I give permission for public access to my Honors paper and for any copying or digitization to be done at the discretion of the College Archivist and/or the College Librarian.

Signed_____

Arishna Chetan Patel

Date _____

Characterizing the functional differences between EZH1 and EZH2 in human embryonic stem cells

Arishna Chetan Patel

Biochemistry and Molecular Biology Program
Rhodes College
Memphis, Tennessee

2016

Submitted in partial fulfillment of the requirements for the
Bachelor of Science degree with Honors in Biochemistry and Molecular Biology

This Honors paper by Arishna Chetan Patel has been read

and approved for Honors in Biochemistry and Molecular Biology.

Dr. Larryn Peterson
Project Advisor

_____

Dr. Loretta Jackson-Hayes
Second Reader

_____

Dr. Rachel Jabaily
Extra-Departmental Reader

_____

Dr. Loretta Jackson-Hayes
Acting Program Director

_____

ACKNOWLEDGEMENTS

I acknowledge the American Lebanese and Syrian Associated Charities (ALSAC) for funding this project. I would like to thank Dr. Catherine Willis, Jonathan Klein, Dr. Xiaoyang Yang, and Dr. Brett Mulvey for not only teaching and mentoring me, but also helping me experimentally. I would also like to thank Yuki Inaba and Helen Xun for experimental help. I owe my deepest gratitude to Dr. Jamy Peng for supervising, editing this thesis, and helping me plan the projects. I would like to thank the Rhodes-St. Jude Summer Plus Fellowship, Dr. Ann Viano, and Dr. Dhammika Muesse for instigating my research. I would also like to thank Matt Jabaily and Rhodes College IT for helping with Adobe Illustrator and my figures. Additionally, I am also grateful for my thesis advisor, Dr. Larryn Peterson, and committee readers, Dr. Loretta Jackson-Hayes and Dr. Rachel Jabaily. Lastly, I would like to acknowledge St. Jude Children's Research Hospital and Rhodes College for making this thesis possible.

CONTENTS

LIST OF TABLES AND FIGURES

Tables

Figures

LIST OF ABBREVIATIONS

ChIP-qPCR: chromatin immunoprecipitation-quantitative polymerase chain reaction

CRISPR: clustered regularly interspaced short palindromic repeats

DMSO: dimethyl sulfoxide

EDTA: ethylenediaminetetraacetic acid

EGTA: ethylene glycol tetraacetic acid

(h/m)ESCs: (human/mouse) embryonic stem cells

EZH1/2: Enhancer of zeste homologue 1/2

gRNA: guide RNA

HEK293T cells: human embryonic kidney 293T cells

HR: homologous recombination

H3K27m3: trimethylation of histone H3 at lysine 27

Indels: insertions and/or deletions

iPS cells: induced pluripotent stem cells

NHEJ: non-homologous end joining repair

NP-40: nonyl phenoxypolytheoxylethanol

PcG proteins: Polycomb-group proteins

PMSF: phenylmethylsulfonyl fluoride (PMSF)

PRC2: Polycomb Repressive Complex 2

RT-qPCR: reverse transcription-quantitative polymerase chain reaction

ABSTRACT

Characterizing the functional differences between
EZH1 and EZH2 in human embryonic stem cells

by

Arishna Chetan Patel

Stem cells are specialized cells that can self-renew and differentiate into multiple cell types during development and regeneration through gene expression changes. These changes in stem cells are achieved, in part, via chromatin modifications, which include histone methylation, phosphorylation, and acetylation. Polycomb Repressive Complex 2 (PRC2) is a protein complex that methylates lysine 27 in histone H3—a histone modification associated with global silencing of gene expression, especially during stem cell development and differentiation. The PRC2 contains an enzymatic subunit, which can be enhancer of zeste homolog 1 or 2 (EZH1 or EZH2). Previous findings suggest that EZH1 and EZH2 serve different roles in embryonic stem cell (ESC) development and differentiation, despite both of them having EED-dependent (another protein subunit of the PRC2 complex) histone methyltransferase activity. Here, I report my investigation of the distinct roles of EZH1 and EZH2 by generating gene mutations separately in human embryonic stem cells by using the CRISPR/Cas9 genome editing technology. The effect of EZH1 and EZH2 mutations on the expression of pluripotency (e.g. *OCT4* and *NANOG*)

and differentiation genes (e.g. *NESTIN*, *Brachyury*, and *GATA4*) was assayed by RT-qPCR. I found that neither EZH1 nor EZH2 mutations affect cell pluripotency; however, EZH2 mutations markedly up-regulate *GATA4* expression, while EZH1 mutations up-regulate *SOX1* and *SOX9* expression.

INTRODUCTION

Stem Cells and Epigenetics

Embryonic stem cells (ESCs), found in the blastocyst during early development, are pluripotent and have the ability to differentiate into three germ layers—endoderm, ectoderm, and mesoderm—*in vivo* and *in vitro* (1). In contrast, adult stem cells mainly serve to maintain tissue homeostasis and are multi- or unipotent. These stem cells can all self-renew and differentiate into terminally differentiated cells during development and regeneration by undergoing gene expression changes, which are achieved, in part, by chromatin modifications, including histone methylation, demethylation and acetylation (2). Generally, acetylation of histones results in an open chromatin state, also known as euchromatin, where transcription is likely occurring. On the other hand, methylation of histones most commonly results in a compact chromatin state, or heterochromatin, where transcription cannot occur since necessary proteins for transcription cannot directly come in contact with the DNA. The study of the heritable gene expression changes and chemical modifications of chromatin is known as epigenetics. Epigenetic regulation of stem cells has been shown to be important for maintaining pluripotency and facilitating differentiation into different cell lineages (3). Additionally, abnormalities in epigenetic mechanisms have implications in cancer tumor formation, by causing deregulated gene expression changes that lead to events such as uncontrolled cell proliferation and defective cell fate programming (4). Therefore, in order to fully understand stem cell regeneration and fate determination, it is important to characterize which modifications influence ESC self-renewal and differentiation, which proteins perform these modifications, and what factors regulate these chromatin modifiers. One such protein

complex responsible for chromatin remodeling and regulating gene expression of key developmental genes is made up of various Polycomb-group (PcG) proteins—Polycomb Repressive Complex 2 (PRC2).

Polycomb Repressive Complex 2

The Polycomb-group (PcG) proteins were discovered in *Drosophila melanogaster* as proteins involved in the repression of the Hox genes, which are transcription factors involved in determining cell identity and differentiation of the anterior-posterior segment during insect development (5-9). PcG proteins were found to be quite conserved in humans (10). They regulate various embryonic development pathways, including cell fate determination and stem cell functions, and affect genes, including tumor suppressors (11-17). PcG proteins form many complexes involved in chromatin modifications, such as the Polycomb Repressive Complex 2 (PRC2). The PRC2 complex is required for ESC differentiation by selectively repressing expression of certain genes, such as DLX5 (neuronal differentiation), BMP3 (bone differentiation), and BMP7 (sex differentiation) (17, 18). The canonical PRC2 complex is composed of three or four main core subunits: SUZ12, EED, Enhancer of zeste homolog 1 or 2 (EZH1/2), and sometimes RbAp46/48 (Figure 1) (17). Its enzymatic activity mainly includes trimethylation of lysine 27 of histone H3 (H3K27m3) via homologous proteins EZH1 or EZH2 and S-adenosyl-L-methionine. (17, 19). These trimethyl marks lead to gene silencing. The PRC2 complex has also been shown to methylate

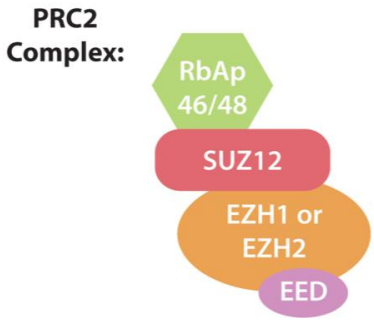


Figure 1: Major components of the Polycomb Repressive Complex 2 (PRC2). PRC2 complexes includes an enzymatic protein, either EZH1 or EZH2.

lysine 9 of histone H3 (H3K9me), which also marks gene silencing (17, 19). The exact mechanism of PRC2-mediated gene silencing is unknown. However, studies suggest that the PRC2 complex and the trimethylated lysine of histone H3 recruit Polycomb Repressive Complex 1 (PRC1), which can bind to the trimethyl mark, as well as the chromatin (20). These cooperative activities of the PRC2 and PRC1 complexes inhibit RNA Polymerase II from binding and transcribing genes—effectively silencing the gene (20).

Since the PRC2 complex is critical to embryonic development, abnormalities in the expression or function of the complex's proteins can inhibit tumor suppressors and activate proto-oncogenes (21). In the case of stem cells, aberrant overexpression of the PRC2 complex proteins leads to unintentional activation of developmental pathways, which improves the cell's ability to proliferate (22). Therefore, deregulation of PRC2 complex components, mainly EZH2, has been implicated in several cancers, such as breast and prostate cancers (21-28). For example, in Adult T cell leukemia/lymphoma, researchers have found an overexpression of the PRC2 complex, which leads to an activation of the NF-kB pathway (22). Abnormal regulation of the NF-kB pathway has been implicated in cancers, as well as autoimmune diseases (29). Additionally, the PRC2 complex is developmentally vital because knockout mutations of EED, EZH2, or SUZ12 all result in embryonic lethality (17).

Differences Between EZH1 and EZH2

Due to the recent discovery and study of EZH1 as another EED-dependent histone methyltransferase, little is known about the mechanistic actions or functions of EZH1 in human ESCs (hESCs) compared to EZH2 (30). The protein sequences of human EZH1

and EZH2 are up to 64% identical (31). Furthermore, studies in mice suggest that EZH1 and EZH2 exhibit different functional roles (32). One study found $EED^{-/-}$ mouse ESCs (mESCs) to have a complete loss of H3K27m3 (as EZH1 and EZH2 are both EED-dependent). In contrast, $EZH2^{-/-}$ mESCs contained detectable, albeit significantly decreased H3K27m3, which is methylated by EZH1 (30). Therefore, EZH1 is able to partially compensate for decreased or no EZH2 function and methylate H3K27 associated with key developmental genes in mESCs (30). Nonetheless, EZH1 does not fully compensate for EZH2 in mESCs, and this provides evidence that EZH1 and EZH2 have different, though partially overlapping, functional roles (30). More recent studies suggest that PRC2 subunits may independently regulate gene expression, suggesting EZH1 or EZH2 may mediate silencing of a different set of genes (33). For example, one study suggests that EZH1 may mediate gene activation by showing that EZH1 physically interacts with the RNA Polymerase II complex in developing muscle cells to promote transcription (34). Thus, this model differs significantly from the canonical model in which EZH1 and EZH2 strictly regulate gene expression through silencing histone marks.

EZH1 is shown to be ubiquitously expressed throughout the organism, while EZH2 expression is higher in proliferating tissues (32). Further, PRC2-EZH2 exhibits higher histone methyltransferase activity compared to PRC2-EZH1 (32). It was also proposed that PRC2-EZH1 suppresses transcription of genes by compacting chromatin and less so by histone H3 lysine 27 methylation (32). Another study showed that EZH1 and EZH2 levels in certain tissues are inversely related: that EZH1 is highly expressed in tissues of the kidney, brain, and skeletal muscle, where little EZH2 expression is detected (35). Therefore, differences in activity and function of EZH1 and EZH2 may also be

correlated to their differential expression patterns. These different proposed functions and activities of EZH1 and EZH2 demand better clarification of the mechanistic activities of these homologs.

