MECHANISMS OF EWING SARCOMA DEVELOPMENT AND THERAPEUTIC RESISTANCE

Nicholas C. Gomez

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Curriculum in Genetics and Molecular Biology.

Chapel Hill
2015

Approved by:

Ian J. Davis

Victoria L. Bautch

Robert J. Duronio

William Y. Kim

Karen L. Mohlke

ABSTRACT

Nicholas C. Gomez: Mechanisms of Ewing sarcoma development and therapeutic resistance.
(Under the direction of Ian J. Davis.)

Many large scale consortia have begun to use high-throughput sequencing technologies to identify key mutations and pathways involved in the etiology of cancers. Identification of specific mutations within tumors can also influence patient-directed therapeutic decisions. Recent evidence implicates the insulin-like growth factor pathway in development of Ewing Sarcoma, a highly malignant bone and soft tissue tumor that primarily affects children and young adults. Despite promising results from preclinical studies of therapies that target this pathway, early phase clinical trials have shown that a significant fraction of patients do not benefit, suggesting that cellular factors determine tumor sensitivity. All cases of Ewing sarcoma are characterized by a translocation between EWSR1 and an ETS-family transcription factor where 85% of the cases result in the EWSR1-FLI1 t(11;22)(q24:q12). This novel transcription factor is retargeted to a subset of repetitive elements in a cell-type specific manner by a previously unknown mechanism. In the chapters that follow, we describe the cellular consequences of PTEN loss in Ewing sarcoma as well as a mechanism for EWSR1-FLI1 retargeting. We demonstrate that Ewing sarcoma cells with PTEN loss exhibited increased transformative properties, as well as reciprocal sensitivity to IGF-1R and mTOR inhibition, therapies currently undergoing testing in clinical trials. In addition, our studies also describe a novel chromatin environment of stem cells in which repetitive elements are enriched in accessible chromatin. Strikingly, a subset of accessible repetitive elements in stem cells is associated with Ewing Sarcoma development. Repetitive elements exhibiting the greatest FAIRE signal in stem cells, is associated

with increased EWSR1-FLI1 binding, suggesting the chromatin environment of stem cells is primed for oncogenesis. Taken together, these studies expand our knowledge of Ewing sarcoma etiology and provide a potential mechanism for therapy resistance. This work serves as the foundation for both expanded preclinical and clinical research for future targeted therapies in Ewing sarcoma.

ACKNOWLEDGMENTS

This work would not have been possible without the support and guidance from my advisor Ian J. Davis. Your mentorship has shaped every aspect of my scientific development. I am indebted to you for the countless discussions and advice over the years. Thanks for the trust and allowing me to bring a new model to the lab. Thanks for creating such a supportive, yet scientifically rigorous lab environment. As I progress through my career I hope that I can emulate the enthusiasm and optimism that you demonstrate even through the toughest of problems.

I would also like to thank the members of the Davis Lab, Mukund Patel, Jeremy Simon, Luke Roode, Jake Troutman, Fang Fang, Andrew McFadden, Aminah Wali, Catherine Fahey, Mariesa Slaughter, and Austin Hepperla. You all have created a unique environment that fosters both intense scientific discussion as well as camaraderie. I thank you for all of your input, ideas, and assistance throughout this entire process. Graduate school was made infinitely times better with the “joke of the day”, excursions to haunted houses, margarita lunches, holiday parties, and genuine friendship.

I am also grateful to the many friends that I have made along way. Specifically, I want to thank Sean Bailey, Daniel Dominguez, Brian Gibbs, Tikvah Hayes, and Les Freeman for always being available to discuss and celebrate science over drinks. Tiki, I’m appreciative of your unique ability to force me to keep moving forward and to always be thinking about the next step.

I am also appreciative of the many people that have mentored me throughout my graduate career. To my committee members, Vicki Bautch, Bob Duronio, Billy Kim, and Karen Mohlke, your scientific and career advice has extremely helpful. Your ideas

and critiques of my science have made me a stronger scientist. To Vicki Bautch, you believed in me from day 1, when I interviewed for graduate school- your support and belief in me has been a significant driver of my confidence. I also want to thank the Initiative for Maximizing Student Diversity and BBSP, specifically Pat Phelps, and Ashalla Freeman for creating an excellent training environment that was both supportive and instrumental to fostering my scientific growth.

Lastly I want to thank my family. To my parents, Frank and Nelly, thanks for instilling confidence and providing a loving and supportive environment and always encouraging me to dream big. I am forever indebted to you and hope to pass these same values to my children someday. To my wife Megan, I appreciate your enormously supportive and positive attitude through both the good and tough times. Without you, I never would have even come close to finishing graduate school. I can be a handful, especially after procrastinating, and am truly thankful for everything you do. To my brother Alex, thanks for always being available to chat and just disconnect from science over video games. Your support and belief in me since we were kids has really propelled me through this process. Lastly, but not least, I want to thank Trotter who always was willing to listen to my troubleshooting. I love you all!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ATAC	Assay for Transposase-Accessible Chromatin
BAC	Bacterial Artificial Chromosome
BM-MSC	Bone Marrow derived Mesenchymal Stem Cells
ChIP	Chromatin Immunoprecipitation
DNase	Deoxyribonuclease
ENCODE	Encyclopedia of DNA Elements
ESC	Embryonic Stem Cell
FAIRE	Formaldehyde-Assisted Isolation of Regulatory Elements
H1-MSC	H1 differentiated Mesenchymal Stem Cells
H&E	Hematoxylin and Eosin
hESC	Human Embryonic Stem Cell
HKC	Human Kidney Cells
HUVEC	Human Umbilical Vein Endothelial Cells
IHC	Immunohistochemistry
IGF1	Insulin-like Growth Factor 1
IGF1R	Insulin-like Growth Factor 1 Receptor
iPSC	Induced Pluripotent Stem Cell
LINE	Long Interspersed Nucleotide Element
MNase	Micrococcal Nuclease
MPC	Multipotent Progenitor Cell
MSC	Mesenchymal Stem Cells
mTOR	Mechanistic Target of Rapomyocin
RPKM	Reads Per Kilobase exon per Million mapped reads
SINE	Short Interspersed Nucleotide Element
TSS	Transcriptional Start Site

CHAPTER 1

INTRODUCTION

Advances in sequencing technologies have led to many discoveries ranging from the understanding of genome-wide regulatory elements such as enhancers [1] to the associations of complex diseases with specific mutations and genomic loci [2][3]. The human genome sequencing project was declared complete in 2003 and has served as the foundation for genomics, a new era in the biological sciences. Advances in sequencing technology have led to sequencing being coupled to previously known biochemical techniques such as chromatin immunopreciptiation (ChIP) [4][5][6], DNaseI hypersensitivity (DNase) [7][8], Micrococcal nuclease digestion (MNase) [9], and RNA isolation [10][11][12]. The field of genomics began to rapidly expand as high-throughput sequencing allowed researchers to analyze information not just for a handful of loci, but now genome-wide, increasing exponentially the amount of available information. The wealth of data gave rise to bioinformaticians who developed new tools and algorithms to handle the massive datasets. Large multi-institutional consortia, such as the Encyclopedia of DNA Elements (ENCODE) and the Epigenome Roadmap, were created and tasked with identifying and understanding the functional elements of normal and disease states of human cells [13][14]. These consortia have directly led to creation of terabytes of publically available data that can be mined by researchers all of the world to answer specific biological questions.

1.1 Embryonic stem cells are derived from the Inner cell mass

Human embryonic stem cells (hESC) are derived from the inner cell mass of the blastula. Post-fertilization, the zygote undergoes rapid cell division transitioning through several stages before becoming the blastocyst around day five. The blastocyst consists of two cellular grups an outer lining of trophoblasts and an inner cell mass. The trophoblasts will eventually go on to form extra-embryonic tissues such as the placenta, chorion, and umbilical cord, while the inner cell mass will form the entire embryo. In 1981, a seminal discovery in stem cell research was made when the first pluripotent cells were isolated from the mouse blastocyst [15][16]. 17 years later the first human ESC (hESC) lines were established [17]. hESCs are unique in the ability to be cultured indefinitely, while still retaining pluripotency: the ability to differentiate into cell types from all three of the embryonic germ layers, endoderm, mesoderm, and ectoderm [17](Figure 1.1). Using well defined chemicals and transcription factors, researchers have been able to differentiate into neurons [18], cardiomyocytes[19][20], and muscle [21] among others [22]. The ethical issues surrounding the study of human development *in-vivo* have prevented advances in understanding of these processes. With the establishment of hESC, these cell lines, and the accompanying ability to differentiate are excellent surrogates for the study of developmental processes including chromatin changes.

1.2 Chromatin Organization

With completion of the human genome project, a new challenge developed - the identification of functional elements in the genome. Of the 3.3 billion bases in the human genome, only 3% encode proteins. The remainder has commonly been referred to as “junk” DNA [23]. Though non-coding, recent studies have began to show the importance and functionality of these regions in evolution [24][25], transcription fac-

tor binding[26][27], and disease risk [28]. Interestingly, nearly 50% of the noncoding human genome, is comprised of repetitive elements [29][30], suggesting a potential importance of the elements in human biology.

Human repetitive elements are often classified into two broad categories, short tandem repeats (also known as microsatellites), represented by simple repeats, and interspersed repeats represented by short and long interspersed nuclear elements (SINEs and LINEs respectively). In addition to SINEs and LINEs, interspersed repeats also consist of long terminal repeats (LTR) and endogenous retroviruses (ERV). These are referred to as retrotransposons, as they have acquired a mechanism to copy and insert themselves into the genome through an RNA intermediate. Retrotransposons have been demonstrated to be active a variety of adult tissues as well as during embryonic development [31][32]. Repetitive elements are also able to in-

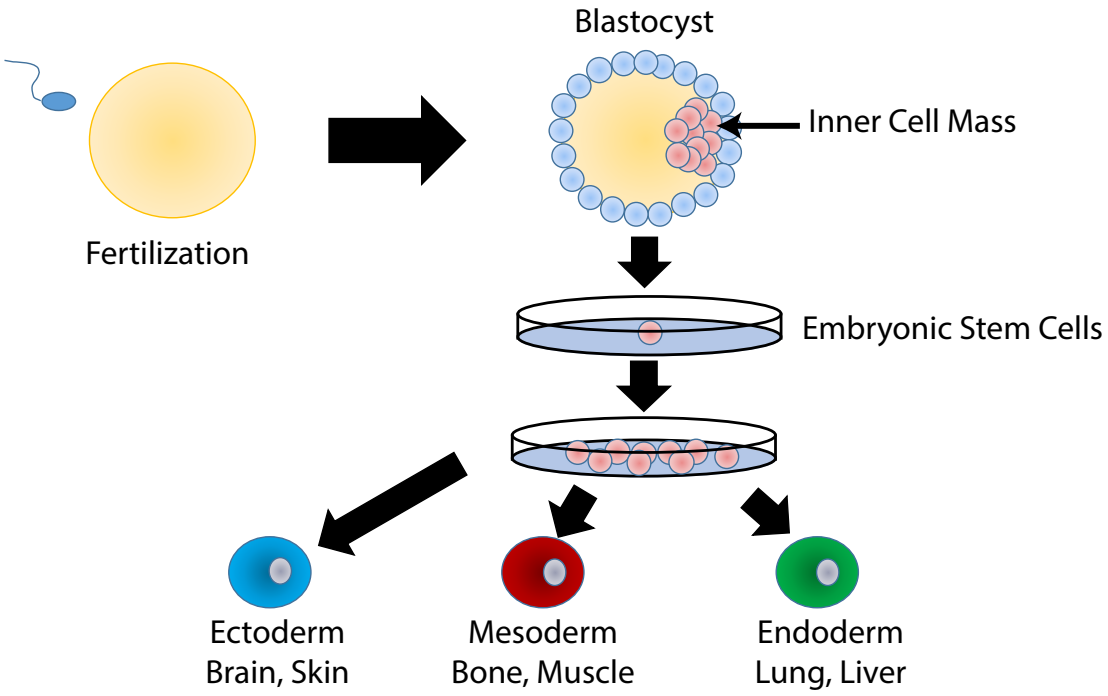


Figure 1.1: Post-fertilization rapid cell division occurs ultimately forming the blastocyst. Embryonic stem cells are derived from the cells of the inner cell mass and can be cultured indefinitely while retaining the ability to differentiate into cell types from the three embryonic germ layers endoderm, mesoderm, and ectoderm.

fluence gene expression by acting as cis-regulatory elements and can also recruit transcription factors [33][34]. Taken together, these data are beginning to shed light on the importance of the noncoding genome and repetitive regions in particular.

In addition to the increased study of functional elements in the genome, advances in sequencing technology have identified new ways to study DNA organization. DNA does not exist as a naked double helix, rather it interacts with proteins. The mixture of proteins and DNA in the nucleus is known as chromatin. Chromatin can be visualized microscopically as tightly compacted chromosomes in dividing cells [35] to diffuse distributed strands with dense heterochromatic regions during interphase [36]. The fundamental protein components comprising chromatin are histones. Core histones, H2A, H2B, H3, and H4, are among the most highly conserved eukaryotic proteins known, suggesting that chromatin evolved from a common ancestor of eukaryotes [37].

The nucleosome, the principal subunit of chromatin, consists of the DNA polymer wrapping twice, apporiximately 147 base pairs (bp), around the combination of two H2A, H2B, H3, and H4 proteins[38][39]. Individual nucleosome core particles are connected by short segments of "linker" DNA ranging from 10-80 bp. Histone H1 binds to the linker region, stabilizing the connecting nucleosome core particles. These fibers can then be folded into a 30-nm structure which results in approximately 50-fold compaction[37]. The polynucleosome structures can be further compacted into higher order structures such as the mitotic chromosome. This compaction, initially started by the nucleosome core particle, allows for nearly two meters of DNA to fit within a six-micron nucleus.

Chromatin, not naked DNA, is the predominant unit within the nucleus. Therefore, nucleosomes serve a role beyond compaction by regulating access to DNA templated processes such as gene expression. In most cases, the eviction, destabilization, or repositioning of nucleosomes is necessary to facilitate access to the underlying DNA.

Chromatin organization and rearrangement has been an intensely studied process, resulting in the elucidation of three general mechanisms. First, chromatin remodeling can occur by specific protein complexes whose tasks are to slide and evict nucleosomes [40]. Second, various post-translational modifications, including acetylation, methylation, ubiquination, phosphorylation, and others, can occur at the histone tails creating biochemical signposts that can act as binding sites for the recruitment of specific proteins [41]. Lastly, alternative histones, such as H2A.Z or H3.3, may replace one or more of the core histones creating a destabilized nucleosome leading to a more permissive environment for regulatory factor binding [42][43].

1.3 Epigenetics and differentiation of ESCs

Histone modifications and variants can persist across generations leading to them being termed “epigenetic”, yet this classification remains controversial [44]. Epigenetic modifications are responsible for the incredible amount of cell-type diversity. Each nucleated cell within the human body contains of the exact same genetic material, yet the chromatin landscape is organized differently in each cell type. This altered organization regulates access of the underlying DNA to transcription factors. The chromatin environment is coordinated by host of factors ranging from DNA methylation to covalent histone modifications. Many of the proteins responsible for regulating chromatin are essential to development, as knockouts in mice have lead to embryonic lethality [45][46][47].

Acetylation, a histone post-translational modification that decreases the affinity of the protein-DNA interaction, is associated with active regions of chromatin by increasing access for regulatory factors [48]. The steady state of histone acetylation is coordinated by both the addition and removal of acetylation by histone acetyltransferases (HAT) and histone deacetylases (HDACS) respectively. In ESCs HATs have been demonstrated to directly regulate the pluripotency transcription factors, OCT4

and NANOG, as well as their downstream targets. Knockdown of specific HATs results in loss of pluripotency, suggesting that HATs, and acetylation, are important for maintaining the pluripotent chromatin environment of the ESC [49][50][51].

Recent evidence has demonstrated that the chromatin environment of ESCs is distinct from differentiated cells. Microscopy of both ESC and differentiated cell nuclei has revealed that ESCs contain less heterochromatic regions [36][52]. Consistent with more "open" chromatin by microscopy, biochemical analysis of ESCs have demonstrated higher levels of acetylated histones with concomitant decrease in repressive histone modifications relative to differentiated cells [5][53]. Consequentially, ESC also demonstrate an increase in basal gene expression relative to differentiated cells [54].

Chromatin remodeling is necessary for proper differentiation [55]. Evidence suggests that stem cell chromatin appears poised for differentiation into various lineages as developmentally important gene promoters as well as enhancers, are bivalent in ESCs. Bivalent regions are marked concomitantly by active and repressive chromatin associated histone modifications [56] [57][58]. As cells terminally differentiate the majority of bivalent genes become monovalent or lose the marks altogether, suggesting an importance in multipotency [59]. In addition, during terminal differentiation, the chromatin environment undergoes an overall increase in H3K9 dimethyl and H3K9 trimethyl, histone marks associated with repressive chromatin, demonstrating a general increase in compaction of the chromatin environment [60][53].

1.4 Stem cells and cancer

Mesenchymal stem cells (MSC) were first identified by Alexander Friedenstein in the 1960's. He detailed a population of cells from mouse bone marrow that could form connective tissue and bones [61]. In 1999, Pittenger et al. demonstrated that human bone marrow contained a distinct population of cells which had the mulitpotent

capacity to differentiate into adipocytes, osteoblasts, and chondrocytes [62]. In vitro these cells are characterized by their ability to adhere to plastic, positive expression of cell surface markers CD73, CD90, CD105, and negative expression of a range of hematopoietic specific cell surface markers including CD34, and CD45 [63][64].

Sarcomas, distinct from carcinomas, are a heterogeneous group of tumors thought to arise from mesenchymal tissues such as bone, fat, muscle, and cartilage [65]. Sarcomas are broadly classified into two groups with distinct features. The first, which includes Ewing sarcoma, is characterized by by the presence of specific translocations while the other group, represented by tumors such as osteosarcoma, is defined by complex karyotypes indicative of chromosomal instability [66][67].

Many studies have begun to link MSCs with the origins of sarcomas. Comparing gene expression profiles of various sarcomas to that of MSCs have demonstrated similarities [68][69]. One study found a correlation in gene expression between stage of MSC adipocyte differentiation and specific subtypes of liposarcoma [70]. Further evidence implicating stem cells as the origin for sarcomas comes from the observation of spontaneous transformation in mouse MSCs [71][72]. Though human MSCs have not been shown to undergo spontaneous transformation [67], transformation can occur upon ectopic expression hTERT, with cells acquiring anchorage independent growth and the ability to form tumors in mice [73][74]. Due to the tissues for sarcomas being mesenchymal in origin, and the data linking MSCs with transformation, it is intriguing to consider MSCs as the origin for many sarcomas.

1.5 Ewing sarcoma

Ewing sarcoma was first described by James Ewing in 1921 as a diffuse endothelioma of the bone that, unlike osteosarcoma, was sensitive to radiation [75]. This disease, Ewing sarcoma, the second most common bone and soft tissue malignancy in children and young adults has an incidence of 2.93 cases per million [76]. Though

rare, Ewing sarcoma is an aggressive disease where 28% of patients present with metastasis and treatment consists of intensive chemotherapy, radiation, and surgery. Nearly 50% of patients will ultimately succumb to the disease within 10 years [76]. The identification of a characteristic fusion between the genes EWSR1 and FLI1 (t(11;22)(q24:q12)) in 1992 greatly enhanced our knowledge of the etiology of this disease and served as foundation for future research projects [77][78]. The translocation results in a novel transcription factor with transformative potential. Since that discovery, all cases of Ewing sarcoma are defined by the presence of a fusion between EWSR1 and an ETS family transcription factor where 85% of the cases are characterized by the specific fusion of EWSR1-FLI1.

Ewing sarcoma had been traditionally classified histologically as a poorly differentiated small round blue cell tumor and was once thought to arise from primitive neuroectodermal cells [79]. Early neural markers as well as neural structural features have been observed in a subset of primary tumors [80][81]. In addition, expression of EWSR1-FLI1 in neuroblastoma cell lines gives cells characteristics of Ewing sarcoma [82].

Despite this evidence suggesting a neuronal precursor, the origin of Ewing sarcoma remains highly controversial. Neural features of Ewing sarcoma are found only in a minority of patients. In addition, EWSR1-FLI1 expression induces neuroectodermal differentiation and activates genes associated with neural differentiation suggesting that the neural features may be a consequence of EWSR1-FLI1 expression rather than features of the cell of origin [83][84].

An alternative hypothesis suggest that Ewing sarcoma develops from a MSC. Knockdown of EWSR1-FLI1 in Ewing sarcoma cells leads to many similarities with MSCs such as gene expression, capacity for multi-lineage differentiation, and expression of MSC cell surface markers [68]. Additionally, a subpopulation of cells, isolated from primary tumors and characterized by CD133 expression, has been demon-

strated to differentiate into adipocytes, osteoblasts, and chondrocytes, a hallmark of MSCs [85].

EWSR1-FLI1 is aberrant transcription factor that is necessary but not sufficient for Ewing sarcoma tumor development. Studies exogenously expressing EWSR1-FLI1 in normal murine and human fibroblasts fail to transform, but instead undergo growth arrest [86][87]. However, both domains of the fusion protein are critical as disruption of either the EWSR1 or FLI1 impairs the functional activity of this protein [78][88]. EWSR1-FLI1 is retargeted to regions distinct from the parental protein FLI1 in a cell-type specific manner [89][90]. Most of EWSR1-FLI1 binding sites act as a transcriptional activator of nearby genes that are important for Ewing sarcoma including IGF1 [91][89][92][93][94]. In addition, binding of EWSR1-FLI1 is able to change the local chromatin environment by increasing FAIRE signal at these target regions [90]. EWSR1-FLI1 binding sites in the cancer are also accompanied by increases in histone post-translational modifications associated with enhancer activity [90][95].

1.6 Chromatin accessibility by FAIRE-seq

Traditionally, chromatin accessibility assays have performed by enzymatic cleavage of DNase in a variety of organisms, including plants [96][7][97][98][99]. An alternative to DNase is formaldehyde assisted isolation of regulatory elements (FAIRE). Instead of relying on DNase-I specific enzymatic cleavage of “open” chromatin, FAIRE utilizes an organic extraction that enriches for a variety regulatory regions in chromatin that are “active” or available and consequently thought to be nucleosome depleted [100][101]. Though both techniques assay chromatin accessibility, FAIRE is more efficient at identifying distal regulatory elements [102]. Applying this technique, we can compare the chromatin environment of stem, differentiated, and cancer cells to gain insight into developmental and disease processes.

1.7 Thesis contributions

The experiments described here show how sequencing serendipitously identified a loss of tumor suppressor and how this loss effected response to therapies (Chapter 2). In addition, the experiments described here will demonstrate that stem cells have a unique chromatin environment that can be exploited by the oncoprotein EWSR1-FLI1 (Chapter 3). The results from these studies have contributed to our understanding of both cancer and stem cell biology as well as providing a potential mechanism for why specific therapies fail a subset of patients (Chapter 4).

This work has been both collaborative and highly interdisciplinary spanning biochemistry, genomics, molecular and cellular biology. In Chapter 2, the Flourescent in-situ hybridization (FISH), colony formation, and chloroquine treaments were conducted by Mukund Patel and Andrew McFadden. I performed the OSI-906 inhibitor experiments, and analyzed the immunohistochemistry (IHC) from the tumor microarray as well as the autophagy results.

In Chapter 3, I performed and anlayzed the FAIRE-seq, RNA-seq, Salt-fractionation, Immunoblot, and flow cytometry for the cell lines tested. Raluca Dumitru cultured and differentiated the hESC. Jeremy Simon, Austin Hepperla, and Fang Fang, assisted with bioinformatics analyses by creating scripts and pipelines related to the project.

CHAPTER 2

WIDESPREAD CHROMATIN ACCESSIBILITY AT REPETITIVE ELEMENTS LINKS STEM CELLS WITH HUMAN CANCER

2.1 Introdution

Early mammalian development necessitates precisely regulated transcriptomic and chromatin changes as cells commit to their terminal fates [156][157][158]. A comprehensive understanding of chromatin remodeling during differentiation may reveal biological pathways that regulate this process and could suggest therapeutic opportunities relevant to cancer-directed and regenerative medicine. Human embryonic stem cells (hESCs), derived from the inner cell mass of human blastocysts, can propagate in vitro and are able to undergo multi-lineage differentiation [17].

Previous studies have explored chromatin dynamics during stem cell differentiation by comparing hESCs to differentiated cells. hESCs are characterized by elevated levels of activation-associated histone post-translational modifications, histone bivalency at developmentally regulated genes, and increased expression of variant histones [56][159][5][53]. Though insightful, histone modification changes represent one of multiple strategies that ultimately regulate the chromatin landscape.

In an effort to comprehensively explore the changes in chromatin organization that accompany differentiation, we utilized FAIRE-seq, an unbiased biochemical assay that enriches for localized regions of nucleosome-depleted (“open”) chromatin [101][100]. Regions identified by FAIRE-seq contain a broad range of regulatory

classes. We applied this technique to compare the chromatin landscape of hESC, primary and in vitro differentiated mesenchymal stem cells and mature cell lines. We identified increased chromatin accessibility at specific classes of repetitive elements in stem cells. These regions harbored distinct histone modifications and underwent chromatin remodeling during differentiation. Chromatin accessibility at repetitive elements in stem cells offered a permissive environment that could be exploited by an oncogenic transcription factor.

2.2 Results

2.2.1 FAIRE selected chromatin from human embryonic stem cells is dominated by repetitive elements

To explore chromatin organization in human embryonic stem cells, we performed FAIRE-seq on undifferentiated H1-ESC (WA01), H7-ESC (WA07), and H9-ESC (WA09) cells. Using bowtie [160], we aligned sequencing reads to the human genome, as previously described [161]. Consistent with published results in a wide range of cells, FAIRE signal was enriched at transcriptional start sites (TSS) and CTCF binding sites in all hESC (Figure 2.1 A)[161]. We also observed signal enrichment at OCT4 and NANOG binding sites, factors critical for the maintenance of pluripotency (Figure 2.1 A) [162][163].

We then identified genomic regions that were unique to stem cells. We compared z-score-transformed FAIRE signal in 500 bp windows to publicly available data from three differentiated cell types (HUVEC, K562, and NHEK) [164]. Of the regions that passed a minimum signal filter, we identified those that demonstrated a significant difference between hESC and the three differentiated cell types ($p \leq 0.01$, t-test). Hierarchical clustering resolved these 12,026 regions into two major groups (Figure 2.2 A). Cluster 1 (C1) consisted of regions with increased FAIRE signal in hESCs, and cluster 2 (C2) contained regions with higher signal in the differentiated cell lines (Fig-

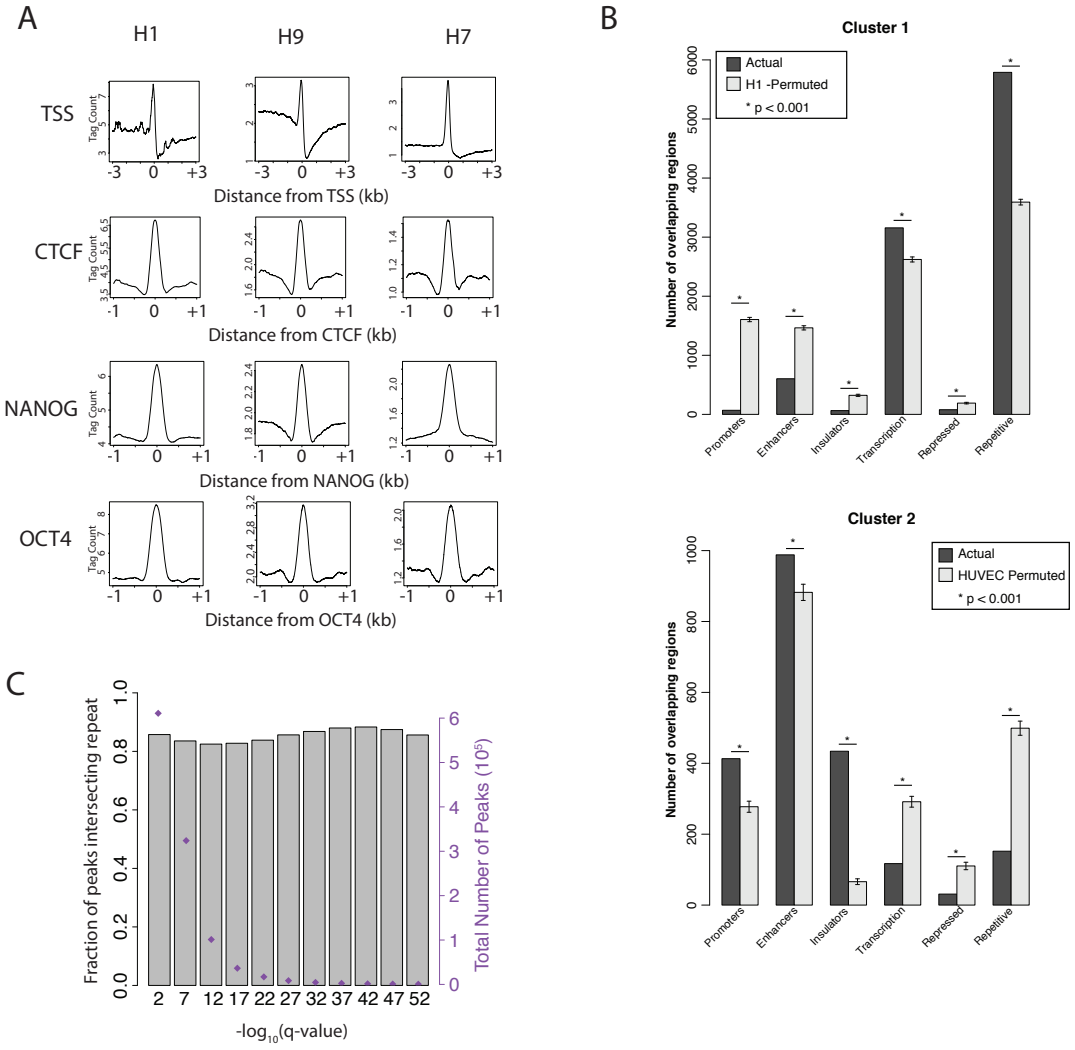


Figure 2.1: ESC FAIRE-seq displays expected signal characteristics as well as enrichment of repetitive elements. (A) FAIRE-seq signal demonstrates enrichment at known regulatory elements. Average FAIRE-seq signal in H1, H9, H7 at transcription start sites (TSS), CTCF, NANOG, and OCT4 binding sites. (B) Repetitive and Transcription categories are specifically enriched in Cluster 1. Permutation results of equal number of regions for either C1 and C2 and their association with ChromHMM classifications. Error bars represent standard deviations of the permuted sets. (C) Repetitive element enrichment is independent of q-value threshold. Enriched regions defined by MACS2 were filtered with increasing q-values and intersected with repetitive elements. Fraction of regions intersecting a repetitive element (gray bar) and number of regions exceeding selected q-values (black box) are shown.

ure 2.2 A and B). The two clusters demonstrated significant differences in location. C1 regions were primarily distal, with a median distance of 39.5 Kb to the nearest TSS. Whereas C2 regions were primarily proximal, with a median distance of 11.4 Kb (Figure 2.2 A). We then annotated the genomic intervals with classifications previously generated by segmentation analyses in H1-ESC, HUVEC, K562, and NHEK (ChromHMM) [164][165][166]. Briefly, segmentation analyses employs a set of histone modifications to develop a classification scheme for genomic regions linked to recognized function. C1 was significantly enriched for transcription and heterochromatic/repetitive states ($p < 0.001$, Figure 2.2 A and Figure 2.1 B). In contrast, C2 was enriched for states such as active and poised promoters, as well as insulators. ($p < 0.001$, Figure 2.2 A and Figure 2.1 B). Interestingly, despite the striking difference in FAIRE signal between cell types, regions in these clusters were similarly classified. Taken together, these data revealed widespread accessible chromatin in stem cells at genomic regions classified as heterochromatic.

To further characterize FAIRE selected chromatin, we identified 610,887 regions with significant signal enrichment (peaks) in H1-ESC, 243,467 in H7-ESC, and 384,162 in H9-ESC (MACS2) [167]. Using a false discovery rate threshold, we selected the top 150,000 peaks for further analysis. The filtered regions were then intersected with repetitive elements defined by RepeatMasker [168], requiring that the site of greatest FAIRE signal was within one bp of a repetitive element. Strikingly, we found that over 82.9%, 94.6% and 94.0% of peak summits identified in H1-ESC, H9-ESC, and H7-ESC, respectively, intersected a repetitive element. The degree of overlap for each hESC was significantly greater than HUVEC, NHEK, K562, and a randomly permuted peak set (Figure 2.2 C). Varying the stringency parameters used to select peaks had no effect on fractional overlap (Figure 2.1 C).

2.2.2 Simple repeats and SINEs are enriched in FAIRE selected chromatin

We then assessed whether the enrichment of repetitive elements was restricted to specific classes. Simple repeats and short interspersed nucleotide elements (SINEs) were selectively enriched among FAIRE peaks in hESC, relative to their genomic prevalence. This pattern was not observed in the three differentiated cell types (Figure 2.2 D).

In further support of repetitive element enrichment, we also found that a large fraction of sequencing reads from each hESC line was discarded during alignment due to redundant genomic mapping (Table 2.1). 83% of unaligned sequences from H1-ESC were repetitive in nature, enriched for SINEs, simple repeats, and LINEs (Table 2.2). In contrast, similar analysis of HUVEC FAIRE identified only 51% of discarded reads as repetitive sequences, a fraction consistent with the abundance of these elements genome-wide.

We then assayed FAIRE signal differences in each repeat class. After normalizing for repeat length and sequencing depth, signal in hESC at simple repeats and SINEs greatly exceeded that of differentiated control cells (Figure 2.2 E). In contrast, signal differences at LINEs were minimal and DNA transposons demonstrated an inverse relationship. Taken together, read, signal and peak-based detection approaches consistently identify the selective enrichment of simple repeats and SINEs by FAIRE in hESCs.

