* and *Egr1* in quiescent cells (Mahadevan et al. 1991; Thomson et al. 1999b). The protein products of immediate early genes, including members of the AP-1 family of transcription factors, then initiate downstream gene expression for re-entry into the cell cycle in G1. While studies have highlighted the importance of ERK signalling in cell proliferation and oncogenesis (Roux and Blenis 2004), an outstanding question is whether ERK signalling potentiates lineage commitment, at least in part, via the direct phosphorylation of chromatin.

As discussed in the Introduction, MSK and RSK families act downstream of EGF/ERK to phosphorylate chromatin and transcription factors. These closely-related families of p90 ribosomal S6 kinases are the closest mammalian paralogs to JIL-1, sharing ~60% and ~50% similarity in the kinase domains. MSK1/2, RSK1/2/3 and are capable of phosphorylating H3S10 *in vitro* (Soloaga et al. 2003; Sassone-Corsi et al. 1999; Hu et al. 2004; Zhao et al. 1995). However, the role of the MSK and RSK families in phosphorylating H3S10 in interphase ESCs has not been addressed, nor has their influence on transcription been studied on a genome-wide scale. Although RSK2 was previously implicated as an H3S10 kinase in ESCs and fibroblasts (Sassone-Corsi et al. 1999), MSKs are likely the major kinases responsible for phosphorylating H3S10 following induction, due to their higher affinity for the substrate, and the persistence of H3S10ph in fibroblasts derived from Coffin-Lowry patients (Soloaga et al. 2003). As described in **Chapter 3,** I found that H3S10ph broadly marks early replicating euchromatin in interphase ESCs. Here, I wished to determine whether MSK and/or RSK family members play a role in this process.

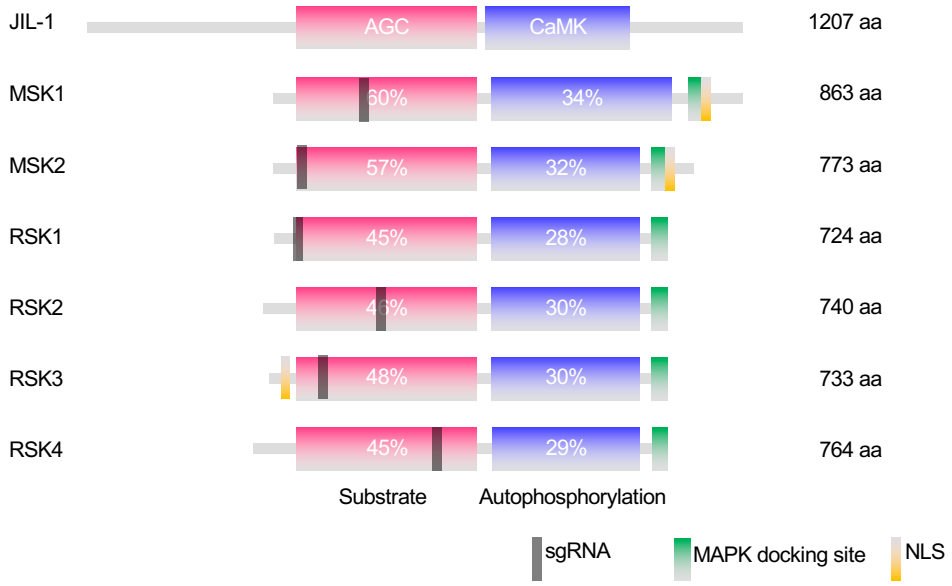
Therefore I systematically surveyed the individual and combined contribution of the p90 ribosomal S6 kinases (*Rps6ka1/2/3/4/5/6*) to interphase H3S10ph in ESCs. I utilized the recently developed Cas9-CRISPR technology to derive single and combinatorial KO ESCs lines, and subsequently conducted H3S10ph ChIP-seq and mRNA-seq on these lines. The depletion of *Msk1/2* or *Rsk1/2/3/4* either singly, or in combination, did not notably impact ESCs cell cycle progression. Surprisingly, *Msk/Rsk* KO did not notably disrupt interphase H3S10ph at gene-dense regions, suggesting the broad euchromatin H3S10ph domains are maintained by alternate kinase(s). While the profile of H3S10ph in interphase was not dramatically altered in any of the mutants analyzed, I identified distinct signatures of H3S10ph loss associated with the individual KOs at specific proximal promoters and distal enhancers. Promoters and enhancers of immediate early and ribosomal biogenesis genes were amongst the regions that showed reduced H3S10ph across the panel of kinase mutants, affirming the previously characterized role of MSK and RSK. *Msk2* ablation showed the greatest loss of interphase H3S10ph at active promoter/enhancers. Unexpectedly, deletion of *Msk2* in ESCs down-regulated a subset of SETDB1/H3K9me3 repressed targets - including imprinted genes, young endogenous retroviruses and germline genes. Lastly, embryoid bodies derived from *Msk* and *Rsk* KO ESCs showed increased propensity for G1-arrest upon LIF-withdrawal, suggesting that MSK/RSK may facilitate ESC differentiation. Altogether, these experiments represent the first systematic study on the functional relevance of MSK/RSK family members on global H3S10ph and transcription in ESCs, and potentially uncover novel crosstalk between MSK2/H3S10ph and SETDB1/H3K9me3 in gene regulation.

5.2 Results

5.2.1 Derivation of kinase knockouts using the CRISPR/Cas9 system

To survey the individual and combined contribution of the kinase family to interphase H3S10ph in ESCs, I utilized the recently developed CRISPR/Cas9 system to derive single and combinatorial KO ESC lines (Hsu et al. 2014). I first isolated several clonal lines from a previously established TT2 ESC line stably expressing the FUCCI reporter (see **Chapter 3**), which allows for cell cycle sorting of cells derived from this line. I chose a clonal line, Fu.7, that showed a cell size distribution and cell cycle timing similar to that of the TT2 parent line. Next, short guide RNAs (sgRNA) were designed for each candidate kinase, including *Msk1/2* and *Rsk1/2/3/4*, that target an exonic region encoding the AGC domain (**Fig. 5.1A**) and cloned into an expression vector encoding Cas9 and puromycin for positive selection. Fu.7 ESCs were transfected with a single sgRNA-Cas9 vector, and selected for puromycin resistance for 48h prior to colony picking. To address potential redundancies within the kinase family, an ESC line, "6SK", that was deleted for all 6 kinases were simultaneously derived. 10-20 clones for each targeted kinase were genotyped using primers flanking the targeted exon, and Sanger sequenced to verify Cas9-mediated mutations. After eliminating clones that were abnormally large and likely polyploid, two or more independent clones for each kinase of interest with distinct mutations predicted to harbor large truncating and/or frameshift mutations were expanded and further characterized.

A



B

Clone	Msk1	Msk2	Rsk1	Rsk2	Rsk3	Rsk4	Other
Fu.7	+/+	+/+	+/+	+/Y	+/+	+/Y	T1, 8, 11
Msk1-1	-/- (4ins)						PuroR
Msk1-3	-/- (1ins)						Xq dup
Msk2-5		-/- (163del)					
Msk2-16		-/- (146del)					Disomic 11
Rsk1-1			-/- (5del)				
Rsk1-2			-/- (Δex3)				
Rsk2-18				-/Y(17del)			
Rsk2-24				-/Y(82del)			
Rsk3-4					-/- (34del)		Y nullsomic
Rsk3-15					-/- (9del)		
Rsk4-4						-/Y (5del)	
Rsk4-15						-/Y (1del)	
6SK1520	-/- (1ins)	N.D.	-/- (C. Het)	-/Y (29del)	-/- (152del)	+/Y	

Figure 5.1 CRISPR targeting strategy for JIL-1 homolog kinases.
A. Protein domain structure of MSK & RSK family members. Drosophila JIL-1 is included for reference. The percentage shown reflects amino acid similarity to JIL-1 in the two kinase domains, with the N-terminal AGC domain, the effector kinase for histone H3, showing relatively higher homology across the kinases. MSKs possess canonical nuclear localization signals (NLS) at the C-terminal tail region, RSK3 also encodes a putative NLS signal at the N-terminus. RSK1/RSK2 are known to localize to the nucleus. Cas9 short guide RNAs (sgRNAs) were designed to target early exons to produce loss of function mutations. The approximate amino acid position of the targeting site is denoted by grey boxes. **B.** Panel of CRISPR-derived KO ESC lines and their associated genotype. N.D. denotes not detected in PCR. CNV differences as inferred by sequencing data is denoted. The parental Fu.7 were trisomic for chr1, 8, and 11, which were also observed in most of the KO lines. In addition, *Msk1*$^{-/-}$ showed amplification of Xq4, *Msk2*$^{-/-}$ reverted to disomy of chr11, and *Rsk3*$^{-/Y}$ lost chrY.

5.2.2 MSK and RSK kinases are dispensable for cell cycling of ESCs

None of the single kinase KO lines showed significant growth or overt morphology defects, consistent with the viability of the *Msk1/2* DKO mouse and the *Rsk1/2/3* TKO mouse (Soloaga et al. 2003; Ananieva et al. 2008; Anjum and Blenis 2008). The 6SK line showed modest morphological differences indicative of differentiation, but no growth defects. Cell cycle analysis further verified that most KO clones had normal cell cycle kinetics and were of comparable size to the parent Fu.7 line (**Fig. 5.2**). Some modest differences were observed in *Msk1* and *Rsk4* KO, but did not notably affect cell viability. Both *Msk1*$^{-/-}$ ESC lines were slightly smaller than their WT counterpart and showed relatively higher proportions of cells in G1 versus G2 as evident by both the FUCCI reporter distribution and DNA staining – suggesting loss of MSK1 leads to faster mitotic cycle progression. *Rsk2*$^{-/Y}$, conversely, were generally larger than WT cells, but showed no difference in cell cycle progression. *Rsk4*$^{-/Y}$ ESCs were also small, with a fraction of cells having <2n DNA content and modest accumulation at early S, suggestive of some replication-associated apoptosis. Loss of MSK2, RSK1, and RSK3 had no effect on cell size or cell cycle distribution. Nevertheless, all KO ESCs were maintained over multiple passages and proliferated at a similar rate to the parental WT line.

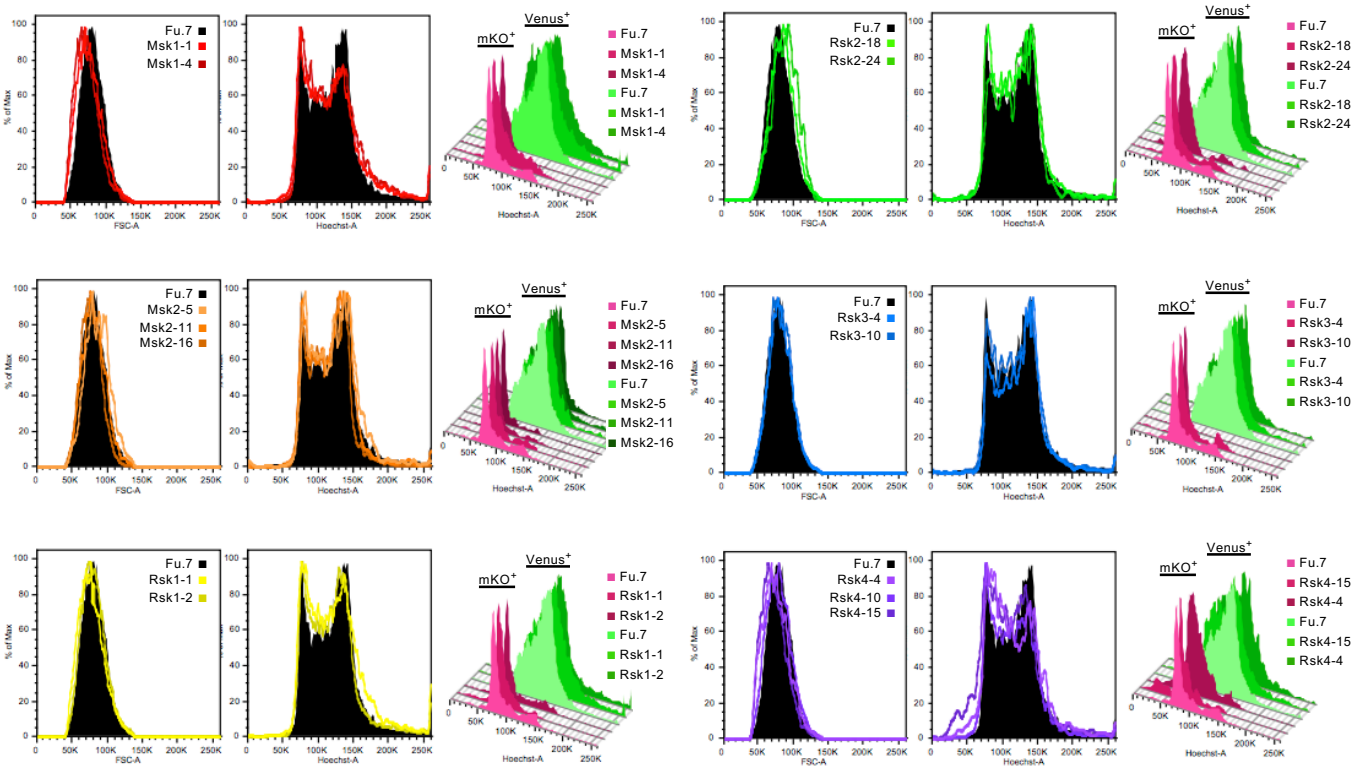


Figure 5.2 *Msk* & *Rsk* KO ESCs show normal cell cycling kinetics.
Flow cytometry analysis of *Msk1/2, Rsk1/2/3/4* mutant ESCs showing overtly normal cell size and cell cycle kinetics in the KOs. Histograms from left to right shows relative cell size in forward scatter (FSC), cell cycle distribution via Hoechst 33342 DNA staining, and the maintenance of FUCCI cell cycle fidelity, as demonstrated by DNA content in the mKO (G1-marker) or Venus (S/G2/M-marker) gated subpopulations, respectively.

5.2.3 Persistence of H3S10ph domains in Msk and Rsk KO

To further characterize the contribution of the kinases on interphase H3S10ph, I sorted 10^6 FUCCI cells in G1 (Venus$^-$) and performed H3S10ph ChIP-seq for each kinase KO line with the corresponding WT Fu.7 line. ChIP-seq revealed only modest changes in global H3S10ph domain structure in interphase in each of the single KO lines as well as the 6KO line, with the megabase-scale H3S10ph domain persisting at early-replicating genomic regions (**Fig. 5.3A**). These results are consistent with western blot analysis of H3S10ph in interphase ESCs, which indicates no discernable depletion in the KO lines (**Fig. 5.3B**). Genome-wide comparisons (tiled 5kb bins)

shows that global G1 H3S10ph levels are similar in all KO lines, with *Msk1*$^{-/-}$ being the most correlated with the WT Fu.7 line (R^2= 0.997) and 6SK1520 being the least correlated (R^2= 0.9) (**Fig. 5.3C**). The discordance observed in 6SK1520 (R^2=0.9), and *Rsk4*$^{-/Y}$ (R^2 = 0.95) is likely due to technical differences during sequencing library construction rather than true biological difference, as these libraries required more PCR amplification rounds than the others to yield material sufficient for sequencing.

5.2.4 MSK and RSK mediate H3S10ph at distinct subsets of promoters and enhancers

To specifically identify regions that showed H3S10ph loss in the KO lines, I compared RPKM over tiled 5kb bins in the genome. Due to the breadth of the domains of H3S10ph in interphase, I identified several karyotypic differences in some of the KO clones – notably a distal amplification of Xq4 in *Msk1*$^{-/-}$, trisomic 11 reversion in *Msk2*$^{-/-}$, and loss of chromosome Y in *Rsk3*$^{-/-}$. To minimize aneuploidy-related artefacts, any overt karyotypic difference as judged by non-equivalent read coverage, including chr11, chrX and chrY, were removed from all downstream analysis. Using a z-score threshold of -0.25, the mutant lines showed variable loss of H3S10ph across the genome, ranging from 1,525 bins in *Rsk2*$^{-/Y}$ to 7,650 bins in *Rsk4*$^{-/Y}$.

To then agnostically determine the functional relevance of these specific regions losing H3S10ph in the KO lines, the called regions were analyzed using GREAT (**Fig 5.3D, E**). Regions identified in *Msk2*$^{-/Y}$, *Rsk1*$^{-/-}$ and *Rsk3*$^{-/-}$, were overwhelmingly found within 50bp of a TSS. Conversely, regions losing H3S10ph in *Msk1*$^{-/-}$, *Rsk4*$^{-/Y}$ or 6SK1520 were generally more distant to any TSS, approximately 500bp or above. *Rsk2*$^{-/-}$ ESCs, interestingly, showed bias for H3S10ph loss at regions both immediate and proximal to the TSS. When oriented directionally

to the TSS, there was also a modest skew for regions downstream of the TSS for loss of H3S10ph in all KO lines. These results suggest that MSK2/RSK1/RSK3 may mediate H3S10ph at TSSs.

