

**Enhancing Direct Conversion of Mouse Fibroblasts to Induced
Hepatocytes**

by

Duygu Payzın

**A Thesis Submitted to the
Graduate School of Sciences and Engineering
in Partial Fulfillment of the Requirements for
the Degree of**

**Master of Science
in
Molecular Biology and Genetics**

Koç University

August 2015

Koç University
Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a master's thesis by

Duygu Payzın

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Committee Members:

Tamer Önder, Ph. D. (Advisor)

Cory D. Dunn, Ph. D.

Şerife Esra Erdal, Ph. D.

Date:

Abstract

Transdifferentiation is the conversion of a differentiated cell to another differentiated cell type by the expression of lineage-specific transcription factors. Compared to directed differentiation of pluripotent stem cells, it is deemed to happen faster and be potentially safer for future clinical applications. However, in many contexts, the fidelity and the efficiency of lineage conversion remains low, which suggest that there are a number of uncharacterized barriers to lineage conversion. We hypothesized that remodeling of the chromatin landscape during transdifferentiation may constitute one such barrier and modulating the activity of chromatin modifier complexes may increase the efficiency of lineage conversion. We have utilized the conversion of mouse fibroblasts to induced hepatocytes (iHeps) as a model system to test this hypothesis and to understand how chromatin modifications change during this process. Specifically, we expressed the two transcription factors *Hnf4 α* and *Foxa1* necessary for the conversion of fibroblasts, and treated the cells with small molecule inhibitors of various chromatin modifiers. We then tested both the efficiency and the quality of conversion at the gene expression level. We observed that removal of the activating H3K79 methylation mark through inhibition of DOT1L, and suppression of the repressive H3K27me3 mark through the inhibition of EZH2 are facilitators of this conversion. Inhibition of these two modifications led to the generation of induced hepatocytes that are more similar to mouse hepatocytes than untreated controls are in terms of hepatocyte marker expression. Furthermore, we found that chemical treatment of cells subsequent to transcription factor expression is more effective than pretreatment. However, none of the chemicals tested were able to fully replace the two transcription factors. These findings provide insight into the chromatin modifications that are important in establishing the identity of engineered hepatocytes.

Özet

Transdiferansiyasyon farklılaşmış bir hücrenin, başka bir hücre tipine ait transkripsiyon faktörlerinin ifadesi ile başka bir hücreye dönüşmesi olarak tanımlanır. İndüklenmiş pluripotent kök hücrelerin farklılaştırılmasına göre bu sürecin daha hızlı ve potansiyel olarak daha güvenli olduğu düşünülmektedir. Ancak birçok açıdan bu dönüşümün verimi düşük kalmaktadır. Bu da soy dönüşümünün henüz tam olarak karakterize edilememiş engelleri olduğunu düşündürmektedir. Biz bu çalışmada bu engellerden birinin kromatinin yeniden modellemesi olabileceği hipotezini kurduk. Bu hipotezi test etmek için fare fibroblastlarının indüklenmiş hepatositlere dönüşümünü ele aldık. Fare fibroblastlarının indüklenmiş hepatositlere dönüştürülmesi için gerekli olan *Hnf4a* ve *Foxa1* transkripsiyon faktörlerini fibroblastlarda ifade ettik ve hücreleri kromatin modifikatörlerinin küçük molekül inhibitörlerine maruz bıraktık. Sonrasında oluşan indüklenmiş hepatositlerin oluşma verimliliğini ve kalitesini gen ifadesi seviyesinde test ettik. Çalışmalarımız sonucunda H3K79 metiltransferaz DOT1L'in ve H3K27 metiltransferaz EZH2'nin inhibisyonunun indüklenmiş hepatosit oluşumunu desteklediğini gözlemledik. Bu iki modifikasyonun inhibisyonu indüklenmiş hepatositlerin normal fare hepatositlerine daha çok benzediğini bulduk. Ayrıca, hücrelerin transkripsiyon faktörlerinin ifadesinden önce yerine sonra küçük moleküllere maruz kalmasının daha etkili olduğunu gözlemledik. Ancak denediğimiz hiçbir küçük molekül bu iki transkripsiyon faktörünün yerine tamamen geçemedi. Bu çalışma ile fibroblasttan programlanmış fare hepatositlerinin kimliklerinin oluşturulmasında önemli olan kromatin modifikasyonlarını inceledik.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my advisor Dr. Tamer Önder for his continuous support, encouragement and guidance. I will always be grateful to him for giving me a second chance in my graduate studies, pushing me to aim higher, and reminding me of why I wanted to become a scientist. His never-ending curiosity for science always kept me interested in research, and helped me to become the researcher I am today.

I would like to thank Dr. Cory Dunn and Dr. Esra Erdal for giving the time and effort to be on my thesis committee and evaluate my work.

I would like to acknowledge the financial support of the Scientific and Technological Research Council of Turkey (TÜBİTAK) and Koç University during the course my graduate study.

I thank the members of the Önder Lab, Ayyub Ebrahimi, Burcu Özçimen, Deniz Uğurlu, Kerem Fidan, Can Aztekin, Eda Süer, Gülnihal Kavaklıoğlu and Gülben Gürhan for the fun and inspiring scientific environment. I am especially grateful to Ayyub for teaching me virtually every technique I know, Kerem for the entertaining notes on my desk lighting up my day, Burcu for always cleaning up after me, Deniz for the organization of lab outings, my cousin Can for the stimulating scientific discussions and far-fetched outlook on life, Eda for the wonderful trip to Sweden, Gülben for being the most cheerful person I know, and Gülnihal for always reminding me there is more to life than the lab. I love you all.

I would also like to thank our summer intern Ali Murat Gali and our undergraduate researcher Elif Kocak for helping me out with the experimental work, and teaching me how to teach in science.

Special thanks to two of my best friends, Nesli Ece Şen and Gülen Saçbüken for their valuable friendship, the good times and fun memories.

I am eternally grateful to Doğuş Doğru, who has always been there for me. Without him these past seven years would have been no more than ordinary. His continuous love and support has kept me sane throughout my masters and reminded me my achievements are nothing without having him there to share it with me.

Finally I thank my parents, Bahriye and Serdar, my sister Özgecan, my grandmothers Seviye and Muzaffer, and my aunts Pınar and Ruhsar for their never-ending patience, love and encouragement. I hope to make you proud in the end.