CRISPR/Cas9 Genome Editing System

In the present study, the genome editing CRISPR/Cas9 system was utilized to knockdown both EZH1 and EHZ2 individually in H9 cell line of human embryonic stem cells (hESCs) to study potential different molecular and cellular roles of EZH1 and EZH2 (36). Knockdown of a gene is a common research approach to understand the gene's function in a cell. Therefore, methods, such as RNA-directed endonucleases, have been developed to allow for efficient and accurate genome targeting that can be used in a multitude of model systems.

CRISPRs, or, clustered, regularly interspaced, short palindromic repeats, are spaced DNA repeats found in many bacterial genomes (37-40). When viral or plasmid DNA enters a bacterium, the DNA is incorporated between the CRISPR sequences (37-40). Once this DNA is transcribed, this CRISPR RNA (crRNA) complexes with another RNA and the Cas9 endonuclease (41). Finally, the Cas9 complex is guided by the crRNA to areas outside of the CRISPR locus where it may find and bind to a complementary sequence (if the viral or plasmid DNA has incorporated into the bacterial genome elsewhere) (41). Complementarity of the crRNAs and the target genomic sequences allows for the Cas9 protein to create a double stranded break in the DNA (Figure 2). DNA repair of the double stranded break introduces deletion or insertion mutations that can result in effective knockout of the gene product (41).

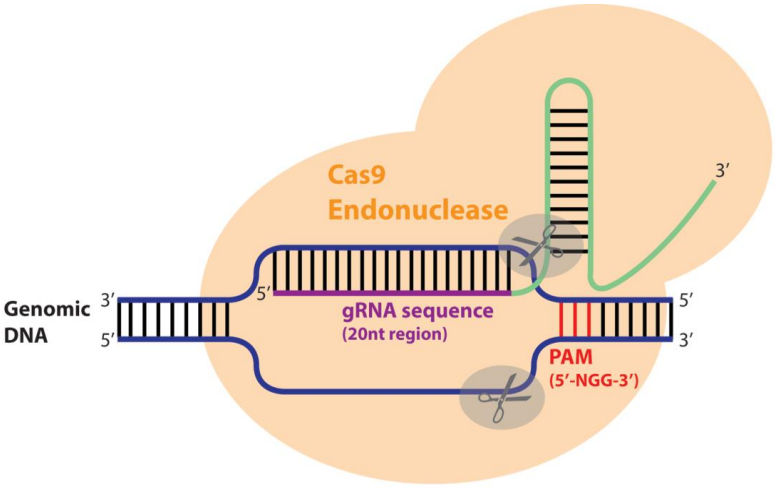


Figure 2: The Cas9 endonuclease complex. The Cas9 endonuclease complex contains the Cas9 protein and an 85-nt CRISPR RNA that contains a 20-nt guide RNA (gRNA) sequence targeting the gene locus of interest. The Cas9-CRISPR complex interacts with the target gene to create double-stranded breaks upstream of the 3-nt PAM sequence.

Researchers were able to modify the CRISPR/Cas9 system for genome editing use in the lab. One modification resulted in the Lenti-CRISPRv2 vector system, which contains a FLAG-tagged Cas9 endonuclease and an 85-nt crRNA that contains a 5' 20-nt guide RNA (gRNA) sequence, which is complementary to a segment of the gene of interest (42) (Figure 3). This vector is packaged into lentivirus, and the target cells are transduced with the virus. Once the vector is inside the target cell and expressed, the Cas9/RNA complex will bind to the gene of interest and introduce double stranded breaks. Double-stranded breaks will initiate DNA repair mechanisms—either homologous recombination (HR) or non-homologous end joining repair (NHEJ). In NHEJ, the ends of DNA are effectively

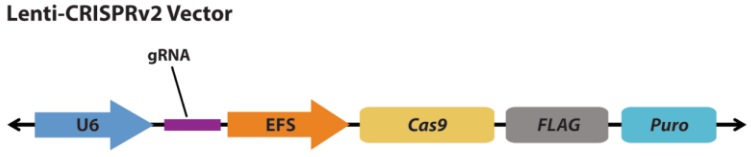


Figure 1: Components of the Lenti-CRISPRv2 vector. It contains two promoters—one upstream of the gRNA sequence and the other upstream of the FLAG-tagged Cas9 coding region. This vector also has a puromycin resistance gene, which serves as a selection marker for cells expressing the vector. Multiple Lenti-CRISPRv2 vectors with different gRNAs targeting EZH1 and EZH2 were cloned for transduction.

ligated back together to often lead to nucleotide insertions or deletions (indels), which can result in frameshift mutations. During translation, frameshift mutations can cause an early stop, resulting in a shortened, non-functional protein and successful knockdown. This process is summarized in Figure 4.

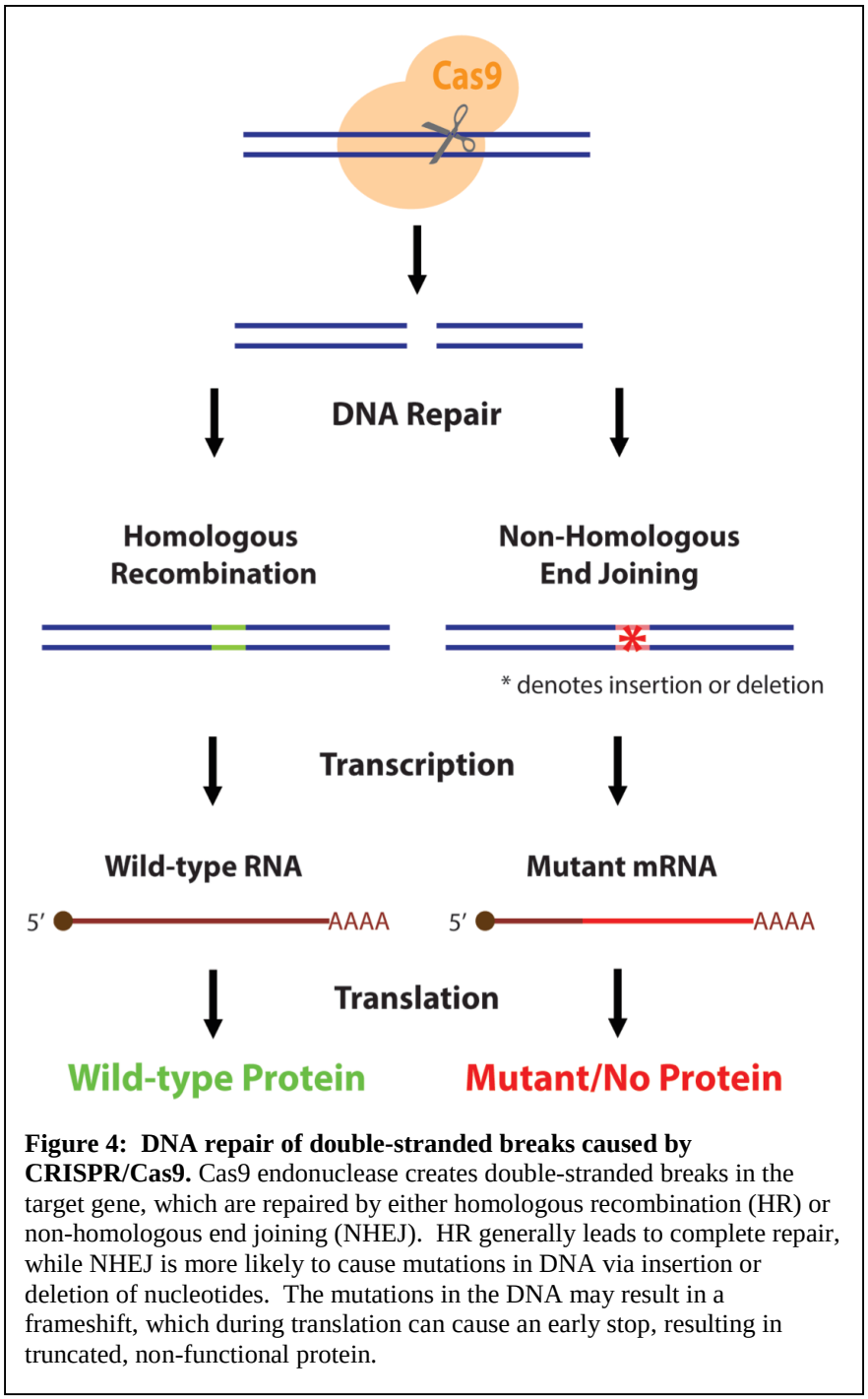


Figure 4: DNA repair of double-stranded breaks caused by CRISPR/Cas9. Cas9 endonuclease creates double-stranded breaks in the target gene, which are repaired by either homologous recombination (HR) or non-homologous end joining (NHEJ). HR generally leads to complete repair, while NHEJ is more likely to cause mutations in DNA via insertion or deletion of nucleotides. The mutations in the DNA may result in a frameshift, which during translation can cause an early stop, resulting in truncated, non-functional protein.

gRNA design is crucial to achieve gene knockdown. Generally, gRNAs are designed to target exons since those constitute the coding region for the protein. Additionally, many gRNAs are designed to target different areas of the same exon (e.g. the 3' and 5' ends). Therefore, cells transduced with multiple Lenti-CRISPRv2 constructs containing different gRNAs may have a higher chance of knockdown because multiple exons or parts of exons are being targeted. Additionally, some gRNAs may not be as effective as others. Therefore, trying multiple gRNAs and combinations of gRNAs is essential.

Motivation

The goal of this study was to identify separate functions of EZH1 and EZH2 in hESC maintenance and differentiation by employing the CRISPR/Cas9 system to generate hESC cells lines. These lines separately contained EZH1 and EZH2 mutations, resulting in protein knockdown (methods summarized in Figure 5). To characterize the cells, the change in mRNA transcript levels of genes involved in pluripotency and germ layer development (e.g., *OCT4*, *T/Brachyury*, *GATA4*, *SOX1*, *SOX9*, and *NESTIN*) were determined for cells with and without protein knockdown for both EZH1 and EZH2 knockdown lines. Further analysis of EZH2 knockdown via chromatin immunoprecipitation-quantitative PCR and histone modification Western blots were also conducted to identify different roles of EZH1 and EZH2. It was hypothesized that the individual knockdown of both EZH1 and EZH2 would affect stem cell pluripotency and induce germ layer cell differentiation; however, cells lacking EZH1 and cells lacking EZH2 would differentiate into cells of different germ layers.

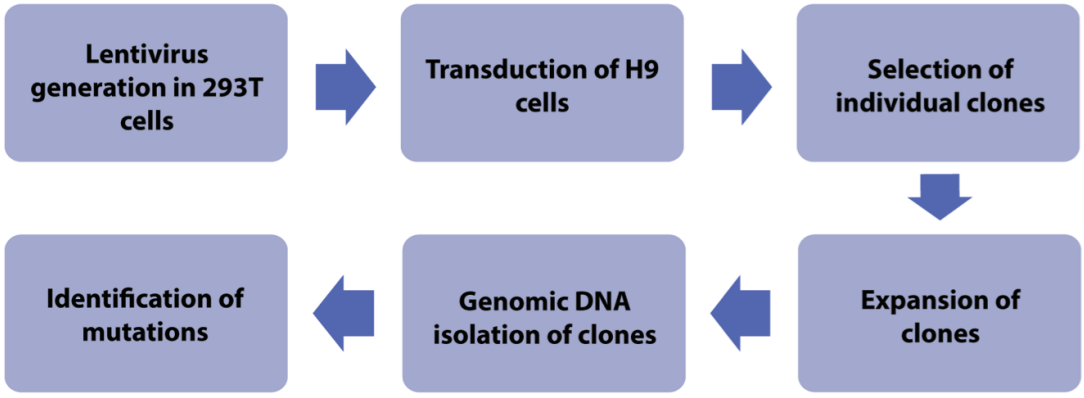


Figure 5: Schematic of knockdown clone generation. Lentiviruses for each Lenti-CRISPRv2 vector were generated to transduce H9 cells. After transduction, individual clones were selected by puromycin and cultured individually. Genomic DNA was isolated from clones, and PCR products were generated from the genomic DNA. PCR products were then ligated into a vector, followed by bacterial transformations. Plasmids were isolated from various bacterial colonies and sequenced by Sanger sequencing.