2.2.3 FAIRE-enriched repeats are shared across hESCs and exhibit distinguishing features

Simple repeats and SINEs consist of thousands to millions of individual regions. Only a small fraction was identified by FAIRE ($<$5% of each class, Table 2.3). We then asked if these regions were consistently identified across hESC lines. Among simple repeats and SINEs with signal in the top quartile a large fraction were shared

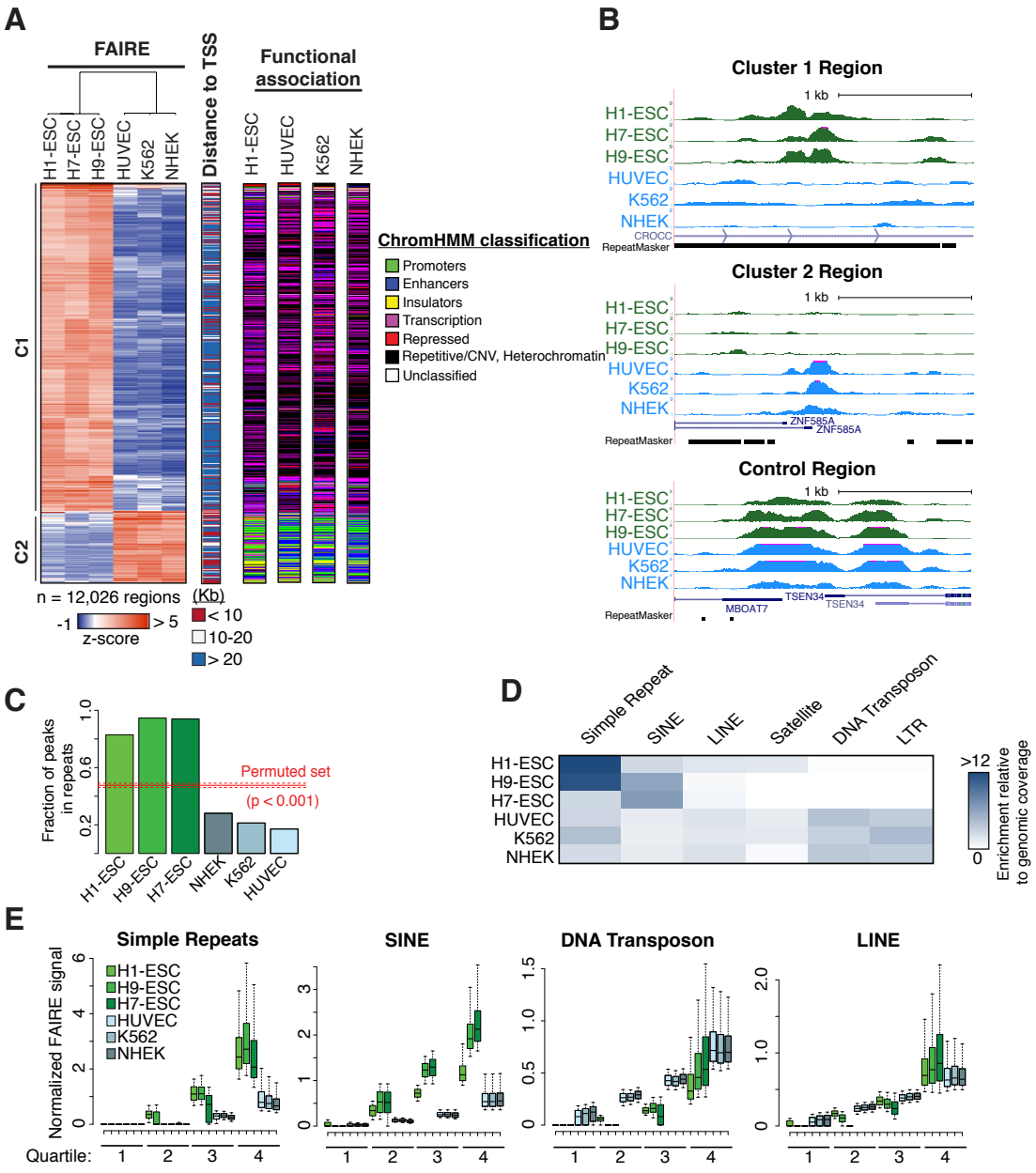


Figure 2.2: FAIRE-seq enriched regions specific to hESC are dominated by simple repeats and SINEs. (A) Heatmap of those regions with significantly different FAIRE enrichment between hESC and control HUVEC, K562, NHEK (500 bp windows, $p \leq 0.01$, t-test, $\text{row}_{max} - \text{row}_{min} > 3$). Regions were assigned classes based on distance to nearest TSS ($<$10 Kb red, 10-20 Kb white, $>$20 Kb blue and by segmentation analysis [164]. (B) Genome browser images of representative regions selected from Cluster 1 (top panel), Cluster 2 (middle panel), and a Control region (bottom panel). (C) Fraction of top 150,000 peak summits that overlapped a repetitive element in hESCs (green) and controls (blue). Fractional overlap with an H1-ESC permuted peak set (red line) and standard deviation (dashed lines) are shown. (D) Heatmap depicting the enrichment of specific classes of repetitive elements in MACS2-identified FAIRE-enriched regions, relative to genomic coverage. (E) Normalized FAIRE signal at Simple Repeats, SINEs, DNA Transposon, and LINE in hESC (green) and control cell lines (blue) plotted by quartile.

Sample	Total Raw Reads (10^6)	Total Reads Passing Tagdust (10^6)	Median Read Quality Score	Reads lost to multiple alignment (10^6)	% lost to multiple alignments	Total Aligned Reads (10^6)	Percent of Raw Reads Successfully Aligned	Percent of TagDust filtered Reads Successfully Aligned
H1-hESC FAIRE Rep 1	86.19	83.83	36.4	23.55	27.32	49.20	57.1	58.7
H1-hESC FAIRE Rep 2	86.44	83.17	36.3	22.11	25.58	50.50	58.4	60.7
H9-hESC FAIRE	80.35	79.96	37.8	23.05	28.69	44.10	54.9	55.2
H7-hESC FAIRE	54.27	54.22	35.8	17.93	33.04	27.59	50.8	50.9
HUVEC FAIRE Rep1	58.84	58.69	29.9	15.65	26.67	39.53	67.2	67.4
HUVEC FAIRE Rep2	48.99	48.77	29.5	12.66	25.96	33.62	68.6	68.9
NHEK FAIRE REP1	66.30	65.97	29.42	11.71	17.75	51.83	78.2	78.6
NHEK FAIRE Rep2	129.99	127.39	29.66	22.74	17.85	98.98	76.2	77.7
K562 FAIRE Rep1	64.28	64.23	27.19	13.50	21.02	48.87	76.0	76.1
K562 FAIRE Rep 2	64.17	64.12	28.51	15.28	23.83	46.84	73.0	73.1

Table 2.1: Read Count and Mapping Statistics

by at least one other hESC line (simple repeat: 48%, 86949 of 179379; SINEs: 59%, 423,819 / 723415; $p < 0.001$ by permutation) (Figures 2.3 A and 2.4 A Figure 2.3 D and 2.4 C). Consistent with a central role of the repetitive segment in mediating chromatin state, we found that for both simple repeats and SINEs FAIRE signal was centered at the repetitive element, rather than extending from flanking regions, and was concordant between the stem cells (Figures 2.3 B, E and 2.4 B and D).

To explore other factors that may influence chromatin status, we asked whether

Sample	Total Length of Unalinged reads (10^6bp)	% of sequence SINE	% of sequence LINE	% of sequence LTR	% of sequence DNA	% of sequence simple repeat	% of sequence Other	Total repetitive bases (10^6bp)	% Repetitive
H1-hESC	1873.89	40.50	21.46	1.82	0.09	9.26	9.98	1557.63	83.1
HUVEC	687.62	22.78	17.05	6.19	1.18	2.29	3.95	367.45	53.4

Table 2.2: Percentage of Masked Bases from Unaligned Reads.

length and G/C content distinguish those repeats that are FAIRE-enriched. For simple repeats, the lengths of FAIRE-enriched and FAIRE-negative sites varied little. However, enriched sites demonstrated a significant skew towards higher G/C content (Figure 2.3 C). The opposite pattern was observed for SINEs. FAIRE-enriched SINEs were significantly longer than others whereas G/C content differed only slightly (Figure 2.3 F). Overall, FAIRE identifies repetitive elements that are common to multiple hESCs and demonstrate shared chromatin patterns and distinct class-specific DNA features.

2.2.4 FAIRE and DNase differ at repetitive regions

Given the abundance of FAIRE-enriched repeats, we were surprised that chromatin accessibility at these sites had not been previously observed. As a comple-

Repeat Class	Total Elements in Repeat Masker	# Found in Peaks	% Found in Open Chromatin
Simple Repeat	417913	16975	4.062
SINE	1793723	26304	1.466
LINE	1498690	28151	1.878
Satellite	9566	529	5.530
DNA Transposon	461751	579	0.125
LTR	717656	544	0.076

Table 2.3: Percentage of Repeats Found in H1-ESC FAIRE (+) Chromatin

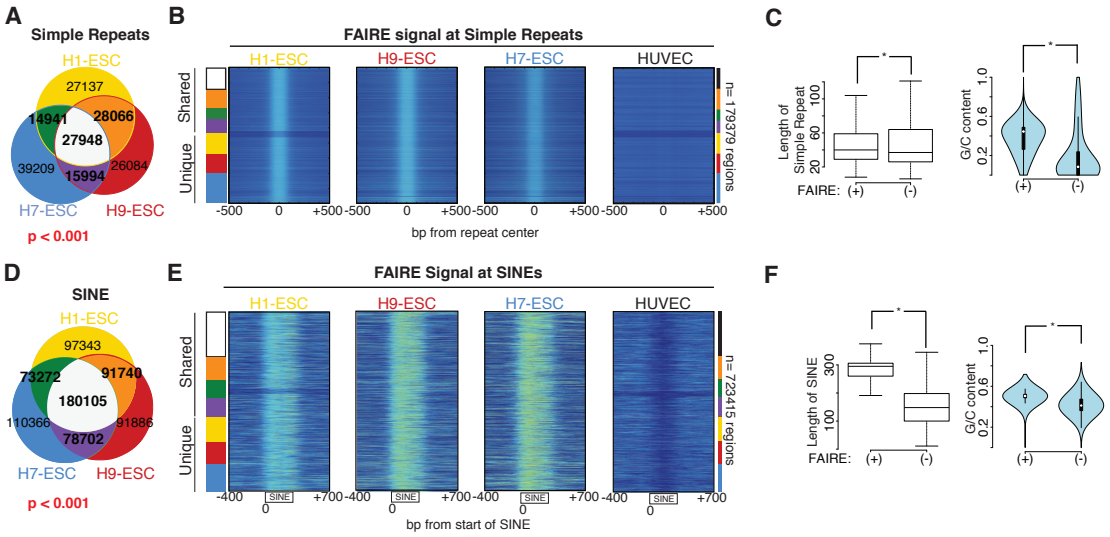
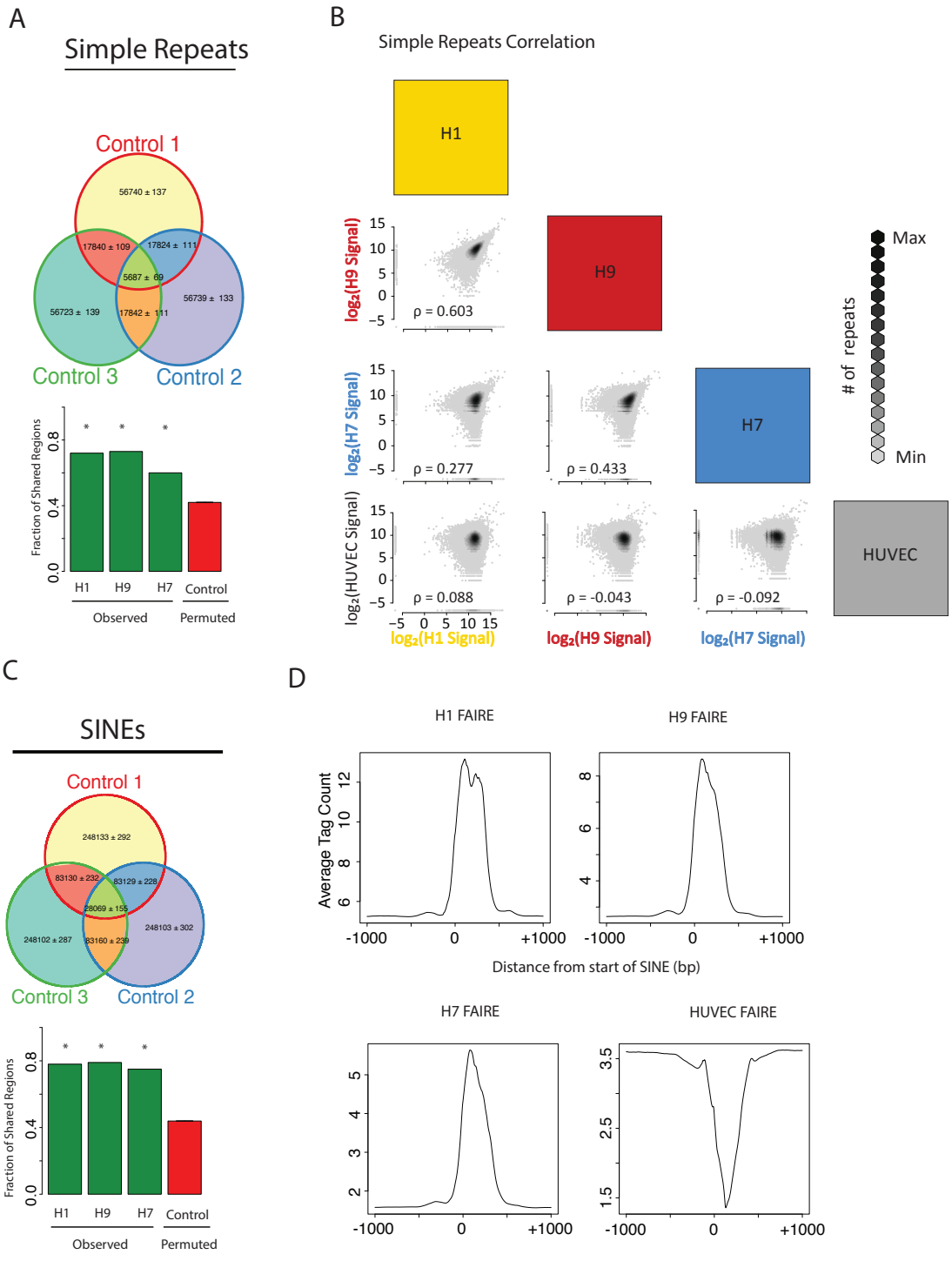


Figure 2.3: hESC share FAIRE signal enrichment at Simple Repeats and SINEs. A) The union set of simple repeats with FAIRE signal in the top quartile (Q4) for hESC are shown. ($p < 0.001$, permutation based on all simple repeats). (B) Heatmap demonstrating FAIRE signal at simple repeats grouped by categories (colors defined in A) with Distance represents bp from the center of the simple repeat. (C) Lengths (left) and G/C content (right) of simple repeats that are FAIRE-enriched ($+$) in all hESC (see panel A, $n = 27{,}948$) and an equal number of randomly selected simple repeats from quartile 1 (FAIRE $-$). (D) The union set of SINEs with FAIRE signal in the top quartile (Q4) for hESC are shown. ($p < 0.001$, permutation based on all SINEs). (E) Heatmap demonstrating FAIRE signal at SINEs grouped by categories (colors defined in D). Distance represents bp from the start of the SINEs. (F) Lengths (left) and G/C content (right) of SINEs that are FAIRE-enriched ($+$) in all hESC (see panel D, $n = 180{,}105$) and an equal number of randomly selected simple repeats from quartile 1 (FAIRE $-$).

mentary approach, we analyzed DNase hypersensitivity data (DNase). In contrast to FAIRE, DNase depends on enzymatic digestion to interrogate chromatin accessibility [7][96]. Leveraging publicly available data, we compared FAIRE and DNase signal at simple repeats and SINEs at FAIRE peaks (from Figure 2.2 C). Surprisingly, these regions lacked DNase signal (Figure 2.5 A). Conversely, the few repeats that demonstrated DNase signal lacked FAIRE enrichment (1099 and 2033 simple repeats and SINEs, respectively). As a control, we examined FAIRE and DNase at transcription start sites (TSS) and CTCF sites. FAIRE and DNase positively correlated at these regions, consistent with published results and confirming the validity of the assays


A
Simple Repeats
Control 1
56740 ± 137
17840 ± 109
17824 ± 111
5687 ± 69
56723 ± 139
17842 ± 111
56739 ± 133
Control 3
Control 2
Fraction of Shared Regions
0.0
0.4
0.8
H1
H9
H7
Control
Observed
Permuted
B
Simple Repeats Correlation
H1
H9
H7
HUVEC
log2(H9 Signal)
log2(H7 Signal)
log2(HUVEC Signal)
log2(H1 Signal)
ρ = 0.603
ρ = 0.277
ρ = 0.433
ρ = 0.088
ρ = -0.043
ρ = -0.092
# of repeats
Max
Min
C
SINEs
Control 1
248133 ± 292
83130 ± 232
83129 ± 228
28069 ± 155
248102 ± 287
83160 ± 239
248103 ± 302
Control 3
Control 2
Fraction of Shared Regions
H1
H9
H7
Control
Observed
Permuted
D
H1 FAIRE
H9 FAIRE
Average Tag Count
-1000
0
+1000
Distance from start of SINE (bp)
H7 FAIRE
HUVEC FAIRE

(Figures 2.5 B and Figure 2.6 A and B).

Because of the discrepancy between FAIRE and DNase at repetitive regions, we then explored nucleosome positioning using published MNase-seq data [169]. By cleaving DNA in the linker region between two nucleosomes, MNase-seq offers insight into the location of nucleosomes. DNase-positive regions, including those in repeats, TSS, and CTCF binding sites, demonstrated decreased MNase signal, consistent with nucleosome depletion (Figure 2.5 A). However, FAIRE-enriched SINEs and shorter simple repeats exhibited the presence of one to two well-positioned nucleosomes. Of note, phased nucleosomes flanked both classes of repeats, similar to patterns observed at other regulatory elements (Figure 2.5 B) [170][171].

To further characterize the relationship between FAIRE and nucleosome positioning at repetitive regions, we examined MNase signal at all simple repeats grouped by the magnitude of FAIRE enrichment. Mnase signal was greatest at regions with highest FAIRE enrichment. Further, regions with the greatest FAIRE signal demonstrated the presence of a single centered nucleosome (Figure 2.5 C). For all regions, we observed symmetrical nucleosome phasing extending beyond the repetitive region. Overall, these data indicate that, in contrast to the recognized association of

---

Figure 2.4: Permutation Results, signal correlation, and signal characteristics. (A) To establish a baseline for comparison, permuted sets of simple repeats were intersected for 1000 iterations and overall relationships are depicted in the Venn diagram (top). Mean numbers of sites for each region $+/-$ standard deviation is shown. Fraction of regions that were in common for at least two ESC lines (H1, H9, H7), or a permuted control set (bottom). (B) hESC FAIRE signal at simple repeats demonstrate a positive correlation. Signal was plotted pair-wise for each cell line and a Spearman correlation measured. (C) To establish a baseline for comparison, permuted sets of SINE were intersected for 1000 iterations and overall relationships are depicted in the Venn diagram (top). Mean numbers of sites for each region $+/-$ standard deviation is shown. Fraction of regions that were in common for at least two ESC lines (H1, H9, H7), or a permuted control set (bottom). (D) Average FAIRE signal at H1, H9, H7, and HUVEC at union set of top quartile FAIRE regions. H1 demonstrate bimodal signal at SINE in top quartile FAIRE regions.

FAIRE with nucleosome depletion, in the context of these regions, FAIRE identifies a chromatin organizational feature characterized by the presence of nucleosomes.

2.2.5 Distinct histone post-translational modifications demarcate repetitive elements

We then asked whether specific histone modifications distinguish the nucleosomes at accessible repetitive elements. We compared H1-ESC ChIP-seq data for a range of histone modifications at FAIRE-enriched and FAIRE-negative sites [172]. We found that FAIRE-enriched simple repeats were marked by specific acetylated histones (Figure 2.7 A). Associated modifications differed from those at FAIRE-enriched SINEs as

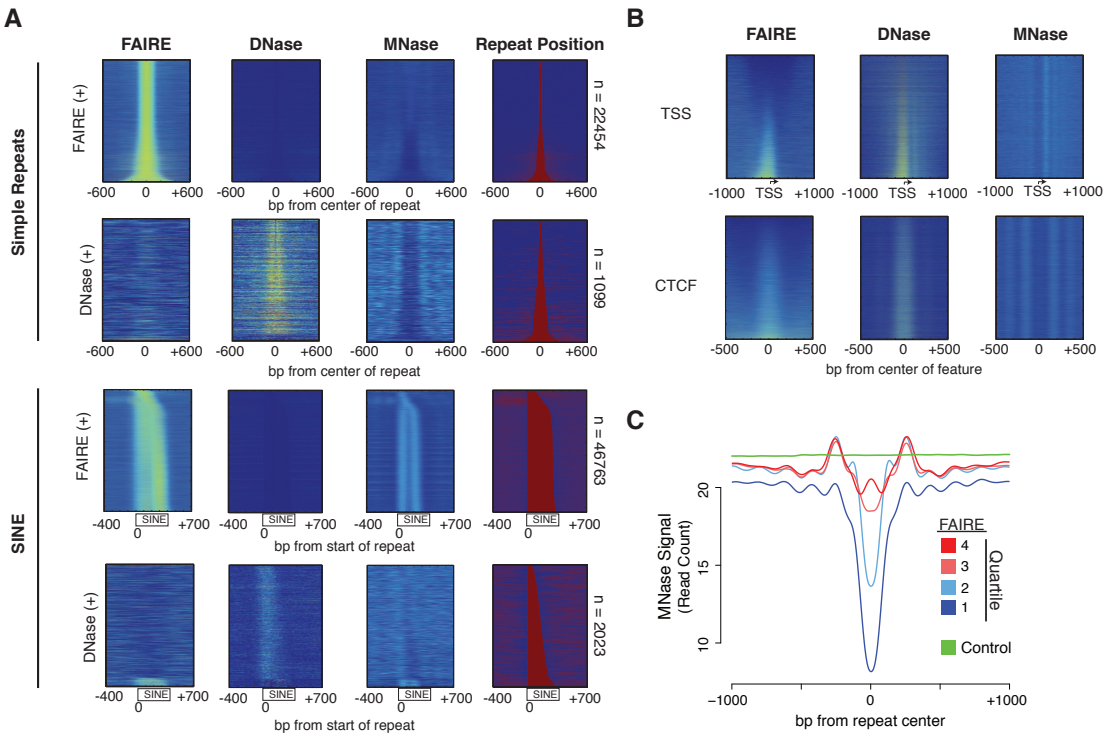


Figure 2.5: Nucleosome-bound repetitive regions are identified by FAIRE. (A) Heatmap representations of H1-ESC FAIRE-seq, DNase, and MNase-seq signal [169] at FAIRE-enriched (FAIRE +) or DNase-enriched (DNase +) simple repeats and SINEs rank ordered by length. For reference, repeat positions (defined by RepeatMasker) are also plotted. (B) Heatmaps of H1-ESC FAIRE-seq, DNase, and MNase-seq signal at transcription start sites (TSS) and CTCF sites. (C) H1-ESC MNase-seq signal at simple repeats grouped by quartiles of FAIRE signal. An equal number of random genomic windows are plotted for comparison (control, green).

well as TSS and CTCF sites (Figure 2.7 A). H3K56ac and H2AK5ac were most associated with simple repeats. Signals for these modifications were centered over the repeat and demonstrated a magnitude similar or greater than that found at TSS and CTCF sites. (Figures 2.7 B and 2.8). H4K8ac and H2A.Z were most associated with SINEs and show subtle but center-weighted enrichment (Figures 2.7 C and 2.9). Overall, these data indicate that FAIRE-enriched simple repeats and SINEs are characterized by distinctly marked nucleosomes.

As an alternative approach to explore chromatin accessibility, we performed salt fractionation of MNase treated nuclei. Salt fractionation separates chromatin based on physical properties [173]. Low salt-soluble regions are enriched for active and highly accessible chromatin, whereas high salt solubilizes the bulk chromatin fraction [174]. Salt fractionation both allows for direct comparisons of nucleosome composition in active chromatin as well as the positioning of individual nucleosomes by high-throughput sequencing. Nucleosomes were extracted from nuclei of H1-ESC and a differentiated control (Human Kidney Cells, HKC) using increasing concentrations

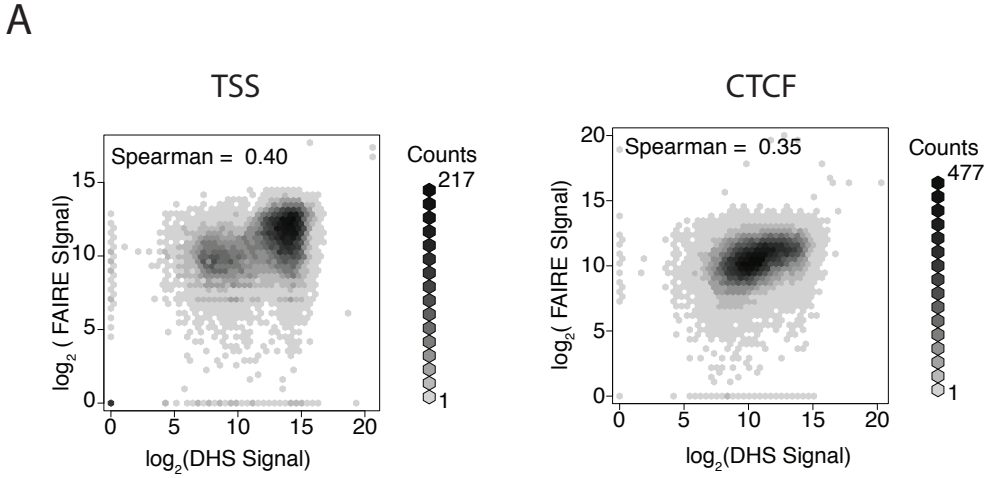


Figure 2.6: FAIRE and DHS correlate at TSS and CTCF sites. (A) Log2 transformed FAIRE and DHS signal at TSS (+/- 300 bp) or CTCF sites (+/- 200 bp) were plotted. Spearman correlation is shown.

of salt. The low salt fraction from both cell types consisted predominantly of mono-nucleosomes whereas the high salt fraction consisted of mostly di-nucleosomes (Figure 2.10), consistent with published results [175]. Histone post-translational modifications associated with each fraction were assayed by immunoblot. As predicted by our informatic analyses, H2AK5ac was significantly enriched in low salt fractions of nucleosomes from stem cells when compared to the differentiated control cells (p-value < 0.05, Figure 2.7 D). This enrichment did not extend to the high salt or insoluble chromatin.

To identify nucleosome positioning at repetitive elements in highly accessible chromatin, we sequenced the DNA in both low and high salt soluble fractions and plotted the signal at simple repeats (Figure 2.7 E). We again identified nucleosome phasing flanking the repeats in both H1-ESC and the differentiated control cells. However, compared with the differentiated cell control, H1-ESC demonstrated an increase in MNase signal at the center of the repeat exclusively in low salt extracted chromatin, indicative of a highly extractable nucleosome. Taken together with the immunoblot and ChIP-seq analysis, these data indicate that specific acetylation is associated with nucleosomal destabilization but not displacement at repetitive elements.

2.2.6 Repetitive regions undergo chromatin remodeling during differentiation

The difference in FAIRE enrichment at repetitive elements in stem and differentiated cells led us to test whether these elements undergo remodeling during differentiation. H1-ESC embryonic stem cells were differentiated in culture towards a mesenchymal lineage (H1-MSC). The mesenchymal lineage consists of mesenchymal stem cells (MSC) that can terminally differentiate into diverse range of cells [62].

Differentiation of hESC to MSC was validated using several approaches. Morphologically, H1-MSCs acquire a fibroblastic appearance in contrast to the spherical colonies of H1-ESC (Figure 2.11 A). The multipotency of H1-MSC was demonstrated

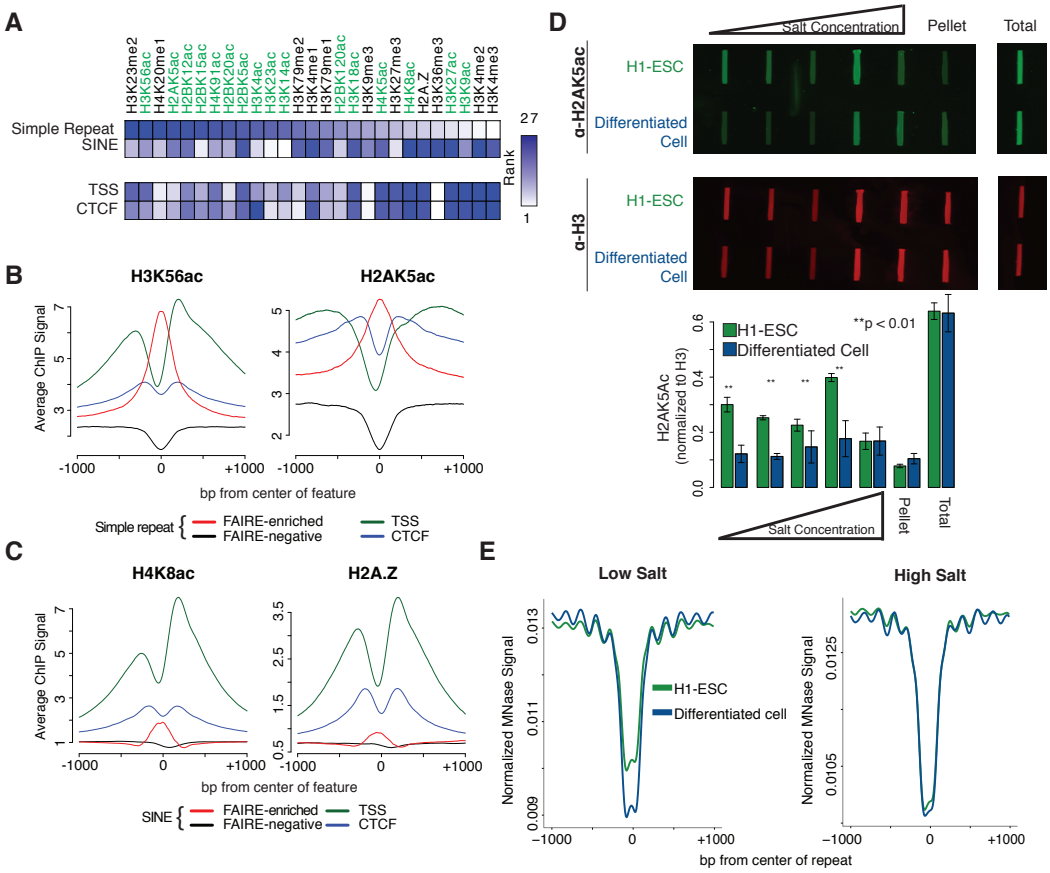


Figure 2.7: Distinct histone modifications characterize FAIRE-enriched repeats. (A) Heatmap of ranked histone posttranslational modifications. Differential ChIP signal comparing FAIRE-enriched (+) with FAIRE-negative repeats (simple repeat +/− 250 bp from center or SINEs start to +300 bp) was rank ordered. For comparison, signal at TSS or CTCF (+/−500 bp) was rank ordered. Histone modification by acetylation is highlighted (green). (B) Mean H1-ESC ChIP signal of selected histone posttranslational modifications at H1-ESC FAIRE-enriched (red line) and FAIRE-negative (black line) simple repeats, and control regions (TSS green line, CTCF blue line). (C) Mean H1-ESC ChIP signal of H4K8ac and H2A.Z at H1-ESC FAIRE-enriched (red line) and FAIRE-negative (black line) SINEs, and control regions (TSS green line, CTCF blue line). (D) Salt fractionated nuclear extracts from H1-ESC and HKC (differentiated cell) were immunoblotted with anti-H2AK5ac (green) and anti-pan-H3 (red). Fluorescence intensity was quantified and normalized to H3. (E) Mean H1-ESC and HKC (differentiated cell) MNase-seq signal from low (left) and high (right) salt fractions at simple repeats. Signal was normalized to reads per million mapped.

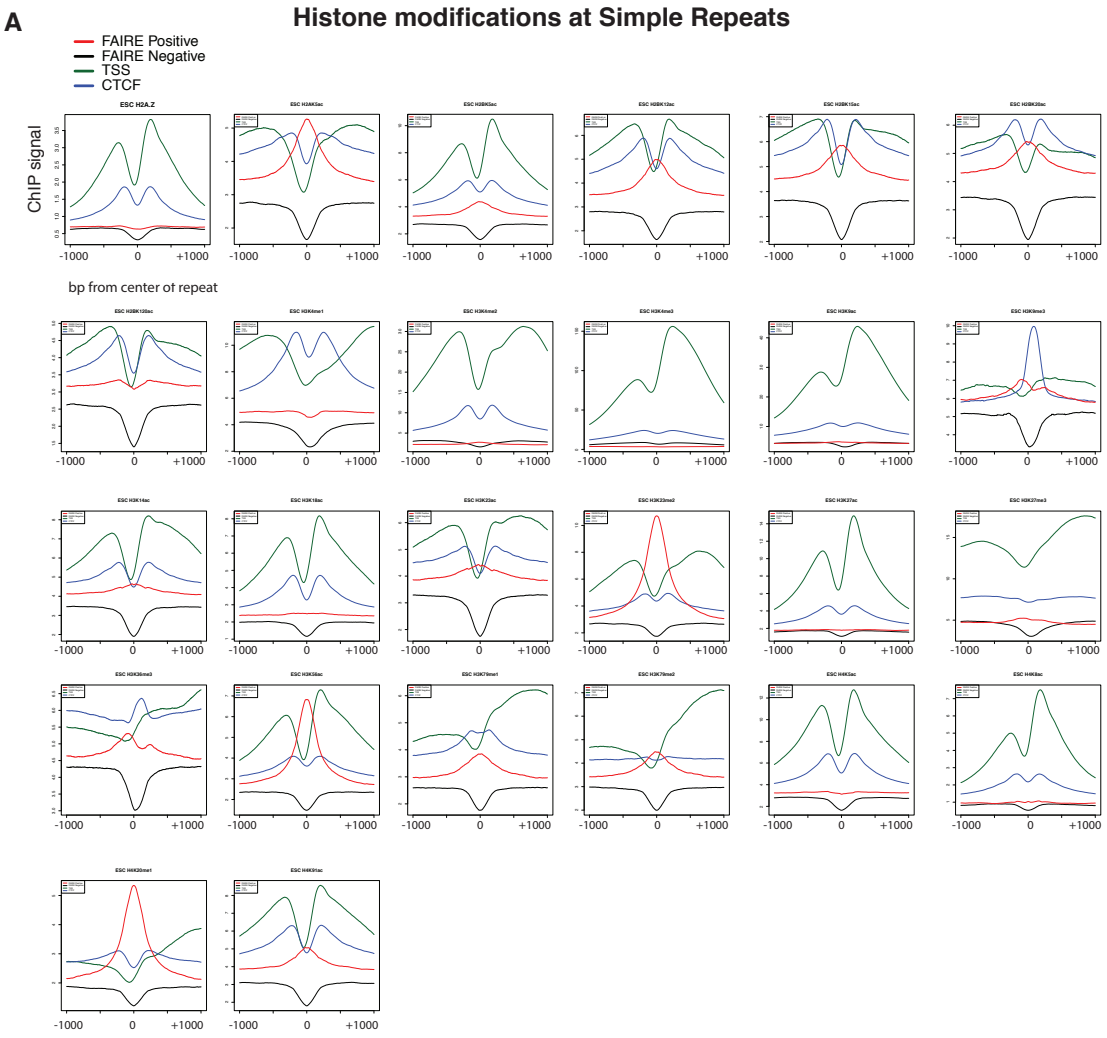


Figure 2.8: Histone modifications at simple repeat. (A) ChIP signal for available histone modifications [172] at FAIRE-positive (red), FAIRE-negative (black) simple repeats, TSS (green), and CTCF binding sites (blue).

by further differentiation into osteoblast and adipocyte lineages, as demonstrated by Alizarin red and Oil Red O staining (Figure 2.11 A). Finally, flow cytometry of H1-MSC identified a robust increase in CD90, CD73, CD105, and CD44, cell surface markers also detected on primary bone marrow-derived MSC (BM-MSC) (Figure 2.11 B). H1-ESC were negative for CD73 and CD105.