Table of Contents

Abstract.....	iii
Özet.....	iv
Acknowledgements.....	v
List of Tables	x
List of Figures.....	xi
Nomenclature.....	xii
Chapter 1	1
INTRODUCTION	1
1.1 Nuclear Transfer	1
1.2 Cell Fusion.....	6
1.3 Transcription Factor Induced Pluripotency	8
1.4 Direct Lineage Conversion.....	11
1.5 Induced hepatocytes	18
1.5.1 Two factor conversion by <i>Hnf4α</i> and <i>Foxa1</i>	18
1.5.2 Three factor conversion by <i>Hnf1α</i> , <i>Foxa3</i> and <i>Gata4</i>	19
1.5.3 HNF4α and FOXA transcription factors	20
1.5.4 Human induced hepatocytes.....	20
1.6 Epigenetics	21
1.6.1 Epigenetics in Stem Cells, Development and Reprogramming to Pluripotency	23
1.6.2 Epigenetics in Lineage Conversion	25
1.6.3 Polycomb Repressive Complex.....	26
1.6.4 Recruitment of and Transcriptional Repression by PRC2	26
1.6.5 EZH2 in pluripotency, differentiation and reprogramming	28
1.6.6 DOT1L	30
1.6.7 DOT1L in pluripotency, differentiation and reprogramming.....	30
1.7 Aim	31
Chapter 2	33
MATERIALS AND METHODS	33
2.1 Cell Culture	33
2.2 DNA Constructs	33
2.3 Small Molecules	34
2.4 Isolation of primary mouse embryonic fibroblasts.....	34
2.5 Generation of Ezh2 Knockout MEF and Dot1l Knockout TTF Cell Lines.....	35
2.6 Production of Viral Particles	35
2.7 Concentration of Viruses	36

2.8	Titerting of Viruses.....	36
2.9	Flow Cytometry.....	36
2.10	Generation of induced hepatocytes	36
2.11	Histone Extraction	37
2.12	Western Blotting.....	37
2.13	Gene Expression Analysis by Quantitative PCR (qPCR)	38
2.14	Periodic Acid-Schiff Staining.....	41
2.15	Immunofluorescence	41
2.16	Immunostaining.....	42
2.17	Chromatin Immunoprecipitation (ChIP)	42
Chapter 3		45
RESULTS		45
3.1	Primary screen identifies Ezh2 and Dot1L inhibitors as hits	45
3.2	Post-infection chemical treatment is slightly more effective than pre-treatment.....	47
3.3	EZH2 and DOT1L inhibitors cannot replace the two transcription factors	49
3.4	Investigating EZH2's effect on fibroblast to hepatocyte conversion	51
3.4.1	DZNep treatment reduces H3K27 trimethylation	51
3.4.2	<i>Ezh2</i> knockdown verifies chemical treatment data	51
3.4.3	<i>Ezh2</i> knockout verifies chemical treatment data	53
3.4.4	Ezh2 overexpression reverses the knockout effect on fibroblast to hepatocyte conversion.....	54
3.4.5	H3K27 residues are less tri-methylated on hepatocyte-specific gene promoters in iHeps.....	59
3.5	Investigating DOT1L's effect on fibroblast to hepatocyte conversion	60
3.5.1	EPZ4777 treatment reduces H3K79 di-methylation	60
3.5.2	<i>Dot1l</i> knockout verifies chemical treatment data	61
3.5.3	DOT1L inhibition does not reduce the expression of fibroblast-specific genes63	
3.5.4	DOT1L inhibition does not reduce the aberrant other lineage-specific gene expression	65
3.5.5	<i>Dot1l</i> knockout does not dramatically affect the expression of EMT Genes	66
Chapter 4		67
DISCUSSION		67
4.1	Possible mechanisms of Ezh2 inhibition on transdifferentiation	67
4.2	Possible mechanisms of Dot1L inhibition on transdifferentiation	69
4.3	Testing Chromatin Modifiers for the Regulation of Cell Conversion in Other Contexts	70

4.4 Future Directions	71
Appendix A: Verification Experiments for Induced Hepatocytes	73
Appendix B: Human Induced Hepatocytes	76
Appendix C: Induced Macrophages	77
Appendix D: Supplemental Methods for Induced Macrophages.....	80
Bibliography	82
Vita	91

List of Tables

Table 1.1 Historically important developments on lineage conversion	16
Table 1.2 Well studied modifications of histones	22
Table 2.1 List of Small Molecules Used	34
Table 2.2 List of all antibodies used for Western Blotting.....	38
Table 2.3 List of primers used for qPCR.....	40
Table 2.4 List of antibodies used for immunofluorescence microscopy	41
Table 2.5 List of antibodies used in chromatin immunoprecipitation experiments.....	43
Table D.1 List of qPCR primers used in induced macrophage experiments.....	81

List of Figures

Figure 1.1 Three approaches to nuclear reprogramming	2
Figure 1.2 Polycomb Repressive Complexes	27
Figure 3.1 Primary screen for small molecule inhibitors	46
Figure 3.2 Verification experiments for small molecules	47
Figure 3.3 Pre-infection vs. post-infection treatment of cells with small molecules.....	48
Figure 3.4 EPZ4777 and DZNep cannot replace the two transcription factors.....	50
Figure 3.5 DZNep treatment reduces global H3K27 trimethylation levels.....	51
Figure 3.6 iHeps derived from <i>Ezh2</i> knock down MEFs show increased expression of mature hepatocyte and epithelial markers.	52
Figure 3.7 iHeps derived from <i>Ezh2</i> knockout MEFs show increased expression of mature hepatocyte and epithelial markers.	54
Figure 3.8 <i>Ezh2</i> overexpression reverses the effect of <i>Ezh2</i> knockout in transdifferentiation.....	57
Figure 3.9 <i>Ezh2</i> knockout MEFs generates more E-cadherin positive iHep colonies.....	58
Figure 3.10 H3K27 tri-methylation is reduced on hepatocytes specific gene promoters in iHeps.	60
Figure 3.11 EPZ4777 reduces global H3K79me2 levels.....	61
Figure 3.12 iHeps derived from <i>Dot1l</i> knockout TTFs more robustly express hepatocyte markers.	62
Figure 3.13 Fibroblast specific genes are expressed in similar levels in DOT1L inhibited and control iHeps.....	64
Figure 3.14 DOT1L inhibition does not reduce the expression of the intestinal marker <i>Cdx2</i>	65
Figure 3.15 <i>Dot1l</i> knockout and wild type iHeps and have similar EMT gene expression levels.....	66
Figure 4.1 Model of <i>Ezh2</i> 's role in conversion of mouse fibroblasts to iHeps.....	68
Figure A.1 Bright field microscopy images of fibroblasts and induced hepatocytes.	73
Figure A.2 Induced hepatocytes express E-cadherin at the protein level.....	74
Figure A.3 Induced hepatocytes express albumin, a key hepatocyte marker.....	74
Figure A.4 Periodic acid-Schiff staining of MEFs and D14 iHeps.	75
Figure B.1 dH1f lineage conversion using mouse transcription factors.....	76
Figure C.1 DOT1L inhibition slightly increases the efficiency of fibroblast to macrophage differentiation.	79