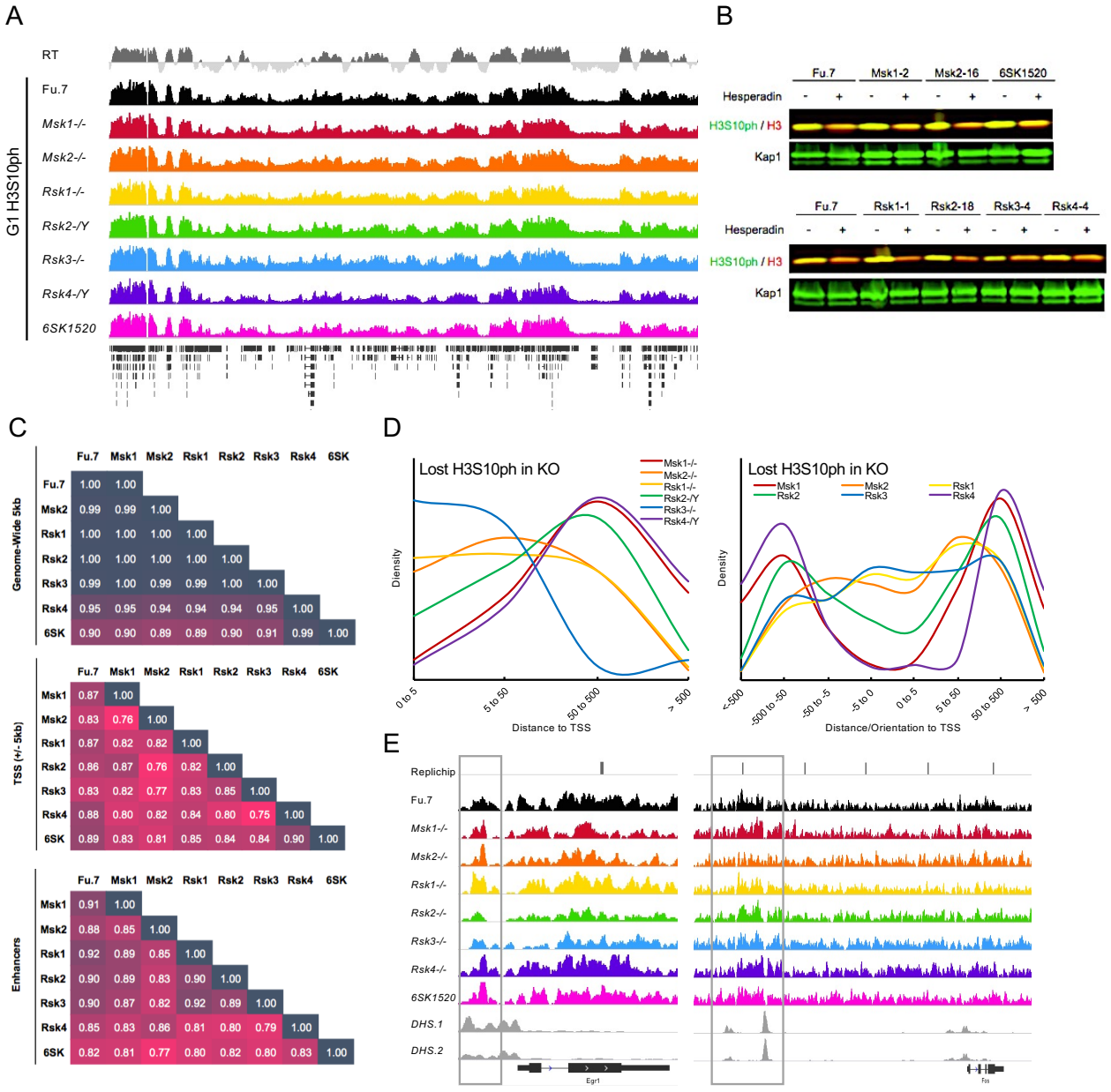


Figure 5.3 Genome-wide characterization of interphase H3S10ph in *Msk* & *Rsk* KO ESCs.
A. Genome browser screenshot of H3S10ph in G1-sorted *Msk/Rsk* KO ESCs compared to WT (Fu.7) showing H3S10ph domains at early-replicating regions are maintained in the single KO and compound mutant line 6SK1520. RT: Replication Timing, with top fraction representing early-replicating regions, and bottom fraction representing

late-replicating regions. **B.** Western blotting of mitotic and interphase H3S10ph in *Msk/Rsk* KO ESCs showing that bulk H3S10ph level does not change in the mutants. For hesperidin treated samples, cells were pre-treated with 200nM hesperidin for 3h to remove mitotic H3S10ph. KAP1 and pan H3 were used as loading controls. **C.** Correlation matrix showing R^2 Pearson correlation values of H3S10ph in each pair-wise comparison of *Msk/Rsk* KOs with WT parent Fu.7 and each other. H3S10ph of *Msk/Rsk* KOs shows high correlation across the entire genome (top), but relatively more discordance at transcription start sites (middle) and at ESC-specific enhancers (bottom). **D.** Regions that lost H3S10ph in the *Msk/Rsk* KOs ($z < -0.25$) are differentially distributed relative to TSSs. *Msk2*$^{-/-}$*, /Rsk1*$^{-/-}$*, Rsk3*$^{-/Y}$ lose H3S10ph proximal to promoters (-50 to 50 bp) whereas *Msk1*$^{-/-}$*, Rsk2*$^{-/Y}$*, Rsk4*$^{-/Y}$ lose H3S10ph at regions distal to promoters (50-500 bp). **E.** Genome browser screenshot showing selective regional loss of H3S10ph at distal enhancers of *Egr1* and *Fos*. DHS denotes DNase hypersensitivity sites.

Immediate early genes are the canonical targets of the nucleosomal response, and most are highly expressed in ESCs. Visual inspection on the genome browser revealed that H3S10ph was indeed lost at the promoter and the distal enhancers of *Egr1* and *c-Fos* in cells lacking RSK2 and RSK3 for the former, and MSK2 the latter (**Fig. 5.3E**).

Pairwise comparison of H3S10ph of KO vs. WT ESCs at all annotated TSS (+/- 5kb) confirmed that a subset of genes lost H3S10ph at their promoter upon kinase depletion (**Fig. 5.4**). Specifically, >1,000 promoters show loss of TSS-associated H3S10ph in *Msk2*$^{-/-}$*, Rsk1*$^{-/-}$ and *Rsk3*$^{-/-}$. *Rsk4*$^{-/Y}$ was the only line that had more promoters gaining H3S10ph compared to loss, perhaps reflecting on compensatory changes by the other kinases. To determine the functions of the kinase genic targets, I performed Gene Ontology (GO) analysis on the subset of promoters with significant H3S10ph loss (**Table 5.1**). Such analysis revealed ontologies associated with cell cycle, chromatin modification and metabolism lost promoter-associated H3S10ph in *Msk2*$^{-/-}$, while loss of H3S10ph was observed at developmental-related promoters in *Rsk1/2/3* KO ESCs. Specifically, *Rsk1*$^{-/-}$ showed loss of promoter-specific H3S10ph in genes associated with morphogenesis and tube development. Genes associated with differentiation lost H3S10ph at the promoter in *Rsk2*$^{-/Y}$, whereas genes involved in neuronal cell development showed H3S10ph loss in *Rsk3*$^{-/-}$. Interestingly, while both MSK/RSK are known to respond to epidermal growth factor,

the *Rsk3*$^{-/-}$ was the only mutant that was significantly enriched for EGF-related terms. From the ChIP-seq data, I conclude that the interphase H3S10ph domain previously profiled in **Chapter 3** is likely a composite of the individual effects exerted by multiple serine kinases, each phosphorylating a small subset of promoters at H3S10.

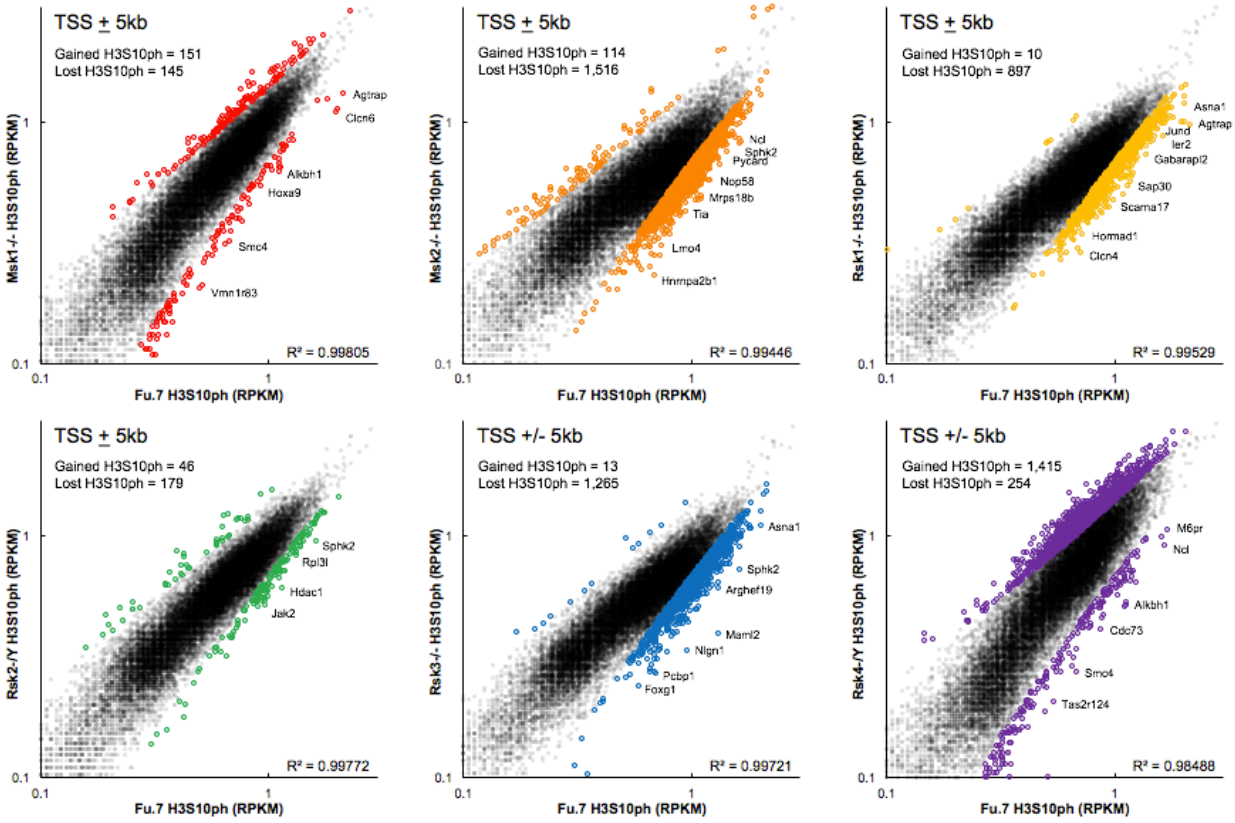


Figure 5.4 Interphase H3S10ph is depleted at a subset of promoters upon depletion of MSK/RSK.
Pair-wise comparison of TSS-associated (+/- 5kb) interphase H3S10ph of *Msk/Rsk* KO to WT Fu.7 parent. Coloured points represent promoters that showed significant change in H3S10ph that exceeds $|z| > 0.25$. R^2 = Pearson correlation.

Table 5.1 Biological processes of genes showing loss of TSS-associated H3S10ph in *Msk2, Rsk1, Rsk2,* and *Rsk3* KO ESCs.

MSK2-/- Gene Ontologies	Fold	p-value*
regulation of chromatin silencing	11.02	7.76E-3
nuclear envelope organization	5.73	4.36E-2
positive regulation of mRNA metabolic process	5.41	2.95E-2
negative regulation of translation	3.9	3.54E-3
regulation of RNA splicing	3.76	4.15E-2
transcription from RNA polymerase II promoter	3.5	1.06E-4

Rsk1-/- Gene Ontologies	Fold	p-value*
negative regulation of translation	5.11	1.53E-3
nucleosome organization	4.93	3.07E-2
rRNA processing	3.83	3.08E-2
chromosome organization	2.85	6.07E-3
mitotic cell cycle	2.55	2.89E-3
embryonic morphogenesis	2.35	2.60E-2
tube development	2.32	1.76E-2

Rsk2-/Y Gene Ontologies	Fold	p-value*
negative regulation of cell differentiation	4.53	4.85E-2

Rsk3-/- Gene Ontologies	Fold	p-value*
cellular response to epidermal growth factor stimulus	8.99	3.41E-02
cell differentiation in spinal cord	7.17	5.07E-04
neuron fate commitment	5.99	1.85E-04
neural crest cell migration	5.64	4.98E-02
ear morphogenesis	4.37	1.25E-03
stem cell population maintenance	4.11	3.10E-03

*Bonferroni-corrected

5.2.5 Msk2 maintains expression of germline-specific genes

To determine whether the change in H3S10ph in interphase is accompanied by a change in transcription, I performed mRNA-seq on the same clones profiled by ChIP-seq (**Fig. 5.5**). In *Msk1, Rsk2, Rsk3* and *Rsk4* KOs, the kinase targeted is amongst the top genes down-regulated – implying efficient nonsense mediated decay and further validating the KO strategy. Changes in gene expression were generally modest in most of the KOs – with only a few dozen genes being differentially expressed using a stringent threshold ($z > |0.8|$). Interestingly, there was significant overlap in genes that were consistently deregulated in most or all KO lines (**Fig. 5.5B**) – including downregulation of *Atp5l*, *Meg3-Rian,* and ribosomal genes such as *Fau, Rpl36, Rps26/27.* Furthermore, while *Egr1* was down-regulated in *Msk2* and *Rsk3* KO lines, and *Fos* was down-regulated in *Msk1/Rsk1/Rsk3* KO lines, other immediate early genes showed in

change in expression in the mutants, possibly due to their low basal expression in WT ESCs (**Fig. 5.5C**).

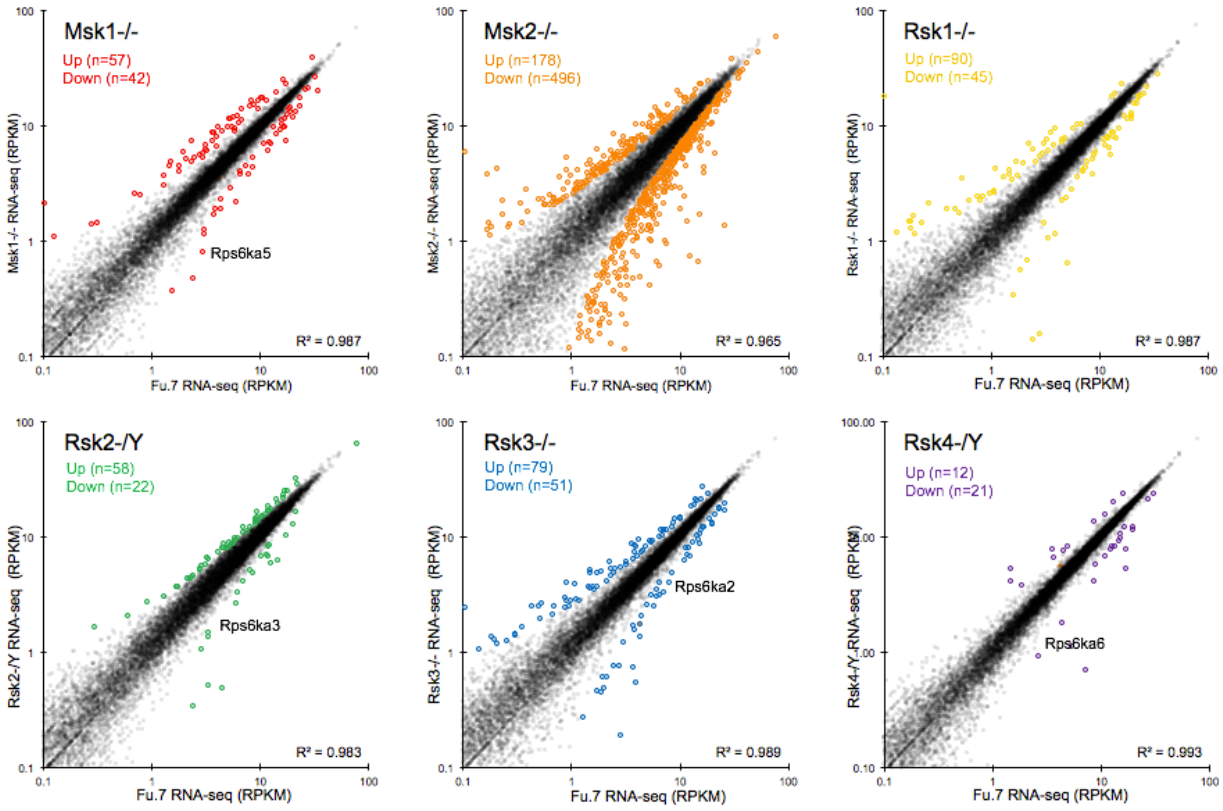


Figure 5.5 Transcriptome analysis of *Msk/Rsk* knockout ESCs.
Pairwise comparison of RNA expression of all annotated genes in kinase KO vs. Fu.7 WT parental line. Coloured points indicate significantly up- or down-regulated ($z > 0.8$). R^2 = Pearson correlation.

Msk2$^{-/-}$ was the notable outlier – showing 10-fold more genes deregulated, even after removal of chr11 genes (as the KO line was no longer trisomic). 496 genes were significantly down-regulated and this subset of genes also showed significant loss of H3S10ph at TSSs, implying that these are direct transcriptional targets of MSK2 (**Fig. 5.6A**). Conversely, the upregulated genes (n=178) showed no significant change in H3S10ph at their TSS and could be secondary effects of *MSK2* deletion. Examination of the top down-regulated genes in *Msk2*$^{-/-}$ revealed striking preference for germline-specific genes (**Fig. 5.6B**). This is further corroborated

using an independent measure of tissue enrichment scoring using Tissue Specific Expression Analysis (TSEA), which showed significant enrichment of testis-specific genes (Dougherty, 2010), including *Fkbp6, Rpl39l, Dazl, Tuba3a* and several members of the *Xlr* family (**Fig. 5.6C**). According to TSEA, of the 497 down-regulated genes, 97 were specific to the testis/ovaries. Conversely, the up-regulated genes in *Msk2*$^{-/-}$ were enriched for various other tissues such as brain and blood vessel – notably, many were specific to the primitive endoderm, including *Gata6*, *Sox7, Gata4* and *Sox17* (Schrode et al. 2014).