FAIRE-seq was then performed on primary BM-MSC and H1-MSC. We identified 25,000 SINEs and 4,700 simple repeats with significantly different FAIRE signal be-

A

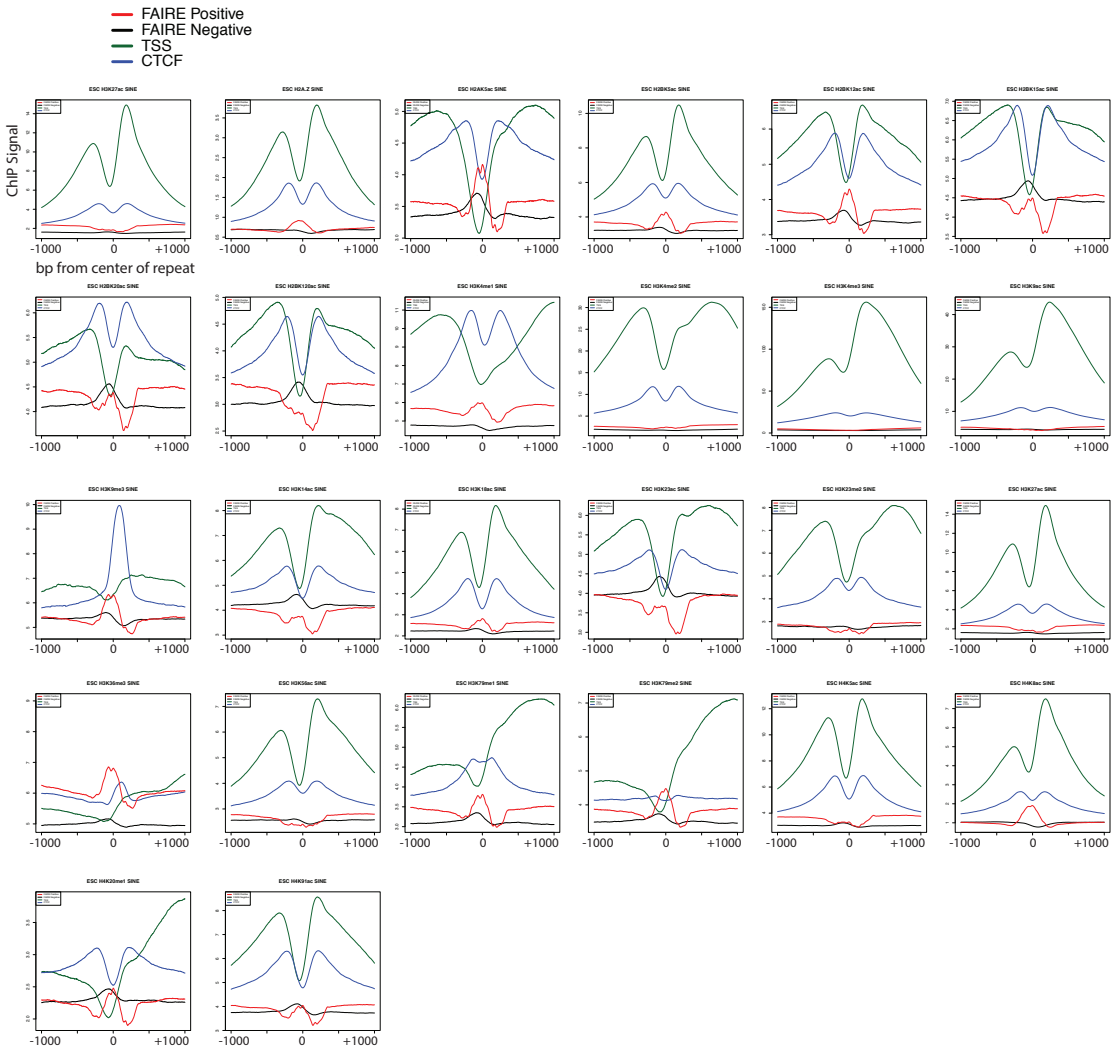


Figure 2.9: Histone modifications at SINE. (A) ChIP signal for available histone modifications [172] at FAIRE-positive (red), FAIRE-negative (black) simple repeats, TSS (green), and CTCF binding sites (blue).

tween hESC and BM-MSC. Unsupervised hierarchical clustering of these regions, together with signal from H1-MSC as well as HUVEC, NHEK, and K562, revealed two main clusters (Figure 2.12 A and E). For SINEs, stem cells clustered together closely whereas differentiated cells segregated into a separate cluster. Notably, the differentiated H1-MSC exhibited greater similarity to primary BM-MSC than to undifferentiated H1-ESC. Overall, virtually all sites that exhibit signal variation demonstrated a progressive decrease in FAIRE enrichment accompanying differentiation. SINEs with differential FAIRE signal were then associated with genes [176]. Of those genes that were also differentially expressed, 95% demonstrate greater expression ($>$ 2-fold) in hESC relative to BM-MSC, compared with 89% for genes associated with invariable SINEs (Figure 2.12 B) The overall skew to greater message abundance in hESC is consistent with higher global transcription levels in these cells [54]. Genes with elevated expression in hESC (Category 1) were linked to terms such as cell cycle whereas those with elevated expression in MSC (Category 2) were linked with

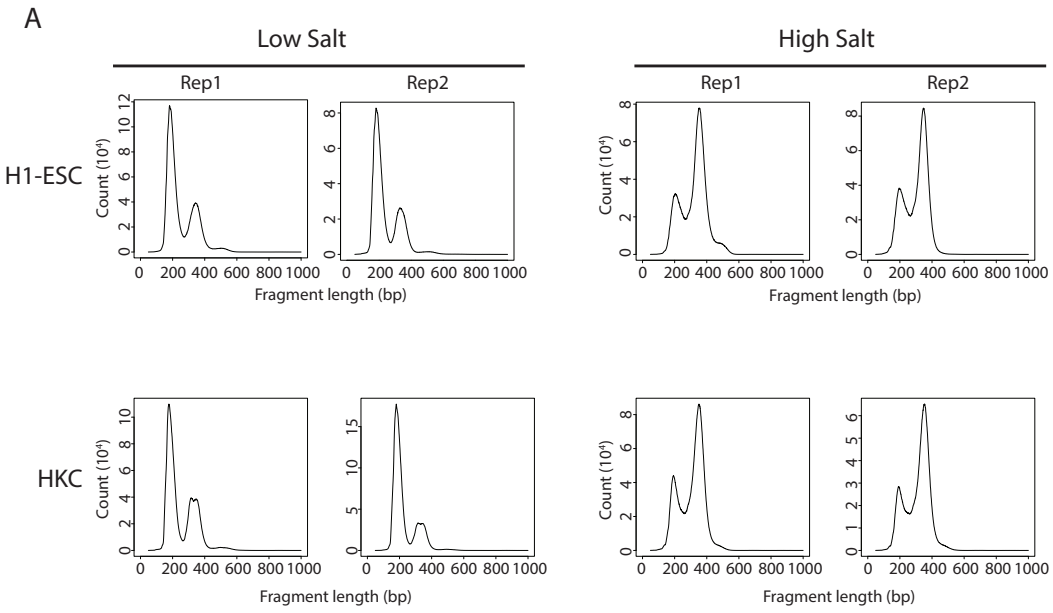


Figure 2.10: Low and High salt fractions enrich for different fragment lengths. (A) Distribution of fragment lengths resulting from paired-end sequencing of MNase of chromatin extracted at low and high salt conditions in H1-ESC and HKC.

terms implicating mesenchymal development such as blood vessel development and response to wounding (Figures 2.12 C and D).

Clustering cell lineages based on FAIRE signal at simple repeats demonstrated a distinct pattern from that observed based on signal at SINEs. hESC clustered together, clearly separated from the differentiated cells. MSC, including bone marrow and H1-derived, clustered closely together but grouped with the differentiated cells. FAIRE signal at simple repeats revealed two patterns. One pattern was similar to that seen for the SINEs with progressively decreasing signal associated with differentiation. The other pattern revealed FAIRE signal greater in MSC lineages compared with either hESC or the differentiated cells (Figure 2.12 E). We again associated these regions with differentially regulated genes. Regions with higher FAIRE signal in hESC or MSC were associated with genes more highly expressed in hESC or MSC, respectively (compared with invariable region-associated genes) (Figure 2.12 F). hESC-associated genes were enriched for gene ontologies related to development, such as tube development and stem cell development, whereas MSC-associated genes were enriched for pathways linked to mesenchymal development, such as blood vessel development, skeletal system development, and response to wounding (Figures 2.12 G, H and Figure 2.9 C).

Taken together, these results suggest that repetitive elements undergo chromatin remodeling during differentiation. Repetitive regions with variable accessibility are associated with changes in lineage-specific gene expression and developmental pathways. However, the pattern of remodeling differs between the two classes of repetitive elements. SINEs primarily become inaccessible during differentiation, whereas a subset of simple repeats become accessible during lineage specification.

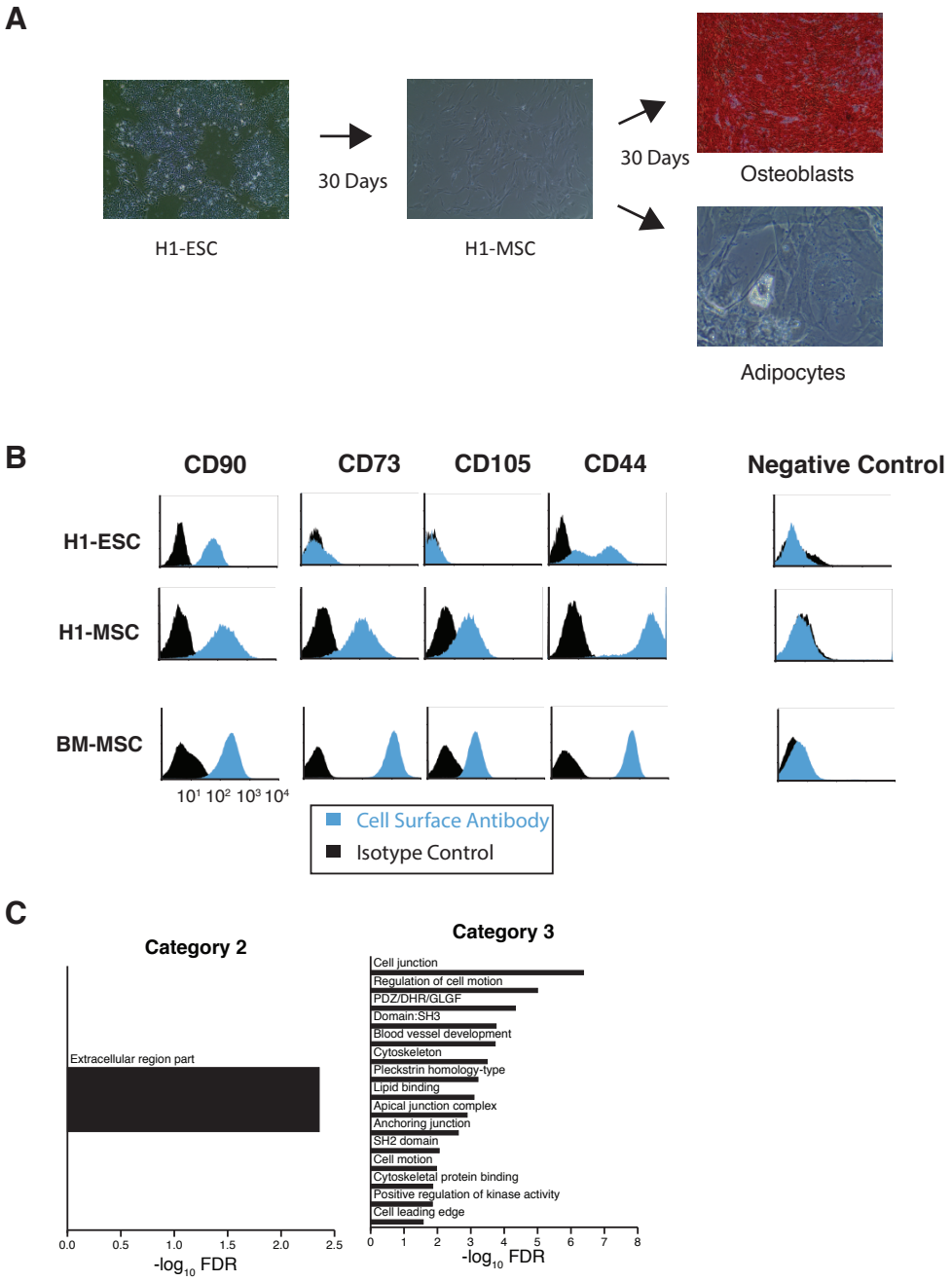
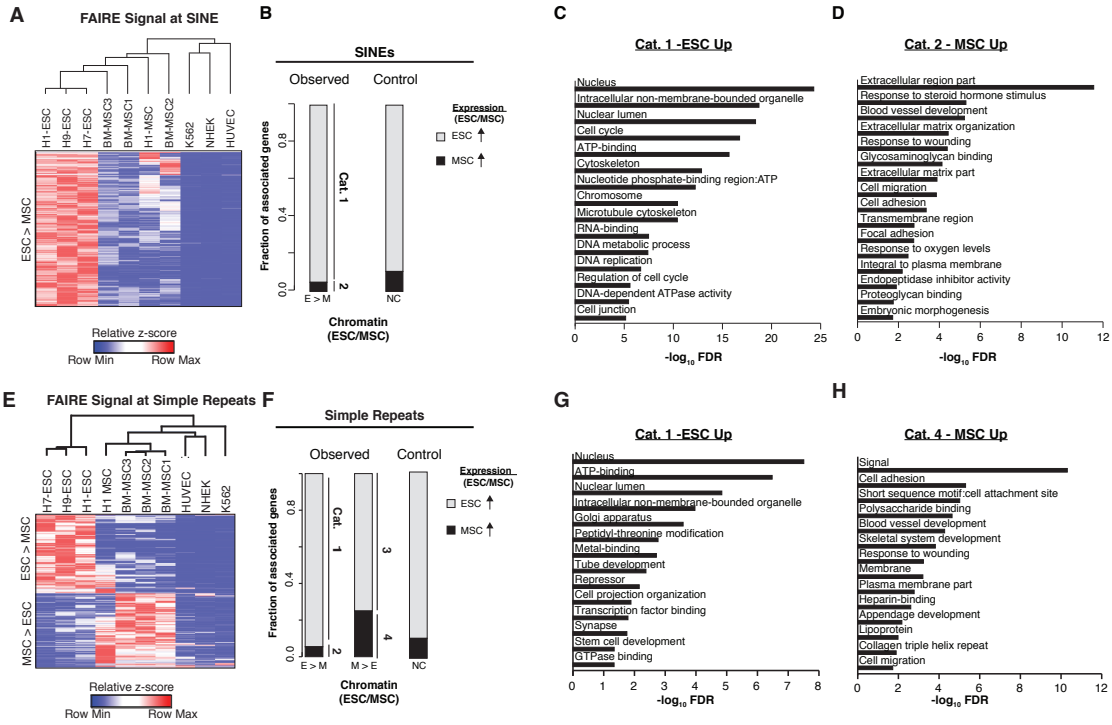


Figure 2.11: Differentiation of H1-ESC to H1-MSC. (A) H1-MSC can differentiate into osteoblasts and adipocytes. Photomicrographs of H1 ESC prior to differentiation (phase contrast), intermediately differentiated H1-MSC (phase contrast) and lineage-specifically differentiated osteoblasts (Alizarin Red) and adipocytes (Oil-Red O). (B) H1-MSC acquire MSC-associated cell surface markers during in vitro differentiation. H1 ESC, H1-MSC and BM-MSC were stained for CD90, CD73, CD105, CD44, and a negative control cocktail and analyzed by flow cytometry. (C) Remaining categories of enriched gene ontologies from Figure 2.12. Category 2 and 3 enriched terms are shown.

2.2.7 Oncogenic transcription coopts stem cell chromatin

Many sarcomas are thought to originate from stem cells of mesenchymal origin [67]. To explore this link, we compared the chromatin environment in stem cells with that in Ewing Sarcoma, the second most common bone malignancy in children and young adults. Ewing sarcoma is characterized by a chromosomal rearrangement that creates a novel transcription factor. We and others have previously shown that the resulting chimeric oncoprotein, EWSR1-FLI1, targets a subset of simple repeats distinct from the parental protein FLI1 [89]. Further, this targeting is cell-type specific [89][90]. This observation led us to hypothesize that a permissive chromatin environment may facilitate this retargeting. We tested for the enrichment of repeat classes in accessible chromatin in tumor cells and primary BM-MSC. Both tumor cells and MSC shared high degree of enrichment at simple repeats, relative to other repetitive element classes (Figure 2.13 A).

Since EWSR1-FLI1 selectively retargets GGAA-containing simple repeats, we ex-



amined FAIRE signal in BM-MSC and Ewing sarcoma cells at all simple repeats containing this motif, clustering these regions based on their signal in the cancer cells (Figure 2.13 B). We found a striking similarity in the pattern of chromatin accessibility between the stem and cancer cells. In BM-MSC, the signal was center-weighted at about half of the regions (Figures 2.13 B and 2.14 A). For others, regions flanking the repeat demonstrated the greatest signal.

To explore the connection between chromatin accessibility and EWSR1-FLI1 targeting, we compared FAIRE signal in BM-MSC with EWSR1-FLI1 ChIP signal from Ewing sarcoma cells. Repeats with the greatest FAIRE signal in BM-MSC demon-

---

Figure 2.12: Repetitive elements undergo extensive chromatin remodeling during differentiation. (A) FAIRE signal from hESC, BM-MSC, H1-MSC, K562, NHEK and HUVEC at SINEs characterized by significantly different FAIRE signal between hESCs and BM-MSCs (t-test $p<0.01$ and $row_{max} - row_{min} > 3$) were z-score transformed and biclustered. Heatmap scale represents relative z-scores. (B) Fraction of genes linked to variable FAIRE at repeats that demonstrate differential gene expression (increased in hESC - gray, category 1; increased in MSC - black, category 2). Differential expression was defined as RPKM $>$ 4 fold change. Control represents an equal number of randomly selected repeats that did not change significantly during differentiation. (C) Gene ontologies enriched for genes in category 1 from (C). Bars indicate $-\log_{10}$FDR. (D) Gene ontologies enriched for genes in category 2 from (C). Bars indicate $-\log_{10}$FDR. (E) FAIRE signal from hESC, BM-MSC, H1-MSC, K562, NHEK and HUVEC at Simple Repeats characterized by significantly different FAIRE signal between hESCs and BM-MSCs (t-test $p<0.01$ and $row_{max} - row_{min} > 3$) were z-score transformed and biclustered. Heatmap scale represents relative z-scores. (F) Fraction of genes linked to variable FAIRE at repeats that demonstrate differential gene expression. Genes associated with increased FAIRE in hESC are further divided, those associated with genes with increased expression in hESC (gray, category 1) and those associated with genes with increased expression in MSC (black, category 2). Genes associated with increased FAIRE in MSC are also further divided, those associated with genes with increased expression in hESC (gray, category 3) and those associated with genes with increased expression in MSC (black, category 4). Differential expression was defined as RPKM $>$ 2 fold change. Control represents an equal number of randomly selected repeats that did not change significantly during differentiation. (G) Gene ontologies enriched for genes identified in category 1 from (G). Bars represent $-\log_{10}$FDR. (H) Gene ontologies enriched for genes identified in category 4 from (G). Bars represent $-\log_{10}$FDR.

strated the greatest ChIP signal in the tumor cells (Figure 2.14 B). Similarly, EWSR1-FLI1 targeted those regions for which the maximal FAIRE signal was over the repeat in BM-MSC ($p < 0.001$, permutation). We then explored the activity of EWSR1-FLI1 on chromatin. We compared the difference in FAIRE signal between BM-MSC and the tumor cells with EWSR1-FLI1 ChIP signal. We found a significant correlation between oncoprotein binding and changes in FAIRE signal (r = 0.74) (Figure 2.13 C). Taken together, these data lend chromatin-based evidence of an MSC origin for these tumors and, further, demonstrate that an MSC chromatin pattern predicts EWSR1-FLI1 oncoprotein targeting.

We then explored chromatin accessibility using enzymatic approaches. Neither DNase-seq data that we generated from BM-MSC nor published DNase and ATAC data from these cells identified signal enrichment at regions ultimately targeted by EWSR1-FLI1 (Figure 2.13 D) [90][95]. The absence of signal is consistent with our previous observation regarding DNase in hESC (Figure 2.5 A). In contrast, in Ewing Sarcoma cells these regions were detected by DNase and ATAC. Moreover, in BM-MSC ATAC enrichment was noted at these sites only after EWSR1-FLI1 was transduced (Figure 2.13 D and [95]). DNase and ATAC signal enrichment was not observed at similar repeats that did not bind EWSR1-FLI1. These data suggest that EWSR1-FLI1 targets nucleosome-destabilized regions ultimately evicting nucleosomes, a feature which can then be detected using the enzymatic approaches of DNase and ATAC.

Since EWSR1-FLI1 did not target all GGAA-containing simple repeats, we asked whether there were other chromatin features that correlated with increased FAIRE signal in BM-MSC and the ability to bind EWSR1-FLI1. Using ChIP from H1-MSC [172] we examined histone modifications at those sites that are targeted by EWSR1-FLI1 in cancer cells. Of histone modifications available for analysis, we noted a subtle increase in enrichment for H3K14ac, H4K91ac, H2BK12ac, all marks enriched in sim-

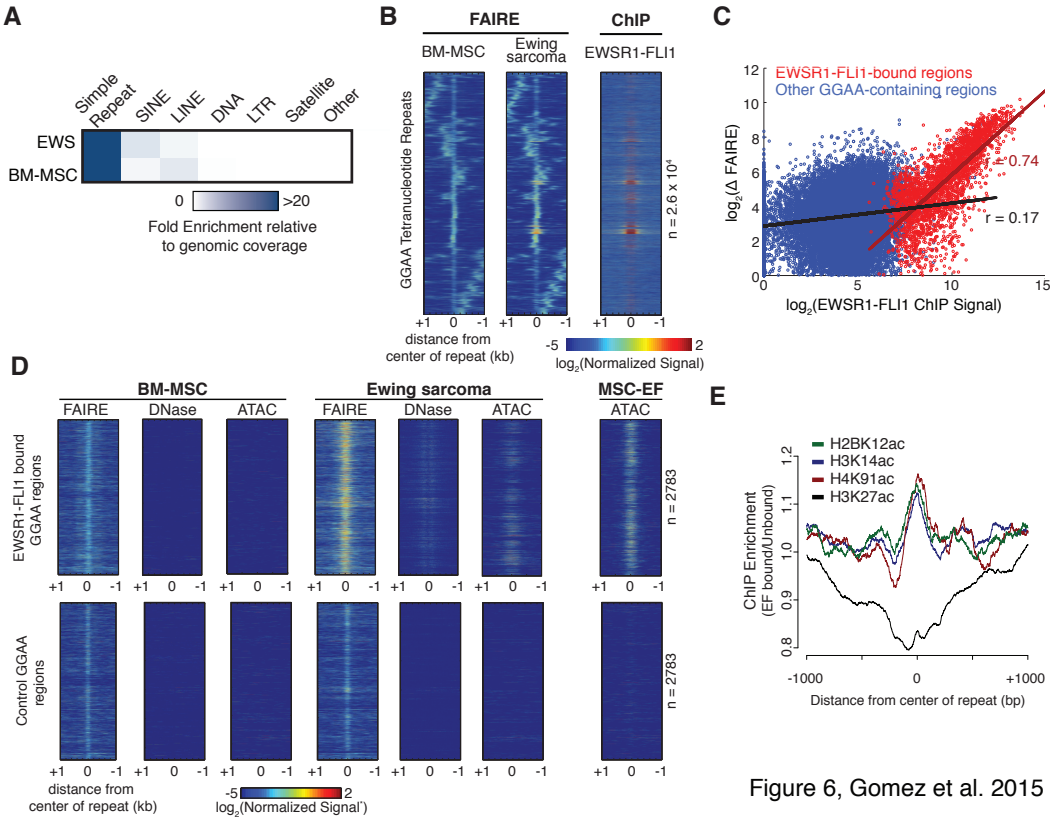


Figure 6, Gomez et al. 2015

Figure 2.13: Cancer exploits unique chromatin environment of stem cells. (A) Heatmap depicting the enrichment of specific classes of repetitive elements in MACS2-identified FAIRE-enriched regions in Ewing Sarcoma (EWS) and BM-MSC chromatin, relative to genomic coverage. (B) Clustered BM-MSC or EWS FAIRE signal at all $(GGAA)_n$-containing simple repeats (left). EWSR1-FLI1 ChIP signal in EWS at $(GGAA)_n$-containing simple repeats (right). (C) Scatterplot of log_2 transformed FAIRE change between BM-MSC and EWS and EWSR1-FLI1 ChIP signal at EWSR1-FLI1 bound (red) or unbound (blue) repeats. Pearson correlation shown. (D) FAIRE, DNase and ATAC signal at EWSR1-FLI1 binding sites in BM-MSC, Ewing Sarcoma (EWS), and MSCs exogenously expressing EWSR1-FLI1 [95]. Distance represents Kb from the center of the repeat. FAIRE and DNase data were normalized to overall read count. ∗ATAC read count was unavailable and consequently not normalized. (E) Fold change of H1-MSC ChIP signal for H3K14ac, H4K91ac, H2BK12ac, and H3K27ac at repeats bound by EWSR1-FLI1 in Ewing sarcoma cells relative to an equal number of repeats that were not bound. Distance represents from center of repeat.

ple repeats in hESC (Figure 2.13 E). These data suggest that chromatin modifications specific to stem cells facilitate EWSR1-FLI1 targeting.

2.3 Discussion

By integrating multiple complementary genome-wide approaches we identified a unique chromatin environment in stem cells marked by accessible chromatin at repetitive DNA sequences. Further, we demonstrate that these features offer a permissive environment for the central oncogenic pathway in Ewing sarcoma.

Though FAIRE-seq revealed the magnitude of this unexpected chromatin signature in hESC, complementary experimental approaches have yielded results consistent with this observation. An examination of chromatin in murine ESC using FAIRE similarly demonstrated variation in regions associated with developmental pathways [177]. Although not addressed in this study, our analysis of these data also demonstrated enrichment of repetitive elements in mESC compared to MEF (data not shown). That hESCs had significantly more FAIRE peaks and a generally lower signal-to-noise ratio than differentiated cells suggests decreased chromatin condensation, consistent with biochemical and microscopic approaches [178][52].

The most characteristic feature associated with the repetitive elements in accessible chromatin was histone acetylation. Variations in histone acetylation have been linked to stem cell differentiation, and nucleosome acetylation can destabilize DNA-nucleosome interactions [179][48]. Interestingly, the sites of acetylation enriched at simple repeats differ from the well-studied H3K27ac and H3K9ac. Segmentation analysis of stem cells has generally categorized repeats as heterochromatic, however these modeling approaches have not included atypical marks, such as H2AK5ac. Indeed, evidence suggests that H2AK5ac enrichment is associated with active regions of chromatin [180]. Given the paucity of available datasets, features other than histone acetylation may also influence chromatin accessibility. As a functional readout

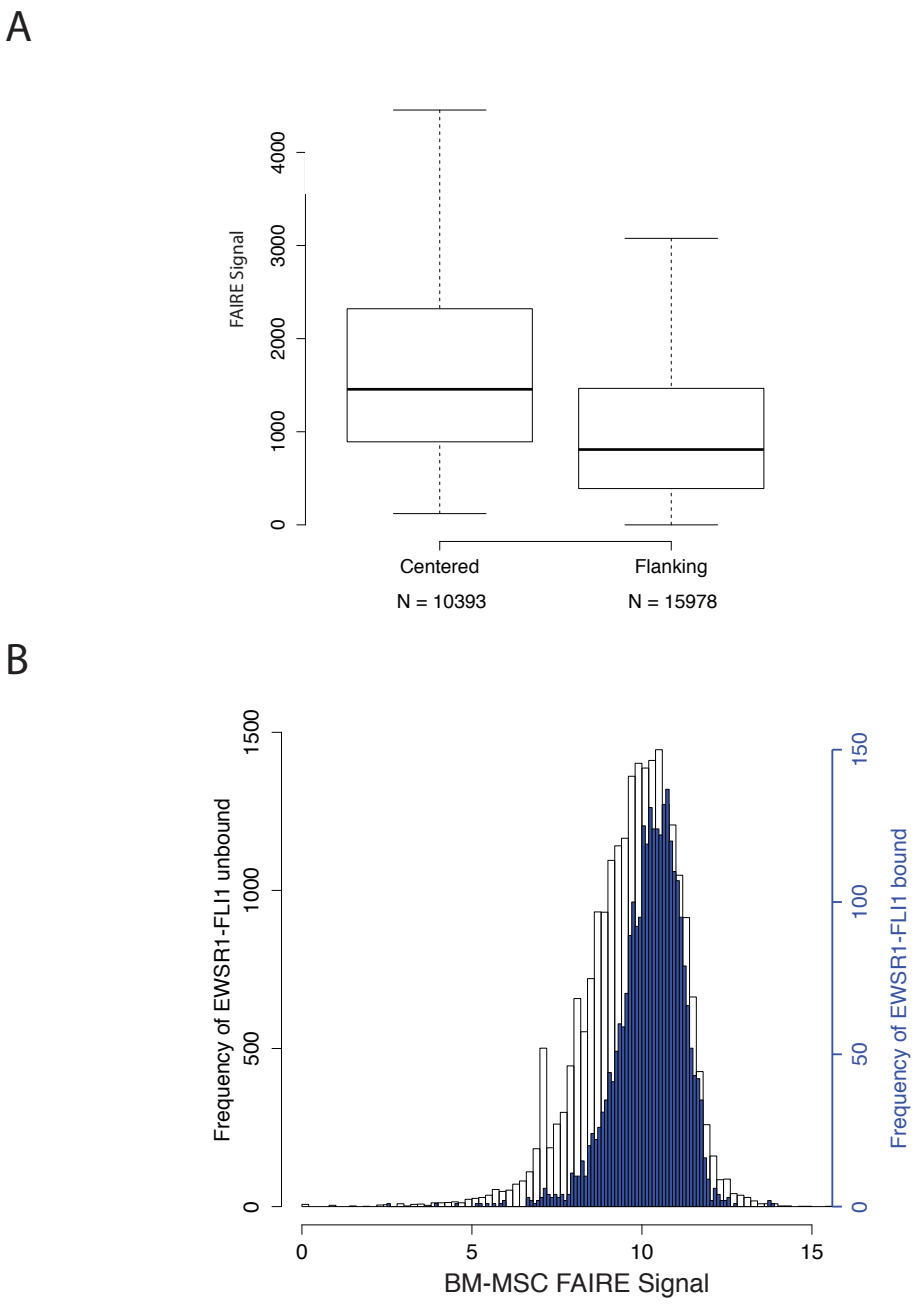


Figure 2.14: Regions with increased FAIRE signal in BM-MSC tend to be center-weighted and associated with EWSR1-FLI1 binding in tumor cells. (A) Repeats were classified as center-weighted signal or flanking signal using a 200 bp sliding window with a 100 bp overlap. Distribution of FAIRE signal at the center of the repeat ($+/-$ 200 bp) was plotted. (B) Histogram of BM-MSC FAIRE summed $+/-$ 100bp from the center of EWSR1-FLI1 at bound (blue) or unbound (white) GGAA-containing simple repeats and log2 transformed.

of chromatin states, the inclusion of FAIRE may increase the power of predictive genomic segmentation.

In support of a functional role for repetitive elements, we observed alterations in chromatin organization that accompany differentiation. Genes associated with SINEs that demonstrated differential accessibility exhibited cell cycle pathway enrichment, consistent with the importance of this pathway in maintaining pluripotency [181]. Similarly, pathways associated with mesenchymal differentiation and function were enriched among regions with gains in accessibility during differentiation. Interestingly, variation in histone posttranslational modifications between induced pluripotent stem cells (iPSC) and ESCs has been inconsistently identified [182][60]. Analysis of iPSC by FAIRE would identify whether chromatin accessibility at repetitive elements is restored during reprogramming and could contribute to chromatin-based insights regarding the reprogramming process.

The observation of phased nucleosomes flanking all repetitive elements was also unexpected. Ordered nucleosomes were observed even in the absence of FAIRE enrichment and in differentiated cells. Stretches of specific repeated sequences can bend DNA which may attract nucleosomes, and DNA sequence content can influence nucleosome position [183][184][185][186]. Our results greatly extend previous reports suggesting that Alu repeats may serve to pattern nucleosomes [187][188].

A striking result of our study was the extensive similarity between MSC chromatin and that of Ewing sarcoma. The shared chromatin pattern strongly supports tumor development from a stem-like population, an observation consistent with studies describing similarities in gene expression and capacity for in vitro differentiation [85][68]. Further, our results offer a mechanistic explanation for the cell-type specific targeting of EWSR1-FLI1 in tumor cells. The absence of accessible chromatin at repetitive elements in differentiated cell types may restrict EWSR1-FLI1 targeting offering an explanation of why this oncogene fails to broadly transform cells [189]. Simple repeats,

when bound by EWSR1-FLI1, gain enhancer activity to regulate the transcription of multiple genes known to be important for Ewing Sarcoma [89][190][92][93][90][94]. Further, germline variation in repetitive element composition has recently been associated with disease risk [191][190]. The differences in location and composition of these repetitive regions relative to critical genes across species may explain the challenge in generating an animal model that faithfully recapitulates features of Ewing sarcoma [192].

Finally, our study offered unexpected technical insights into FAIRE. Previous studies have noted discrepancies between FAIRE and DNase, particularly at distal regulatory elements [102]. However, the biochemical differences characterizing those regions that are enriched by FAIRE but not detected by DNase have not been identified. Compared with DNase and ATAC, FAIRE-seq seems unique in its ability to identify unstable nucleosome-bound regions. Resulting from chromatin organizational differences or histone acetylation, these destabilized nucleosomes may not survive the biochemical extraction process of FAIRE. In a similar fashion, unstable H2A.Z/H3.3 containing nucleosomes have been found at regions deemed “nucleosome depleted” [193]. In contrast, DNase and ATAC depend on exposed DNA for enzymatic cleavage. Consistent with this difference, DNase and ATAC data from Ewing Sarcoma indicates nucleosome eviction. Nucleosome eviction was also reflected by quantitative gains in FAIRE signal. Strategies that explore chromatin organization yield distinct insights. Apparent differences between these methods may indicate specific states that influence chromatin accessibility.

Overall, we identify a link between stem cell-specific chromatin features at repetitive elements and cancer development. Because of their abundance these elements may broadly influence nucleosome positioning and chromatin remodeling during differentiation. Multiple mechanisms result in variation in repeat element structure and location. How these factors converge to alter chromatin organization will continue to

enhance our understanding of development and disease.

2.4 Materials and Methods

2.4.1 Cell Culture and Isolation of BM-MSC

The human embryonic stem cell line H1-ESC, H7-ESC, H9-ESC was obtained from WiCell Research Institute (Madison, WI). H1 hESC were maintained undifferentiated on 6-well plates coated with growth factor reduced Matrigel (BD Biosciences) in mTeSR1 media (StemCell Technologies) and the media was changed daily. Cells were passaged every three days with 0.5mM EDTA.

BM-MSCs were derived from human primary bone marrow (IRB exemption 09-0127). Media was added to bone marrow and centrifuged at 1500rpm to pellet cells. Pellet was resuspended and separated using a Ficoll gradient by centrifugation for 30 minutes. Mononuclear cells were collected and washed 1X in PBS. Cells were resuspended in growth media (DMEM low glucose supplemented with 10% selected serum) and seeded on plates. After 3-4 days cells were trypsinized, passed through a cell filter, and incubated with CD11B/MAC-1 (BD Pharmingen #555386) and CD45 (BD Pharmingen #555481). Cells were washed 2X and resuspended with streptavidin coated beads (MACS) according to manufacturer's recommendations. Cells were washed 2X and ran through a MS-column (MACS). Negative fraction (Unlabeled cells) consisting of BM-MSCs was expanded in growth media with twice weekly media changes.

