The role of transcription and splicing on histone H3 lysine 4 trimethylation dynamics

By

Veronica Lau

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

Department of Biochemistry and Medical Genetics

University of Manitoba

Winnipeg

ABSTRACT

Histone H3 lysine 4 trimethylation (H3K4me3) is a histone post-translational modification that is normally located near the transcription start site and first 5' splice site of a gene. Enzymes are involved in regulating the addition and removal of H3K4me3 along the length of the gene body. However, little is understood about these changes during transcriptional induction. Additionally, the effects of splicing inhibition have not been examined in an endogenous system. We hypothesized that transcription and pre-mRNA splicing at the 5' coding region plays a role in H3K4me3 dynamics. Chromatin immunoprecipitation assays ± splicing (isoginkgetin) or transcription (5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole) inhibitors were done to determine the H3K4me3 levels along the immediate early gene *Fos*. The gene was stimulated with epidermal growth factor, and monitored at various times. H3K4me3 levels increased and decreased along the induced *Fos* gene. Inhibition of transcription or splicing prevented these changes in H3K4me3 levels during transcription, suggesting that splicing may influence H3K4me3 levels indirectly. The significance of these finding is that it provides a better understanding of how H3K4me3 is regulated during transcription and splicing of a gene.

ACKNOWLEDGEMENTS

I am grateful for this opportunity given to me. First, I would like to thank my supervisor and mentor Dr. James Davie for all his support and help throughout the process. His knowledge, experience and guidance was greatly appreciated and helped me overcome many problems throughout my studies. His enthusiasm and passion for understanding science was very motivating, and inspirational.

I would also like to thank my committee members, Dr. Kirk McManus and Dr. Yvonne Myal. Their time, suggestion and advice was greatly appreciated.

In addition, I would like to thank the current and past members of the Davie lab. Carolina Gonzalez-Zuluaga, Aleksandar Ilic, Sanzida Jahan, Cheryl Peltier, Kiran Sharma and Ife Adewumi, helped me learn many techniques throughout my studies. A special thanks to Aleksandar Ilic, Sanzida Jahan and Ife Adewumi, who have all been incredibly helpful through my studies. I would also like to thank Debbie Tsuyuki and Wayne Xu for their help in sequencing.

Finally, I would like to thank my family, my parents Samantha and Stephen Lau, my sister Marintha, and Basil. For having faith in me in times when I, myself questioned it. I would also like to thank Luther for his patience, support and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF COPYRIGHT FIGURES FOR WHICH PERMISSION WAS OBTAINED

LIST OF ABBREVIATION

^{o}C	Degrees Celsius
A_{260}	Absorbance at 260 nm
A_{280}	Absorbance at 280 nm
ASH2L	Set1/Ash2 histone methyltransferase complex subunit ASH2
ATAC	Ada-Two-A-containing 2A
ATCC	American type culture collection
bp	Base pair
$CaCl_2$	Calcium chloride
cDNA	Complementary DNA
CFP1	CXXC finger protein 1
CHD1	Chromodomain-helicase-DNA-binding protein 1
ChIP	Chromatin immunoprecipitation
ChIP-seq	Chromatin immunoprecipitation sequencing
CO_2	Carbon Dioxide
CpG	Cytosine-phosphate-guanine
ddH_2O	Double distilled water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DPY-30	Protein dpy-30 homolog
DRB	Dichloro-1-beta-D-ribofuanosylbenzimidazole
Dusp1	Dual specificity phosphatase 1

Dusp2	Dual specificity phosphatase 2
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetracetic acid
EGF	Epidermal growth factor
ENCODE	Encyclopedia of DNA elements
Erasers	Enzymes that remove post-translational modifications
ERK	Extracellular signal-regulated kinase
Fos	FBS osteosarcoma oncogene
Fosl1	Fos-like antigen 1
FBS	Fetal bovine serum
H2B	Histone H2B
uH2B	Ubiquitinated histone H2B
H3K4	Histone H3 lysine 4
H3K4me3	Histone H3 lysine 4 trimethylation
H3R2me2s	Symmetric dimethylation of H3R2
HCl	Hydrochloric acid
HDAC	Histone deacetylase
hr	Hours
HRP	Horse radish peroxidase
IgG	Immunoglobulin G
ING4	Inhibitor of growth 4
Jun	Jun proto-oncogene
KAT	Lysine acetyltransferase

KCl	Potassium chloride
KMD	Lysine demethylase
KDM5A/B/C/D/E	lysine demethylase family 5 member A/B/C/D/E
KMT	Lysine methyltransferase
LiCl	Lithium chloride
MEM	Minimum essential media
min	Minutes
ml	Milliliter
MLL	Mixed-lineage leukemia
mM	Millimolar concentration
MNase	Micrococcal nuclease
MSK	Mitogen- and stress-activated protein kinase
MYC	Myc proto-oncogene protein
NaCl	Sodium chloride
$NaHCO_3$	Sodium bicarbonate
ng	Nanogram
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffer saline
PIPES	Piperazine-N,N'-bis(2-ethanesulfonic acid)
RAS MAPK	RAS mitogen-activated protein kinase
RBBP5	Retinoblastoma-binding protein 5
Reader	Non-histone proteins recruited by histone post-translational modifications
Recruiters	Recruiting molecules that directs SETD1A/B for trimethylation of H3K4

RNA	Ribonucleic acid
RNase	Ribonuclease
RT-qPCR	Reverse transcriptase quantitative PCR
SAGA	Spt-Ada-Gcn5 acetyltransferase
SDC	Sodium deoxycholate
SDS	Sodium dodecyl sulfate
SETD1A/B	Set domain containing 1 A/B
snRNP	Small nuclear ribonucleoprotein particles
TAF3	TATA-Box Binding Protein Associated Factor 3
TBE	Tris borate EDTA
TBST	Tris-buffer saline solution with tween
TSS	Transcription start site
U	Units
ug	Microgram
uH2B	Ubiquitinated H2B
ul	Microliter
uM	Micromolar concentration
V	Volts
WDR5	WD repeat-containing protein 5
WRAD complex	Methyltransferase complex containing WDR5, RBBP5, ASH2L and DPY30
Writers	Enzymes that add post-translational modifications
xg	Times gravity

1. INTRODUCTION

1.1 Chromatin Structure and Histone Post-translational Modification

Eukaryotic nuclear DNA is packaged into chromatin. The structural unit of the chromatin is the nucleosome which consists of 147 bp of DNA wrapped around a histone octamer. The core histones that comprises the histone octamer are H2A, H2B, H3 and H4. Each core histone has a similar basic structure and have N-terminal tails that radiate outside of the nucleosome. These N-terminal tails are capable of interacting with linker DNA, proteins and nearby nucleosomes (Figure 1) (Davie et al. 2015; Howe et al. 2017; Madar et al. 1997; Maze et al. 2014).

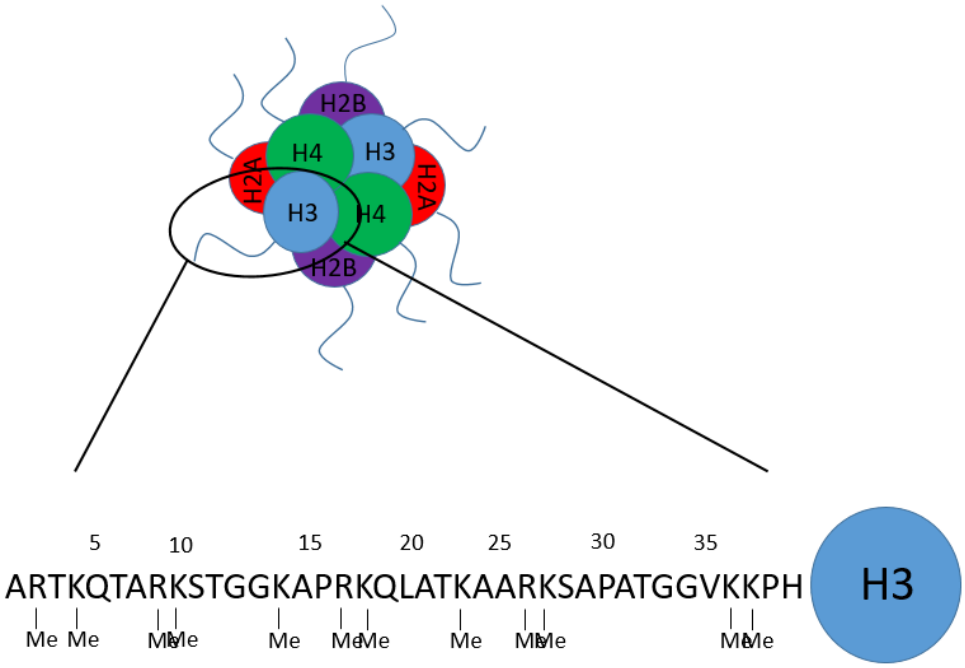


Figure 1. The histone octamer in nucleosome structure.

A) Basic histone structure of the histone octamer. The N-terminal tails radiate out of the histone octamer. B) Magnification of the histone H3. The potential methylation sites are indicated along the N-terminal tail of the histone H3. Figure was adapted from Figure 1 by Davie et al. 2015.

A variety of histone post-translational modifications can occur along the core histones. These modifications include acetylation, phosphorylation and methylation, which are all important in gene expression and epigenetic regulation. Epigenetics is the study of meiotically and mitotically heritable alterations in gene expression without mutations or alterations to the underlying DNA sequence. Epigenetic marks are Although epigenetic marks are established in early development, they are dynamic and are greatly impacted by environmental and intrinsic factors (Delcuve, Rastegar and Davie 2009; Tsai and Baylin, 2011). Histone post-translational modifications are

reversible and are regulated by a variety of enzymes. Post-translational modifications can be added on by specific enzymes ("writer") such as lysine methyltransferases (KMT), and lysine acetyltransferases (KAT), or removed by enzymes ("erasers") like histone deacetylase (HDAC) and lysine demethylase (KMD) (Davie et al. 2015; Allis et al. 2007). Histone post-translational modifications can also promote the recruitment of non-histone proteins ("readers") to specific chromatin regions. Histone post-translational modifications can be associated with transcriptionally active regions or transcriptionally inactive region dependent on the type of post-translational modification present (Davie et al. 2015; Turner 2014).

1.2 Histone H3K4 Trimethylation (H3K4me3)

1.2.1 H3K4me3 genomic position and distribution

Methylation of histone H3 lysine 4 (H3K4) can be the addition of one, two or three methyl groups. Histone H3 lysine 4 trimethylation (H3K4me3) has been associated with transcriptionally active genes (Howe et al. 2017; Santos-Rosa et al. 2002). Due to the close association of the H3K4me3 mark with the transcription start site (TSS), it was believed that H3K4me3 was associated with transcription (Sims et al. 2007; Howe et al. 2017). However, one study which used *in vitro* Chromatin immunoprecipitation (ChIP) assays indicated that the H3K4me3 mark did not directly affect transcription (Pavri et al. 2006). Instead, with the demonstration through functional transcription assays, active transcription was required for the establishment of H3K4me3 levels, suggesting that H3K4me3 was activated downstream of transcription initiation or during reinitiation (Pavri et al. 2006; Sims et al. 2007).

This finding was further supported by two other studies (Bieberstein et al. 2012; Edmunds et al. 2008). Using the ChIP sequencing (ChIP-seq) data available from the ENCODE Consortium; all human cell line tracks were compared. What was observed was that the H3K4me3 peak was more downstream than expected, and the highest level of H3K4me3 was located at the first 5' splice site (Bieberstein et al. 2012) (Figure 2). One study used ChIP assays to map the levels of H3K4me3 along the mouse *Fos* gene using 10T1/2 cells and determined that the H3K4me3 peak was in the first intron (Edmunds et al. 2008). With the position of H3K4me3 at the 5' splice site, a relationship between splicing and H3K4me3 was suspected. This relationship was tested with the establishment of transgenic cell lines and the use of the splicing inhibitor spliceostatin A. Spliceostatin A reduced splicing by 50%, and the level of H3K4me3 was lost at the 5' splice site (Beiberstein et al. 2012).

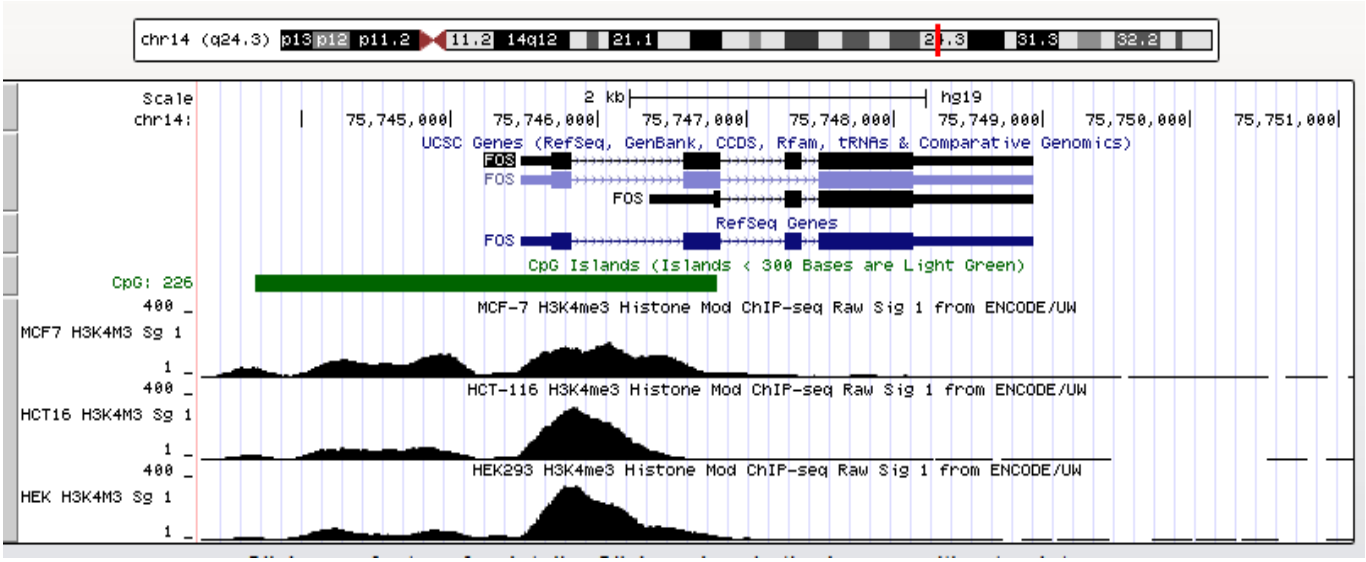


Figure 2. Distribution of H3K4me3 along the *FOS* gene.

Levels of H3K4me3 are shown above for the gene *FOS* for the human cell lines HCT116, MCF7 and HEK293. Note the highest peak of H3K4me3 was observed at the exon 1/ intron 1 junction in both HCT116 and HEK293 cells. However, the highest peak was not observed at this junction in MCF7 cells. Arrow heads along the *FOS* introns on the gene map indicates the direction of transcription. Data were obtained from ENCODE *(http://genome.ucsc.edu/ENCODE/dataMatrix/encodeDataMatrixHuman.html).*

H3K4me3 location has also been indicated to be associated with CpG islands. During the introduction of artificial CpG clusters into the mouse genome, sites originally not associated with H3K4me peaks displayed H3K4me3 associated nucleosomes (Thomson et al. 2010; Davie et al. 2015). The SET1 complex is a methyltransferase of H3K4me3, and contains the subunit CXXC finger protein 1 (CFP1). CFP1 binds to unmethylated CpG island. Knockdown of Cfp1 in mouse embryonic stem cells caused two changes in H3K4me3 locations; there was a significant loss of

H3K4me3 levels on genes associated with CpG islands, and ectopic location of H3K4me3 at other regulatory regions (Clouaire et al. 2012). Although both CpG islands and 5' splice sites are associated with H3K4me3 location, both are not necessary for the formation of H3K4me3 sites (Thomson et al. 2010; Davie et al. 2015).

1.2.2 H3K4me3 regulators: writers, recruiters, erasers and readers

The methyltransferases responsible for the methylation of H3K4 is highly conserved in species from *Saccharomyces cerevisiae* to humans (Bernstein et al. 2005; van Nuland et al. 2013; Shilatifard 2012). In eukaryotes, there are methyltransferases involved in the methylation of H3K4; Set domain containing 1B and 1A (SETD1B and SETD1A) and the mixed-lineage leukemia MLL family (MLL1-4). SETD1A and SETD1B are the main enzymes responsible for trimethylation, while MLL1-4 are more involved in monomethylation and dimethylations (Shilatifard 2012; Shinsky et al. 2015; van Nuland et al. 2013; Davie et al. 2015). Each methyltransferase enzyme consists of the core subunits of ASH2L (Absent, small or homeotic-like), WDR5 (WD Repeat Domain 5), RBBP5 (Retinoblastoma Binding Protein 5) and DPY-30 (Dumpy-30) and is referred to as the WRAD complex. The WRAD complex is involved in regulation of enzymatic activity, and is associated with the catalytic subunit of the methyltransferases (Steward et al. 2006; Patel et al. 2011; van Nuland et al. 2013). SETD1A and SETD1B contain the protein CFP1 within their complexes, and SETD1B has the additional protein biorientation of chromosomes in cell division protein 1-like (BOD1L) which is unique to the SETD1B complex (van Nuland et al. 2013) (Figure 3).

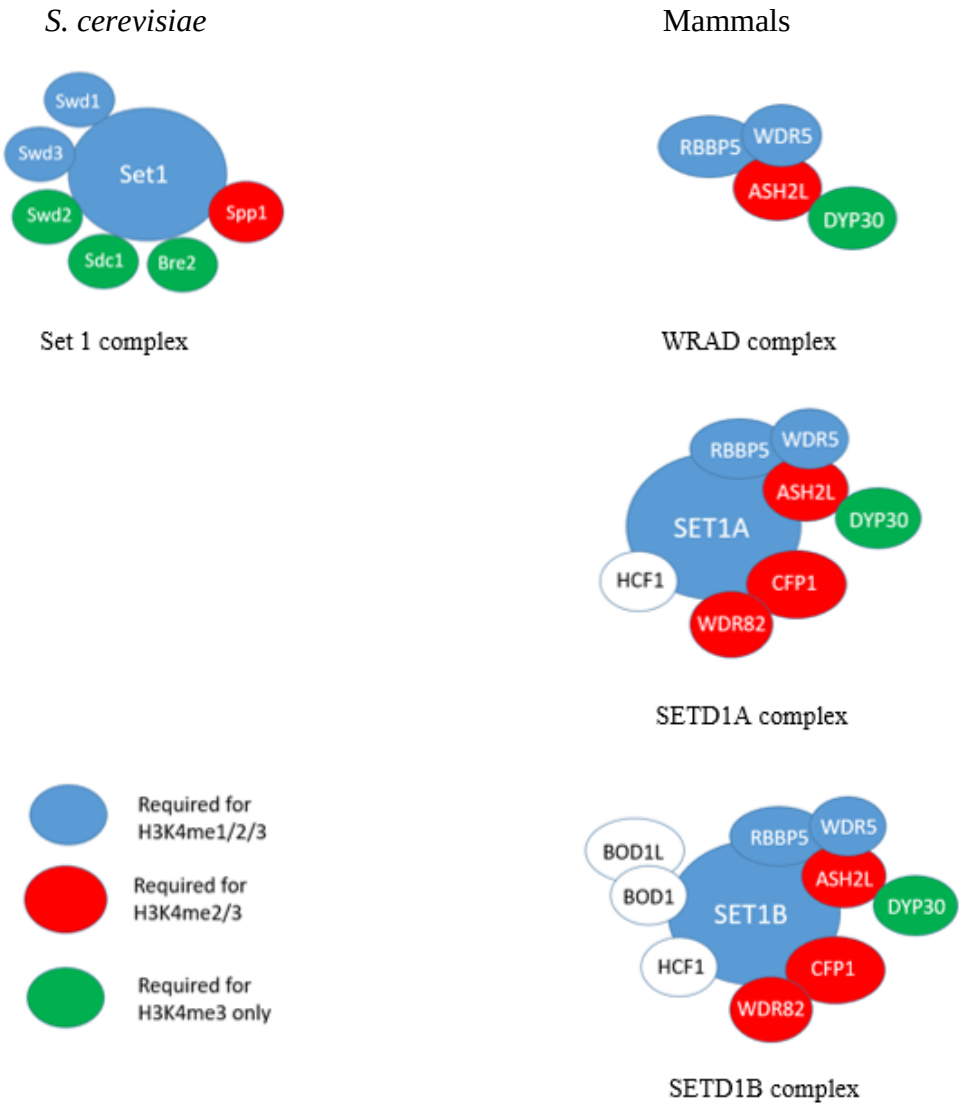


Figure 3. Composition summary of the Set1 complex in yeast, and the SETD1A/B complex in mammals, which contains the WRAD complex.

The subunits are colored blue, red or green indicating those are required for H3K4me1/2/3, H3K4me2/3 or H3K4me3. Figure is based on Figure 1 from Howe et al. 2016, and Figure 6B from van Nuland et al. 2013.

The location of SETD1A and SETD1B for the trimethylation of H3K4 is regulated by a variety of recruiting molecules. These “recruiters” of SETD1A and SETD1B include CFP1, ubiquitinated H2B (uH2B), and histone H3 arginine 2 dimethylation symmetric (H3R2me2s). As described previously, CFP1 is a component of the SETD1A and SETD1B complex, and is involved in the placement of H3K4me3 marks on unmethylated CpG islands (Clouaire et al. 2012). Ring finger protein 20 mediated uH2B was originally thought to be required for the recruitment of H3K4me3 through the SETD1A and SETD1B complexes (Shilatifard 2012). However, it was later discovered that H3K4me3 levels were unaffected after global loss of uH2B level during myogenic differentiation. Downregulated genes in myotubes, showed a parallel decline in H3K4me3 and uH2B levels, suggesting that the impact of uH2B on H3K4me3 was gene specific (Vethantham et al. 2012). H3R2me2s is colocalized with H3K4me3 at the TSS and is involved in the prevention of heterochromatin formation. WDR5 in the WRAD complex is recruited by H3R2me2s through interacting with the W40 domain of the WDR5 (Jahan and Davie 2015).

There are four main KDMs, which are all part of the KDM5 family, that are responsible for H3K4me3 demethylation (5A/JARID1A/RBP2, 5B/JARID1B/PLU1, 5C/JARID1C/SMCX and 5D/JARID1D/SMCY). The recruitment of the KDM5 family to regulatory region is due to various interactions to transcription factors (Davie et al. 2015). Members of the KDM5 family are associated with androgen receptors in prostate cancer cells, myc proto-oncogene protein (MYC) and HDAC1/2 (Davie et al. 2015). Luciferase reporter assays were done to determine the association of androgen receptor and KDM5, and the transfection of KDM5B led to a dose dependent enhancement in androgen receptor activity (Xiang et al. 2007). The relationship between KDM5 and MYC was determined through co-immunoprecipitation assays which

indicated requirement of Transcription Factor AP 2 gamma C-terminal for MYC and KDM5B (Wong et al. 2012). Through *in vivo* immunoprecipitation and *in vitro* pull down assays, KDM5B was found to co-precipitate and bind directly to HDAC1/2 (Barrett et al. 2007). KDM5 can also undergo post-translational modification such as phosphorylation, ubiquitination and acetylation, for further regulation KDM5 location and/or activity (Hornbeck et al. 2015).