MATERIALS AND METHODS

Generation of DNA Constructs for CRISPR/Cas9

Plasmid DNAs were constructed for CRISPR/Cas9 using the protocol provided by AddGene (adapted from Zhang lab). Briefly, various guide RNA oligonucleotides (gRNAs) that complement endogenous EZH1 sequences at exons 6 and 7 and EZH2 sequences at exons 6, 7, and 9 were designed and obtained by the lab prior to my arrival (see appendix A). The gRNA 5'-3' and complementing 3'-5' strands were annealed and individually ligated into the LentiCRISPRv2 vector (AddGene Plasmid #52961) by T4 DNA ligase (New England Biolabs). The ligation reactions were transformed into NEB Stable Competent *E. coli* to propagate the constructs. Plasmid DNAs were purified from the resulting clones with the GeneJET Plasmid Miniprep Kit (Thermo Scientific #K0503), according to manufacturer's protocol. Individual gRNA insertion into the vector was confirmed by Sanger sequencing conducted by the St. Jude Hartwell Center. The constructs were stored at -20°C. LentiCRISPR-v2 was a gift from Feng Zhang (43).

Lentivirus Generation in 293T Cells and Virus Titer

The LentiCRISPR-v2 gRNA constructs were used to make lentiviruses in human embryonic kidney 293T (HEK293T) cells based on a protocol by Tiscornia et al. (44). HEK293T cells were acquired from ATCC (11268). Briefly, to generate lentiviruses, about 4.5 million plated HEK293T cells at 70% confluency were transfected with the Lenti-CRISPRv2 plasmid and packaging vectors. For transfection, a mixture of 4 μg each of pVSVL, pREV, pHDL (virus packaging vectors), 8 μg of the LentiCRISPRv2 gRNA construct, and 33 μL of Lipofectamine 3000 reagent (Thermo Scientific) in a total volume of 1 mL with Opti-MEM media was added to the cells. After approximately 24

hours post-transfection, the transfection media was replaced with mTeSR media (Stemcell Technologies 05875) to harvest the virus for 36 hours. The supernatant (which contains the virus) was collected, and more mTeSR was added to the cells for an additional 36 hours to further harvest virus. The virus collections were pooled and filtered through a 45-μm membrane.

The virus titer was determined using the Quant-X 1 Step qRT-PCR Kit and Lenti-X qRT-PCR Titration Components (Clontech 638317; Clontech 631236) according to manufacturer's protocol. To prepare the samples for titer, virus was treated with DNaseI, diluted with H_2O, and denatured at 95°C for five minutes. This mixture was used with RT-qPCR reactions along with 2X Quant-X Buffer, Lenti-X Forward Primer, Lenti-X Reverse Primer, Quant-X Enzyme, RT Enzyme Mix, and RNase-free H_2O per reaction for titer quantitation. Thermocycler conditions followed manufacturer's protocol. This was done in duplicate. Cycle threshold (C_T) values, or the number of qPCR cycles required for the fluorescent signal to overcome the threshold, for the various virus titers were around 20-21 for optimal virus production. EZH2 lentivirus creation and titer were performed by Jamy Peng. See Appendix B for EZH1 lentivirus titer calculations.

H9 Cell Culture

H9 human embryonic stem cells were obtained from WiCell (WA09). All clones were maintained in mTeSR media with 0.325 μg/mL puromycin to select only for cells expressing CRISPR/Cas9. Cells were passaged when about 70% confluent with mTeSR (no puromycin until 24 hours post-passage). Cell media was changed every day. Cell stocks were frozen in Knockout Serum with 10% dimethyl sulfoxide (DMSO) and stored either at -80°C or in liquid nitrogen.

Transduction of H9 ESCs

H9 ESCs were transduced with the lentivirus (individual gRNA viruses or a combination of gRNA viruses). Infected cells were selected by puromycin resistance, and resistant colonies (referred to as clones) were picked for further culturing. Approximately 200,000 H9 ESCs were mixed with 1 mL of lentiviruses (can contain different gRNAs) in mTeSR media containing 8 μg/mL polybrene and plated in two single wells of a 6-well culture plate. After three hours, the cells were transduced again with fresh media containing 8 μg/mL polybrene and lentivirus. After another three hours, the media and lentivirus were replaced with fresh mTeSR. Twenty-four hours later, the cells were treated with 0.325 μg/mL puromycin in mTeSR to select for transduced cells. Five days after culturing, small cell clones (presumably originated from single cells) were picked by pipet tip and placed in separate wells of a 24-well plate. After two weeks, the cells were amplified, frozen in stocks, and their genomic DNA was extracted using the PureLink Genomic DNA Mini Kit (Life Technologies K1820-02) according to manufacturer's protocol. EZH2 transductions and clone picking were done by Jamy Peng.

Genomic DNA Isolation and Mutation Identification of Clones

Specific PCR primers (see Appendix A) were used to amplify the regions of the genomic DNA specific to the exon targeted by the gRNAs. Each reaction contained Phusion High-Fidelity PCR Master Mix (1X Phusion High-Fidelity Buffer, 200 μM dNTP, 3% DMSO, 2X Phusion DNA Polymerase), 4 μM reverse and forward primer mix, and 150 ng of genomic DNA. The PCR products of each sample were purified using the GeneJET Gel Extraction and DNA Cleanup Micro Kit (Thermo Scientific #K0832) according to manufacturer's protocol.

The purified PCR products were ligated into the pJET vector via the CloneJET PCR Cloning Kit vector using the Blunt-End Cloning Protocol provided by the manufacturer (Thermo Scientific #K1232). Each ligation reaction contained 2X Reaction Buffer, 25 ng PCR product, pJET1.2 (vector), T4 DNA ligase, and dH_2O. These products were used to transform lab-made TOP10 bacteria and plated on carbenicillin agar plates. Between 8 and 12 colonies were picked and cultured in TB-carbenicillin media. The plasmids were extracted using the GeneJET Plasmid Miniprep Kit according to manufacturer's protocol (Thermo Scientific #K0503). Next, the plasmids were sent for Sanger Sequencing at the St. Jude Hartwell Center. The sequencing results were analyzed using "ApE-A Plasmid Editor" program to determine if there were any mutations, and if so, the nature of the mutations (insertion or deletion) and translation outcome from the mutated sequence(s). Mutation analysis of some EZH2 clones was done by Yuki Inaba.

Reverse Transcriptase-Quantitative PCR (RT-qPCR)

RNA was extracted from the knockdown clones using the GeneJET RNA Purification Kit (Thermo Scientific #K0732) according to manufacturer's protocol. Samples were then treated with DNaseI to remove any DNA, and concentrated using the GeneJET RNA Cleanup and Concentration Micro Kit (Thermo Scientific #K0842) according to manufacturer's protocol. Reverse transcription (RT) was performed for each RNA sample (in triplicate). Each reaction contained 300 ng of RNA, Reverse Transcriptase (Thermo Scientific #4311235), 1X RT Buffer, 1X dNTP, 1X Random Primers, RNase Inhibitor, and dH_2O. The cDNA was used for quantitative PCR (qPCR) to assay for different developmental and pluripotency gene expression levels in duplicate. These

genes include: *β-ACT* (housekeeping gene), *OCT4*, *NANOG*, *NESTIN*, *SOX1*, *T/Brachyury*, *SOX9*, *CXCR4*, and *GATA4*. Each clone had six qPCR replicates to analyze by quantitative PCR (qPCR; 3 RTs x 2 qPCR samples per clone). Each qPCR reaction included iTaq Sybr Green Mix, 4 μg target gene forward and reverse primers, 0.6 μg cDNA, and dH_2O. The C_T values from each qPCR were normalized using β-ACT and then compared to the values of the control clones to determine statistically significantly up-regulation, down-regulation, or no change in that specific gene's expression compared to the control. Statistical significance was determined using a one-tailed, homoscedastic Student's *t* test.

Protein Extract Preparation and Western Blotting

Clone cell pellets were harvested for cellular fractionation. The cytoplasmic fraction was removed using Buffer A (10mM Tris-HCl pH 8.0, 10 mM KCl, 1.5 mM $MgCl_2$, 340 mM sucrose, 10% glycerol, 0.1% TritonX-100, 1 mM dithiothreitol, and 1X Protease Inhibitor). Next, the remaining cell pellet was treated with Buffer D (40 mM Tris-HCl pH 8.0, 0.5 mM EDTA, 300 mM KCl, 10% glycerol, 0.1% TritonX-100, 1mM dithiothreitol, and 1X Protease Inhibitor) to isolate the nuclear fraction. Lastly, this supernatant was diluted by 2-fold using equal volume of deionized water. Approximately 100 μg of protein from each nuclear extract preparation were loaded onto a 6% SDS-PAGE gel for electrophoresis and subsequent Western blotting. The proteins from the gel were transferred onto a nitrocellulose membrane via semi-wet transfer (1 hour at 14V). The blot was blocked in 2% Bovine Serum A (BSA) in HEPM with 0.002% sodium azide and incubated in primary antibody dilutions, anti-EZH1 (1:1000) or anti-EZH2 (1:1000) and anti-GAPDH (1:1000) overnight at 4°C with agitation. The primary

antibodies were detected by dye-conjugated secondary antibodies using infrared imaging. Using the background-subtracted intensity signal values quantitated from the "ImageStudio" program, relative intensity level (to the control sample) for each protein band was determined. The formula of calculating relative level is $\frac{\frac{\text{Clone signal}}{\text{GAPDH signal}}}{\frac{\text{Control signal}}{\text{GAPDH signal}}}$. See Appendix C for calculations.

Histone extracts for the EZH2 clones were obtained by incubating the remaining chromatin pellets from the cell fractionation in 0.1 M HCl overnight at 4°C. Approximately 5 μg of protein per each clone were loaded for electrophoresis and western blotting as described above. The blot was probed using anti-H3K27m3 (1:1000) or anti-H3 (1:1000) antibodies. Imaging and quantification were done according to description above. Western blot depicted in results was run by Jamy Peng (Figure 10A).

Chromatin Immunoprecipitation-Quantitative PCR (ChIP-qPCR)

Cell pellets of each clone were harvested on ice and cross-linked with 1.1% formaldehyde in PBS. Cross-linking was quenched using 0.125 M glycine in PBS. Next, the cells were washed in PBS and lysed with a lysis buffer containing: 50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% nonyl phenoxypolytheoxylethanol (NP-40), 0.25% TritonX-100, 1X Protease Inhibitor, and 1mM phenylmethylsulfonyl fluoride (PMSF). Next, the supernatant was removed. The pellets were then treated with another lysis buffer containing: 10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1X Protease Inhibitor, and 1 mM PMSF. Again, the supernatant was removed. Next, the pellets were treated with a third lysis buffer containing: 10 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 EGTA, 0.1% Na-Deoxycholate, 0.5% N-Lauroylsarcosine, 1X Protease Inhibitor, and 1 mM

PMSF. This sample was sonicated twice using the Biorupter ® PICO and Minichiller (4°C) with the following program: 15 seconds ON, 45 seconds OFF for 5 cycles. Chromatin DNA and protein concentrations of each lysate were determined using a NanoDrop 2000 UV-Vis Spectrophotometer.