Nomenclature

4-OHT	4-Hydroxytamoxifen
BCA	Bicinchoninic acid assay
ac	acetylation
cDNA	Complementary DNA
ChIP	Chromatin immunoprecipitation
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dpc	Days post coitum
E	Embryonic day
EGF	Epidermal growth factor
EMT	Epithelial-mesenchymal transition
ES cell	Embryonic stem cell
ESC	Embryonic stem cells
EtOH	Ethanol
EV	Empty vector
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
GFP	Green fluorescent protein
GRN	Gene regulatory network
HF	Hnf4 α and Foxa1
HGF	Hepatocyte growth factor
HRP	Horseradish peroxidase
iHep	Induced hepatocytes
iMac	Induced macrophages
iPSC	Induced pluripotent stem cells

iPSC	Induced pluripotent stem cell
LDL	Low-density lipoprotein
me	Methylation
MEF	Mouse embryonic fibroblast
MET	Mesenchymal–epithelial transition
mRNA	Messenger RNA
P	Passage
PAS	Periodic acid–Schiff
PBS	Phosphate buffered saline
PEG	Polyethylene glycol
qPCR	quantitative (real-time) polymerase chain reaction
RNA	Ribonucleic acid
rpm	Revolutions per minute
RT	Room temperature

Chapter 1

INTRODUCTION

Traditionally, development was thought to be the irreversible process of an embryo dividing to generate all of the cells of an adult organism. Initially, this single cell embryo has the potential to differentiate into all of the lineages in the adult, but gradually loses this potential. The idea that specialized somatic cells are not irreversibly differentiated was first shown by nuclear transfer experiments. The generation of a fertile adult frog by transferring differentiated nuclei of a somatic cell into enucleated frog oocytes led to the initial principles of nuclear reprogramming [1-3]. The plasticity of cell fate was further demonstrated by the expression of four factors that allowed reprogramming of terminally differentiated cells to a pluripotent state [4]. More recently, direct conversion of lineage as a new technique of reprogramming has received attention. In the introduction, I will discuss the historical development of the most commonly used three approaches of reprogramming, which are nuclear transfer, cell fusion, and transcription factor transduction. These approaches are illustrated in Figure 1.1.

1.1 Nuclear Transfer

Nuclear transfer, or cloning, is the transfer of a nucleus from a specialized somatic cell to an enucleated oocyte. This procedure results in the formation of a whole organism, which has identical genetic material to the nucleus donor cell. Nuclear transfer experiments are especially significant in the context of reprogramming, because they have revealed that genes necessary to generate an organism are preserved in specialized cells, and these genes can be expressed through exposure to different conditions. This possibility indicates that cell specialization can be reversed.

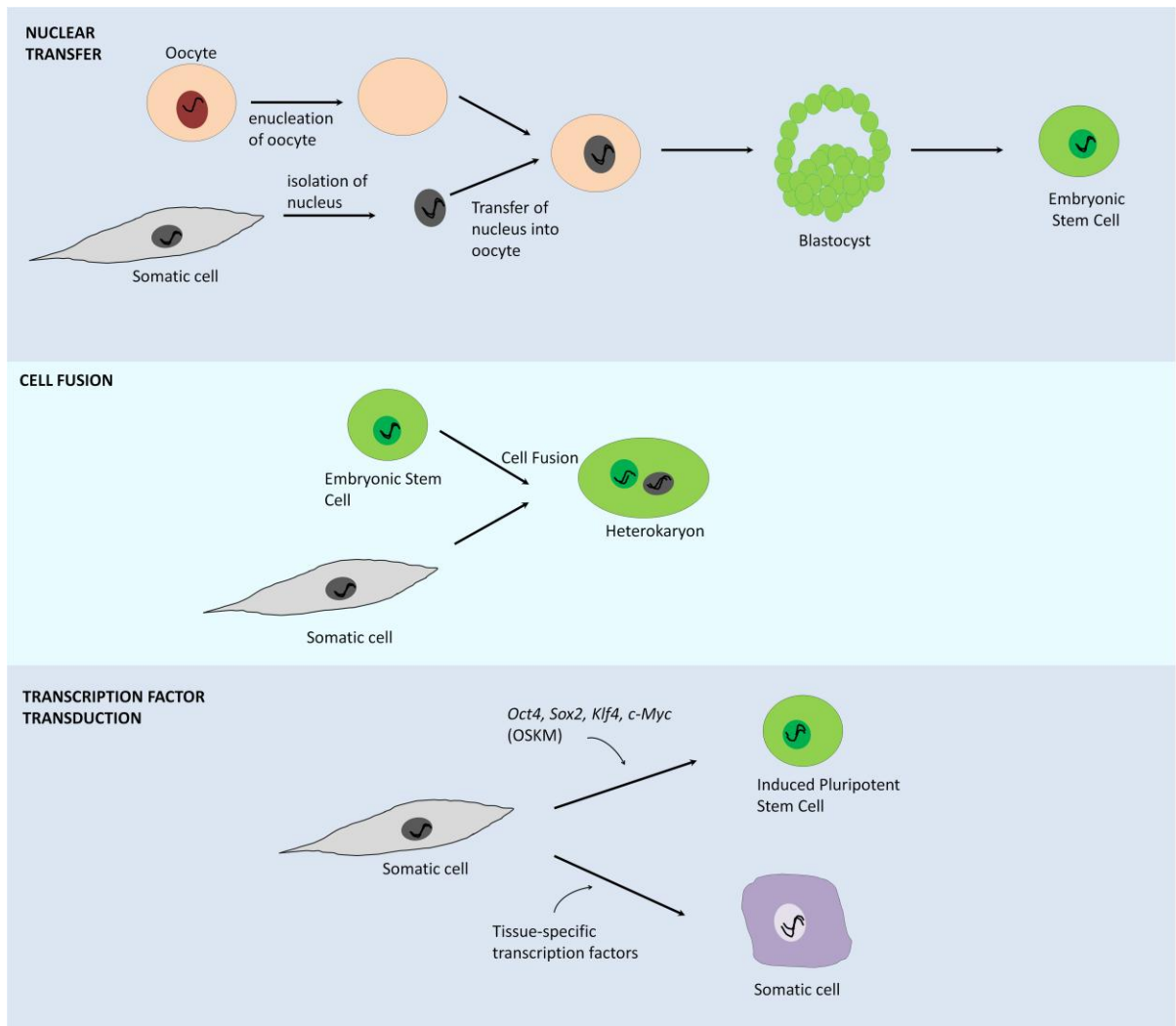


Figure 1.1 Three approaches to nuclear reprogramming

The three common approaches to reprogramming are nuclear transfer (upper panel), cell fusion (middle panel) and transduction of transcription factors (lower panel). Figure adapted from reference [5].

Many of the earliest studies in nuclear transfer were conducted in frog species, especially the *Xenopus*. The reasons behind this choice are because their eggs are larger compared to other species, and can be manipulated easier [6]. Moreover, *Xenopus* frogs become

sexually mature in only a year, produce eggs year-round, and they are easy to maintain in the laboratory [7].