HKC were maintained in DMEM high glucose supplemented with 10% serum + 1% nonessential amino acids, 1% L-glutamine, 1% pen/strep. Media was changed every 2 days and cells were passaged 2 times a week.

2.4.2 Differentiation of H1 MSCs to osteoblasts

H1-MSC were seeded at 6 X 10^3 cells/cm^2 in Mesencult Osteogenic Stimulatory Kit (StemCell Technologies) and were maintained according to the manufacturer's instructions. Medium was changed every three days and osteogenic differentiation was visualized using Alizarin Red R staining one month after the beginning of differentiation.

2.4.3 Differentiation of H1 MSCs to adipocytes

H1-MSC were seeded at 5.5 X 10^5 cells per 10 cm dish or 9.5 X 10^4 cells/well of a 6-well plate in StemPro Adipogenesis differentiation kit (Life Technologies) and were maintained according to the manufacturer's instructions. The presence of lipid droplets was confirmed after one month of differentiation using Oil Red O staining.

2.4.4 Flow Cytometry

H1, H1-MSC, and BM-MSCs were stained according to manufacture's recommended conditions using the BD Stem Flow Human MSC Analysis Kit (#562245).

2.4.5 FAIRE-Seq Analysis

Chromatin from hESC, H1-MSC, and BM-MSC was isolated by fixing cells in 1% final concentration of formaldehyde for five minutes and quenched using 125mM glycine. Nuclei were isolated by dounce and sonicated to average fragment size of approximately 400bp. FAIRE was preformed and sequencing libraries were generated from DNA enriched by FAIRE and sequenced (Illumina). 50-bp single end reads were filtered using TagDust (Lassmann et al., 2009) and aligned to hg19 using Bowtie [160]. FAIRE peaks were called using MACS2 [167]. Default parameters were used except –shiftsize=67 and –nomodel to better account for the FAIRE fragment length. To narrow our peak calls to the most confident set, we selected the $\log_{10}$(q-value)

threshold for the 150,000th peak and selected the regions that were $\geq$ to that value. Varying the q-value threshold had no effect on the fraction of peak summits $+/-$ 1bp that intersected a repetitive element. Data are available from GEO (Accession number GSE75172). FASTQ files from HUVEC, NHEK, and K562 FAIRE were obtained from ENCODE and processed as above.

Z-score transformed wiggle files were created using the average FAIRE signal per base-pair over the genome, excluding chromM and chromY for each sample. Genome browser shots of representative clusters were visualized using UCSC genome browser. For quartile analysis average Z-score was found for each repetitive element, defined be RepeatMasker in the specific class. Repeats were then binned into quartiles and signal plotted for hESCs and differentiated cell types.

The respective repeats in the top quartile of simple repeats and SINEs in hESCs were selected an intersected between H1-ESC, H7-ESC, H9-ESC. Simple repeats $+/-$ 500 bp were centered and FAIRE signal plotted. Signal was normalized for reads per million mapped and log2 transformed. Repeats that were shared among all three of the ESC were then compared to an equal number of randomly selected repeats from quartile 1 (FAIRE-negative) for DNA characteristics. G/C content was measured using KentUtils from ENCODE. FAIRE-enriched (FAIRE $+$) simple repeats and SINE were identified by intersecting FAIRE peak summits $+/-$ 1 bp with repetitive elements defined by RepeatMasker. Normalized FAIRE signal was again plotted around the repetitive element. For repeat position (Figure 3A) a wiggle file was created where repetitive regions had scores. This was then plotted to demonstrate the position of the repeat relative to FAIRE and DNase signals. TSS were downloaded from ENCODE and CTCF sites were from NHEK cells from ENCODE.

Average FAIRE signal for TSS, CTCF, NANOG, and OCT4 sites were plotted in H1-ESC, H9-ESC, and H7-ESC (Figure 2.1). CTCF sites were derived from CTCF Peaks called in NHEK cells from ENCODE. NANOG binding sites were determined by

peak analysis from H1 ChIP-seq from the ENCODE Consortium. OCT4 binding sites were determined from peak calls in H9 data previously generated (GEO accession number GSM545204). Z-scores were used to assay chromatin changes of repetitive elements during differentiation. Average z-score for each repeat in simple repeats and SINE was determined. Statistically different repeats were determined between two groups: H1-ESC, H7-ESC, H9-ESC, and BM-MSC (p-value < 0.01 for SINE and p-value < 0.05 for simple repeats, t-test). In order to filter out background signal we required that each repeat pass a z-score threshold ($row_{max} - row_{min} > 3$).

2.4.6 DNase-Seq

DNase I hypersensitivity was performed on primary BM-MSC as previously described [96] and sequenced (Illumina). Resulting sequencing reads were in-silico clipped to 20 bp and aligned to the hg19 using BWA. DNase data for H1-ESC was downloaded from the ENCODE.

2.4.7 Analysis of unaligned reads

Raw sequencing reads that bowtie failed to align to the genome, or were determined to have multiple mapping sites (m=4) were captured. Reads were then converted into FASTA files. Identical sequences were collapsed and resulting file was assayed for repetitive elements using RepeatMasker.

2.4.8 Genomic Window Analysis and segmentation

Z-scores were calculated for each nucleotide position excluding chromosomes M and Y. Scores were then averaged over 500bp non-overlapping windows. Significance was assayed between hESCs and HUVEC, NHEK, and K562 by t-test (p-value < 0.01). To filter out regions of background signal we required that a maximum – minimum value $>$ three. Data was clustered using Cluster 3.0 (`http://bonsai.hgc.`

jp/~mdehoon/software/cluster/software.htm) and visualized using Java Treeview [194] or GENE-E (Website reference : http://www.broadinstitute.org). Significantly altered windows were then intersected with ChromHMM segmentation coordinates obtained from ENCODE (http://hgdownload.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeBroadHmm) using BEDTools [195]. The segmentation-based association with the greatest fractional window coverage was selected for those regions mapping to more than one segmentation. The 15 segmentations patterns were then merged as follows:

Promoters: 1 Active Promoter, 2 Weak Promoter, 3 Poised Promoter Enhancers: 4 Strong Enhancer, 5 Strong Enhancer, 6 Weak Enhancer, 7 Weak Enhancer Insulators: 8 Insulator Transcription: 9 Txn Transition, 10 Txn Elongation, 11 Weak Txn Repressed: 12 Repressed Heterochromatic/Repetitive/CNV: 13 Heterochrom/lo, 14 Repetitive/CNV, 15 Repetitive/CNV

2.4.9 Segmentation permutation

An equal number of random FAIRE windows was selected from H1 or HUVEC passing a threshold where at least one sample had a z-score of 3. Regions were then assigned a classification as before. We conducted 1000 iterations.

2.4.10 Simple repeat and SINE permutation

103,345 of simple repeats and 442,460 SINEs, representing the number found in a quartile, was randomly selected from all available repetitive elements three times, representing the three samples. These regions were then intersected, and the shared regions were calculated. 1000 iterations were conducted.

2.4.11 Correlation Analysis

The union set of the top quartile of FAIRE signal at simple repeats and SINE was summed in H1, H7, H9, and HUVEC. Log_2 transformed signal was then plotted against each respective cell type and spearman correlation measured. For TSS and CTCF, FAIRE and DHS signal was summed $+/-$ 300bp from the center of TSS or $+/-$ 200bp CTCF. Log_2 transformed signal was then plotted against each respective cell type and spearman correlation measured

2.4.12 Histone modification analysis at repetitive elements

Available H1 histone modifications were downloaded from the Epigenome Roadmap. Average signal was plotted for FAIRE positive simple repeats and SINE, FAIRE negative simple repeats and SINE, TSS, and CTCF sites.

2.4.13 RNA-Sequencing and analysis

RNA was isolated from BM-MSCs using Trizol according to manufacturer?s directions. Ribosomal RNA depletion and preparation for sequencing were performed as described previously [161]. Sequencing reads were aligned to hg19 using TopHat [196]. RPKM were calculated and used for comparison against H1-ESC. H1-ESC RNA-seq data was downloaded from the ENCODE Consortium.

2.4.14 Salt Fractionation

H1-ESCs and HKCs were cultured in normal growth conditions. 10 million cells were counted using the Bio-Rad TC20 and treated as previously described [175] with the following modifications. 5 U of MNase I was used to digest nuclei. Extractions were performed at 40 mM, 80 mM, 160 mM, 320 mM, and 640 mM NaCl in a volume of 50 μL.

2.4.15 Salt Fractionation MNase-Seq

Nucleosome protected DNA was isolated from successive treatments of low (160 mM) and high (640 mM) salt as previously described above. Resulting DNA fragments were run on a gel and size excluded from 50-500 bp. DNA was purified using a gel extraction kit (Qiagen) with slight modification. 6 volumes Buffer QG and 2 volumes of isopropanol were used and samples allowed to dissolve at RT. Purified DNA was library prepped according to manufacturer's instructions (Illumina) and paired-end sequenced.

Paired reads were filtered using TagDust and aligned to the human genome (hg19) using bowtie with max insertion size of 1000 and Maq rounding disabled. Distribution of MNase fragment sizes was calculated from paired-read lengths. There are two peaks clearly shown, indicating the co-existence of mono-nucleosomes and di-nucleosomes. A wiggle file was created using DANPOS v2.1.3 [197]. DANPOS is designed mainly for mono-nucleosome MNase-Seq data, so we preprocessed our sequencing data to account for DANPOS center-weighted mapping for reads from di-nucleosomes. According to the fragment size distribution (Figure 2.10), fragments greater than 266 bp were considered from di-nucleosomes and others were from mono-nucleosomes. The empirical distribution of mono-nucleosomes was estimated and used to simulate the fragment length when splitting di-nucleosomes into mononucleosomes in silico. After preprocessing, the fragment size distribution meets the assumption of DANPOS. The resulting wiggle file was used to plot nucleosome density at simple repeats.

CHAPTER 3

PTEN DEFICIENCY MEDIATES RECIPROCAL RESPONSE TO IGF1 AND MTOR INHIBITION[1]

3.1 Introduction

Ewing sarcoma is a malignant bone and soft tissue tumor primarily affecting children and young adults. Despite intensive chemotherapy, surgery and radiation therapy approximately 50% of patients ultimately succumb to the disease. Ewing sarcoma is characterized by chromosomal translocations that fuse a member of the TET family to one of a subset of ETS transcription factors [103][77]. Eighty to eighty-five percent of Ewing Sarcoma tumors contain t(11;22)(q24;q12) generating an in-frame fusion of *EWSR1* to *FLI1*[77]. The resulting chimeric EWS-FLI1 protein is a potent transcriptional modulator that regulates multiple genes implicated in malignant transformation [92][94].

Several lines of evidence support a role for the insulin-like growth factor (IGF) pathway in the development of Ewing sarcoma. EWS-FLI1 regulates *IGF1* in Ewing sarcoma cell lines and is induced by EWS-FLI1 in mesenchymal stem cells [90][104][105]. IGF-1 and its receptor (IGF-1R) are expressed in tumors, and IGF-1 expression in cell lines leads to autocrine activation [106][107]. IGF-1 signaling is necessary for the

[1]This chapter previously appeared as an article in Molecular Cancer Research. The original citation is as follows: Patel M, Gomez NC, McFadden AW, Moats-Staats BM, Wu S, Rojas A, Sapp T, Simon JM, Smith SV, Kaiser-Rogers K, Davis IJ. PTEN deficiency mediates a reciprocal response to IGFI and mTOR inhibition. Mol Cancer Res. 2014 Nov;12(11):1610-20.

survival and proliferation of Ewing sarcoma cells [108][109], transformation of murine fibroblasts by EWS-FLI [110] as well as for normal bone development [111]. The promising results of preclinical trials targeting IGF pathway in Ewing Sarcoma has made it an attractive therapeutic target [112][113][114][115]. However, studies of IGF-1 and IGF-1R inhibitors in early phase clinical trials have shown a limited response rate [116][117][118]. A biomarker predictive of individuals who may respond to IGF1-mediated treatment remains to be identified [119][120]. IGF-1 bound to IGF-1R initiates a signaling cascade through the PI3K pathway resulting in phosphorylation of downstream targets including AKT. Phosphorylation of AKT at serine-473 (S473) and threonine-308 (T308) promotes cell cycle progression, cell survival, migration, and metabolism through differential interactions with multiple substrates including mTOR [121][122]. Signaling through the PI3K pathway is attenuated by PTEN through dephosphorylation of PIP_3[123]. The loss of PTEN results in increased accumulation of PIP3 and AKT activation, which has been associated with poor clinical outcomes [124][125][126]. The loss or mutation of PTEN has been demonstrated in a range of cancers [124][125][126][127][128]; however, the function of PTEN in Ewing sarcoma has yet to be investigated. Here we describe PTEN loss in Ewing sarcoma and its consequences on IGF and mTOR signaling, as well as on biochemical responses to small molecule inhibitors. PTEN deficiency augments PI3K signaling to AKT while diminishing cellular responsiveness to IGF inhibition. Interestingly, PTEN loss enhances sensitivity to autophagy induced by mTOR inhibition. Together these data suggest how PTEN loss may influence the response to biological therapies in Ewing sarcoma.

3.2 Materials and Methods

3.2.1 Fluorescent *In Situ* Hybridization

The RP11-383D9 (D9) and RP11-846G17 (G17) BACs were obtained from the Children's Hospital Oakland Research Institute. Bacterial cultures of both BACs were grown in LB with 25 μg/mL chloramphenicol and DNA extracted using Qiagen Plasmid Midi Kit with slight modifications (10 mL of Buffer P1, P2, and P3 and DNA was eluted in 1 mL increments using prewarmed Buffer QF at 65°C). Probes were made using 1 μg of BAC DNA by nick translation (Abbott Laboratories, cat #32-801300) with Red-dUTP (Abbott, cat # 02N34-050) according to manufacturer's protocol. A Chromosome 10 centromeric probe (CEP, Abbott Laboratories) was used as a control. Cell lines were trypsinized, washed, and then resuspended in a small volume of PBS. 10 mL of KCl at 37°C was added dropwise with gentle agitation for the first 2 mL. After adding KCl, the solution was mixed and placed in a 37°C water bath for 12 min after which 1 mL of fresh cold 3:1 methanol:acetic acid (fixative) was added. Cells were collected by centrifugation (10 min, 1000 RPM) and the pellet was resuspended in 10 mL of fresh cold fixative which was added dropwise with gentle agitation for the first 2 mL and incubated at room temperature for 10 min. This process was repeated twice. BAC and CEP probes were then hybridized to each cell line before imaging. PTEN and CEP signals were manually counted from at least 20 nuclei in five separate fields.

3.2.2 Cell culture and Antibodies

Unless otherwise indicated, EWS502, EWS894, and RD-ES cell lines were cultured in RPMI supplemented with 15% fetal bovine serum. A673 and MHH-ES-1 cell lines were cultured in RPMI supplemented with 10% fetal bovine serum. SK-ES cells were cultured in McCoy's 5A supplemented with 15% fetal bovine serum. SK-N-

MC cells were cultured in DMEM supplemented with 10% fetal bovine serum, 2 mM L-glutamine and 1X nonessential amino acids. EWS502 and EWS894 were kindly provided by Dr. Jonathan Fletcher (Brigham and Women's Hospital, Boston) and A673 by Dr. Stephen Lessnick (Univ. of Utah). Other cell lines were obtained from the DSMZ (Braunschweig, Germany). HUVEC cells (Lifeline Technologies) were cultured in Vasculife Basal Media (Lifeline Technologies) supplemented with 10% FBS. CD99 (clone 12E7, Ready-to-use, PA0559, Leica Microsystems) and PTEN antibodies (clone 138G6, 9559S, Cell Signaling Technology) were used for IHC and IF. AKT (#4691), pAKT T308 (#2965), pAKT S473 (#4060) and LC3B (#3868), cleaved PARP (#5625) were used for immunoblotting (Cell Signaling Technology).

3.2.3 Cell Proliferation, Apoptosis, Soft agar, and Autophagy

pLL5.0-PTEN (which expresses PTEN), pLL5.0-shPTEN (which expresses an shRNA directed at PTEN 5'-GTATAGAGCGTGCAGATAG-3') and pLL5.0-shNS (which expresses a non-specific shRNA as a control) were kindly provided by Dr. James Bear (UNC-Chapel Hill). Lentivirus was produced as previously described [129]. EWS502 cells were transduced with either pLL5.0-PTEN or vector control lentivirus in the presence of polybrene (6 μg/mL) for 3 hours, after which media was changed and the cells split for proliferation, soft agar assays. Cells were stained with trypan blue and counted using a hemocytometer to assay proliferation. For soft agar, 0.6% agar was used as the base layer and 0.5% agar as the top layer. The plates were counted manually using ImageJ (NIH). Apoptosis was assessed using the Annexin V-Cy3 Apoptosis Detection Kit (Sigma-Aldrich) according to the manufacturer's protocol. Flow cytometry was performed using the CyAn ADP (Beckman-Coulter). For assessment of autophagy, three days after lentiviral transduction A673 and EWS502 cells were split 1:3 and treated with 20 μM chloroquine for 3 hours or chloroquine followed by 10 ng/mL temsirolimus (LC Laboratories) for 20 hrs. Cells were lysed in

CHAPS buffer and extracts were separated by SDS-PAGE.

3.2.4 IGF-1 inhibition

Cells were treated with NVP-AEW541 (Cayman Chemical) and OSI-906 (ChemieTek) at the indicated concentrations. Prior to treatment with IGF-1, cells were kept in serum-free media for two hours in combination with the IGF-1 inhibitor. Cells were then treated with IGF-1 (Cell Signaling Technologies) for 15 min and lysed in RIPA buffer (25 mM Tris-HCl pH 7.6, 150 mM NaCl, 1% NP-40, 0.1% SDS) supplemented with 200 mM $NaVO_4$ and 50 mM NaF. Cell extracts were separated by SDS-PAGE and blotted with anti-phospho AKT and imaged (LiCor). For assessment of cell viability, EWS502 cells were transduced with lentiviral pLL5.0-PTEN or pLL5.0 as a vector control. 24 hours post infection the cells were treated with NVP-AEW541 in complete media. Viability was assayed 72 hours following NVP-AEW541 treatment using WST-1 (Roche).

3.2.5 Tissue microarray (TMA) and Cell Line Array (CLA) construction

Pellets from the Ewing sarcoma cell lines were fixed in 10% buffered formalin (SF98-4, Fisher Scientific) for 16-24 hours, washed twice in 70% ethanol, clotted in 2% low-melting agarose (Fisher), and then embedded in paraffin wax. Blocks were sectioned and stained with hematoxylin and eosin (H&E, Hematoxylin 7211, Eosin 7111, Richard-Allan). Three 1 mm cores were removed and embedded into recipient CLA block. For TMA construction, Ewing sarcoma cases ($n = 25$) and controls (breast carcinoma, and PTEN-deleted sarcoma) were selected from The University of North Carolina Surgical Pathology and St. Jude Children's Research Hospital archives under an IRB-approved protocol. Hematoxylin and eosin (H&E) stained slides were re-reviewed and representative areas of tumor were marked for coring. TMA blocks, containing triplicate 0.6 mm cores per case were constructed. TMA and CLA blocks

were cut into 4 and 5 micron sections respectively and placed on positively charged glass slides.

3.2.6 Immunohistochemistry (IHC) and Immunofluorescence (IF)

TMA and CLA slides were stained with CD99 and PTEN antibodies (Bond fully-automated slide staining system, Leica Microsystems). Slides were deparaffinized (Bond, AR9222) and hydrated in wash solution (Bond, AR9590). Epitope retrieval (pH 9.0, AR9640, Bond) was performed followed by a peroxide blocking step (Bond DS9800). CD99 and PTEN (1:400) antibodies were incubated for 15 and 30 minutes, respectively then secondary antibody was applied (polymer, Bond DS9800). Chromogenic detection with 3,3'-diaminobenzidine (DAB) and hematoxylin was performed (Polymer Refine Detection, DS9800, Bond). Stained slides were dehydrated and mounted. For fluorescent detection, the TSA-Cy5 reagent (PerkinElmer), Hoechst 33258 (Invitrogen) and ProLong Gold antifade reagent (Molecular Probes) were used.

3.2.7 Imaging and digital image analysis

IHC stained TMA sections were digitally imaged (Aperio ScanScope XT, Aperio Technologies). High-resolution DAPI and Cy5 IF images were obtained (Aperio ScanScope FL). For digital images from IHC slides, Aperio's Cytoplasmic algorithm was used to determine the percentage and intensity of cells positive for PTEN or CD99. A PTEN-deleted tumor control was used to set the negative/low positive intensity threshold for the PTEN stained TMA slide. IF signal was quantified (Definiens Tissue Studio, version 3.6).

3.3 Results

3.3.1 A subset of Ewing Sarcomas lack PTEN

We recently reported widespread alterations in chromatin structure and histone modifications in Ewing sarcoma cells using high-throughput sequencing [90]. Although the experiments performed were intended to detect nucleosome-depleted regions of chromatin, background signal from Formaldehyde Assisted Isolation of Regulatory Elements (FAIRE-seq) typically covers the remainder of the genome mappable by short sequencing read and thus offers a genome-wide sampling of DNA content. Unexpectedly, we observed an approximately 1 Mb region on chromosome 10 that demonstrated a nearly complete loss of FAIRE-seq signal, which we hypothesized to indicate homozygous deletion (Figure 3.1 A). The potential deletion encompassed several genes including the terminal exons of *PTEN* (Figure 3.1 A).

Since deletion of *PTEN* had yet to be detected in Ewing sarcoma using high throughput sequencing approaches, we verified this deletion by fluorescence in situ hybridization (FISH) using two probes that overlap the PTEN locus, as well as a control centromeric probe. One probe (G17) is fully contained within the deleted region whereas half of the second probe (D9) was predicted to hybridize outside the deletion (Figure 3.1 A). Probes were hybridized to seven Ewing sarcoma cell lines (EWS502, EWS894, A673, MHH-ES-1, SK-ES, RD-ES-1, SK-N-MC) and one control cell line (HUVEC). The absence of signal from the G17 probe in EWS502 cells confirmed a homozygous deletion at this region (Figure 3.1 B). Signal from the D9 probe was detected which likely results from hybridization to the retained region centromeric to the deletion. Signal was observed for both probes in the other Ewing sarcoma and control cell lines. However, EWS894 and SK-N-MC cells exhibited PTEN/centromeric probe ratios not equal to one suggesting other cytogenetic aberrations involving the long arm or centromere of chromosome 10 (Figure 3.1 B, C). EWS894 had two copies

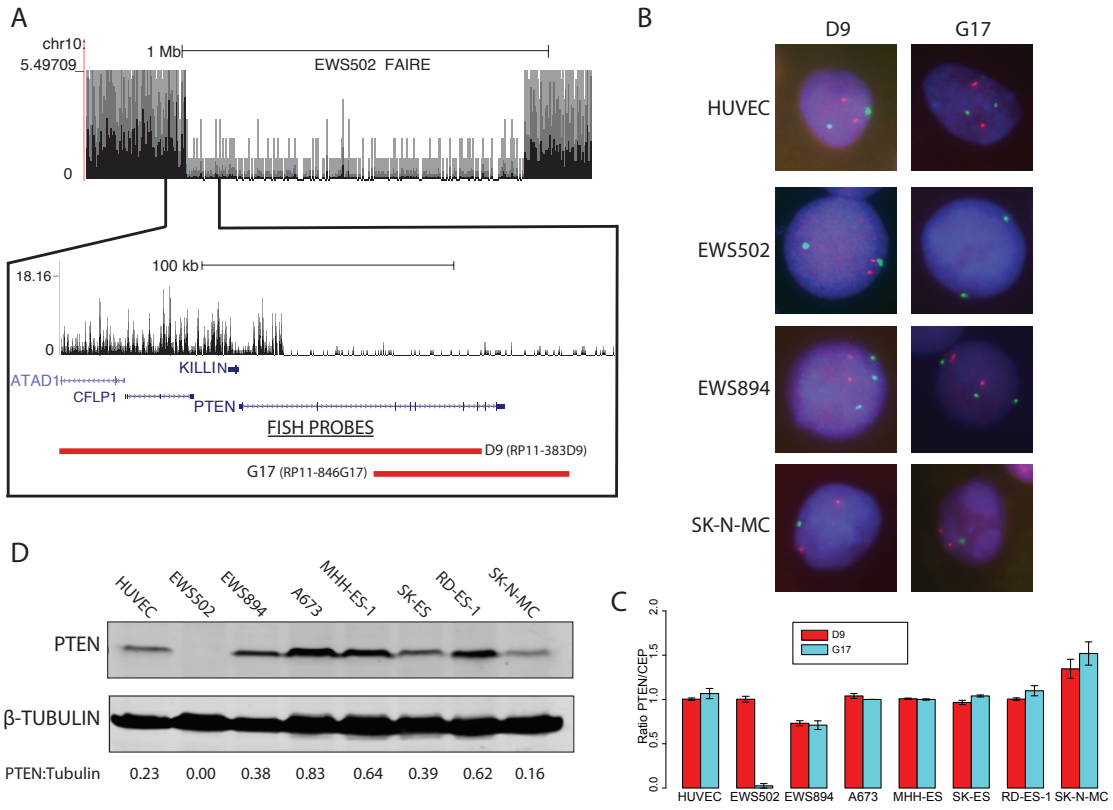


Figure 3.1: Identification of PTEN deletion in Ewing Sarcoma. A) FAIRE-seq derived high throughput sequencing tag density around the PTEN locus in EWS502. Red bars indicate the regions of hybridization for the FISH probes. Only genes located within the deletion are shown. B) PTEN FISH for representative Ewing sarcoma cell lines and a control (HUVEC). Chromosome 10 centromeric probe (green) and PTEN BAC probes (red) are shown. C) Observed ratio of PTEN probe to centromeric probe (CEP) signal. Error bars indicate standard deviation of PTEN/CEP probe ratio from five unique fields counting a minimum of 20 cells per field (with the exception of SK-ES in which 20 nuclei were analyzed). D) Immunoblot of PTEN in Ewing sarcoma cells. Extracts of EWS502, EWS894, A673, MHH-ES-1, SK-ES, RD-ES, and SK-N-MC and a control cell line, HUVEC were blotted with anti-PTEN antibody. Tubulin was used as a loading control.

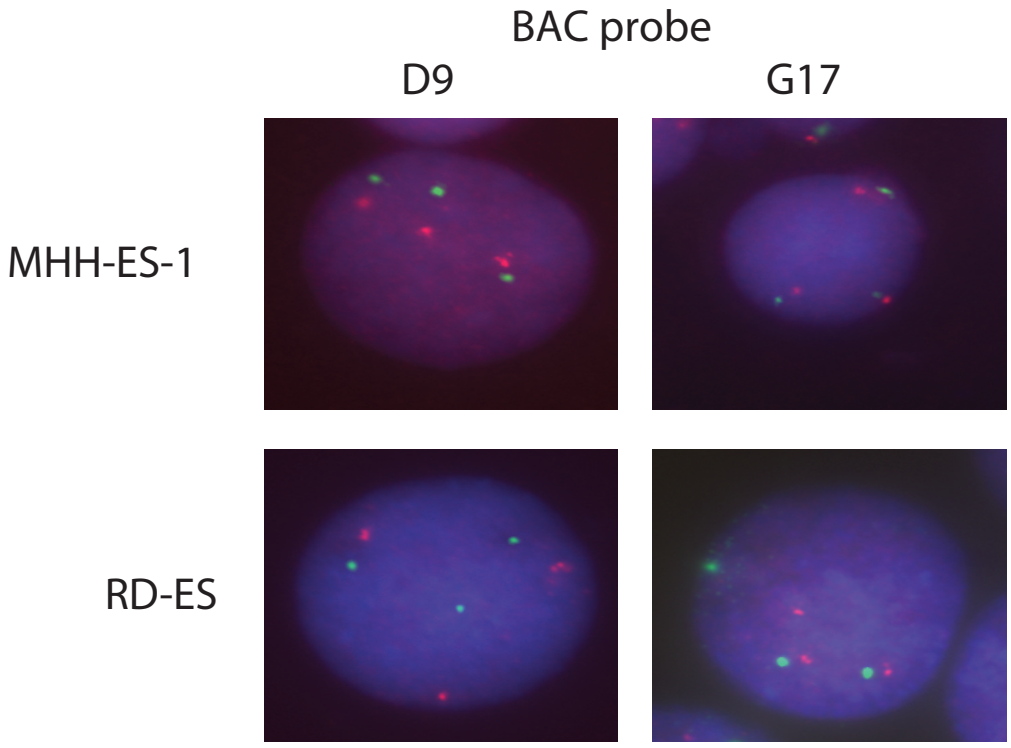


Figure 3.2: MHH-ES-1 and RD-ES are triploid for chromosome 10. Interphase nuclei of indicated cell lines were hybridized with two BACS (D9 and G17) targeted to the PTEN locus (red) and centromeric probe (green).

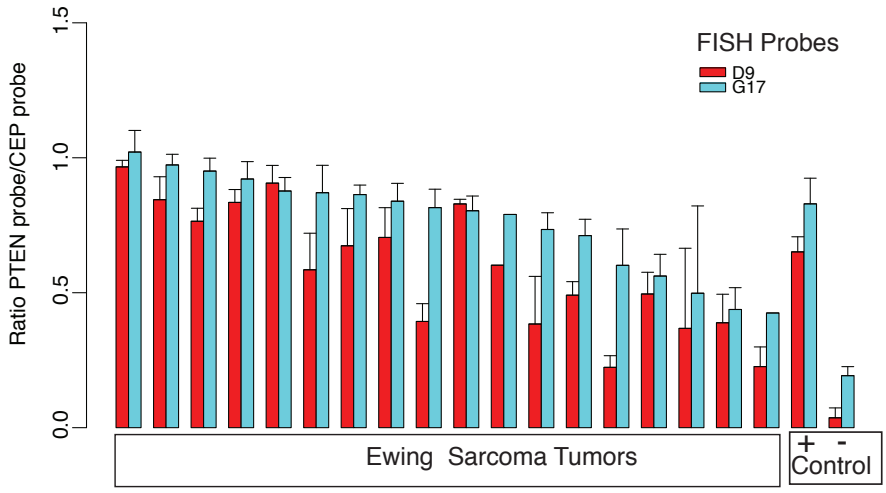


Figure 3.3: Sections of the Ewing sarcoma tissue microarray were hybridized with D9 and G17 PTEN BACS as well as a chromosome 10 centromeric control (CEP). The PTEN/centromeric ratio is plotted. Breast cancer ($+$ control) and undifferentiated sarcoma ($-$ control) are shown.

of the PTEN locus but three copies of the centromeric probe whereas one copy of the PTEN locus and two copies of the centromeric probe were detected in SK-N-MC. The PTEN/centromeric probe ratio was equivalent for the remaining cell lines; MHH-ES-1 and RD-ES exhibited triploidy of chromosome 10 (Figure 3.2). Consistent with PTEN chromosomal loss, PTEN protein was absent in EWS502 whereas other Ewing sarcoma cell lines showed variable levels (Figure 3.1 D).

In order to address whether PTEN was similarly lost in primary Ewing sarcoma tumor, we generated a tissue microarray consisting of 25 tumors diagnosed as Ewing Sarcoma during clinical evaluation. The samples were re-reviewed prior to microarray generation, and tumor-specific regions were selected for core preparation.

Each tumor was represented in triplicate at random positions on the array. Microarray sections were hybridized to both FISH probes. For the 20 tumors from which FISH signal was interpretable, homozygous loss was not observed however copy number varied across tumors (Figure 3.3). Since PTEN expression can be affected by mechanisms other than deletion, we analyzed PTEN protein levels by immunofluorescence (IF) and immunohistochemistry (IHC). A Ewing sarcoma cell line array was generated to validate antibody-mediated detection of PTEN. PTEN detection by IHC and IF on the cell line array quantitatively matched detection by Western blotting (r^2

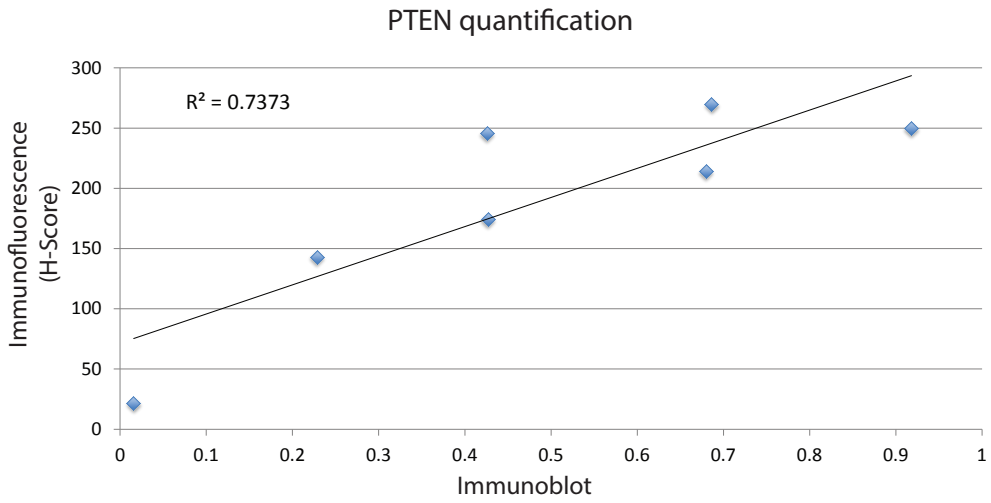


Figure 3.4: PTEN immunofluoresence is concordant with immunoblotting. Immunofluorescence for PTEN was performed on the CLA with anti-PTEN antibody and quantified. Extracts of each cell line were separated by SDS-PAGE, immunoblotted with anti-PTEN antibody and quantified. The average H-score was plotted against the corresponding immunoblot quantification normalized to tubulin (arbitrary units). Best fit line is shown.

$= 0.74$, Figure 3.4). Because of the diverse age of the samples that contributed to the primary tumor array and the evolving criteria for Ewing sarcoma diagnosis, we performed IHC and IF for CD99 as confirmation of diagnosis and as a quality control. IHC and IF for CD99 as well as PTEN were highly concordant (Figure 3.5). After eliminating CD99 negative tumors and those with poor staining 15 tumors remained. A wide range of PTEN expression was detected by IF among the Ewing sarcoma samples. Three tumors demonstrated significantly reduced signal when compared to a PTEN-expressing control breast carcinoma sample and a PTEN-deficient undifferentiated sarcoma (Figure 3.6 A). Histological examination suggested that non-tumor cells confounded accurate PTEN quantification. We attempted IF for CD99 to specifically identify tumor cells, but due to technical constraints co-staining of PTEN and CD99 was not possible. However, using CD99 IHC in adjacent sections, we confirmed the IF results. We observed that for one additional tumor (tumor 2, Figure 3.6 A and B) 55% of the cells did not demonstrate PTEN signal (Figure 3.7). Remaining PTEN expression in this sample may be related to CD99-negative non-tumor cells or tumor heterogeneity (Figure 3.6 B). These data suggest that PTEN expression is reduced in approximately 25% (4 of 15) of Ewing sarcomas, and that the loss of PTEN is primarily through mechanisms other than large genomic losses. This observation is consistent with other tumors in which PTEN expression is lost due to gene silencing or focal deletions [130][131][132][133](32-35).