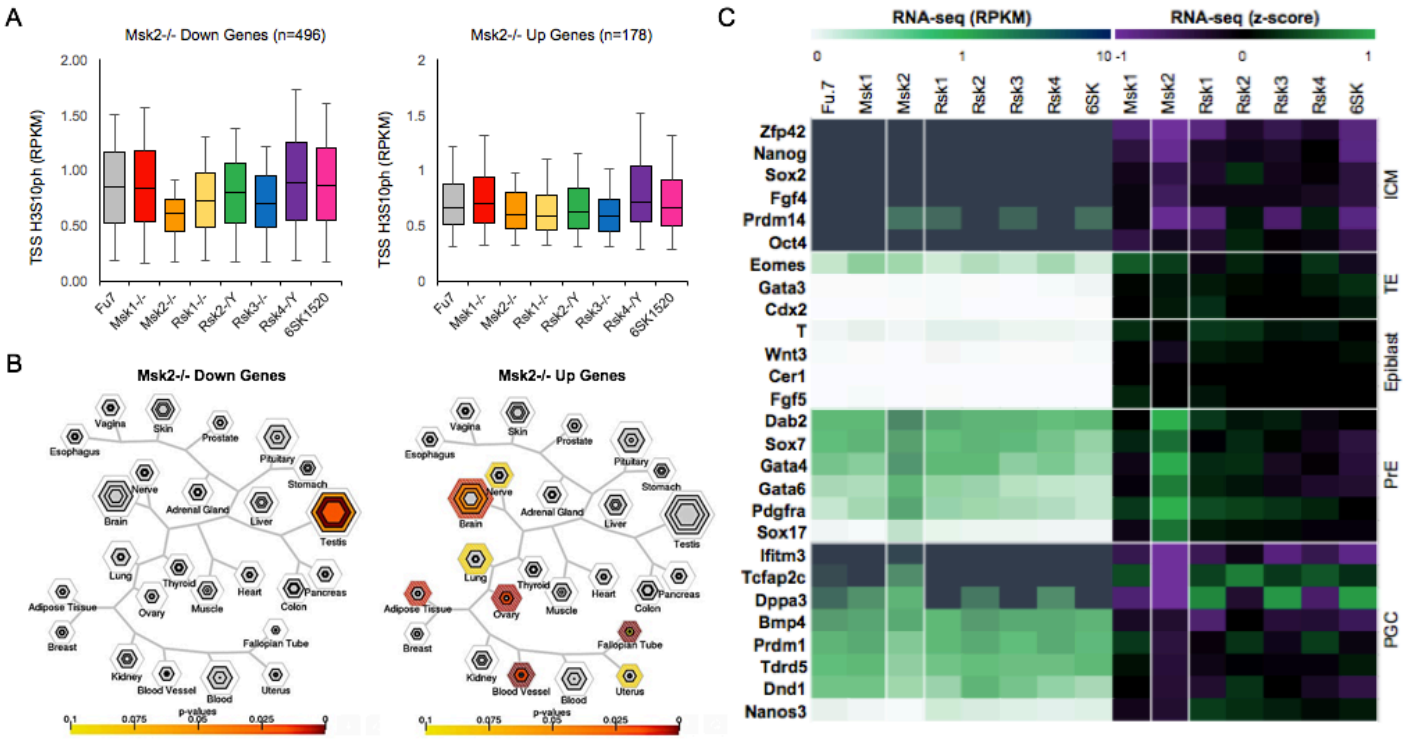


Figure 5.6 MSK2 mediates the expression of testis-specific genes in male ESCs.
A. Genes down-regulated in *Msk2*$^{-/-}$ ESCs concurrently show loss of TSS-associated H3S10ph. Box and whisker plot depicting TSS-associated H3S10ph RPKM in all KO and parental Fu.7 for genes down- or up-regulated in *Msk2*$^{-/-}$ KO ESCs. **B.** Genes down-regulated in *Msk2*$^{-/-}$ are highly enriched for testis-specific expression. TSEA analysis of genes down- or up-regulated in *Msk2*$^{-/-}$ KO ESCs. **C.** MSK2 mediates expression of ICM- and PGC-markers and suppresses primitive endoderm (PrE) markers. Expression analysis for embryonic lineage markers in the WT Fu.7 and KO ESCs. ICM: inner cell mass, TE: trophectoderm, PGC: primordial germ cells.

A small number of imprinted genes – including *Grb10* and the *Meg3/Rian* cluster, were also deregulated in *Msk2*$^{-/-}$ and several other KOs (**Fig. 5.7**). The disruption of imprinted genes is partially attributable to *Zfp57*, which was moderately down (< 1.5-fold) in all the KO lines, and is essential for maintenance of transcriptional repression of imprinted loci (Quenneville et al. 2011). Several of the germline genes or imprinted genes were also downregulated in *Rsk3*$^{-/-}$, and therefore are unlikely to be artefacts of differences in karyotype. Interestingly, the germline genes identified as targets of MSK2 bears some resemblance to our previously identified targets of SETDB1-mediated repression, with overlapping genes including *Tuba3a, Gpa33* and *Rpl39* (Karimi et al. 2011). Indeed, a small subset of MSK2-induced genes was concurrently marked with H3K9me3 and H3K9me3S10ph at the promoter. The convergence of SETDB1 and MSK2 is otherwise minimal, as most of the 184 direct repressed targets of SETDB1 (marked with H3K9me3 in WT and are upregulated upon SETDB1 deletion) are not significantly perturbed in the *Msk2*$^{-/-}$. Of the overlapping genes that are also significantly downregulated in the *Msk2*$^{-/-}$, however, all 17 showed testis-specific expression.

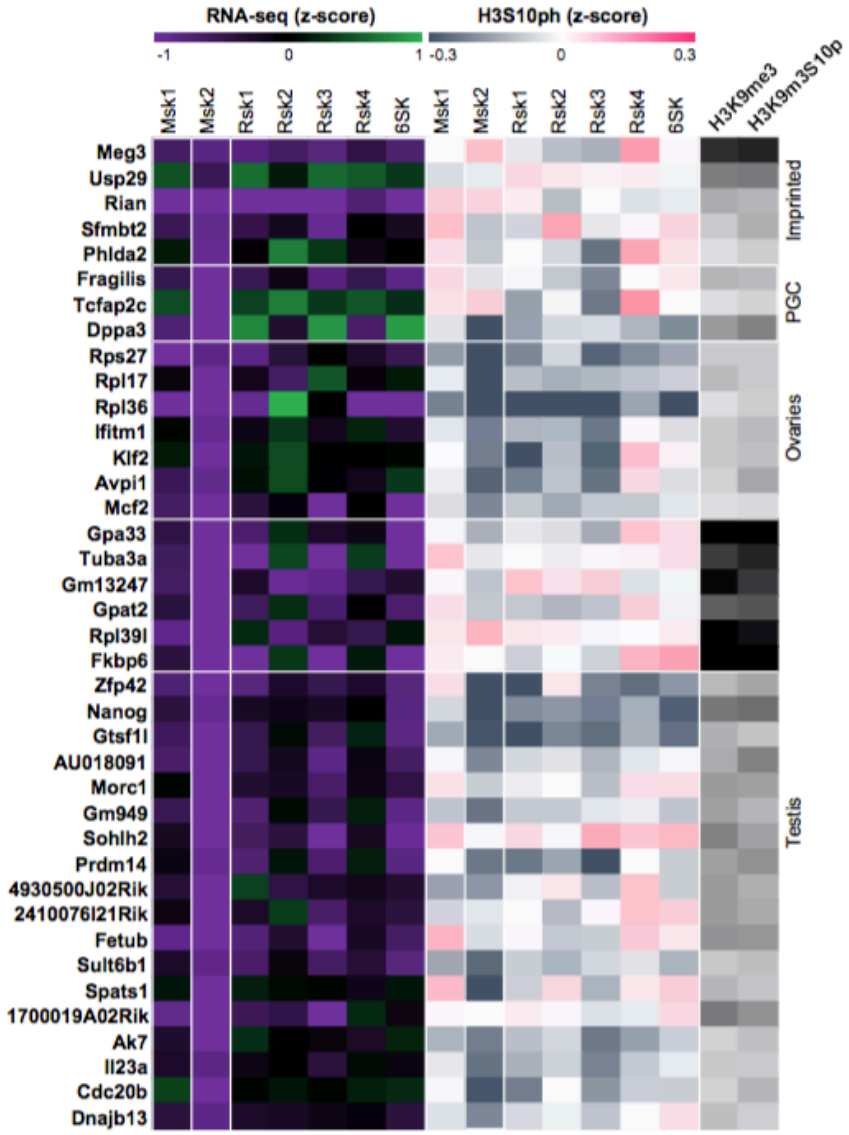


Figure 5.7 Imprinted & germline-specific genes down-regulated in *Msk2*$^{-/-}$ ESCs.
Heatmap of transcriptional and H3S10ph change in the kinase KO to WT Fu.7 for the subset of imprinted genes and germline-specific genes that were significantly down-regulated in *Msk2*$^{-/-}$ ESCs. Note that a significant fraction of *MSK2*-regulated genes are concurrently marked with H3K9me3 and H3K9me3S10ph.

5.2.6 Class I & II ERVs are up-regulated in MSK2$^{-/-}$ *ESCs*

As there was an overlap between the genic targets of Msk2 and SETDB1 I wondered whether SETDB1-repressed ERVs were also aberrantly expressed in the *Msk2*$^{-/-}$ line. Strikingly, class I and II ERVs, such as IAPs and ETns, were significantly upregulated ($z > 2$) in *Msk2*$^{-/-}$ only (**Fig.**

5.8). However, upregulation was modest relative to that observed in Setdb1 conditional deletion in ESCs, ~2-3 fold upregulation of IAPs in *Msk2*$^{-/-}$ compared to ~8-10 fold in *Setdb1* cKO. Taken together, the RNA-seq and ChIP-seq data supports the role of multiple independent H3S10 kinases acting at distinct genomic regions in ESCs in interphase, with MSK2 being necessary for H3S10ph deposition and expression of imprinted and germline-specific genes.

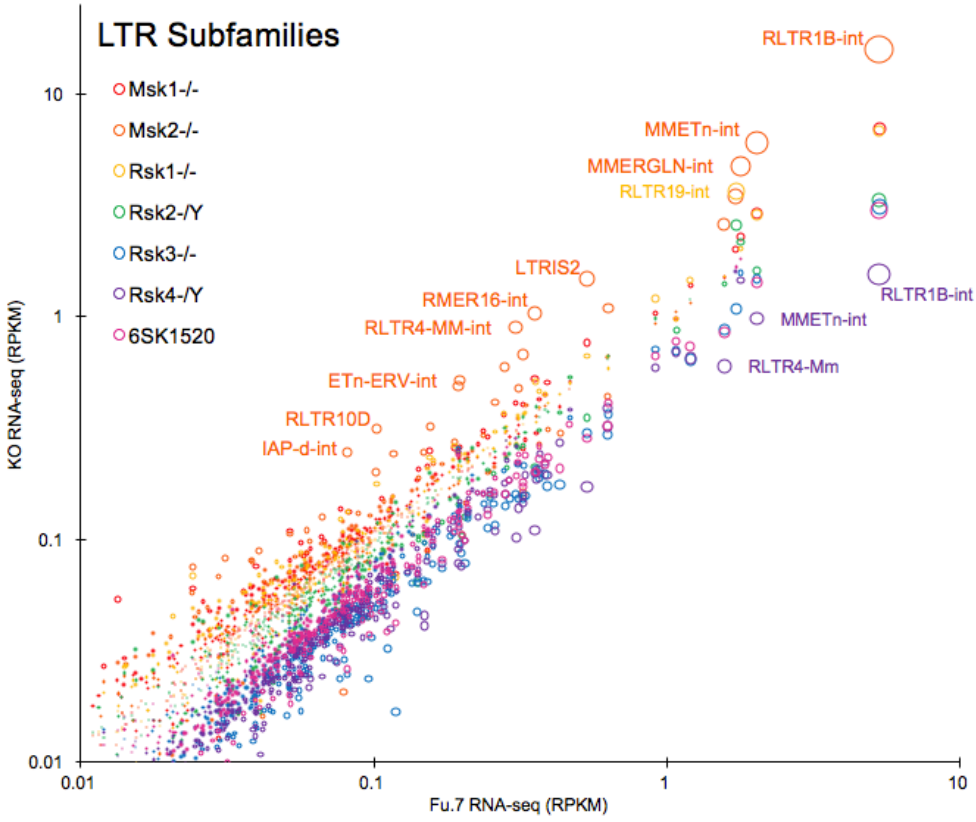


Figure 5.8 Class I & II ERVs are uniquely upregulated in *Msk2*$^{-/-}$ ESCs.
Pairwise comparison of WT Fu.7 vs. kinase KO agglomerated RPKM for individual ERV subfamilies. Points are coloured according to KO line, and are sized to reflect |z-score|. Note that upregulation of ERVs are exclusively observed in *Msk2*$^{-/-}$ ESCs.

5.2.7 Embryoid bodies derived from Msk2 and Rsk KO ESCs show distinct morphologies

To explore the possibility that MSK or RSK may mediate ESC differentiation potential, I analyzed embryoid bodies (EBs) from WT and the KO ESCs from days 0 – 5 following LIF

withdrawal (**Fig. 5.9**). I minimized the heterogeneity intrinsically associated with ESC differentiation by seeding 12 x 10^3 cells in a 96-well plates with round bottoms. Embryoid bodies generated were consistent in size and morphology across the 12 technical replicates. At day 1 after LIF withdrawal, most KO EBs were of a similar size to the WT parent, with the notable exception of *Msk2* and *Rsk1* KO lines, which were small compared to Fu.7 and other KO lines. By day 3 post LIF withdrawal, except for *Msk1*$^{-/-}$, all other KO clones shows considerably smaller EB colonies suggesting defects in cell proliferation. I also noted that while the WT and *Msk1*$^{-/-}$ clones maintained high expression of Geminin-Venus (green), all other KOs were only positive for the G1 marker, mKO, suggesting that the proliferation defect was due to G1 arrest. *Msk2*$^{-/-}$, in particular, developed complex EB morphology by day 4 with a greater tendency for cystic EBs or the formation of hollow tube-projections from the main colony. In comparison, the Fu.7 parent and other kinase KO lines generally formed solid EBs. These gross morphological differences suggest that MSK/RSK may have a functional role in dictating lineage commitment upon differentiation.

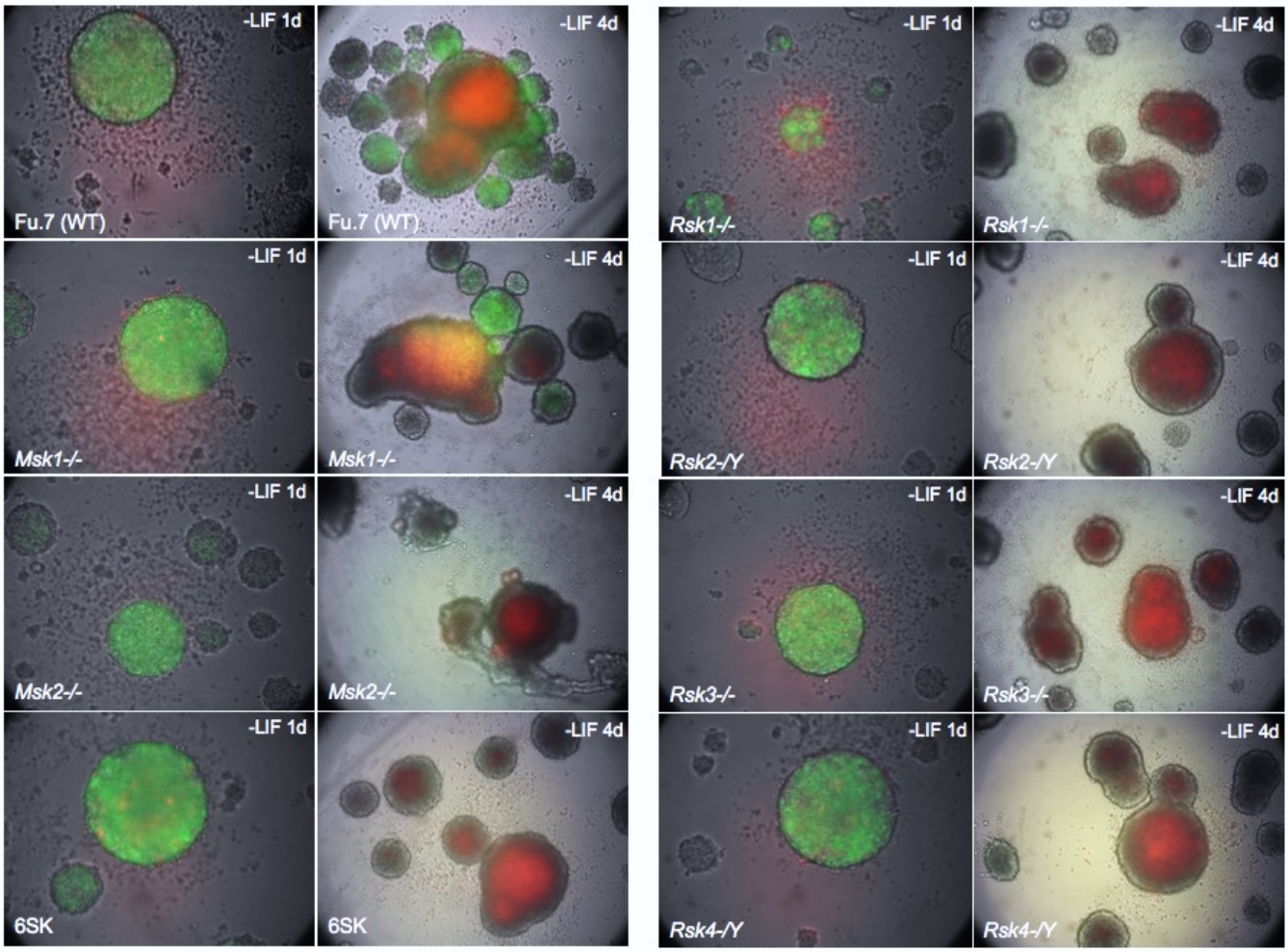


Figure 5.9 Embryoid bodies of *Msk/Rsk* KO FUCCI ESCs.
Live imaging of FUCCI expression in differentiated embryoid bodies at 1 and 4 days post-LIF withdrawal. Green denotes S/G2/M marker Geminin-Venus, Red represents G1 marker Cdt1-mKO2. Note that at day 1 in the differentiation, EBs from KO ESCs are similar in size to WT parent (Fu.7), with the exception of *Rsk1*$^{-/-}$, and are mostly expressing the S/G2/M marker (green). After 4 days of LIF withdrawal, *Msk2* & *Rsk1/2/3/4* KO EBs are smaller than WT, and are mostly expressing the G1 marker (red), indicative of increased G1-arrest during differentiation.