H3K4me3 is a histone post-translational modification that is recognized by many "readers". These include proteins such as those involved with transcription (e.g. SAGA-associated factor 29, TAF3 and ING4) and those involved with splicing (e.g. CHD1) (Davie et al. 2015). SAGA-associated factor 29 is a component of the lysine acetyltransferase SAGA (Spt-Ada-Gcn5 acetyltransferase) and ATAC (Ada-Two-A-containing 2A) co-activator complexes. During endoplasmic reticulum stress, SAGA binds to H3K4me3 on promoter regions of endoplasmic reticulum stress genes, and promotes cell survival (Schram et al. 2013). TATA-Box Binding Protein Associated Factor 3 (TAF3) is a component of the transcription factor IID, which is a basal transcription factor involved in the recruitment of RNA polymerase II. Through stable isotope labeling by amino acid in cell culture, the plant homeo-domain finger of TAF3 was shown to bind to H3K4me3. Furthermore, increased expression of TAF3 caused enhanced transcription through stimulation of ASH2L, a component of SET1/MLL (Vermeulen et al. 2007). ING4 (inhibitor of growth 4) is a tumor suppressor protein that couples with the KAT complex, HBO1. HBO1 is known to acetylate mostly histone H4. Through peptide arrays, it was determined that ING4 bound strongly to H3K4me3 (Hung et al. 2009). Through further investigation, it was determined that ING4 was required to bind to H3K4me3 to promote cell death (Hung et al. 2009). CHD1 is a chromatin remodeler that is involved in transcription elongation. Through peptide affinity columns, CHD1 was shown to

bind to H3K4me3. It was also observed through co-immunoprecipitation studies that CHD1 had a strong interaction with splicing factor 3a and splicing factor 3b subunits of U2 small nuclear ribonucleoprotein particles (snRNP), a component of the spliceosome. When levels of CHD1 or H3K4me3 were diminished, splicing factor 3a activity was reduced, suggesting that CHD1 may act as a "bridge" to recruit the spliceosome to H3K4me3 sites (Sims et al. 2007) (Figure 4)

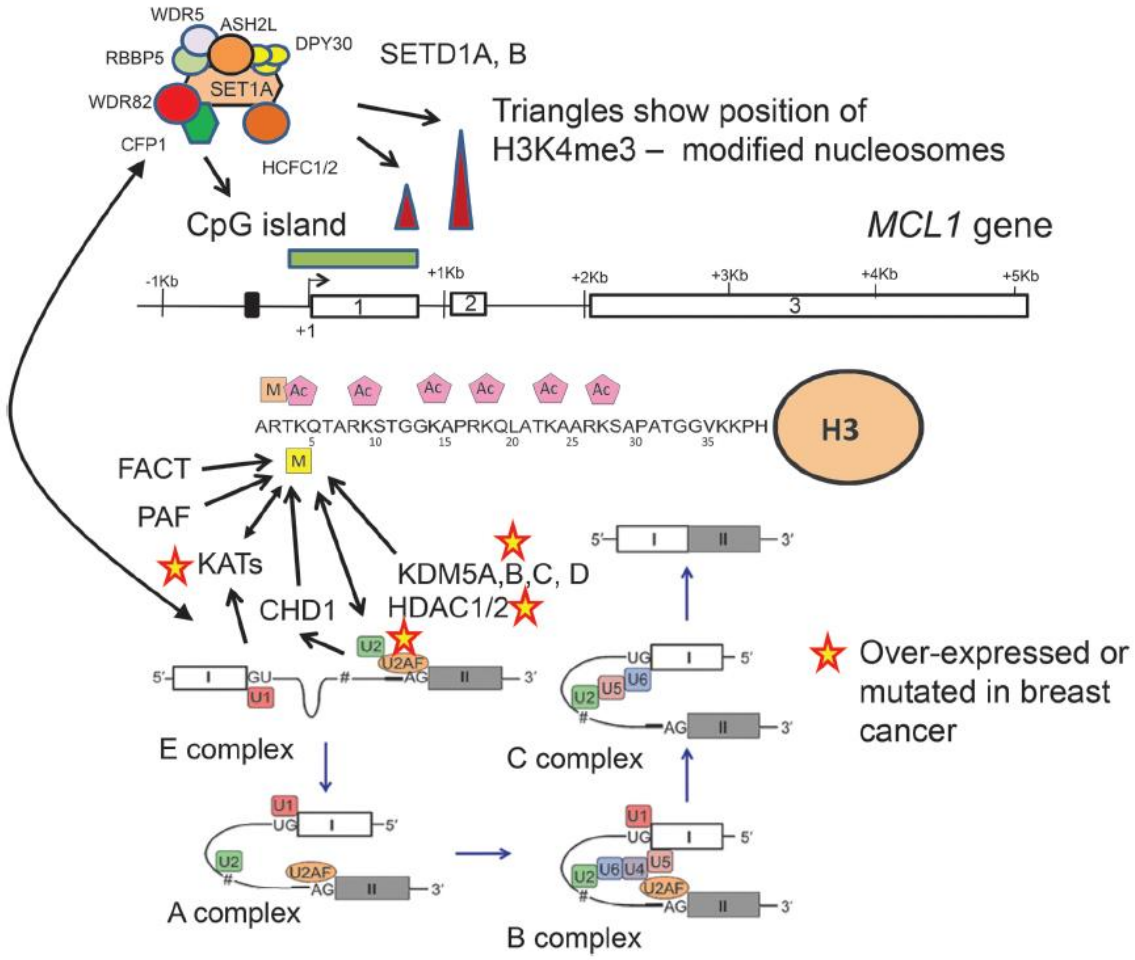


Figure 4. Model for the regulation of H3K4me3 on the *MCL1* gene.

SETD1A/B are directed to the CpG islands by CXXC1 (CFP1), and the first 5' splice site. SETD1A/B then trimethylates H3K4 at these two regions. H3K4me3 recruits readers such as FACT, PAF, KATs, CHD1 and U2 snRNP (Sims et al. 2007). KDM5A/B/C/D associates with HDAC1/2. U2 snRNP, CHD1, U2AF65 and KAT2B bind to the transcript and form the RNA platform dialogue with the nucleosomes modified by H3K4me3, allowing formation of the pre-mRNA splicing events (Beiberstein et al. 2012; Sims et al. 2007). The figure was reproduced from Davie et al. 2015, with permission from NRC Research Press.

1.2.3 H3K4me3 dynamics

Although H3K4me3 distributions along the genome are stable in both mitotic cells and interphasic cells, H3K4me3 levels increase and decrease with the activation and cessation of transcription (Terrenoire et al. 2010; Edmunds et al. 2008). A previous study determined the distribution and levels of H3K4me3 along the *Fos* gene in epidermal growth factor (EGF) induced 10T1/2 cells at various time points (0, 15, 30, 60 and 120 min) (Edmunds et al. 2008). Although H3K4me3 levels were observed to be associated with transcription, in the time course ChIP assays, H3K4me3 levels did not decline even after transcription of the *Fos* gene had ceased. In *S. cerevisiae*, studies using the *GAL10* gene indicated that after the dissociation of the methyltransferase Set1 and polymerase II, H3K4me3 levels remained high. H3K4me3 levels returned to basal levels after 5 hours, suggesting that the loss of the H3K4me3 levels was not due to demethylation, rather the turnover of histone H3 (Ng et al. 2003). This finding was also supported in a global analysis on *S. cerevisiae* and meiotic progression. Stable global mRNA levels had declined within the hour; however, H3K4me3 levels decreased gradually until 6 hours (Borde et al. 2009). Interestingly, not all genes showed similar H3K4me3 responses when upregulated or downregulated, indicating that H3K4me3 dynamics may differ depending on the meiotic potential of the associated gene (Borde et al. 2009).

1.2.4 H3K4me3 dysregulation and cancer

There are many cancers that have been shown to have aberrant expression of writers and erasers of H3K4me3. These cancers include colorectal cancer, breast cancer, leukemia and prostate

cancer. SETD1A is considered one of the main KMT of H3K4me3, and has been shown to be aberrantly expressed in colorectal cancer and in breast cancer. In many colorectal cancer cell lines, SETD1A was observed to be overexpressed. When SETD1A was silenced, there was a global loss of H3K4me3 in the colorectal cells, and smaller cell colonies were formed which were measured by colony formation assays. This indicated that SETD1A was involved in the metastatic proliferation (Salz et al. 2014). In breast cancer, knockdown of SETD1A causes a reduction in cell invasiveness in breast cancer cell lines, and a reduction in metastasis in nude mice. However, global levels of H3K4me3 were not affected (Salz et al. 2015). In recent years, a rearrangement of the methyltransferase MLL has been associated to the development of leukemia. The rearrangement in the MLL protein leads to a fusion with a gene partner and the loss of the SET-domain H3K4me3 methyltransferase component of MLL. This consequently leads to the methylation of histone H3 lysine 79 rather than H3K4me3, and results in the overexpression of certain MLL-target genes (Liu and Wang 2016).

KDMs have also been implicated in cancer development. In both breast cancer and prostate cancer, KDM5B levels were dysregulated. Levels of KDM5B were determined to be underexpressed in estrogen receptor negative and triple negative breast cancer cells, and overexpressed in estrogen receptor positive breast cancer cells (Klein et al. 2014). Through RNA sequencing on KDM5B knockdown cells, it was determined that KDM5B was involved in both gene activation and gene repression. KDM5B is involved in downregulating immune response and cell proliferation genes, suggesting that low expression of KDM5B in triple negative and estrogen receptor negative breast cancer cells would result in an overexpression of certain genes involved in immune response and cell proliferation (Klein et al. 2014). KDM5A expression levels may also be associated with

significant increase of metastasis in estrogen receptor negative breast cancer, but weakly associated with metastasis in estrogen receptor positive breast cancer (Cao et al. 2014). KDM5B levels were observed to be upregulated in prostate cancer, with metastatic prostate cancers expressing the highest level of KDM5B when compared to benign and prostate cancer (Xiang et al. 2007).

1.3 H3K4me3 and Transcription

Although H3K4me3 is located near the TSS, and levels of H3K4me3 are higher in highly expressed genes, H3K4me3 is not required for transcription (Pavri et al. 2006; Clouaire et al. 2012; Jiang et al. 2011). CFP1 is an important component of SETD1A and SETD1B, the methyltransferases responsible for global H3K4me3 levels. Knockdown of Cfp1 in mouse embryonic stem cells led to a global loss of H3K4me3; however, there was little change in transcription (Clouaire et al. 2012). Similar results were observed in embryonic mouse cells with the depleted Dpy-30, a component of the WRAD complex. Downregulation of Dpy-30 led to global loss of H3K4me3. However, the effect on gene expression was minimal (Jiang et al. 2011). Interestingly, in endogenous retinoic acid receptor beta 2 promoters, H3K4me3 levels increased following transcription activation, suggesting that H3K4me3 may impact gene expression downstream of transcription (Pavri et al. 2006; Edmunds, Mahadevan and Clayton 2008).

1.3.1 H3K4me3 and transcription consistency

Broad distribution of H3K4me3 on key regulatory regions had been previously observed in multiple cell types such as hair follicle stem cells and hematopoietic stem cells. A meta-analysis of H3K4me3 ChIP-seq data was performed to determine the biological significance of the breadth of H3K4me3 levels. The broadest domains of H3K4me3 were identified in genes involved in cell identity and function, which were conserved in both humans and mice. These domains were also associated with transcriptional consistency (Benayoun et al. 2014), as well as region with intense peaks of histone dimethylated histone H3 lysine 79. This post-translational modification corresponds with transcription elongation (Chen et al. 2015). This finding supports the role of H3K4me3 broad domain in transcriptional consistency, as increased transcription elongation and pausing is linked with transcription consistency (Benayoun et al. 2014). Through analysis of divergence patterns in yeast species, H3K4me3 was coupled with essential genes that are periodically expressed. In contrast, histone H3 lysine 9 acetylation was higher in genes that were considered non-essential. This suggested that H3K4me3 may be associated with gene stability, while histone H3 lysine 9 acetylation was linked to gene variability (Mosesson et al. 2014).

1.3.2 H3K4me3 and transcriptional memory

Transcription frequencies were measured between daughter cells and mother cells in *Dictyostelium* cells. In related mother and daughter cells, there was very little differences in transcription frequency, compared to random mother/daughter cell pairs. Transcription frequency was further tested using granddaughter cells, and similar findings were achieved with daughter/granddaughter

cells. When SET1 H3K4 methyltransferase was mutated, the transcriptional frequency between mother/daughter pairs was lost, suggesting that H3K4 methylation was involved in inheritance of the transcriptional state (Muramoto et al. 2010).

Using the *GAL10* gene in *S. cerevisiae*, H3K4me3 levels were determined to persist even after the dissociation of the SET1 and RNA polymerase II. The interpretation of this finding was that the prolonged induced H3K4me3 levels acted as a molecular memory for the recent activation of transcription (Ng et al. 2003). The gradual decrease of H3K4me3 in a global analysis of *S. cerevisiae* study was also observed in a large subset of genes. It was speculated that this may act as an adaptive mechanism to allow for genes to undergo rapid transition back to meiotic growth (Borde et al. 2009). However, since H3K4me3 levels returned to basal level after a few hours, H3K4me3 levels in *S. cerevisiae* were observed to not be inherited across cell generations (Ng et al. 2003). A study in *Arabidopsis* also indicated that H3K4me3 levels may act as a memory mark for transcription. During repetitive stress, elevated levels of transcript was observed, along with higher levels of H3K4me3. In non-stress states the levels of H3K4me3 were lower than those exposed to repetitive stress, indicating that the H3K4me3 mark was retained during the state of stress recovery (Ding, Fromm and Avramova 2012).

1.3.3 H3K4me3 and transcription termination

H3K4me3 is normally associated with recruitment of transcription factors involved in transcription activation. However, a study has shown that H3K4me3 may also be involved in transcription termination. The Nrd1-Nab3-Sen1 pathway in *S. cerevisiae,* is responsible for the termination of

cryptic unstable transcripts and small nucleolar RNAs. Termination that is Nrd1 dependent occur in the proximal promoter region, which coincides with the position of H3K4me3. To test if H3K4me3 levels were associated with Nrd1 termination, the methyltransferase SET1 was deleted. Upon deletion of SET1, termination defects were detected (Terzi et al. 2011). The SET1 complex has many domains. To determine which domain was responsible for the recruitment of Nrd1, multiple SET1 deletion mutants were generated. Those with deletion of the RNA recognition motif domains showed similar phenotypes as the set1 deleted mutants. These domains are responsible for H3K4me3 but not H3K4me or H3K4me2, indicating that H3K4me3 was relevant in efficient Nrd1 regulated termination (Terzi et al. 2011). Although H3K4me3 does not regulate transcription, it is apparent that H3K4me3 has various roles involved in transcription integrity.

1.4 H3K4me3 and Pre-mRNA Splicing

1.4.1 Spliceosome formation

Splicing is the removal of introns within a transcript and the adjoining of exons during maturation of RNA. In mammalian cells, splicing occurs in a stepwise manner and is normally coupled with transcription (Figure 5). The spliceosome is a multi-complex responsible for splicing in mammals, and is composed of snRNPs and auxiliary proteins (Davie et al. 2015; Chen and Manley 2009). The spliceosome consists of five snRNPs (U1, U2, U4, U5 and U6). U1 snRNP initiates the formation of the spliceosome by recognizing 5'splice site. This is followed by the binding of the auxillary factor U2 (U2AF) heterodimer to the 3' terminal AG and the polypyrimidine tract, and the splicing factor 1 (SF1) to the branch point. The binding of these three factors does not require

ATP, and form E complex. The pre-spliceosomal A complex is formed when U2 snRNP replaces SF1 at the branch point, and is ATP-dependent (Chen and Manley 2009). The tri-snRNP (U4/U6-U5) complex is then recruited and interacts with multiple protein factors to forms B complex. C complex is catalytically active and is formed through remodeling of various protein-protein, RNA-protein and RNA-RNA interaction (Chen and Manley 2009; Braunschweig et al. 2013).

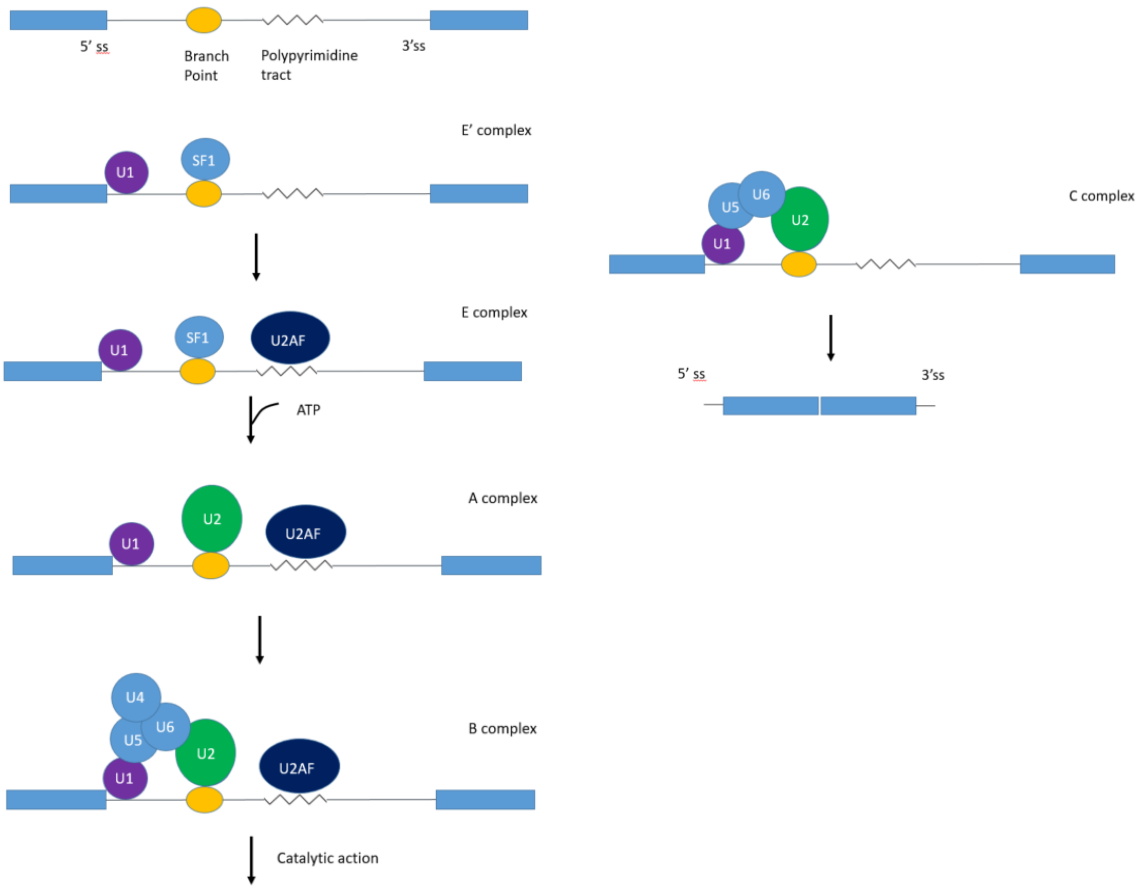


Figure 5. Stepwise progression for the assembly of the catalytic spliceosome complex.

Figure is based on box 1 of Chen and Manley (2009), and Figure 1 from Davie et al. (2015). In E' complex, splicing factor 1 (SF1) binds to the branch point followed by U1 snRNP (U1) binding to the 5' splice site. The E complex is formed when U2 auxillary factor (U2AF) binds to the polypyrimidine tract. With the use of ATP, the A complex is formed when U2 snRNP (U2) replaces SF1. The tri-snRNP complex (U4, U5, U6) is recruited to form B complex. After protein modification, B complex forms the catalytically active C complex, which is capable of splicing introns.

1.4.2 Promoter mediated and chromatin mediated splicing models

Splicing is dependent on the promoter architecture and chromatin changes. There are two models that have been proposed on how promoter architecture influences splicing rate and efficiency (Braunschweig et al. 2014). The "recruitment model" is dependent on the capability of the promoter to recruit splicing factors to the transcription. Promoters that can recruit these factors will facilitate splicing. However, those that fail to recruit the splicing factors, will be included in the final transcript. The "kinetic model" is dependent on how much time the splicing factors have, to associate and recognize each splice site. The rate of elongation of the RNA polymerase II is dependent on the promoter architecture. If the elongation time is longer, splicing factors have more time to associate with a variety of splice sites. However, if the elongation time is short, splicing factors will favor stronger splice sites rather than weaker ones and promote exon exclusion (Braunschweig et al. 2014).

Both the recruitment model and kinetic model can also be mediated by chromatin changes. In the chromatin mediated recruitment model, splicing repressors are recruited by adaptor proteins that associate with the nucleosome at the alternative exon splice site, and promote exon exclusion. In the chromatin mediated kinetic model, transcription elongation rate can be altered through chromatin mediated changes. Hyperacetylation of histone marks have been associated with the relaxing of chromatin structure. This phenomenon increases the accessibility of transcription factors and splicing factors to bind to the DNA, consequently increasing transcription elongation rate and promoter exon exclusion (Gómez Acuña et al. 2012; Khan et al. 2014).

1.4.3 Nucleosome occupancy and pre-mRNA splicing

Multiple studies have shown that nucleosome occupancy is higher in exons than introns (Andersson et al. 2009; Schwartz, Moshorer and Ast 2009; Spies et al. 2009; Tilgner et al. 2009). Through the analysis of compiled genome-wide nucleosome occupancy analyses and nucleosome ChIP-seq, studies have indicated that nucleosome occupancy is associated with splicing (Spies et al. 2009; Tilgner et al. 2009). Exons that were isolated by long introns had significantly higher nucleosome enrichment compared to exons that were more clustered (Spies et al. 2009). Both studies also observed an inverse relationship between nucleosome occupancy and the strength of the splice site. This suggested that nucleosome positioning may play a role in exon retention during alternative splicing (Spies et al. 2009; Tilgner et al. 2009). An inverse relationship between strong splice site was also correlated with the histone post-translational modifications histone H3 lysine 36 trimethylation and histone H3 lysine 27 dimethylation (Spies et al. 2009).

1.4.4 H3K4me3 and pre-mRNA alternative splicing

Studies have indicated that H3K4me3 may play a role in alternative splicing. A study used paired-end RNA sequencing to determine the effects of KDM5 on alternative splicing (He and Kidder 2017). What they observed was that depletion of KDM5 lead to upregulation and downregulation of alternatively spliced exons. In this study, they determined that normally, H3K4me3 levels were high near the TSS and low in the gene body. Conversely, KDM5 levels were shown to be high in the gene body and low near the TSS. When KDM5 was depleted, H3K4me3 levels were low near the TSS, and high in the gene body. This lead to the speculation that a decrease in H3K4me3 levels

at the promoter, and increase H3K4me3 levels in the gene body would cause increase in the release of RNA polymerase II pausing and slower elongation. Based on the integrative aspect of the recruitment and kinetic model of splicing, slower transcription elongation would be associated with exon inclusion, but splicing would also be influenced by the proteins associated with the spliceosome, thus causing a differential pattern of alternative splicing (He and Kidder 2017).

In another study, H3K4me3 ChIP-seq was performed on the colorectal cancer cell line HCT116. They noticed an intense peak of H3K4me3 at the second exon of the *MCL1* gene. *MCL1* gene has two forms; a long form and a short form. The long form is the result of inclusion of exon 2 and is anti-apoptotic, while in the short form, exon 2 is excluded and is pro-apoptotic. In cancer cells, alternative splicing of the *MCL1* gene favored the formation of the long form (Khan et al. 2014). H3K4me3 is known to also recruit acetyltransferases such as KAT7 and KAT2. KAT2 is a component of the SAGA complex, while KAT7 is part of the ING4 complex which has been observed to directly bind to H3K4me3 (Schram et al. 2013; Hung et al. 2009). When HCT116 cells were treated with HDAC inhibitors, increase of the short form of *MCL1* was detected. This finding suggested that recruitment of KAT by H3K4me3 caused hyperacetylation and increased transcription elongation rate, which promoted exon 2 exclusion (Khan et al. 2014; Khan et al. 2016).

1.4.5 H3K4me3 and splicing efficiency

There has been some contradiction on whether pre-mRNA splicing is related to H3K4me3. Some studies have shown that H3K4me3 levels and splicing were related (Sims et al. 2007; Beiberstein

et al. 2012), while another study has shown otherwise (Teoh and Sharrocks, 2014). The methyltransferase component ASH2 was depleted by small interfering RNA, and H3K4me3 levels were measured by subcellular fractionation. Splicing efficiency was measured by determining the change in ratio of non-spliced mRNA to total mRNA. Using the inducible gene *IRF1*, depletion of H3K4me3 lead to a decrease in splicing efficiency of *IRF1* transcript (Sims et al. 2007).

Another study used small interfering RNA to deplete the components of the WRAD complex; RBBP5, WDR5, and ASH2L individually. Using the inducible gene *FOS*, splicing efficiency was measured when RBBP5, WDR5 or ASH2L was depleted. Although a decrease in splicing efficiency was observed in all knock down experiments, depletion of ASH2L did not cause a significant decrease. H3K4me3 levels were also measured in cells with depleted RBBP5. Contrast to other findings, this study showed that depletion of RBBP5 did not alter H3K4me3 levels (Teoh and Sharrocks 2014). However, this study looked only at the levels of H3K4me3 upstream of the TSS. Based on a previous study (Beiberstein et al. 2012), H3K4me3 levels peak at the 5' splice site. It is speculated that if H3K4me3 levels were measured downstream of the TSS, a decrease in H3K4me3 may have been observed, and would correlate with the decrease in splicing efficiency of the *FOS* transcript (Davie et al. 2015).

1.5 Immediate Early Genes

Immediate early genes are protein synthesis independent, and the transcription of these genes occur rapidly when an inducer is added to the serum starved cells, and have transient expression. The mRNA produced have low stability, which is promoted by the protein zinc finger protein 36 (Amit et al. 2007). The zinc finger protein 36 is a RNA binding protein that is involved with transcription

degradation of immediate early genes (Amit et al. 2007).Consequently, the basalexpression levels of immediate early gene transcripts are minimal and are well regulated (Lau and Nathan 1987).