3.3.2 PTEN loss in Ewing sarcoma augments AKT signaling

To determine the effect of PTEN loss on AKT signaling across Ewing sarcoma cell lines we examined phosphorylation at S473 and T308. Phosphorylation of these sites is indicative of AKT activation [121][122]. Among the cell lines tested, EWS502 had the highest level of pAKT (Figure 3.8 A). Low levels of S473 phosphorylation was also observed in EWS894, SK-ES, and RD-ES-1 cells. T308 phosphorylation

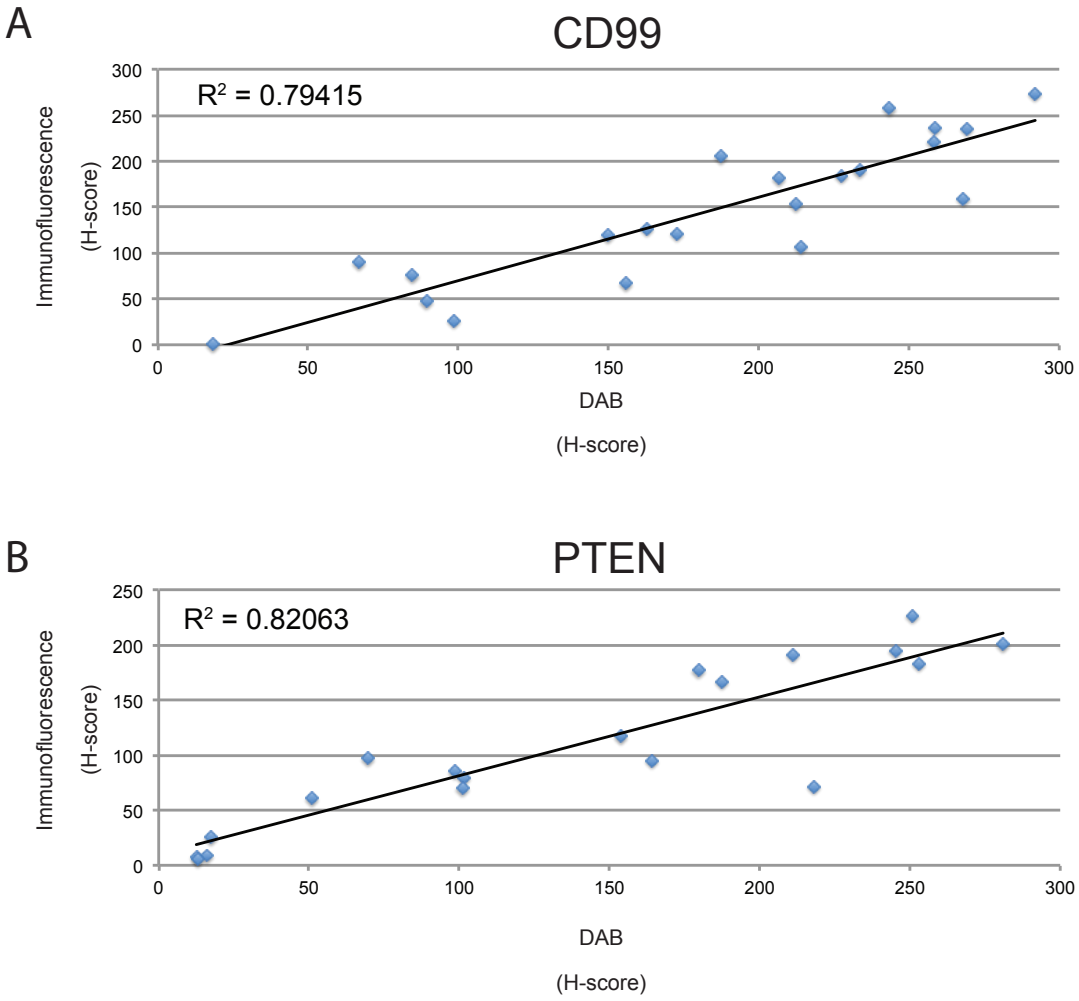


Figure 3.5: Immunofluoresence and immunohistochemical staining are highly concordant. IF or DAB staining for CD99 (A) and PTEN (B) was performed on adjacent TMA sections. Staining was quantified and the average H-score for both IF and DAB was plotted.

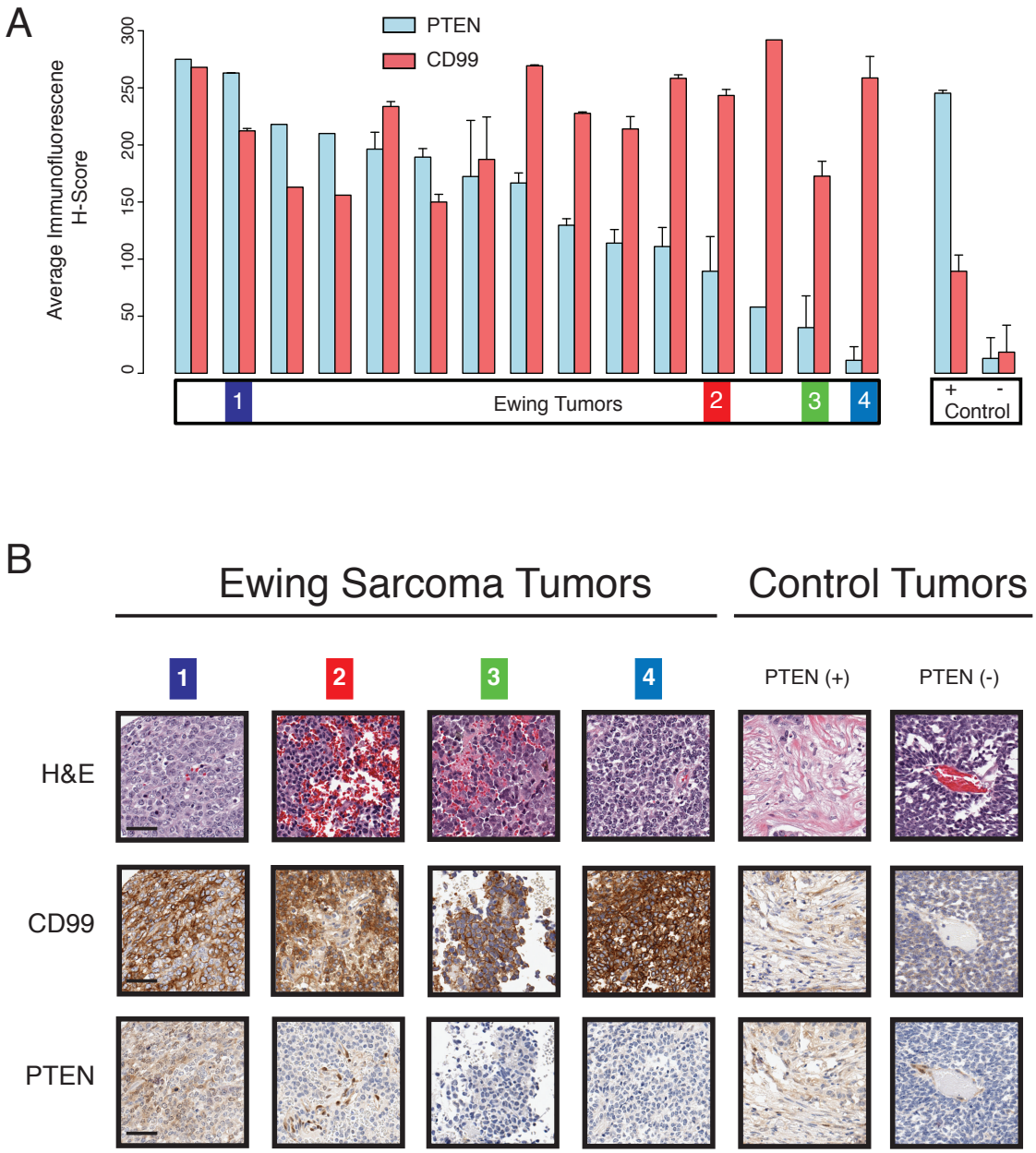


Figure 3.6: Loss of PTEN expression in primary Ewing Sarcoma. A) PTEN and CD99 immunofluorescence in Ewing sarcoma. Average immunofluorescence (H-score) for PTEN (blue bars) and CD99 (red bars) in Ewing sarcoma and control tumors. Positive $(+)$ and negative $(-)$ PTEN control tumors are shown to the right. Error bars represent standard deviation of H-scores between replicate cores. Absence of error bars indicates single core available for analysis. B, Representative hematoxylin and eosin (H&E), CD99, and PTEN diaminobenzidine (DAB) staining. Immunohistochemistry for selected Ewing sarcoma and control tumors (indicated in A) is shown. Scale bar, 50 μm.

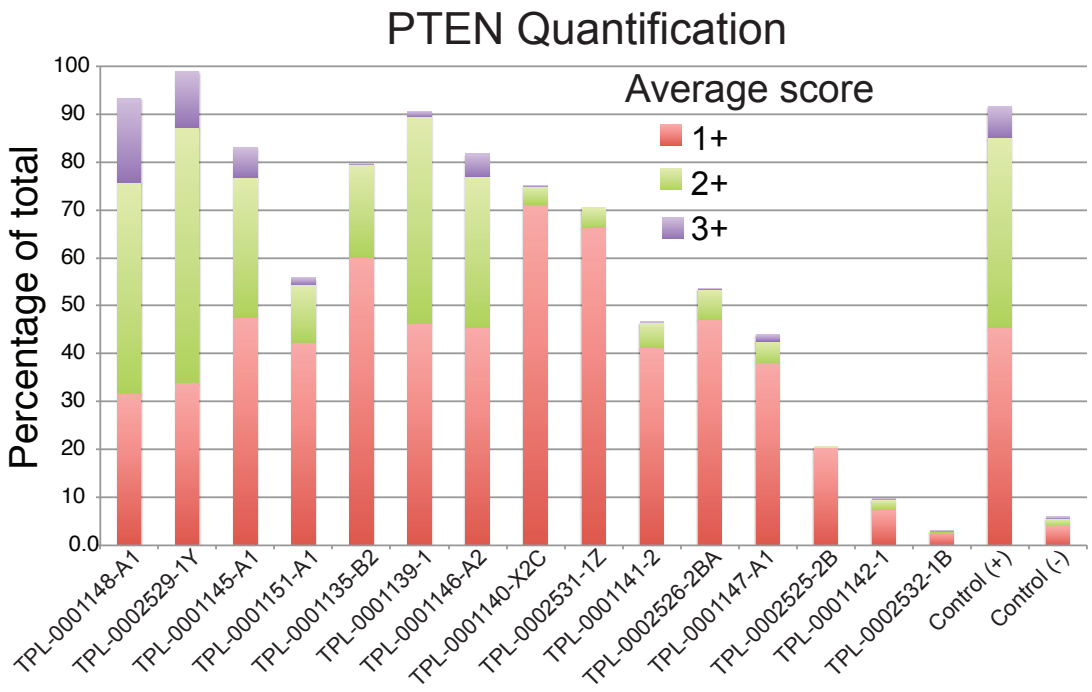


Figure 3.7: Quantification of Cytoplasmic PTEN. TMA sections were immunohistochemically stained for PTEN using DAB. Cells were imaged and quantified. Cytoplasmic staining was scored between 0 to 3+, and scores from the replicates were averaged. The percentage of cells of each score was plotted.

was limited to EWS502. PDK1-associated phosphorylation of T308 is associated with full AKT activation [134][135] and was only observed in the absence of PTEN suggesting that AKT activation is augmented by PTEN loss. We then ectopically expressed PTEN in EWS502 cells to test the association between PTEN levels and activated AKT. Increasing PTEN was associated with a progressive decrease in pAKT at S473 and T308 (Figure 3.8 B) suggesting that AKT activation in EWS502 is due in part to PTEN deficiency. To test whether PTEN loss altered IGF-1 sensitivity, we examined dose-dependent stimulation by IGF-1 under serum-free conditions. AKT demonstrated baseline phosphorylation in all Ewing sarcoma cell lines. IGF-1 stimulation resulted in further AKT activation. However, there was no difference in IGF-1

IC50 (Figure 3.9). These data indicate that PTEN levels influence AKT activation but do not result in enhanced sensitivity to IGF-1.

The cellular effects of PTEN loss in Ewing sarcoma were examined by testing the effect of PTEN on cellular proliferation and anchorage-independent growth. PTEN was transduced into EWS502 cells and expression was confirmed by immunoblotting. PTEN expression resulted in significantly decreased cellular proliferation (Figure 3.8 C). To address whether the reduction in cell proliferation following PTEN expression could be attributed to increased apoptosis, we assayed annexin V reactivity by flow cytometry and observed a significant increase relative to control cells (Figure 3.8 D). We also observed a similar increase in cleaved PARP (Figure 3.10). Anchorage-independent growth as assayed by colony formation in soft agar was also greatly diminished (Figure 3.8 E). Taken together, these data demonstrate that PTEN loss enhances cellular properties associated with transformation in Ewing sarcoma cells.

3.3.3 PTEN loss decreases sensitivity to IGF-1 inhibition

Since clinical trials of IGF-1-targeted inhibitors have demonstrated robust but limited patient responses, we asked whether PTEN loss might mitigate the effect of these compounds in Ewing sarcoma cells. Ewing sarcoma cells were treated with two IGF-1R inhibitors, NVP-AEW541[136] and OSI-906 [137]. NVP-AEW541 has been tested for Ewing sarcoma whereas OSI-906 is an investigational agent for a variety of cancers [113][114][137][138][139][140]. Cells cultured in serum-free media were pretreated with these inhibitors prior to stimulation by IGF-1. PTEN loss was associated with increased IC50 to the IGF-1 inhibitors as measured by AKT activation (Figure 3.11 A). This differential sensitivity was detectable by phosphorylation at both S473 and T308. Interestingly, intermediate sensitivity to these inhibitors was observed for EWS894 and SK-ES, both of which demonstrated lower PTEN levels and detectable pAKT-S473. We then examined the effect of PTEN expression on

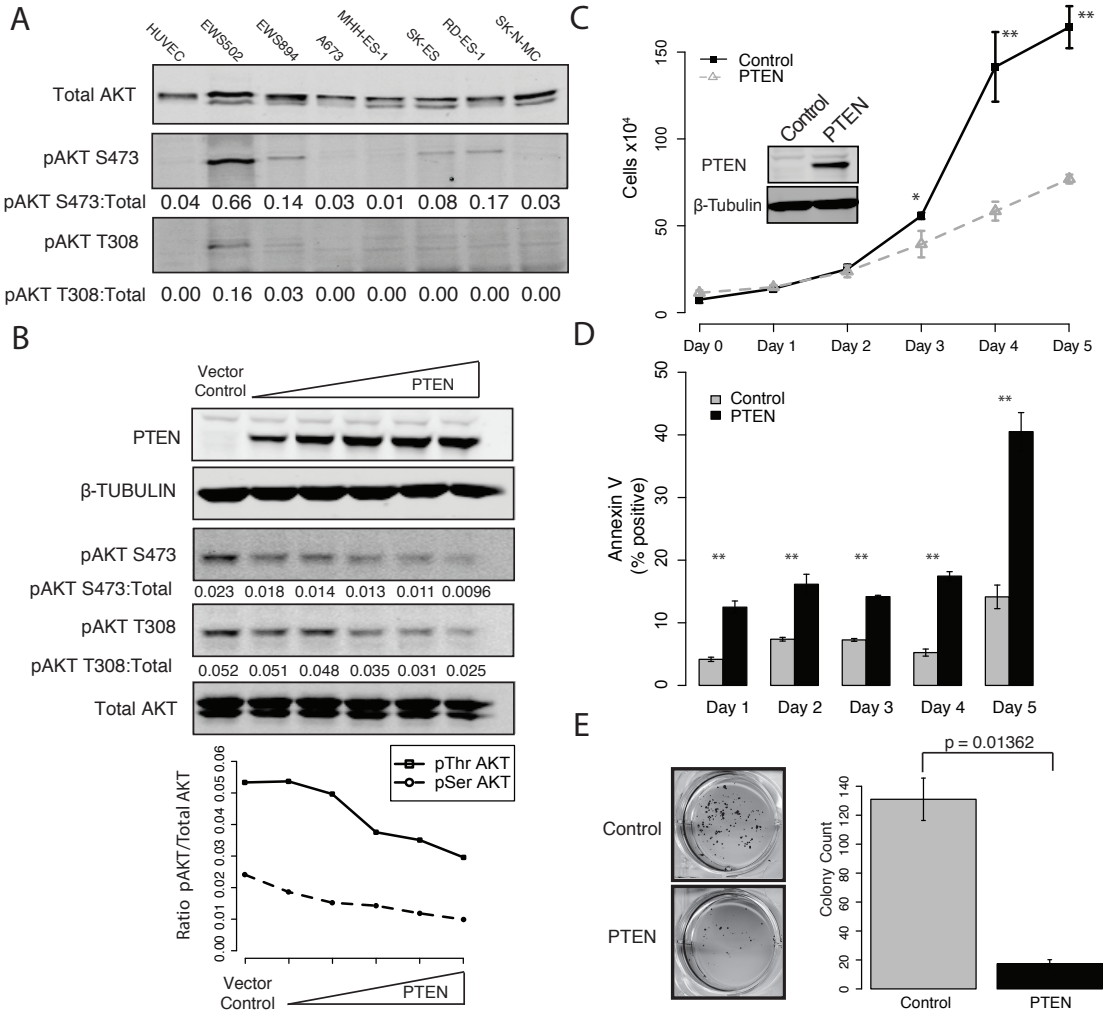
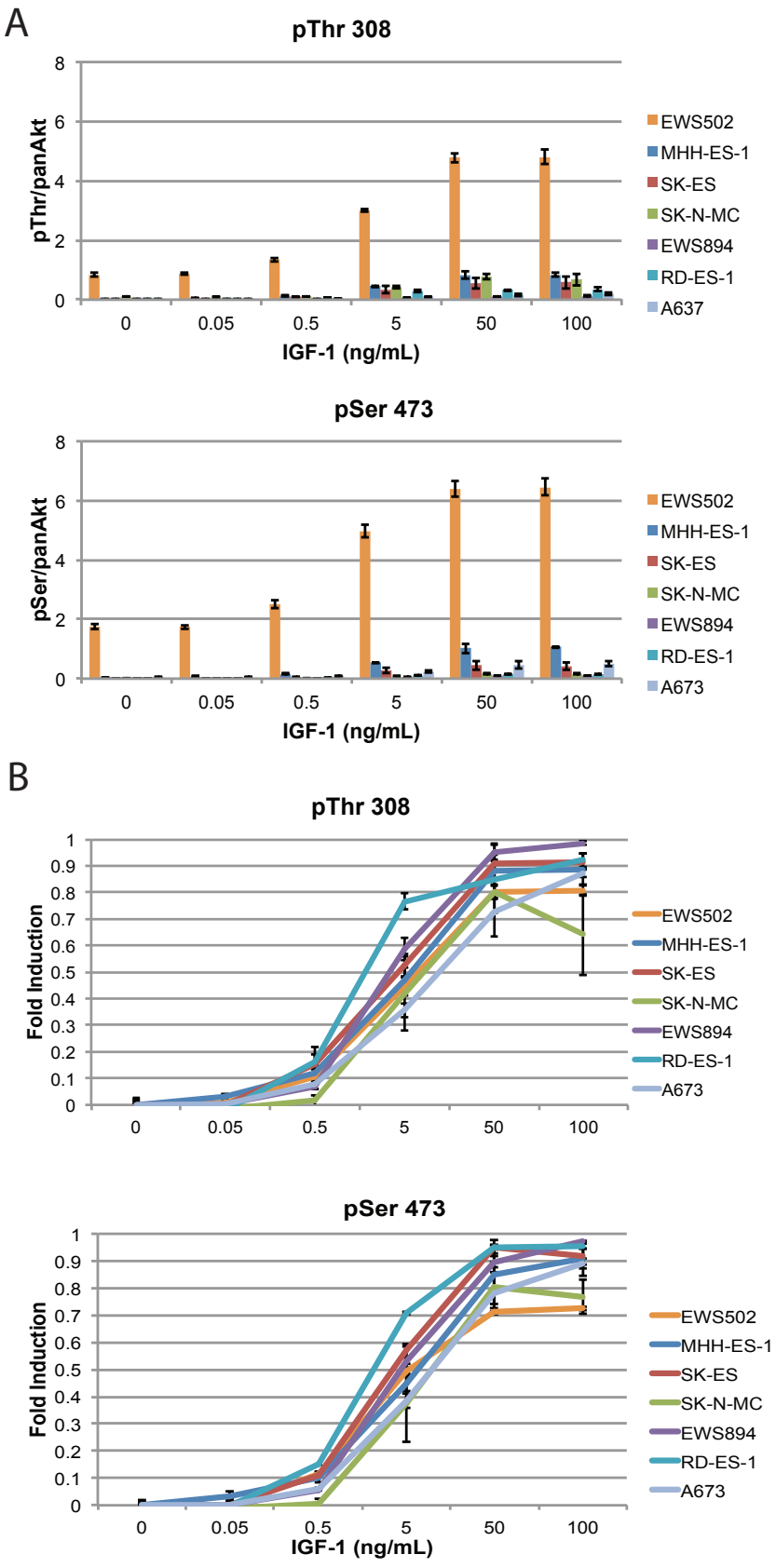


Figure 3.8: PTEN loss enhances AKT signaling promoting transformation. A) Immunoblot for total and phospho-AKT. Ewing cell lines and HUVEC under normal growth conditions were immunoblotted for phospho-AKT at S473 and T308. B) PTEN expression abrogates phospho-AKT. EWS502 were transduced with increasing amounts of PTEN-expressing lentivirus. Extracts were immunoblotted for PTEN, AKT, pAKT S473, pAKT T308, and tubulin. Phospho- and total AKT were quantified. The ratio of pAKT/AKT is shown (bottom). C) PTEN expression reduces cell proliferation. EWS502 were transduced with PTEN on day 0 and cells were counted daily. D) PTEN expression increases apoptosis. Annexin V staining in EWS502 cells transduced with PTEN (black) or a control vector (grey) were analyzed by flow cytometry. Percentages of annexin positive cells are shown. E) PTEN expression reduces colony formation. EWS502 transduced with PTEN or control vector were plated in soft agar. Colonies were stained with MTT for visualization (left) and quantified (right). Colonies greater than 1 mm in size were counted. For each panel, error bars represent standard error between triplicates. $\cdot$ and $\cdot\cdot$ indicate $p < 0.05$ and $p < 0.01$ respectively (two-tailed T-test).


A
pThr 308
pThr/panAkt
IGF-1 (ng/mL)
EWS502
MHH-ES-1
SK-ES
SK-N-MC
EWS894
RD-ES-1
A637
pSer 473
pSer/panAkt
A673
B
Fold Induction

IGF-1R inhibition focusing on NVP-AEW541 due to its selectivity for IGF-1R [136]. Transduced PTEN resulted in enhanced sensitivity for NVP-AEW541 with an IC50 approximating the other PTEN-expressing Ewing sarcoma cells (Figure 3.11 B). The enhanced sensitivity for NVP-AEW541 was associated with increased cellular toxicity (Figure 3.11 C). These data suggest that PTEN loss in Ewing sarcoma diminishes the efficacy of IGF-1R inhibitors on PI3K signaling as well as viability.

3.3.4 PTEN loss enhances response to temsirolimus

AKT signaling acts on the mTOR pathway to influence multiple cellular processes including autophagy [141][142]. In light of the emerging role of mTOR inhibition in Ewing sarcoma treatment, we examined the relationship between PTEN loss and autophagic response to the mTOR inhibitor, temsirolimus. Ewing sarcoma cells were treated with temsirolimus and autophagy was assayed by quantification of LC3BII, a protein localized to autophagosome membranes that is generated during autophagy [143]. The assay was performed in the presence of chloroquine to inhibit lysosomal processing and thus enable assessment of autophagy without ongoing degradation. PTEN-expressing Ewing sarcoma cells (A673) demonstrated minimal LC3BII induction in response to chloroquine or to chloroquine and temsirolimus (Figure 3.12 A). In contrast, EWS502 cells demonstrated a modest induction of LC3BII in response to chloroquine, but this response was significantly increased by temsirolimus (Figure 3.12 B). Since EWS502 but not A673 cells demonstrated induction of temsirolimus-induced autophagy, we examined the effect of modulating PTEN. Silencing PTEN in

---

Figure 3.9: IGF-1 sensitivity is not influenced by the presence of PTEN. Cell lines were serum starved for two hours followed by treatment with IGF-1 treatment at the indicated concentrations for 15 min. Cells were lysed, immunoblotted for pAKT and quantified. Phospho-specific signals were normalized to total AKT. A) Quantification of phospho-T308 and phospho-S473 are shown. B) Fold induction for phosphor-T308 and phospho-S473 was calculated relative to the maximum.

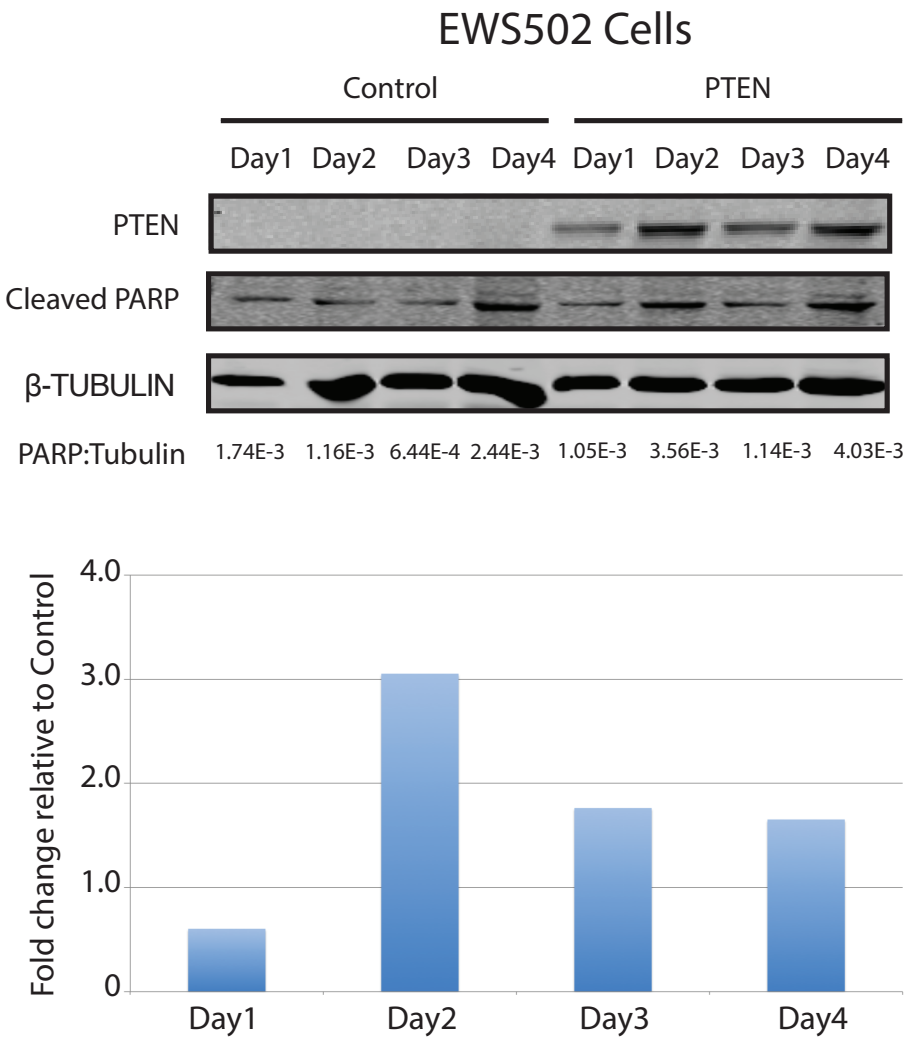
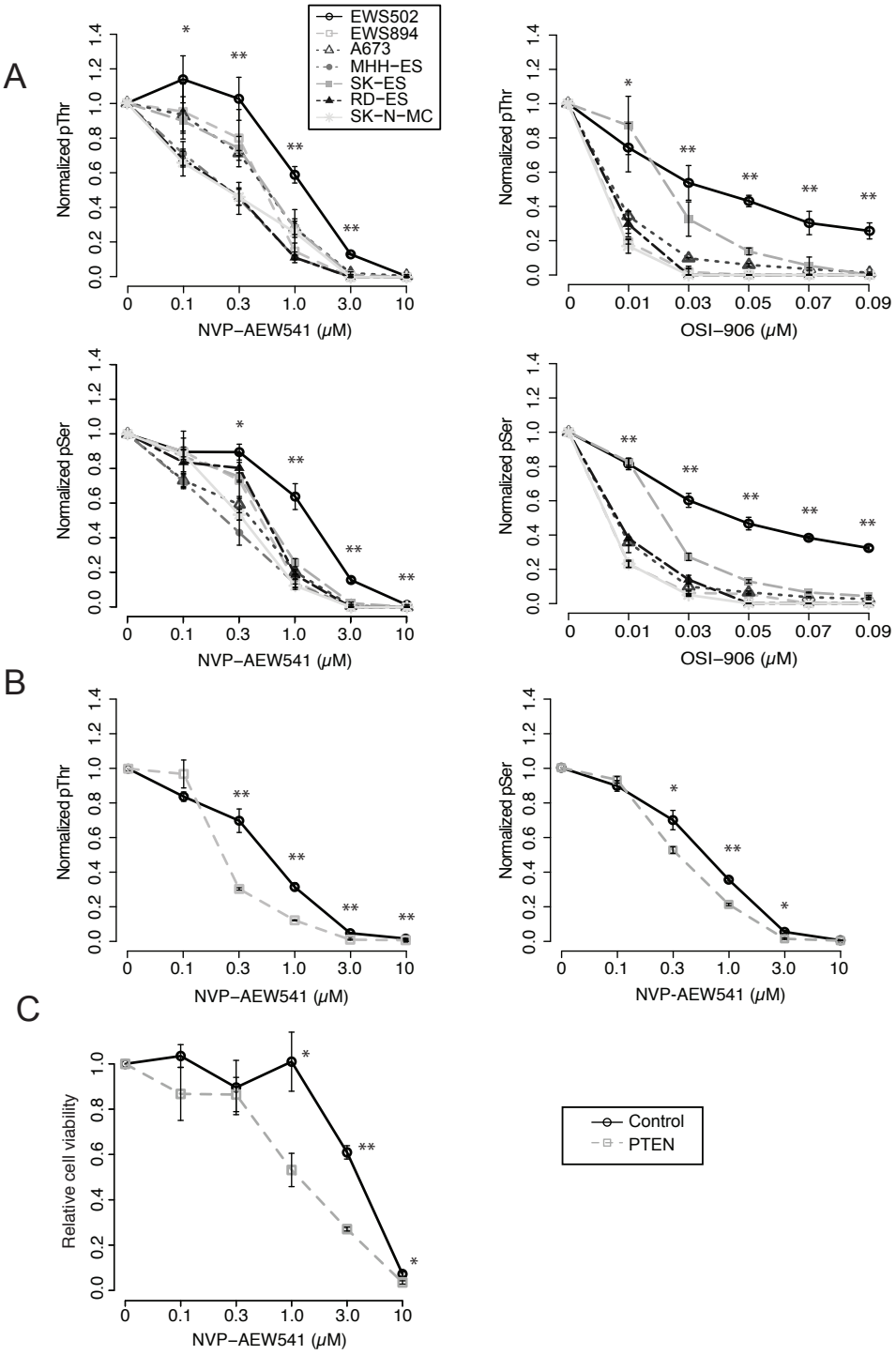


Figure 3.10: Expression of PTEN increases cleaved PARP. EWS502 cells were transduced with lentivirus expressing either PTEN or GFP as a control. Lysates were immunoblotted for cleaved PARP. Cleaved PARP levels relative to control are shown. Immunoblots are normalized to tubulin.


A
B
C
EWS502
EWS894
A673
MHH-ES
SK-ES
RD-ES
SK-N-MC
Normalized pThr
Normalized pSer
NVP-AEW541 (μM)
OSI-906 (μM)
Relative cell viability
Control
PTEN

A673 cells augmented the autophagic response to temsirolimus whereas exogenous PTEN expression in EWS502 eliminated the effect of temsirolimus (but not chloroquine) (Figures 3.12 A and B). We then examined the effect of inhibiting autophagy with chloroquine on cellular viability. Interestingly, treatment with chloroquine attenuated the toxic effects of temsirolimus in the absence of PTEN but this difference was lost when PTEN was expressed. (Figure 3.13). Together these experiments demonstrate that PTEN expression in Ewing sarcoma cells influences autophagic response to temsirolimus with PTEN loss associated with increased responsiveness to mTOR inhibition. Further, the induction of autophagy by temsirolimus is associated with decreased viability, suggesting that autophagy partially mediates the effects of temsirolimus.

3.4 Discussion

The unexpected identification of PTEN deletion in a Ewing Sarcoma cell line led us to explore the status of PTEN in primary tumors. Although we were unable to detect a similar deletion in other cell lines or a set of primary tumors using FISH, quantitative assessment of PTEN expression by IHC and IF suggested that approximately 25%

Figure 3.11: PTEN modulates sensitivity to IGF-1R inhibitors. A) PTEN loss decreases sensitivity to IGF-1R inhibitors. Ewing Sarcoma cells were treated with NVP-AEW541 (left) or OSI-906 (right) at the indicated concentrations for 2 hours in serum-free media and then stimulated with IGF-1 (5 ng/mL final concentration) for 15 min. Extracts were immunoblotted for pAKT at Thr 308 (top) and S473 (bottom) and results were quantified. Relative inhibition was calculated by normalizing pAKT signal to mock treatment (zero concentration). B) PTEN expression increases sensitivity to IGF-1R inhibition. EWS502 transduced with PTEN (dotted grey) or a control vector (solid black) were exposed to NVP-AEW541 for 2 hours in serum-free media and then stimulated with IGF-1 (5 ng/mL final concentration) for 15 min. Relative inhibition was calculated as above. C) PTEN expression increases the cellular toxicity associated with NVP-AEW541 treatment. EWS502 cells were transduced as in B and treated with NVP-AEW541 for 72 hours at indicated concentrations and assayed for viability. For each panel, error bars represent standard error between replicates. $\cdot$ and $\cdot\cdot$ indicate $p < 0.05$ and $p < 0.01$ respectively (two-tailed T-test).

of Ewing sarcoma tumors are PTEN deficient. Small deletions and other mutations undetectable by FISH, in addition to gene silencing, remain alterative mechanisms that result in PTEN loss in Ewing sarcoma. However, our observation of PTEN loss is consistent with a recent study that used high resolution SNP arrays to examine copy number variation in Ewing sarcoma and observed PTEN deletion in 14% of the tumorsLynn:2013gy. We found that PTEN deficiency leads to enhanced AKT activation associated with decreased apoptosis, increased proliferation, and anchorage-independent growth. Enhanced properties associated with cellular transformation in Ewing sarcoma could result in a more aggressive tumor phenotype. Intriguingly, ETS deregulation may cooperate with PTEN loss to accelerate tumorigenesis [144]. Several lines of evidence indicate that mTOR contributes to PTEN-dependent negative feedback regulation of AKT (reviewed in [145][146]). The loss of PTEN in Ewing sarcoma may be one mechanism mediating hyperactivation of AKT even in the absence of growth factors such as IGF-1. In addition to potentially contributing to a more transformed phenotype, hyperactivation of AKT may decrease sensitivity of Ewing sarcoma cells to chemotherapy [147][148].

We have demonstrated that loss of PTEN decreases sensitivity to IGF-1R inhibition, as measured by AKT phosphorylation. Of the limited number of available cell lines tested, there were varying degrees of response to IGF-1R inhibition. An intermediate effect was seen in two cell lines with reduced PTEN expression and increased AKT phosphorylation. These findings are consistent with a prior study demonstrating that PTEN silencing in cultured glioblastoma decreased response to NVP-AEW541[149].