For the immunoprecipitation, 300 µg of protein from each lysate were added to sets of protein A and G dynabeads that were preconjugated for at least 2 hours with either anti-H3K27m3 antibody or non-specific IgG antibody (per clone sample). Each IP was washed quickly in RIPA Buffer (50 mM HEPES-KOH pH 7.5, 500 mM LiCl, 1 mM EDTA, 1% NP-40, and 0.7% Na-Deoxycholate) 5 times and once with 1X TE Buffer (10 mM Tris and 1 mM EDTA). DNA from each IP was eluted using ChIP Elution Buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, and 1% SDS). Each IP elution was reverse-crosslinked at 72°C for six hours with 5 µL of Protease K from the ReliaPrep FFPE gDNA Miniprep System (Promega A2352). DNA was isolated from these samples using the ReliaPrep FFPE gDNA Miniprep System (Promega A2352) and the concentrations were found using the Qubit dsDNA High Sensitivity Assay Kit (Life Technologies Q32851). The ChIP samples for each knockdown clone were analyzed using qPCR for the following genes—*β-ACT*, *OCT4*, *GATA4*, *SOX1*, *SOX2*, and *PAX6*. Each reaction included iTaq Sybr Green Mix, target gene forward and reverse primer mix, recovered ChIP DNA, and dH2O. Each reaction was run in duplicate. The C_T values were normalized using *β-ACT* and compared to control (IgG) values. A one-tailed, homoscedastic Student's *t* test was used for statistical analyses.

RESULTS AND DISCUSSION

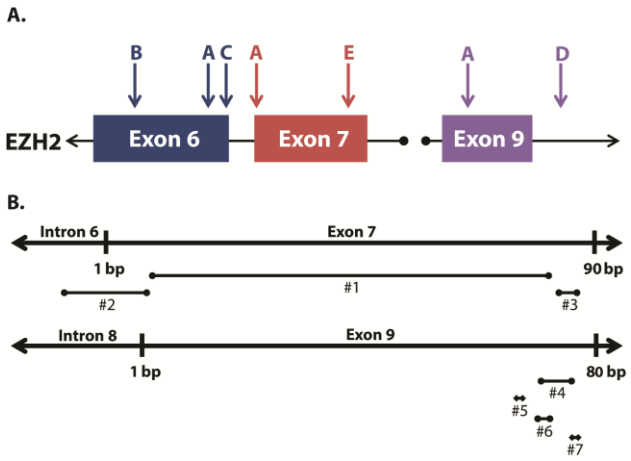
Utilizing the CRISPR/Cas9 System for Genome Editing in hESCs

Knockdown of EZH2

Three transductions with lentiviruses containing different gRNAs targeting EZH2 exons 6, 7, and 9 were performed (Figures 6A). Seventeen clones were generated from the three transductions (Table F6-2). Of the EZH2 exon 6 gRNAs A, B, and C (6-ABC) clones, only clone 2 showed over 50% knockdown—55.9% by Western blot. Therefore, 6-ABC clones were not further characterized. For the exon 7 gRNAs A and E (7-AE) transduction clones, all four clones showed 85% or greater knockdown. 7-AE Clones 1 (97.5% knockdown), 2 (89.9% knockdown), and 3 (97.8% knockdown) were chosen for further studies. For the transduction with gRNAs A and D (9-AD), all seven clones except clone 2 showed greater than 85% knockdown. Only clones 3 (88.4% knockdown), 6 (99.9% knockdown), and 7 (98.2% knockdown) were chosen for further studies since they had the greatest and most consistent knockdown with repeated analysis (Figure 6C).

The above clones chosen for further characterization were sequenced to identify the mutations caused by the DNA repair mechanism following CRISPR/Cas9-mediated DNA damage (Figure 6B & Table F6-1). All of the exon 7 clones showed an approximate 76 bp deletion starting around base pair 11 (if base pair 1 is the first of exon 7). Sequencing of clone 7-AE2 found that 33.3% of the cells contained an 18 bp deletion starting eight base pairs before the exon and also another 2 bp deletion starting at base pair 88. Sequencing results for the 9-AD clones were not as uniform nor as consistent between clone types; however, some cells in each clone type showed an approximate 5

bp deletion near base pair 70 (if base pair 1 is the first of exon 9). Twenty percent of the 9-AD3 clone cells showed a 1 bp insertion between base pairs 64 and 65. Additionally, 40% of the 9-AD6 clone cells show two small (4 bp and about 2 bp) deletions starting at base pairs -20 (intron 8) and 606, respectively. For the 9-AD7 clones, 66.7% showed a 5 bp insertion, TGGGT, between base pairs 72 and 73. In all, these data verified that the knockdown of EZH2 protein seen in both 7-AD and 9-AD clones, visualized by Western blot, is due to specific mutations caused by defective DNA repair of the EHZ2 gene (Figure 6).



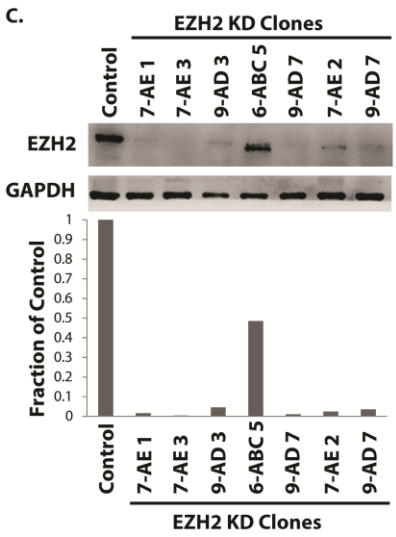


Table F6-1: Summary of EZH2 clone mutations

		EZH2 Ex. 7 Knockdown Clones		
#	Mutation	7-AE1	7-AE2	7-AE3
1	~76 bp del.	100%	33.3%	100%
2	18 bp del.	—	66.7%	—
3	2 bp del.	—	66.7%	—

		EZH2 Ex. 9 Knockdown Clones		
#	Mutation	9-AD3	9-AD6	9-AD7
4	~5 bp del.	80%	60%	33.3%
5	1 bp ins	20%	—	—
6	~2 bp del.	—	40%	—
7	5 bp ins.	—	—	67.7%

Table F6-2: Summary of generated EZH2 clones

EZH2 Transduction	# of Clones	Clones with Mutations	Clones with >80% KD
Ex.6-ABC	6	0	0
Ex.7-AE	4	4	4
Ex.9-AD	7	7	6

Figure 6: CRISPR/Cas9-mediated knockdown of EZH2. A. Schematic of gRNAs targeting the exons 6, 7 and 9 of the EZH2 gene locus. The arrows indicate the areas targeted by each specific gRNA (represented by letters). **B.** Schematic of the types of mutations found in either EZH2 exon 7 or exon 9 in the knockdown clones. The lines indicate the areas of each exon involved in the mutation. Lines with circle heads denote deletions, while lines with diamond heads denote insertions. The numbers correlate to the indels listed in Table F6-1. **C.** Representative Western blot and quantification of EZH2 protein levels relative to the control cells. GAPDH is a loading control. **Table F6-1.** List of the various indels and their frequencies found in the EZH2 clones. The percentages represent the number of cells from each clone which indicated the given mutation. **Table F6-2.** Summary table of clones generated by three transductions of EZH2 gRNA-containing Lenti-CRISPR virus combinations. In total, 17 clones were produced and screened, 11 had mutations, and 10 showed greater than 80% knockdown of EZH2 protein.

Knockdown of EZH1

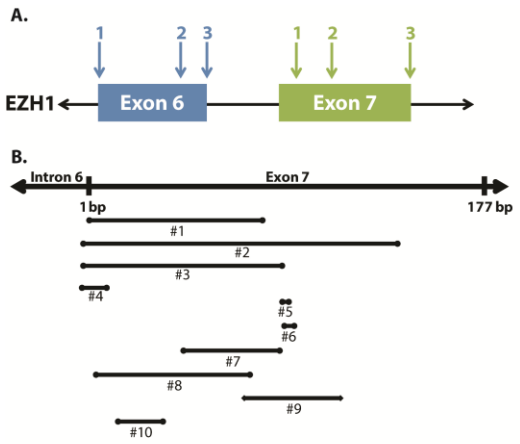
Unlike the clone characterization procedure followed for EZH2 knockdown clones, EZH1 clones were first screened by sequencing to identify mutations. The reason for this change was to decrease the amount of cell culturing for initial screen for knockdown. The amount of cells needed for genomic DNA extraction is far less than the amount of cells needed for nuclear protein extraction. Additionally, identifying the clones with

mutations first allowed for a better prediction of whether protein knockdown occured. Therefore, the EZH1 clones were first screened by mutations, then promising clones for protein knockdown based on their specific mutations were screened via Western blotting.

Three transductions with lentiviruses containing different gRNAs targeting EZH1 exons 6 and 7 were performed (Figures 7A). Greater than 25 clones were generated (Table F7-2). For the exon 6 gRNAs 1, 2 and 3 (Ex.6-123) transduction, greater than three clones were generated; however, no mutations (only wild-type sequences) were identified. For the ten clones produced from the exon 7 gRNAs 1, 2, and 3 (Ex.7-123) transduction, six clones (denoted 7-clone)—7-2, 7-3, 7-4, 7-5, 7-6, and 7-7—contained various mutations in EZH1 exon 7 (Figure 7B & Table F7-1). The most common mutation was a 92-98 bp deletion starting about eight base pairs before exon 7 to position 91 within exon 7. This specific mutation, in most clones, would lead to an early stop codon during translation and result in truncated EZH1 protein. In general, five of the Ex.7-123 clones contained mutations that would result in an early stop codon.

The third transduction contained a combination of all six EZH1 gRNA lentiviruses targeting exon 6 and 7 (Ex.6-123 & Ex.7-123). This transduction yielded 12 clones (denoted 67-clone); however, none contained mutations in exon 6, and only three clones had mutations in exon 7. Of these three clones, two also contained a 92-98 bp deletion spanning the early parts of exon 7. Only one clone, 67-11, had a mutation that would lead to an early stop codon during translation. As summarized in Figure 7B and Table F7-1, most of the knockdown clones contained a multitude of mutations within each clone type, suggesting that the clones were mosaic (containing different genotypes).

After looking at the nature of mutations in the clones, four were selected for characterization by Western blot—7-2, 7-3, 7-4, and 67-11 (Figure 7C). Clones 7-2 (26% knockdown), 7-3 (52% knockdown), and 7-4 (13% knockdown) all showed some degree of knockdown of EZH1 protein, relative to control. However, both replicates of clone 67-11 did not indicate knockdown of EZH1. In fact, this clone seemed to have more EZH1 expression than the control. Clones 7-2, 7-3, and 7-4 were chosen for further analysis.