1.5.1 Immediate early gene activation and regulation

Immediate early genes can be activated by two mitogen-activated protein kinase (MAPK) signaling pathways; the RAS MAPK signaling pathway, and the p38 MAPK signaling pathway. The RAS MAPK signaling pathway is activated by growth factors, while the p38 MAPK signaling pathway is activated by stress or ultraviolet light (Healy, Khan and Davie 2013). Growth factors capable of activating the RAS MAPK signaling pathway include epidermal growth factor, fibroblast growth factors, platelet-derived growth factors and nerve growth factors (Kouhara et al. 1997; Yao and Cooper 1995). These growth factors cause the phosphorylation of a cascade of proteins such as the extracellular signal-regulated kinase (ERK) and the mitogen- and stress-activated protein kinase (MSK), which in turn activates the immediate early genes through the activation of regulatory proteins (Healy, Khan and Davie 2013) (Figure 6).

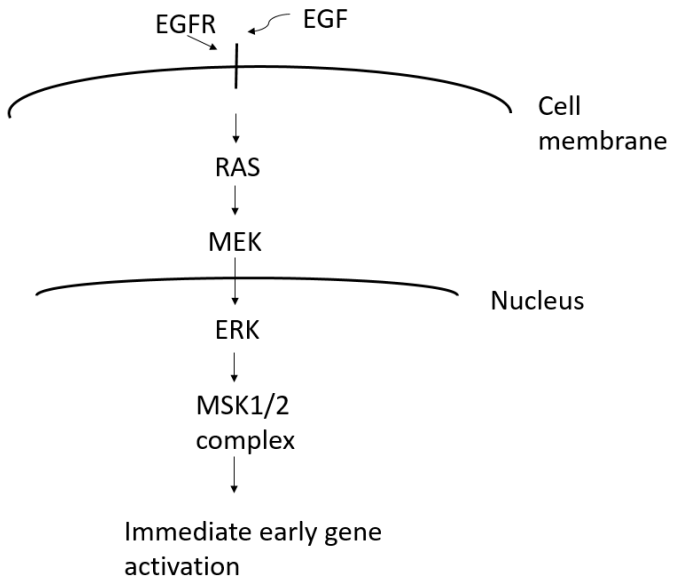


Figure 6. Simple schematic of the RAS MAPK signaling pathway with EGF activation

Schematic indicates the activation of the RAS MAPK signaling pathway with EGF activation through the binding of the EGF receptor (EGFR). This causes a cascade affect, phosphorylating the following proteins (RAS, MEK, ERK and MSK1/2 complex) downstream of the pathway which in turns activate the immediate early genes.

When MSK1/2 are phosphorylated, these enzymes phosphorylate multiple transcription factors responsible for regulating immediate early genes expression such as cAMP response element-binding protein, and nuclear factor kappa-light-chain-enhancer of activated B cells. In addition, phosphorylated MSK1/2 phosphorylates histone H3 serine 10 and H3 serine 28, which recruits the protein 14-3-3. This protein is capable of binding to components that facilitate the open conformation at the promoter region of immediate early genes, promoting transcription (Healy, Khan and Davie 2013).

1.5.2 Immediate early gene *Fos*

The immediate early gene *Fos* expresses the transcription factor Fos and is involved in many biological processes, which are dependent on the stimulus. Some processes that are regulated by the *Fos* gene include apoptosis, cell differentiation, and immune response (Karin, Liu and Zandi 1997). The transcription factor Fos forms a heterodimer with the transcription factor family Jun to form the complex, activator protein 1 (AP-1) (Chiu et al. 1988). AP-1 is formed by the interaction of Jun and Fos through a leucine zipper motif, and the affinity of the complex is dependent on the differential binding of the transcription factors (Scheurmann et al. 1989). AP-1 can also be formed by Jun homodimers (Healy, Khan and Davie 2013). A recent study has indicated that Fos is involved in the regulation of lipid synthesis responsible for membrane biogenesis, independent of AP-1 activity. In this process, Fos interacts with phosphatidylinositol-4-monophosphate 5-kinase to promote phosphatidylinositol 4,5-bisphosphate formation (Ferrero et al. 2014).

The regulation of transcriptional activity of the FOS transcription factor can occur at two levels; the activation and the expression of FOS (Healy, Khan and Davie 2013). *Fos* activation is regulated by the temporal activation of the RAS MAPK and p38 MAPK signaling pathways, that can activate regulatory proteins. FOS is also capable of autorepression by binding to the regulatory region, serum response element at the upstream promoter region. However, it is only functional when in heterodimer form (Lucibello et al. 1989). The stability of *Fos* expression can also be regulated by the RAS MAPK signaling pathway. ERK1/2 can phosphorylate multiple region of the C-terminal transactivation domain of the FOS protein and increase expression stability. It has

been suggested that this phosphorylation increases DNA binding affinity and is responsible for the translocation of the Fos protein from the cytoplasm to the nucleus (Monje et al. 2003).

1.5.3 *Fos* dysregulation in cancer

The RAS MAPK signaling pathway is often dysregulated in cancer, consequently leading to the transcription factor FOS to be constitutively activated. Overexpression of the *Fos* has been associated with neoplastic transformation, increased anchorage-independent growth and increased invasiveness (Healy, Khan and Davie 2013). Deletion or downregulation of *Fos* expression through mouse models or inhibition regulatory proteins protected cells from neoplastic transformation, and have been observed in cancer cells from liver, breast and colorectal cancer (Castellaro et al. 2015; Peng et al. 2016; Bakiri et al. 2017). Knockout of *Fos* in fibroblast lead to a decrease in anchorage-independent growth and invasiveness. In addition, changes in cell morphology was also observed (Hu et al. 1994). These observations were further validated using human fibroblast cells transformed with *v-Fos*. This transformation resulted in increased invasiveness. However, no change in anchorage dependence was observed (Scott et al. 2004). Interestingly, although ectopic expression of the *Fos* gene is associated with tumorigenesis, mutation of the gene itself is rare in cancer (Healy, Khan and Davie 2013).

2. RATIONALE, HYPOTHESIS AND PROJECT AIMS

2.1 Rationale

H3K4me3 is a mark of transcriptionally active genes, that is involved in the recruitment of many non-histone chromosomal proteins (e.g. CHD1, KAT, and HDAC). Thus, H3K4me3 is a post-translational modification that is involved in gene regulation and organization (Davie et al. 2015). In certain cancers, such as breast cancer, colorectal and prostate cancer, aberrant expression and/or function of SETD1A, KDM5B, and/or H3K4me3 have been indicated (Cao et al. 2014; Klein et al. 2014; Salz et al. 2014; Salz et al. 2015; Xiang et al. 2007). Although there is evidence suggesting that aberrant expression of proteins involved in regulating H3K4me3 levels may contribute to cancer, very little is known about the dynamics of H3K4me3 itself in mammalian cells during transcriptional induction. One paper reported that H3K4me3 levels increased in response to the transcription of immediate early genes, *Fos* and *Jun* (Edmunds, Mahadevan and Clayton 2008). However, when transcription ceased, H3K4me3 levels remained high. Immediate early genes can be induced through the activation of the RAS MAPK signaling pathway or by the stress activated pathway, p38 (Healy et al. 2012; Healy, Khan and Davie 2013). The RAS MAPK signaling pathway can be induced with the binding of EGF to the EGF receptor. This binding initiates the signaling pathway and the phosphorylation of the final kinase in the pathway, ERK (Healy et al. 2012). Phosphorylated ERK (Phospho-ERK) can activate the kinase MSK which is involved in the stimulation of immediate early genes through phosphorylation of H3 at serine 10 and serine 28 (Healy et al. 2012; Healy, Khan and Davie 2013). To date, when and if H3K4me3 levels return to basal level after induction of immediate early genes have not been reported in

mammalian cells. Studies done in yeast have observed that H3K4me3 levels return to basal level hours after cessation of transcription, suggesting that the decrease in H3K4me3 levels are most likely due to histone H3 turnover.

H3K4me3 is a post-translational modification that is associated with active transcription and splicing. Depletion of subunits of the WRAD complex has been observed to impair splicing efficiency (Teoh and Sharrocks 2014; Sims et al. 2007). Inhibition of splicing also influenced H3K4me3 levels at the 5' splice site, and caused a decrease in H3K4me3 levels in *FOS* constructs (Beiberstein et al. 2012). Little is known about the mechanism of how splicing inhibition affects H3K4me3. It is speculated that splicing inhibitors may prevent the binding of SETD1A/1B to the RNA (Davie et al. 2015). To date, the effects of splicing inhibition on H3K4me3 levels have not been determined during endogenous expression of the *Fos* gene. The effects of splicing inhibition on H3K4me3 dynamics has also not been studied to our knowledge.

For the study being proposed, the cell line that will be used is the mouse fibroblast cell line 10T1/2, and the induction of the immediate early gene *Fos* will be used as the gene model. This model was chosen as the induction of the *Fos* gene has been previously done in 10T1/2 cells (Edmunds, Mahadevan and Clayton 2008). To be consistent with previously published data, the same model was used.

2.2 Hypothesis

Transcription and pre-mRNA splicing at the 5' end of first exon of the *Fos* gene play a role in H3K4me3 dynamics.

2.3 Project Aims

2.3.1 Aim 1: Determine the dynamics of H3K4me3 levels in EGF-induced, serum-starved 10T1/2 cells

Whether transcriptionally increased H3K4me3 levels return to basal values after transcription has ceased has not been determined in mammalian cells. Thus, expression levels of the inducible gene *Fos* after treatment of EGF was monitored in the immortal mouse fibroblast cell line 10T1/2. Serum starved cells will be treated with EGF (50ng/ml) for the time points 0, 15, 30 and 60 min, as a previous study indicated that this concentration was sufficient to induce immediate early genes (Edmunds, Mahadevan and Clayton 2008). Expression levels of both *Fos* and *Jun* will be monitored through RT-qPCR, using serum starved 10T1/2 treated with EGF at the time points mentioned above. ChIP assays will be used to determine the changes in H3K4me3 levels in cross-linked, serum starved 10T1/2 treated with EGF at the time points previously mentioned. This will then be followed by ChIP-seq experiments to determine the genomic distribution and location of H3K4me3 levels in the quiescent 10T1/2 cells induced with EGF at selected time points. The data will then be analyzed with the help of Dr. Wayne Xu.

2.3.2 Aim 2: Determine if the inhibition of splicing interferes with the dynamics of H3K4me3 in immediate early gene *Fos*

Inhibition of splicing has been shown to decrease H3K4me3 levels at the 5' splice site of the *Fos* gene in transgenic cell lines (Bieberstein et al. 2012). However, the effects of splicing inhibition on H3Kme3 has not been tested at the 5' splice site of endogenous *Fos* gene, nor the effects of splicing inhibition on H3K4me3 dynamics during transcription induction. Serum starved 10T1/2 cells will be incubated with isoginkgetin, a splicing inhibitor that prevents the formation of the spliceosome B complex by inhibiting the stable recruitment of the tri-snRNP complex (O'Brien et al. 2008). Reverse transcriptase PCRs will be done to determine the appropriate dose of isoginkgetin and incubation time needed to inhibit splicing. Once this is determined, ChIP assays for H3K4me3 will be done with cross-linked, serum starved 10T1/2 cells treated with isoginkgetin and EGF at the time points 0 and 30 min. Assays will be done to determine the impact of splicing inhibition on H3K4me3 during induction along the 5' splice site of the *Fos* gene. ChIP assays for H3K4me3 will also be done with serum starved 10T1/2 cells treated with transcription inhibitor DRB (5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole) (25 ug/ml for 10 min) and stimulated with EGF at the time points previously stated. These experiments will determine whether inhibition of splicing would have similar effects on H3K4me3 dynamic as inhibition of transcription.

3. MATERIALS AND METHODS

3.1 Tissue Culture

3.1.1 Cell Culture and Treatments

Immortalized mouse fibroblast 10T1/2 were purchased from American Type Culture Collection (ATCC) and grown in complete alpha minimum essential media (MEM) (Invitrogen) supplemented with 10% fetal bovine serum (FBS) (Gibco) and 1% antibiotic-antimycotic (Invitrogen). The RAS MAPK signaling pathway was induced by serum starving 10T1/2 cells (90-100% confluent) for 24 hr in 0.1% FBS and 1% antibiotic-antimycotic, followed by epidermal growth factor EGF (50 ng/ml, Gibco) treatment at various time points (0, 15, 30 or 60 min). For splicing inhibition studies, serum-starved 10T1/2 cells were treated with isoginkgetin (30 uM, ChemFaces) for 3 hr prior to EGF treatment. To test the effects of transcription on the dynamics of H3K4me3 the transcriptional inhibitor DRB (25 ug/ml, Sigma) was added for 10 min before EGF treatment.

3.1.2 Cell Passaging

Once cells had reached a confluency of 60-70%, the media was aspirated from the plate and washed with 1xphosphate-buffer saline (PBS) (wash buffer) (137 mM NaCl, 2.7mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4). The wash buffer was aspirated, followed by the addition of 0.5 – 3 ml of TripLE Express enzyme (Gibco) with incubation at 37^oC in 5% CO_2 for 5 min. Cell detachment

was monitored using an Olympus CK40 microscope and once sufficient number of cells had detached, 5-10 ml of media was added to the cell culture plate to neutralize the trypsin reaction. The solution was transferred to a 15 ml centrifuge tube (FroggaBio) and centrifuged at 300xg for 3 min at room temperature. The supernatant was aspirated and the cell pellet was resuspended in an appropriate volume of media. The resuspended cells were seeded onto new cell culture plates.

3.1.3 Cell Storage and Revival

For cell storage, cells were grown to a confluency of 80-90%. The cells were washed with 1xPBS and 1-3 ml of TripLE Express trypsin (Gibco) was added to the cells and incubated at 37^{o}C at 5% CO_2 for 5 min. Once the cells had detached from the plate, the appropriate volume of media was added and the solution was transferred to a 15 ml centrifuge tube and centrifuged at 300xg for 3 min. The supernatant was aspirated, and the cell pellet was resuspended in a solution of 95% FBS (Gibco) and 5% dimethyl sulfoxide (DMSO) (Sigma). The resuspended cells were transferred to a cryovial (Thermo Scientific) and stored at -80^{o}C. For long term storage, cells were transferred to liquid nitrogen.

To revive cells from cryovials, the cryovials were placed in 37^{o}C water and allowed to thaw. Once thawed, the solution of 95% FBS (Gibco), 5% DMSO (Sigma) and cells were transferred into a 15 ml centrifuge tube with 9 ml of complete media. The 15 ml tube was then centrifuged at 300xg for 3 min. The supernatant was aspirated, and the cell pellet was resuspended in an appropriate volume of media and seeded onto a 60 mM cell culture plate. After 24 hr of seeding the cells, the old media was aspirated and fresh media was added to the cell culture plate. Once the cells had reached a

confluency of 60-70%, the cells were passaged and transferred to a 100 mm cell culture plate. Once cells were restored, they were passaged at least three times before they were used for experimental analyses.

3.2 Ethanol Fixation and Flow Cytometry Preparation

Media was aspirated and the cells (90-100% confluent) were washed once with 1xPBS. An appropriate amount of 1xPBS was then added before the cells were carefully scraped off the cell culture plate with a cell scraper (Fisher Scientific). The scraped cells in 1xPBS were collected in a 15 ml centrifuge tube (FroggaBio) and centrifuged at 300xg for 3 min. The supernatant was aspirated, and the cell pellet was resuspended in 2 ml of 1xPBS. The resuspended cell solution was then added dropwise into 4 ml of 95% ethanol while the solution was continuously mixed at a medium speed. Once all the cells were added to the ethanol, the cells were fixed and stored at 4^{o}C. For long storage, the fixed cells were stored at -20^{O}C.

To prepare cells for flow cytometry the fixed cells were centrifuged at 500xg for 5 min at 4^{o}C. The supernatant was aspirated, and the cell pellet was resuspended in a 2 ml solution of PBS with ribonuclease (RNase) A (with a final concentration of 5 ug/ml) and incubated for 30 min at 37^{o}C. The cells were then stained with 1 ul of a 10 ug/ul stock of ethidium bromide (Invitrogen). The cell solution was then mixed and wrapped in aluminum foil. Flow cytometry and analysis was done by the Flow Cytometry Core Facility at the University of Manitoba.

3.3 SDS PAGE and Western Blot

Nuclear extract (preparation will be discussed in section 3.6.4) was dissolved in reducing sodium dodecyl sulfate (SDS) loading buffer (187.5 mM Tris-HCl pH 6.8, 10% glycerol, 3% SDS, 0.2% bromophenol blue and 5% β-mercaptoethanol, with each percentage based on volume) which reduced and denatured the extract. The extract was then boiled for 10 min and resolved on a 15% polyacrylamide gel, using the apparatus dimension of 1.5mm. The SDS polyacrylamide gel electrophoresis (PAGE) was run at 150V for 45-60 min or until the SDS loading dye had ran to the bottom of the gel. The proteins on the polyacrylamide gel were then transferred on to a 0.45 uM nitrocellulose membrane (Biorad) at 100V for 75 min at 4^{o}C. The membrane was stained with Ponceau S (0.5% Ponceau S and 5% acetic acid) for 1-2 min to measure the transfer efficiency and then rinsed with ddH20 to remove the excess stain. Following rinsing the membrane was blocked with 5% bovine serum albumin (BSA) powder in 0.1% tween in Tris-buffer saline (TBST) (for anti-ERK and anti-phospho-ERK) or 5% skim powder in 0.1% TBST (for anti-H3K4me3) for 1h at room temperature and incubated with the primary antibody anti-ERK (Cell Signaling), anti-phospho-ERK (Cell Signaling) or anti-H3K4me3 (Abcam) overnight at 4^{o}C (Table 1). The next day the membrane was washed 3 times in 0.1% TBST (10 min/wash) and incubated with the horseradish peroxidase-conjugated anti-rabbit IgG secondary antibody (Bio-Rad) at a dilution of 1/5000 in 5% BSA in 0.1% TBST or 5% skim milk in 0.1% TBST for 1 hr at room temperature. The membrane was washed 3 more times with 1xTBST (10 min/wash) before ERK, phospho-ERK or H3K4me3 was detected on the membrane using Lighting Plus-ECL (Perkin Elmer) on Hyperfilm ECL (Amersham) according to the manufacturer's instructions.

Table 1. List of antibodies used for SDS PAGE and ChIP experiments

Antibody Name	Catalog number	Type	Company	Antibody amount used
p44/42MAPK (Erk1/2)	9102	Rabbit Polyclonal	Cell Signaling	1/1000
Phospho-p44/42 MAPK(Erk1/2) (Thr202/Tyr204) (D13.14.4E)	4370	Rabbit Monoclonal	Cell Signaling	1/2000
Anti-Histone H3 (tri methyl K4)	ab8580	Rabbit Polyclonal	Abcam	1/1000
Normal Rabbit IgG	12-370	Rabbit Polyclonal	Millipore	1ul/A260
Goat Anti-Rabbit IgG (H+L)-HRP Conjugate	1706515		Biorad	1/5000

3.4 RNA Isolation and RT-qPCR

Total RNA was extracted from 10T1/2 cells using the RNeasy Plus Mini Kit (Qiagen) as instructed by the manufacturer and eluted in 40 ul nuclease free water (Fisher Scientific). The purity and concentration of RNA was assessed by the A_{260}/A_{280} and A_{260}/A_{230} ratios and measured by Nanodrop 2000 (ThermoFisher Scientific). The isolated RNA was treated with DNase (Promega) using the instructions provided by the manufacturer. First-strand cDNA synthesis was performed using the iScript Reverse Transcription Supermix (Biorad) according to manufacturer instructions. RT-PCRs were performed for *Gapdh*, *Fos* and *Jun.* RT-PCR products for *Fos* were resolved on a 1.5% agarose gel for the splicing inhibition studies.

Table 2. List of Primers used for gene expression studies using RT-qPCR

Primer	Sequences	Amplicon size
Fos (2605) Forward *Fos* (2605) Reverse	5'- GAG CTG GTG CAT TAC AGA GA -3' 5'- AAC AGG TAA GGT CCT CCC TA – 3'	77 bp
Jun (1115) Forward *Jun* (1115) Reverse	5' – CCA AGA ACT CGG ACC TTC TCA – 3' 5' – GGT GAT GTG CCC ATT GCT – 3'	104 bp
Gapdh Forward *Gapdh* Reverse	5' – TTC CGT GTT CCT ACC CCC AAT GTG T – 3' 5' – GGA GTT GCT GTT GAA GTC GCA GGA G – 3'	174 bp

Table 3. List of Primers used for splicing studies using RT-qPCR

Primer	Sequence	Amplicon size
Fos Exon junction 3/4 (1749) Forward	5' – CAG CCA AGT GCC GGA AT – 3'	204bp (unspliced)
Fos Exon junction 3/4 (1749) Reverse	5' – GCA ACG CAG ACT TCT CAT CT – 3'	78 bp (spliced)

3.5 Agarose Gel preparation

Agarose gels of 1.5% were prepared by using the appropriate weight of agarose (Invitrogen) which is dissolved in 1x TBE buffer (100 mM Tris base, 100 mM boric acid and 2 mM EDTA). The solution was heated in the microwave for 1-2 min and left to cool before the addition of 1xGelStar nucleic acid stain (Lonza) to the liquid gel. The gel was left to solidify, and placed in 1xTBE for the agarose gel electrophoresis.

3.6 ChIP, ChIP Optimizations and ChIP Sequencing

3.6.1 Cross-linking Conditions

Cycling 10T1/2 cells (90-100% confluent) were washed twice with 1xPBS and cross-linked with 1% formaldehyde (Fisher Scientific) for 10 mins at room temperature on rotation. The cross-linking reaction was stopped by adding 1.25 mM glycine to a final concentration of 125 uM glycine and incubating for 5 min at room temperature with rotation. The cells were washed 2 more times with 1xPBS and were collected in 1xPBS, transferred to a 15 ml falcon tube (FroggerBio) and centrifuged at 300xg for 3 min. The cell pellet (4x140 mm plates of 10T1/2 cells) was stored at -80^{o}C until used for experimental analysis.

3.6.2 Micrococcal Nuclease Digestion

Cross-linked cell pellets of cycling 10T1/2 cells were resuspended in cell lysis buffer (5 mM PIPES pH 8, 85 mM KCl and 0.5% NP-40) supplemented with protease and phosphatase inhibitors (Roche) for 10 min at 4^{o}C with rotation, and was centrifuged at 2000xg for 5 min at 4^{o}C. This step was repeated one more time to release the nuclei. The nuclear pellet was resuspended in 0.5-1 ml of MNase digestion buffer (10 mM Tris-HCl pH 7.5, 0.25 M sucrose and 75 mM NaCl) supplemented with protease and phosphatase inhibitors (Roche), and the A_{260} was measured. $CaCl_2$ was added to the solutions to a final concentration of 3 mM and were incubated at 37^{o}C for 10 min with rotation. Immediately after, 2.5 U of MNase (Worthington Biochemical Corporation) is added per A_{260} of DNA A_{260} was and incubated at 37^{o}C with rotation. Aliquots of 100-200 ul were

collected from the MNase treated nuclear lysate at the various time points (5, 10, 20, 30 and 40 min), and EDTA pH 8 to a final concentration of 5 mM was added to each aliquot to stop the MNase reaction. Nuclei were lysed with 10% SDS to a final concentration of 0.5% SDS and incubated at room temperature for 1 hr. The lysate was purified, and the DNA fragments were resolved on a 1.5% agarose gel, or run on the bioanalyzer. The bioanalyzer was ran by Aleksander Ilic. The MNase treated lysates were added to a well which contained a sieving polymer matrix with fluorescent dye. Each well was connected to a 16-pin electrode which drove the voltage gradient. A ladder was ran in an adjacent well for reference in data analysis.

3.6.3 Dot Blot Assay

Two microliters of the unmodified H3 (Abcam; N-terminus 1-16), H3K4me (Abcam), H3K4me2 (Abcam) and H3K4me3 (Abcam) peptides (0.25 ug/ul) were spotted on a 0.45 uM nitrocellulose membrane (Biorad) and baked at 65^{o}C for 15 min. The membrane was blocked with 5% skim milk powder in 0.1% TBST (50 mM Tris-HCl pH 7.6, 150 mM NaCl and 0.1% Tween-20) for 1 hr at room temperature on the (VWR, Model 200), followed by an overnight incubation at 4^{o}C of the primary antibody anti-H3K4me3 (Abcam) at a dilution of 1/1000 in 5% skim milk in 0.1% TBST on a rotator (Boekel Scientific, Model 260200). The membrane was washed 3 times with 0.1% TBST (10 min/wash) at room temperature on the rotating platform before incubation with the horseradish peroxidase-conjugated anti-rabbit IgG secondary antibody (Biorad) at a dilution of 1/5000 in 5% skim milk in 0.1% TBST for 1 hr at room temperature. The membrane was washed 3 more times with 1xTBST (10 min/wash) before H3K4me3 was detected on the membrane using

Lighting Plus-ECL (Perkin Elmer) on Hyperfilm ECL (Amersham) according to the manufacturer's instructions.