PTEN loss led to increased sensitivity to temsirolimus treatment as marked by the activation of autophagy. Autophagy is a metabolic recycling process in which cellular components are broken down in times of stress to maintain metabolic homeostasis. The role of autophagy in cancer is complex. Our results suggest that autophagy is re-

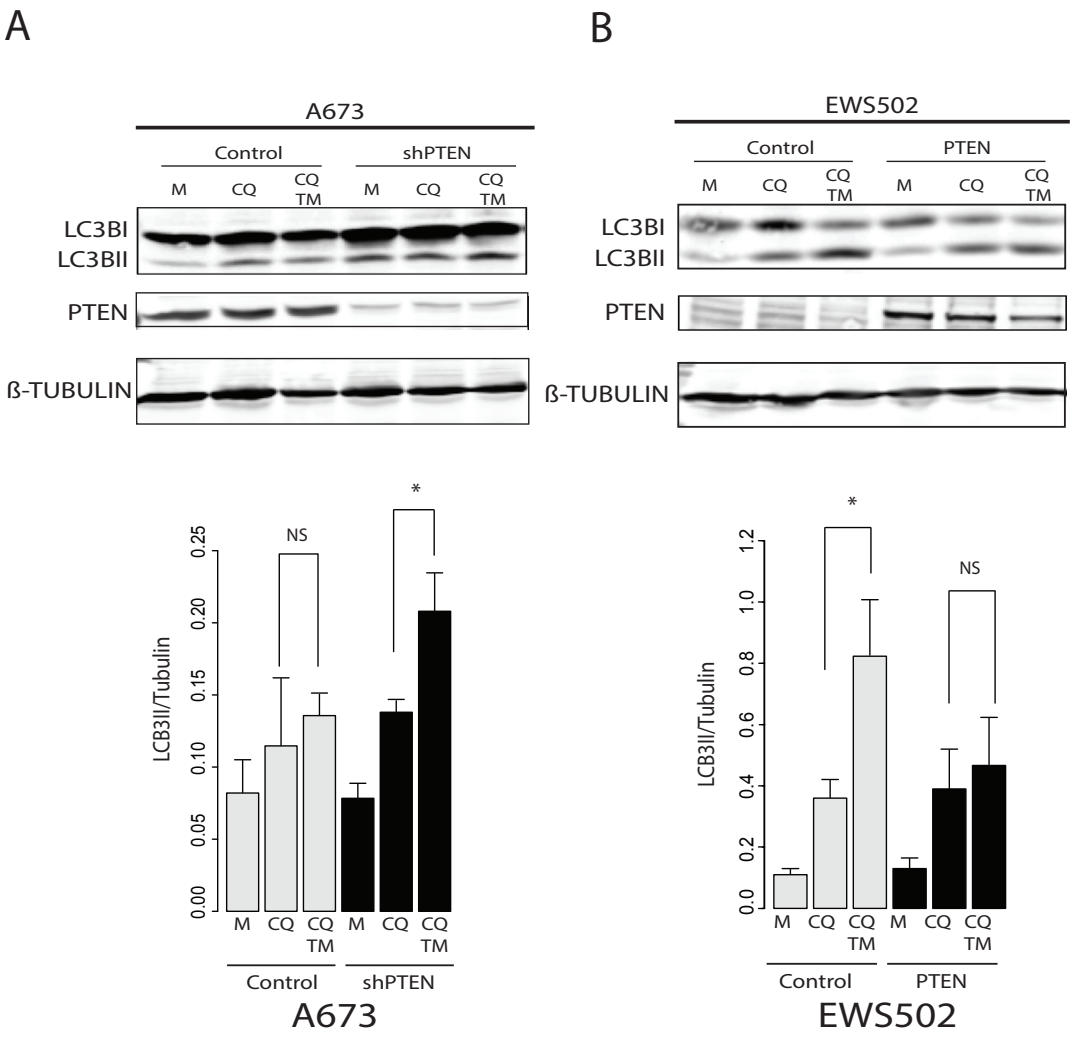


Figure 3.12: PTEN loss potentiates temsirolimus-induced autophagy. A) PTEN silencing enhances induction of autophagy in response to temsirolimus. A673 cells transduced with a PTEN-shRNA (shPTEN) or nonspecific control (shNS) were treated with chloroquine (CQ) alone or chloroquine and temsirolimus (10 ng/mL) for 20 hours (CQ/TM). Chloroquine was added 3 hours prior to the initiation of temsirolimus treatment. Cell extracts were immunoblotted for LC3B, PTEN, and tubulin (top). LC3BII bands were quantified and normalized to tubulin (bottom). B) PTEN expression abrogates induction of autophagy in response to temsirolimus treatment. EWS502 cells transduced with exogenous PTEN or a control vector (GFP) were treated with chloroquine (CQ) alone or chloroquine and temsirolimus (10 ng/mL) (CQ/TM) as described above. Cell extracts were immunoblotted and quantified as above. $\cdot$ and $\cdot\cdot$ indicate $p < 0.05$ and $p < 0.01$ respectively by two-tailed T-test.

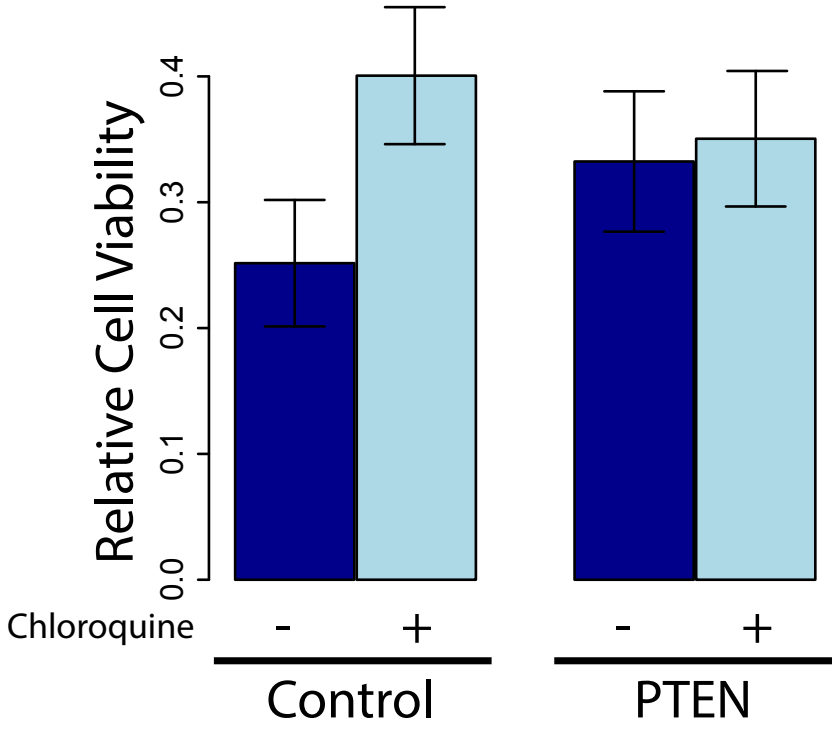


Figure 3.13: PTEN expression reduces the autophagy-mediated viability effects of temsirolimus. EWS502 cells were lentivirally transduced with either PTEN or GFP control, plated at a density of 5 x 10^3 cells/well and allowed to adhere for 24 hours. Cell viability was assayed by WST-1 72 hours after treatment with either temsirolimus (1 ng/mL), or temsirolimus and chloroquine (20 μM). To assess the specific effect of temsirolimus on cell viability results were normalized to DMSO or DMSO $+$ chloroquine, as indicated.

quired to mediate the cell viability effects of mTOR inhibition by temsirolimus. These data are in agreement with studies indicating that induction of excessive autophagy can lead to cell death [150][151]. mTOR inhibitors may constitute a promising therapeutic class for cancers lacking functional PTEN by inducing autophagy-mediated apoptosis.

PTEN deficiency renders cells less sensitive to IGF-1R inhibition but increases autophagic response to mTOR inhibition. The differential response to AKT/mTOR pathway manipulation has therapeutic implications. The promise of personalized therapy for cancer depends on the identification of genetic alterations in specific tumors. The limited efficacy of IGF-1R inhibition offers an opportunity for the application of relevant biomarkers. Our results indicate that loss of PTEN expression may diminish the therapeutic response of Ewing sarcoma to IGF-1R inhibitors. However, our study also suggests a reciprocal interaction between PI3K/AKT signaling and autophagy. Whereas PTEN loss decreased sensitivity to IGF-1R inhibition, it enhanced sensitivity to temsirolimus. These data suggest that patients who are unresponsive to IGF-1R inhibition may benefit from mTOR inactivation. The application of PTEN expression as a biomarker to future clinical trial would be needed to directly assess this possibility. Due to interactions between the IGF-1R and mTOR pathways, combination of IGF-1R- and mTOR-directed therapies are being evaluated in preclinical and early phase clinical trials with evidence of efficacy [152][153][154][155]. The ability to identify and apply relevant prognostic biomarkers during the selection of biologically active therapies may greatly increase the possibility of therapeutic benefit.

CHAPTER 4

CONCLUSIONS AND FUTURE DIRECTIONS

We had originally intended to assay the chromatin changes that accompany differentiation, but were surprised to find a previously unreported widespread enrichment of repetitive elements in stem cell chromatin. Following up on this observation, we confirmed that two classes of repetitive elements, simple repeats and SINEs, were significantly enriched in FAIRE-positive chromatin of three independent hESC lines. We demonstrated that FAIRE-positive repetitive elements are marked by acetylated histones and associated with genes involved in lineage specific developmental pathways. The high concordance of signal across three cell lines, and specific histone acetylation enrichment suggests that these regions are playing a role in general stem cell biology. Though, acetylation have previously been shown to be important for maintaining pluripotency, these studies have generally focused on pan acetylation [210][211]. These data, coupled with the our observation of repetitive element chromatin closing during differentiation, suggests an intriguing hypothesis that open chromatin at repetitive elements is necessary for the multipotency of these cells. Treating hESC with inhibitors targeting the specific HATs responsible for the acetylation found at repetitive elements could lead to more condensed chromatin resulting in spontaneous differentiation. Alternatively, the condensed chromatin environment could have the opposite effect and reduce the efficiency of differentiation, suggesting that the chromatin environment at these repetitive elements are important for downstream de-

velopmental processes. Regardless of the outcome, these experiments represent a targeted approach to interrogate the importance of specific histone acetylation marks and repetitive elements in hESC physiology.

We demonstrated chromatin remodeling at repetitive elements during differentiation of hESC to MSCs. Both, *in-vitro* differentiated MSCs and primary BM-MSC retained chromatin accessibility at repetitive elements. However, HUVECs and other differentiated cell types lacked FAIRE signal suggesting that these regions are further remodeled during terminal differentiation. To answer this hypothesis, we have begun to assay chromatin accessibility in primary adipose tissue to identify the status of repetitive elements of *in-vivo* terminally differentiated cells. Through a special collaboration with a surgeon at UNC, we have received excess adipose tissue from surgical discards. I then dissociate the tissue using collagenase, isolate and purify the adipocytes through centrifugation. Due to the low density of lipids, adipocytes float as a layer of cells after centrifugation. We can then isolate a pure population of adipocytes and perform FAIRE-seq specifically looking at chromatin accessibility of repetitive elements. Currently, we are attempting to optimize the parameters of FAIRE for this cell type.

Both hESCs and BM-MSCs demonstrated chromatin accessibility at repetitive elements suggesting a signature for multipotency and stemness. Recently, Yamanaka et al. were able to “reprogram” terminally differentiated cells by exogenously expressing a set of four transcription factors [212][213]. These cells reverted to a more stem-like state having the ability to differentiate into cells from the three germ-layers, as well as form chimeric mice, leading to the term induced pluripotent stem cells (iPSCs) [212]. Therefore, we hypothesize that iPSCs undergo chromatin remodeling resulting in accessibility at repetitive elements. Assaying iPSC chromatin, using FAIRE-seq, at various stages during reprogramming would give insight into repetitive element chromatin remodeling. Couple these experiments with RNA-seq, and we could identify

the specific chromatin remodelers that are activated at the time point when repetitive elements become accessible, providing biological insight into how and when these regions are remodeled. Further, inhibition of the identified remodeler(s) using an inducible CRISPR knockout, would allow us to determine if chromatin remodeling at these regions is necessary for complete reprogramming. Taken together, iPSCs are an excellent model to understand repetitive element remodeling from the perspective of reprogramming into a stem cell as would add further evidence for repetitive element accessibility as a signature of stemness.

Consistent with other studies, our results suggest that the chromatin environment at repetitive elements is influencing gene expression [214]. We hypothesize that the repetitive element itself is crucial for this activity. To test this, we would delete specific repetitive elements using site specific CRISPR and assay RNA and protein with RT-qPCR and immunoblotting respectively, in knockouts and wild-type. Additionally, we could prevent the acetylation of these elements by targeting the specific HATs using RNAi or CRISPR techniques. Both of these approaches would help elucidate the role of repetitive elements in gene expression. Ultimately, the term “junk” DNA is rapidly being debunked and repetitive elements likely have functional significance within the human genome. In fact, evidence from our lab and others have previously demonstrated the repetitive elements can serve as transcription factor binding sites as well as act as enhancer elements, suggesting an importance in other cell types including cancer [90][215][216].

Though the origin of Ewing sarcoma continues to be debated, our results suggest a stem-like precursor. Strikingly, both BM-MSCs and Ewing Sarcoma demonstrated a marked similarity in enrichment of chromatin accessibility at a subset of simple repeats important for EWSR1-FLI1 targeting. These results are one of the first to demonstrate similarities in the chromatin environment between MSCs and Ewing sarcoma and add to the growing evidence suggesting MSCs as the cell of origin for this

disease. However, our results have not definitively ruled out a neuronal precursor. In fact, we hypothesize that a neuronal precursor would also have increased chromatin accessibility at repetitive elements leading to oncoprotein retargeting. It would be beneficial for future experiments to assay chromatin accessibility of repetitive elements in these cell types. In addition, using ChIP-seq to assay binding sites of EWSR1-FLI1 in neuronal cells, may demonstrate a significant overlap in binding sites with BM-MSCs. It may be that both sides of the debate are correct, and Ewing sarcoma arises after a translocation in either neuronal or mesenchymal stem cells due to the permissive chromatin environment.

Though we noted an enrichment of FAIRE signal at all GGAA in BM-MSC, those regions that had increased signal in BM-MSCs tended to be the the ones that were targeted by EWSR1-FLI1 in the cancer (Figure 2.14). The repeats targeted by EWSR1-FLI1 also demonstrated a mild enrichment in acetylated histones relative to non-targets. This suggests that a subset of these regions in BM-MSCs are primed for exploitation of the fusion protein. Further studies will need to be conducted in order to understand why these specific regions are targeted. Repetitive elements have been demonstrated to form alternative DNA structures. One hypothesis is that these regions share a common structure that distinguishes them from non-targeted regions. Previous studies have determined that 10-14 (GGAA) motifs are required for optimal EWSR1-FLI1 binding [90]. Perhaps these regions form an alternative DNA structure that is recognized by EWSR1-FLI1, and those regions having more or less motifs do not properly form the correct structure. Using algorithms designed to identify alternate structures, we could look for differences between EWSR1-FLI1 bound and unbound GGAA repeats.

In addition to Ewings, many other sarcomas, such as liposarcomas, alveolar sarcoma, and chondrosarcoma, are characterized by the presence of translocations creating novel transcription factors [66]. A fundamental question remaining is if these

translocations are similarly retargeted away from their canonical motif. Using bioinformatics, we can scan the underlying DNA sequence of repetitive elements and other FAIRE-positive regions of stem cells, looking for both the canonical and motifs mimicking that of the parental transcription factor. The bioinformatics predictions would then be validated using lentiviral transduction of fusion proteins in MSCs and performing ChIP-seq. Ultimately, the results from these experiments would demonstrate that the chromatin environment of stem cells is permissive for a range of fusion protein retargeting, suggesting that the chromatin environment of stem cells is associated with sarcoma development.

Interestingly, at sites where EWSR1-FLI1 is bound, cancer cells demonstrated an increase in FAIRE signal relative to BM-MSCs. This suggests that binding of EWSR1-FLI1 is increasing chromatin accessibility at these regions. However, neither EWSR1 or FLI1 have been associated with chromatin modifying activity, leading us to hypothesize that EWSR1-FLI1 is cooperating with other proteins to remodel chromatin. The simplest technique to assay partners of EWSR1-FLI1 would be to perform an immunoprecipitation and use mass spectrometry to identify potential protein interactions. Though theoretically simple, we and others have failed to co-purify EWRS1-FLI1, suggesting that the protein is inherently unstable. An alternative approach is to fuse EWSR1-FLI1 with a biotin ligase that biotynilates proximal and interacting proteins [217]. Using this method, we could directly pull down proximal/interacting partners of EWSR1-FLI1 using streptavidin and submit these results for mass spectrometry, looking for potential chromatin modifiers. Currently, our lab is using a novel high-throughput screening approach to identify compounds that affect chromatin accessibility at EWSR1-FLI1 binding sites. Utilizing both approaches would be beneficial as they are complimentary. The biotin assay would identify interacting partners and the compound screen would identify those proteins that affected chromatin accessibility at EWSR1-FLI1 binding sites, narrowing the list of potential targets. The

proteins from the intersecting set of these two approaches could lead to novel targeted therapies for Ewing sarcoma.

Techniques using high-throughput sequencing are powerful for conducting both discovery and hypothesis driven research. The sheer amount of information from these experiments enables researchers to use the data for secondary projects. Though we originally intended to identify regulatory regions in Ewing sarcoma cells, background signal from FAIRE-seq typically covers the remainder of the genome, allowing a genome-wide sampling of DNA content. Unexpectedly, we identified an approximate 1 MB deletion encompassing the tumor suppressor *PTEN*. Until this study, the consequences of PTEN loss had not yet been studied in Ewing sarcoma. Of the 7 Ewing sarcoma cell lines test, only EWS502 contained a PTEN loss validated by both fluorescent in-situ hybridization and immunoblotting. However, we demonstrated that 25% (4/15) of patient tumors demonstrated loss of PTEN protein expression, a ratio similar to that of and other previously described Ewing sarcoma recurrent mutations *CDKN2A* (10%-30%)[198][199] and *TP53* (3%-14%) [200][199].

Though, *TP53* and *CDKN2A* have recently been implicated as prognostic biomarkers, large scale analysis of primary Ewing tumors revealed no significant difference in event free survival for patients with these mutations [201]. Therefore, reliable biomarkers remain to be identified. Loss of PTEN resulted in a significant increase in activated AKT, leading to an enhancement of cellular transformative properties such as increased proliferation and colony formation, as well as decreased apoptosis, results of consistent with other AKT activation studies [202][203][204]. The enhancement of cellular transformative properties in PTEN null cells, suggests that patients lacking PTEN may have more aggressive disease. However, to date, no studies have associated PTEN status with clinical outcomes in Ewing sarcoma. The Children's Oncology Group, as well as many of our collaborators, have access to a significant number of primary Ewing sarcoma tumors and corresponding clinical data. Using

these resources, we could create a tissue microarray (TMA) of all available tumors. Using IHC and analysis similar to our study, we would identify PTEN null tumors and associate them with common clinical features such as time to relapse, disease free survival, and incidence of metastasis looking for statistically significant differences in PTEN null tumors. These studies would lend necessary evidence for PTEN as a biomarker in Ewing sarcoma and serve to help stratify patients into risk categories.

Despite the preclinical success of IGF-1R inhibitors, they fail a significant fraction of patients for unknown reasons [116][117][118]. Identifying these reasons, would allow clinicians to tailor therapies based on predicted response, saving patients valuable time, and more importantly, preventing patients from experiencing debilitating side effects from ineffective drugs. Our study demonstrated reciprocal responses to IGF-1R and mTOR inhibition in PTEN null cells. Cells lacking PTEN have increased levels of activated AKT even in the presence of potent IGF-1R inhibitors OSI-906 [137] and NVP-AEW541[205] (Figures 3.8 and 3.11). After exogenous expression of PTEN, cells became more responsive to inhibitor treatments, suggesting that IGF-1R-mediated inhibition of the AKT pathway requires functional PTEN. Our study also noted a significant increase in response to mTOR inhibition in PTEN null cells. mTOR a downstream target of AKT, is responsible for many cellular processes including autophagy [206]. Inhibition of mTOR, in cells with PTEN loss, was associated with an increase in autophagy and a concomitant decrease in cell viability consistent with studies in other cell lines [207] including cancer cells [150][151]. Taken together, our preclinical results suggest that patients with PTEN loss may not respond to IGF-R1 inhibitors. Alternatively, PTEN null patients may be more sensitive to therapies targeting PI3K, mTOR, and/or other downstream IGF-1R effectors. To test this idea, patient tumors, whose PTEN status has been identified, would be cultured *in-vitro* and subjected to various inhibitor treatments [208]. Similar to our study, therapies that reduce activated AKT and cell viability would be identified as potential treatment options. The

results of these experiments could be brought to the clinic and may predict the most effective therapies and treatment combinations for individual patient tumors.

Though, our experiments suggest that PTEN plays a crucial role in response to therapies, these studies were all done using *in-vitro* cell culture. In order to prove useful in a clinical setting it is essential that we demonstrate similar results *in-vivo*. Despite not having a faithful mouse model of Ewing sarcoma, our lab has previously demonstrated successful tumor formation using Ewing sarcoma cell lines in immunodeficient mice (unpublished data). To assay PTEN function with tumor progression and therapeutic response, CRISPR or lentiviral transduction of a doxycycline inducible PTEN would be engineered in EWS502 cells. The cells would then be injected into mice where tumor formation would occur. Our results suggest the expression of PTEN in this cell line is sufficient to prevent soft-agar formation, though a second study contradicts this finding [209]. This contradiction may be due to differing levels of PTEN protein. Niemeyer et al. used an antibiotic selection method, enriching for cells that could tolerate PTEN expression. It is possible that low levels of PTEN may lower activated AKT, yet still allow for anchorage-independent growth, while higher levels of PTEN cause apoptosis seen in our study. Controlled induction of PTEN at various time points in an *in-vivo* system would help resolve the controversy. We hypothesize that if expression of PTEN prevents tumor formation by increasing apoptosis, early induction should prevent tumor formation, while late induction should result in significant decrease of tumor size. Our results also suggest that PTEN expression enhances sensitivity to IGF-R1 inhibition while loss increases sensitivity to mTOR inhibition. Using the inducible system described above, we could allow tumor formation and induce PTEN expression while simultaneously treating with IGF-1R/mTOR inhibitors. We hypothesize that the group expressing PTEN and treated with inhibitor should have significantly decreased tumor size relative to both untreated and PTEN expression alone with the reciprocal results for mTOR inhibitors. Determin-

ing if PTEN expression modulates responsiveness to IGF-1R inhibtiors *in-vivo*, is an important step before moving into a clinical setting. Preclinical and early-phase clinical trials utilizing dual IGF-1R and mTOR inhibitors have begun to promising results [152][153][154][155]. However our study suggests that these therapeutic strategies may improve patient outcomes if stratified by PTEN status.

The preceding research has made scientific advances considered both translational and basic. Our results have suggested a potential mechanism for IGF-1R failure, as well as an alternative treatment in a subset of patients. Our study is one example demonstrating how loss of regulatory proteins can ultimately lead to attenuated drug sensitivity. Future studies should compare the genotypes of those patients who respond and compare them to those who the therapy fails. Our results suggest that informative insights into biology can be gained from those patients whose treatments fail them. Ultimately, the majority of the patients suffering from Ewing sarcoma are children, and the side effects from these therapies can be debilitating. Therefore, prescribing treatments with increased predicted response is vital for patient well-being. In addition, our study described a unique chromatin environment in stem cells that can be exploited by cancer. Taken together, our results suggest an intriguing hypothesis that stem cells can become cancerous. This impacts many facets of biology including regeneration and tissue repair, where stem cells are being manipulated and used as potential treatments. The recent discovery a stem cell specific chromatin remodeler lends further evidence for unique biological factors defining stemness [218]. If the chromatin environment is primed for oncogenic retargeting, it is not a far stretch to think that other biological factors may also be permissive for oncogenesis. Together, these studies have furthered our knowledge regarding therapy and the etiology of Ewing sarcoma. Though much work is still needed, we are on the cusp of identifying novel targeted therapeutics which have implications for patients who desperately need improved therapies.

BIBLIOGRAPHY

[1] A. Visel, M. J. Blow, Z. Li, T. Zhang, J. A. Akiyama, A. Holt, I. Plajzer-Frick, M. Shoukry, C. Wright, F. Chen, V. Afzal, B. Ren, E. M. Rubin, and L. A. Pennacchio, "ChIP-seq accurately predicts tissue-specific activity of enhancers.," *Nature*, vol. 457, pp. 854–858, Feb. 2009. 1

[2] P. R. Burton, D. G. Clayton, L. R. Cardon, N. Craddock, P. Deloukas, A. Duncanson, D. P. Kwiatkowski, M. I. McCarthy, W. H. Ouwehand, N. J. Samani, J. A. Todd, P. Donnelly, J. C. Barrett, D. Davison, D. Easton, D. Evans, H.-T. Leung, J. L. Marchini, A. P. Morris, C. C. A. Spencer, M. D. Tobin, A. P. Attwood, J. P. Boorman, B. Cant, U. Everson, J. M. Hussey, J. D. Jolley, A. S. Knight, K. Koch, E. Meech, S. Nutland, C. V. Prowse, H. E. Stevens, N. C. Taylor, G. R. Walters, N. M. Walker, N. A. Watkins, T. Winzer, R. W. Jones, W. L. McArdle, S. M. Ring, D. P. Strachan, M. Pembrey, G. Breen, D. S. Clair, S. Caesar, K. Gordon-Smith, L. Jones, C. Fraser, E. K. Green, D. Grozeva, M. L. Hamshere, P. A. Holmans, I. R. Jones, G. Kirov, V. Moskvina, I. Nikolov, M. C. O'Donovan, M. J. Owen, D. A. Collier, A. Elkin, A. Farmer, R. Williamson, P. McGuffin, A. H. Young, I. N. Ferrier, S. G. Ball, A. J. Balmforth, J. H. Barrett, D. T. Bishop, M. M. Iles, A. Maqbool, N. Yuldasheva, A. S. Hall, P. S. Braund, R. J. Dixon, M. Mangino, S. Stevens, J. R. Thompson, F. Bredin, M. Tremelling, M. Parkes, H. Drummond, C. W. Lees, E. R. Nimmo, J. Satsangi, S. A. Fisher, A. Forbes, C. M. Lewis, C. M. Onnie, N. J. Prescott, J. Sanderson, C. G. Mathew, J. Barbour, M. K. Mohiuddin, C. E. Todhunter, J. C. Mansfield, T. Ahmad, F. R. Cummings, D. P. Jewell, J. Webster, M. J. Brown, G. M. Lathrop, J. Connell, A. Dominiczak, C. A. B. Marcano, B. Burke, R. Dobson, J. Gungadoo, K. L. Lee, P. B. Munroe, S. J. Newhouse, A. Onipinla, C. Wallace, M. Xue, M. Caulfield, M. Farrall, A. Barton, T. B. i. R. Genetics, G. BRAGGS, I. N. Bruce, H. Donovan, S. Eyre, P. D. Gilbert, S. L. Hider, A. M. Hinks, S. L. John, C. Potter, A. J. Silman, D. P. M. Symmons, W. Thomson, J. Worthington, D. B. Dunger, B. Widmer, T. M. Frayling, R. M. Freathy, H. Lango, J. R. B. Perry, B. M. Shields, M. N. Weedon, A. T. Hattersley, G. A. Hitman, M. Walker, K. S. Elliott, C. J. Groves, C. M. Lindgren, N. W. Rayner, N. J. Timpson, E. Zeggini, M. Newport, G. Sirugo, E. Lyons, F. Vannberg, A. V. S. Hill, L. A. Bradbury, C. Farrar, J. J. Pointon, P. Wordsworth, M. A. Brown, J. A. Franklyn, J. M. Heward, M. J. Simmonds, S. C. L. Gough, S. Seal, B. C. S. C. UK, M. R. Stratton, N. Rahman, M. Ban, A. Goris, S. J. Sawcer, A. Compston, D. Conway, M. Jallow, K. A. Rockett, S. J. Bumpstead, A. Chaney, K. Downes, M. J. R. Ghori, R. Gwilliam, S. E. Hunt, M. Inouye, A. Keniry, E. King, R. McGinnis, S. Potter, R. Ravindrarajah, P. Whittaker, C. Widden, D. Withers, N. J. Cardin, T. Ferreira, J. Pereira-Gale, I. B. Hallgrimsdóttir, B. N. Howie, Z. Su, Y. Y. Teo, D. Vukcevic, D. Bentley, and A. Compston, "Genome-wide association study of 14,000 cases of seven com-

mon diseases and 3,000 shared controls," *Nature*, vol. 447, pp. 661–678, June 2007. 1

[3] T. A. Manolio and F. S. Collins, "The HapMap and genome-wide association studies in diagnosis and therapy.," *Annu. Rev. Med.*, vol. 60, pp. 443–456, 2009. 1

[4] A. Barski, S. Cuddapah, K. Cui, T.-Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev, and K. Zhao, "High-resolution profiling of histone methylations in the human genome.," *Cell*, vol. 129, pp. 823–837, May 2007. 1

[5] T. S. Mikkelsen, M. Ku, D. B. Jaffe, B. Issac, E. Lieberman, G. Giannoukos, P. Alvarez, W. Brockman, T.-K. Kim, R. P. Koche, W. Lee, E. Mendenhall, A. O'Donovan, A. Presser, C. Russ, X. Xie, A. Meissner, M. Wernig, R. Jaenisch, C. Nusbaum, E. S. Lander, and B. E. Bernstein, "Genome-wide maps of chromatin state in pluripotent and lineage-committed cells," *Nature*, vol. 448, pp. 553–560, July 2007. 1, 6, 11

[6] G. Robertson, M. Hirst, M. Bainbridge, M. Bilenky, Y. Zhao, T. Zeng, G. Euskirchen, B. Bernier, R. Varhol, A. Delaney, N. Thiessen, O. L. Griffith, A. He, M. Marra, M. Snyder, and S. Jones, "Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing.," *Nature Methods*, vol. 4, pp. 651–657, Aug. 2007. 1

[7] A. P. Boyle, S. Davis, H. P. Shulha, P. Meltzer, E. H. Margulies, Z. Weng, T. S. Furey, and G. E. Crawford, "High-resolution mapping and characterization of open chromatin across the genome," *Cell*, vol. 132, pp. 311–322, Jan. 2008. 1, 9, 19

[8] J. R. Hesselberth, X. Chen, Z. Zhang, P. J. Sabo, R. Sandstrom, A. P. Reynolds, R. E. Thurman, S. Neph, M. S. Kuehn, W. S. Noble, S. Fields, and J. A. Stamatoyannopoulos, "Global mapping of protein-DNA interactions in vivo by digital genomic footprinting.," *Nature Methods*, vol. 6, pp. 283–289, Apr. 2009. 1

[9] D. E. Schones, K. Cui, S. Cuddapah, T.-Y. Roh, A. Barski, Z. Wang, G. Wei, and K. Zhao, "Dynamic regulation of nucleosome positioning in the human genome.," *Cell*, vol. 132, pp. 887–898, Mar. 2008. 1

[10] R. Lister, R. C. O'Malley, J. Tonti-Filippini, B. D. Gregory, C. C. Berry, A. H. Millar, and J. R. Ecker, "Highly integrated single-base resolution maps of the epigenome in Arabidopsis.," *Cell*, vol. 133, pp. 523–536, May 2008. 1

[11] A. Mortazavi, B. A. Williams, K. McCue, L. Schaeffer, and B. Wold, "Mapping and quantifying mammalian transcriptomes by RNA-Seq.," *Nature Methods*, vol. 5, pp. 621–628, July 2008. 1

[12] U. Nagalakshmi, Z. Wang, K. Waern, C. Shou, D. Raha, M. Gerstein, and M. Snyder, "The transcriptional landscape of the yeast genome defined by RNA sequencing.," *Science (New York, NY)*, vol. 320, pp. 1344–1349, June 2008. 1

[13] ENCODE Project Consortium, “A user’s guide to the encyclopedia of DNA elements (ENCODE).,” *PLoS biology*, vol. 9, p. e1001046, Apr. 2011. 1

[14] A. Kundaje, W. Meuleman, J. Ernst, M. Bilenky, A. Yen, A. Heravi-Moussavi, P. Kheradpour, Z. Zhang, J. Wang, M. J. Ziller, V. Amin, J. W. Whitaker, M. D. Schultz, L. D. Ward, A. Sarkar, G. Quon, R. S. Sandstrom, M. L. Eaton, Y. C. Wu, A. R. Pfenning, X. Wang, M. Claussnitzer, Y. Liu, C. Coarfa, R. A. Harris, N. Shoresh, C. B. Epstein, E. Gjoneska, D. Leung, W. Xie, R. D. Hawkins, R. Lister, C. Hong, P. Gascard, A. J. Mungall, R. Moore, E. Chuah, A. Tam, T. K. Canfield, R. S. Hansen, R. Kaul, P. J. Sabo, M. S. Bansal, A. Carles, J. R. Dixon, K. H. Farh, S. Feizi, R. Karlic, A. R. Kim, A. Kulkarni, D. Li, R. Lowdon, G. Elliott, T. R. Mercer, S. J. Neph, V. Onuchic, P. Polak, N. Rajagopal, P. Ray, R. C. Sallari, K. T. Siebenthall, N. A. Sinnott-Armstrong, M. Stevens, R. E. Thurman, J. Wu, B. Zhang, X. Zhou, A. E. Beaudet, L. A. Boyer, P. L. De Jager, P. J. Farnham, S. J. Fisher, D. Haussler, S. J. Jones, W. Li, M. A. Marra, M. T. McManus, S. Sunyaev, J. A. Thomson, T. D. Tlsty, L. H. Tsai, W. Wang, R. A. Waterland, M. Q. Zhang, L. H. Chadwick, B. E. Bernstein, J. F. Costello, J. R. Ecker, M. Hirst, A. Meissner, A. Milosavljevic, B. Ren, J. A. Stamatoyannopoulos, T. Wang, and M. Kellis, “Integrative analysis of 111 reference human epigenomes,” *Nature*, vol. 518, pp. 317–330, Feb. 2015. 1

[15] M. J. Evans and M. H. Kaufman, “Establishment in culture of pluripotential cells from mouse embryos.,” *Nature*, vol. 292, pp. 154–156, July 1981. 2

[16] G. R. Martin, “Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 78, pp. 7634–7638, Dec. 1981. 2

[17] J. A. Thomson, “Embryonic Stem Cell Lines Derived from Human Blastocysts,” *Science (New York, NY)*, vol. 282, pp. 1145–1147, Nov. 1998. 2, 11

[18] L. Gerrard, L. Rodgers, and W. Cui, “Differentiation of human embryonic stem cells to neural lineages in adherent culture by blocking bone morphogenetic protein signaling.,” *Stem cells (Dayton, Ohio)*, vol. 23, pp. 1234–1241, Oct. 2005. 2

[19] V. A. Maltsev, J. Rohwedel, J. Hescheler, and A. M. Wobus, “Embryonic stem cells differentiate in vitro into cardiomyocytes representing sinusnodal, atrial and ventricular cell types.,” *Mechanisms of development*, vol. 44, pp. 41–50, Nov. 1993. 2

[20] S. J. Kattman, A. D. Witty, M. Gagliardi, N. C. Dubois, M. Niapour, A. Hotta, J. Ellis, and G. Keller, “Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines.,” *Cell stem cell*, vol. 8, pp. 228–240, Feb. 2011. 2

[21] J. Rohwedel, V. Maltsev, E. Bober, H. H. Arnold, J. Hescheler, and A. M. Wobus, "Muscle cell differentiation of embryonic stem cells reflects myogenesis in vivo: developmentally regulated expression of myogenic determination genes and functional expression of ionic currents.," *Developmental biology*, vol. 164, pp. 87–101, July 1994. 2

[22] C. E. Murry and G. Keller, "Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development.," *Cell*, vol. 132, pp. 661–680, Feb. 2008. 2

[23] E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, R. Funke, D. Gage, K. Harris, A. Heaford, J. Howland, L. Kann, J. Lehoczky, R. LeVine, P. McEwan, K. McKernan, J. Meldrim, J. P. Mesirov, C. Miranda, W. Morris, J. Naylor, C. Raymond, M. Rosetti, R. Santos, A. Sheridan, C. Sougnez, N. Stange-Thomann, N. Stojanovic, A. Subramanian, D. Wyman, J. Rogers, J. Sulston, R. Ainscough, S. Beck, D. Bentley, J. Burton, C. Clee, N. Carter, A. Coulson, R. Deadman, P. Deloukas, A. Dunham, I. Dunham, R. Durbin, L. French, D. Grafham, S. Gregory, T. Hubbard, S. Humphray, A. Hunt, M. Jones, C. Lloyd, A. McMurray, L. Matthews, S. Mercer, S. Milne, J. C. Mullikin, A. Mungall, R. Plumb, M. Ross, R. Shownkeen, S. Sims, R. H. Waterston, R. K. Wilson, L. W. Hillier, J. D. McPherson, M. A. Marra, E. R. Mardis, L. A. Fulton, A. T. Chinwalla, K. H. Pepin, W. R. Gish, S. L. Chissoe, M. C. Wendl, K. D. Delehaunty, T. L. Miner, A. Delehaunty, J. B. Kramer, L. L. Cook, R. S. Fulton, D. L. Johnson, P. J. Minx, S. W. Clifton, T. Hawkins, E. Branscomb, P. Predki, P. Richardson, S. Wenning, T. Slezak, N. Doggett, J. F. Cheng, A. Olsen, S. Lucas, C. Elkin, E. Uberbacher, M. Frazier, R. A. Gibbs, D. M. Muzny, S. E. Scherer, J. B. Bouck, E. J. Sodergren, K. C. Worley, C. M. Rives, J. H. Gorrell, M. L. Metzker, S. L. Naylor, R. S. Kucherlapati, D. L. Nelson, G. M. Weinstock, Y. Sakaki, A. Fujiyama, M. Hattori, T. Yada, A. Toyoda, T. Itoh, C. Kawagoe, H. Watanabe, Y. Totoki, T. Taylor, J. Weissenbach, R. Heilig, W. Saurin, F. Artiguenave, P. Brottier, T. Bruls, E. Pelletier, C. Robert, P. Wincker, D. R. Smith, L. Doucette-Stamm, M. Rubenfield, K. Weinstock, H. M. Lee, J. Dubois, A. Rosenthal, M. Platzer, G. Nyakatura, S. Taudien, A. Rump, H. Yang, J. Yu, J. Wang, G. Huang, J. Gu, L. Hood, L. Rowen, A. Madan, S. Qin, R. W. Davis, N. A. Federspiel, A. P. Abola, M. J. Proctor, R. M. Myers, J. Schmutz, M. Dickson, J. Grimwood, D. R. Cox, M. V. Olson, R. Kaul, C. Raymond, N. Shimizu, K. Kawasaki, S. Minoshima, G. A. Evans, M. Athanasiou, R. Schultz, B. A. Roe, F. Chen, H. Pan, J. Ramser, H. Lehrach, R. Reinhardt, W. R. McCombie, M. de la Bastide, N. Dedhia, H. Blöcker, K. Hornischer, G. Nordsiek, R. Agarwala, L. Aravind, J. A. Bailey, A. Bateman, S. Batzoglou, E. Birney, P. Bork, D. G. Brown, C. B. Burge, L. Cerutti, H. C. Chen, D. Church, M. Clamp, R. R. Copley, T. Doerks, S. R. Eddy, E. E. Eichler, T. S. Furey, J. Galagan, J. G. Gilbert, C. Harmon, Y. Hayashizaki, D. Haussler, H. Hermjakob, K. Hokamp, W. Jang, L. S. Johnson, T. A. Jones, S. Kasif, A. Kaspryzk, S. Kennedy, W. J. Kent, P. Kitts, E. V.