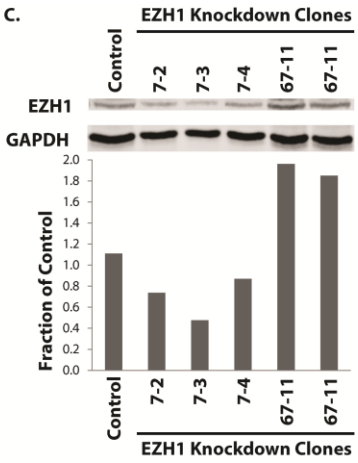


Table F7-1: Summary of EZH1 clone mutations

		EZH1 Ex. 7 Knockdown Clones							
#	Mutation	7-2	7-3	7-4	7-5	7-6	7-8	67-3	67-5	67-11
1	85 bp del.	✓								
2	107-118 bp del.	✓								✓
3	92-98 bp del.	✓	✓		✓	✓		✓		✓
4	~12 bp del.	✓								
5	1 bp del.		✓							
6	4-10 bp del.		✓	✓	✓					
7	44 bp del.				✓					
8	72-77 bp del.				✓		✓		✓	
9	~42 bp ins.								✓	
10	24 bp del.								✓	

✓indicates presense of mutation in clone's cell population

Table F7-2: Summary of generated EZH1 clones

EZH1 Transduction	# of Clones	Clones with Mutations	Clones with Early Stop Mutations
Ex.6-123	3+	0	0
Ex.7-123	10	6	5
Ex.6-123 & Ex.7-123	12	3	1

Figure 7: CRISPR/Cas9-mediated knockdown of EZH1. A. Schematic of gRNAs targeting exons 6 and 7 of the EZH1 gene locus. The arrows indicate the areas of targeted by each specific gRNA (represented by letters). **B.** Schematic of the types of mutations found in EZH1 exon 7 in the knockdown clones. The lines indicate the areas of the exon involved in the mutations. Lines with circle heads denote deletions, while lines with diamond heads denote insertions. The numbers correlate to the indels listed in Table F6-1. **C.** Representative Western blot and quantifications of EZH1 protein levels relative to the control cells. GAPDH is a loading control. **Table F7-1.** List of the various indels and their frequencies found in the EZH1 clones. The check mark indicates that the specific mutation was found at least once in the clone's cell population. **Table F7-2.** Summary table of clones produced by three transductions of EZH1 gRNA-containing Lenti-CRISPR virus combinations. Greater than 25 clones were generated and screened. Of those, nine contained mutations and six had mutations resulting in an early stop codon during translation.

CRISPR/Cas9 system successfully produces gene-specific mutant cells

The CRISPR/Cas9 system can be utilized to engineer target-specific genomic mutations, some of which can result in gene knockdown. However, while many have consistently reported ease of use and efficiency of CRISPR/Cas9 in mammalian cells, it was found that successful knockdown in human embryonic stem cells not only requires ample troubleshooting and time, but also is highly dependent on the genomic region(s) of interest (45). CRISPR/Cas9-mediated indels were engineered at both EZH2 and EZH1 loci. Nonetheless, the nature of the type(s) of indel(s) created varies greatly between the two genes. Results indicate that the specific mutations found at both exons 7 and 9 of EZH2 are found in multiple clones, and the variability of mutations is fairly conserved (Figure 6B & Table F6-1). Additionally, two clones, 7-AE1 and 7-AE3, indicated that the same mutations occurred in both alleles of EZH2 since 100% of the cells tested in each clone showed identical mutations.

In contrast, engineered mutant clones of EZH1 genes often contained more than three types of mutations (Figure 7B & Table F7-1). This is likely due to clone mosaicism, or a population with varied genomes (even though each cell of a single clone should theoretically have originated from the same cell). However, since the cells were not separated into single cells post-transfection and were merely hand-picked once cell clumps began to form, there is some uncertainty whether each cell in each clump came from the same original cell. Therefore, a clone may have contained cells with different genotypes. The EZH1 clones, compared to the EZH2 clones, showed a greater variability in the types of mutations found, as well as the occurrence of each mutation in each clone

type (Tables F6-1 & F7-1). At least four of the tested EZH1 clones were mosaic as they had greater than three types of mutations occurring at an EZH1 locus (Figure 7B).

Along with clonal mosaicism, other difficulties of CRISPR/Cas9 included the low yield of clones containing gene disruptions and the excessive culturing. This technical challenge was partially due to the model system being hESCs, which are resistant to genetic manipulations such as transgene integration/expression. The functional importance of the gene of interest to stem cell pluripotency is also a key factor: EZH1 mutations may likely negatively influence stem cell functions (detailed below) to inhibit formation of knockout cell clone. One consequence is that more than half of the picked clones (after each transduction) did not survive. Furthermore, of the EZH1 clones that did survive, only a select few had frameshift mutations that could lead to gene knockdown (Table F7-2). These results suggest an active selection against cells containing EZH1 knockout.

These studies have also demonstrated that gRNA design is crucial for CRISPR/Cas9 success. Exon 6 gRNAs did not result in any mutations in the respective genes since only wild-type sequences were identified in all clones. Additionally, the location of gRNA target and the location of gene disruptions are indeed correlated; however, they are not always identical. Many mutations occurred near the genomic regions targeted by gRNAs. For example, EZH2 9-AD clones showed different gene disruptions around the 608^{th} bp of the exon, near the location for where gRNA D targets (Figures 6A & B).

Perhaps the greatest difference between the EZH2 and EZH1 clones is the efficiency of knockdown as quantified by protein levels. While the EZH2 protein levels

in most of the exon 7 and 9 clones indicated near knockout or greater than 80% knockdown (Figure 6C & Table F6-2), the EZH1 clones did not show greater than 50% knockdown by Western blot, even though mutations resulting in early translational stop of the gene were identified (Figures 7C & Table F7-2). One explanation for the low knockdown efficiency is that the mutation of one allele, which can be transcribed, does not lead to an early translational stop. A mosaic clone population would also result in lower knockdown efficiency. Another explanation is that post-transcriptional mechanisms, such as gene-splicing, can skip exons to generate small proteins excluding the mutated region (46). Nevertheless, these proteins still remain functional.

Lastly, there may be less calculated knockdown than that is actually present due to the method of quantifying knockdown—Western blotting, which only provides information about protein abundance. Western blots also do not give any information about functionality; therefore, if the protein was still being produced and it was not functional, then the cells would still have functional knockdown/knockout. Therefore, although the CRISPR/Cas9 system can be used to generate cells with site-specific mutant genes, the experimental process is fraught with numerous technical challenges.

Pluripotency and Germ Layer Marker Analysis of EZH1 and EZH2 Knockdown Cells

Knockdown of EZH2 or EZH1 does not result in loss of pluripotency of ESCs

RT-qPCR assays were conducted to compare the level of mRNAs isolated from control, EZH2 knockdown, and EZH1 knockdown clones. More specifically, mRNA levels of pluripotency and various known differentiation markers were tested. The following EZH2 clones were selected: 7-AE1, 7-AE2, 7-AE3, 9-AD3, 9-AD6, and 9-AD7. The

following EZH1 clones were selected: 7-2, 7-3, and 7-4. The mRNA levels of two transcription factors involved in pluripotency—*OCT4* and *NANOG*—were quantified. Three EZH2 knockdown clones, 7-AE2, 9-AD3, and 9-AD6, showed a statistically significant up-regulation of *OCT4*, while all but clone 9-AD3 showed a statistically significant down-regulation of *NANOG*, compared to control (Figure 8A). Nonetheless, the general slight down-regulation of *OCT4* is not enough to conclude that the EZH2 knockdown cells are no longer pluripotent, especially since *NANOG* expression is not decreased. The unaffected *OCT4* and *NANOG* levels in the EZH1 knockdown clones, compared to control, also indicate that the clones are pluripotent (Figure 8B). Only clone 7-2 shows statistically significant up-regulated *OCT4* when compared to control, while none of the clones have any changes in *NANOG* expression.

PcG proteins are important in maintaining cell fate, plasticity, and renewal (47). However, the results indicate that loss of EZH2 does not directly affect pluripotency of hESCs, as seen by the unaffected level of pluripotency markers, OCT4 and NANOG, in the knockdown cells compared to the control (Figures 8A & B). Therefore, while EZH2 function may be necessary for initial establishment of pluripotency, it is not needed for maintenance of pluripotency, as shown in induced pluripotent stem (iPS) cells (48). Additionally, while literature suggests that EZH1 is needed to maintain mouse embryonic stem cell pluripotency and self-renewal, the mutations created in this study are not sufficient to impact pluripotency marker expression (30). Other studies suggest that EZH1 plays a more prominent role in adult human stem cells by repressing master regulators of senescence and early, unscheduled differentiation (49). Thus, it is possible that the reason why the reduction of EZH1 does not result in significant changes in

embryonic stem cells is because the enzyme is not the major player in embryonic stem cell identity.

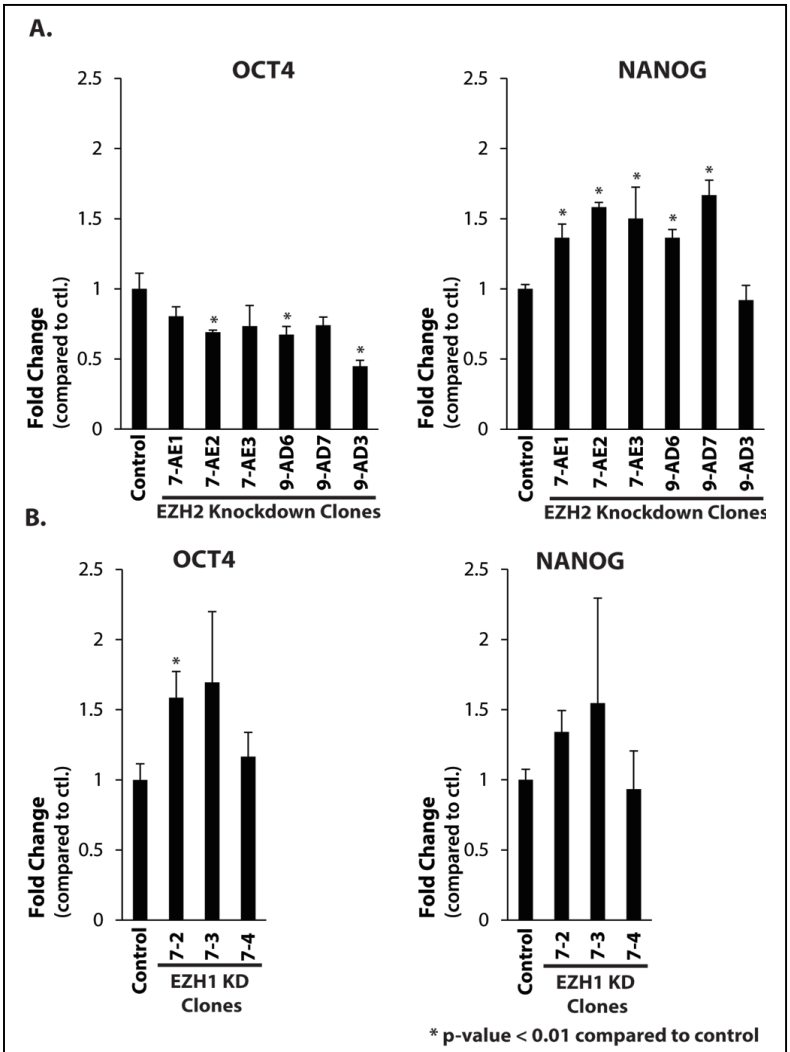


Figure 8: Pluripotency marker levels in knockdown clones. Quantification graphs of RT-qPCR assays examining fold-expression levels (compared to control after normalization to β-actin) of pluripotency gene markers, OCT4 and NANOG for **A.** EZH2 knockdown clones and **B.** EZH1 knockdown clones**.** *denotes p-value <0.01 compared to the control.

Knockdown of EZH2 and EZH1 result in up-regulation of different developmental genes

mRNA levels of various germ layer genes in both EZH1 and EZH2 knockdown clones were also tested via RT-qPCR to understand whether the clones were directing towards a specific germ layer with the loss of their respective gene (Figure 9A).