3.6.4 ChIP and testing immunoprecipitation efficiency

Cross-linked cell pellets were resuspended in cell lysis buffer (5 mM PIPES pH 8, 85 mM KCl and 0.5% NP-40) supplemented with protease and phosphatase inhibitors for 10 min at 4^{o}C with rotation, and was centrifuged at 2000xg for 5 min at 4^{o}C. This step was repeated one more time to release the nuclei. The nuclear pellet was resuspended in 0.5-1 ml of MNase digestion buffer (10 mM Tris-HCl pH 7.5, 0.25 M sucrose and 75 mM NaCl) supplemented with protease and phosphatase inhibitors, and the A_{260} was measured using spectrophotometer. $CaCl_2$ was added to the solutions to a final concentration of 3 mM and were incubated at 37^{o}C for 10 min with rotation. Immediately after, MNase (Worthington Biochemical Corporation) at a final concentration of 2.5U/A260 was added, and the nuclei were incubated at 37^{o}C for 20 min with rotation. The MNase reaction was stopped with the addition of EDTA pH 8 to a final concentration of 5 mM. Nuclei were lysed with 10% SDS to a final concentration of 0.5% SDS and incubated at room temperature for 1 hr. The lysate was diluted 5 fold in RIPA buffer (10 mM Tris-HCl pH 8, 1% Triton-X-100, 0.1% SDS and 0.1% SDC) supplemented with protease and phosphatase inhibitors (Roche). The nuclear lysate was further lysed by passing it through a 22g needle 10-20 times. The lysed nuclei were centrifuged at 2000xg for 5 min and the supernatant was collected. The supernatant was then pre-cleared using 60 ul of protein A/G Plus agarose beads (Santa Cruz) for every 1 ml of lysate and incubated for 1h at 4^{o}C with rotation. The beads were pelleted down using pulse centrifugation and the A_{260} was measured once again. Twelve A_{260} of the lysate was aliquoted and 1 ug/A_{260} of

the antibody H3K4me3 (Abcam) or normal rabbit pre-immune IgG (Millipore) was added into the lysate and incubated overnight at 4^{o}C with rotation. Some lysate (1.2A_{260}) was also saved for input. The next day, 7 ul/A_{260} of Dynabeads Protein G (Invitrogen) were added to the lysate and incubated for 2h at 4^{o}C with rotations. The beads were washed twice with rotation with 1 ml of each of the following wash buffers: low salt wash buffer (0.1% SDS, 1% Triton-X-100, 2mM EDTA, 20 mM Tris-pH 8.1 and 150 mM NaCl), high salt wash buffer (0.1% SDS, 1% Triton-X-100, 2 mM EDTA, 20 mM Tris-pH 8.1 and 500 mM NaCl), LiCl wash buffer (250mM LiCl, 1% NP-40, 1% deoxycholate, 1 mM EDTA and 10 mM Tris-HCl pH 8.1) and 1xTE buffer (10 mM Tris-HCl pH 7.5 and 1 mM EDTA).

To test the immunoprecipitation efficiency, 5A_{260} of the lysate is collected instead of 12 A_{260} and 1 ug/ A_{260} or 1.5 ug/A_{260} of anti-H3K4me3 antibody was added. An aliquot of 0.5 A_{260} is saved for input. Another aliquot of 0.5 A_{260} is taken after the lysate has been incubated with the magnetic beads, and is the immunodepleted fraction. After the washes, the antibody bound chromatin, input and immunodepleted fraction was reduced in SDS loading buffer and a SDS PAGE and western blot was run as described previously.

For ChIP assays, the antibody bound chromatin was eluted from the magnetics beads using 200-250 ul of elution buffer (1% SDS and 100 mM $NaHCO_3$) and incubated at room temperature with rotation for 30 min. The antibody bound chromatins along with the input samples were put in a 65^{o}C water bath overnight to reverse DNA-protein crosslinking. RNase A (final concentration of 0.02 ug/ml) was added to each sample the next day, and incubated at 37^{o}C for 30 min with rotation.

Proteinase K (final concentration of 0.5ug/ml) treatment was added after and incubated for 1 hr at 55^{o}C. The Input and ChIP DNA were purified using the QIAquick PCR Purification Kit (Qiagen).

3.6.5 ChIP-sequencing

Input and ChIP DNA were quantified using the Qubit 2.0 Fluorometer (Life Technologies), and the size and quality of the DNA was analyzed using the Bioanalyzer (Agilent Technologies). Around 20 ng of input and ChIP DNA was used to prepare libraries for sequencing. ChIP-seq library preparation was done by Aleksandar Ilic and was prepared according to the manufacturer instructions of the NEB Ultra II DNA library prep kit. The DNA ends were repaired and dA tailing was done. This was followed by adaptor ligation and DNA size selection. The DNA libraries were amplified using the Techne ^{3}Prime thermocycler (Fisher Scientific) and samples were validated using qPCR. After the libraries were validated, samples were sent to Genome Quebec for sequencing using the Illumina Hiseq 4000 platform and was done 100 based paired end sequencing. The data was analyzed by Dr. Wayne Xu (bioinformatician).

3.7 DNA purification and qPCR

Input and ChIP DNA was purified using the QIAquick PCR purification kit (Qiagen) as instructed by the manufacturer and eluted with 30 ul nuclease free water (Fisher Scientific). The concentrations were measured by the Qubit 2.0 Fluorometer (Life Technologies). The size and quality of the input and ChIP DNA were analyzed using the high sensitivity DNA kit and 2000 Bioanalyzer (Agilent Technologies) which was run by Aleksandar Ilic. H3K4me3 levels at specific

locations along the *Fos* genes using amplicons were quantified using qPCR, which were performed with the CFX96 Real-Time PCR Detection System (Biorad). Each qPCR reaction was prepared in a 0.2 ml PCR tube (Sarstedt) according to the SYBR® Green Master Mix (Biorad) manufacturer. Equal amounts of 1.0 ng of ChIP or input DNA was added to each reaction tube to prepare a total volume of 20 ul, whereas the negative control with the use of the Normal Rabbit IgG immunoprecipitated DNA was added based on equal volume to ChIP DNA. The volume of ChIP DNA was determined based on the volume required to obtain 1.0 ng of ChIP DNA.

Table 4. List of Primers used for ChIP experiments using qPCR

Primer	Sequences	Amplicon size
Fos exon 1 (164) Forward *Fos* exon 1(164) Reverse	5'-ACT ACG AGG CGT CAT CCT-3' 5'- CCT GTG TGT TGA CAG GAG AG-3'	117 bp
Fos intron 1(749) Forward *Fos* intron 1(749) Reverse	5'-CGC AGA CGT CAG GGA TAT TT-3' 5'-AGC CTT TCT CCT CCA TCC TA-3'	95 bp
Fos intron 1/exon 2(1001) Forward *Fos* intron 1/exon 2 (1001) Reverse	5'- CAC CTC CGC TTT CCT CTT T-3' 5'- TCT GGG CTG GTG GAG AT-3'	107 bp
Fos exon 2/intron 2 (1222) Forward *Fos* exon 2/intron 2(1222) Reverse	5'- AAT GGT GAA GAC CGT GTC AG-3' 5'- GCC CAC AAA GGT CCA GAA T-3'	93 bp
Fos exon 3 (1700) Forward *Fos* exon 3 (1700) Reverse	5'-GAA GAG GAA GAG AAA CGG AGA AT-3' 5'-CTT GGA GTG TAT CTG TCA GCT C-3'	97 bp
Fos exon 4 (2550) Forward *Fos* exon 4 (2550) Reverse	5'-CCC TGT GAG CAG TCA GAG A-3' 5'-GGG AAC CTT CGA GGG AAG A-3'	105 bp
Fos negative (3370) Forward *Fos* negative (3370) Reverse	5'-GCG ACC ACC TTC TTG CT-3' 5'-ATC TTG ACA AAC TGG TCT CCT C-3'	79 bp

4. RESULTS

4.1 Optimizing conditions for the induction of immediate early genes

To accurately study the induction of immediate early genes, cells must be synchronized in the G_0/G_1 phase. Flow cytometry was performed by the Flow Cytometry Core at the University of Manitoba with ethanol fixed cycling 10T1/2 and 24 hr serum starved 10T1/2 cells. Cells were considered synchronized after 24 hr of serum starvation as >90% of the cell population was arrested in the G_0/G_1 phase (Figure 7). Due to the default setting of the flow cytometry program, 10T1/2 cells are stated as being diploid. However, these cells are aneuploid (2N= 81).

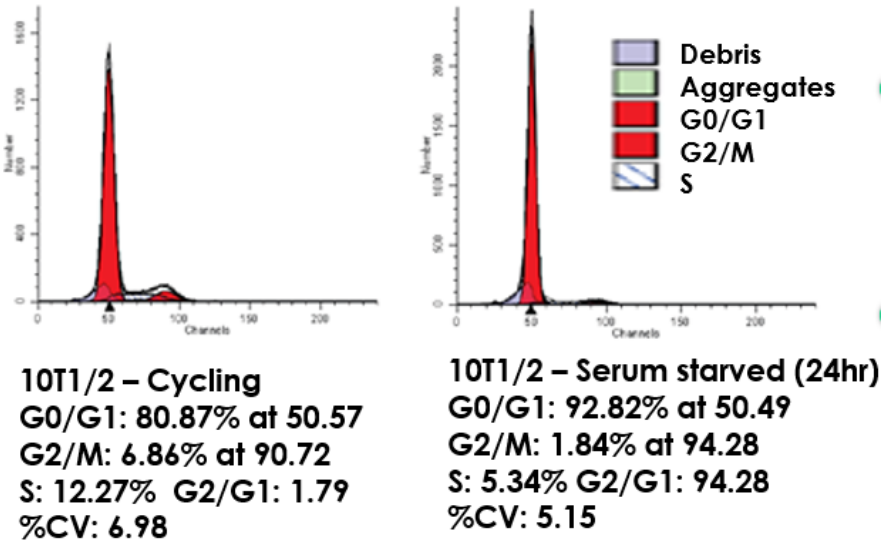


Figure 7. Flow cytometry cell cycle analyses on ethanol fixed cycling 10T1/2 and 24 hr serum-starved 10T1/2.

Cells serum starved for 24 hr showed more cells arrested in the G1 phase with less cell undergoing replication in the S phase compared to the cycling 10T1/2 cells. Analysis was done in the core Flow Cytometry facility at the University of Manitoba. The first peak at channel 50 represents the cells in G0/G1 phase. The second peak at channel 100 represent the cell population in G2/M.

To induce the RAS MAPK signaling pathway, serum starved 10T1/2 cells were induced with EGF (50 ng/ml) at various time points (0, 15, 30 or 60 min). Cellular extracted protein was resolved by SDS-PAGE and analyzed using western blots with the anti-ERK or anti-phospho-ERK antibody. Levels of phospho-ERK were measured to monitor the activation of the RAS MAPK signaling pathway, and ERK levels were used as a control. The RAS MAPK signaling pathway was most active after EGF stimulation for 15 min, and began to decline at 30 min (Figure 8).

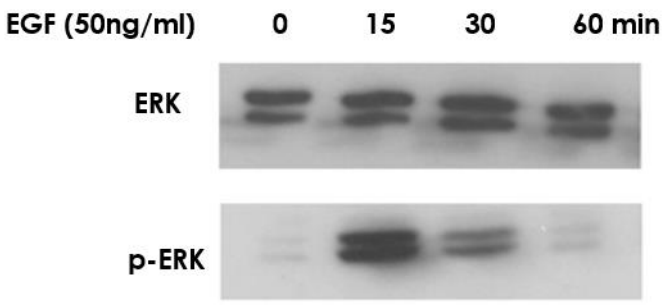


Figure 8. Levels of phospho-ERK increased after 15 min of EGF stimulation in quiescent 10T1/2 mouse cells.

Western blot analysis was done on EGF (50 ng/ml) induced nuclear extracted proteins at the various time points of 0, 15, 30 and 60 min, which were immuno-stained with phospho-ERK and ERK antibodies. Equal volume of input nuclear extract (0.5 A_{260}) was loaded to each lane. ERK levels were used as a loading control.

After the successful induction of the RAS MAPK signaling pathway, RT-qPCRs were done to measure the gene expression of the immediate early genes *Fos* and *Jun* after EGF stimulation. Previous RT-qPCR studies demonstrated that *Fos* and *Jun* were activated within the first hour of induction (Edmunds, Mahadevan and Clayton 2008). The results produced in our study showed similar trends in *Fos* and *Jun* activation as previously published (Edmunds, Mahadevan and Clayton 2008). Both genes were induced within the first hour of EGF induction, and the expressions peaked at 30 min post-EGF treatment before declining at 60 min (Figure 9).

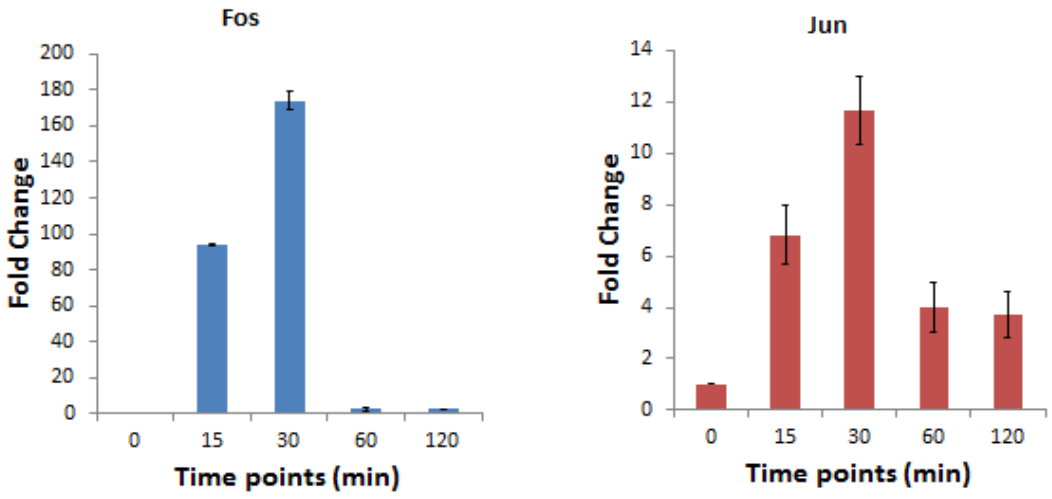


Figure 9. RT-qPCRs for the *Fos* and *Jun* gene at various EGF induction time points (0, 15, 30 and 60 min).

Fold change was normalized to *Gapdh* and the average fold change was calculated and plotted using three independent experiments. Error bar represents the standard error of the means.

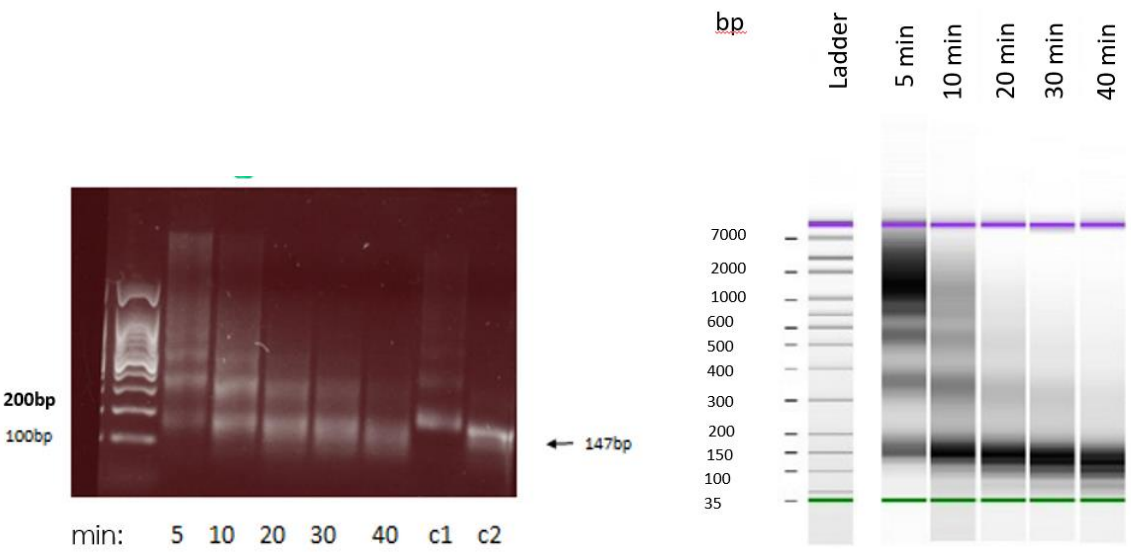
4.2 H3K4me3 ChIP optimizations

Prior to performing ChIP experiments, multiple variables and conditions had to be optimized. These experiments included optimizing MNase conditions, validating the antibody specificity, and the immunoprecipitation efficiency.

Since H3K4me3 is a histone post-translational modification that is located on the nucleosome, to provide a better resolution for the ChIP assays, MNase was used instead of sonication for DNA fragmentation. The digestion time for MNase differs from cell line to cell line. Thus, MNase digestion for 10T1/2 cells had to be optimized. Aliquots of nuclear lysate were collected at various

time (5, 10, 20, 30 and 40 min) points during the MNase digestion time course. The DNA was purified and resolved on a 1% agarose gel and ran on the bioanalyzer. The time point for MNase incubation was determined to be 20 min for 10T1/2 cells as the DNA was predominantly mononucleosome length (146 bp) at the 20 min (Figure 10), which was indicated by the results provided by the bioanalyzer.

A) B)



C)

Time point (min)	DNA fragment size (bp)	Percent (%)
5	170	20
10	155	35
20	149	66
30	141	73
40	132	78

Figure 10. MNase digested 10T1/2 DNA is predominantly mononucleosomes after 20 min.

A) Purified DNA at various MNase incubation times (5, 10, 20, 30 and 40 min) was loaded onto a 1.5% agarose gel. The size of a mononucleosome DNA detected was 147 bp. Mononucleosomes from chicken erythrocytes were used as reference (c1 and c2). B) Bioanalyzer results using the same purified DNA at various MNase incubation times. The bioanalyzer was run by Dr. Deborah Tsuyuki. C) Table of the percentage of the most predominant DNA fragment size for the various MNase incubation time points.

To test the antibody specificity, dot blot using different epitope peptides were done. Because the post-translational modification of interest is H3K4me3, antibodies were tested against the monomethylated, dimethylated and trimethylated forms of H3K4. The unmodified H3 peptide was also tested in the assay. In order for the antibody to be considered specific, it must recognize at least 75% of the mark of interest (Egelhofer et al. 2011). This can be tested using dot blot assays. If the antibody recognizes the peptide of interest with an affinity of more than 75% of the signal detected to bind the peptide of interest, the antibody is specific. Equal volumes of peptides (0.5 ng) were pipetted onto a nitrocellulose membrane, and analyzed using a western blot (Figure 11). ChIP experiments would proceed only if the antibody was considered specific. By doing a dot blot, and visually comparing the antibody binding to the peptide, the anti-H3K4me3 antibody from Abcam was determined to specifically bind to the H3K4me3 epitope.

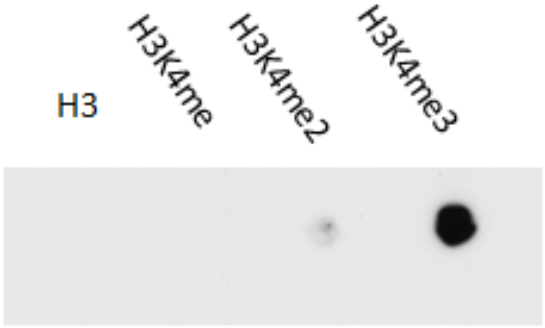


Figure 11. H3K4me3 antibody is specific to the H3K4me3 peptide.

Equal volumes of peptide (0.5ng) were loaded onto a nitrocellulose membrane and incubated with anti-H3K4me3 antibody. The specificity was analyzed using western blot.

The antibody is further tested before ChIP assays were done. To test the efficiency of the antibody to pull down the protein of interest, immunoprecipitation experiments were done and resolved with SDS PAGE and analyzed using western blot. Since the purpose of testing the immunoprecipitation efficiency was for ChIP assays, immunoprecipitation of H3K4me3 was done under the same conditions as those used for ChIP assays. Nuclear lysate of cross-linked cycling 10T1/2 cells were used and 5 A_{260} of DNA was aliquoted for the immunoprecipitation and the negative control sample, IgG (normal rabbit IgG). Input and the immunodepleted fraction of 0.5 A_{260} were also collected. The different ratios (1 ug/A_{260} and 1.5 ug/ A_{260}) of anti-H3K4me3 antibody to DNA was used for the immunoprecipitation experiments and dissolved in SDS loading buffer. SDS PAGE and western blot were used to measure the immunoprecipitation efficiency. The antibody was considered to be efficient for immunoprecipitation if 50% of the input had been immunoprecipitated by the anti-H3K4me3 antibody (ENCODE consortium) (Figure 12). The anti-H3K4me3 from Abcam indicated immunoprecipitation efficiency for both ratio of 1 ug/A_{260} and 1.5 ug/A_{260} of anti-H3K4me3 to DNA. Since both ratios of showed similar immunoprecipitation results, the ratio of 1 ug/A_{260} was used for ChIP experiments.

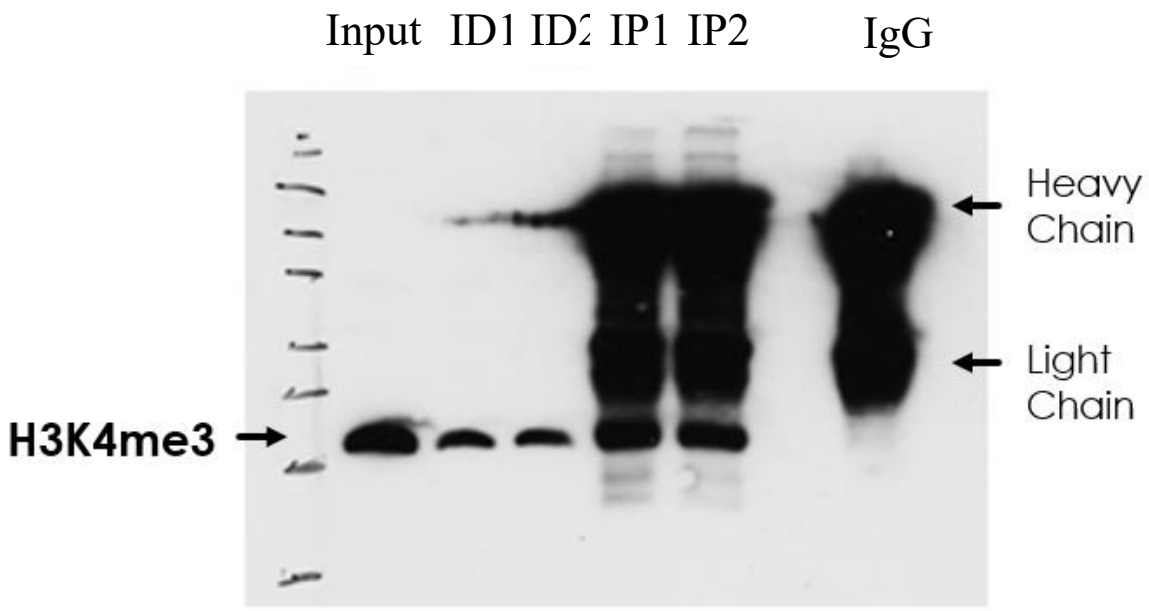


Figure 12. H3K4me3 antibody efficiently immunoprecipitated the H3K4me3 protein. H3K4me3 immunoprecipitation was performed on cycling 10T1/2 under ChIP conditions.

The ratios of antibody to DNA of 1 ug/A_{260} (1) and 1.5 ug/A_{260} (2) were used to determine which ratio would be the most efficient for immunoprecipitation (IP). Levels of the immunodepleted (ID) was also measured. Both antibody to DNA ratios of 1 ug/A_{260} and 1.5 ug/A_{260} showed similar immunoprecipitation.

4.3 H3K4me3 levels increase and decrease with the transcriptional activity of the *Fos* gene

To validate and determine where H3K4me3 peaked along the *Fos* gene in our study, ChIP assays were done on formaldehyde cross-linked cycling 10T1/2 cells to map the levels of H3K4me3 along

the *Fos* gene. Since H3K4me3 levels were known to be more concentrated around the first exon than within the gene body, a primer located far down the gene body was considered a negative control for H3K4me3 levels. ChIP assay using anti-normal rabbit IgG isotype antibody was also used as a negative control. The results were similar to previously published data (Edmunds, Mahadevan and Clayton 2008), and the H3K4me3 levels peaked at the first intron and declined further down the gene body (Figure 13).