One of the first experiments where the developmental potential of cells was investigated was conducted in 1928 in the salamander. Upon pinching off a single cell from a salamander embryo at the 16-cell stage, it was shown that both the pinched off cell and the remaining cells generated a complete organism in the same egg [8]. With this experiment, cells were shown to be totipotent at least during the first four divisions. However, the results did not eliminate the possibility that cells could undergo genetic changes following this stage, when they become specialized.

The definitive test of preservation of the genome during development was first carried out in 1952, where the nucleus of a live cell was transplanted successfully to an egg. In this experiment, oocytes of the frog *Rana pipiens* were enucleated, and nuclei derived from blastocysts were transplanted to these oocytes [9], which resulted in swimming tadpoles. One problem with this technique was that it was limited to the plasticity of the cell from which the nuclei were derived. When nuclei from more differentiated tissues were used, they resulted in abnormal tadpoles [10]. This led to the notion that cells lose their plasticity irreversibly as they differentiate and become specialized.

These initial conclusions were later challenged by experiments in the model organism *Xenopus laevis*. Using this species, differentiated epithelial cells from the frog intestine were successfully transferred into UV irradiated oocytes, which generated adult frogs [11, 12]. The adult frogs were healthy and fertile. Thus, the species used in nuclear transfer was shown to be important. It was reported through these findings that differentiation did not cause stable changes in the genome. Interestingly, these conclusions were found to be controversial. The first concern was that the efficiency was only a mere 1% when differentiated intestinal cells were used, compared to the efficiency of 30% upon use of embryonic cells as nuclear donors, a significant difference suggestive of stem cell

contamination [5]. Furthermore, the limitation of the technique to this species was an important problem.

In later years, nuclei obtained from the blood cells of the adult *R. pipiens* were used for nuclear transfer to generate tadpoles [13]. This finding confirmed the previous one with the *Xenopus*, indicating the preservation of the genome. There were only a few exceptions to this principle, such as cells that produce antibodies and undergo genome modifications to generate different proteins.

It was thirty years later that nuclear transfer was achieved in mammals. In 1997, the famous Dolly the sheep was cloned through the transfer of a mammary cell to an enucleated oocyte [14]. The reason for the long gap was technical difficulties, which were overcome by using nuclei from cells that had exited the cell cycle by serum starvation. A year later, mice were cloned using a similar technique [15]. The controversy of a possible stem cell contamination was finally resolved in 2002, by cloning experiments using mouse B cells or T cells as the donor [16]. The gene rearrangement in the immunoglobulin locus in B cells, and T cell receptor rearrangements allowed determining the nuclear origin of the generated organism. All of the tissues of the generated adult mice contained one of these arrangements.

Using somatic cells in nuclear transfer (SCNT) is not restricted to oocytes, it can also be used in zygotes [17]. This possibility implies that the required factors are still present in the fertilized egg. In nuclear transfer, it is also important to note that the entire nuclei are not removed from the eggs, rather, only the chromosomes and the spindle are removed at the MII stage [18]. This way, the cell is at a stage where the nuclear membrane is broken down, and the required factors that are normally located in the nucleus are in the cytosol.

Since the cloning of sheep and mice, a range of other animals such as dogs, goats and wolves were cloned by somatic cell nuclear transfer [19]. So far, primates have not yet

been cloned by SCNT. Importantly, nuclei derived from 16 year old frozen tissues were used for nuclear transfer [20]. Such advances raise the possibility to use frozen cells for therapeutic purposes.

One obstacle in nuclear cloning is the low efficiency of the conversion process. However, embryonic stem cells derived from nuclear transferred blastocysts (ntES cells), can be obtained at significantly higher efficiencies. For example in mice, the very low efficiency of 1% increases to 20% when ES cells are derived from the blastocysts [21]. This advantage has been used in the derivation of nuclear-transferred ES cells from non-human primates to show proof of principle that nuclear cloning can be achieved in primates [22]. Recently, nuclei obtained from human skin cells were transplanted to oocytes without enucleation [23]. This led to the generation of triploid human blastocysts. Embryonic stem cells were derived from these blastocysts, and they were similar to normal embryonic stem cells. This suggests that in principle, reprogramming of human somatic cells may be possible until the blastocyst stage.

A major problem in cloning by nuclear transfer is the severe defects in some of the produced organisms. Mostly studied in mice, some of these defects are aberrant gene expression, elongated telomeres, obesity, weak immune systems, inclination to cancer, and short lifespan [19]. These issues limit the efficiency of nuclear transfer to 1-2%.

The use of nuclear-transfer embryonic stem (ntES) cells as nuclei donors increases the efficiency to 20% [21]. Interestingly, the embryonic stem cell state significantly improves nuclear transfer. This phenomenon hints at the possibility of an epigenetic memory in the donor cell that hinders reprogramming [24]. Epigenetic factors that regulate the activity of genes seem to be remaining in the donor nuclei, and may be influencing the fidelity of reprogramming. This calls for deeper understanding and appreciation of mechanisms involved in gene regulation and epigenetic memory. Overcoming these obstacles will potentially make way for the use of ntES cells for clinical purposes.

1.2 Cell Fusion

Fusion of a cell with another cell can result in either hybrids or heterokaryons. Hybrids form when the nuclei of the initial cells fuse together. They proliferate, whereas heterokaryons do not. Instead, heterokaryons contain multiple nuclei as separate entities, and are short-lived. In cell fusion experiments, heterokaryons were mostly preferred to hybrids, because hybrid cells, since they are dividing, constantly undergo chromosomal rearrangement and aneuploidy when different species are used for the fusion [5]. Heterokaryons, on the other hand, because the nuclei do not fuse, circumvent this problem.

Initial fusion experiments demonstrated that when a mouse fibroblast was fused with a rat hepatocyte, a hepatocyte-specific enzyme, tyrosine aminotransferase could no longer be produced [25]. The same was true for melanin production when mouse fibroblasts were fused with hamster melanocytes [26]. These experiments indicated that gene expression is regulated by *trans*-acting repressors, in addition to *cis*-acting DNA elements. In later experiments, fusion of a normal cell with a tumor cell resulted in the formation of a cell with a normal state, where tumor formation was inhibited [27]. This can be attributed to the activity of a tumor suppressor originally expressed from the non-tumor nucleus, which may have been lost in the tumor cell.

Some studies also revealed that upon fusion, genes that were previously silent in the original cells were activated. These were definitively demonstrated by experiments using heterokaryons originating from cells of different species. Use of different species allowed the identification of the origin of a gene product, and determined which nuclear program was dominant over the other. Such experiments also challenged the idea that the differentiated state of specialized cells is irreversible. For example, in experiments where mouse muscle cells were fused with human amniotic cells to generate heterokaryons, a number of human muscle proteins were synthesized [28]. In this experiment, mouse muscle cells were specifically selected because they are multinucleate and already post-

mitotic, aiding the formation of a heterokaryon instead of a hybrid. Human amniotic cells were chosen because they were embryonic, hence possibly more plastic. Interestingly, DNA replication was not necessary for gene activation [28]. Overall, this study, through the expression of muscle genes in non-muscle cells, demonstrated that previously silent genes can be activated upon exposure to activating factors.