The ectoderm markers tested were *NESTIN* and *SOX1* (Figure 9B). NESTIN levels in the EZH2 clones 7-AE1, 9-AD3, 9-AD6, and 9-AD7 were statistically up-regulated compared to the control; however, the magnitude of the increase was small. In the EZH1 clones, *NESTIN* expression was only statistically up-regulated in clone 7-3 by about 3.5-fold. Therefore, in both EZH2 and EZH1 knockdown clones, the trend of *NESTIN* expressions remained unchanged. *SOX1* transcript levels in the EZH2 knockdown clones also stayed statistically unchanged compared to the control. In contrast, *SOX1* expression in the EZH1 knockdown clones was significantly and markedly up-regulated in all the clones (Figure 9B). Clones 7-2, 7-3, and 7-4 showed about 5-fold, 63-fold, and 14-fold up-regulation, respectively.

T/Brachyury and *SOX9* were the mesoderm markers tested in control, EZH2 knockdown, and EZH1 knockdown cells (Figure 9C). *T/Brachyury* levels in the EZH2 knockdown cells mainly remained similar to the control, with clones 7-AE2 and 7AE-3 statistically different from control. The EZH1 knockdown clones had statistically similar expression levels of *T/Brachyury* relative to control. Similarly, *SOX9* expression in the EZH2 knockdown clones was comparable to control. In contrast, all EZH1 knockdown clones had statistically significant up-regulation of *SOX9* compared to control, with approximately 2-fold increase in clones 7-2 and 7-4 and nearly an 8-fold increase in clone 7-3 (Figure 9C).

The endoderm markers tested in the knockdown clones were *CXCR4* and *GATA4* (Figure 9D). *CXCR4* transcript levels in the EZH2 knockdown clones were similar to the control, but *GATA4* expression significantly increased by about 27-fold in the clones compared to the control. Clone 7-AE2, which showed a near 50-fold increase in *GATA4* compared to control, was the only clone that did not show statistically significant change in expression due to high variance in qPCR values. Nonetheless, *GATA4* expression, in general, was increased by the knockdown of EZH2. Both *CXCR4* and *GATA4* levels in the EZH1 knockdown cells did not statistically differ from those of the control. Furthermore, *GATA4* transcript was not detected in clone 7-3.

The transcript levels, in both EZH2 and EZH1 knockdown cells, for many of the assayed germ layer markers remained unchanged compared to the control. However, with the knockdown of EZH2, the expression of *GATA4* increased by greater than 25-fold, suggesting that EZH2 is critical in the regulation of *GATA4* (Figure 9D). However, two *SOX* genes, *SOX1* and *SOX9*, were significantly up-regulated with the reduction of EZH1 (Figures 9B & C). This finding suggests that EZH1 directly regulates expression of both *SOX* genes independent of EZH2. While *SOX1* was characterized as an ectoderm marker and *SOX9* as a mesoderm marker, studies have shown that both *SOX* genes are also neural stem cell markers (50). Because EZH1 has increased expression in the brain, the function of EZH1 may involve regulation and maintenance of neural stem cells, which is why there is an increase in *SOX1* and *SOX9* expression in the absence of EZH1 (35, 51).

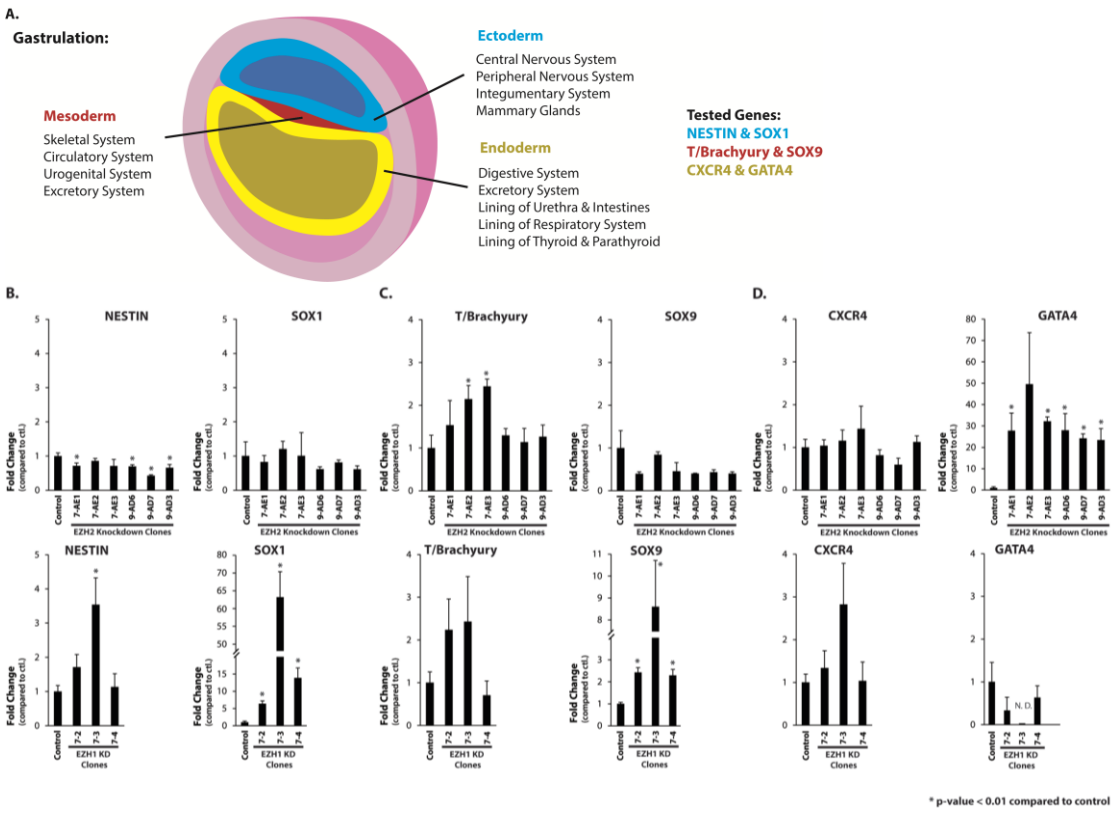


Figure 9: Germ layer marker levels in knockdown clones. A. Cartoon depiction of embryo in gastrulation, during which three germ layers—ectoderm, mesoderm, and endoderm—form. Listed are the organ systems originated from each germ layer. **B.** Quantification graphs of qPCR assays examining fold-expression levels (compared to control after normalization to *β-Actin*) of ectoderm gene markers, NESTIN and SOX1 in EZH2 knockdown clones. **C.** Quantification of qPCR assays for mesoderm gene markers—*T/Brachyury* and *SOX9*. **D.** Quantification of qPCR assays for endoderm gene markers—*CXCR4* and *GATA4*. N.D. indicates not detectable. *denotes p-value <0.01 compared to the control.

Knockdown of EZH2 results in global and gene-specific loss of H3K27m3

The global levels of histone H3 lysine 27 trimethylation (H3K27m3) were assayed by Western blot in the control and the EZH2 knockdown clones. In all the EZH2 knockdown clones, H3K27m3 was almost completely eliminated, to less than 10% of the control (Figure 10A). Next, the levels of H3K27m3 at specific gene loci were determined via the chromatin immunoprecipitation-quantitative PCR (ChIP-qPCR) assay

in wild-type, control, 7-AE1, 7-AE2, 7-AE3, and 9-AD7 (Figure 10B). The gene loci included a housekeeping gene (*β-ACT*), a pluripotency marker (*OCT4*), and four germ layer markers (*PAX6*, *SOX2*, *SOX1*, and *GATA4*). *B-ACT* and *OCT4* served as negative controls, as they should not be repressed and bound by H3K27m3. In general, H3K27m3 at all the assayed germ layer markers were significantly decreased compared to the wild-type and control lines. More specifically, the data for all three ectoderm markers tested, *PAX6*, *SOX2*, and *SOX1*, suggest that there is a loss of H3K27m3 (shown as the fraction of input) in all the EZH2 knockdown clones when compared to control. *SOX1* shows the greatest change among the ectoderm markers compared to control—about 15-fold decrease. *PAX6* shows about a 2-fold decrease, while *SOX2* shows about a 4-fold decrease. For *GATA4*, an endoderm marker, all clones also showed about a 25-fold decrease in H3K27m3 level. The H3K27m3 levels in the wild-type and control samples were similar at all the genes tested, suggesting that the control cells are truly serving as a control. These results from the ChIP-qPCR indicate that EZH2 knockdown significantly reduced PRC2 activities at all tested developmental genes.

The findings indicate that knockdown of EZH2 in hESCs caused a global loss of H3K27m3; however, some H3K27m3 still remains (Figure 10A). The residual H3K27m3 may be due to activities of EZH1 at specific gene loci (30). Therefore, characterization of these gene targets would aid in understanding differential roles between EZH1 and EZH2 in hESC pluripotency and differentiation. *GATA4* was the only germ layer marker assayed that showed a significant increase in transcript levels in the EZH2 knockdown cells (Figure 9D). Therefore, EZH1 does not trimethylate H3K27 in histones associated with these genes in the absence of EZH2. These findings suggest

alternative mechanism(s), i.e. epigenetic marks such as H3K9 methylation, may be involved in gene repression. Moreover, gene suppression of *GATA4* solely relies on EZH2 activities. A past study has shown that the PRC2 complex is attributed to directly methylate *GATA4* at lysine 299 (52). This methylation represses the transcriptional activity of *GATA4* by lowering the interaction of this gene locus with p300, a master regulator of developmental gene activation (52). *GATA4* and potential other genes as relying solely on EZH2 activities have strong implications because EZH2, not EZH1, is found to be recurrently mutated in a wide range of cancers (21-28).

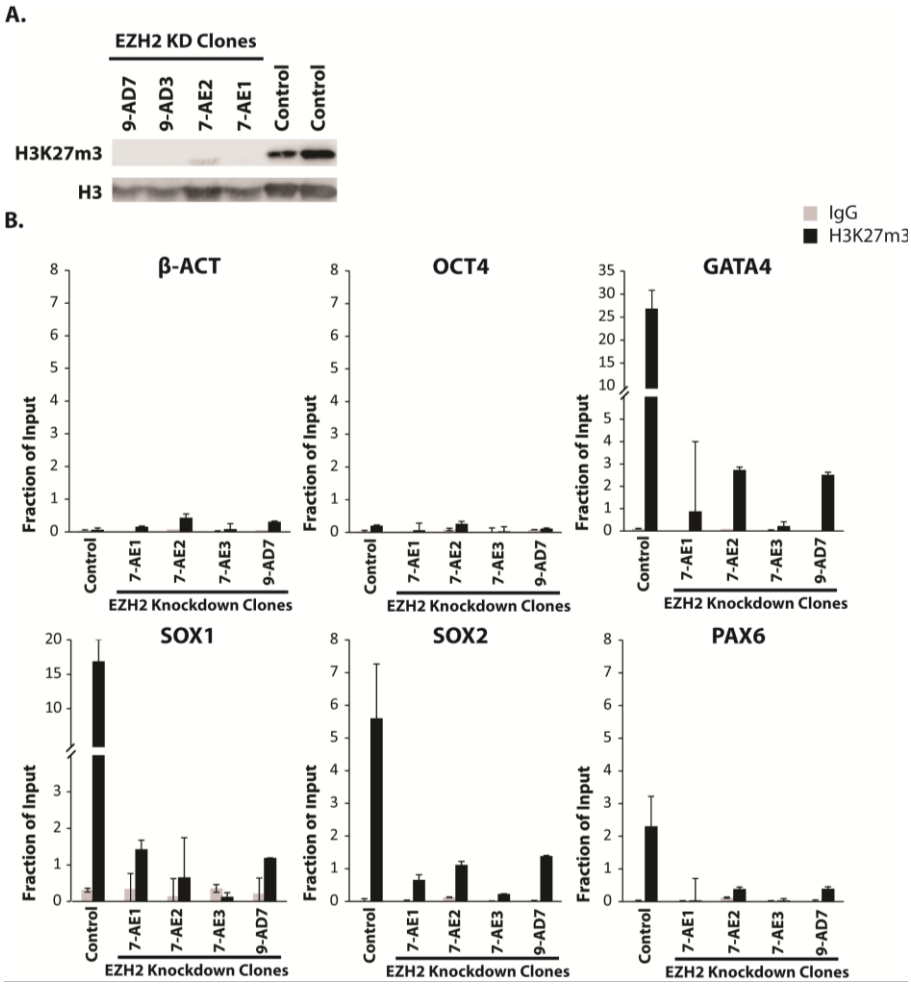


Figure 10: Global and gene-specific levels of H3K27m3. A. Western blot showing global levels of trimethylated lysine 27 of histone H3 (H3K27m3) and of histone H3 in knockdown and control cells. H3K27m3 levels in knockdown cells are drastically decreased in all EZH2 knockdown clones compared to control. **B.** ChIP-qPCR quantification graphs showing enrichment of H3K27m3 (in black, as fraction of input) in H9 wild-type, control, and EZH2 knockdown clone cells at various gene loci.