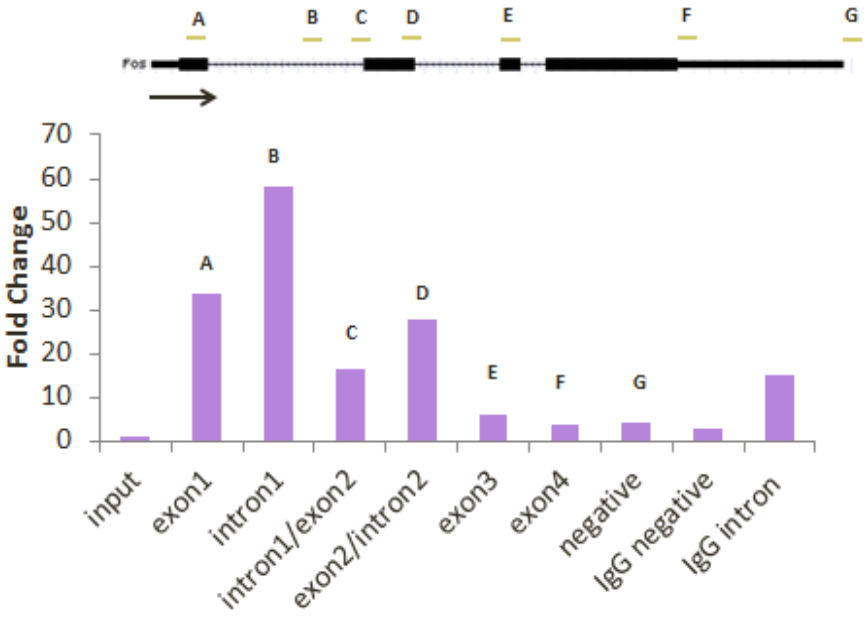


Figure 13. H3K4me3 levels peak at the first intron and decline further along the gene body.

The position of the amplicons used for the real-time PCR is presented above the *Fos* gene schematic. Exons are represented as filled black boxes, and the arrow indicates the direction of gene. Formaldehyde cross-linked mononucleosomes were prepared and used for ChIP and immunoprecipitated for H3K4me3 using anti-H3K4me3 antibody. Immunoprecipitation with normal rabbit IgG antibody was used as a negative control. The average fold change from one individual biological repeat with PCR triplicates is plotted.

To determine when and if H3K4me3 levels diminished during the induction of the *Fos* gene, ChIP assays for H3K4me3 were done on formaldehyde cross-linked EGF stimulated 10T1/2 cells. Quiescent 10T1/2 cells were induced with EGF for varying amounts of time (0, 15, 30 or 60 min). For these experiments, only the intron 1 region was used to determine the changes in H3K4me3 levels, and the exon 3 region was used as a negative control. The results showed that H3K4me3 levels peaked after 15 min of EGF induction before declining at 30 min (Figure 14). These results differed from the previous study, which did not show any decline in H3K4me3 levels along the induction of the *Fos* gene (Edmunds, Mahadevan and Clayton 2008). This result suggests that H3K4me3 levels change with the transcriptional activity of the *Fos* gene.

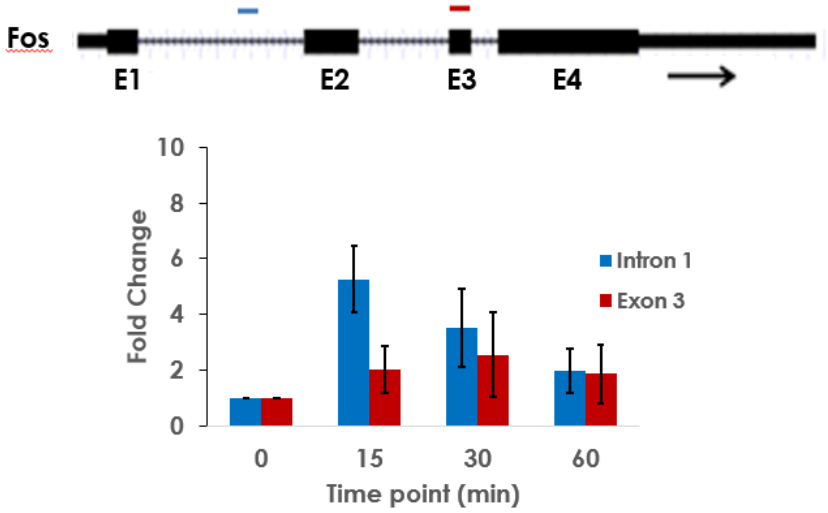


Figure 14. H3K4me3 dynamics along the *Fos* gene.

The position of the amplicons are indicated above along the *Fos* gene map. E1, E2, E3 and E4 represent the location of the of exons. Formaldehyde cross-linked mononucleosomes were prepared and used for ChIP assays using the anti-H3K4me3 antibody in EGF (50 ng/ul) induced 10T1/2 cells at various time points (0, 15, 30 or 60 min). H3K4me3 ChIP assay was normalized to both input and the unstimulated sampled (0 min). The average fold change from three independent experiments was plotted. Error bars represent the standard error mean.

4.4 H3K4me3 ChIP-seq

Prior to doing ChIP-seq experiments, multiple variables were optimized. After validating the antibody (Figure 11) and the immunoprecipitation efficiency (Figure 12) of the antibody used for the H3K4me3 ChIP experiments, ChIP assays were performed. For the H3K4me3 ChIP-seq, ChIP assays were performed on cycling 10T1/2 and EGF-induced 10T1/2 at the various time points 0,

15 and 30 min. The EGF induction after 60 min was not included as we were more focused on the dynamics of H3K4me3 in immediate early genes (Figure 15).

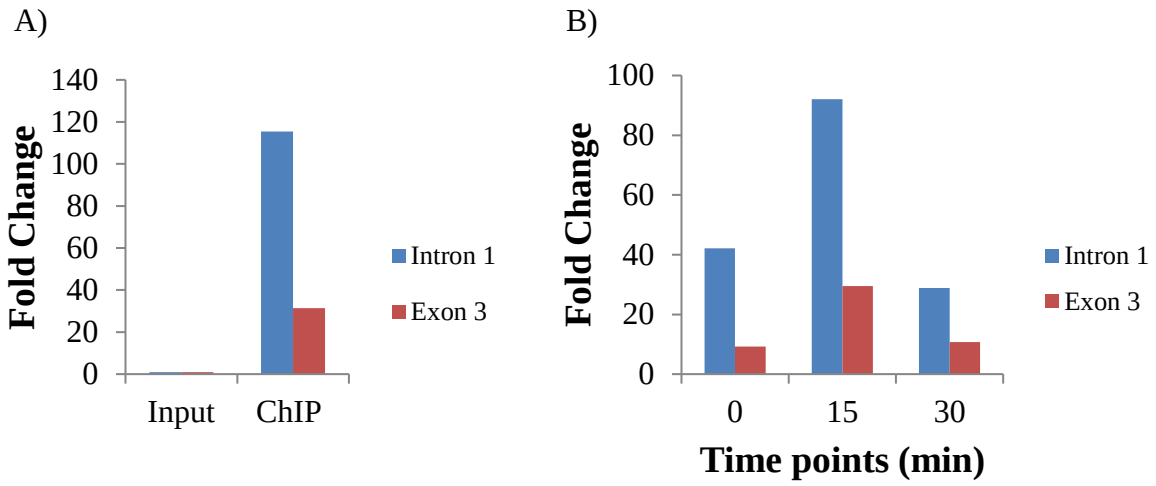
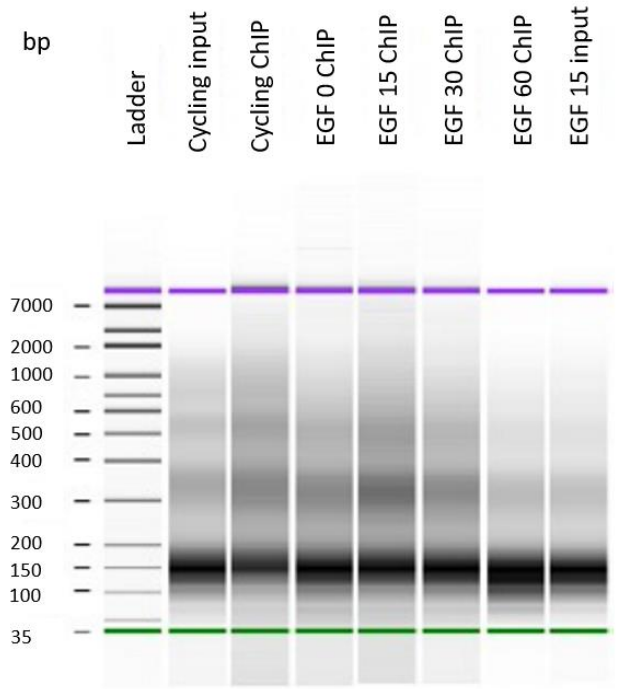


Figure 15. ChIP assay tests for H3K4me3 ChIP-seq.

A) H3K4me3 ChIP assay using cycling 10T1/2, cells and B) H3K4me3 ChIP assays for quiescent 10T1/2 cells stimulated with EGF at various time points (0, 15 and 30 min). Each qPCR was normalized to input, and the average fold change of three PCR triplicates was plotted. Note the scale for each graph is different.

Once ChIP assays for the samples were successful, the concentration of the DNA was measured, and the appropriate size and quantity of the DNA was assessed with the bioanalyzer, which was done by Aleksandar Ilic (Figure 16).

A)



B)

Samples	DNA fragment size (Bp)	Percentage (%)
Cycling Input	154	61
Cycling ChIP	160	40
EGF 0 Input	143	72
EGF 15 Input	144	73
EGF 0 ChIP	152	54
EGF 15 ChIP	155	47
EGF 30 ChIP	151	54

Figure 16. Bioanalyzer tests for H3K4me3 ChIP-seq.

A) Bioanalyzer electrophoresis runs for the H3K4me3 ChIP-seq samples. Order of samples are (from left – right) cycling input, cycling ChIP, EGF 0 ChIP, EGF 15 ChIP, EGF 30 ChIP, EGF 0 input, and EGF 15 input. The bioanalyzer was run by Aleksandar Ilic. B) Table of the percentage of the most predominant DNA fragment size for the various H3K4me3 ChIP-seq samples. All samples showed a high percentage of mononucleosome size DNA fragments. The EGF 15 input was used to construct the H3K4me3 ChIP-seq libraries.

Before the input and ChIP DNA libraries were prepared, the samples had to be mostly mononucleosome sized DNA, and have a minimum amount of DNA (10-20ng). ChIP-seq libraries were prepared by Aleksandar Ilic, and amplified using Techne ^{3}Prime thermocycler. Upon completion of the ChIP-seq libraries, the library products were validated by qPCR to determine if the library preparation had been done successfully (Figure 17) before the samples were sent to Genome Quebec for Illumina Hiseq sequencing.

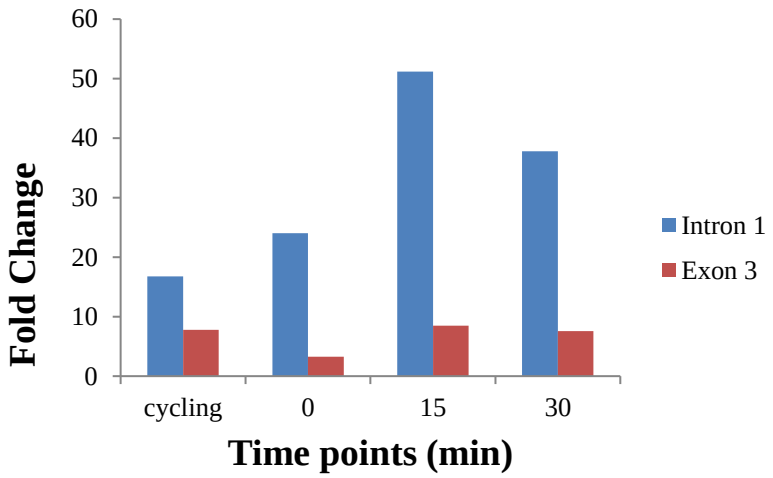


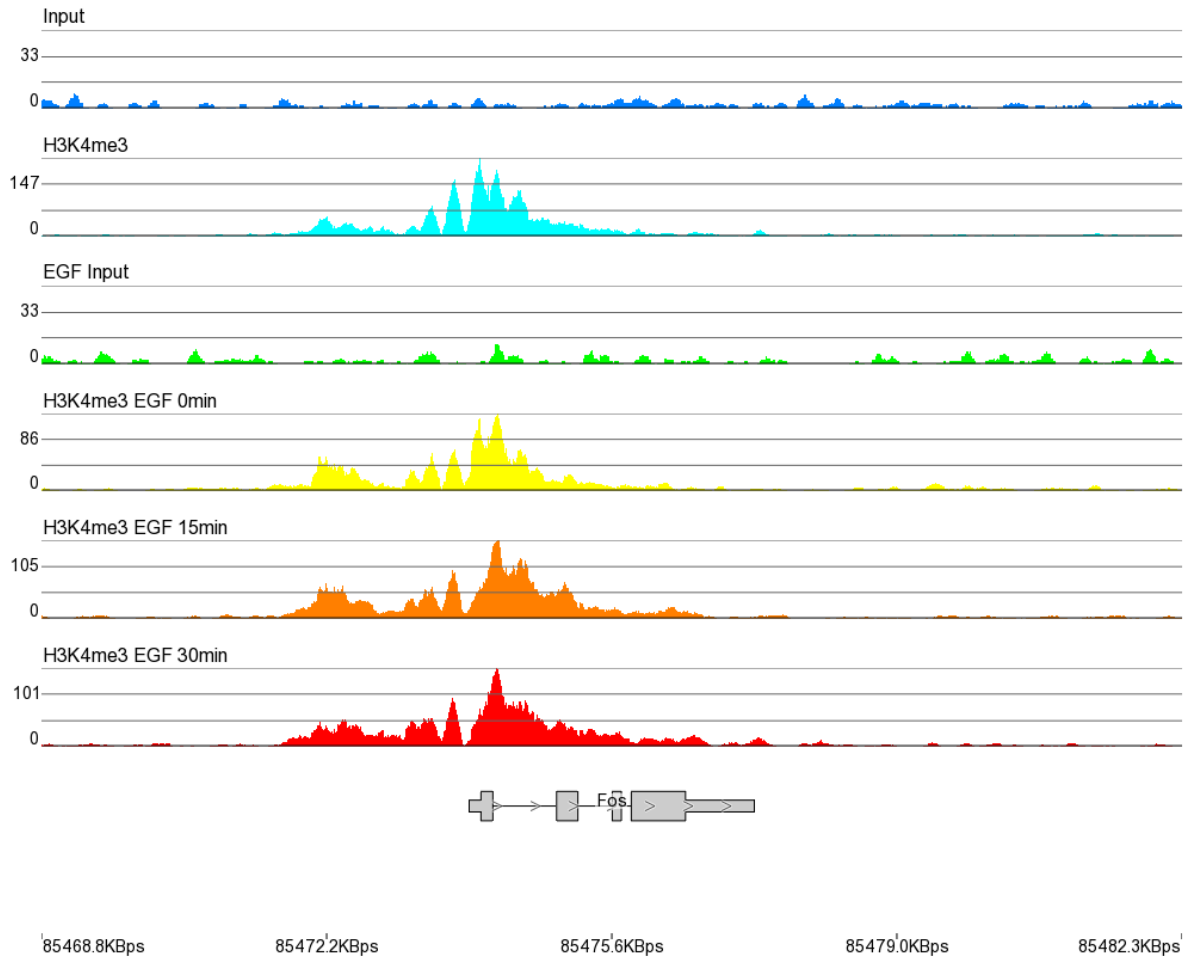
Figure 17. qPCR tests for H3K4me3 ChIP seq libraries.

ChIP seq libraries were prepared using the NEB Ultra II DNA library prep kit and amplified using emulsion PCR. Test qPCRs were normalized to the input. The average fold changes of PCR triplicates are plotted above.

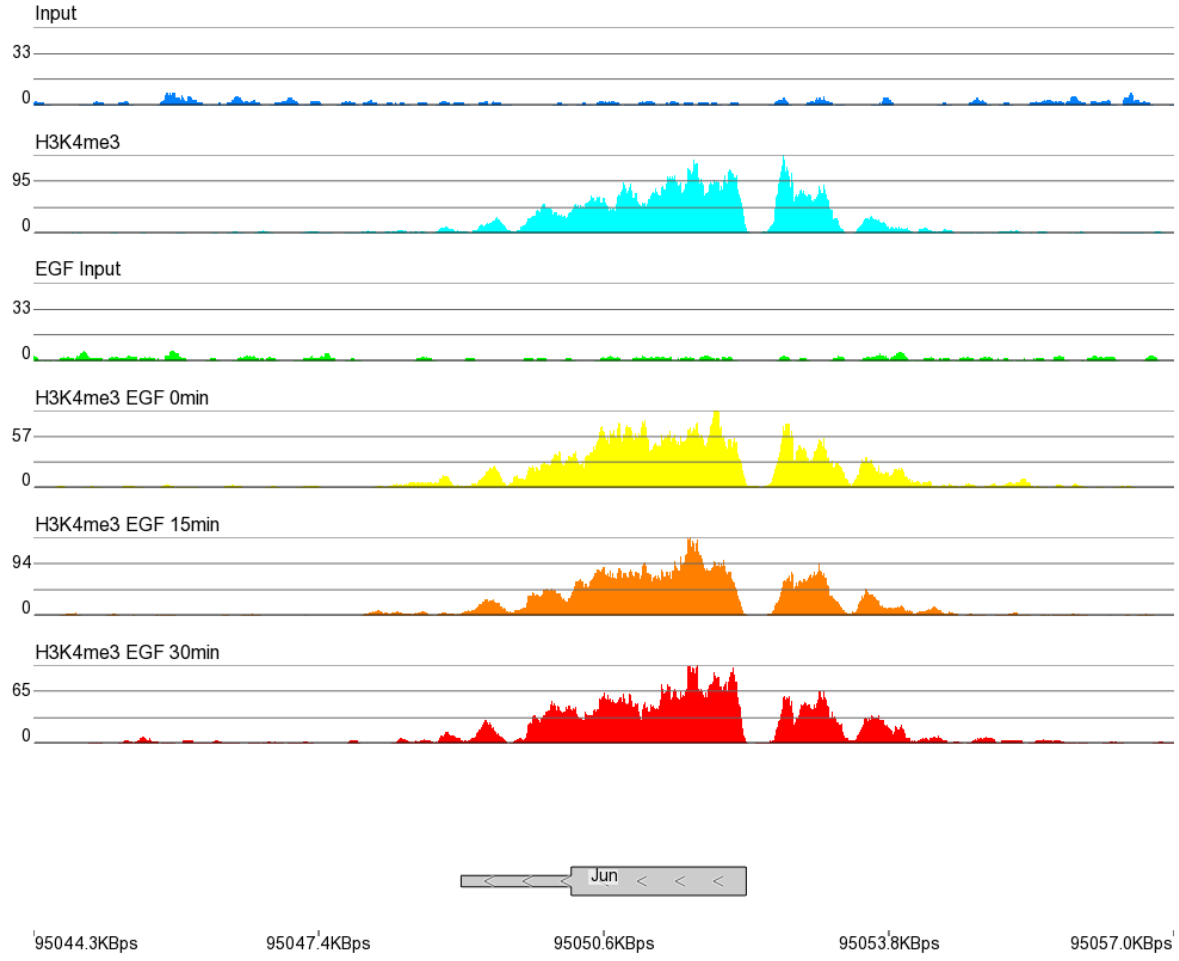
The H3K4me3 ChIP sequence track for the *Fos* gene is shown in Figure 18A. I also present the H3K4me3 tracks of other immediate early genes (*Jun*, *Dusp1*, *Dusp2* and *Fosl1*). The distribution

of H3K4me3 along these genes following stimulation of the cells with EGF is shown in Figure 18. Because ChIP-seq data are not quantifiable, it cannot be used to determine if there was an increase in H3K4me3 levels following 15 min of EGF induction. However, trends within each ChIP-seq track could be the observed. The breadth of H3K4me3 distribution did not change when the immediate early genes were induced. H3K4me3 levels peaked at the 5'splice site for all the following genes that were analyzed. Interestingly, at 15 min induction of EGF, a new peak was observed at the intron 1 region of the *Fos* gene, which was not present in both ChIP-seq tracks for 0 and 30 min induced EGF. Formation of new peaks in the intron 1 region during induction was also observed for the *Fosl1* gene.

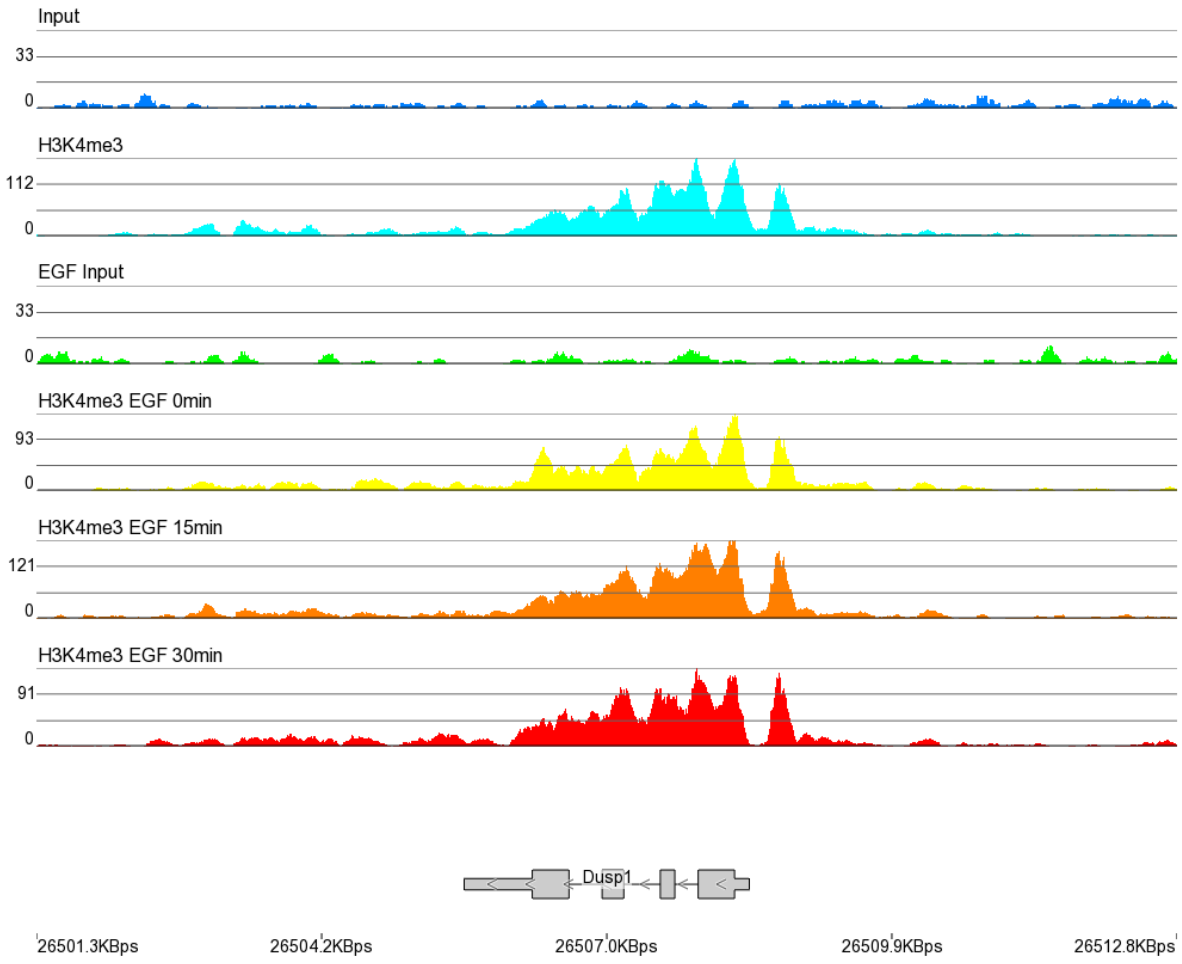
A)


Input
33
0
H3K4me3
147
0
EGF Input
33
0
H3K4me3 EGF 0min
86
0
H3K4me3 EGF 15min
105
0
H3K4me3 EGF 30min
101
0
Fos
85468.8KBps
85472.2KBps
85475.6KBps
85479.0KBps
85482.3KBps