Similar experiments were also performed by fusing mouse muscle cells with human fibroblasts, hepatocytes and keratinocytes that arise from the mesoderm, endoderm and ectoderm, respectively. These experiments showed that silent genes could be activated in the original cell, which might have originated from any of the three germ layers [28, 29]. In the same system, the relative dosage of the unidentified factors was found to determine the dominant cell type, meaning which genes would be activated or repressed [30]. Overall, cell fusion experiments for somatic cells indicated to the plasticity of cell fate.

The fusion of pluripotent cells with differentiated cells allowed reprogramming of specialized cells to pluripotency. Upon fusion of a female mouse embryonic germ cell with a differentiated thymocyte derived from a male mouse, a tetraploid pluripotent cell was generated [31]. Analysis of the methylation status of the genes by restriction digestion with a methylation-sensitive enzyme revealed that a number of imprinted and non-imprinted genes from the thymocyte were demethylated. Similarly, fusion of an embryonic stem cell from a male mouse with a thymocyte from a female mouse, where the thymocyte expressed a reporter under the control of the pluripotency gene *Oct4*, led to the activation of the reporter [32]. In the same system, genes on the inactive X chromosome of the thymocyte were found to be re-activated. Dissimilar to fusion with germ cells, this fusion did not result in the demethylation of the imprinted genes.

In later experiments, different mouse species were used to trace the origin of DNA and RNA in the fused cells. In one study, ES cells derived from *Mus musculus domesticus* were fused with thymocytes derived from *Mus musculus molossinus* [33]. In the resulting

hybrid, the somatic cell genome became hyperacetylated at histone H3 and H4, and hyper-di- and tri-methylated at histone H3 lysine 4 (H3K4), which are epigenetic marks of a pluripotent state.

Heterokaryons of different species offer a more detailed study of reprogramming than heterokaryons of same species, because changes in gene expression can be determined for each species separately. This advantage led to a study where human B cells were fused with mouse embryonic stem cells [34]. The study revealed that in a single day, human *OCT4* and *NANOG* gene promoters were demethylated, and the genes were activated.

Another important use of heterokaryons was to better understand the mechanisms that are essential to initiate reprogramming to pluripotency. In these experiments, gain- and loss-of function techniques were used to identify factors that are necessary for reprogramming. Such a technique was used in a study where *Oct4* deficient mouse embryonic stem cells were fused with human B cells to a heterokaryon [34]. In this case, B cells were not programmed to pluripotency, highlighting the requirement of OCT4. Similarly, through heterokaryon experiments with mouse embryonic stem cells, the enzyme activation-induced cytidine deaminase (AID) was shown to be required for DNA demethylation to reprogram human fibroblasts to pluripotency [35]. These studies demonstrate how cell fusion can be used to gain insight into the obstacles of somatic cell reprogramming.

1.3 Transcription Factor Induced Pluripotency

Even though it was shown through cell fusion experiments that embryonic stem cells and embryonic germ cells possess the factors that are necessary to reprogram differentiated cells to pluripotency, the identification of the master regulators of embryonic stem cells was a challenge. It was initially thought that reprogramming was so complex that it required about a hundred factors [5]. In 2006, it was shown that the ectopic expression of four transcription factors can reprogram mouse fibroblasts to pluripotency [4]. A pool of 24 genes which were known to be expressed in ES cells were delivered in retroviral

vectors to fibroblasts, and cells were tested for their expression of a reporter gene under the control of *Fbx15*, a gene specific to ES cells [36]. Removal of factors from the 24 factor pool one-by-one revealed that only four factors, namely OCT4, SOX2, KLF4 and c-MYC, were sufficient to induce pluripotency. These cells were termed induced pluripotent stem cells (iPSCs). They were remarkable because they were capable of forming teratomas upon injection into mice, which contain cells of all three embryonic lineages.

Initially, there was much controversy on the validity of the findings, because these cells were unable to generate adult chimeric mice when injected to a mouse blastocyst. However, it was later shown that these cells can indeed generate chimeras [37], and even contribute to germ cells [38].

An important aspect of the reprogrammed cells was that the exogenous retroviral genes were silenced upon reprogramming [4]. Therefore, the maintenance of the pluripotent cells was solely dependent on the successfully activated endogenous pluripotency gene network. Only a year later, human fibroblasts were also reprogrammed to iPSCs, using the same four factors [39].

Since the generation of iPSCs by the use of four factors, technical developments in iPSC generation allowed substitution of the four factors with other factors, or even small molecules. The idea that a set of factors are required, however, remains the same. The originating cell type has also been changed to a variety of different cells, such as beta-cells, keratinocytes and hepatocytes [40]. However, the efficiency of reprogramming is different between these cell types, which indicates that changes in cell fate are dependent on the context. The efficiency is significantly higher in the reprogramming of stem or progenitor cells than in differentiated cells, presumably because they are more plastic.

Although it has almost been ten years since fibroblasts have been converted to induced pluripotent stem cells, the mechanism through which reprogramming occurs is not yet fully known. Reprogramming appears to be a stochastic event that happens at low efficiencies even though the cell successfully expresses the four factors ectopically. This issue was best demonstrated in an elegant experiment where iPS cells generated by factors expressed from doxycycline inducible vectors were used to produce a transgenic mouse [41]. From this mouse, fibroblasts were isolated and factor expression was once again induced by doxycycline. In this second reprogramming, the efficiency increased about 100 fold compared to the first reprogramming event. Still, considering that these cells once had the properties to be successfully reprogrammed once, the efficiency was much too low. This hints at a random, or at least yet unknown nature of reprogramming [42].

The three approaches to reprogramming, namely nuclear transfer, cell fusion, and transcription-factor transduction, all reveal an important aspect of altering cell fate: There is a balance of regulators that promote and maintain the cell toward a specific lineage, and if this balance is tipped towards another direction by exposure to a massive amount of pluripotency factors, cells can change their epigenome in a significant way, and reprogram to a pluripotent state.

As revealed by these reprogramming experiments, reprogramming to pluripotency involves the demethylation of promoters of pluripotency genes, and the activation of embryonic stem cell-specific genes. The exact molecular mechanisms of how the initial transcriptional and epigenetic memory is erased and how a new epigenetic landscape is established are still not fully understood. Revealing these inner mechanisms will aid in the use of these cells as therapeutic agents.

1.4 Direct Lineage Conversion

Nuclear reprogramming is not restricted to the de-differentiation of specialized cells. This term simply refers to the conversion of a cell into a different cell type. Reprogramming cells without de-differentiation is termed lineage conversion; the direct conversion of a differentiated, specialized cell into another specialized cell.