5.3 Discussion

5.3.1 JIL-1 homologs are dispensable for interphase H3S10ph domains in ESCs

In **Chapter 3**, I described the pervasive interphase H3S10ph domain structure in ESCs, reminiscent of JIL-1-mediated H3S10ph at euchromatin polytene bands. MSKs and RSKs represent the closest mammalian orthologs to the chromosomal H3S10 kinase JIL-1, sharing ~40-60% homology at the N-terminal AGC kinase domain (Jin et al. 1999). MSK1/2 and RSK2 have reported kinase activity towards H3S10 *in vitro* (Soloaga et al. 2003; Sassone-Corsi et al. 1999), and are necessary for inducing phosphorylation of H3 upon EGF stimulation *in vivo*. While the JIL-1 homologs MSK1 and MSK2 are reported to be the major H3S10 kinases in mammalian cells, genetic ablation of these kinases resulted in surprisingly modest effects on H3S10ph domain structure in interphase or absolute levels of H3S10ph in ESC. Instead, in murine ESCs, MSKs and RSKs are necessary for deposition of H3S10ph over only a small proportion of the genome, specifically at a subset of promoters and enhancers. This is consistent with reports detailing the persistence of inducible H3S10ph in *Msk1/2* DKO MEFs and RSK2-deficient fibroblasts derived from Coffin-Lowry patients (Duncan et al. 2006; Soloaga et al. 2003). This may be explained by the specific divergence of JIL-1 in *Drosophila* - JIL-1 possesses a 270kDa N-terminal domain not found in any MSK or RSK family members that is necessary for H3S10ph and euchromatin interbanding (Li et al. 2013). Additionally, unlike its mammalian homologs, JIL-1 possesses autophosphorylation activity *in vitro* (Li et al. 2013), suggesting that JIL-1 may have independently acquire the ability to become the sole H3S10 kinase in flies. Nevertheless, this leaves unanswered the question of which kinases maintain H3S10ph in interphase in mammals.

5.3.2 Downregulation of germline genes in Msk2$^{-/-}$ cells and the link with SETDB1/H3K9me3

Ablation of MSK2 leads to loss of H3S10ph at specific promoters and enhancers, but not necessarily to decreased expression of the associated genes– suggesting that this mark is not necessary to maintain expression. Amongst the genes and ERVs that are deregulated in *Msk2*$^{-/-}$ and other knockout lines, many are regulated by SETDB1-dependent H3K9me3 in ESCs (Karimi et al. 2011; Leung et al. 2014). I postulate that MSK2-deposited H3S10ph at germline promoters may counteract the repressive effects of SETDB1/H3K9me3 to maintain low basal expression. Thus, the absence of MSK2-mediated H3S10ph may allow for more robust silencing effects of SETDB1, effectively ablating the expression of these genes (**Fig. 5.10**). Consequently, the enhanced repression at the germline promoters may titrate SETDB1 away from its native targets, leading to modest derepression of ERVs. As SETDB1/H3K9me3 is only targeted to a minority of the germline genes that are down-regulated in *Msk2*$^{-/-}$ ESCs, ectopic H3K9me3 accumulation is unlikely to explain the profound loss of expression of ~100 germline genes. Silencing of a master regulator of gametogenesis, such as *Fragilis* or *Stella*/*Dppa3*, could consequently lead to the indirect silencing of downstream germline genes. It is important to note that the relationship between MSK2 and SETDB1 is speculative and will require follow-up investigation of SETDB1-dependent repression in *Msk2*$^{-/-}$ ESCs.

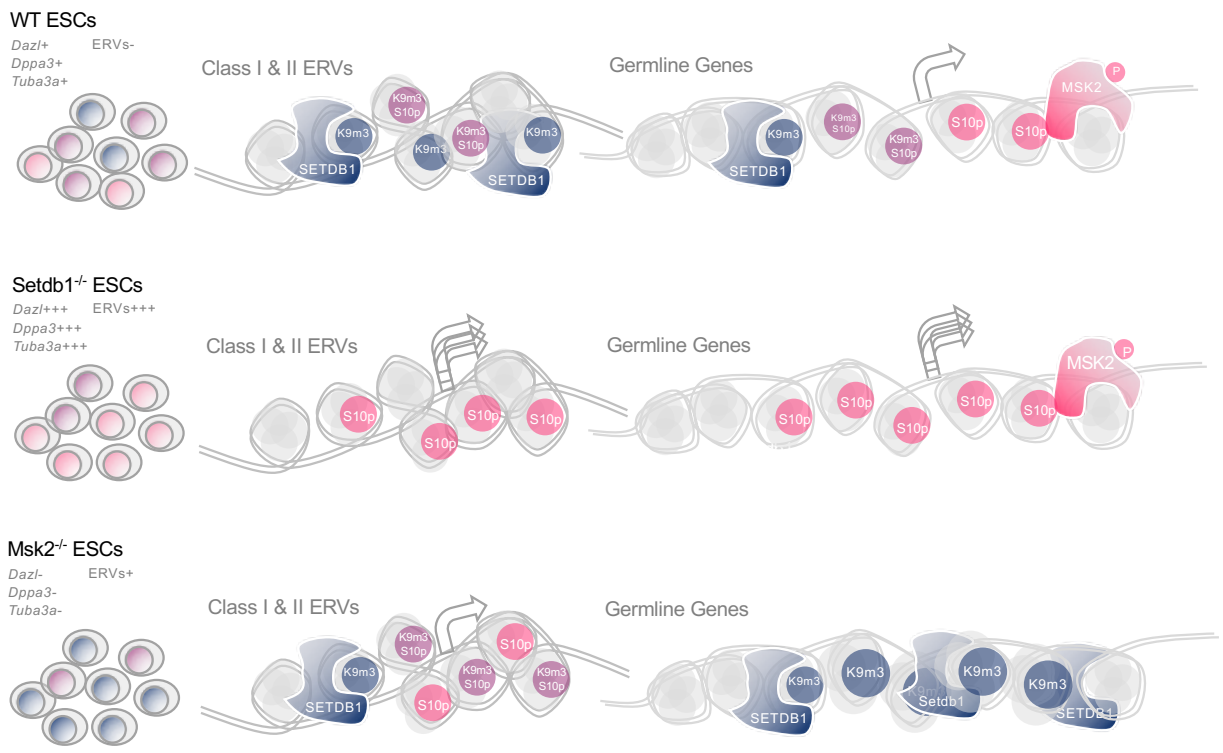


Figure 5.10 Model for crosstalk of MSK2-mediated H3S10ph and SETDB1-mediated H3K9me3 at ERVs and at the germline promoters.
ESCs cultured in serum are maintained in the metastable state where cells express germline lineage markers, including *Dazl,* in a heterogeneous manner. The stochastic expression of germline-specific genes is mediated by MSK2-dependent H3S10ph and SETDB1-dependent H3K9me3, respectively, in WT ESCs. In *Setdb1*$^{-/-}$ ESCs, H3K9me3 is depleted from both ERVs and germline promoters, therefore strongly de-repressing the transcription of both. Conversely, upon MSK2 ablation, H3S10ph is lost from the germline promoters which liberate non-H3S10 phosphorylated nucleosomes for SETDB1 recruitment and robust repression of these genes. The increase in SETDB1 occupancy at genic promoters comes at an expense for recruitment to the ERVs, leading to attenuated proviral silencing.

5.3.3 MSK2 in the maintenance of ESC pluripotency

MSK and RSK are known as "downstream" kinases of ERK signalling. Given the genes dysregulated and the morphology of EBs, it is possible that MSK2 mediates the maintenance of pluripotency and eventual germline potential of ESCs, by inhibiting conversion to the extraembryonic lineage. It is perhaps surprising that the *Msk2* and *Rsk1/2/3/4* mutants exhibit differentiation defects in ESCs, as one might expect these distal effector kinases to phenocopy ERK mutants, which are impaired in development of neural, mesodermal, and endodermal

lineages (Kunath et al. 2007; Hamilton and Brickman 2014). Previously, MSK1/2 have been reported to initiate a negative feedback loop to abort ERK signalling in response to inflammatory cytokines (Anaieva, 2008), and may exert similar antagonistic effects in ESCs. Interestingly, male-to-female sex reversal phenotypes have been observed for the p38 MAPK (Gierl, 2012), which acts upstream of MSKs. Thus, MSK2 may predominantly signal through p38, rather than ERK in ESCs to promote basal expression of testis-specific genes.

Differences in the transcriptome observed in *Msk2*$^{-/-}$ ESCs require confirmation with an independently derived clone, as the clone characterized underwent chr11 trisomy reversion. Alternatively, one could use the same clone and perform rescue with a MSK2 WT and kinase-dead expression constructs. It is also possible that the observations presented here reflect artefactual divergence of individual ESC clones, given that culturing can impart epigenetic changes on imprinted genes and affect germline transmission (Carstea, 2009). The molecular phenotypes observed in this study are likely specific to cultured ESCs *in vitro,* as *Msk1/2* DKO mice have been previously derived and are reported to be viable and fertile under pathogen-free conditions (Wiggins, 2002; Ananieva, 2008). Unfortunately, single *Msk2* KO mice have not been described. The single KO phenotype may require re-evaluation, however, as a study reported severe deficiency in blastocyst formation in morpholino-mediated MSK2 knockdown in one-cell mouse zygotes, showing micronuclei mitotic arrest at 4- to 8-cell stage (Chavez, 2014), and may point to a role for maternally deposited MSK2 in the early zygote. Nevertheless, the observation of germline marker loss in *Msk2*$^{-/-}$ ESC *in vitro* may uncover novel crosstalk with SETDB1/H3K9me3, and have significant implications for the utility of ESCs in *ex vivo* models of germ cell development (Hayashi et al. 2011; 2012) and chimeric mice generation.

6. CONCLUDING REMARKS & FUTURE DIRECTIONS

6.1 Interphase H3S10ph in mouse embryonic stem cells

Despite being widely documented amongst diverse eukaryotes, the study of histone H3 serine 10 phosphorylation has largely been limited to small-scale studies, and the relevance of this mark in transcriptional regulation in mammalian cells is not well understood. In order to understand the functional relevance of this mark, I combined a novel method for purification of cells in distinct phases of the cell cycle with genome-wide analyses of H3S10ph and transcription, revealing novel roles for this mark in interphase mammalian cells.

In **Chapter 3**, I showed that an Aurora B-independent kinase deposits H3S10ph at gene-dense regions, covering up to 30% of the genome in G1 ESCs. While some cell cycle fluctuations in H3S10ph enrichment were observed, predominantly at enhancers, H3S10ph levels were largely preserved through interphase at gene-dense regions. Such H3S10ph domains show remarkable correlation with domains of early DNA replication timing (RT) and anti-correlation with H3K9me2 and LADs (Yokochi et al. 2009; Peric-Hupkes et al. 2010). Disruption of H3S10ph by expression of non-phosphorylatable H3.3S10A resulted in ectopic accumulation of H3K9me2 in adjacent H3S10ph-enriched euchromatic regions, mimicking the phenotype observed in *Drosophila* JIL-1 kinase mutants (Wang et al. 2001). Conversely, interphase H3S10ph domains expanded into flanking unmarked regions on the order of ~100kb in each direction in $Glp^{-/-}$ ESCs, revealing that H3S10ph is restricted by GLP-mediated H3K9me2 in ESCs. Strikingly, spreading of H3S10ph at RT transition regions (TTRs) is accompanied by

aberrant strand-biased transcription initiation of genes and repetitive elements co-oriented with the replication. Indicating that H3K9me2 plays a critical role in establishing repressive chromatin on the leading strand arm of the replication fork.

I showed that H3S10ph is also present in interphase MEFs, but is restricted to intragenic regions of actively transcribing genes. The depletion of interphase H3S10ph at early-replicating regions in MEFs is concurrent with accumulation of H3K9me2, indicating that antagonism between H3S10ph and H3K9me2 also manifests in MEFs. These results are consistent with the model that H3S10ph functions to specify higher order euchromatin structure, and displace lamin-tethered H3K9me2/3 readers and writers to promote chromatin detachment from the nuclear scaffold for transcription and replication.

In **Chapter 4**, I found that H3S10ph coexists with H3K9me3 in interphase ESCs. Unlike H3S10ph alone, such phosphorylation of H3K9me3S10 is dependent upon Aurora B kinase activity, suggesting that while most of the Aurora-dependent H3S10ph is dephosphorylated by PP1 by telophase, nucleosomes containing H3K9me3S10ph are apparently excluded from such phosphatase activity, leading to the persistence of this mark into G1. Interestingly, while the phosphorylation of H3S10 on H3K9me3 marked tails was previously shown to displace HP1, a mechanism known as the "phospho-methyl switch" (Fischle et al. 2005), I found that class I and II ERVs are strongly enriched for H3K9me3, H3K9me3S10ph and HP1 in G1-sorted cells. These paradoxical observations are most likely explained by distinct nucleosomes bearing H3K9me3 or H3K9me3S10ph in G1 nuclei, with only the former bound by HP1. Regardless, *in vitro* experiments aimed at addressing whether PP1 is inhibited by specific histone PTMs, such

as H3K9me3, are clearly warranted. Alternatively, PP1 may be occluded by H3K9me3S10ph reader proteins, such as ATRX or NP95 (Kunowska et al. 2015), which could be addressed by studying ESCs deficient in such readers.

Unlike HP1 proteins, both ATRX and NP95 are functionally relevant for SETDB1-dependent silencing of ERVs in ESCs (Sadic et al. 2015; Elsässer et al. 2015; Sharif et al. 2016). Given that NP95 plays a critical role in maintenance DNA methylation, and ATRX plays an important role in deposition of the histone variant H3.3, H3K9me3S10ph likely promotes both DNA methylation maintenance and nucleosome remodeling activities at ERVs (Noh et al. 2014; Rothbart et al. 2012). Whether H3K9me3S10ph-binding promotes retention of ATRX or UHRF1/NP95 at these elements during mitosis would be interesting to pursue in future studies.

In **Chapter 5**, I reasoned that interphase H3S10ph at gene-dense regions may be deposited by a JIL-1 homologous kinase in ESCs, as interphase H3S10ph domains bear a striking resemblance to JIL-1 occupancy in *Drosophila* (Regnard et al. 2011). MSK and RSK are the closest mammalian orthologs to JIL-1, sharing ~40-60% similarity in the effector AGC kinase domain relevant for H3 phosphorylation, and several MSK/RSK family members have been shown to mediate MAPK-induced H3S10ph in mammalian cells (Soloaga et al. 2003; Sassone-Corsi et al. 1999). To ascertain the functional role of *Msk1/2* and *Rsk1/2/3/4*, I derived individual and combinatorial KOs in FUCCI-ESC lines using CRISPR/Cas9 and characterized interphase H3S10ph and transcription in the KOs and corresponding WT parent line. H3S10ph domains in G1-sorted cells were not notably different from WT in any of the individual or

combinatorial KO lines, indicating that euchromatic H3S10ph is not maintained by MSKs or RSKs in interphase ESCs.

Nevertheless, deletion of *Msk2, Rsk1, Rsk2* or *Rsk3* caused depletion of H3S10ph at a small number of promoters and distal enhancers, including those associated with immediate early genes and translation-related functions. Consistent with the modest change in H3S10ph, expression analysis further revealed that only a few dozen genes are deregulated in each of the KOs, except for *Msk2*$^{-/-}$ cells. *Msk2*$^{-/-}$ ESCs showed notable deregulation of imprinted loci, including *Meg3-Rian*, and striking decrease in expression of testis-specific genes. Interestingly, MSK2 depletion also resulted in up-regulation of primitive endoderm markers, including *Gata6,* suggesting that MSK2 inhibits the differentiation of ESCs towards extraembryonic lineages *in vitro*. Finally, while deletions of *Msk* & *Rsk* genes did not notably impact cell cycle progression in ESCs, following LIF-withdrawal, most embryoid bodies arising in the KO lines showed increased G1-arrest and were significantly smaller than those in the WT control– suggesting that these kinases may be required for efficient differentiation of ESCs into the appropriate germ layer lineages.