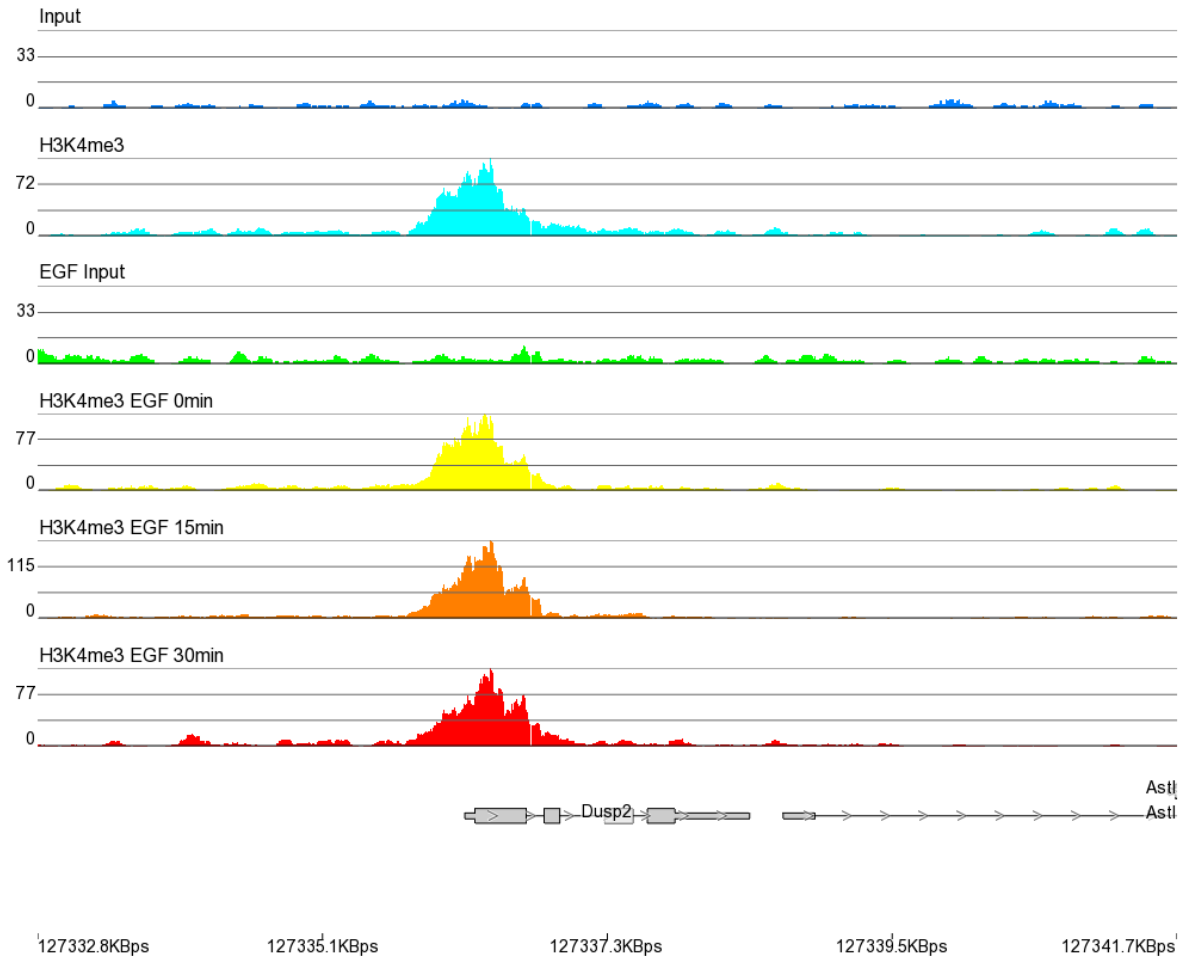
B)


Input
33
0
H3K4me3
95
0
EGF Input
33
0
H3K4me3 EGF 0min
57
0
H3K4me3 EGF 15min
94
0
H3K4me3 EGF 30min
65
0
Jun
95044.3KBps
95047.4KBps
95050.6KBps
95053.8KBps
95057.0KBps

C)


Input
33
0
H3K4me3
112
0
EGF Input
33
0
H3K4me3 EGF 0min
93
0
H3K4me3 EGF 15min
121
0
H3K4me3 EGF 30min
91
0
Dusp1
26501.3KBps
26504.2KBps
26507.0KBps
26509.9KBps
26512.8KBps

D)


Input
33
0
H3K4me3
72
0
EGF Input
33
0
H3K4me3 EGF 0min
77
0
H3K4me3 EGF 15min
115
0
H3K4me3 EGF 30min
77
0
Dusp2
AstI
AstI
127332.8KBps
127335.1KBps
127337.3KBps
127339.5KBps
127341.7KBps

E)

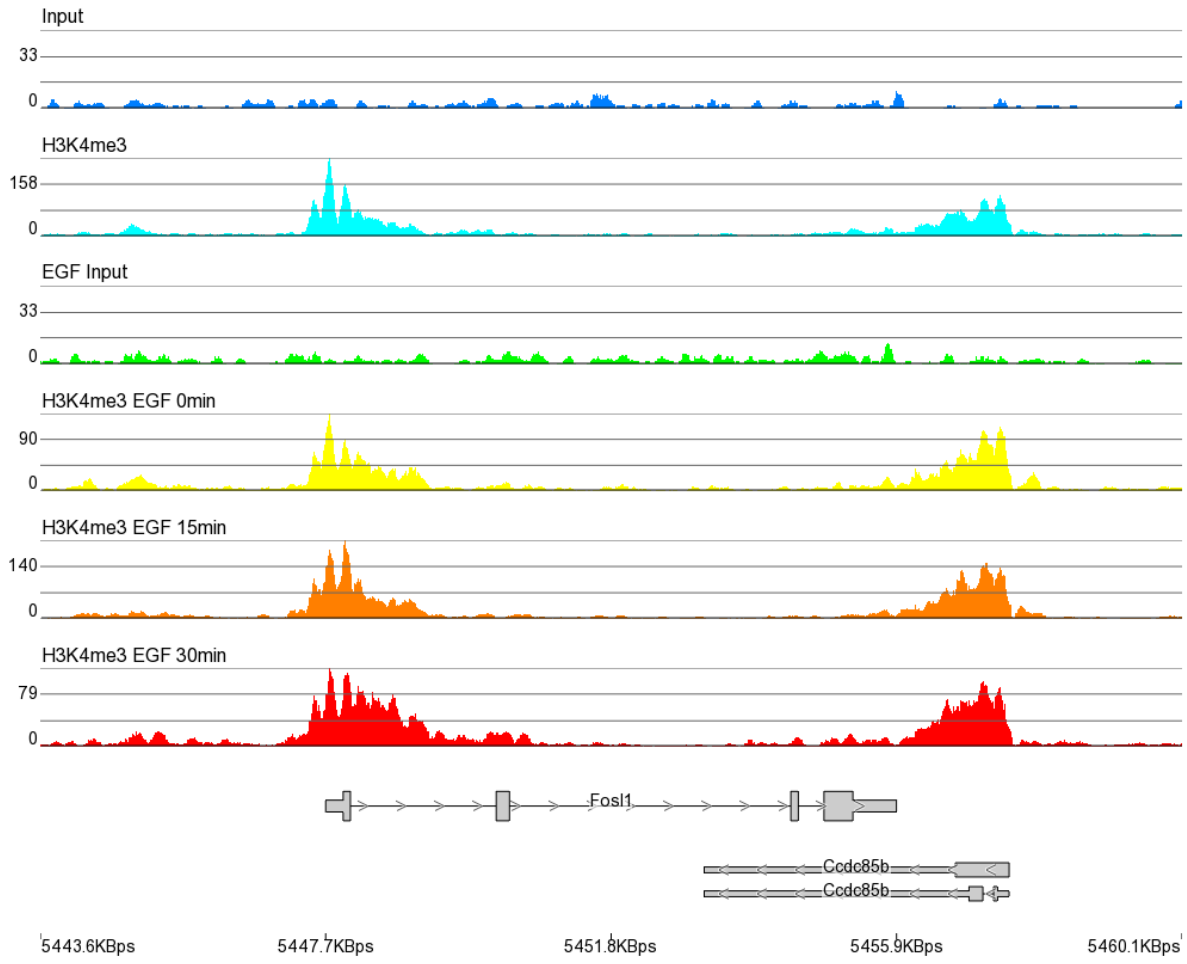


Figure 18. Alignment of input (blue and green) and H3K4me3 ChIP seq peaks from 10T1/2 formaldehyde cross-linked cells under cycling conditions (light blue) and EGF induction at time points 0 (yellow), 15 (orange) and 30 (red) min.

The H3K4me3 ChIP seq peaks co-localized at the 5' splice site region of the immediate early genes *Fos* (A), *Jun* (B), *Dusp1* (C), *Dusp2* (D), and *Fosl1* (E). Y axis represents the reads per kilobase per million mapped. ChIP-seq data is based on one experiment.

4.5 Splicing inhibitor isoginkgetin and transcription inhibitor DRB affect the dynamics of H3K4me3

The splicing inhibitor isoginkgetin was used for our studies to determine whether splicing inhibition affects H3K4me3 dynamics. Since no studies have been published that used isoginkgetin on 10T1/2 cells, the dose response of isoginkgetin on 10T1/2 cells was determined. The concentration for isoginkgetin was first determined by using various concentrations (50, 100 and 300 uM) with a constant incubation time of 30 min. Since the presence of pre-existing spliced RNA would interfere with the ratio of unspliced to spliced variants produced by isoginkgetin, 10T1/2 cells were synchronized in the G_0/G_1 phase. The synchronized cells were incubated with the splicing inhibitor and induced with EGF for 30 min as mRNA transcript levels were the highest at 30 min for the *Fos* gene. The RNA was isolated; reverse transcribed and amplified using PCR. The cDNA products were purified and ran on a 1.5% agarose gel to determine if the products were spliced or unspliced (Figure 19). DMSO was also used as a vehicle control. The results showed evidence of splicing inhibition after the use of 100 uM of isoginkgetin with 30 min incubation. Using 300 uM of isoginkgetin, it became more prominent that splicing inhibition had occurred. Some unspliced transcript was also observed in the cDNA that had not been stimulated by EGF (EGF 0) after splicing inhibition.

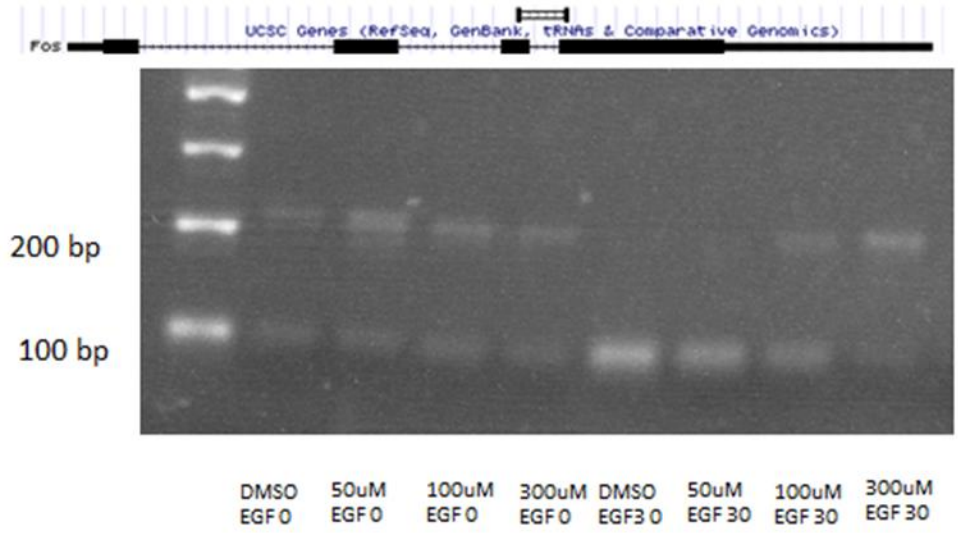


Figure 19. Splicing of the *Fos* transcript is partially inhibited with 100 uM of isoginkgetin after 30 min incubation.

The above map shows the location of the amplicon on the *Fos* gene used to detect the unspliced transcript (204 bp) and the spliced transcript (78 bp). Serum starved 10T1/2 Cells were treated with splicing inhibitor isoginkgetin at various concentrations (50 uM, 100 uM, and 300 uM) for 30 min. After isoginkgetin incubation, the Fos transcript was either unstimulated (EGF 0) or induced with EGF for 30 min (EGF 30). DMSO was used as a vehicle control.

Thus, a lower dose of isoginkgetin (30 uM) was tested using various incubation times that were longer than 30 min. The incubation times tested were 3, 8 and 24 hr (Figure 20). At 3 hr incubation of 30 uM isoginkgetin splicing inhibition was observed. However, the decline in *Fos* transcript was still observed so another dose response was done for isoginkgetin using various concentrations of 5, 10 and 20 uM were used with a 3 hr incubation time (Figure 21). The results indicated that the concentration of isoginkgetin was too low, and no splicing inhibition was observed. However,

transcription continued to be inhibited. From these dose response results, we decided to use the concentration of 30 uM of isoginkgetin with a 3 hr incubation for further experimental analyses.

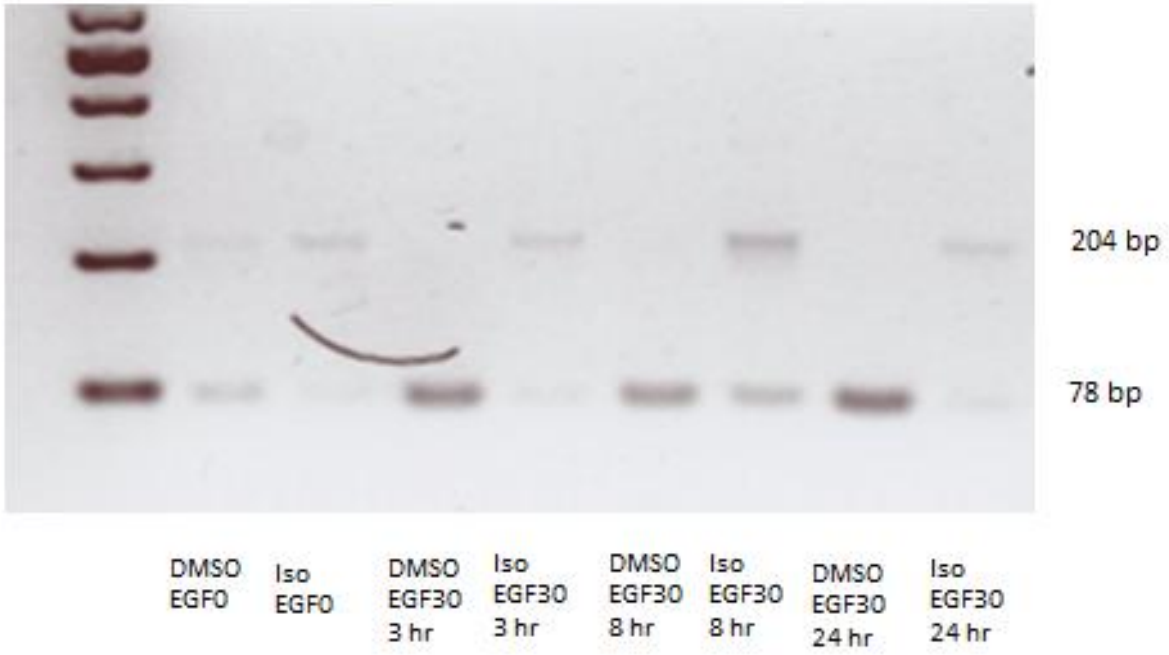


Figure 20. Splicing is inhibited with 30 uM of isoginkgetin after a 3 hr incubation.

Multiple incubation times (3 hr, 8 hr and 24 hr) were tested using 30 uM of isoginkgetin (Iso). After isoginkgetin incubation, the Fos transcript was either unstimulated (EGF 0) or induced with EGF for 30 min (EGF 30). The unspliced transcript (204 bp) and the spliced transcript (78 bp) at different time points were detected. DMSO was used as a vehicle control.

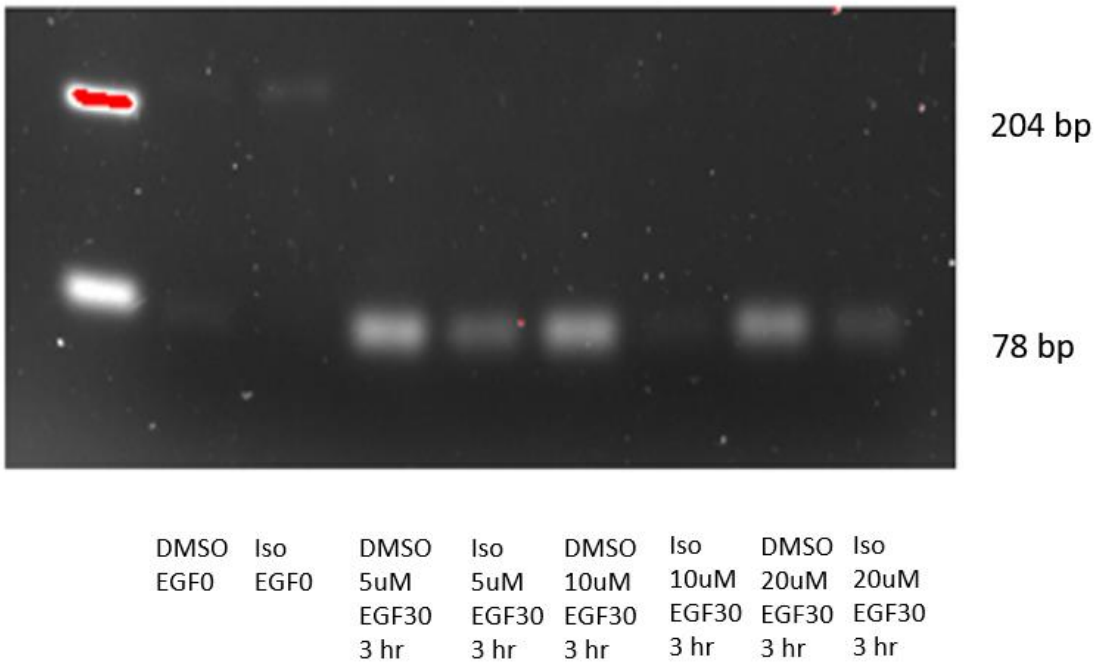


Figure 21. For splicing inhibition using isoginkgetin (Iso), at least 30 uM is required with an incubation time of 3 hr.

Treated cDNA was amplified and loaded on a 1.5% agarose gel. Various concentrations (5, 10 and 20 uM) of isoginkgetin were tested with a constant isoginkgetin incubation time of 3 hr. After isoginkgetin treatment, the Fos transcript was either unstimulated (EGF 0) or induced with EGF for 30 min (EGF 30). The unspliced transcript (204 bp) and the spliced transcript (78 bp) were observed. DMSO was used as a vehicle control.

To determine if splicing inhibition plays a role in the dynamics of H3K4me3 levels along the *Fos* gene, H3K4me3 ChIP assays were done on isoginkgetin-treated 10T1/2 cells. Quiescent 10T1/2 cells were treated with isoginkgetin (30 uM) and incubated for 3 hr before cells were stimulated

with EGF (Figure 22). From these results a trend was observed with a decrease in H3K4me3 levels at 15 min. However, when the assay was normalized to exon 3, H3K4me3 levels decreased at both 15 min and 30 min time points. This suggested that splicing may play a role in H3K4me3 dynamics. However, because the splicing inhibitor also inhibited transcription, H3K4me3 ChIP assays were done with DRB treated 10T1/2 cells.

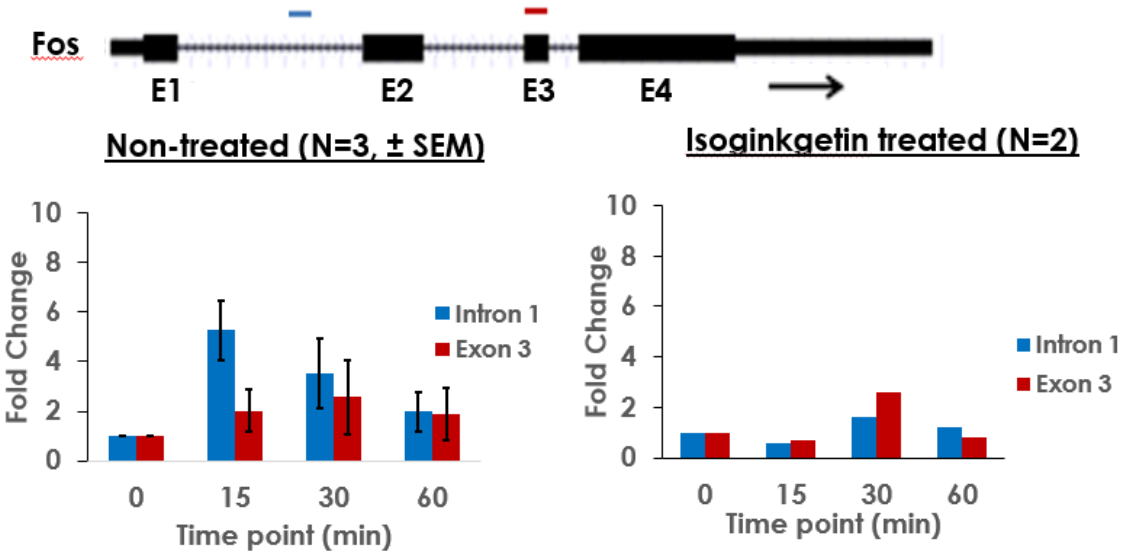


Figure 22. Splicing inhibition causes decline in H3K4me3 levels after 15 min of EGF stimulation. Quantitative PCR of H3K4me3 ChIP assays treated with isoginkgetin (30 uM for 3 hr) stimulated with EGF at various time points (0, 15, 30 and 60 min) are plotted above. The map above shows the relative location of the amplicons used (blue is intron 1, red is exon 3), with the exons labeled E1, E2, E3 and E4. H3K4me3 ChIP assays not treated with isoginkgetin (left) are compared to the isoginketin treated H3K4me3 ChIP assays (right). Both non-treated and isoginkgetin treated H3K4me3 ChIP assays were normalized to the input and EGF 0. The average fold change of three independent experiments (N=3) were plotted for the non-treated H3K4me3 ChIP assays. Error bars represent the standard error mean (SEM). Two independent experiments (N=2) were plotted for the isoginkgetin treated samples.

Prior to these H3K4me3 ChIP assay experiments, RT-qPCR was done to determine at which concentration and incubation time DRB would inhibit transcription. Since H3K4me3 ChIP assay

experiments using DRB has been previously published (Edmunds, Mahadevan and Clayton 2008), the concentration (25 ug/ml) and incubation time (10 min) used previously was tested (Figure 23). The additional immediate early gene *Jun* was tested to determine if transcription of other immediate early genes had been inhibited. Using these conditions, our results showed transcription inhibition occurred after quiescent cells were incubated for 10 min with DRB followed by 30 min of EGF induction.

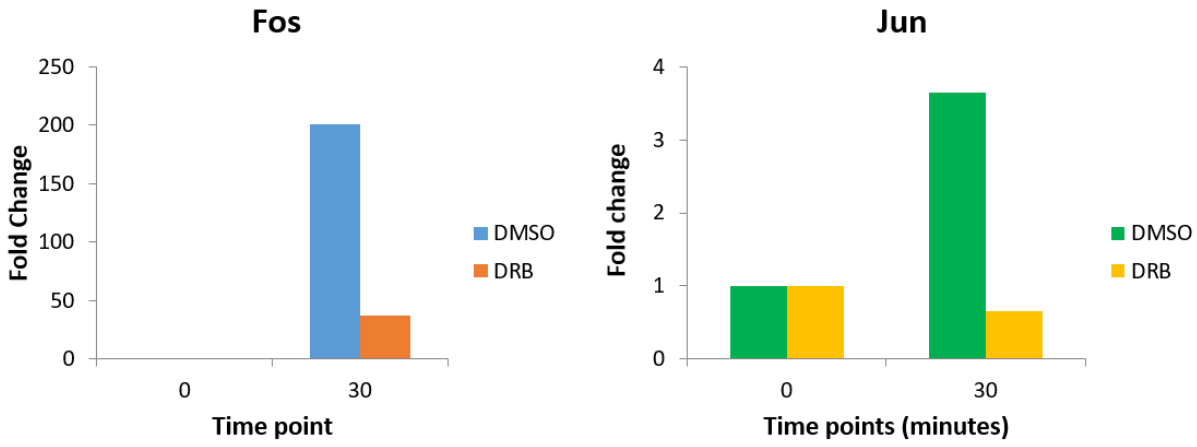


Figure 23. Transcription of the *Fos* gene in 10T1/2 cells are inhibited with DRB (25 ug/ml) with 10 min incubation time.