The possibility of experimentally manipulating cells to change lineages raises the question whether these also occur naturally. One rare example of natural transdifferentiation is the regeneration of the eye lens in newts and salamanders [43, 44]. In this event, termed Wolffian regeneration, once the lens of the eye is removed, cells found in the dorsal iris of the eye begin proliferation. They first lose their pigmentation and then differentiate to generate a new lens. The original cells of the dorsal iris are epithelial cells and are pigmented, whereas the newly formed lens is made up of transparent keratinocytes.

Physiological lineage conversion is also found in mouse. During embryonic development, endothelial cells are capable of generating blood cells. This was shown in a long-term single cell tracing experiment of mesodermal cells that are responsible for the formation of endothelial and blood cells [45]. A small number of endothelial cells lost their tight junctions, became round, and expressed erythroid and monocytic markers.

Lineage conversion is also considered to take a role in pathology, specifically in epithelial-mesenchymal transition (EMT) during tumor formation [46]. This was shown for the EMT events that occur in breast cancers. As in normal epithelial-mesenchymal transitions, transcription factors associated with EMT, such as *SNAI1*, *SNAI2* and *TWIST* were shown to be activated [46, 47]. It is possible that many other such cell conversions exist in nature, and the advances in lineage tracing techniques will allow their discovery in the near future.

In *Drosophila melanogaster*, a phenomenon termed transdetermination occurs in the imaginal discs [48]. Imaginal discs are structures in the larvae that are destined to become appendages in the adult fly. If these imaginal discs are dissociated and serially transplanted to different sites within the larvae, they switch their fate. A cell that was originally programmed to form the leg, when transplanted to another position, can form a wing instead. Although not fully understood then, with the current knowledge today, it is known that the disk cells alter their fate when exposed to different environmental cues, which is achieved through the expression of transcription factors.

Transdetermination in *Drosophila* can also be achieved simply by the ectopic expression of the transcription factors specific to an imaginal disk. For example, the exogenous expression of the gene *eyeless* resulted in the formation of eyes on the legs, wings and antennae of the fruit fly [49]. In another study, the homeotic gene *Antennapedia* controlled by a promoter of a heat-shock gene was overexpressed in larvae. The result was a striking change in the body of the larvae, with legs growing in place of the antennae [50]. These initial experiments were not able to determine whether the ectopic expression of factors can only induce lineage conversion in progenitor cells, or can also reprogram differentiated cells. Still, they were significant because they demonstrated the plasticity of the differentiated state.

The first reported case of direct lineage conversion in mouse is the conversion of fibroblasts to muscle, adipocytes and chondrocytes by treating the cells with 5-AZA-cytidine (AzaC) [51]. Shortly after this finding, it was found that AzaC conveyed its effect by inhibiting DNA methylation [52]. It was later identified that fibroblasts can be directly converted into myoblast like cells by 5-AZA-cytidine or the expression of transcription factor *MyoD* [53, 54].

The blood system also presented evidence for lineage conversion. Exogenous expression of GATA1, a transcription factor associated with erythroids and megakaryocytes, led to

the expression of erythroid-megakaryocyte markers in monocytes [55, 56]. At the same time, this reduced the expression of monocytic markers. This conversion could also happen in reverse. The expression of the transcription factor PU.1 in an erythroid-megakaryocytic cell line converted the cells to monocytes, repressing GATA1 [57]. This was an important finding that showed transcription factors not only promote a specific cell fate, but also repress alternative cell fates.

One concern with the conclusions drawn from these experiments was that they did not address whether lineage switches can also be induced in mature hematopoietic cells. Later experiments in the blood system also revealed that fully specialized cells can also change lineages, as demonstrated by the conversion of mature B-cells into macrophages through the expression of C/EBP α [58]. The efficiency of conversion was close to 100% upon expression of the same transcription factor in committed B-cell [58] and T-cell progenitors [59]. In a similar logic, loss of C/EBP α in hematopoietic cells led to the loss of granulocytes and macrophages, and resulted in the proliferation of immature cells [60].

The experiments where fates of the blood cells were changed are important in understanding the principle in lineage conversion, but they do not have an obvious therapeutic advantage. One of the first cell types with possible therapeutic applications was generated in 2009 by the *in vivo* transdifferentiation of adult mouse pancreatic exocrine cells to insulin producing β -cells [61]. In this study, over a thousand transcription factors were screened by *in situ* hybridization in pancreas to determine where they are expressed. Twenty of these factors were found to be expressed in mature β -cells and their precursors. Nine of these 20 genes resulted in different β -cell phenotypes upon mutation. These 9 candidate factors were transduced to the pancreas of immune deficient mice, and 3 of these factors, NGN3, PDX1 and MAFA were shown to result in an increase in beta cell production. PDX1 is known to be important for early pancreatic development [62] and glucose homeostasis in β -cells [63]. NGN3 is a transcription factor that promotes neuronal lineages, and it regulates pancreatic endocrine lineages during

development [64]. MAFA functions in insulin secretion [65]. The generated β -cells by this approach were also shown to be functional in a mouse model of insulin deficiency, where the *in vivo* transdifferentiation induction led to the mitigation of the symptoms.

Another lineage conversion with potential therapeutic applications was the generation of brown fat cells. These cells are responsible to produce heat from chemical energy, making them important in the body's defense against cold. In 2008, brown fat cells were induced from mouse muscle progenitor cells by expression of the transcription factor PRDM16 [66]. A year later, C/EBP β transcription factor was determined to be the partner of PRDM16, and their expression together resulted in the conversion of both human and mouse fibroblasts to induced brown fat cells [67].

Transdifferentiation of mouse fibroblasts to beating cardiomyocytes is another example of direct lineage conversion with therapeutic potential. The transcription factor genes GATA4, MEF2C and TBX5 that induce this conversion were identified from a pool of factors that were expressed in cardiomyocytes and were previously linked to cardiac development [68]. Notably, the formed cells were electrophysiologically similar to mature ventricular cardiomyocytes, and could beat in culture. Such developments are especially important, because organ transplantation is the only available treatment to end-stage heart failure.

One crucial question to address in lineage conversion was whether transduction of transcription factors can induce transdifferentiation between cells of different germ layers. An early attempt answer to this question came from experiments where the expression of C/EBP α and PU.1 converted mouse fibroblasts to macrophage-like cells [69]. However, these cells, upon inactivation of the exogenously expressed factors, converted back to their mesenchymal state. This implied the instability of this particular conversion. A more definitive answer to the question was revealed by the conversion of mouse fibroblasts to induced neurons by the expression of transcription factors ASCL1,

MYT1L, and BRN2 [70]. The generated neurons were functional, could generate action potentials and form synapses in culture. Remarkably, the engineered neurons were able to maintain their new state even when the exogenous transcription factors were inactivated. ASCL1 is known to be a transcription factor that regulates differentiation of stem cells of the central and peripheral nervous system [71]. BRN2 is expressed primarily in cortical progenitors generating glutamatergic neurons [72]. MYT1L is expressed in the central and the peripheral nervous system, as well as early post-mitotic neurons [73, 74]. In this conversion, ASCL1 is known to be the master regulator; it first binds to chromatin, and then recruits the other factors to induce neuronal reprogramming [75]. It was also recently shown to be sufficient to induce neuronal lineages by itself [76]. Since this landmark conversion, human fibroblasts have also been converted to functional neurons [77-79].