6.2 A strategy for identifying the "interphase H3S10ph kinase"

The modest effect on interphase H3S10ph observed in the *Msk* & *Rsk* KO ESC lines, detailed **in Chapter 5,** indicate that broad H3S10ph associated with gene-dense regions are not dependent on any known JIL-1 homologs in ESCs – suggesting a novel serine kinase is phosphorylating 30% of the genome in ESCs. A systematic kinase screen would be the most efficient strategy to isolate the relevant kinase(s) for interphase H3S10ph. While kinase inhibitor panels have

previously been generated and extensively characterized (Karaman et al. 2008), a robust and high-throughput readout of interphase H3S10ph levels remains to be developed.

Intriguingly, the Kimura group showed that the Fab region of the monoclonal H3S10ph antibody (CMA311) used in our study reports cell cycle dynamics of H3S10ph when injected into living pre-implantation embryos, demonstrating the feasibility of a phospho sensor *in vivo* (Hayashi-Takanaka et al. 2009). Subsequently, they demonstrated that variable regions of IgG heavy and light chains can be cloned from the antibody-secreting hybridoma, and genetically introduced as fluorescent modification-specific intracellular antibodies, or "mintbodies", to create living histone PTM reporter transgenic cells and organisms (Sato et al. 2013). The mintbodies can theoretically be co-expressed alongside FUCCI, assuming compatibilities between the different fluorophores, to distinguish interphase from mitotic H3S10ph. Given the efficiency and low-cost of CRISPR/Cas-9 mediated deletions, the H3S10ph sensor could be coupled with systematic KO of non-essential serine kinases in ESCs. Identification of the interphase H3S10ph kinase using a reporter-driven screening strategy would yield important insights into the functional relevance of this mark in regulating global chromatin structure.

6.3 Physiological roles of interphase H3S10ph in mammals

6.3.1 Role of H3K9me3 & H3S10ph in germline development

Previous work in the Lorincz lab revealed that ~40 germline-specific genes are dependent on both SETDB1-dependent H3K9me3 and DNA methylation for transcriptional repression in ESCs (Karimi et al. 2011; Leung et al. 2014). In **Chapter 5**, I found that many of these germline genes lacked the canonical signatures of KRAB-Zfp-dependent silencing, namely KAP1 and HP1

binding at the H3K9me3 marked promoter regions. Furthermore, transcriptional repression of these genes is not dependent on either KAP1 or HP1s in ESCs (Rowe et al. 2010; Maksakova et al. 2013). As SETDB1 does not have intrinsic DNA binding activity, this repressive KMT is likely recruited to these specific genomic regions via a yet undescribed targeting pathway, perhaps via known transcriptional repressors, such as E2F6. In **Chapter 5**, I showed that the basal transcription of many of the same germline-specific genes is regulated by MSK2, possibly through deposition of H3S10ph at the promoter to counteract SETDB1-dependent repression. Taken together, these results suggest the phospho-methyl antagonism may converge at these specific genomic loci encoding essential genes for gametogenesis.

Further characterization of the pathways governing the appropriate expression of these germline genes is clearly warranted. One caveat related to such studies in ESCs is that many germline genes are stochastically expressed in a heterogeneous manner, and epigenetic deregulation at these genes can be an artefact of tissue culture conditions. For example, repression of *Dazl* in ESCs is due in part to aberrant DNA methylation accumulation *in vitro*, a consequence of low TET activity in the absence of vitamin C (Blaschke et al. 2013). Given that *in vivo* studies have conclusively demonstrated the necessity of both DNA methylation and SETDB1-dependent H3K9me3 in germ cell development (Liu et al. 2014; Hargan-Calvopina et al. 2016), it is possible that MSK2 also plays an important role in the developing germline. Interestingly, while SETDB1 depletion leads to severe loss of E13.5 PGCs (Liu et al. 2014), the molecular basis of this failure of germline development remains obscure. One possibility is that loss of SETDB1 results in subversion of germ cells to the extraembryonic trophectoderm lineage

(Dodge et al. 2004; Yuan et al. 2009); similarly, whether MSK2 plays a role in germline development by suppressing differentiation to the primitive endoderm is best addressed *in vivo*.

6.3.2 Interphase H3S10ph in the zygotic pronuclei

Notably, H3S10ph is detectable by immunofluorescence during interphase of the first two mitotic cycles (**Fig. 6.1**) (Ribeiro-Mason et al. 2012; Hayashi-Takanaka et al. 2009; Huang et al. 2007; Teperek-Tkacz et al. 2010). Upon fertilization, the totipotent one cell embryo undergoes dramatic nuclear remodeling beginning with protamine to histone H3.3 exchange, global DNA demethylation on the paternal pronucleus, and finally zygotic genome activation at the 2-cell stage (Hemberger et al. 2009). As the maternal and paternal genomes remain in separate pronuclei (PN), asymmetric histone PTMs can be readily discernable up until the 4-cell stage. Such asymmetry of H3K9me2/3 is visible in the first two zygotic divisions. While maternal pronuclei are marked with pre-existing H3K9me2 and H3K9me3 on pericentromeric and centromeric satellites, respectively, paternal pronuclei lack H3K9me altogether (Santos et al. 2005). Consistent with its role as a H3K9me3 reader, localization of HP1β in the zygote matches that of the H3K9me3 asymmetry, and is only found on the maternal pericentromeric heterochromatin (Martin et al. 2006). HP1α is absent from both pronuclei at the zygotic stage, and is only loaded onto pericentromeric heterochromatin symmetrically at late S in a replication-independent manner (Meglicki et al. 2014). Conversely, ATRX is asymmetrically transmitted to the zygote via binding to the maternal pericentromeric heterochromatin (La Fuente et al. 2015), possibly through recognition of H3K9me3S10ph. This maternally-deposited ATRX is required for silencing of major satellite transcription on the maternal pronucleus and to prevent illegitimate centromere recombination in the early zygote.

In contrast to H3K9me2/3, H3S10ph is observed on both the maternal and paternal pronuclei, but is more pronounced for the paternal perinucleolar heterochromatin. This interphase H3S10ph first accumulates in the nuclear space in G1 (PN1/2), is diluted during S (PN3), and eventually re-accumulates on the pericentromeric rings in late S (PN4/5) (Ribeiro-Mason et al. 2012; Lan et al. 2017). This pattern of enrichment is detected exclusively during the first two cell cycles with H3S10ph restricted to mitosis in subsequent cell divisions. It is possible the zygotic interphase H3S10ph inhibits binding of HP1, and thereby prevents premature H3K9me2/3 deposition on the paternal pronucleus to mediate the extensive chromatin reprogramming necessary for zygotic genome activation.

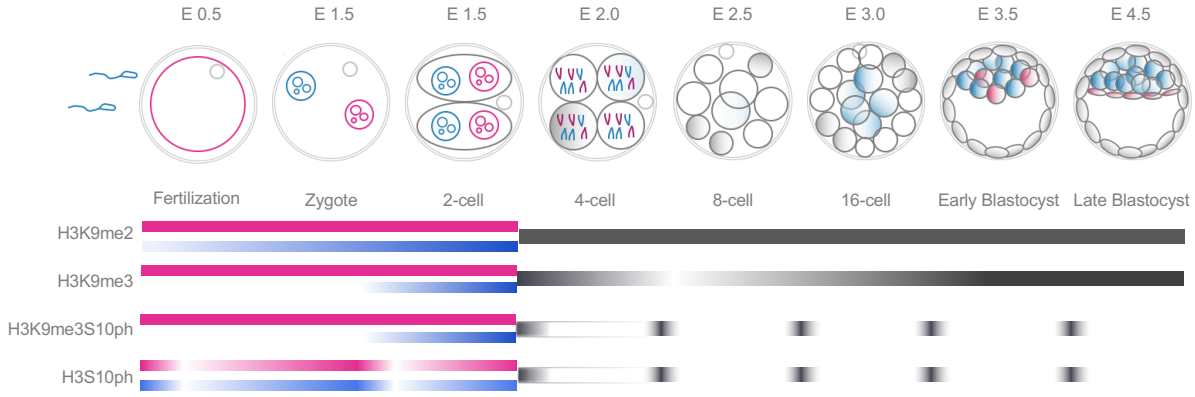


Figure 6.1 H3K9me & H3S10ph dynamics in early mammalian embryo.
Kinetics of H3K9me2, H3K9me3 and H3S10ph in the cleavage stages of mammalian embryonic development, depicted as a heatmap. At fertilization, the maternal and paternal genomes bear distinct epigenetic modifications, including H3K9me2/3. The maternal pronucleus is marked with H3K9me2/3, whereas this mark is absent on the paternal pronucleus and does not appear until the late zygote for H3K9me2, and 2-cell for H3K9me3. During the first two zygotic cell cycles, interphase H3S10ph has been reported on both parental pronuclei, but the significance of this signal is not yet determined. Beginning from the 4-cell stage, H3S10ph then becomes restricted to mitosis.

6.3.3 Inducible H3S10ph in neurological and immunological responses

The nucleosomal response, which describes the transient phosphorylation of chromatin upon mitogenic or stress stimulation, seems to be a common response to environmental cues. However, inducible H3S10ph has largely been studied under serum-starved conditions *in vitro*, and the biological relevance of this phenomenon is not well understood.

Crosio et al. were the first to demonstrate that night-time exposure of mice to light induces a strong increase of H3S10ph-positive neurons in the hypothalamus (Crosio et al. 2000). Notably, this stimulatory H3S10ph was highly specific to a small population of hypothalamic neurons termed the hypothalamic suprachiasmatic nucleus, which is essential for circadian rhythm – and was the first conclusive evidence that phosphorylation of chromatin is relevant for neuron-related functions. Subsequent work has also shown that MSK1-dependent H3S10ph accumulates in the brain of cocaine-injected mice (Stipanovich et al. 2008; Brami-Cherrier 2005). This dopamine response mediated by MSK1 was limited to the reward and motor centre within the ventral striatum, and not the surrounding cortex, and is functionally important for locomotor sensitization. These studies point to H3S10ph as an essential signalling component of neuronal stimulation. Indeed, H3S10ph is a reliable biomarker for stimulated regions of the brain. An obvious avenue to pursue in future research is to identify the stimulus-specific genomic targets of induced H3S10ph genome-wide, to reveal key downstream targets necessary for cognitive or other neurological changes and perhaps to examine the crosstalk between this mark and H3K9me in these regions.

In addition to neural-associated functions, H3S10ph also orchestrates transcription of pro-inflammatory cytokines in dendritic cells as a part of p38-dependent innate immune response against bacterial pathogens (Saccani et al. 2001; Josefowicz et al. 2016). Interestingly, some pathogens, including *Listeria monocytogenes* and *Shigella flexneri,* interfere with the induction of inflammatory response through dephosphorylation of basal H3S10ph (Arbibe et al. 2006; Hamon et al. 2007). In contrast to the mitotic arrest caused by *Helicobacter pylori,* which causes global loss of H3S10ph (Fehri et al. 2009)*,* pathogenic dephosphorylation of H3S10 occurs specifically at key genomic loci encoding cytokines, such as *Cxcl2*, and is potentially important in preventing recruitment of leukocytes *in vivo* (Arbibe et al. 2006). G9a-mediated H3K9me2 was previously shown to suppress production of interferons in dendritic cells (Fang et al. 2012a), raising the possibility that crosstalk between H3S10ph and H3K9me2 is relevant for dynamic regulation of innate immunity. Furthermore, deletion of *G9a* confers a protective role against RNA viruses, whereas *Msk1/2 DKO* mice exhibit increased susceptibility to endotoxin shock (Fang et al. 2012a; Ananieva et al. 2008). These opposing phenotypes are reminiscent of aberrant chromatin changes left unchecked in the absence of H3K9me or H3S10ph, as suggested by my work in ESCs. Investigating the role of histone modifiers in appropriating the magnitude of response towards pathogens may yield novel therapeutic targets for infectious diseases and inflammatory disorders.

6.4 Conclusions

My thesis work presents the first genome-wide characterization of H3S10ph and reveals that this mark is more pervasive in interphase than previously believed, with potential implications in the regulation of both transcription and DNA replication. My study is consistent with the model that H3S10ph promotes euchromatin structure by antagonizing methylation of the adjacent H3K9, and in turn prevent establishment of repressive chromatin structure. Furthermore, I specifically define replication timing transition regions and germline genes as genomic hotspots for the phospho-methyl interplay in ESCs. Future work directed at identification of the kinase(s) responsible for deposition of H3S10ph in interphase, and characterization of this mark *in vivo* should yield insights into the role of interphase H3S10ph in transcriptional regulation and establishment of euchromatic domains.

BIBLIOGRAPHY

Aagaard L, Schmid M, Warburton P, Jenuwein T. 2000. Mitotic phosphorylation of SUV39H1, a novel component of active centromeres, coincides with transient accumulation at mammalian centromeres. *Journal of Cell Science* **113 (Pt 5)**: 817–829.

Abe Y, Sako K, Takagaki K, Hirayama Y, Uchida KSK, Herman JA, DeLuca JG, Hirota T. 2016. HP1-Assisted Aurora B Kinase Activity Prevents Chromosome Segregation Errors. *Developmental Cell* **36**: 487–497.

Adams RR, Maiato H, Earnshaw WC, Carmena M. 2001. Essential roles of Drosophila inner centromere protein (INCENP) and aurora B in histone H3 phosphorylation, metaphase chromosome alignment, kinetochore disjunction, and chromosome segregation. *The Journal of Cell Biology* **153**: 865–880.

Ahmad K, Henikoff S. 2002. The Histone Variant H3.3 Marks Active Chromatin by Replication-Independent Nucleosome Assembly. *Mol Cell* **9**: 1191–1200.

Ainsztein AM, Kandels-Lewis SE, Mackay AM, Earnshaw WC. 1998. INCENP centromere and spindle targeting: identification of essential conserved motifs and involvement of heterochromatin protein HP1. *The Journal of Cell Biology* **143**: 1763–1774.

Akhtar W, de Jong J, Pindyurin AV, Pagie L, Meuleman W, de Ridder J, Berns A, Wessels LFA, van Lohuizen M, van Steensel B. 2013. Chromatin Position Effects Assayed by Thousands of Reporters Integrated in Parallel. *Cell* **154**: 914–927.

Allan RS, Zueva E, Cammas F, Schreiber HA, Masson V, Belz GT, Roche D, Maison C, Quivy J-P, Almouzni G, et al. 2012. An epigenetic silencing pathway controlling T helper 2 cell lineage commitment. *Nature* **487**: 249–253.

Allshire RC, Ekwall K. 2015. Epigenetic Regulation of Chromatin States in Schizosaccharomyces pombe. *Cold Spring Harbor Perspectives in Biology* **7**: a018770–26.

Ananieva O, Darragh J, Johansen C, Carr JM, McIlrath J, Park JM, Wingate A, Monk CE, Toth R, Santos SG, et al. 2008. The kinases MSK1 and MSK2 act as negative regulators of Toll-like receptor signaling. *Nat Immunol* **9**: 1028–1036.

Andrews FH, Gatchalian J, Krajewski K, Strahl BD, Kutateladze TG. 2016. Regulation of Methyllysine Readers through Phosphorylation. *ACS Chem Biol* **11**: 547–553.

Anest V, Hanson JL, Cogswell PC, Steinbrecher KA, Strahl BD, Baldwin AS. 2003. A nucleosomal function for IkappaB kinase-alpha in NF-kappaB-dependent gene expression. *Nature* **423**: 659–663.

Angell RR, Jacobs PA. 1975. Lateral asymmetry in human constitutive heterochromatin. *Chromosoma* **51**: 301–310.

Anjum R, Blenis J. 2008. The RSK family of kinases: emerging roles in cellular signalling. *Nat Rev Mol Cell Biol* **9**: 747–758.

Arbibe L, Kim DW, Batsche E, Pedron T, Mateescu B, Muchardt C, Parsot C, Sansonetti PJ. 2006. An injected bacterial effector targets chromatin access for transcription factor NF-κB to alter transcription of host genes involved in immune responses. *Nat Immunol* **8**: 47–56.

Arthur JS, Cohen P. 2000. MSK1 is required for CREB phosphorylation in response to mitogens in mouse embryonic stem cells. *FEBS Letters* **482**: 44–48.

Auclair G, Borgel J, Sanz LA, Vallet J, Guibert S, Dumas M, Cavelier P, Girardot M, Forné T, Feil R, et al. 2016. EHMT2 directs DNA methylation for efficient gene silencing in mouse embryos. *Genome Research* **26**: 192–202.

Aucott R, Bullwinkel J, Yu Y, Shi W, Billur M, Brown JP, Menzel U, Kioussis D, Wang G, Reisert I, et al. 2008. HP1-β is required for development of the cerebral neocortex and neuromuscular junctions. *The Journal of Cell Biology* **183**: 597–606.

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research* **37**: W202–W208.

Bannister AJ, Kouzarides T. 2011. Regulation of chromatin by histone modifications. *Nature Publishing Group* **21**: 381–395.