RT-qPCR using DRB treated 10T1/2 cells unstimulated (0 min) and stimulated (30 min) with EGF. DMSO was used as a vehicle control. Average fold changes of PCR triplicates are plotted.

With the concentration and time incubation being optimized, H3K4me3 ChIP assay with DRB treated 10T1/2 cells were done (Figure 24). When the qPCR results are normalized to the input

and the EGF unstimulated sample (0 min), it can be observed that H3K4me3 levels are inhibited when transcription is inhibited.

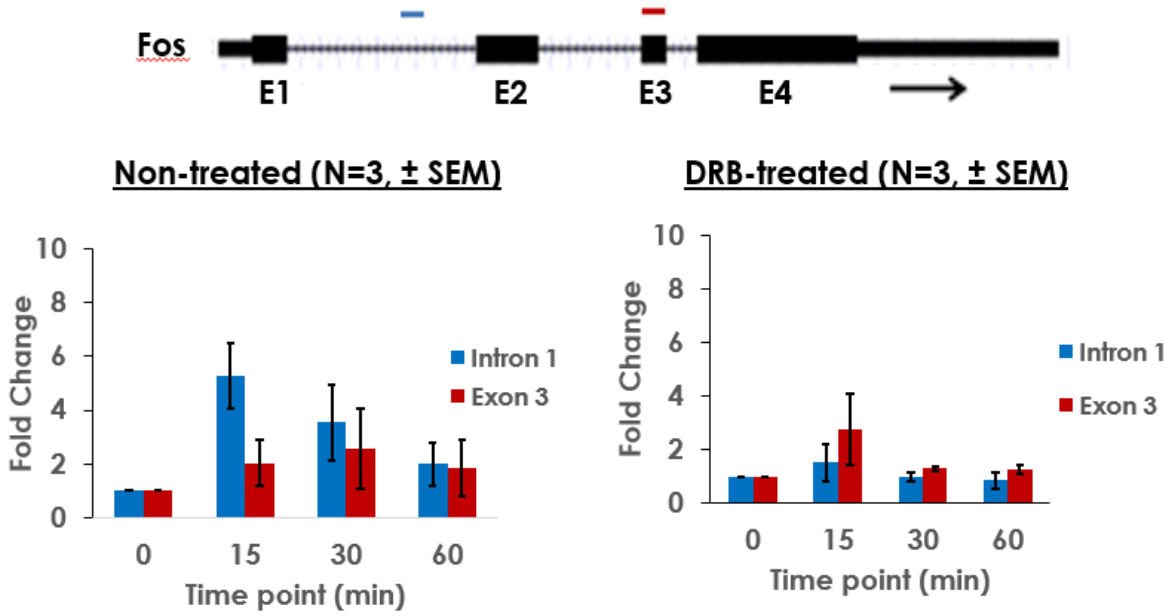


Figure 24. Inhibition of *Fos* transcription depletes H3K4me3 levels.

Quantitative PCRs of H3K4me3 ChIP assays non-treated (left) and treated with DRB (25 ug/ml for 10 min) and stimulated with EGF at various time points (0, 15, 30 and 60 min) are plotted above (right). The map above shows the relative location of the amplicons used (blue is intron 1, red is exon 3), with the exons labeled E1, E2, E3 and E4. The H3K4me3 ChIP assays were normalized to the input DNA and the unstimulated sample (0 min). Average fold change of three independent experiments (N=3) for both the non-treated and DRB treated are plotted. Error bars represent standard error mean (SEM).

The ChIP analysis data of H3K4me3 levels for each treatment (H3K4me3 with isoginkgetin, H3K4me3 with DRB and H3K4me3 without inhibitors) were compiled, and changes in H3K4me3 levels between treatments were compared (Figure 25). The qPCR results that were compiled were normalized to the input and the EGF unstimulated sample (0 min).

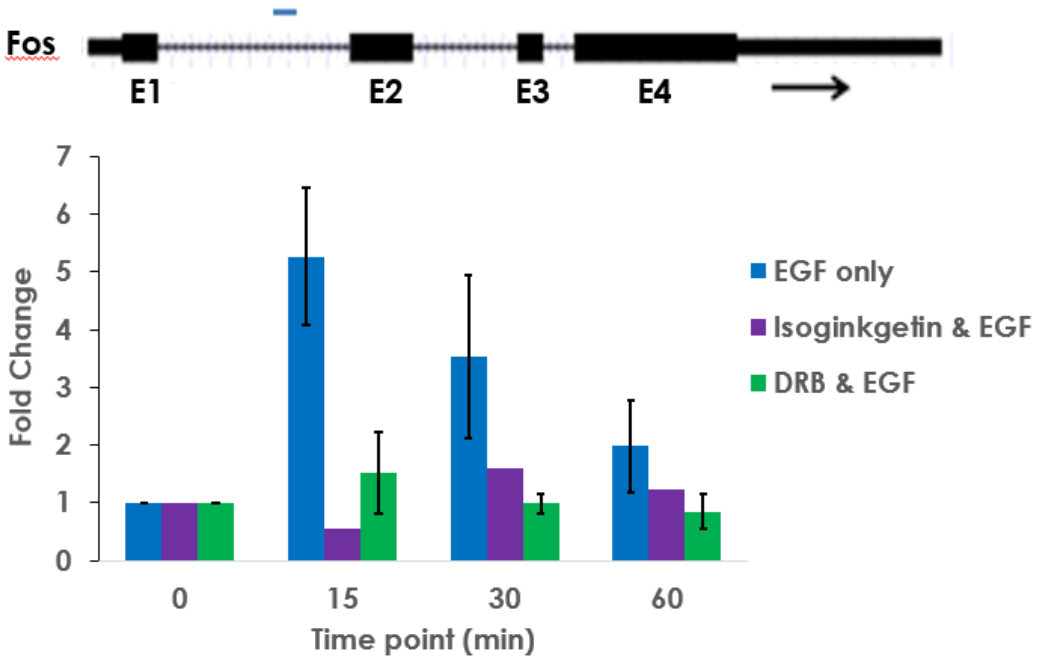


Figure 25. Compiled ChIP assays done for H3K4me3 levels for the *Fos* gene.

Quantitative PCRs of H3K4me3 ChIP assays treated with either DRB (25 ug/ml for 10 min) (green), isoginkgetin (30 mM for 3 hr) (purple) or untreated with inhibitors (blue) followed by stimulation with EGF at various time points (0, 15, 30 and 60 min) are plotted above. The map above shows the relative location of the amplicons used for intron1, with the exons labeled E1, E2, E3 and E4. H3K4me3 ChIP assays were normalized to the input DNA and the unstimulated sample (0 min). Average fold change of three independent experiments are plotted for the DRB treated, and untreated with inhibitor ChIP assays. Average fold change of two independent isoginkgentin ChIP assays for H3K4me3 plotted and were calculated based on the average of PCR triplicates. Error bars represent the standard error mean.

5. DISCUSSION

5.1 H3K4me3 dynamics along the transcriptional activation of the *Fos* gene

From the ChIP result for EGF induced 10T1/2 cells at time points 15, 30, 60 min and unstimulated (EGF 0), H3K4me3 levels increased and decreased with the transcriptional activity of the *Fos* gene. The *Fos* gene transcripts levels were the highest at 30 min, while H3K4me3 levels were the highest at 15 min. Consistent with previous findings, human *FOS* transcription occurred at 30 min (Edmunds, Mahadevan and Clayton 2008; Teoh and Sharrocks 2014). However, the increase of H3K4me3 levels prior to the highest level of mRNA expression is not consistent with what was previously observed (Edmunds, Mahadevan and Clayton 2008). Interestingly, depletion of RBBP5 in EGF induced Hela cells resulted in an increase in unspliced *FOS* transcript at 15 min (Teoh and Sharrocks 2014). When comparing the H3K4me3 ChIP data from this study to the changes of pre-mRNA levels of *FOS* during induction with EGF, H3K4me3 dynamics were similar to the changes in pre-mRNA expression along the induction of the *FOS* gene (Teoh and Sharrocks 2014). A previous study determined through mapping of the methyltransferase Set1 occupancy, that Set1 binding was concentrated at 5' coding regions associated with active polymerase II activity. Co-immunoprecipitation experiments indicated that Set1 was associated with RNA polymerase II when phosphorylated at serine 5, which is associated with the initiating form of the RNA polymerase, but not with the RNA polymerase II when serine 2 was phosphorylated which is associated with the elongating form of the RNA polymerase (Ng et al. 2003). The paper further tested the association of Set1 occupancy and H3K4me3 levels. From their findings, they determined that localized H3K4me3 levels had an apparent relationship with Set1 occupancy (Ng et al. 2003).

Together the results shown in the study done by Ng et al. (2003), and the observation that H3K4me3 levels peaked when pre-mRNA levels were the highest provides further support that H3K4me3 dynamics may be transcription dependent. However, since steady state of mRNA expression of *Fos* was measured, when transcription initiation and elongation occurs cannot be determined through these experiments, and further studies must be done.

The ChIP assay results of H3K4me3 also revealed that H3K4me3 levels decreased alongside the cessation of transcription of the *Fos* gene. This finding is not consistent with previous studies (Edmunds, Mahadevan and Clayton 2008; Ng et al. 2003; Borde et al. 2009). However, H3K4me3 is a post-translational modification that is observed to be affected by stress and may be involved in transcription memory (Ding, Fromm and Avramova 2011). In *Arabidopsis*, when exposed to repetitive stress, H3K4me3 levels has been shown to increase and were retained during the recovery phase, when transcription was at basal level (Ding, Fromm and Avramova 2011). The *Fos* gene is an immediate early gene and can be also activated by stress stimuli via the p38 pathway. It is possible that the difference in results compared to the previous study (Edmunds, Mahadevan and Clayton 2008) may be due to differences in environmental factors, and potential stress accrued during experiments. Borde et al. (2009) did a global analysis of H3K4me3 in *S. cerevisiae* and determined that majority of genes meiotically downregulated (73%) showed a gradual decrease of H3K4me3 levels. However, not all genes displayed the same trend of H3K4me3 levels, suggesting that changes in H3K4me3 levels are gene dependent.

Due to normalization issues, the ChIP-seq data for H3K4me3 could not be measured quantitatively during the different induction time (Bonhoure et al. 2014; Chen et al. 2016). Issues with normalization of ChIP-seq are due to inaccurate assumptions with the data. When normalizing ChIP-seq data to each other, it is inaccurate as the amount of DNA for each sample varies. In addition, changes in local signals may change the total signals in a sample (Chen et al. 2016). Interesting, although an increase of H3K4me3 level was detected at 15 min in ChIP assays, there was no change in H3K4me3 distribution breadth during *Fos* induction. This suggested that changes in H3K4me3 levels occurred only at specific regions along the gene body. However, an increase in H3K4me3 intensity was observed at the intron 1 region of the *Fos* gene with induction of EGF for 15 min that was not present in both the unstimulated (EGF 0) and induction of EGF for 30 min samples. The increase of H3K4me3 levels at the intron 1 region was also observed during the induction of the *Fosl1* gene after EGF treatment for 30 min.

When mapping the distance of the intron 1 (+749 - +844) amplicon used for H3K4me3 ChIP assays to predicted branch points in the intron 1 region, the amplicon was around 100-200 bp (+947, and +999) from the branch point. These observations suggest a relationship between H3K4me3 and the branch point in the intron 1 region. Since no expansion of H3K4me3 domains was observed in ChIP-seq during EGF stimulation, the reason for increased H3K4me3 levels along the gene body is most likely not due to an increase in nucleosome numbers. Other possible outcomes that may result in the increase in H3K4me3 levels are methylation of the other histone H3 in the histone octamer and/or an increase in *Fos* alleles that are modified by H3K4me3 (Figure 26).

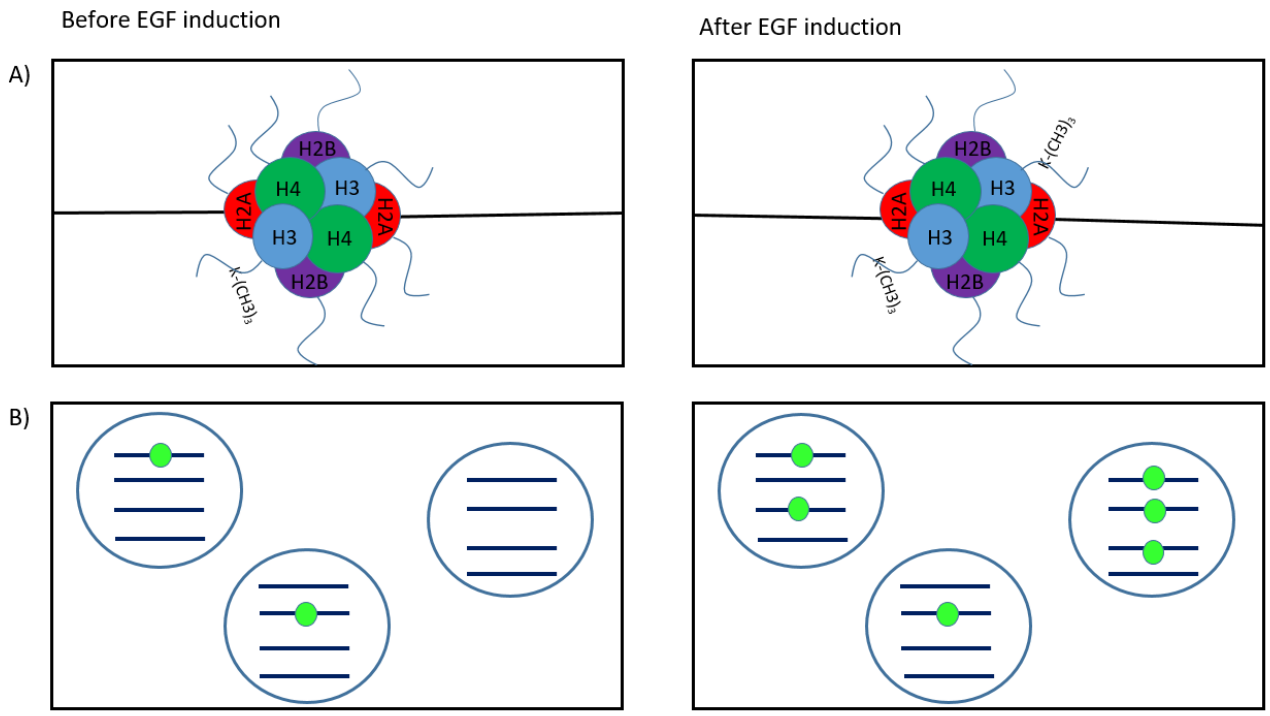


Figure 26. Possible mechanisms of how H3K4me3 levels increase in 10T1/2 cells during EGF induction of the *Fos* gene.

A) Addition of methylation at lysine 4 (K-$(CH_3)_3$) may occur on the other histone H3 N-terminal tail present on the same histone octamer after induction of EGF. B) Each 10T1/2 cell has on average 4 alleles (navy line) in each cell (blue outline circle) for a specific gene. The *Fos* alleles may have trimethylation of H3K4 (light green circle), independent to other alleles present in the same cell or in other cells. Increase of H3K4me3 levels after EGF induction may be caused by the increase in trimethylation at H3K4 at individual alleles.

The nucleosome unit contains a histone octamer. The histone octamer core consists of H2A and H2B dimers, and H3-H4 tetramers (Davie et al. 2015). It is possible that the increase in H3K4me3 levels may be due to the addition of a methyl group to the other histone H3 N-terminal tail (Figure 24A). Additionally, in the mouse cell line 10T1/2, the chromosome number is around 81 (2N), and each cell would contain on average 4 alleles for each gene (Khan et al. 2017). In one study it was observed that in a cell population, there was differentiational levels of the histone post-translational modification histone H3 serine 10 phosphorylation and histone H3 serine 28 phosphorylation on each allele during serum starvation, following stimulation by EGF or 12-O-tetradecanoylphorbol-13-acetate in 10T/12 mouse cells and CCD 1070SK human fibroblast cells (Khan et al. 2017). Alleles that associate with different histone post-translational marks are called epi-alleles. Using fluorescence in situ hybridization, they determined that histone H3 serine 10 phosphorylation and serine 28 phosphorylation more frequently were not colocalized, and that there was heterogenous levels of these histone post-translational modification epi-alleles for the gene *Jun* (Khan et al. 2017). It is possible that H3K4me3 levels may also vary between epi-alleles of the *Fos* gene in 10T1/2 cells. Thus, the increase in H3K4me3 levels may due to the increase in trimethylation occurring on H3K4 at individual epi-alleles. This would result in an increase in average H3K4me3 levels observed through ChIP assays and ChIP-seq (Figure 24B). It is also possible that a combination of both trimethylation of histone H3K4 on the other histone H3 in a nucleosome, and trimethylation of histone H3K4 at different epi-alleles may be responsible in the increase of H3K4me3 levels observed.

The other mechanism of interest with H3K4me3 dynamics was how H3K4me3 levels decrease along a gene body. Three models were proposed that may be responsible for the decline of

H3K4me3 levels. The first model suggests that the decrease in H3K4me3 levels are due to the turnover of the histone H3 in the histone octamer. The second model suggests that KDMs are responsible for the demethylation of the H3K4. Finally, the last model proposed that both KDMs and HDACs and KATs interact with each other to decrease the levels of H3K4me3. In one study indicated a relationship with histone H3 lysine 4 acetylation and H3K4me3 in breast cancer progression (Messier et al. 2016). While histone H3 lysine 4 acetylation was more prevalent in early stages of breast cancer, H3K4me3 was associated with more aggressive and metastatic forms of breast cancer (Messier et al. 2016). In addition, KDM5B has been observed to bind with HDAC1, HDAC4, HDAC5 and HDAC7 through co-immunoprecipitation assays (Barrett et al. 2007). In the third model, I propose that a KDM5 family member demethylates H3K4me3 following the cessation of *Fos* gene transcription. The associated HDAC activity would keep the demethylated H3 in an unacetylated state. Only after the KDM5-HDAC activity was absent from the targeted nucleosome, the H4K4 of that nucleosome could be potentially acetylated, preventing methylation.

5.2 Inhibition of splicing and H3K4me3 dynamics

Dose response curves for isoginkgetin, revealed a diminished signal from the samples treated with isoginkgetin followed by EGF stimulation for 30 min compared to the vehicle control. This indicated that less RNA transcript was being made, suggesting that the splicing inhibitor isoginkgetin may also be affecting transcription. The coupled inhibition of splicing and transcription is expected as pre-mRNA splicing has been determined to mostly occur cotranscriptionally (Braunschweig et al. 2013; Shukla and Oberdoerffer 2012). In the dose curve

analysis of isoginkgetin, the vehicle controls revealed two slight bands (204 bp for unspliced mRNA and 78 bp for spliced mRNA) on the agarose gel. A possible reason for this is the cells are being synchronized at G_0 phase and thus caused low levels of transcription and inefficiency in splicing.

I observed that treating cells with isoginkgetin before EGF prevented the increase in H3K4me3 at the Fos gene intron 1. (Figure 23). In addition, ChIP assays for H3K4me3 treated with DRB ± EGF also prevented an increase in H3K4me3 levels at intron 1 of the *Fos* gene (Figure 23). Results obtained from this study were consistent to previous results done in *FOS* gene constructs, which used a different splicing inhibitor. H3K4me3 levels were depleted after treatment with splicing inhibitor spliceostatin A at the 5' splice site (Bieberstein et al. 2012). H3K4me3 levels at this site were also depleted after incubation with transcription inhibitor DRB (Bieberstein et al. 2012; Edmunds, Mahadevan and Clayton 2008). However, because of the coupled relationship between transcription and pre-mRNA splicing, it is difficult to determine if H3K4me3 levels are associated with transcription, splicing or both.

One study has also shown that inhibition of splicing affects RNA polymerase II serine 2 phosphorylation (Koga, Hayashi and Kaida 2015). Using splicing inhibitors pladienolide B and spliceostatin A, phosphorylation of the C-terminal domain on the RNA polymerase II was examined during splicing inhibition. Using western blot for total RNA polymerase II, RNA polymerase II serine 2 phosphorylation, and RNA polymerase II serine 5 phosphorylation, it was observed that splicing inhibition decreased levels of RNA polymerase II serine 2 phosphorylation but not RNA polymerase II serine 5 phosphorylation. These results were not observed when cells

were treated with a transcription inhibitor (actinomycin D) and translation inhibitor (cycloheximide), providing more evidence that splicing inhibition was responsible for the decrease in RNA polymerase II serine 2 phosphorylation. These results were validated using an alternative method of splicing inhibition using antisense morpholino oligo, which also provided similar results (Koga, Hayashi and Kaida 2015).

5.3 Conclusion and Significance

This study determined that H3K4me3 levels increase and decrease along the transcription and cessation of transcription of the *Fos* gene. These changes in H3K4me3 level may be dependent on transcription activation. Although inhibition of splicing decreased H3K4me3 levels to basal level, based on our studies it is difficult to determine whether the decrease in H3K4me3 levels is due to splicing inhibition, or splicing inhibition decreases H3K4me3 levels indirectly by inhibiting transcription.

These results may aid in providing the foundation on understanding H3K4me3 dynamics, which may have a broad implication in many diseases such as breast cancer, colorectal cancer, leukemia and prostate cancer which have been observed to display dysregulation in H3K4me3 writers and erasers. Since H3K4me3 is associated with transcriptionally active genes, increase in levels of H3K4me3 along specific gene may be an indication of those genes being overexpressed. Thus it can act as a marker for determining which genes are being transcribed at a specific time. In addition, mutation causing disruption of splicing is the cause of at least 30% of diseases in humans (Merkhofer, Hu and Johnson 2014). Since splicing occurs mostly cotranscriptionally, it may also

have implications in developmental defects (Merkhofer, Hu and Johnson 2014). H3K4me3 levels are associated with transcription and splicing. By understanding how H3K4me3 is associated with these mechanisms, it will provide better knowledge of how diseases dysregulates transcription and splicing, and may provide possible mechanisms as to how to alter H3K4me3 levels to revert a cancerous phenotype back to a non-cancerous one.

5.4 Future Direction

To further investigate the dynamics of H3K4me3, the ChIP-seq data requires further validation, and will be repeated. However, due to the difficulty of normalization in ChIP-seq, data analyses through ChIP-seq is not quantifiable (Bonhoure et al. 2014; Chen et al. 2016). To quantify the ChIP-seq data, a spike adjustment procedure will be done prior to DNA purification of the ChIP DNA. A specific ratio of foreign DNA is added with the ChIP sample. The ratio the foreign DNA must remain constant during data analysis to confirm linear amplification has occurred (Chen et al. 2016). This can then be used to determine the genomic levels of H3K4me3 along the transcriptional induction of all immediate early genes in the mouse cell line 10T1/2.

To determine if the increase of H3K4me3 is due to trimethylation of histone H3K4 at different epi-alleles, immune-fluorescence in situ hybridization assays will be done on serum starved and EGF stimulated (at 15 min) 10T1/2 cells. Immediate early genes will be probed for, and the antibody for H3K4me3 will be used to measure the colocalization of H3K4me3 to immediate early genes. If there is an increase in colocalization of the immediate early gene to H3K4me3 levels after transcriptional induction, it would be suspected that increase in H3K4me3 levels is due to the

presence of more H3K4me3 associating with epi-alleles after EGF stimulation. To determine if the increase of H3K4me3 levels are due to the trimethylation of the second histone H3 present in the nucleosome, an assay that can detect individual nucleosome must be utilized. This assay must also be able to differentiate a nucleosome containing one histone H3 trimethylated at lysine 4 from a nucleosome with both H3K4 trimethylated. Unfortunately, there is no assay that can fulfill these criteria, and limits the ability to determine if addition of trimethylation at lysine 4 on the other histone H3 in a nucleosome causes H3K4me3 levels to increase.

There are many enzymes involved in regulating H3K4me3 such as SETD1A/1B, MLL and KDM5B proteins. However, when and where these enzymes are recruited along the immediate early gene body following induction is not known. To provide further insight on the dynamics of H3K4me3 during induction of immediate early genes, dual cross-linking ChIP assays to determine the timing and location of H3K4me3 regulatory enzymes across the gene body of EGF-induced *Fos* and *Jun* gene will be done. Mouse 10T1/2 cells will be dual cross-linked with formaldehyde and dithiobis (succinimidyl propionate) to crosslink DNA to protein, as well as protein to protein (Zeng et al. 2006). Others and our research group have reported that dual cross-linking is more efficient than formaldehyde alone to monitor the association of regulatory proteins across genes (Zeng et al. 2006). After cross-linking, ChIP assays will be done as described previously. Furthermore, the dynamics of H3K4me3 will be assessed in cancer cell lines such as colorectal cancer HCT-116, and breast cancer cell lines MCF-7, and MDA-MB-231 cells. Optimization steps and ChIP assays will be performed as previously discussed.