Since the initial reports of lineage conversion, similar approaches have been used to convert differentiated cell types into a number of clinically relevant cell types. A list of the most important developments in lineage conversion is provided in Table 1.1.

Table 1.1 Historically important developments on lineage conversion

M: mouse, C: chicken, H: human. Adapted from reference [80].

Year	Conversion From	Target Cell Type	Species	Reference
1987	Fibroblasts	Myoblast	M	[54]
1995	Myelomonocytic cell	Megakaryocyte-erythrocyte precursors	C	[55]
2006	Fibroblasts	Induced pluripotent stem cells	M	[4]
2007	Fibroblasts	Induced pluripotent stem cells	H	[39]
2008	Pancreatic exocrine cells	β cells	M	[61]
2010	Fibroblasts	Neurons	M	[70]
	Fibroblasts	Cardiomyocytes	M	[68]
	Fibroblasts	Blood progenitors	H	[81]
2011	Fibroblasts	Neuronal progenitors	M	[82]
	Fibroblasts	Dopaminergic cells	M	[83]
	Fibroblasts	Motor neurons	M	[84]
	Fibroblasts	Hepatocytes	M	[85]
	Fibroblasts	Neurons	H	[86]
	Fibroblasts	Motor neurons	H	[84]
	Fibroblasts	Dopaminergic cells	H	[83]
2012	Fibroblasts	Neural stem cells	H	[87]
	Fibroblasts	Embryonic Sertoli-like cells	M	[88]
2013	Fibroblasts	Endothelial cells	M	[89]
	Fibroblasts	Hepatic stem cells	M	[90]
	Fibroblasts	Hemogenic precursors	M	[91]
2014	Pancreatic acinar cells	α and δ cells	M	[92]
	Fibroblasts	Thymic epithelial cells	M	[93]
	Lymphoid, myeloid progenitors, myeloid effector cells	Hematopoietic stem cells	M	[94]
	Endothelial cells	Blood progenitors	H	[95]
	Fibroblasts	Hepatocytes	H	[96]
	Fibroblasts	Neural crest cells	H	[97]
	Skin epithelial stem cells	Limbal progenitor cells	H	[98]

Direct lineage conversion as a mode of reprogramming has several implications. First, a high level of expression of reprogramming factors is required to override the key gene regulatory networks that are readily operating in the initial cell. Overexpression of these factors is believed to induce an open epigenetic state that allows the conversion to specific cell types. However, this may also result in an intermediate cell type, with identities of both the original and target cell [99, 100]. Second, the direct nature of identity conversion shortens the time required to establish the target cell gene network. Generally, the first changes in cell phenotype occur within a few days of expression of the reprogramming factors. This is in contrast to the generation of induced pluripotent stem cells, which takes about three weeks by the use of Yamanaka factors. Considering the additional weeks-months required to obtain the target cell types from iPSCs, direct lineage conversion is a much quicker strategy to generate the desired cell types. A third aspect to consider is safety. The use of iPSCs poses serious safety concerns [101, 102]. Due to the use of integrating viruses to deliver the reprogramming factors, these cells hold the potential of tumor formation. Compared to reprogramming to pluripotency, direct lineage conversion relies on fewer factors. Furthermore, cells are not, at any point of conversion, pluripotent. Therefore, they are thought to have a lesser risk of teratoma formation in future potential applications. Another notable aspect of direct lineage conversion is that conversion does not depend on cell proliferation [70, 79, 83, 103]. Conversely, generation of iPSCs depend heavily on cell proliferation [104].

In summary, advantages of direct lineage conversion over directed differentiation of induced pluripotent stem cells include the shorter time required for conversion, higher efficiencies, independence from cell proliferation and lower risk for tumorigenicity.

1.5 Induced hepatocytes

1.5.1 Two factor conversion by *Hnf4a* and *Foxa1*

The direct conversion of mouse fibroblasts to induced hepatocyte-like cells (iHeps) was achieved by two separate groups in 2011. The first protocol by Sekiya and Suzuki used combination of *Hnf4a* with *Foxa1*, *Foxa2* or *Foxa3* [85]. These factors were identified by an initial screen of 12 candidate factors that were known to be involved in hepatocyte differentiation during development [105, 106]. Out of the pool of 12 factors, only the loss of *Hnf4a* led to significant decrease in the expression levels of the two prominent hepatic-lineage markers, albumin and alpha-fetoprotein. Investigating the cooperative effects of *Hnf4a* with the remaining factors, Sekiya and Suzuki identified *Foxa1*, *Foxa2* or *Foxa3* as partners of *Hnf4a* to be used in this lineage conversion.

The generated iHeps cells share many features with mature hepatocytes, including the expression of key hepatocyte marker genes, such as albumin (*Alb*) and alpha-1 antitrypsin (*Aat*), reduced expression of mesenchymal marker genes such as Vimentin (*Vim*), ability to store glycogen, uptake low-density lipoprotein, secrete albumin, and metabolize drugs [85]. Most importantly, iHeps can successfully reconstitute hepatic tissues of fumarylacetoacetate hydrolase (*Fah*)-deficient mice, in a model of hereditary tyrosinaemia type I [107]. The efficiency of this conversion is relatively low, about 0.1% as calculated by % E-cadherin-positive cell clusters [85].

One divergence of induced hepatocytes from mature hepatocytes is their aberrant expression of the embryonic hepatocyte marker alpha-fetoprotein. Another is their ability to proliferate indefinitely in culture. However, after transplantation into *Fah*^{-/-} mouse livers, iHeps stop proliferation, as do mature hepatocytes [85]. Furthermore, induced hepatocytes, unlike induced pluripotent stem cells, do not silence the exogenous genes once they are reprogrammed. However, the endogenous genes are activated and cells are

no longer dependent on the expression of exogenous genes, as illustrated by doxycycline removal in an inducible expression system [85].

In later studies, an analysis of the gene regulatory networks of iHeps, mature hepatocytes, and cells of several lineages revealed that iHeps generated by this protocol bear little identity to hepatocytes [108]. Interestingly, iHeps were found to aberrantly express *Cdx2*, an intestinal marker. iHeps were also shown to functionally engraft mice colon, revealing its progenitor, rather than terminally differentiated state [108]. In the same study, *Cdx2* was knocked down with shRNAs, and shown to result in a moderate increase in expression of hepatocyte-specific genes *Alb*, *Afp*, and *Aat* [108]. Clearly, the engineered cells carry characteristics of other specialized and progenitor cells types.