CONCLUSION

The CRISPR/Cas9 genome editing system was utilized to successfully introduce sequence specific mutations in hESCs at two different genes—EZH1 and EZH2 (Figures 6 & 7). While the mutations introduced in EZH2 were rather effective in causing protein knockdown, the case was not the same in the EZH1 clones. In fact, it was found that mutations introduced in one specific EZH1 clone were not ubiquitous among all cells of the clones. Additionally, some cells of the clone populations even retained a wild-type genotype. Therefore, the degree of protein knockdown caused by CRISPR/Cas9-mediated mutations is highly dependent on, not only the specific gene, but also the target sequence.

Using these cell lines, the effect on cell pluripotency due to gene knockdown was characterized. The individual knockdown of EZH1 or EZH2 in hESCs does not result in the loss of cell pluripotency, suggesting that the PRC2 complex may not be the only complex involved with hESC fate determination (Figure 8). However, the loss of EZH2 results in *GATA4* (endodermal marker) up-regulation, while the reduction of EZH1 results in *SOX1* (ectodermal marker) and *SOX9* (mesodermal marker) up-regulation (Figures 9B-D). Thus, EZH1 and EZH2 are responsible for regulating different developmental genes and may potentially regulate differentiation into different germ layers.

Lastly, the loss of EZH2 results in the loss of H3K27m3, not only globally, but also at very specific developmental genes (Figures 10A & B). This suggests that EZH1 does not compensate fully for the loss of EZH2, further corroborating the idea that EZH1 and EZH2 do not serve the same roles in hESCs despite being homologues.

In conclusion, separate EZH1 and EZH2 knockdown cells were created utilizing the CRISPR/Cas9 system. Contrary to what was hypothesized, the individual loss of EZH1 or EZH2 did not affect hESC pluripotency; however, the loss did result in up-regulation of different germ layer-specific genes. Therefore, both EZH1 and EZH2 are responsible for regulating different genes. Additionally, the loss of EZH2 showed a global and gene specific reduction of the silencing mark, H3K27m3, indicating that EZH1 did not fully compensate for the loss of EZH2. Even with this reduction in H3K27m3, many developmental genes remained silenced. Thus, the PRC2 complex may not be the only epigenetic regulator involved in the repression and activation of key genes during embryonic development.

REFERENCES

1. **Prochazkova M, Chavez M, Prochazka J, Felfy H, Mushegyan V, Klein O.** 2014. Chapter 18 - Embryonic Versus Adult Stem Cells. Stem Cell Biology and Tissue Engineering in Dental Sciences, p 249-262. *In* Vishwakarma A, Sharpe P, Shi S, Ramalingam M (ed), Stem Cell Biology and Tissue Engineering in Dental Sciences, 1st ed, Academic Press, London, UK.
2. **Karlic R, Chung H, Lasserre J, Vlahovicek K, Vingron M. 2010.** Histone modification levels are predictive for gene expression. Proc Natl Acad Sci U S A **107**:2926-2931.
3. **Wutz A.** 2013. Epigenetic regulation of stem cells: the role of chromatin in cell differentiation. Adv Exp Med Biol **786**:307-328.
4. **Goldthwaite C**. Are Stem Cells Involved in Cancer? U.S. Department of Health and Human Services. Chapter 9.
5. **Buechle D, Struhl G, Muller J**. 2001. Polycomb group proteins and heritable silencing of Drosophila Hox genes. Development **128**:993-1004.
6. **Lewis EB.** 1978. A gene complex controlling segmentation in Drosophila. Nature **276**:565-570.
7. **Struhl G.** 1981. A gene product required for correct initiation of segmental determination in Drosophila. Nature **293**:36-41.
8. **Duncan IM.** 1982. Polycomblike: a gene that appears to be required for the normal expression of the bithorax and antennapedia gene complexes of Drosophila melanogaster. Genetics **102**:49-70.

9. **Zink B, Paro R.** 1989. In vivo binding pattern of a trans-regulator of homoeotic genes in Drosophila melanogaster. Nature **337**:468-471.
10. **Levine SS, Weiss A, Erdjument-Bromage H, Shao Z, Tempst P, Kingston RE.** 2002. The core of the polycomb repressive complex is compositionally and functionally conserved in flies and humans. Mol Cell Biol **22**:6070-6078.
11. **Kerppola TK.** 2009. Polycomb group complexes—many combinations, many functions. Trends Cell Biol **19**:692-704.
12. **Schwartz YB, Kahn TG, Nix DA, Li XY, Bourgon R, Biggin M, Pirrotta V.** 2006. Genome-wide analysis of Polycomb targets in Drosophila melanogaster. Nat Genet **38**:700-705.
13. **Jürgens G.** 1985. A group of genes controlling the spatial expression of the bithorax complex in Drosophila. Nature **316**:153-155.
14. **Moazed D, O'Farrell PH.** 1992. Maintenance of the engrailed expression pattern by Polycomb group genes in Drosophila. Development **116**:805-810.
15. **Bracken AP, Helin K.** 2009. Polycomb group proteins: navigators of lineage pathways led astray in cancer. Nat Rev Cancer **9**:773-784.
16. **Gieni RS, Hendzel MJ.** 2009. Polycomb group protein gene silencing, non-coding RNA, stem cells, and cancer. Biochem Cell Biol **87**:711-746.
17. **Margueron R, Reinberg D.** (2011). The Polycomb complex PRC2 and its mark in life. Nature **469**:343-349.
18. **Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K.** 2006. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev **20**:1123-1136.

19. **Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D.** 2002. Histone methyltransferase activity associated with human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev **16**:2893-2905.
20. **Simon J, Kingston R.** 2009. Mechanisms of Polycomb gene silencing: knowns and unknowns. Nat Rev Mol Cell Biol **10**:697-708.
21. **Wang W, Qin J, Voruganti S, Nag S, Zhou J, Zhang R.** 2015. Polycomb Group (PcG) Proteins and Human Cancers: Multifaceted Functions and Therapeutic Implications. Med Res Rev **35**:1220-1267.
22. **Richly H, Aloia L, Di Croce L.** 2011. Roles of the Polycomb group proteins in stem cells and cancer. Cell Death Dis **2**:e204.
23. **Sparmann A, van Lohuizen M.** 2006. Polycomb silencers control cell fate, development and cancer. Nat Rev Cancer **6**:846-856.
24. **Jacobs JJ, van Lohuizen M.** 2002. Polycomb repression: from cellular memory to cellular proliferation and cancer. Biochim Biophys Acta **1602**:151-161.
25. **Gil J, Bernard D, Peters G.** 2005. Role of polycomb group proteins in stem cell self-renewal and cancer. DNA Cell Biol **24**:117-125.
26. **Valk-Lingbeek ME, Bruggeman SW, van Lohuizen M.** 2004. Stem cells and cancer; the polycomb connection. Cell **118**:409-418.
27. **Hormaeche I, Licht JD.** 2007. Chromatin modulation by oncogenic transcription factors: new complexity, new therapeutic targets. Cancer Cell **11**:475-478.
28. **Sauvageau M, Sauvageau G.** 2010. Polycomb group proteins: multi-faceted regulators of somatic stem cells and cancer. Cell Stem Cell **7**:299–313.

29. **Karin M, Cao Y, Greten FR, Li ZW.** 2002. NF-kappaB in cancer: from innocent bystander to major culprit. Nat Rev Cancer **2**:301-310.

30. **Shen X, Liu Y, Hsu Y, Fujiwara Y, Kim J, Mao X, Yuan G, Orkin S.** 2008. EZH1 Mediates Methylation on Histone H3 Lysine 27 and Complements EZH2 in Maintaining Stem Cell Identity and Executing Pluripotency. Mol Cell **32**:491-502.

31. **Zhang Z Schwartz S, Wagner L, Miller W.** 2000. A greedy algorithm for aligning DNA sequences. J Comput Biol **7**:203-214.

32. **Margueron R, Li G, Sarma K, Blais A, Zavadil J, Woodcock CL, Dynlacht BD, Reinberg D.** 2008. Ezh1 and Ezh2 Maintain Repressive Chromatin through Different Mechanisms. Mol Cell **32**:503-518.

33. **Xu J, Shao Z, Li D, Xie H, Kim W, Huang J, Taylor J, Pinello L, Glass K, Jaffe J, Yuan C, Orkin S.** 2015. Developmental Control of Polycomb Subunit Composition by GATA Factors Mediates a Switch to Non-Canonical Functions. Mol Cell **57**:304-316.

34. **Mousavi K, Zare H, Wang AH, Sartorelli T.** 2012. Polycomb protein Ezh2 promotes RNA polymerase II elongation. Mol Cell **45**:255-262.

35. **Laible G, Wolf A, Dorn R, Reuter G, Nislow C, Lebersorger A, Popkin D, Pillus L, Jenuwein T.** 1997. Mammalian homologues of the Polycomb-group gene Enhancer of zeste mediate gene silencing in Drosophila heterochromatin and at S.cerevisiae telomeres. EMBO J **16**:3219-3232.

36. **Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E.** 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science **337**:816-821.

37. **Wiedenheft B, Stemberg SH, Doudna JA.** 2012. RNA-guided genetic silencing systems in bacteria and archaea. Nature **482**:331-338.

38. **Fineran PC, Charpentier E.** 2012. Memory of viral infections by CRISPR-Cas adaptive immune systmes: acquisition of new information. Virology 434:202-209.

39. **Horvath P, Barrangou R.** 2012. CRISPR/Cas, the immune system of bacteria and archaea. Science **327**:167-170.

40. **Barrangaou R, Fremaux C, Deveau H, Richards M, Boyaval R, Moineau S, Romero D, Horbath P.** 2007. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. Science **315**:1709-1712.

41. **Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, Charpentier E.** 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature **471**:602-607.

42. **Reis A, Hornblower B, Robb B, Tzertzinis G.** 2014. CRISPR/Cas9 and Targeted Genome Editing: A New Era in Molecular Biology. NEB Expressions **1**:3-6.

43. **Sanjana NE, Shalem O, Zhang F.** 2014. Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods **11**:783-784

44. **Tiscornia G, Singer O, and Verman I.** 2006. Production and purification of lentiviral vectors. Nat Protoc **1**:241-245.

45. **Liang X, Potter J, Kumar S, Zou Y, Quintanilla R, Sridharan M, Carte J, Chen W, Roark N, Ranganathan S, Ravinder N, Chesnut JD.** 2015. Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. J Biotechnol **208**:44-53.