Barratt MJ, Hazzalin CA, Cano E, Mahadevan LC. 1994. Mitogen-stimulated phosphorylation of histone H3 is targeted to a small hyperacetylation-sensitive fraction. *Proc Natl Acad Sci USA* **91**: 4781–4785.

Bénit L, Lallemand JB, Casella JF, Philippe H, Heidmann T. 1999. ERV-L elements: a family of endogenous retrovirus-like elements active throughout the evolution of mammals. *J Virol* **73**: 3301–3308.

Bickmore WA, van Steensel B. 2013. Genome Architecture: Domain Organization of Interphase Chromosomes. *Cell* **152**: 1270–1284.

Bischoff JR, Anderson L, Zhu Y, Mossie K, Ng L, Souza B, Schryver B, Flanagan P, Clairvoyant F, Ginther C, et al. 1998. A homologue of Drosophila aurora kinase is oncogenic and amplified in human colorectal cancers. *EMBO J* **17**: 3052–3065.

Blaschke K, Ebata KT, Karimi MM, Zepeda-Martínez JA, Goyal P, Mahapatra S, Tam A, Laird DJ, Hirst M, Rao A, et al. 2013. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature*.

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**: 256–263.

Bode AM. 2005. Inducible Covalent Posttranslational Modification of Histone H3. *Science's STKE* **2005**: re4–re4.

Bowman GD, Poirier MG. 2015. Post-Translational Modifications of Histones That Influence Nucleosome Dynamics. *Chem Rev* **115**: 2274–2295.

Bradbury EM, Inglis RJ, Matthews HR, Sarner N. 1973. Phosphorylation of Very-Lysine-Rich Histone in Physarum polycephalum. *The FEBS Journal* **33**: 131–139.

Brami-Cherrier K. 2005. Parsing Molecular and Behavioral Effects of Cocaine in Mitogen- and Stress-Activated Protein Kinase-1-Deficient Mice. *Journal of Neuroscience* **25**: 11444–11454.

Breuer K, Foroushani AK, Laird MR, Chen C, Sribnaia A, Lo R, Winsor GL, Hancock REW, Brinkman FSL, Lynn DJ. 2012. InnateDB: systems biology of innate immunity and beyond--recent updates and continuing curation. *Nucleic Acids Research*.

Brind'Amour J, Liu S, Hudson M, Chen C, Karimi MM, Lorincz MC. 2015. An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations. *Nature Communications* **6**: 6033.

Cai W, Wang C, Li Y, Yao C, Shen L, Liu S, Bao X, Schnable PS, Girton J, Johansen J, et al. 2014. Genome-wide analysis of regulation of gene expression and H3K9me2 distribution by JIL-1 kinase mediated histone H3S10 phosphorylation in Drosophila. *Nucleic Acids Research*.

Cammas F, Janoshazi A, Lerouge T, Losson R. 2007. Dynamic and selective interactions of the transcriptional corepressor TIF1β with the heterochromatin protein HP1 isotypes during cell differentiation. *Differentiation* **75**: 627–637.

Carmena M, Wheelock M, Funabiki H, Earnshaw WC. 2012. The chromosomal passenger complex (CPC): from easy rider to the godfather of mitosis. *Nat Rev Mol Cell Biol* **13**: 789–803.

Castel AL, Cleary JD, Pearson CE. 2010. Repeat instability as the basis for human diseases and as a potential target for therapy. 1–6.

Chen ES, Zhang K, Nicolas E, Cam HP, Zofall M, Grewal SIS. 2008. Cell cycle control of centromeric repeat transcription and heterochromatin assembly. *Nature* **451**: 734–737.

Chen P, Zhao J, Wang Y, Wang M, Long H, Liang D, Huang L, Wen Z, Li W, Li X, et al. 2013. H3.3 actively marks enhancers and primes gene transcription via opening higher-ordered chromatin. *Genes Dev* **27**: 2109–2124.

Chen RH, Sarnecki C, Blenis J. 1992. Nuclear localization and regulation of erk-and rsk-encoded protein kinases. *Molecular and Cellular Biology* **12**: 915–927.

Cheung P, Tanner KG, Cheung WL, Sassone-Corsi P, Denu JM, Allis CD. 2000. Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation. *Mol Cell* **5**: 905–915.

Choi Y-S, Karelina K, Alzate-Correa D, Hoyt KR, Impey S, Arthur JS, Obrietan K. 2012. Mitogen- and stress-activated kinases regulate progenitor cell proliferation and neuron development in the adult dentate gyrus. *J Neurochem* **123**: 676–688.

Chwang WB, Arthur JS, Schumacher A, Sweatt JD. 2007. The Nuclear Kinase Mitogen- and Stress-Activated Protein Kinase 1 Regulates Hippocampal Chromatin Remodeling in Memory Formation. *Journal of Neuroscience* **27**: 12732–12742.

Coffin GS, Siris E, Wegienka LC. 1966. Mental Retardation With Osteocartilaginous Anomalies. *Am J Dis Child* **112**: 205–213.

Collins RE, Northrop JP, Horton JR, Lee DY, Zhang X, Stallcup MR, Cheng X. 2008. The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules. *Nature Structural & Molecular Biology* **15**: 245–250.

Core LJ, Waterfall JJ, Lis JT. 2008. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**: 1845–1848.

Crosio C, Cermakian N, Allis CD, Sassone-Corsi P. 2000. Light induces chromatin modification in cells of the mammalian circadian clock. *Nat Neurosci* **3**: 1241–1247.

Crosio C, Fimia GM, Loury R, Kimura M, Okano Y, Zhou H, Sen S, Allis CD, Sassone-Corsi P. 2002. Mitotic phosphorylation of histone H3: spatio-temporal regulation by mammalian Aurora kinases. *Molecular and Cellular Biology* **22**: 874–885.

Deng H, Bao X, Cai W, Blacketer MJ, Belmont AS, Girton J, Johansen J, Johansen KM. 2008. Ectopic histone H3S10 phosphorylation causes chromatin structure remodeling in Drosophila. *Development* **135**: 699–705.

Dodge JE, Kang YK, Beppu H, Lei H, Li E. 2004. Histone H3-K9 Methyltransferase ESET Is Essential for Early Development. *Molecular and Cellular Biology* **24**: 2478–2486.

Domcke S, Bardet AF, Ginno PA, Hartl D, Burger L, Schübeler D. 2015. Competition between DNA methylation and transcription factors determines binding of NRF1. *Nature* **528**: 575–579.

Dougherty JD, Schmidt EF, Nakajima M, Heintz N. 2010. Analytical approaches to RNA profiling data for the identification of genes enriched in specific cells. *Nucleic Acids Research* **38**: 4218–4230.

Duan Q, Chen H, Costa M, Dai W. 2008. Phosphorylation of H3S10 blocks the access of H3K9 by specific antibodies and histone methyltransferase. Implication in regulating chromatin dynamics and epigenetic inheritance during mitosis. *J Biol Chem* **283**: 33585–33590.

Dummler BA. 2005. Functional Characterization of Human RSK4, a New 90-kDa Ribosomal S6 Kinase, Reveals Constitutive Activation in Most Cell Types. *J Biol Chem* **280**: 13304–13314.

Duncan EA, Anest V, Cogswell P, Baldwin AS. 2006. The kinases MSK1 and MSK2 are required for epidermal growth factor-induced, but not tumor necrosis factor-induced, histone H3 Ser10 phosphorylation. *J Biol Chem* **281**: 12521–12525.

Ecco G, Cassano M, Kauzlaric A, Duc J, Coluccio A, Offner S, Imbeault M, Rowe HM, Turelli P, Trono D. 2016. Transposable Elements and Their KRAB-ZFP Controllers Regulate Gene Expression in Adult Tissues. *Developmental Cell* **36**: 611–623.

Elgin SCR, Reuter G. 2013. Position-Effect Variegation, Heterochromatin Formation, and Gene Silencing in Drosophila. *Cold Spring Harbor Perspectives in Biology* **5**: a017780–a017780.

Elsässer SJ, Noh K-M, Diaz N, Allis CD, Banaszynski LA. 2015. Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. *Nature* **522**: 240–244.

Eustermann S, Yang J-C, Law MJ, Amos R, Chapman LM, Jelinska C, Garrick D, Clynes D, Gibbons RJ, Rhodes D, et al. 2011. Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. *Nature Structural & Molecular Biology* **18**: 777–782.

Falconer E, Hills M, Naumann U, Poon SSS, Chavez EA, Sanders AD, Zhao Y, Hirst M, Lansdorp PM. 2012. DNA template strand sequencing of single-cells maps genomic rearrangements at high resolution. *Nat Meth* **9**: 1107–1112.

Fang TC, Schaefer U, Mecklenbrauker I, Stienen A, Dewell S, Chen MS, Rioja I, Parravicini V, Prinjha RK, Chandwani R, et al. 2012a. Histone H3 lysine 9 di-methylation as an epigenetic signature of the interferon response. *Journal of Experimental Medicine*.

Fang TC, Schaefer U, Mecklenbrauker I, Stienen A, Dewell S, Chen MS, Rioja I, Parravicini V, Prinjha RK, Chandwani R, et al. 2012b. Histone H3 lysine 9 di-methylation as an epigenetic signature of the interferon response. *J Exp Med* **209**: 661–669.

Fazzio TG, Panning B. 2010. Condensin complexes regulate mitotic progression and interphase chromatin structure in embryonic stem cells. *The Journal of Cell Biology* **188**: 491–503.

Fehri LF, Rechner C, Janssen S, Mak TN, Holland C, Bartfeld S, Brüggemann H, Meyer TF. 2009. Helicobacter pylori-induced modification of the histone H3 phosphorylation status in gastric epithelial cells reflects its impact on cell cycle regulation. *epigenetics* **4**: 577–586.

Feng YQ, Lorincz MC, Fiering S, Greally JM, Bouhassira EE. 2001. Position Effects Are Influenced by the Orientation of a Transgene with Respect to Flanking Chromatin. *Molecular and Cellular Biology* **21**: 298–309.

Filippova GN, Thienes CP, Penn BH, Cho DH, Hu YJ, Moore JM, Klesert TR, Lobanenkov VV,

Tapscott SJ. 2001. CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nat Genet* **28**: 335–343.

Fischle W, Tseng BS, Dormann HL, Ueberheide BM, Garcia BA, Shabanowitz J, Hunt DF, Funabiki H, Allis CD. 2005. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* **438**: 1116–1122.

Flicek P, Aken BL, Ballester B, Beal K, Bragin E, Brent S, Chen Y, Clapham P, Coates G, Fairley S, et al. 2009. Ensembl's 10th year. *Nucleic Acids Research* **38**: D557–D562.

Forneris F, Binda C, Vanoni MA, Battaglioli E, Mattevi A. 2005. Human Histone Demethylase LSD1 Reads the Histone Code. *Journal of Biological Chemistry* **280**: 41360–41365.

Franz H, Mosch K, Soeroes S, Urlaub H. 2010. Multimerization and H3K9me3 binding are required for CDYL1b heterochromatin association. *Journal of Biological Chemistry* **285**: 11754–11754.

Frietze S, O'Geen H, Blahnik KR, Jin VX, Farnham PJ. 2010. ZNF274 Recruits the Histone Methyltransferase SETDB1 to the 3′ Ends of ZNF Genes ed. P.-A. Defossez. *PLoS ONE* **5**: e15082–16.

Fritsch L, Robin P, Mathieu JRR, Souidi M, Hinaux H, Rougeulle C, Harel-Bellan A, Ameyar-Zazoua M, Ait-Si-Ali S. 2010. A Subset of the Histone H3 Lysine 9 Methyltransferases Suv39h1, G9a, GLP, and SETDB1 Participate in a Multimeric Complex. *Mol Cell* **37**: 46–56.

Fry CJ, Shogren-Knaak MA. 2004. Histone H3 amino-terminal tail phosphorylation and acetylation: synergistic or independent transcriptional regulatory marks? *Cold Spring Harbor Symposia on Quantitative Biology* **69**: 1–8.

Fuller BG, Lampson MA, Foley EA, Rosasco-Nitcher S, Le KV, Tobelmann P, Brautigan DL, Stukenberg PT, Kapoor TM. 2008. Midzone activation of aurora B in anaphase produces an intracellular phosphorylation gradient. *Nature* **453**: 1132–1136.

Garrick D, Fiering S, Martin DIK, Whitelaw E. 1998. Repeat-induced gene silencing in mammals. *Nat Genet* **18**: 56–59.

Gilbert N, Boyle S, Fiegler H, Woodfine K, Carter NP, Bickmore WA. 2004. Chromatin Architecture of the Human Genome. *Cell* **118**: 555–566.

Goertz MJ, Wu Z, Gallardo TD, Hamra FK, Castrillon DH. 2011. Foxo1 is required in mouse spermatogonial stem cells for their maintenance and the initiation of spermatogenesis. *J Clin Invest* **121**: 3456–3466.

Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W, et al. 2008. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**: 948–951.

Guo XW, Th'ng JP, Swank RA, Anderson HJ, Tudan C, Bradbury EM, Roberge M. 1995. Chromosome condensation induced by fostriecin does not require p34cdc2 kinase activity and histone H1 hyperphosphorylation, but is associated with enhanced histone H2A and H3 phosphorylation. *EMBO J* **14**: 976–985.

Gurley LR, Walters RA, Tobey RA. 1974. Cell cycle-specific changes in histone phosphorylation associated with cell proliferation and chromosome condensation. *The Journal of Cell Biology* **60**: 356–364.

Hamilton WB, Brickman JM. 2014. Erk Signaling Suppresses Embryonic Stem Cell Self-Renewal to Specify Endoderm. *Cell Rep* **9**: 2056–2070.

Hamon MA, Batsche E, Régnault B, Tham TN, Seveau S, Muchardt C, Cossart P. 2007. Histone modifications induced by a family of bacterial toxins. *Proc Natl Acad Sci USA* **104**: 13467–13472.

Hargan-Calvopina J, Taylor S, Cook H, Hu Z, Lee SA, Yen M-R, Chiang Y-S, Chen P-Y, Clark AT. 2016. Stage-Specific Demethylation in Primordial Germ Cells Safeguards against Precocious Differentiation. *Developmental Cell* **39**: 75–86.

Hathaway NA, Bell O, Hodges C, Miller EL, Neel DS, Crabtree GR. 2012. Dynamics and memory of heterochromatin in living cells. *Cell* **149**: 1447–1460.

Hayakawa T, Haraguchi T, Masumoto H, Hiraoka Y. 2003. Cell cycle behavior of human HP1 subtypes: distinct molecular domains of HP1 are required for their centromeric localization during interphase and metaphase. *Journal of Cell Science* **116**: 3327–3338.

Hayashi K, Ogushi S, Kurimoto K, Shimamoto S, Ohta H, Saitou M. 2012. Offspring from oocytes derived from in vitro primordial germ cell-like cells in mice. *Science* **338**: 971–975.

Hayashi K, Ohta H, Kurimoto K, Aramaki S, Saitou M. 2011. Reconstitution of the Mouse Germ Cell Specification Pathway in Culture by Pluripotent Stem Cells. *Cell* **146**: 519–532.

Hayashi-Takanaka Y, Yamagata K, Nozaki N, Kimura H. 2009. Visualizing histone modifications in living cells: spatiotemporal dynamics of H3 phosphorylation during interphase. *The Journal of Cell Biology* **187**: 781–790.

Healy S, Khan P, He S, Davie JR. 2012. Histone H3 phosphorylation, immediate-early gene expression, and the nucleosomal response: a historical perspective. *Biochem Cell Biol* **90**: 39–54.

Hemberger M, Dean W, Reik W. 2009. Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. *Nat Rev Mol Cell Biol* **10**: 526–537.

Henikoff S. 1998. Conspiracy of silence among repeated transgenes. *Bioessays* **20**: 532–535.

Hiragami-Hamada K, Soeroes S, Nikolov M, Wilkins B, Kreuz S, Chen C, La Rosa-Velázquez

De IA, Zenn HM, Kost N, Pohl W, et al. 2016. Dynamic and flexible H3K9me3 bridging via HP1β dimerization establishes a plastic state of condensed chromatin. *Nature Communications* **7**: 11310.

Hiratani I, Ryba T, Itoh M, Rathjen J, Kulik M, Papp B, Fussner E, Bazett-Jones DP, Plath K, Dalton S, et al. 2010. Genome-wide dynamics of replication timing revealed by in vitro models of mouse embryogenesis. *Genome Research* **20**: 155–169.

Hsu JY, Sun ZW, Li X, Reuben M, Tatchell K, Bishop DK, Grushcow JM, Brame CJ, Caldwell JA, Hunt DF, et al. 2000. Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. *Cell* **102**: 279–291.

Hsu PD, Lander ES, Zhang F. 2014. Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* **157**: 1262–1278.

Hu Y, Fang X, Dunham SM, Prada C, Stachowiak EK, Stachowiak MK. 2004. 90-kDa Ribosomal S6 Kinase Is a Direct Target for the Nuclear Fibroblast Growth Factor Receptor 1 (FGFR1): ROLE IN FGFR1 SIGNALING. *Journal of Biological Chemistry* **279**: 29325–29335.

Huang J-C, Lei Z-L, Shi L-H, Miao Y-L, Yang J-W, Ouyang Y-C, Sun Q-Y, Chen D-Y. 2007. Comparison of histone modifications in in vivo and in vitro fertilization mouse embryos. *Biochemical and Biophysical Research Communications* **354**: 77–83.

Huvet M, Nicolay S, Touchon M, Audit B, d'Aubenton-Carafa Y, Arneodo A, Thermes C. 2007. Human gene organization driven by the coordination of replication and transcription. *Genome Research* **17**: 1278–1285.