Koonin, I. Korf, D. Kulp, D. Lancet, T. M. Lowe, A. McLysaght, T. Mikkelsen, J. V. Moran, N. Mulder, V. J. Pollara, C. P. Ponting, G. Schuler, J. Schultz, G. Slater, A. F. Smit, E. Stupka, J. Szustakowski, D. Thierry-Mieg, J. Thierry-Mieg, L. Wagner, J. Wallis, R. Wheeler, A. Williams, Y. I. Wolf, K. H. Wolfe, S. P. Yang, R. F. Yeh, F. Collins, M. S. Guyer, J. Peterson, A. Felsenfeld, K. A. Wetterstrand, A. Patrinos, M. J. Morgan, P. de Jong, J. J. Catanese, K. Osoegawa, H. Shizuya, S. Choi, Y. J. Chen, J. Szustakowki, and International Human Genome Sequencing Consortium, "Initial sequencing and analysis of the human genome.," *Nature*, vol. 409, pp. 860–921, Feb. 2001. 2

[24] G. Bourque, B. Leong, V. B. Vega, X. Chen, Y. L. Lee, K. G. Srinivasan, J. L. Chew, Y. Ruan, C. L. Wei, H. H. Ng, and E. T. Liu, "Evolution of the mammalian transcription factor binding repertoire via transposable elements," *Genome Research*, vol. 18, pp. 1752–1762, Nov. 2008. 2

[25] U. J, "Alu elements: at the crossroads between disease and evolution," *Biochemical Society Transactions*, vol. 41, pp. 1532–1535, Dec. 2013. 2

[26] M. Kamal, X. Xie, and E. S. Lander, "A large family of ancient repeat elements in the human genome is under strong selection.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, pp. 2740–2745, Feb. 2006. 3

[27] C. B. Lowe, G. Bejerano, and D. Haussler, "Thousands of human mobile element fragments undergo strong purifying selection near developmental genes.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, pp. 8005–8010, May 2007. 3

[28] D. A. Kleinjan and V. van Heyningen, "Long-range control of gene expression: emerging mechanisms and disruption in disease.," *Am J Hum Genet*, vol. 76, pp. 8–32, Jan. 2005. 3

[29] A. P. J. de Koning, W. Gu, T. A. Castoe, M. A. Batzer, and D. D. Pollock, "Repetitive Elements May Comprise Over Two-Thirds of the Human Genome," *PLoS genetics*, vol. 7, p. e1002384, Dec. 2011. 3

[30] T. J. Treangen and S. L. Salzberg, "Repetitive DNA and next-generation sequencing: computational challenges and solutions," *Nature reviews. Genetics*, vol. 13, pp. 36–46, Nov. 2011. 3

[31] V. P. Belancio, A. M. Roy-Engel, R. R. Pochampally, and P. Deininger, "Somatic expression of LINE-1 elements in human tissues," *nar.oxfordjournals.org*. 3

[32] H. Kano, I. Godoy, C. Courtney, M. R. Vetter, G. L. Gerton, E. M. Ostertag, and H. H. Kazazian, "L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism.," *Genes & development*, vol. 23, pp. 1303–1312, June 2009. 3

[33] G. J. Faulkner, Y. Kimura, C. O. Daub, S. Wani, C. Plessy, K. M. Irvine, K. Schroder, N. Cloonan, A. L. Steptoe, T. Lassmann, K. Waki, N. Hornig, T. Arakawa, H. Takahashi, J. Kawai, A. R. R. Forrest, H. Suzuki, Y. Hayashizaki, D. A. Hume, V. Orlando, S. M. Grimmond, and P. Carninci, "The regulated retrotransposon transcriptome of mammalian cells.," *Nature Genetics*, vol. 41, pp. 563–571, May 2009. 4

[34] L. Crepaldi, C. Policarpi, A. Coatti, W. T. Sherlock, B. C. Jongbloets, T. A. Down, and A. Riccio, "Binding of TFIIIC to sine elements controls the relocation of activity-dependent neuronal genes to transcription factories.," *PLoS genetics*, vol. 9, no. 8, p. e1003699, 2013. 4

[35] A. M. Wobus, H. Holzhausen, P. Jäkel, and J. Schöneich, "Characterization of a pluripotent stem cell line derived from a mouse embryo.," *Experimental cell research*, vol. 152, pp. 212–219, May 1984. 4

[36] K. Ahmed, H. Dehghani, P. Rugg-Gunn, E. Fussner, J. Rossant, and D. P. Bazett-Jones, "Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo.," *PloS one*, vol. 5, no. 5, p. e10531, 2010. 4, 6

[37] G. Felsenfeld and M. Groudine, "Controlling the double helix.," *Nature*, vol. 421, pp. 448–453, Jan. 2003. 4

[38] R. D. Kornberg, "Chromatin structure: a repeating unit of histones and DNA.," *Science (New York, NY)*, vol. 184, pp. 868–871, May 1974. 4

[39] K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond, "Crystal structure of the nucleosome core particle at 2.8 A resolution.," *Nature*, vol. 389, pp. 251–260, Sept. 1997. 4

[40] B. R. Cairns, Y. J. Kim, M. H. Sayre, B. C. Laurent, and R. D. Kornberg, "A multisubunit complex containing the SWI1/ADR6, SWI2/SNF2, SWI3, SNF5, and SNF6 gene products isolated from yeast.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 91, pp. 1950–1954, Mar. 1994. 5

[41] B. D. Strahl and C. D. Allis, "The language of covalent histone modifications," *Nature*, vol. 403, pp. 41–45, Jan. 2000. 5

[42] K. Ahmad and S. Henikoff, "Centromeres are specialized replication domains in heterochromatin.," *The Journal of cell biology*, vol. 153, pp. 101–110, Apr. 2001. 5

[43] C. Jin, C. Zang, G. Wei, K. Cui, W. Peng, K. Zhao, and G. Felsenfeld, "H3.3/H2A.Z double variant—[ndash]—containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions," *Nature Genetics*, vol. 41, pp. 941–945, Aug. 2009. 5

[44] M. Ptashne, "Epigenetics: core misconcept," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, pp. 7101–7103, Apr. 2013. 5

[45] S. Bultman, T. Gebuhr, D. Yee, C. La Mantia, J. Nicholson, A. Gilliam, F. Randazzo, D. Metzger, P. Chambon, G. Crabtree, and T. Magnuson, "A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes.," *Molecular cell*, vol. 6, pp. 1287–1295, Dec. 2000. 5

[46] D. O'Carroll, S. Erhardt, M. Pagani, S. C. Barton, M. A. Surani, and T. Jenuwein, "The polycomb-group gene Ezh2 is required for early mouse development.," *Molecular and cellular biology*, vol. 21, pp. 4330–4336, July 2001. 5

[47] E. Li, T. H. Bestor, and R. Jaenisch, "Targeted mutation of the DNA methyltransferase gene results in embryonic lethality.," *Cell*, vol. 69, pp. 915–926, June 1992. 5

[48] M. Shogren-Knaak, H. Ishii, J.-M. Sun, M. J. Pazin, J. R. Davie, and C. L. Peterson, "Histone H4-K16 acetylation controls chromatin structure and protein interactions.," *Science (New York, NY)*, vol. 311, pp. 844–847, Feb. 2006. 5, 35

[49] T. G. Fazzio, J. T. Huff, and B. Panning, "An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity.," *Cell*, vol. 134, pp. 162–174, July 2008. 6

[50] X. Li, L. Li, R. Pandey, J. S. Byun, K. Gardner, Z. Qin, and Y. Dou, "The histone acetyltransferase MOF is a key regulator of the embryonic stem cell core transcriptional network.," *Cell stem cell*, vol. 11, pp. 163–178, Aug. 2012. 6

[51] X. Zhong and Y. Jin, "Critical roles of coactivator p300 in mouse embryonic stem cell differentiation and Nanog expression.," *The Journal of biological chemistry*, vol. 284, pp. 9168–9175, Apr. 2009. 6

[52] M. A. Ricci, C. Manzo, M. F. García-Parajo, M. Lakadamyali, and M. P. Cosma, "Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo.," *Cell*, vol. 160, pp. 1145–1158, Mar. 2015. 6, 35

[53] B. Wen, H. Wu, Y. Shinkai, R. A. Irizarry, and A. P. Feinberg, "Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells.," *Nature Genetics*, vol. 41, pp. 246–250, Jan. 2009. 6, 11

[54] S. Efroni, R. Duttagupta, J. Cheng, H. Dehghani, D. J. Hoeppner, C. Dash, D. P. Bazett-Jones, S. Le Grice, R. D. G. McKay, K. H. Buetow, T. R. Gingeras, T. Misteli, and E. Meshorer, "Global transcription in pluripotent embryonic stem cells.," *Cell stem cell*, vol. 2, pp. 437–447, May 2008. 6, 28

[55] A. Adamo, B. Sesé, S. Boue, J. Castaño, I. Paramonov, M. J. Barrero, and J. C. Izpisua Belmonte, "LSD1 regulates the balance between self-renewal and

differentiation in human embryonic stem cells.," *Nature cell biology*, vol. 13, pp. 652–659, June 2011. 6

[56] B. E. Bernstein, T. S. Mikkelsen, X. Xie, M. Kamal, D. J. Huebert, J. Cuff, B. Fry, A. Meissner, M. Wernig, K. Plath, R. Jaenisch, A. Wagschal, R. Feil, S. L. Schreiber, and E. S. Lander, "A bivalent chromatin structure marks key developmental genes in embryonic stem cells.," *Cell*, vol. 125, pp. 315–326, Apr. 2006. 6, 11

[57] M. P. Creyghton, A. W. Cheng, G. G. Welstead, T. Kooistra, B. W. Carey, E. J. Steine, J. Hanna, M. A. Lodato, G. M. Frampton, P. A. Sharp, L. A. Boyer, R. A. Young, and R. Jaenisch, "Histone H3K27ac separates active from poised enhancers and predicts developmental state," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, pp. 21931–21936, Dec. 2010. 6

[58] G. E. Zentner, P. J. Tesar, and P. C. Scacheri, "Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions.," *Genome Research*, vol. 21, pp. 1273–1283, Aug. 2011. 6

[59] C. A. Gifford, M. J. Ziller, H. Gu, C. Trapnell, J. Donaghey, A. Tsankov, A. K. Shalek, D. R. Kelley, A. A. Shishkin, R. Issner, X. Zhang, M. Coyne, J. L. Fostel, L. Holmes, J. Meldrim, M. Guttman, C. Epstein, H. Park, O. Kohlbacher, J. Rinn, A. Gnirke, E. S. Lander, B. E. Bernstein, and A. Meissner, "Transcriptional and Epigenetic Dynamics during Specification of Human Embryonic Stem Cells," *Cell*, May 2013. 6

[60] R. D. Hawkins, G. C. Hon, L. K. Lee, Q. Ngo, R. Lister, M. Pelizzola, L. E. Edsall, S. Kuan, Y. Luu, S. Klugman, J. Antosiewicz-Bourget, Z. Ye, C. Espinoza, S. Agarwahl, L. Shen, V. Ruotti, W. Wang, R. Stewart, J. A. Thomson, J. R. Ecker, and B. Ren, "Distinct epigenomic landscapes of pluripotent and lineage-committed human cells.," *Cell stem cell*, vol. 6, pp. 479–491, May 2010. 6, 37

[61] A. J. Friedenstein, I. I. Piatetzky-Shapiro, and K. V. Petrakova, "Osteogenesis in transplants of bone marrow cells," 6

[62] M. F. Pittenger, A. M. Mackay, S. C. Beck, R. K. Jaiswal, R. Douglas, J. D. Mosca, M. A. Moorman, D. W. Simonetti, S. Craig, and D. R. Marshak, "Multilineage potential of adult human mesenchymal stem cells," *Science (New York, NY)*, vol. 284, pp. 143–147, Apr. 1999. 7, 24

[63] A. Uccelli, L. Moretta, and V. Pistoia, "Mesenchymal stem cells in health and disease.," *Nature reviews. Immunology*, vol. 8, pp. 726–736, Sept. 2008. 7

[64] M. Dominici, K. Le Blanc, I. Mueller, I. Slaper-Cortenbach, F. Marini, D. Krause, R. Deans, A. Keating, D. Prockop, and E. Horwitz, "Minimal criteria for defining

multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement.," *Cytotherapy*, vol. 8, no. 4, pp. 315–317, 2006. 7

[65] V. Y. Jo and C. D. M. Fletcher, "WHO classification of soft tissue tumours: an update based on the 2013 (4th) edition.," *Pathology*, vol. 46, pp. 95–104, Feb. 2014. 7

[66] L. J. Helman and P. Meltzer, "Mechanisms of sarcoma development.," *Nature Reviews Cancer*, vol. 3, pp. 685–694, Sept. 2003. 7, 75

[67] R. Rodriguez, R. Rubio, and P. Menendez, "Modeling sarcomagenesis using multipotent mesenchymal stem cells," *Cell research*, vol. 22, pp. 62–77, Jan. 2012. 7, 31

[68] F. Tirode, K. Laud-Duval, A. Prieur, B. Delorme, P. Charbord, and O. Delattre, "Mesenchymal stem cell features of Ewing tumors," *Cancer cell*, vol. 11, pp. 421–429, May 2007. 7, 8, 37

[69] A. M. Cleton-Jansen, J. K. Anninga, I. H. Briaire-de Bruijn, S. Romeo, J. Oosting, R. M. Egeler, H. Gelderblom, A. H. M. Taminiau, and P. C. W. Hogendoorn, "Profiling of high-grade central osteosarcoma and its putative progenitor cells identifies tumourigenic pathways.," *British Journal of Cancer*, vol. 101, pp. 1909–1918, Dec. 2009. 7

[70] I. Matushansky, E. Hernando, N. D. Socci, T. Matos, J. Mills, M. A. Edgar, G. K. Schwartz, S. Singer, C. Cordon-Cardo, and R. G. Maki, "A developmental model of sarcomagenesis defines a differentiation-based classification for liposarcomas," *The American journal of pathology*, vol. 172, pp. 1069–1080, Apr. 2008. 7

[71] J. Tolar, A. J. Nauta, M. J. Osborn, A. Panoskaltsis Mortari, R. T. McElmurry, S. Bell, L. Xia, N. Zhou, M. Riddle, T. M. Schroeder, J. J. Westendorf, R. S. McIvor, P. C. W. Hogendoorn, K. Szuhai, L. Oseth, B. Hirsch, S. R. Yant, M. A. Kay, A. Peister, D. J. Prockop, W. E. Fibbe, and B. R. Blazar, "Sarcoma derived from cultured mesenchymal stem cells.," *Stem cells (Dayton, Ohio)*, vol. 25, pp. 371–379, Feb. 2007. 7

[72] A. B. Mohseny, K. Szuhai, S. Romeo, E. P. Buddingh, I. Briaire-de Bruijn, D. de Jong, M. van Pel, A.-M. Cleton-Jansen, and P. C. W. Hogendoorn, "Osteosarcoma originates from mesenchymal stem cells in consequence of aneuploidization and genomic loss of Cdkn2.," *The Journal of Pathology*, vol. 219, pp. 294–305, Nov. 2009. 7

[73] N. Serakinci, P. Guldberg, J. S. Burns, B. Abdallah, H. Schrødder, T. Jensen, and M. Kassem, "Adult human mesenchymal stem cell as a target for neoplastic transformation.," *Oncogene*, vol. 23, pp. 5095–5098, June 2004. 7

[74] Y. Shima, T. Okamoto, T. Aoyama, K. Yasura, T. Ishibe, K. Nishijo, K. R. Shibata, Y. Kohno, K. Fukiage, S. Otsuka, D. Uejima, T. Nakayama, T. Nakamura, T. Kiyono, and J. Toguchida, "In vitro transformation of mesenchymal stem cells by oncogenic H-rasVal12.," *Biochemical and biophysical research communications*, vol. 353, pp. 60–66, Feb. 2007. 7

[75] J. Ewing, "Diffuse endothelioma of bone," *Proceeding of New York Pathological Society*, vol. 21, p. 17, 1921. 7

[76] N. Esiashvili, M. Goodman, and R. B. Marcus, "Changes in Incidence and Survival of Ewing Sarcoma Patients Over the Past 3 Decades," *Journal of Pediatric Hematology/Oncology*, vol. 30, pp. 425–430, June 2008. 7, 8

[77] O. Delattre, J. Zucman, B. Plougastel, C. Desmaze, T. Melot, M. Peter, H. Kovar, I. Joubert, P. de Jong, G. Rouleau, and et al, "Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours," *Nature*, vol. 359, pp. 162–165, Sept. 1992. 8, 46

[78] M. L. G. S. L. L. L. B. L. B. C. L. O. D. J. Z. G. T. C. T. D. W A May, "Ewing sarcoma 11;22 translocation produces a chimeric transcription factor that requires the DNA-binding domain encoded by FLI1 for transformation.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 90, p. 5752, June 1993. 8, 9

[79] A. O. Cavazzana, J. S. Miser, J. Jefferson, and T. J. Triche, "Experimental evidence for a neural origin of Ewing's sarcoma of bone.," *The American journal of pathology*, vol. 127, pp. 507–518, June 1987. 8

[80] C.-H. Suh, N. G. Ordóñez, J. Hicks, and B. Mackay, "Ultrastructure of the Ewing's sarcoma family of tumors.," *Ultrastruct Pathol*, vol. 26, pp. 67–76, Mar. 2002. 8

[81] S. Navarro, M. González-Devesa, A. Ferrández-Izquierdo, T. J. Triche, and A. Llombart-Bosch, "Scanning electron microscopic evidence for neural differentiation in Ewing's sarcoma cell lines.," *Virchows Archiv*, vol. 416, no. 5, pp. 383–391, 1990. 8

[82] C. J. Rorie, V. D. Thomas, P. Chen, H. H. Pierce, J. P. O'Bryan, and B. E. Weissman, "The Ews/Fli-1 fusion gene switches the differentiation program of neuroblastomas to Ewing sarcoma/peripheral primitive neuroectodermal tumors," *Cancer Research*, vol. 64, pp. 1266–1277, Feb. 2004. 8

[83] M. A. Teitell, A. D. Thompson, P. H. Sorensen, H. Shimada, T. J. Triche, and C. T. Denny, "EWS/ETS fusion genes induce epithelial and neuroectodermal differentiation in NIH 3T3 fibroblasts," *Lab Invest*, vol. 79, pp. 1535–1543, Dec. 1999. 8

[84] S. Hu-Lieskovan, J. Zhang, L. Wu, H. Shimada, D. E. Schofield, and T. J. Triche, "EWS-FLI1 fusion protein up-regulates critical genes in neural crest development and is responsible for the observed phenotype of Ewing's family of tumors," *Cancer Research*, vol. 65, pp. 4633–4644, June 2005. 8

[85] M.-L. Suvà, N. Riggi, J.-C. Stehle, K. Baumer, S. Tercier, J.-M. Joseph, D. Suvà, V. Clément, P. Provero, L. Cironi, M.-C. Osterheld, L. Guillou, and I. Stamenkovic, "Identification of cancer stem cells in Ewing's sarcoma," *Cancer Research*, vol. 69, pp. 1776–1781, Mar. 2009. 9, 37

[86] B. Deneen and C. T. Denny, "Loss of p16 pathways stabilizes EWS/FLI1 expression and complements EWS/FLI1 mediated transformation.," *Oncogene*, vol. 20, pp. 6731–6741, Oct. 2001. 9

[87] S. L. Lessnick, C. S. Dacwag, and T. R. Golub, "The Ewing's sarcoma oncoprotein EWS/FLI induces a p53-dependent growth arrest in primary human fibroblasts.," *Cancer cell*, vol. 1, pp. 393–401, May 2002. 9

[88] S. Lessnick, B. S. Braun, C. T. Denny, and W. A. May, "Multiple domains mediate transformation by the Ewing's sarcoma EWS/FLI-1 fusion gene," *Oncogene*, vol. 10, pp. 423–431, 1995. 9

[89] K. Gangwal, S. Sankar, P. C. Hollenhorst, M. Kinsey, S. C. Haroldsen, A. A. Shah, K. M. Boucher, W. S. Watkins, L. B. Jorde, B. J. Graves, and S. L. Lessnick, "Microsatellites as EWS/FLI response elements in Ewing's sarcoma.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, pp. 10149–10154, July 2008. 9, 31, 38

[90] M. Patel, J. M. Simon, M. D. Iglesia, S. B. Wu, A. W. McFadden, J. D. Lieb, and I. J. Davis, "Tumor-specific retargeting of an oncogenic transcription factor chimera results in dysregulation of chromatin and transcription.," *Genome Research*, vol. 22, pp. 259–270, Feb. 2012. 9, 31, 33, 38, 46, 52, 74, 75

[91] L. Cironi, N. Riggi, P. Provero, N. Wolf, M.-L. Suva, D. Suva, V. Kindler, and I. Stamenkovic, "IGF1 is a common target gene of Ewing's sarcoma fusion proteins in mesenchymal progenitor cells.," *PloS one*, vol. 3, no. 7, p. e2634, 2008. 9

[92] M. Kinsey, R. Smith, and S. L. Lessnick, "NR0B1 is required for the oncogenic phenotype mediated by EWS/FLI in Ewing's sarcoma," *Mol Cancer Res*, vol. 4, pp. 851–859, Nov. 2006. 9, 38, 46

[93] W. Luo, K. Gangwal, S. Sankar, K. M. Boucher, D. Thomas, and S. L. Lessnick, "GSTM4 is a microsatellite-containing EWS/FLI target involved in Ewing's sarcoma oncogenesis and therapeutic resistance.," *Oncogene*, vol. 28, pp. 4126–4132, Nov. 2009. 9, 38

[94] R. Smith, L. A. Owen, D. J. Trem, J. S. Wong, J. S. Whangbo, T. R. Golub, and S. L. Lessnick, "Expression profiling of EWS/FLI identifies NKX2.2 as a critical target gene in Ewing's sarcoma," *Cancer cell*, vol. 9, pp. 405–416, May 2006. 9, 38, 46

[95] N. Riggi, B. Knoechel, S. M. Gillespie, E. Rheinbay, G. Boulay, M. L. Suvà, N. E. Rossetti, W. E. Boonseng, O. Oksuz, E. B. Cook, A. Formey, A. Patel, M. Gymrek, V. Thapar, V. Deshpande, D. T. Ting, F. J. Hornicek, G. P. Nielsen, I. Stamenkovic, M. J. Aryee, B. E. Bernstein, and M. N. Rivera, "EWS-FLI1 Utilizes Divergent Chromatin Remodeling Mechanisms to Directly Activate or Repress Enhancer Elements in Ewing Sarcoma," *Cancer cell*, vol. 26, pp. 668–681, Nov. 2014. 9, 33, 34

[96] G. E. Crawford, I. E. Holt, J. Whittle, B. D. Webb, D. Tai, S. Davis, E. H. Margulies, Y. Chen, J. A. Bernat, D. Ginsburg, D. Zhou, S. Luo, T. J. Vasicek, M. J. Daly, T. G. Wolfsberg, and F. S. Collins, "Genome-wide mapping of DNase hypersensitive sites using massively parallel signature sequencing (MPSS)," *Genome Research*, vol. 16, pp. 123–131, Jan. 2006. 9, 19, 42

[97] L. Song and G. E. Crawford, "DNase-seq: A High-Resolution Technique for Mapping Active Gene Regulatory Elements across the Genome from Mammalian Cells," *Cold Spring Harbor Protocols*, vol. 2010, pp. pdb.prot5384–pdb.prot5384, Feb. 2010. 9

[98] M. S. Wilken, J. A. Brzezinski, A. La Torre, K. Siebenthall, R. Thurman, P. Sabo, R. S. Sandstrom, J. Vierstra, T. K. Canfield, R. S. Hansen, M. A. Bender, J. Stamatoyannopoulos, and T. A. Reh, "DNase I hypersensitivity analysis of the mouse brain and retina identifies region-specific regulatory elements.," *Epigenetics & chromatin*, vol. 8, p. 8, 2015. 9

[99] W. Zhang and J. Jiang, "Genome-wide mapping of DNase I hypersensitive sites in plants.," *Methods in Molecular Biology*, vol. 1284, pp. 71–89, 2015. 9

[100] J. M. Simon, P. G. Giresi, I. J. Davis, and J. D. Lieb, "Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA.," *Nature protocols*, vol. 7, pp. 256–267, Feb. 2012. 9, 11

[101] P. G. Giresi, J. Kim, R. M. McDaniell, V. R. Iyer, and J. D. Lieb, "FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin.," *Genome Research*, vol. 17, pp. 877–885, June 2007. 9, 11

[102] L. Song, Z. Zhang, L. L. Grasfeder, A. P. Boyle, P. G. Giresi, B.-K. Lee, N. C. Sheffield, S. Gräf, M. Huss, D. Keefe, Z. Liu, D. London, R. M. McDaniell, Y. Shibata, K. A. Showers, J. M. Simon, T. Vales, T. Wang, D. Winter, Z. Zhang, N. D. Clarke, E. Birney, V. R. Iyer, G. E. Crawford, J. D. Lieb, and T. S. Furey, "Open chromatin defined by DNaseI and FAIRE identifies regulatory elements

that shape cell-type identity.," *Genome Research*, vol. 21, pp. 1757–1767, Oct. 2011. 9, 38

[103] P. H. Sorensen, S. L. Lessnick, D. Lopez-Terrada, X. F. Liu, T. J. Triche, and C. T. Denny, "A second Ewing's sarcoma translocation, t(21;22), fuses the EWS gene to another ETS-family transcription factor, ERG.," *Nature Genetics*, vol. 6, pp. 146–151, Feb. 1994. 46

[104] D. Herrero-Martín, D. Osuna, J. L. Ordóñez, V. Sevillano, A. S. Martins, C. Mackintosh, M. Campos, J. Madoz-Gúrpide, A. P. Otero-Motta, G. Caballero, A. T. Amaral, D. H. Wai, Y. Braun, M. Eisenacher, K.-L. Schaefer, C. Poremba, and E. de Alava, "Stable interference of EWS–FLI1 in an Ewing sarcoma cell line impairs IGF-1/IGF-1R signalling and reveals TOPK as a new target," *British Journal of Cancer*, vol. 101, pp. 80–90, June 2009. 46

[105] N. Riggi, L. Cironi, P. Provero, M.-L. Suva, K. Kaloulis, C. Garcia-Echeverria, F. Hoffmann, A. Trumpp, and I. Stamenkovic, "Development of Ewing's Sarcoma from Primary Bone Marrow–Derived Mesenchymal Progenitor Cells," *cancerres.aacrjournals.org*. 46

[106] K. Scotlandi, S. Benini, M. Sarti, M. Serra, P.-L. Lollini, D. Maurici, P. Picci, M. C. Manara, and N. Baldini, "Insulin-like growth factor I receptor-mediated circuit in Ewing's sarcoma/peripheral neuroectodermal tumor: a possible therapeutic target.," *Cancer Research*, vol. 56, pp. 4570–4574, Oct. 1996. 46

[107] D. Yee, R. E. Favoni, G. S. Lebovic, F. Lombana, D. R. Powell, C. P. Reynolds, and N. Rosen, "Insulin-like growth factor I expression by tumors of neuroectodermal origin with the t(11;22) chromosomal translocation. A potential autocrine growth factor.," *The Journal of clinical investigation*, vol. 86, pp. 1806–1814, Dec. 1990. 46

[108] K. Scotlandi, S. Benini, P. Nanni, P.-L. Lollini, G. Nicoletti, L. Landuzzi, M. Serra, M. C. Manara, P. Picci, and N. Baldini, "Blockage of insulin-like growth factor-I receptor inhibits the growth of Ewing's sarcoma in athymic mice.," *Cancer Research*, vol. 58, pp. 4127–4131, Sept. 1998. 47

[109] F. van Valen, W. Winkelmann, and H. Jurgens, "Type I and type II insulin-like growth factor receptors and their function in human Ewing's sarcoma cells.," *J Cancer Res Clin Oncol*, vol. 118, no. 4, pp. 269–275, 1992. 47

[110] J. A. Toretsky, T. Kalebic, V. Blakesley, D. LeRoith, and L. J. Helman, "The insulin-like growth factor-I receptor is required for EWS/FLI-1 transformation of fibroblasts.," *The Journal of biological chemistry*, vol. 272, pp. 30822–30827, Dec. 1997. 47

[111] S. Yakar, C. J. Rosen, W. G. Beamer, C. L. Ackert-Bicknell, Y. Wu, J.-L. Liu, G. T. Ooi, J. Setser, J. Frystyk, Y. R. Boisclair, and D. LeRoith, "Circulating levels

of IGF-1 directly regulate bone growth and density.," *The Journal of clinical investigation*, vol. 110, pp. 771–781, Sept. 2002. 47

[112] A. S. Martins, C. Mackintosh, D. H. Martín, M. Campos, T. Hernández, J. L. Ordóñez, and E. de Alava, "Insulin-like growth factor I receptor pathway inhibition by ADW742, alone or in combination with imatinib, doxorubicin, or vincristine, is a novel therapeutic approach in Ewing tumor.," *Clinical cancer research : an official journal of the American Association for Cancer Research*, vol. 12, pp. 3532–3540, June 2006. 47

[113] K. Scotlandi, M. C. Manara, G. Nicoletti, P.-L. Lollini, S. Lukas, S. Benini, S. Croci, S. Perdichizzi, D. Zambelli, M. Serra, C. Garcia-Echeverria, F. Hofmann, and P. Picci, "Antitumor activity of the insulin-like growth factor-I receptor kinase inhibitor NVP-AEW541 in musculoskeletal tumors.," *Cancer Research*, vol. 65, pp. 3868–3876, May 2005. 47, 61

[114] M. C. Manara, L. Landuzzi, P. Nanni, G. Nicoletti, D. Zambelli, P.-L. Lollini, C. Nanni, F. Hofmann, C. Garcia-Echeverria, P. Picci, and K. Scotlandi, "Preclinical in vivo study of new insulin-like growth factor-I receptor–specific inhibitor in Ewing's sarcoma.," *Clinical cancer research : an official journal of the American Association for Cancer Research*, vol. 13, pp. 1322–1330, Feb. 2007. 47, 61

[115] E. A. Kolb, R. Gorlick, R. Lock, H. Carol, C. L. Morton, S. T. Keir, C. P. Reynolds, M. H. Kang, J. M. Maris, C. Billups, M. A. Smith, and P. J. Houghton, "Initial testing (stage 1) of the IGF-1 receptor inhibitor BMS-754807 by the pediatric preclinical testing program.," *Pediatr Blood Cancer*, vol. 56, pp. 595–603, Apr. 2011. 47

[116] A. S. Pappo, S. R. Patel, J. Crowley, D. K. Reinke, K. P. Kuenkele, S. P. Chawla, G. C. Toner, R. G. Maki, P. A. Meyers, R. Chugh, K. N. Ganjoo, S. M. Schuetze, H. Juergens, M. G. Leahy, B. Geoerger, R. S. Benjamin, L. J. Helman, and L. H. Baker, "R1507, a monoclonal antibody to the insulin-like growth factor 1 receptor, in patients with recurrent or refractory Ewing sarcoma family of tumors: results of a phase II Sarcoma Alliance for Research through Collaboration study," *Journal of Clinical Oncology*, vol. 29, pp. 4541–4547, Dec. 2011. 47, 78

[117] H. Juergens, N. C. Daw, B. Geoerger, S. Ferrari, M. Villarroel, I. Aerts, J. Whelan, U. Dirksen, M. L. Hixon, D. Yin, T. Wang, S. Green, L. Paccagnella, and A. Gualberto, "Preliminary efficacy of the anti-insulin-like growth factor type 1 receptor antibody figitumumab in patients with refractory Ewing sarcoma.," *Journal of Clinical Oncology*, vol. 29, pp. 4534–4540, Dec. 2011. 47, 78