46. **Clancy S.** 2008. RNA splicing: introns, exons, and spliceosome. Nat Edu **1**:31.

47. **Goldberg AD, Allis CD, Bernstein E.** 2007. Epigenetics: a landscape takes shape. Cell **128**:635-638.

48. **Ding X, Wang X, Sontag S, Qin J, Wanek P, Lin Q, Zenke M.** 2014. The polycomb protein EZH2 impacts on induced pluripotent stem cell generation. Stem Cells Dev **23**:931-940.

49. **Hidalgo I, Herrera-Merchan A, Ligos JM, Carramolino L, Nuñez J, Martinez F, Dominquez O, Torres M, Gonzalez S.** 2012. Ezh1 Is Required for Hematopoietic Stem Cell Maintenance and Prevents Senescence-like Cell Cycle Arrest. Cell Stem Cell **11**:649-662.

50. **Alcock J, Lowe J, England T, Bath P, Sottile V.** 2009. Expression of Sox1, Sox2 and Sox9 is maintained in adult human cerebellar cortex. Neurosci Lett **450**:114-116.

51. **Henriquez B, Bustos FJ, Aguilar R, Becerra A, Simon F, Montecino M, van Zundert B.** 2013. Ezh1 and Ezh2 differentially regulate PSD-95 gene transcription in developing hippocampal neurons. Mol Cell Neurosci **57**:130-143.

52. **He A, Shen Z, Ma Q, Cao J, von Gise A, Zhou P, Wang G, Marquez VE, Orkin SH, Pu WT.** 2012. PRC2 directly methylates GATA4 and represses its transcriptional activity. Genes Dev **26**:37-42.

APPENDIX A: PRIMER SEQUENCES

Table A1: EZH1 gRNAs

Exon 6	Forward	Reverse
1	cacc GTTTCAGGTAGAAGATGAGA	aaac TCTCATCTTCTACCTGAAAC
2	cacc g AGAAGATGAGACTTTTATTG	aaac CAATAAAAGTCTCATCTTCTc
3	cacc GTCCATGGTGAAGAAGGTAG	aaac CTACCTTCTTCACCATGGAC
Exon 7		
1	cacc g CACTAATCAGAACGGATCCA	aaac TGGATCCGTTCTGATTAGTGc
2	cacc g ACTCAGATGAGGAGGAGGAA	aaac TTCCTCCTCCTCATCTGAGTc
3	cacc g CTATTGAAGGTACGTAGCAC	aaac GTGCTACGTACCTTCAATAGc

Table A2: EZH2 gRNAs

Exon 6	Forward	Reverse
A	cacc GATCTGGAGGATCACCGAGA	aaac TCTCGGTGATCCTCCAGATC
B	cacc GTGGAGTTGGTGAATGCCCT	aaac AGGGCATTCACCAACTCCAC
C	cacc gTACACAGAATCCTAATAATC	aaac GATTATTAGGATTCTGTGTAc
D	cacc gTTTGATGCGTTTCAGAATGT	aaac ACATTCTGAAACGCATCAAAc
E	cacc gCATCATCATTATATTGACCA	aaac TGGTCAATATAATGATGATGc
Exon 7		
A	cacc gATAAAGAAAGCCGCCCACCT	aaac AGGTGGGCGGCTTTCTTTATc
B	cacc gCAGAAGGAAATTTCCGAGGT	aaac ACCTCGGAAATTTCCTTCTGc
C	cacc gACAGGGCAAGATTGCCTCAA	aaac TTGAGGCAATCTTGCCCTGTc
D	cacc GCCCTTATCTGGAAACATTG	aaac CAATGTTTCCAGATAAGGGC
E	cacc GGGCACAGCAGAAGAACTAA	aaac TTAGTTCTTCTGCTGTGCCC
Exon 9		
A	cacc GTTCTTCCGCTTATAAGTGT	aaac ACACTTATAAGCGGAAGAAC
B	cacc gTAGTGCCTAATTATTGCAGA	aaac TCTGCAATAATTAGGCACTAc
C	cacc GCTCTAGACAACAAACCTTG	aaac CAAGGTTTGTTGTCTAGAGC
D	cacc gAAGTACCCTCTGCAATAATT	aaac AATTATTGCAGAGGGTACTTc
E	cacc gTGGTAACACTGTGGTCCACA	aaac TGTGGACCACAGTGTTACCAc

Table A3: Primer Sequences for PCR from Genomic DNA

EZH1	Forward	Reverse
Exon 6	TTCTGCAGTGTGTACCCCAT	CTTGGCCCCATTTGAGGTTT
Exon 7	CAAGAAGCACCAACAGCCTT	GCAATAAGGCGAAACTCCGT
EZH2		
Exon 7	CTTGCAAACACTGTAAACAAAGTGTAG	CTGTCAGATCCATTTTCTATTTGAGGG
Exon 9	TAGAGGGATATGGCTATGAGAGGG	TTAATGGTTGGAGGAGGAGGAATG

Table A4: RT-qPCR Primer Sequences

	Forward	Reverse
β-ACT	ACCATGGATGATGATATCGC	TCATTGTAGAAGGTGTGGTG
OCT4	GACAACAATGAGAACCTTCAGGAGA	CTGGCGCCGGTTACAGAACCA
NANOG	GCAATGGTGTGACGCAGAAGG	AGGTTCCCAGTCGGGTTCA
NESTIN	CTGCTACCCTTGAGACACCTG	GGGCTCTGATCTCTGCATCTAC
SOX1	CTGACACCAGACTTGGGTTTT	AAAGTGGGCTTCGCCTCTC
T	TATGAGCCTCGAATCCACATAGT	CCTCGTTCTGATAAGCAGTCAC

SOX9	GCTCTGGAGACTTCTGAACGA	CCGTTCTTCACCGACTTCCT
GATA4	GTGTCCCAGACGTTCTCAGTC	GGGAGACGCATAGCCTTGT
CXCR4	ACTACACCGAGGAAATGGGCT	CCCACAATGCCAGTTAAGAAGA

Table A5: ChIP-qPCR Primer Sequences

	Forward	**Reverse**
β-ACT	CTTGCCGACTTCAGAGCAAC	CCCAACACCACACTCTACCT
OCT4	AAGCTGCCCACCTAACTTCT	GCCAGGGTCTCTCTTTCTGT
GATA4	ACAGTTCCTCCCACGCATAT	TGATACATGGTCCCTGCGAG
SOX1	TCCGTCTGAATTCCTCTCCG	CAGGTCGGTCTCCATCATCA
SOX2	CCAAGATGCACAACTCGGAG	GCTTAGCCTCGTCGATGAAC
PAX6	ACCCACTAATCACTCCGCAA	AAACGGACCAATTGCACCAG

APPENDIX B: LENTIVIRUS TITER

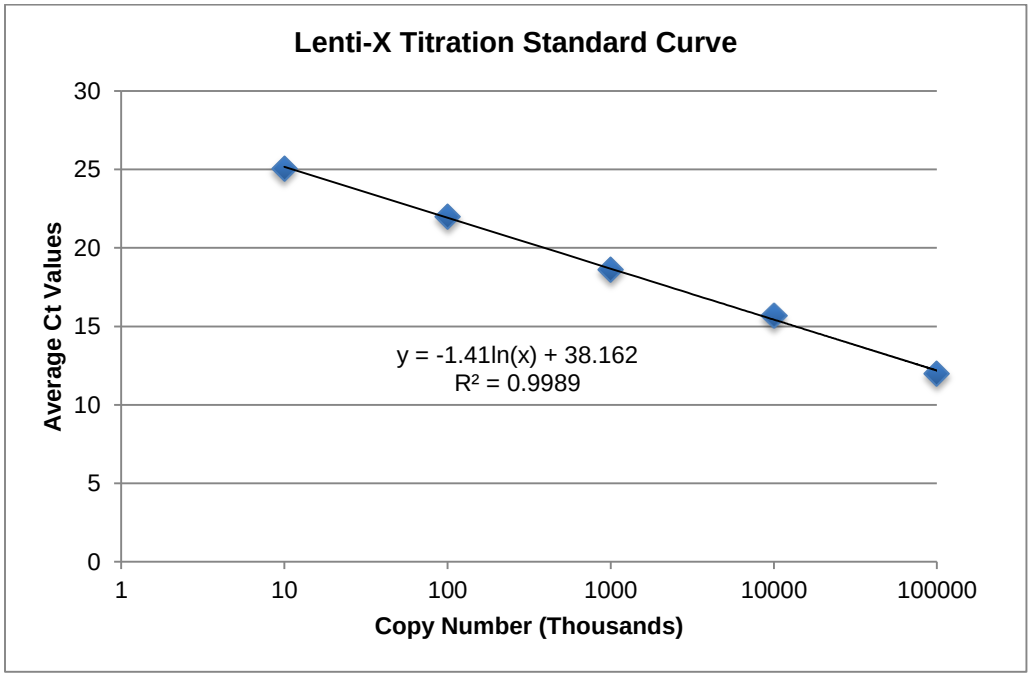


Figure B1: Lenti-X Titration Standard Curve. Standard curve generated to predict lentiviral copy numbers from qPCR C_T values.

Table B1: Titer qPCR C_T Values and Calculations

Sample	Copy Number from Formula	C_T Values	Average C_T	Copies/ml
EZH1 6.1 (1:20 rep 1)	2.44E+05	20.70	20.67	2.44E+05
EZH1 6.1 (1:20 rep 2)		20.64		
EZH1 6.2 (1:20 rep 1)	1.00E+05	22.21	21.92	1.00E+05
EZH1 6.2 (1:20 rep 2)		21.64		
EZH1 6.3 (1:20 rep 1)	1.55E+05	21.33	21.31	1.55E+05
EZH1 6.3 (1:20 rep 2)		21.30		
EZH1 7.1 (1:20 rep 1)	3.10E+05	20.28	20.33	3.10E+05
EZH1 7.1 (1:20 rep 2)		20.38		
EZH1 7.2 (1:20 rep 1)	1.49E+05	21.36	21.36	1.49E+05
EZH1 7.2 (1:20 rep 2)		21.37		
EZH1 7.3 (1:20 rep 1)	1.49E+05	21.59	21.37	1.49E+05
EZH1 7.3 (1:20 rep 2)		21.15		

APPENDIX C: WESTERN BLOT CALCULATIONS

Table C1: EZH2 Knockdown Western Blot Calculations

Clone	EZH2 Intensity	GAPDH Intensity	EZH2/GAPDH	Fraction of Control	% KD
Control	0.0626	0.253	0.247	1	0.0
7-AEI	0.000581	0.148	0.00392	0.0159	98.4
7-AE3	0.000154	0.152	0.00101	0.00409	99.6
9-AD3	0.000661	0.0589	0.0112	0.0454	95.5
6-ABC5	0.0117	0.0975	0.120	0.485	51.5
9-AD7	0.000351	0.139	0.00252	0.0102	99.0
7-AE2	0.00117	0.195	0.00600	0.0242	97.6
9-AD7	0.00186	0.211	0.00882	0.0356	96.4

Table C2: EZH1 Knockdown Western Blot Calculations

Clone	EZH1 Intensity	GAPDH Intensity	EZH1/GAPDH	Fraction of Control	% KD
Control	0.112	49.3	0.00227	1.00	0.0
7-2	0.0516	34.2	0.00151	0.66	33.6
7-3	0.0305	31.3	0.000974	0.43	57.1
7-4	0.0705	39.6	0.00178	0.78	21.6
67-11-1	0.171	42.6	0.00401	1.77	-76.7
67-11-2	0.153	40.4	0.00379	1.67	-66.7