Jacobs SA. 2002. Structure of HP1 Chromodomain Bound to a Lysine 9-Methylated Histone H3 Tail. *Science* **295**: 2080–2083.

Jacques P-É, Jeyakani J, Bourque G. 2013. The majority of primate-specific regulatory sequences are derived from transposable elements. *PLoS Genet* **9**: e1003504.

Jin Y, Wang Y, Walker DL, Dong H, Conley C, Johansen J, Johansen KM. 1999. JIL-1: a novel chromosomal tandem kinase implicated in transcriptional regulation in Drosophila. *Mol Cell* **4**: 129–135.

Johansen KM, Johansen J. 2006. Regulation of chromatin structure by histone H3S10 phosphorylation. *Chromosome Res* **14**: 393–404.

Josefowicz SZ, Shimada M, Armache A, Li CH, Miller RM, Lin S, Yang A, Dill BD, Molina H, Park H-S, et al. 2016. Chromatin Kinases Act on Transcription Factors and Histone Tails in Regulation of Inducible Transcription. *Mol Cell* **64**: 347–361.

Jung HR, Pasini D, Helin K, Jensen ON. 2010. Quantitative mass spectrometry of histones H3.2 and H3.3 in Suz12-deficient mouse embryonic stem cells reveals distinct, dynamic post-

translational modifications at Lys-27 and Lys-36. *Mol Cell Proteomics* **9**: 838–850.

Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J. 2005. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res* **110**: 462–467.

Karaman MW, Herrgard S, Treiber DK, Gallant P, Atteridge CE, Campbell BT, Chan KW, Ciceri P, Davis MI, Edeen PT, et al. 2008. A quantitative analysis of kinase inhibitor selectivity. *Nature Biotechnology* **26**: 127–132.

Karimi MM, Goyal P, Maksakova IA, Bilenky M, Leung D, Tang JX, Shinkai Y, Mager DL, Jones S, Hirst M, et al. 2011. DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell* **8**: 676–687.

Karmodiya K, Krebs AR, Oulad-Abdelghani M, Kimura H, Tora L. 2012. H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics* **13**: 424.

Kassiotis G. 2014. Endogenous Retroviruses and the Development of Cancer. *The Journal of Immunology* **192**: 1343–1349.

Kent WJ, Zweig AS, Barber G, Hinrichs AS, Karolchik D. 2010. BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* **26**: 2204–2207.

Khan DH, Healy S, He S, Lichtensztejn D. 2017. Mitogen-induced distinct epi-alleles are phosphorylated at either H3S10 or H3S28 depending on H3K27 acetylation. *Molecular Biology of ...*.

Kind J, Pagie L, de Vries SS, Nahidiazar L, Dey SS, Bienko M, Zhan Y, Lajoie B, de Graaf CA, Amendola M, et al. 2015. Genome-wide Maps of Nuclear Lamina Interactions in Single Human Cells. *Cell* 1–15.

Kind J, Pagie L, Ortabozkoyun H, Boyle S, de Vries SS, Janssen H, Amendola M, Nolen LD, Bickmore WA, van Steensel B. 2013. Single-Cell Dynamics of Genome-Nuclear Lamina Interactions. *Cell* **153**: 178–192.

Kind J, van Steensel B. 2010. Genome–nuclear lamina interactions and gene regulation. *Current Opinion in Cell Biology* **22**: 320–325.

Kubicek S, O'Sullivan RJ, August EM, Hickey ER, Zhang Q, Teodoro ML, Rea S, Mechtler K, Kowalski JA, Homon CA, et al. 2007. Reversal of H3K9me2 by a Small-Molecule Inhibitor for the G9a Histone Methyltransferase. *Mol Cell* **25**: 473–481.

Kunath T, Saba-El-Leil MK, Almousailleakh M, Wray J, Meloche S, Smith A. 2007. FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development* **134**: 2895–2902.

Kunowska N, Rotival M, Yu L, Choudhary J, Dillon N. 2015. Identification of protein complexes that bind to histone H3 combinatorial modifications using super-SILAC and weighted correlation network analysis. *Nucleic Acids Research* **43**: 1418–1432.

La Fuente De R, Baumann C, Viveiros MM. 2015. ATRX contributes to epigenetic asymmetry and silencing of major satellite transcripts in the maternal genome of the mouse embryo. *Development* **142**: 1806–1817.

Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T. 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**: 116–120.

Lan J, Lepikhov K, Giehr P, Walter J. 2017. Histone and DNA methylation control by H3 serine 10/threonine 11 phosphorylation in the mouse zygote. *Epigenetics & Chromatin* 1–19.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC. 2001. Initial sequencing and analysis of the human genome. *Nature* **412**: 565–566.

Lau LF, Nathans D. 1987. Expression of a set of growth-related immediate early genes in BALB/c 3T3 cells: coordinate regulation with c-fos or c-myc. *Proc Natl Acad Sci USA* **84**: 1182–1186.

Laugel-Haushalter V, Paschaki M, Marangoni P, Pilgram C, Langer A, Kuntz T, Demassue J, Morkmued S, Choquet P, Constantinesco A, et al. 2014. RSK2 Is a Modulator of Craniofacial Development ed. Y. Gibert. *PLoS ONE* **9**: e84343–15.

Law MJ, Lower KM, Voon HPJ, Hughes JR, Garrick D, Viprakasit V, Mitson M, De Gobbi M, Marra M, Morris A, et al. 2010. ATR-X Syndrome Protein Targets Tandem Repeats and Influences Allele-Specific Expression in a Size-Dependent Manner. *Cell* **143**: 367–378.

Lehnertz B, Ueda Y, Derijck AAHA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AHFM. 2003. Suv39h-Mediated Histone H3 Lysine 9 Methylation Directs DNA Methylation to Major Satellite Repeats at Pericentric Heterochromatin. *Current Biology* **13**: 1192–1200.

Leung D, Du T, Wagner U, Xie W, Lee AY, Goyal P, Li Y, Szulwach KE, Jin P, Lorincz MC, et al. 2014. Regulation of DNA methylation turnover at LTR retrotransposons and imprinted loci by the histone methyltransferase Setdb1. *Proceedings of the National Academy of Sciences*.

Leung DC, Dong KB, Maksakova IA, Goyal P, Appanah R, Lee S, Tachibana M, Shinkai Y, Lehnertz B, Mager DL, et al. 2011. Lysine methyltransferase G9a is required for de novo DNA methylation and the establishment, but not the maintenance, of proviral silencing. *Proceedings of the National Academy of Sciences* **108**: 5718–5723.

Leung DC, Lorincz MC. 2011. Silencing of endogenous retroviruses: when and why do histone marks predominate? *Trends in Biochemical Sciences* 1–7.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.

Li Y, Cai W, Wang C, Yao C, Bao X, Deng H, Girton J, Johansen J, Johansen KM. 2013. Domain Requirements of the JIL-1 Tandem Kinase for Histone H3 Serine10 Phosphorylation and Chromatin Remodeling in Vivo. *J Biol Chem*.

Lienert F, Mohn F, Tiwari VK, Baubec T, Roloff TC, Gaidatzis D, Stadler MB, Schübeler D. 2011. Genomic Prevalence of Heterochromatic H3K9me2 and Transcription Do Not Discriminate Pluripotent from Terminally Differentiated Cells ed. W. Reik. *PLoS Genet* **7**: e1002090.

Liu CWY, Wang R-H, Dohadwala M, Schönthal AH, Villa-Moruzzi E, Berndt N. 1999. Inhibitory Phosphorylation of PP1α Catalytic Subunit during the G1/S Transition. *Journal of Biological Chemistry* **274**: 29470–29475.

Liu S, Brind'Amour J, Karimi MM, Shirane K, Bogutz A, Lefebvre L, Sasaki H, Shinkai Y, Lorincz MC. 2014. Setdb1 is required for germline development and silencing of H3K9me3-marked endogenous retroviruses in primordial germ cells. *Genes Dev* **28**: 2041–2055.

Lowry B, Miller JR, Fraser FC. 1971. A new dominant gene mental retardation syndrome: association with small stature, tapering fingers, characteristic facies, and possible hydrocephalus. *American Journal of Diseases of …*.

Luger K, Mäder AW, Richmond RK, Sargent DF. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**: 251–260.

Macfarlan TS, Gifford WD, Driscoll S, Lettieri K, Rowe HM, Bonanomi D, Firth A, Singer O, Trono D, Pfaff SL. 2012. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **184**: 1286–19.

Mager DL, Stoye JP. 2015. Mammalian Endogenous Retroviruses. *Microbiology Spectrum* **3**: 1–20.

Mahadevan LC, Willis AC, Barratt MJ. 1991. Rapid histone H3 phosphorylation in response to growth factors, phorbol esters, okadaic acid, and protein synthesis inhibitors. *Cell* **65**: 775–783.

Maksakova IA, Goyal P, Bullwinkel J, Brown JP, Bilenky M, Mager DL, Singh PB, Lorincz MC. 2011. H3K9me3-binding proteins are dispensable for SETDB1/H3K9me3-dependent retroviral silencing. *Epigenetics & Chromatin* **4**: 12.

Maksakova IA, Romanish MT, Gagnier L, Dunn CA, van de Lagemaat LN, Mager DL. 2006. Retroviral Elements and Their Hosts: Insertional Mutagenesis in the Mouse Germ Line. *PLoS Genet* **2**: e2.

Maksakova IA, Thompson PJ, Goyal P, Jones SJ, Singh PB, Karimi MM, Lorincz MC. 2013. Distinct roles of KAP1, HP1 and G9a/GLP insilencing of the two-cell-specific retrotransposonMERVL in mouse ES cells. *Epigenetics & Chromatin* **6**: 1–1.

Martin C, Beaujean N, Brochard V, Audouard C, Zink D, Debey P. 2006. Genome restructuring in mouse embryos during reprogramming and early development. *Dev Biol* **292**: 317–332.

Martin E, Betuing S, Pages C, Cambon K, Auregan G, Deglon N, Roze E, Caboche J. 2011. Mitogen- and stress-activated protein kinase 1-induced neuroprotection in Huntington's disease: role on chromatin remodeling at the PGC-1-alpha promoter. *Human Molecular Genetics* **20**: 2422–2434.

Mateescu B, England P, Halgand F, Yaniv M, Muchardt C. 2004. Tethering of HP1 proteins to chromatin is relieved by phosphoacetylation of histone H3. *EMBO Rep* **5**: 490–496.

Matsui T, Leung D, Miyashita H, Maksakova IA, Miyachi H, Kimura H, Tachibana M, Lorincz MC, Shinkai Y. 2010. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature* **464**: 927–931.

McCarthy EM, McDonald JF. 2004. Long terminal repeat retrotransposons of. *Genome Biol* **5**: R14–8.

McClintock B. 1963. Further studies of gene-control systems in maize. *Carnegie Inst Wash Year Book*.

McDowell TL, Gibbons RJ, Sutherland H, ORourke DM, Bickmore WA, Pombo A, Turley H, Gatter K, Picketts DJ, Buckle VJ, et al. 1999a. Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes. *Proc Natl Acad Sci USA* **96**: 13983–13988.

McDowell TL, Gibbons RJ, Sutherland H, ORourke DM, Bickmore WA, Pombo A, Turley H, Gatter K, Picketts DJ, Buckle VJ, et al. 1999b. Localization of a putative transcriptional regulator(ATRX) at pericentromeric heterochromatin andthe short arms of acrocentric chromosomes. *PNAS* 1–6.

Meglicki M, Teperek-Tkacz M, Borsuk E. 2014. Appearance and heterochromatin localization of HP1α in early mouse embryos depends on cytoplasmic clock and H3S10 phosphorylation. *Cell Cycle* **11**: 2189–2205.

Meshorer E, Yellajoshula D, George E, Scambler PJ, Brown DT, Misteli T. 2006. Hyperdynamic Plasticity of Chromatin Proteins in Pluripotent Embryonic Stem Cells. *Developmental Cell* **10**: 105–116.

Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, Thomas PD. 2017. PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Research* **45**: D183–D189.

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim T-K, Koche RP, et al. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**: 553–560.

Montes de Oca R, Andreassen PR, Wilson KL. 2014. Barrier-to-Autointegration Factor influences specific histone modifications. *Nucleus* **2**: 580–590.

Morgan HD, Sutherland HG, Martin DI, Whitelaw E. 1999. Epigenetic inheritance at the agouti locus in the mouse. *Nat Genet* **23**: 314–318.

Morrissy AS, Garzia L, Shih DJH, Zuyderduyn S, Huang X, Skowron P, Remke M, Cavalli FMG, Ramaswamy V, Lindsay PE, et al. 2016. Divergent clonal selection dominates medulloblastoma at recurrence. *Nature* **529**: 351–357.

Mouse Genome Sequencing Consortium, Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, et al. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**: 520–562.

Muller HJ. 1930. Types of visible variations induced by X-rays inDrosophila. *J Genet* **22**: 299–334.

Musselman CA, Lalonde M-E, Côté J, Kutateladze TG. 2012. Perceiving the epigenetic landscape through histone readers. *Nature Structural & Molecular Biology* **19**: 1218–1227.

Musselman CA, Mansfield RE, Garske AL, Davrazou F, Kwan AH, Oliver SS, O'Leary H, Denu JM, Mackay JP, Kutateladze TG. 2009. Binding of the CHD4 PHD2 finger to histone H3 is modulated by covalent modifications. *Biochem J* **423**: 179–187.

Naumova N, Imakaev M, Fudenberg G, Zhan Y, Lajoie BR, Mirny LA, Dekker J. 2013. Organization of the mitotic chromosome. *Science* **342**: 948–953.

Ng J-H, Kumar V, Muratani M, Kraus P, Yeo J-C, Yaw L-P, Xue K, Lufkin T, Prabhakar S, Ng H-H. 2013. In Vivo Epigenomic Profiling of Germ Cells Reveals Germ Cell Molecular Signatures. *Developmental Cell*.

Ng SS, Kavanagh KL, McDonough MA, Butler D, Pilka ES, Lienard BMR, Bray JE, Savitsky P, Gileadi O, Delft von F, et al. 2007. Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. *Nature* **448**: 87–91.

Noh K-M, Maze I, Zhao D, Xiang B, Wenderski W, Lewis PW, Shen L, Li H, Allis CD. 2014. ATRX tolerates activity-dependent histone H3 methyl/phos switching to maintain repetitive element silencing in neurons. *Proc Natl Acad Sci USA* 201411258.

Nozawa R-S, Nagao K, Masuda H-T, Iwasaki O, Hirota T, Nozaki N, Kimura H, Obuse C. 2010. Human POGZ modulates dissociation of HP1α from mitotic chromosome arms through Aurora B activation. *Nat Cell Biol* **12**: 719–727.

O'Sullivan RJ, Karlseder J. 2010. Telomeres: protecting chromosomes against genome instability. 1–11.

Osborne CS, Chakalova L, Mitchell JA, Horton A, Wood AL, Bolland DJ, Corcoran AE, Fraser P. 2007. Myc Dynamically and Preferentially Relocates to a Transcription Factory Occupied by Igh ed. Susan M Gasser. *Plos Biol* **5**: e192–10.

Peaston AE, Evsikov AV, Graber JH, de Vries WN, Holbrook AE, Solter D, Knowles BB. 2004. Retrotransposons Regulate Host Genes in Mouse Oocytes and Preimplantation Embryos. *Developmental Cell* **7**: 597–606.

Peric-Hupkes D, Meuleman W, Pagie L, Bruggeman SWM, Solovei I, Brugman W, Gräf S, Flicek P, Kerkhoven RM, van Lohuizen M, et al. 2010. Molecular Maps of the Reorganization of Genome-Nuclear Lamina Interactions during Differentiation. *Mol Cell* **38**: 603–613.

Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schöfer C, Weipoltshammer K, Pagani M, Lachner M, Kohlmaier A, et al. 2001. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**: 323–337.

Poleshko A, Katz RA. 2014. Specifying peripheral heterochromatin during nuclear lamina reassembly. *Nucleus* **5**: 32–39.

Pontecorvo G. 1944. Structure of Heterochromatin. *Nature* **153**: 365–367.

Pope BD, Ryba T, Dileep V, Yue F, Wu W, Denas O, Vera DL, Wang Y, Hansen RS, Canfield TK, et al. 2014. Topologically associating domains are stable units of replication-timing regulation. *Nature* **515**: 402–405.

Pruitt KD, Maglott DR. 2001. RefSeq and LocusLink: NCBI gene-centered resources. *Nucleic Acids Research* **29**: 137–140.

Qi X, Li T-G, Hao J, Hu J, Wang J, Simmons H, Miura S, Mishina Y, Zhao G-Q. 2004. BMP4 supports self-renewal of embryonic stem cells by inhibiting mitogen-activated protein kinase pathways. *Proc Natl Acad Sci USA* **101**: 6027–6032.

Quenneville S, Verde G, Corsinotti A, Kapopoulou A, Jakobsson J, Offner S, Baglivo I, Pedone PV, Grimaldi G, Riccio A, et al. 2011. In Embryonic Stem Cells, ZFP57/KAP1 Recognize a Methylated Hexanucleotide to Affect Chromatin and DNA Methylation of Imprinting Control Regions. *Mol Cell* **44**: 361–372.

Rathert P, Dhayalan A, Murakami M, Zhang X, Tamas R, Jurkowska R, Komatsu Y, Shinkai Y, Cheng X, Jeltsch A. 2008. Protein lysine methyltransferase G9a acts on non-histone targets. *Nat Chem Biol* **4**: 344–346.

Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, Schmid M, Opravil S, Mechtler K,

Ponting CP, Allis CD, et al. 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**: 593–599.

Regnard C, Straub T, Mitterweger A, Dahlsveen IK, Fabian V, Becker PB. 2011. Global Analysis of the Relationship between JIL-1 Kinase and Transcription ed. B. Van Steensel. *PLoS Genet* **7**: e1001327.

Reyskens KMSE, Arthur JSC. 2016. Emerging Roles of the Mitogen and Stress Activated Kinases MSK1 and MSK2. *Front Cell Dev Biol* **4**: 2426–8.

Rhead B, Karolchik D, Kuhn RM, Hinrichs AS, Zweig AS, Fujita PA, Diekhans M, Smith KE, Rosenbloom KR, Raney BJ, et al. 2009. The UCSC Genome Browser database: update 2010. *Nucleic Acids Research* **38**: D613–D619.

Ribeiro-Mason K, Boulesteix C, Fleurot R, Aguirre-Lavin T, Adenot P, Gall L, Debey P, Beaujean N. 2012. H3S10 phosphorylation marks constitutive heterochromatin during interphase in early mouse embryos until the 4-cell stage. *J Reprod Dev* **58**: 467–475.

Rossetto D, Avvakumov N, Côté J. 2012. Histone phosphorylation: a chromatin modification involved in diverse nuclear events. *epigenetics* **7**: 1098–1108.

Rothbart SB, Dickson BM, Raab JR, Grzybowski AT, Krajewski K, Guo AH, Shanle EK, Josefowicz SZ, Fuchs SM, Allis CD, et al. 2015. An Interactive Database for the Assessment of Histone Antibody Specificity. *Mol Cell* **59**: 502–511.

Rothbart SB, Krajewski K, Nady N, Tempel W, Xue S, Badeaux AI, Barsyte-Lovejoy D, Martinez JY, Bedford MT, Fuchs SM, et al. 2012. Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nature Structural & Molecular Biology*.

Roux PP, Blenis J. 2004. ERK and p38 MAPK-activated protein kinases: a family of protein kinases with diverse biological functions. *Microbiology and Molecular Biology Reviews* **68**: 320–344.

Rowe HM, Friedli M, Offner S, Verp S, Mesnard D, Marquis J, Aktas T, Trono D. 2013. De novo DNA methylation of endogenous retroviruses is shaped by KRAB-ZFPs/KAP1 and ESET. *Development* **140**: 519–529.

Rowe HM, Jakobsson J, Mesnard D, Rougemont J, Reynard S, Aktas T, Maillard PV, Layard-Liesching H, Verp S, Marquis J, et al. 2010. KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* **463**: 237–240.

Saccani S, Pantano S, Natoli G. 2001. p38-dependent marking of inflammatory genes for increased NF-κB recruitment. *Nat Immunol* **3**: 69–75.

Sadic D, Schmidt K, Groh S, Kondofersky I, Ellwart J, Fuchs C, Theis FJ, Schotta G. 2015. Atrx promotes heterochromatin formation at retrotransposons. *EMBO Rep* **16**: 836–850.

Sahlén P, Abdullayev I, Ramsköld D, Matskova L, Rilakovic N, Lötstedt B, Albert TJ, Lundeberg J, Sandberg R. 2015. Genome-wide mapping of promoter-anchored interactions with close to single-enhancer resolution. *Genome Biol* 1–13.

Sakaue-Sawano A, Kurokawa H, Morimura T, Hanyu A, Hama H, Osawa H, Kashiwagi S, Fukami K, Miyata T, Miyoshi H, et al. 2008. Visualizing Spatiotemporal Dynamics of Multicellular Cell-Cycle Progression. *Cell* **132**: 487–498.

Santos F, Peters AH, Otte AP, Reik W, Dean W. 2005. Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. *Dev Biol* **280**: 225–236.

Sasagawa Y, Nikaido I, Hayashi T, Danno H, Uno KD, Imai T, Ueda HR. 2013. Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. *Genome Biol* **14**: R31.

Sassone-Corsi P, Mizzen CA, Cheung P, Crosio C, Monaco L, Jacquot S, Hanauer A, Allis CD. 1999. Requirement of Rsk-2 for epidermal growth factor-activated phosphorylation of histone H3. *Science* **285**: 886–891.

Sato Y, Mukai M, Ueda J, Muraki M, Stasevich TJ, Horikoshi N, Kujirai T, Kita H, Kimura T, Hira S, et al. 2013. Genetically encoded system to track histone modification in vivo. *Sci Rep* **3**: 57–7.

Sawicka A, Seiser C. 2012. Histone H3 phosphorylation – A versatile chromatin modification for different occasions. *Biochimie* **94**: 2193–2201.

Schrode N, Saiz N, Di Talia S, Hadjantonakis A-K. 2014. GATA6 Levels Modulate Primitive Endoderm Cell Fate Choice and Timing in the Mouse Blastocyst. *Developmental Cell* **29**: 454–467.

Schultz DC. 2002. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev* **16**: 919–932.

Schultz J. 1936. Variegation in Drosophila and the inert chromosome regions.

Sharif J, Endo TA, Nakayama M, Karimi MM, Shimada M, Katsuyama K, Goyal P, Brind'Amour J, Sun M-A, Sun Z, et al. 2016. Activation of Endogenous Retroviruses in Dnmt1-/- ESCs Involves Disruption of SETDB1-Mediated Repression by NP95 Binding to Hemimethylated DNA. *Stem Cell* 1–15.

Sharif J, Muto M, Takebayashi S-I, Suetake I, Iwamatsu A, Endo TA, Shinga J, Mizutani-Koseki Y, Toyoda T, Okamura K, et al. 2007. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450**: 908–912.

Shinkai Y, Tachibana M. 2011. H3K9 methyltransferase G9a and the related molecule GLP. *Genes Dev* **25**: 781–788.

Shogren-Knaak MA. 2003. A Native Peptide Ligation Strategy for Deciphering Nucleosomal Histone Modifications. *Journal of Biological Chemistry* **278**: 15744–15748.

Skourti-Stathaki K, Kamieniarz-Gdula K, Proudfoot NJ. 2014. R-loops induce repressive chromatin marks over mammalian gene terminators. *Nature*.

Slotkin RK, Martienssen R. 2007. Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* **8**: 272–285.

Smith SB, Cui Y, Bustamante C. 1996. Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science* **271**: 795–799.

Soloaga A, Thomson S, Wiggin GR, Rampersaud N, Dyson MH, Hazzalin CA, Mahadevan LC, Arthur JSC. 2003. MSK2 and MSK1 mediate the mitogen- and stress-induced phosphorylation of histone H3 and HMG-14. *EMBO J* **22**: 2788–2797.

Sousa Martins JP, Liu X, Oke A, Arora R, Franciosi F, Viville S, Laird DJ, Fung JC, Conti M. 2016. DAZL and CPEB1 regulate mRNA translation synergistically during oocyte maturation. *Journal of Cell Science* **129**: 1271–1282.

Stavridis MP, Lunn JS, Collins BJ, Storey KG. 2007. A discrete period of FGF-induced Erk1/2 signalling is required for vertebrate neural specification. *Development* **134**: 2889–2894.

Stipanovich A, Valjent E, Matamales M, Nishi A, Ahn J-H, Maroteaux M, Bertran-Gonzalez J, Brami-Cherrier K, Enslen H, Corbillé A-G, et al. 2008. A phosphatase cascade by which rewarding stimuli control nucleosomal response. *Nature* **453**: 879–884.

Stocking C, Kozak CA. 2008. Endogenous retroviruses. *Cell Mol Life Sci* **65**: 3383–3398.

Stoye JP, Coffin JM. 1988. Polymorphism of murine endogenous proviruses revealed by using virus class-specific oligonucleotide probes. *J Virol* **62**: 168–175.

Strahl BD, Allis CD. 2000. The language of covalent histone modifications. *Nature* **403**: 41–45.

Strathdee D, Ibbotson H, Grant SGN. 2006. Expression of Transgenes Targeted to the Gt(ROSA)26Sor Locus Is Orientation Dependent ed. P. Fraser. *PLoS ONE* **1**: e4–9.

Sun FL, Haynes K, Simpson CL, Lee SD, Collins L, Wuller J, Eissenberg JC, Elgin SCR. 2004a. cis-Acting Determinants of Heterochromatin Formation on Drosophila melanogaster Chromosome Four. *Molecular and Cellular Biology* **24**: 8210–8220.

Sun FL, Haynes K, Simpson CL, Lee SD, Collins L, Wuller J, Eissenberg JC, Elgin SCR. 2004b. cis-Acting Determinants of Heterochromatin Formation on Drosophila melanogaster Chromosome Four. *Molecular and Cellular Biology* **24**: 8210–8220.

Tachibana M. 2005. Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes Dev* **19**: 815–826.

Tachibana M, Matsumura Y, Fukuda M, Kimura H, Shinkai Y. 2008. G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *EMBO J* **27**: 2681–2690.

Tachibana M, Sugimoto K, Nozaki M, Ueda J, Ohta T, Ohki M, Fukuda M, Takeda N, Niida H, Kato H, et al. 2002. G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev* **16**: 1779–1791.

Takada Y, Naruse C, Costa Y, Shirakawa T, Tachibana M, Sharif J, Kezuka-Shiotani F, Kakiuchi D, Masumoto H, Shinkai YI, et al. 2011. HP1 links histone methylation marks to meiotic synapsis in mice. *Development* **138**: 4207–4217.

Takeuchi A, Mishina Y, Miyaishi O, Kojima E, Hasegawa T, Isobe K-I. 2003. Heterozygosity with respect to Zfp148 causes complete loss of fetal germ cells during mouse embryogenesis. *Nat Genet* **33**: 172–176.

Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ. 2007. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nature Structural & Molecular Biology* **14**: 1025–1040.

Teperek-Tkacz M, Meglicki M, Pasternak M, Kubiak JZ, Borsuk E. 2010. Phosphorylation of histone H3 serine 10 in early mouse embryos: active phosphorylation at late S phase and differential effects of ZM447439 on first two embryonic mitoses. *Cell Cycle* **9**: 4674–4687.

Thompson PJ, Dulberg V, Moon K-M, Foster LJ, Chen C, Karimi MM, Lorincz MC. 2015. hnRNP K coordinates transcriptional silencing by SETDB1 in embryonic stem cells. *PLoS Genet* **11**: e1004933.

Thompson PJ, Macfarlan TS, Lorincz MC. 2016. Long Terminal Repeats: From Parasitic Elements to Building Blocks of the Transcriptional Regulatory Repertoire. *Mol Cell* **62**: 766–776.

Thomson S, Clayton AL, Hazzalin CA, Rose S, Barratt MJ, Mahadevan LC. 1999a. The nucleosomal response associated with immediate-early gene induction is mediated via alternative MAP kinase cascades: MSK1 as a potential histone H3/HMG-14 kinase. *EMBO J* **18**: 4779–4793.

Thomson S, Mahadevan LC, Clayton AL. 1999b. MAP kinase-mediated signalling to nucleosomes and immediate-early gene induction. *Semin Cell Dev Biol* **10**: 205–214.

Tomas-Zuber M, Mary JL, Lamour F, Bur D, Lesslauer W. 2001. C-terminal Elements Control Location, Activation Threshold, and p38 Docking of Ribosomal S6 Kinase B (RSKB). *Journal of Biological Chemistry* **276**: 5892–5899.

Treangen TJ, Salzberg SL. 2012. Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat Rev Genet* **13**: 36–46.

Van Hooser A, Goodrich DW, Allis CD, Brinkley BR, Mancini MA. 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.

Velasco G, Hube F, Rollin J. 2010. Dnmt3b recruitment through E2F6 transcriptional repressor mediates germ-line gene silencing in murine somatic tissues.

Vermeulen L, Berghe WV, Beck IME, De Bosscher K, Haegeman G. 2009. The versatile role of MSKs in transcriptional regulation. *Trends in Biochemical Sciences* **34**: 311–318.

Vogel MJ, Guelen L, de Wit E, Hupkes DP, Loden M, Talhout W, Feenstra M, Abbas B, Classen AK, van Steensel B. 2006. Human heterochromatin proteins form large domains containing KRAB-ZNF genes. *Genome Research* **16**: 1493–1504.

Walter J, Schermelleh L, Cremer M, Tashiro S, Cremer T. 2003. Chromosome order in HeLa cells changes during mitosis and early G1, but is stably maintained during subsequent interphase stages. *The Journal of Cell Biology* **160**: 685–697.

Wang C, Cai W, Li Y, Deng H, Bao X, Girton J, Johansen J, Johansen KM. 2012. The epigenetic H3S10 phosphorylation mark is required for counteracting heterochromatic spreading and gene silencing in Drosophila melanogaster. *Journal of Cell Science* **124**: 4309–4317.

Wang F, Higgins JMG. 2012. Histone modifications and mitosis: countermarks, landmarks, and bookmarks. *Trends Cell Biol* 1–10.

Wang Y, Zhang W, Jin Y, Johansen J, Johansen KM. 2001. The JIL-1 tandem kinase mediates histone H3 phosphorylation and is required for maintenance of chromatin structure in Drosophila. *Cell* **105**: 433–443.

Wei Y, Mizzen CA, Cook RG, Gorovsky MA, Allis CD. 1998. Phosphorylation of histone H3 at serine 10 is correlated with chromosome condensation during mitosis and meiosis in Tetrahymena. *Proc Natl Acad Sci USA* **95**: 7480–7484.

Wei Y, Yu L, Bowen J, Gorovsky MA, Allis CD. 1999. Phosphorylation of Histone H3 Is Requiredfor Proper Chromosome Condensation and Segregation. *Cell, Vol , ,* 1–11.

Wen B, Wu H, Shinkai Y, Irizarry RA, Feinberg AP. 2009. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat Genet* **41**: 246–250.

Wiggin GR, Soloaga A, Foster JM, Murray-Tait V, Cohen P, Arthur JSC. 2002. MSK1 and MSK2 Are Required for the Mitogen- and Stress-Induced Phosphorylation of CREB and ATF1 in Fibroblasts. *Molecular and Cellular Biology* **22**: 2871–2881.

Wolf D, Goff SP. 2009. Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature* **458**: 1201–1204.

Yamamizu K, Fujihara M, Tachibana M, Katayama S, Takahashi A, Hara E, Imai H, Shinkai Y, Yamashita JK. 2012. Protein Kinase A Determines Timing of Early Differentiation through Epigenetic Regulation with G9a. *Cell Stem Cell* **10**: 759–770.

Yang X, Matsuda K, Bialek P, Jacquot S, Masuoka HC, Schinke T, Li L, Brancorsini S, Sassone-Corsi P, Townes TM, et al. 2004. ATF4 is a substrate of RSK2 and an essential regulator of osteoblast biology; implication for Coffin-Lowry Syndrome. *Cell* **117**: 387–398.

Ying Q-L, Wray J, Nichols J, Batlle-Morera L, Doble B, Woodgett J, Cohen P, Smith A. 2008. The ground state of embryonic stem cell self-renewal. *Nature* **453**: 519–523.

Yokochi T, Poduch K, Ryba T, LU J, Hiratani I, Tachibana M, Shinkai Y, Gilbert DM. 2009. G9a selectively represses a class of late-replicating genes at the nuclear periphery. *Proceedings of the National Academy of Sciences* **106**: 19363–19368.

Younesy H, Möller T, Lorincz MC, Karimi MM, Jones SJM. 2015. VisRseq: R-based visual framework for analysis of sequencing data. *BMC Bioinformatics* **16 Suppl 11**: S2.

Yuan P, Han J, Guo G, Orlov YL, Huss M, Loh Y-H, Yaw L-P, Robson P, Lim B, Ng H-H. 2009. Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells. *Genes Dev* **23**: 2507–2520.

Yue F, Cheng Y, Breschi A, Vierstra J, Wu W, Ryba T, Sandstrom R, Ma Z, Davis C, Pope BD, et al. 2014. A comparative encyclopedia of DNA elements in the mouse genome. *Nature* **515**: 355–364.

Zamiri B, Mirceta M, Bomsztyk K, Macgregor RB Jr., Pearson CE. 2015. Quadruplex formation by both G-rich and C-rich DNA strands of the C9orf72(GGGGCC)8•(GGCCCC)8 repeat: effect of CpG methylation. *Nucleic Acids Research* gkv1008–10.

Zhang W, Deng H, Bao X, Lerach S, Girton J, Johansen J, Johansen KM. 2006. The JIL-1 histone H3S10 kinase regulates dimethyl H3K9 modifications and heterochromatic spreading in Drosophila. *Development* **133**: 229–235.

Zhao Y, Bjørbaek C, Weremowicz S, Morton CC, Moller DE. 1995. RSK3 encodes a novel pp90rsk isoform with a unique N-terminal sequence: growth factor-stimulated kinase function and nuclear translocation. *Molecular and Cellular Biology* **15**: 4353–4363.

Zhou H, Liu L-H, Zhang H, Lei Z, Lan Z-J. 2010. Expression of zinc finger protein 105 in the testis and its role in male fertility. *Mol Reprod Dev* **77**: 511–520.

Zippo A, Serafini R, Rocchigiani M, Pennacchini S, Krepelova A, Oliviero S. 2009. Histone crosstalk between H3S10ph and H4K16ac generates a histone code that mediates transcription elongation. *Cell* **138**: 1122–1136.

Zylicz JJ, Dietmann S, Günesdogan U, Hackett JA. 2015. Chromatin dynamics and the role of G9a in gene regulation and enhancer silencing during early mouse development. *eLife*.