To determine how H3K4me3 levels are depleted, ChIP assays must be done for KDMs, histone variant H3.3 and H3K4ac. ChIP assays for the variant H3.3. would provide information as to whether the decrease in H3K4me3 levels is due to the turnover of H3 in the histone octamer. H3.3 is replication independent, and would be expected that this variant would be the predominant histone H3 variant to be incorporated into histone octamers during immediate early gene induction. If the H3.3 ChIP assays indicated that H3.3. distribution correlated with the H3K4me3 dynamics, inhibition studies on H3.3 would be done to determine if H3K4me3 levels would remain high once histone turnover was inhibited. To test if KDMs were responsible for the decrease in H3K4me3 levels, a similar experimental design of histone turnover would be utilized. Dual cross-linking ChIP assays for the KDMs would be done to determine the distribution of these enzymes along the genes, followed by H3K4me3 ChIP assays after KDM inhibition treatment. If KDMs were required for H3K4me levels to decrease, it would be expected the KDM inhibition would cause H3K4me3 levels to remain high along a gene. For the final model proposed on the mechanism of how H3K4me3 levels decrease, dual cross-linking KDM ChIP assays would be done alongside histone H3 lysine 4 acetylated ChIP assays. If the final model was the correct model for H3K4me3 depletion, it would be expected that the distribution of histone H3 lysine 4 acetylation would be similar to the distribution of the KDMs along the gene.

Since inhibition splicing directly affects transcription, it is difficult to measure if splicing inhibition does affect H3K4me3 dynamics, or if it is due to inhibition of transcription. To determine if splicing does affect H3K4me3 dynamics, isolation of splicing inhibition from transcription would be required. This is possible with the use of bacterial artificial chromosomes. The bacterial artificial chromosome would contain construct of the *Fos* gene with a mutated branch point,

polypyrimidine region of the intron 1, or mutated splice sites, and would then be transfected into 10T1/2 cells. The mutated splicing regions would prevent splicing from occurring at intron 1, and if splicing inhibition does affect H3K4me3 levels during gene induction, the levels of H3K4me3 would decrease along intron 1, which would be detected through H3K4me3 ChIP assays. Another method to test whether splicing inhibition affects H3K4me3 levels would be the use of morpholino oligomers. The oligomers would be designed to recognize and bind to the branch point or the polypyrimidine tract region in the first intron of the *Fos* gene. This would also prevent the spliceosome formation at intron 1. H3K4me3 ChIP assays would be done along the *Fos* gene containing the morpholino oligomers to determine how splicing inhibition affects H3K4me3 levels.

6. REFERENCES

Allis, C. David, Shelley L. Berger, Jacques Cote, Sharon Dent, Thomas Jenuwien, Tony Kouzarides, Lorraine Pillus et al. "New nomenclature for chromatin-modifying enzymes." Cell 131, no. 4 (2007): 633-636.

Amit, Ido, Ami Citri, Tal Shay, Yiling Lu, Menachem Katz, Fan Zhang, Gabi Tarcic et al. "A module of negative feedback regulators defines growth factor signaling." Nature genetics 39, no. 4 (2007): 503.

Andersson, Robin, Stefan Enroth, Alvaro Rada-Iglesias, Claes Wadelius, and Jan Komorowski. "Nucleosomes are well positioned in exons and carry characteristic histone modifications." Genome research 19, no. 10 (2009): 1732-1741.

Bakiri, Latifa, Rainer Hamacher, Osvaldo Graña, Ana Guío-Carrión, Ramón Campos-Olivas, Lola Martinez, Hans P. Dienes, Martin K. Thomsen, Sebastian C. Hasenfuss, and Erwin F. Wagner. "Liver carcinogenesis by FOS-dependent inflammation and cholesterol dysregulation." Journal of Experimental Medicine (2017): jem-20160935.

Barrett, Angela, Samantha Santangelo, Keith Tan, Steve Catchpole, Kevin Roberts, Bradley Spencer-Dene, Debbie Hall et al. "Breast cancer associated transcriptional repressor PLU-1/JARID1B interacts directly with histone deacetylases." International journal of cancer 121, no. 2 (2007): 265-275.

Benayoun, Bérénice A., Elizabeth A. Pollina, Duygu Ucar, Salah Mahmoudi, Kalpana Karra, Edith D. Wong, Keerthana Devarajan et al. "H3K4me3 breadth is linked to cell identity and transcriptional consistency." Cell 158, no. 3 (2014): 673-688.

Bernstein, Bradley E., Michael Kamal, Kerstin Lindblad-Toh, Stefan Bekiranov, Dione K. Bailey, Dana J. Huebert, Scott McMahon et al. "Genomic maps and comparative analysis of histone modifications in human and mouse." Cell 120, no. 2 (2005): 169-181.

Bieberstein, Nicole I., Fernando Carrillo Oesterreich, Korinna Straube, and Karla M. Neugebauer. "First exon length controls active chromatin signatures and transcription." Cell reports 2, no. 1 (2012): 62-68.

Bonhoure, Nicolas, Gergana Bounova, David Bernasconi, Viviane Praz, Fabienne Lammers, Donatella Canella, Ian M. Willis et al. "Quantifying ChIP-seq data: a spiking method providing an internal reference for sample-to-sample normalization." Genome research 24, no. 7 (2014): 1157-1168.

Borde, Valérie, Nicolas Robine, Waka Lin, Sandrine Bonfils, Vincent Géli, and Alain Nicolas. "Histone H3 lysine 4 trimethylation marks meiotic recombination initiation sites." The EMBO journal 28, no. 2 (2009): 99-111.

Braunschweig, Ulrich, Serge Gueroussov, Alex M. Plocik, Brenton R. Graveley, and Benjamin J. Blencowe. "Dynamic integration of splicing within gene regulatory pathways." Cell 152, no. 6 (2013): 1252-1269.

Cao, Jian, Zongzhi Liu, William KC Cheung, Minghui Zhao, Sophia Y. Chen, Siew Wee Chan, Carmen J. Booth, Don X. Nguyen, and Qin Yan. "Histone demethylase RBP2 is critical for breast cancer progression and metastasis." Cell reports 6, no. 5 (2014): 868-877.

Castellaro, Andrés M., Alfredo Tonda, Hugo H. Cejas, Héctor Ferreyra, Beatriz L. Caputto, Oscar A. Pucci, and German A. Gil. "Oxalate induces breast cancer." BMC cancer 15, no. 1 (2015): 761.

Chen, Kaifu, Zheng Hu, Zheng Xia, Dongyu Zhao, Wei Li, and Jessica K. Tyler. "The overlooked fact: fundamental need for spike-in control for virtually all genome-wide analyses." Molecular and cellular biology 36, no. 5 (2016): 662-667.

Chen, Kaifu, Zhong Chen, Dayong Wu, Lili Zhang, Xueqiu Lin, Jianzhong Su, Benjamin Rodriguez et al. "Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes." Nature genetics (2015).

Chen, Mo, and James L. Manley. "Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches." Nature reviews Molecular cell biology 10, no. 11 (2009): 741-754.

Chiu, Robert, William J. Boyle, Jennifer Meek, Tod Smeal, Tony Hunter, and Michael Karin. "The c-fos protein interacts with c-junAP-1 to stimulate transcription of AP-1 responsive genes." Cell 54, no. 4 (1988): 541-552.

Clouaire, Thomas, Shaun Webb, Pete Skene, Robert Illingworth, Alastair Kerr, Robert Andrews, Jeong-Heon Lee, David Skalnik, and Adrian Bird. "Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells." Genes & development 26, no. 15 (2012): 1714-1728.

Davie, James R., Wayne Xu, and Genevieve P. Delcuve. "Histone H3K4 trimethylation: dynamic interplay with pre-mRNA splicing 1." Biochemistry and Cell Biology 94, no. 1 (2015): 1-11.

Delcuve, Geneviève P., Mojgan Rastegar, and James R. Davie. "Epigenetic control." Journal of cellular physiology 219, no. 2 (2009): 243-250.

Ding, Yong, Michael Fromm, and Zoya Avramova. "Multiple exposures to drought 'train' transcriptional responses in Arabidopsis." Nature Communications 3 (2012): 740.

Edmunds, John W., Louis C. Mahadevan, and Alison L. Clayton. "Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation." The EMBO journal 27, no. 2 (2008): 406-420.

Egelhofer, Thea A., Aki Minoda, Sarit Klugman, Kyungjoon Lee, Paulina Kolasinska-Zwierz, Artyom A. Alekseyenko, Ming-Sin Cheung et al. "An assessment of histone-modification antibody quality." Nature structural & molecular biology 18, no. 1 (2011): 91.

Ferrero, Gabriel O., Marianne L. Renner, Germán A. Gil, Lucia Rodríguez-Berdini, and Beatriz L. Caputto. "c-Fos-activated synthesis of nuclear phosphatidylinositol 4, 5-bisphosphate [PtdIns (4, 5) P2] promotes global transcriptional changes." Biochemical Journal 461, no. 3 (2014): 521-530.

Gómez Acuña, Luciana I., Ana Fiszbein, Mariano Alló, Ignacio E. Schor, and Alberto R. Kornblihtt. "Connections between chromatin signatures and splicing." Wiley Interdisciplinary Reviews: RNA 4, no. 1 (2013): 77-91.

He, Runsheng, and Benjamin L. Kidder. "H3K4 demethylase KDM5B regulates global dynamics of transcription elongation and alternative splicing in embryonic stem cells." Nucleic Acids Research (2017).

Healy, Shannon, Protiti Khan, and James R. Davie. "Immediate early response genes and cell transformation." Pharmacology & therapeutics 137, no. 1 (2013): 64-77.

Healy, Shannon, Protiti Khan, Shihua He, and James R. Davie. "Histone H3 phosphorylation, immediate-early gene expression, and the nucleosomal response: a historical perspective 1 1 This article is part of Special Issue entitled Asilomar Chromatin and has undergone the Journal's usual peer review process." Biochemistry and Cell Biology 90, no. 1 (2012): 39-54.

Hornbeck, Peter V., Bin Zhang, Beth Murray, Jon M. Kornhauser, Vaughan Latham, and Elzbieta Skrzypek. "PhosphoSitePlus, 2014: mutations, PTMs and recalibrations." Nucleic acids research 43, no. D1 (2015): D512-D520.

Howe, Françoise S., Harry Fischl, Struan C. Murray, and Jane Mellor. "Is H3K4me3 instructive for transcription activation?." BioEssays 39, no. 1 (2017): 1-12.

Hu, E. R. D. I. N. G., E. Mueller, S. Oliviero, V. E. Papaioannou, R. Johnson, and B. M. Spiegelman. "Targeted disruption of the c-fos gene demonstrates c-fos-dependent and-independent pathways for gene expression stimulated by growth factors or oncogenes." The EMBO journal 13, no. 13 (1994): 3094.

Hung, Tiffany, Olivier Binda, Karen S. Champagne, Alex J. Kuo, Kyle Johnson, Howard Y. Chang, Matthew D. Simon, Tatiana G. Kutateladze, and Or Gozani. "ING4 mediates crosstalk

between histone H3 K4 trimethylation and H3 acetylation to attenuate cellular transformation." Molecular cell 33, no. 2 (2009): 248-256.

Jahan, Sanzida, and James R. Davie. "Protein arginine methyltransferases (PRMTs): role in chromatin organization." Advances in biological regulation 57 (2015): 173-184.

Jiang, Hao, Abhijit Shukla, Xiaoling Wang, Wei-yi Chen, Bradley E. Bernstein, and Robert G. Roeder. "Role for Dpy-30 in ES cell-fate specification by regulation of H3K4 methylation within bivalent domains." Cell 144, no. 4 (2011): 513-525.

Karin, Michael, Zheng-gang Liu, and Ebrahim Zandi. "AP-1 function and regulation." Current opinion in cell biology 9, no. 2 (1997): 240-246.

Khan, Dilshad H., Carolina Gonzalez, Charlton Cooper, Jian-Min Sun, Hou Yu Chen, Shannon Healy, Wayne Xu et al. "RNA-dependent dynamic histone acetylation regulates MCL1 alternative splicing." Nucleic acids research 42, no. 3 (2014): 1656-1670.

Khan, Dilshad H., Carolina Gonzalez, Nikesh Tailor, Mohammad K. Hamedani, Etienne Leygue, and James R. Davie. "Dynamic Histone Acetylation of H3K4me3 Nucleosome Regulates MCL1 Pre-mRNA Splicing." Journal of cellular physiology (2016).

Khan, Dilshad H., Shannon Healy, Shihua He, Daniel Lichtensztejn, Ludger Klewes, Kiran L. Sharma, Veronica Lau, Sabine Mai, Geneviève P. Delcuve, and James R. Davie. "Mitogen-induced distinct epi-alleles are phosphorylated at either H3S10 or H3S28 depending on H3K27 acetylation." Molecular Biology of the Cell (2017): mbc-E16.

Klein, Brianna J., Lianhua Piao, Yuanxin Xi, Hector Rincon-Arano, Scott B. Rothbart, Danni Peng, Hong Wen et al. "The histone-H3K4-specific demethylase KDM5B binds to its substrate and product through distinct PHD fingers." Cell reports 6, no. 2 (2014): 325-335.

Koga, Mitsunori, Megumi Hayashi, and Daisuke Kaida. "Splicing inhibition decreases phosphorylation level of Ser2 in Pol II CTD." Nucleic acids research (2015): gkv740.

Kouhara, H., Y. R. Hadari, T. Spivak-Kroizman, J. Schilling, D. Bar-Sagi, I. Lax, and J. Schlessinger. "A lipid-anchored Grb2-binding protein that links FGF-receptor activation to the Ras/MAPK signaling pathway." Cell 89, no. 5 (1997): 693-702.

Lau, Lester F., and Daniel Nathans. "Expression of a set of growth-related immediate early genes in BALB/c 3T3 cells: coordinate regulation with c-fos or c-myc." Proceedings of the National Academy of Sciences 84, no. 5 (1987): 1182-1186.

Liu, Qing, and Ming-wei Wang. "Histone lysine methyltransferases as anti-cancer targets for drug discovery." Acta Pharmacologica Sinica (2016).

Lucibello, Frances C., Carolin Lowag, Manfred Neuberg, and Rolf Müller. "trans-repression of the mouse c-fos promoter: a novel mechanism of Fos-mediated trans-regulation." Cell 59, no. 6 (1989): 999-1007.

Mader, Armin W., and Karolin Luger. "Crystal structure of the nucleosome core particle at 2.8 A resolution." Nature: International weekly journal of science 389, no. 6648 (1997): 251-260.

Maze, Ian, Kyung-Min Noh, Alexey A. Soshnev, and C. David Allis. "Every amino acid matters: essential contributions of histone variants to mammalian development and disease." Nature Reviews Genetics 15, no. 4 (2014): 259-271.

Merkofer, Evan C., Peter Hu, and Tracy L. Johnson. “Introduction to cotranscriptional RNA splicing.” Spliceosomal Pre-mRNA splicing: Methods and Protocols (2014): 83-96

Messier, Terri L., Jonathan AR Gordon, Joseph R. Boyd, Coralee E. Tye, Gillian Browne, Janet L. Stein, Jane B. Lian, and Gary S. Stein. "Histone H3 lysine 4 acetylation and methylation dynamics define breast cancer subtypes." Oncotarget 7, no. 5 (2016): 5094.

Monje, Paula, Maria Julia Marinissen, and J. Silvio Gutkind. "Phosphorylation of the carboxyl-terminal transactivation domain of c-Fos by extracellular signal-regulated kinase mediates the transcriptional activation of AP-1 and cellular transformation induced by platelet-derived growth factor." Molecular and cellular biology 23, no. 19 (2003): 7030-7043.

Mosesson, Yaron, Yoav Voichek, and Naama Barkai. "Divergence and selectivity of expression-coupled histone modifications in budding yeasts." PloS one 9, no. 7 (2014): e101538.

Muramoto, Tetsuya, Iris Müller, Giles Thomas, Andrew Melvin, and Jonathan R. Chubb. "Methylation of H3K4 Is required for inheritance of active transcriptional states." Current Biology 20, no. 5 (2010): 397-406.

Ng, Huck Hui, François Robert, Richard A. Young, and Kevin Struhl. "Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity." Molecular cell 11, no. 3 (2003): 709-719.

O'Brien, K., Matlin, A.J., Lowell, A.M. and Moore, M.J., 2008. The biflavonoid isoginkgetin is a general inhibitor of Pre-mRNA splicing. Journal of Biological Chemistry, 283(48), pp.33147-33154.

Patel, Anamika, Valarie E. Vought, Venkatasubramanian Dharmarajan, and Michael S. Cosgrove. "A novel non-SET domain multi-subunit methyltransferase required for sequential nucleosomal histone H3 methylation by the mixed lineage leukemia protein-1 (MLL1) core complex." Journal of Biological Chemistry 286, no. 5 (2011): 3359-3369.

Pavri, Rushad, Bing Zhu, Guohong Li, Patrick Trojer, Subhrangsu Mandal, Ali Shilatifard, and Danny Reinberg. "Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II." Cell 125, no. 4 (2006): 703-717.

Peng, Cong, Weiqi Zeng, Juan Su, Yehong Kuang, Yijin He, Shuang Zhao, Jianglin Zhang et al. "Cyclin dependent kinase 2 (CDK2) is a key mediator for EGF-induced cell transformation mediated through the ELK4/c-Fos signaling pathway." Oncogene 35, no. 9 (2016): 1170.

Salz, Tal, Changwang Deng, Christine Pampo, Dietmar Siemann, Yi Qiu, Kevin Brown, and Suming Huang. "Histone methyltransferase hSETD1A is a novel regulator of metastasis in breast cancer." Molecular Cancer Research 13, no. 3 (2015): 461-469.

Salz, Tal, Guangyao Li, Frederic Kaye, Lei Zhou, Yi Qiu, and Suming Huang. "hSETD1A regulates Wnt target genes and controls tumor growth of colorectal cancer cells." Cancer research 74, no. 3 (2014): 775-786.

Santos-Rosa, Helena, Robert Schneider, Andrew J. Bannister, Julia Sherriff, Bradley E. Bernstein, NC Tolga Emre, Stuart L. Schreiber, Jane Mellor, and Tony Kouzarides. "Active genes are tri-methylated at K4 of histone H3." Nature 419, no. 6905 (2002): 407-411.

Schram, Andrea W., Roy Baas, Pascal WTC Jansen, Anne Riss, Laszlo Tora, Michiel Vermeulen, and H. Th Marc Timmers. "A dual role for SAGA-associated factor 29 (SGF29) in ER stress survival by coordination of both histone H3 acetylation and histone H3 lysine-4 trimethylation." PloS one 8, no. 7 (2013): e70035.

Schuermann, Marcus, Manfred Neuberg, John B. Hunter, Thomas Jenuwein, Rolf-P. Ryseck, Rodrigo Bravo, and Rolf Müller. "The leucine repeat motif in Fos protein mediates complex formation with Jun/AP-1 and is required for transformation." Cell 56, no. 3 (1989): 507-516.

Schwartz, Schraga, Eran Meshorer, and Gil Ast. "Chromatin organization marks exon-intron structure." Nature structural & molecular biology 16, no. 9 (2009): 990-995.

Scott, Linda A., J. Keith Vass, E. Kenneth Parkinson, David AF Gillespie, Joseph N. Winnie, and Bradford W. Ozanne. "Invasion of normal human fibroblasts induced by v-Fos is independent of proliferation, immortalization, and the tumor suppressors p16INK4a and p53." Molecular and cellular biology 24, no. 4 (2004): 1540-1559.

Shilatifard, Ali. "The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis." Annual review of biochemistry 81 (2012): 65-95.

Shinsky, Stephen A., Kelsey E. Monteith, Susan Viggiano, and Michael S. Cosgrove. "Biochemical reconstitution and phylogenetic comparison of human SET1 family core complexes involved in histone methylation." Journal of Biological Chemistry 290, no. 10 (2015): 6361-6375.

Shukla, Sanjeev, and Shalini Oberdoerffer. "Co-transcriptional regulation of alternative pre-mRNA splicing." Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms 1819, no. 7 (2012): 673-683.

Sims, R.J., Millhouse, S., Chen, C.F., Lewis, B.A., Erdjument-Bromage, H., Tempst, P., Manley, J.L. and Reinberg, D., 2007. Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. Molecular cell, 28(4), pp.665-676.

Spies, Noah, Cydney B. Nielsen, Richard A. Padgett, and Christopher B. Burge. "Biased chromatin signatures around polyadenylation sites and exons." Molecular cell 36, no. 2 (2009): 245-254.

Steward, Melissa M., Jung-Shin Lee, Aisling O'Donovan, Matt Wyatt, Bradley E. Bernstein, and Ali Shilatifard. "Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes." Nature structural & molecular biology 13, no. 9 (2006): 852-854.

Teoh, Peik, and Andrew Sharrocks. "WDR5, ASH2L, and RBBP5 control the efficiency of FOS transcript processing." Cellular and Molecular Biology Letters 19, no. 2 (2014): 215-232.

Terrenoire, Edith, Fiona McRonald, John A. Halsall, Paula Page, Robert S. Illingworth, A. Malcolm R. Taylor, Val Davison, Laura P. O'Neill, and Bryan M. Turner. "Immunostaining of modified histones defines high-level features of the human metaphase epigenome." Genome biology 11, no. 11 (2010): R110.

Terzi, Nihal, L. Stirling Churchman, Lidia Vasiljeva, Jonathan Weissman, and Stephen Buratowski. "H3K4 trimethylation by Set1 promotes efficient termination by the Nrd1-Nab3-Sen1 pathway." Molecular and cellular biology 31, no. 17 (2011): 3569-3583.

Thomson, John P., Peter J. Skene, Jim Selfridge, Thomas Clouaire, Jacky Guy, Shaun Webb, Alastair RW Kerr et al. "CpG islands influence chromatin structure via the CpG-binding protein Cfp1." Nature 464, no. 7291 (2010): 1082-1086.

Tilgner, Hagen, Christoforos Nikolaou, Sonja Althammer, Michael Sammeth, Miguel Beato, Juan Valcárcel, and Roderic Guigó. "Nucleosome positioning as a determinant of exon recognition." Nature structural & molecular biology 16, no. 9 (2009): 996-1001.

Tsai, Hsing-Chen, and Stephen B. Baylin. "Cancer epigenetics: linking basic biology to clinical medicine." Cell research 21, no. 3 (2011): 502-517.

Turner, Bryan M. "Nucleosome signalling; an evolving concept." Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms 1839, no. 8 (2014): 623-626.

van Nuland, Rick, Arne H. Smits, Paschalina Pallaki, Pascal WTC Jansen, Michiel Vermeulen, and H. Th Marc Timmers. "Quantitative dissection and stoichiometry determination of the human

SET1/MLL histone methyltransferase complexes." Molecular and cellular biology (2013): MCB-01742.

Vermeulen, Michiel, Klaas W. Mulder, Sergei Denissov, WWM Pim Pijnappel, Frederik MA van Schaik, Radhika A. Varier, Marijke PA Baltissen, Henk G. Stunnenberg, Matthias Mann, and H. Th Marc Timmers. "Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4." Cell 131, no. 1 (2007): 58-69.

Vethantham, Vasupradha, Yan Yang, Christopher Bowman, Patrik Asp, Jeong-Heon Lee, David G. Skalnik, and Brian D. Dynlacht. "Dynamic loss of H2B ubiquitylation without corresponding changes in H3K4 trimethylation during myogenic differentiation." Molecular and cellular biology 32, no. 6 (2012): 1044-1055.

Wong, Ping-Pui, Fabrizio Miranda, KaYi V. Chan, Chiara Berlato, Helen C. Hurst, and Angelo G. Scibetta. "Histone demethylase KDM5B collaborates with TFAP2C and Myc to repress the cell cycle inhibitor p21cip (CDKN1A)." Molecular and cellular biology 32, no. 9 (2012): 1633-1644.

Xiang, Yang, Ziqi Zhu, Gang Han, Xiaolei Ye, Bo Xu, Zhouchun Peng, Yuanjun Ma et al. "JARID1B is a histone H3 lysine 4 demethylase up-regulated in prostate cancer." Proceedings of the National Academy of Sciences 104, no. 49 (2007): 19226-19231.

Yao, Ryoji, and Geoffrey M. Cooper. "Requirement for phosphatidylinositol-3 kinase in the prevention of apoptosis by nerve growth factor." Science 267, no. 5206 (1995): 2003.

Zeng, P., Christopher R. Vakoc, Z. Chen, Gerd A. Blobel, and Shelley L. Berger. "In vivo dual cross-linking for identification of indirect DNA-associated proteins by chromatin immunoprecipitation." Biotechniques 41, no. 6 (2006): 694.