[118] S. Malempati, B. Weigel, A. M. Ingle, C. H. Ahern, J. M. Carroll, C. T. Roberts, J. M. Reid, S. Schmechel, S. D. Voss, S. Y. Cho, H. X. Chen, M. D. Krailo, P. C. Adamson, and S. M. Blaney, "Phase I/II trial and pharmacokinetic study of cixutumumab in pediatric patients with refractory solid tumors and Ewing

sarcoma: a report from the Children's Oncology Group.," *Journal of Clinical Oncology*, vol. 30, pp. 256–262, Jan. 2012. 47, 78

[119] A. O'Neill, N. Shah, N. Zitomersky, M. Ladanyi, N. Shukla, A. Uren, D. Loeb, and J. Toretsky, "Insulin-like growth factor 1 receptor as a therapeutic target in ewing sarcoma: lack of consistent upregulation or recurrent mutation and a review of the clinical trial literature.," *Sarcoma*, vol. 2013, p. 450478, 2013. 47

[120] N. Shukla, J. Schiffman, D. Reed, I. J. Davis, R. B. Womer, S. L. Lessnick, E. R. Lawlor, and C. O. G. E. S. B. Committee, "Biomarkers in Ewing Sarcoma: The Promise and Challenge of Personalized Medicine. A Report from the Children's Oncology Group," *Frontiers in Oncology*, vol. 3, p. 141, 2013. 47

[121] I. Vivanco and C. L. Sawyers, "The phosphatidylinositol 3-Kinase AKT pathway in human cancer.," *Nature Reviews Cancer*, vol. 2, pp. 489–501, July 2002. 47, 57

[122] J. Downward, "PI 3-kinase, Akt and cell survival.," *Seminars in cell & developmental biology*, vol. 15, pp. 177–182, Apr. 2004. 47, 57

[123] T. Maehama and J. E. Dixon, "The Tumor Suppressor, PTEN/MMAC1, Dephosphorylates the Lipid Second Messenger, Phosphatidylinositol 3,4,5-Trisphosphate," *Journal of Biological Chemistry*, vol. 273, pp. 13375–13378, May 1998. 47

[124] P. L. Depowski, S. I. Rosenthal, and J. S. Ross, "Loss of expression of the PTEN gene protein product is associated with poor outcome in breast cancer.," *Mod Pathol*, vol. 14, pp. 672–676, July 2001. 47

[125] C. P. Hsu, T. Y. Kao, W. L. Chang, S. Nieh, H. L. Wang, and Y. C. Chung, "Clinical significance of tumor suppressor PTEN in colorectal carcinoma," *European Journal of Surgical Oncology (EJSO)*, vol. 37, pp. 140–147, Feb. 2011. 47

[126] K. Sircar, M. Yoshimoto, F. A. Monzon, I. H. Koumakpayi, R. L. Katz, A. Khanna, K. Alvarez, G. Chen, A. D. Darnel, A. G. Aprikian, F. Saad, T. A. Bismar, and J. A. Squire, "PTEN genomic deletion is associated with p-Akt and AR signalling in poorer outcome, hormone refractory prostate cancer.," *The Journal of Pathology*, vol. 218, pp. 505–513, Aug. 2009. 47

[127] J. Li, C. Yen, D. Liaw, K. Podsypanina, S. Bose, S. I. Wang, J. Puc, C. Miliaresis, L. Rodgers, R. McCombie, S. H. Bigner, B. C. Giovanella, M. Ittmann, B. Tycko, H. Hibshoosh, M. H. Wigler, and R. Parsons, "PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer.," *Science (New York, NY)*, vol. 275, pp. 1943–1947, Mar. 1997. 47

[128] S. I. Wang, J. Puc, J. Li, J. N. Bruce, P. Cairns, D. Sidransky, and R. Parsons, "Somatic mutations of PTEN in glioblastoma multiforme.," *Cancer Research*, vol. 57, pp. 4183–4186, Oct. 1997. 47

[129] D. A. Rubinson, C. P. Dillon, A. V. Kwiatkowski, C. Sievers, L. Yang, J. Kopinja, D. L. Rooney, M. M. Ihrig, M. T. McManus, F. B. Gertler, M. L. Scott, and L. Van Parijs, "A lentivirus-based system to functionally silence genes in primary mammalian cells, stem cells and transgenic mice by RNA interference," *Nature Genetics*, vol. 33, pp. 401–406, Mar. 2003. 49

[130] L. Poliseno, L. Salmena, L. Riccardi, A. Fornari, M. S. Song, R. M. Hobbs, P. Sportoletti, S. Varmeh, A. Egia, G. Fedele, L. Rameh, M. Loda, and P. P. Pandolfi, "Identification of the miR-106b 25 microRNA cluster as a proto-oncogenic PTEN-targeting intron that cooperates with its host gene MCM7 in transformation.," *Science signaling*, vol. 3, no. 117, p. ra29, 2010. 57

[131] Y. Tay, L. Kats, L. Salmena, D. Weiss, S. M. Tan, U. Ala, F. Karreth, L. Poliseno, P. Provero, F. Di Cunto, J. Lieberman, I. Rigoutsos, and P. P. Pandolfi, "Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs.," *Cell*, vol. 147, pp. 344–357, Oct. 2011. 57

[132] N. Konishi, M. Nakamura, M. Kishi, M. Nishimine, E. Ishida, and K. Shimada, "Heterogeneous methylation and deletion patterns of the INK4a/ARF locus within prostate carcinomas.," *The American journal of pathology*, vol. 160, pp. 1207–1214, Apr. 2002. 57

[133] Y. Yamaguchi, T. Takabatake, S. Kakinuma, Y. Amasaki, M. Nishimura, T. Imaoka, K. Yamauchi, Y. Shang, T. Miyoshi-Imamura, H. Nogawa, Y. Kobayashi, and Y. Shimada, "Complicated biallelic inactivation of Pten in radiation-induced mouse thymic lymphomas.," *Mutation research*, vol. 686, pp. 30–38, Apr. 2010. 57

[134] N. Chalhoub, G. Zhu, X. Zhu, and S. J. Baker, "Cell type specificity of PI3K signaling in Pdk1- and Pten-deficient brains.," *Genes & development*, vol. 23, pp. 1619–1624, July 2009. 60

[135] A. Mora, D. Komander, D. M. F. van Aalten, and D. R. Alessi, "PDK1, the master regulator of AGC kinase signal transduction.," *Seminars in cell & developmental biology*, vol. 15, pp. 161–170, Apr. 2004. 60

[136] C. Garcia-Echeverria, M. A. Pearson, A. Marti, T. Meyer, J. Mestan, J. Zimmermann, J. Gao, J. Brueggen, H.-G. Capraro, R. Cozens, D. B. Evans, D. Fabbro, P. Furet, D. G. Porta, J. Liebetanz, G. Martiny-Baron, S. Ruetz, and F. Hofmann, "In vivo antitumor activity of NVP-AEW541-A novel, potent, and selective inhibitor of the IGF-IR kinase.," *Cancer cell*, vol. 5, pp. 231–239, Mar. 2004. 61, 64

[137] M. J. Mulvihill, A. Cooke, M. Rosenfeld-Franklin, E. Buck, K. Foreman, D. Landfair, M. O'Connor, C. Pirritt, Y. Sun, Y. Yao, L. D. Arnold, N. W. Gibson, and Q.-S. Ji, "Discovery of OSI-906: a selective and orally efficacious dual inhibitor of the IGF-1 receptor and insulin receptor," *Future Medicinal Chemistry*, vol. 1, pp. 1153–1171, Sept. 2009. 61, 78

[138] D. J. Kuhn, Z. Berkova, R. J. Jones, R. Woessner, C. C. Bjorklund, W. Ma, R. E. Davis, P. Lin, H. Wang, T. L. Madden, C. Wei, V. Baladandayuthapani, M. Wang, S. K. Thomas, J. J. Shah, D. M. Weber, and R. Z. Orlowski, "Targeting the insulin-like growth factor-1 receptor to overcome bortezomib resistance in preclinical models of multiple myeloma.," *Blood*, vol. 120, pp. 3260–3270, Oct. 2012. 61

[139] S. A. Flanigan, T. M. Pitts, T. P. Newton, G. N. Kulikowski, A. C. Tan, M. C. McManus, A. Spreafico, M. I. Kachaeva, H. M. Selby, J. J. Tentler, S. G. Eckhardt, and S. Leong, "Overcoming IGF1R/IR resistance through inhibition of MEK signaling in colorectal cancer models.," *Clinical cancer research : an official journal of the American Association for Cancer Research*, vol. 19, pp. 6219–6229, Nov. 2013. 61

[140] M. L. Kuijjer, E. F. P. Peterse, B. E. W. M. van den Akker, I. H. Briaire-de Bruijn, M. Serra, L. A. Meza-Zepeda, O. Myklebost, A. B. Hassan, P. C. W. Hogendoorn, and A.-M. Cleton-Jansen, "IR/IGF1R signaling as potential target for treatment of high-grade osteosarcoma.," *BMC Cancer*, vol. 13, p. 245, 2013. 61

[141] L. Asnaghi, P. Bruno, M. Priulla, and A. Nicolin, "mTOR: a protein kinase switching between life and death.," *Pharmacological research*, vol. 50, pp. 545–549, Dec. 2004. 64

[142] S. S. Yea and D. A. Fruman, "Achieving cancer cell death with PI3K/mTOR-targeted therapies.," *Annals of the New York Academy of Sciences*, vol. 1280, pp. 15–18, Mar. 2013. 64

[143] Y. Kabeya, N. Mizushima, T. Ueno, A. Yamamoto, T. Kirisako, T. Noda, E. Kominami, Y. Ohsumi, and T. Yoshimori, "LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing.," *The EMBO Journal*, vol. 19, pp. 5720–5728, Nov. 2000. 64

[144] Y. Chen, P. Chi, S. Rockowitz, P. J. Iaquinta, T. Shamu, S. Shukla, D. Gao, I. Sirota, B. S. Carver, J. Wongvipat, H. I. Scher, D. Zheng, and C. L. Sawyers, "ETS factors reprogram the androgen receptor cistrome and prime prostate tumorigenesis in response to PTEN loss," *Nature Medicine*, vol. 19, pp. 1023–1029, Aug. 2013. 68

[145] D. A. Guertin and D. M. Sabatini, "Defining the role of mTOR in cancer.," *Cancer cell*, vol. 12, pp. 9–22, July 2007. 68

[146] C. Nardella, Z. Chen, L. Salmena, A. Carracedo, A. Alimonti, A. Egia, B. Carver, W. Gerald, C. Cordon-Cardo, and P. P. Pandolfi, "Aberrant Rheb-mediated mTORC1 activation and Pten haploinsufficiency are cooperative oncogenic events," *Genes & development*, vol. 22, pp. 2172–2177, Aug. 2008. 68

[147] J. A. Toretsky, M. Thakar, A. E. Eskenazi, and C. N. Frantz, "Phosphoinositide 3-hydroxide kinase blockade enhances apoptosis in the Ewing's sarcoma family of tumors.," *Cancer Research*, vol. 59, pp. 5745–5750, Nov. 1999. 68

[148] S. Hofbauer, G. Hamilton, G. Theyer, K. Wollmann, and F. Gabor, "Insulin-like growth factor-I-dependent growth and in vitro chemosensitivity of Ewing's sarcoma and peripheral primitive neuroectodermal tumour cell lines.," *Eur J Cancer*, vol. 29A, no. 2, pp. 241–245, 1993. 68

[149] D. Hägerstrand, M. B. Lindh, C. Peña, C. Garcia-Echeverria, M. Nistér, F. Hofmann, and A. Ostman, "PI3K/PTEN/Akt pathway status affects the sensitivity of high-grade glioma cell cultures to the insulin-like growth factor-1 receptor inhibitor NVP-AEW541.," *Neuro-oncology*, vol. 12, pp. 967–975, Sept. 2010. 68

[150] T. Kanzawa, L. Zhang, L. Xiao, I. M. Germano, Y. Kondo, and S. Kondo, "Arsenic trioxide induces autophagic cell death in malignant glioma cells by upregulation of mitochondrial cell death protein BNIP3.," *Oncogene*, vol. 24, pp. 980–991, Feb. 2005. 71, 78

[151] S. Turcotte, D. A. Chan, P. D. Sutphin, M. P. Hay, W. A. Denny, and A. J. Giaccia, "A molecule targeting VHL-deficient renal cell carcinoma that induces autophagy.," *Cancer cell*, vol. 14, pp. 90–102, July 2008. 71, 78

[152] E. A. Kolb, R. Gorlick, J. M. Maris, S. T. Keir, C. L. Morton, J. Wu, A. W. Wozniak, M. A. Smith, and P. J. Houghton, "Combination testing (Stage 2) of the Anti-IGF-1 receptor antibody IMC-A12 with rapamycin by the pediatric preclinical testing program.," *Pediatr Blood Cancer*, vol. 58, pp. 729–735, May 2012. 71, 80

[153] R. T. Kurmasheva, L. Dudkin, C. Billups, L. V. Debelenko, C. L. Morton, and P. J. Houghton, "The insulin-like growth factor-1 receptor-targeting antibody, CP-751,871, suppresses tumor-derived VEGF and synergizes with rapamycin in models of childhood sarcoma.," *Cancer Research*, vol. 69, pp. 7662–7671, Oct. 2009. 71, 80

[154] A. Naing, P. LoRusso, S. Fu, D. S. Hong, P. Anderson, R. S. Benjamin, J. Ludwig, H. X. Chen, L. A. Doyle, and R. Kurzrock, "Insulin growth factor-receptor (IGF-1R) antibody cixutumumab combined with the mTOR inhibitor temsirolimus in patients with refractory Ewing's sarcoma family tumors," *Clinical cancer research : an official journal of the American Association for Cancer Research*, vol. 18, pp. 2625–2631, May 2012. 71, 80

[155] G. K. Schwartz, W. D. Tap, L.-X. Qin, M. B. Livingston, S. D. Undevia, B. Chmielowski, M. Agulnik, S. M. Schuetze, D. R. Reed, S. H. Okuno, J. A. Ludwig, V. Keedy, P. Rietschel, A. S. Kraft, D. Adkins, B. A. Van Tine, B. Brockstein, V. Yim, C. Bitas, A. Abdullah, C. R. Antonescu, M. Condy, M. A. Dickson, S. D. Vasudeva, A. L. Ho, L. A. Doyle, H. X. Chen, and R. G. Maki, "Cixutumumab and temsirolimus for patients with bone and soft-tissue sarcoma: a

multicentre, open-label, phase 2 trial.," *Lancet Oncol*, vol. 14, pp. 371–382, Apr. 2013. 71, 80

[156] J. R. Dixon, I. Jung, S. Selvaraj, Y. Shen, J. E. Antosiewicz-Bourget, A. Y. Lee, Z. Ye, A. Kim, N. Rajagopal, W. Xie, Y. Diao, J. Liang, H. Zhao, V. V. Lobanenkov, J. R. Ecker, J. A. Thomson, and B. Ren, "Chromatin architecture reorganization during stem cell differentiation," *Nature*, vol. 518, pp. 331–336, 02 2015. 11

[157] K. Kurimoto, Y. Yabuta, K. Hayashi, H. Ohta, H. Kiyonari, T. Mitani, Y. Moritoki, K. Kohri, H. Kimura, T. Yamamoto, Y. Katou, K. Shirahige, and M. Saitou, "Quantitative Dynamics of Chromatin Remodeling during Germ Cell Specification from Mouse Embryonic Stem Cells," *Cell stem cell*, vol. 16, pp. 517–532, May 2015. 11

[158] J. A. Wamstad, J. M. Alexander, R. M. Truty, A. Shrikumar, F. Li, K. E. Eilertson, H. Ding, J. N. Wylie, A. R. Pico, J. A. Capra, G. Erwin, S. J. Kattman, G. M. Keller, D. Srivastava, S. S. Levine, K. S. Pollard, A. K. Holloway, L. A. Boyer, and B. G. Bruneau, "Dynamic and coordinated epigenetic regulation of developmental transitions in the cardiac lineage," *Cell*, vol. 151, pp. 206–220, Sept. 2012. 11

[159] G. R. Kafer, S. A. Lehnert, M. Pantaleon, P. L. Kaye, and R. J. Moser, "Expression of genes coding for histone variants and histone-associated proteins in pluripotent stem cells and mouse preimplantation embryos," *Gene Expression Patterns*, vol. 10, pp. 299–305, Sept. 2010. 11

[160] B. Langmead, C. Trapnell, M. Pop, and S. L. Salzberg, "Ultrafast and memory-efficient alignment of short DNA sequences to the human genome," *Genome Biology*, vol. 10, no. 3, p. R25, 2009. 12, 40

[161] J. M. Simon, K. E. Hacker, D. Singh, A. R. Brannon, J. S. Parker, M. Weiser, T. H. Ho, P. F. Kuan, E. Jonasch, T. S. Furey, J. F. Prins, J. D. Lieb, W. K. Rathmell, and I. J. Davis, "Variation in chromatin accessibility in human kidney cancer links H3K36 methyltransferase loss with widespread RNA processing defects," *Genome Research*, vol. 24, pp. 241–250, Feb. 2014. 12, 44

[162] L. A. Boyer, T. I. Lee, M. F. Cole, S. E. Johnstone, S. S. Levine, J. P. Zucker, M. G. Guenther, R. M. Kumar, H. L. Murray, R. G. Jenner, D. K. Gifford, D. A. Melton, R. Jaenisch, and R. A. Young, "Core transcriptional regulatory circuitry in human embryonic stem cells," *Cell*, vol. 122, pp. 947–956, Sept. 2005. 12

[163] Y. H. Loh, Q. Wu, J. L. Chew, V. B. Vega, W. Zhang, X. Chen, G. Bourque, J. George, B. Leong, J. Liu, K. Y. Wong, K. W. Sung, C. W. Lee, X. D. Zhao, K. P. Chiu, L. Lipovich, V. A. Kuznetsov, P. Robson, L. W. Stanton, C. L. Wei, Y. Ruan, B. Lim, and H. H. Ng, "The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells," *Nature Genetics*, vol. 38, pp. 431–440, Apr. 2006. 12

[164] E. P. Consortium, “An integrated encyclopedia of DNA elements in the human genome,” *Nature*, vol. 489, pp. 57–74, Sept. 2012. 12, 14, 16

[165] J. Ernst and M. Kellis, “Discovery and characterization of chromatin states for systematic annotation of the human genome,” *Nature Biotechnology*, vol. 28, pp. 817–825, July 2010. 14

[166] J. Ernst and M. Kellis, “ChromHMM: automating chromatin-state discovery and characterization,” *Nature Methods*, vol. 9, pp. 215–216, Feb. 2012. 14

[167] Y. Zhang, T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li, and X. S. Liu, “Model-based analysis of ChIP-Seq (MACS),” *Genome Biology*, vol. 9, no. 9, p. R137, 2008. 14, 40

[168] A. F. A. Smit, R. Hubley, and P. Green, “RepeatMasker Open-4.0,” 2015. 14

[169] J. A. West, A. Cook, B. H. Alver, M. Stadtfeld, A. M. Deaton, K. Hochedlinger, P. J. Park, M. Y. Tolstorukov, and R. E. Kingston, “Nucleosomal occupancy changes locally over key regulatory regions during cell differentiation and reprogramming,” *Nature Communications*, vol. 5, p. 4719, Aug. 2014. 21, 22

[170] Y. Fu, M. Sinha, C. L. Peterson, and Z. Weng, “The insulator binding protein CTCF positions 20 nucleosomes around its binding sites across the human genome,” *PLoS genetics*, vol. 4, no. 7, p. e1000138, 2008. 21

[171] T. Orvis, A. Hepperla, V. Walter, S. Song, J. Simon, J. Parker, M. D. Wilkerson, N. Desai, M. B. Major, D. N. Hayes, I. J. Davis, and B. Weissman, “BRG1/SMARCA4 inactivation promotes non-small cell lung cancer aggressiveness by altering chromatin organization,” *Cancer Research*, vol. 74, pp. 6486–6498, Nov. 2014. 21

[172] B. E. Bernstein, J. A. Stamatoyannopoulos, J. F. Costello, B. Ren, A. Milosavljevic, A. Meissner, M. Kellis, M. A. Marra, A. L. Beaudet, J. R. Ecker, P. J. Farnham, M. Hirst, E. S. Lander, T. S. Mikkelsen, and J. A. Thomson, “The NIH Roadmap Epigenomics Mapping Consortium,” *Nat Biotech*, vol. 28, no. 10, pp. 1045–1048, 2010. 22, 26, 27, 33

[173] M. M. Sanders, “Fractionation of nucleosomes by salt elution from micrococcal nuclease-digested nuclei,” *The Journal of cell biology*, vol. 79, pp. 97–109, Oct. 1978. 23

[174] S. Henikoff, J. G. Henikoff, A. Sakai, G. B. Loeb, and K. Ahmad, “Genome-wide profiling of salt fractions maps physical properties of chromatin.,” *Genome Research*, vol. 19, pp. 460–469, Mar. 2009. 23

[175] S. S. Teves and S. Henikoff, “Salt fractionation of nucleosomes for genome-wide profiling.,” *Methods in Molecular Biology*, vol. 833, pp. 421–432, 2012. 24, 44

[176] C. Y. McLean, D. Bristor, M. Hiller, S. L. Clarke, B. T. Schaar, C. B. Lowe, A. M. Wenger, and G. Bejerano, "GREAT improves functional interpretation of cis-regulatory regions," *Nature Biotechnology*, vol. 28, pp. 495–501, May 2010. 28

[177] M. Murtha, F. Strino, Z. Tokcaer-Keskin, N. Sumru Bayin, D. Shalabi, X. Xi, Y. Kluger, and L. Dailey, "Comparative FAIRE-seq Analysis Reveals Distinguishing Features of the Chromatin Structure of Ground State- and Primed-Pluripotent Cells," *Stem cells (Dayton, Ohio)*, vol. 33, pp. 378–391, Jan. 2015. 35

[178] E. Meshorer and T. Misteli, "Chromatin in pluripotent embryonic stem cells and differentiation.," *Nature Reviews Molecular Cell Biology*, vol. 7, pp. 540–546, July 2006. 35

[179] J. H. Lee, S. R. Hart, and D. G. Skalnik, "Histone deacetylase activity is required for embryonic stem cell differentiation," *Genesis*, vol. 38, pp. 32–38, Jan. 2004. 35

[180] S. Cuddapah, R. Jothi, D. E. Schones, T. Y. Roh, K. Cui, and K. Zhao, "Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains," *Genome Research*, vol. 19, pp. 24–32, Jan. 2009. 35

[181] K. A. U. Gonzales, H. Liang, Y.-S. Lim, Y.-S. Chan, J.-C. Yeo, C.-P. Tan, B. Gao, B. Le, Z.-Y. Tan, K.-Y. Low, Y.-C. Liou, F. Bard, and H.-H. Ng, "Deterministic Restriction on Pluripotent State Dissolution by Cell-Cycle Pathways," *Cell*, vol. 162, pp. 564–579, July 2015. 37

[182] M. G. Guenther, G. M. Frampton, F. Soldner, D. Hockemeyer, M. Mitalipova, R. Jaenisch, and R. A. Young, "Chromatin structure and gene expression programs of human embryonic and induced pluripotent stem cells," *Cell stem cell*, vol. 7, pp. 249–257, Aug. 2010. 37

[183] I. Brukner, M. Dlakic, A. Savic, S. Susic, S. Pongor, and D. Suck, "Evidence for opposite groove-directed curvature of GGGCCC and AAAAA sequence elements," *Nucleic Acids Research*, vol. 21, pp. 1025–1029, Feb. 1993. 37

[184] C. H. Hsieh and J. D. Griffith, "The terminus of SV40 DNA replication and transcription contains a sharp sequence-directed curve," *Cell*, vol. 52, pp. 535–544, Feb. 1988. 37

[185] P. T. Lowary and J. Widom, "New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning," *Journal of Molecular Biology*, vol. 276, pp. 19–42, Feb. 1998. 37

[186] A. Valouev, S. M. Johnson, S. D. Boyd, C. L. Smith, A. Z. Fire, and A. Sidow, "Determinants of nucleosome organization in primary human cells," *Nature*, vol. 474, pp. 516–520, June 2011. 37

[187] E. W. Englander and B. H. Howard, "Nucleosome positioning by human Alu elements in chromatin," *The Journal of biological chemistry*, vol. 270, pp. 10091–10096, Apr. 1995. 37

[188] Y. Tanaka, R. Yamashita, Y. Suzuki, and K. Nakai, "Effects of Alu elements on global nucleosome positioning in the human genome," *BMC Genomics*, vol. 11, p. 309, 2010. 37

[189] L. A. Owen and S. L. Lessnick, "Identification of target genes in their native cellular context: an analysis of EWS/FLI in Ewing's sarcoma," *Cell cycle (Georgetown, Tex.)*, vol. 5, pp. 2049–2053, Sept. 2006. 37

[190] T. G. Grunewald, V. Bernard, P. Gilardi-Hebenstreit, V. Raynal, D. Surdez, M. M. Aynaud, O. Mirabeau, F. Cidre-Aranaz, F. Tirode, S. Zaidi, G. Perot, A. H. Jonker, C. Lucchesi, M. C. Le Deley, O. Oberlin, P. Marec-Berard, A. S. Veron, S. Reynaud, E. Lapouble, V. Boeva, T. Rio-Frio, J. Alonso, S. Bhatia, G. Pierron, G. Cancel-Tassin, O. Cussenot, D. G. Cox, L. M. Morton, M. J. Machiela, S. J. Chanock, P. Charnay, and O. Delattre, "Chimeric EWSR1-FLI1 regulates the Ewing sarcoma susceptibility gene EGR2 via a GGAA microsatellite," *Nature Genetics*, vol. 47, pp. 1073–1078, Sept. 2015. 38

[191] S. Beck, B. E. Bernstein, R. M. Campbell, J. F. Costello, D. Dhanak, J. R. Ecker, J. M. Greally, J. P. Issa, P. W. Laird, K. Polyak, B. Tycko, P. A. Jones, and A. C. E. T. Force, "A blueprint for an international cancer epigenome consortium. A report from the AACR Cancer Epigenome Task Force," *Cancer Research*, vol. 72, pp. 6319–6324, Dec. 2012. 38

[192] P. P. Lin, M. K. Pandey, F. Jin, S. Xiong, M. Deavers, J. M. Parant, and G. Lozano, "EWS-FLI1 Induces Developmental Abnormalities and Accelerates Sarcoma Formation in a Transgenic Mouse Model," *Cancer Research*, vol. 68, pp. 8968–8975, Nov. 2008. 38

[193] C. Jin, C. Zang, G. Wei, K. Cui, W. Peng, K. Zhao, and G. Felsenfeld, "H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions," *Nature Genetics*, vol. 41, pp. 941–945, Aug. 2009. 38

[194] A. J. Saldanha, "Java Treeview–extensible visualization of microarray data," *Bioinformatics (Oxford, England)*, vol. 20, pp. 3246–3248, Nov. 2004. 43

[195] A. R. Quinlan and I. M. Hall, "BEDTools: a flexible suite of utilities for comparing genomic features," *Bioinformatics (Oxford, England)*, vol. 26, pp. 841–842, Mar. 2010. 43

[196] C. Trapnell, L. Pachter, and S. L. Salzberg, "TopHat: discovering splice junctions with RNA-Seq.," *Bioinformatics (Oxford, England)*, vol. 25, pp. 1105–1111, May 2009. 44

[197] K. Chen, Y. Xi, X. Pan, Z. Li, K. Kaestner, J. Tyler, S. Dent, X. He, and W. Li, "DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing," *Genome Research*, vol. 23, pp. 341–351, Feb. 2013. 45

[198] H. Kovar, G. Jug, D. N. Aryee, A. Zoubek, P. Ambros, B. Gruber, R. Windhager, and H. Gadner, "Among genes involved in the RB dependent cell cycle regulatory cascade, the p16 tumor suppressor gene is frequently lost in the Ewing family of tumors.," *Oncogene*, vol. 15, pp. 2225–2232, Oct. 1997. 77

[199] T. Tsuchiya, K. Sekine, S. Hinohara, T. Namiki, T. Nobori, and Y. Kaneko, "Analysis of the p16INK4, p14ARF, p15, TP53, and MDM2 genes and their prognostic implications in osteosarcoma and Ewing sarcoma.," *Cancer Genet Cytogenet*, vol. 120, pp. 91–98, July 2000. 77

[200] H. Kovar, A. Auinger, G. Jug, D. Aryee, A. Zoubek, M. Salzer-Kuntschik, and H. Gadner, "Narrow spectrum of infrequent p53 mutations and absence of MDM2 amplification in Ewing tumours.," *Oncogene*, vol. 8, pp. 2683–2690, Oct. 1993. 77

[201] D. M. Lerman, M. J. Monument, E. McIlvaine, X. Q. Liu, D. Huang, L. Monovich, N. Beeler, R. G. Gorlick, N. M. Marina, R. B. Womer, J. A. Bridge, M. D. Krailo, R. L. Randall, and S. L. Lessnick, "Tumoral TP53 and/or CDKN2A alterations are not reliable prognostic biomarkers in patients with localized Ewing sarcoma: a report from the Children's Oncology Group," *Pediatr Blood Cancer*, vol. 62, pp. 759–765, May 2015. 77

[202] L. Héron-Milhavet, C. Franckhauser, V. Rana, C. Berthenet, D. Fisher, B. A. Hemmings, A. Fernandez, and N. J. C. Lamb, "Only Akt1 is required for proliferation, while Akt2 promotes cell cycle exit through p21 binding.," *Molecular and cellular biology*, vol. 26, pp. 8267–8280, Nov. 2006. 77

[203] A. Brunet, A. Bonni, M. J. Zigmond, M. Z. Lin, P. Juo, L. S. Hu, M. J. Anderson, K. C. Arden, J. Blenis, and M. E. Greenberg, "Akt Promotes Cell Survival by Phosphorylating and Inhibiting a Forkhead Transcription Factor," *Cell*, vol. 96, pp. 857–868, Mar. 1999. 77

[204] M. Y. Follo, L. Manzoli, A. Poli, J. A. McCubrey, and L. Cocco, "PLC and PI3K/Akt/mTOR signalling in disease and cancer," *Advances in Biological Regulation*, vol. 57, pp. 10–16, Jan. 2015. 77

[205] C. Garca-Echeverra, M. A. Pearson, A. Marti, T. Meyer, J. Mestan, J. Zimmermann, J. Gao, J. Brueggen, H.-G. Capraro, R. Cozens, D. B. Evans, D. Fabbro, P. Furet, D. G. Porta, J. Liebetanz, G. Martiny-Baron, S. Ruetz, and F. Hofmann, "In vivo antitumor activity of NVP-AEW541—A novel, potent, and selective inhibitor of the IGF-IR kinase," *Cancer cell*, vol. 5, pp. 231–239, Mar. 2004. 78

[206] J. Kim, M. Kundu, B. Viollet, and K.-L. Guan, "AMPK and mTOR regulate autophagy through direct phosphorylation of Ulk1," *Nature cell biology*, vol. 13, pp. 132–141, Jan. 2011. 78

[207] M. S. Neshat, I. K. Mellinghoff, C. Tran, B. Stiles, G. Thomas, R. Petersen, P. Frost, J. J. Gibbons, H. Wu, and C. L. Sawyers, "Enhanced sensitivity of PTEN-deficient tumors to inhibition of FRAP/mTOR," vol. 98, pp. 10314–10319, Aug. 2001. 78

[208] H. Gao, J. M. Korn, S. Ferretti, J. E. Monahan, Y. Wang, M. Singh, C. Zhang, C. Schnell, G. Yang, Y. Zhang, O. A. Balbin, S. Barbe, H. Cai, F. Casey, S. Chatterjee, D. Y. Chiang, S. Chuai, S. M. Cogan, S. D. Collins, E. Dammassa, N. Ebel, M. Embry, J. Green, A. Kauffmann, C. Kowal, R. J. Leary, J. Lehar, Y. Liang, A. Loo, E. Lorenzana, E. Robert McDonald, M. E. McLaughlin, J. Merkin, R. Meyer, T. L. Naylor, M. Patawaran, A. Reddy, C. Röelli, D. A. Ruddy, F. Salangsang, F. Santacroce, A. P. Singh, Y. Tang, W. Tinetto, S. Tobler, R. Velazquez, K. Venkatesan, F. Von Arx, H. Q. Wang, Z. Wang, M. Wiesmann, D. Wyss, F. Xu, H. Bitter, P. Atadja, E. Lees, F. Hofmann, E. Li, N. Keen, R. Cozens, M. R. Jensen, N. K. Pryer, J. A. Williams, and W. R. Sellers, "High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response.," *Nature Medicine*, vol. 21, pp. 1318–1325, Nov. 2015. 78

[209] B. F. Niemeyer, J. K. Parrish, N. S. Spoelstra, T. Joyal, J. K. Richer, and P. Jedlicka, "Variable Expression of PIK3R3 and PTEN in Ewing Sarcoma Impacts Oncogenic Phenotypes," *PloS one*, vol. 10, no. 1, p. e0116895, 2015. 79

[210] B. Feng, J.-H. Ng, J.-C. D. Heng, and H.-H. Ng, "Molecules that promote or enhance reprogramming of somatic cells to induced pluripotent stem cells.," *Cell stem cell*, vol. 4, pp. 301–312, Apr. 2009. 72

[211] S. Kobayakawa, K. Miike, M. Nakao, and K. Abe, "Dynamic changes in the epigenomic state and nuclear organization of differentiating mouse embryonic stem cells.," *Genes to cells : devoted to molecular & cellular mechanisms*, vol. 12, pp. 447–460, Apr. 2007. 72

[212] K. Takahashi and S. Yamanaka, "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.," *Cell*, vol. 126, pp. 663–676, Aug. 2006. 73

[213] K. Takahashi, K. Tanabe, M. Ohnuki, M. Narita, T. Ichisaka, K. Tomoda, and S. Yamanaka, "Induction of pluripotent stem cells from adult human fibroblasts by defined factors.," *Cell*, vol. 131, pp. 861–872, Nov. 2007. 73

[214] Y.-D. Jung, K. Ahn, Y.-J. Kim, J.-H. Bae, J.-R. Lee, and H.-S. Kim, "Retroelements: molecular features and implications for disease," *Genes & Genetic Systems*, vol. 88, no. 1, pp. 31–43, 2013. 74

[215] G. Kunarso, N.-Y. Chia, J. Jeyakani, C. Hwang, X. Lu, Y.-S. Chan, H.-H. Ng, and G. Bourque, "Transposable elements have rewired the core regulatory network of human embryonic stem cells," *Nature Genetics*, vol. 42, pp. 631–634, June 2010. 74

[216] V. V. Lunyak, G. G. Prefontaine, E. Núñez, T. Cramer, B.-G. Ju, K. A. Ohgi, K. Hutt, R. Roy, A. García-Díaz, X. Zhu, Y. Yung, L. Montoliu, C. K. Glass, and M. G. Rosenfeld, "Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis.," *Science (New York, NY)*, vol. 317, pp. 248–251, July 2007. 74

[217] K. J. Roux, D. I. Kim, M. Raida, and B. Burke, "A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells," *The Journal of cell biology*, vol. 196, pp. 801–810, Mar. 2012. 76

[218] L. Ho, J. L. Ronan, J. Wu, B. T. Staahl, L. Chen, A. Kuo, J. Lessard, A. I. Nesvizhskii, J. Ranish, and G. R. Crabtree, "An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, pp. 5181–5186, Mar. 2009. 80