1.5.2 Three factor conversion by *Hnf1α*, *Foxa3* and *Gata4*

In the same year, Huang *et al.* also reported a second method to convert mouse fibroblasts to induced hepatocytes. This conversion involved the forced expression of transcription factors *Hnf1α*, *Foxa3* and *Gata4* out of a screen of 14 transcription factors known to be critical for liver development [109].

iHeps generated by this protocol express key hepatocyte specific genes, store glycogen, metabolize drugs, and uptake LDL. They also reconstitute hepatic tissues when injected to *Fah*^{-/-} mouse livers. The efficiency of this protocol is higher compared to the Sekiya and Suzuki protocol; 23% of the cells that gain epithelial morphology by day 14 are positive for albumin [109]. However, the production of iHep cells requires deletion of p19^{Arf}, due to cell death and growth arrest observed during the conversion process [109]. This is most likely a major contributor to the high efficiency of hepatic conversion. Deletion of p19^{Arf} is an important shortcoming of this protocol, because it relies on disruption of a tumor suppressor, potentially risking the safety of the technique.

1.5.3 HNF4 α and FOXA transcription factors

FOXA (forkhead box A) transcription factors are expressed in fetal and adult liver, and some other tissues belonging to the endoderm [110, 111]. FOXA proteins are crucial to regulate almost all genes specific to the liver. *Foxa2* expression is first seen around E6.5, in the primitive streak and in the node. *Foxa1* begins to be expressed at E7-E8, prior to organogenesis. *Foxa3* is expressed first at E8-E9, with its expression constricted to the midgut and hindgut [112-115]. In the developing mouse, *Foxa1* and *Foxa3* are also expressed in some neural tissues and the notochord, which are of ectoderm and mesoderm lineage, respectively, but in the adult, *Foxa* genes are only expressed in endoderm lineages [111]. Homozygous loss of *Foxa1* causes defects in glucagon expression in the pancreas, and death after birth [116]. Loss of *Foxa2* leads to lethality following gastrulation, with severe irregularities in the formation of the neural tube and in foregut morphogenesis [117, 118]. Homozygous *Foxa3* mutations result in significant decrease in liver specific gene expression, but not a dramatic phenotype [119].

HNF4 (hepatocyte nuclear factor 4) is a transcription factor belonging to nuclear receptor superfamily. It has two isoforms: *Hnf4 α* and *HNF4 γ* . *Hnf4* loss results in embryonic lethality prior to gastrulation [120]. This is caused by defects of the visceral endoderm. Lethality can be overcome by tetraploid chimera [121], but in this case, crucial hepatocyte specific genes cannot be activated [122]. Targets for *Hnf4 α* include alpha-1 antitrypsin, transthyretin genes, and *Hnf1 α* .

1.5.4 Human induced hepatocytes

Recently, two protocols of human fibroblast conversion to induced hepatocytes were published. One protocol expresses *HNF1A*, *HNF4A* and *HNF6* as the fate conversion factors, and *ATF5*, *PROX1*, and *CEBPA* as maturation factors, along with p53-siRNA and *C-MYC* to promote cell proliferation [123]. By day 25 of conversion, about 90% of the induced cells become albumin-positive as determined by flow cytometry. These cells possess drug metabolism capabilities, and can functionally reconstitute livers of Tet-uPA

(urokinase-type plasminogen activator)/*Rag2*^{-/-} mice. The second protocol delivers *FOXA3*, *HNF1A*, and *HNF4A* to human fibroblasts, and requires SV40 large T antigen to induce proliferation [96]. By day 12, about 20% of cells are double positive for albumin and alpha-1-antitrypsin as determined by immunofluorescence staining and flow cytometry. Induced hepatocytes generated by this protocol are expandable in culture, and display gene expression profiles similar to mature hepatocytes. Furthermore, the obtained cells perform well functionally; they extend survival of mice with concanavalin-A-induced acute liver injury, or fumarylacetoacetate dehydrolase deficiency [96].

1.6 Epigenetics

Epigenetics is the study of external modifications of DNA that regulates its activity. The naked DNA is wrapped around proteins with modifications, which make the DNA more or less accessible to the transcriptional machinery at any given time. This packaging is dynamic, and can be altered according to developmental and environmental cues.

Two main types of epigenetic modifications are DNA methylation and histone modification. DNA methylation takes place at the C5 position of cytosines, and DNA methylation in the context of CpG dinucleotides is associated with transcriptionally inactive regions. Histone modifications take place at the protruding N-terminal ends of core histone proteins H2A, H2B, H3 and H4, and the linker histone H1. Histone modifications function either as docking sites for other regulatory proteins to be recruited, or as neutralizers of the positive charge of histones, leading to the relaxation of chromatin to allow transcription [124]. The most common post translational modifications of histones are acetylation (maintained by histone acetyltransferases and deacetylases), methylation (maintained by lysine methyltransferases and lysine demethylases), ubiquitylation (maintained by ubiquitin ligases and deubiquitylases), and phosphorylation (maintained by kinases and phosphatases).

Combinations of different modifications can mean different things in terms of transcriptional regulation, but there are a number of modifications that are known to

convey specific effects on transcription. For example, acetylation of lysine residues on histones relinquishes the interactions of positively charged lysine residues with the negatively charged phosphate groups in DNA, allowing the relaxation of chromatin. This relief allows DNA to become accessible to other regulators of transcription. Histone methylation is different; it can have different regulatory functions depending on the modified residue. For example, tri-methylation of histone H3 lysine 27 maintained by the Polycomb Repressive Complex 2 is a transcriptionally repressive mark [125], whereas tri-methylation of histone H3 lysine 4 by MLL family of lysine methyl transferases is a transcriptionally activating mark [126]. Table 1.2 lists some of the more common histone modifications and the functions of these marks.

Table 1.2 Well studied modifications of histones

Table adapted from [127].

Modification	Function	Location of Mark
H3K4me1	activation	Enhancers and downstream of transcription start sites
H3K4me2	activation	Promoters and enhancers
H3K4me3	activation	Promoters and transcription start sites
H3K9ac	activation	Promoters
H3K9me1	activation	5' end of genes
H3K9me3	repression	Constitutive heterochromatin and repetitive elements
H3K27ac	activation	Enhancers and promoters
H3K27me3	repression	Repressive domains and silent developmental genes
H3K36me3	activation	Transcribed regions of genes, especially 3' regions after intron 1
H3K79me2	activation	Transcription-associated mark, especially 5' end of genes
H4K20me1	activation	5' end of genes

Besides histone modifications, regulation of chromatin is also achieved through the function of chromatin remodeling complexes. These complexes harness the energy obtained from ATP hydrolysis to exchange histones, to change nucleosome positions, or to entirely remove nucleosomes. The well known families of such complexes are SWI/SNF, ISWI, CHD and INO80 complexes [128].

1.6.1 Epigenetics in Stem Cells, Development and Reprogramming to Pluripotency

Embryonic stem cells have unique chromatin features. Compared to differentiated cells, ES cells have more relaxed chromatin, even in the constitutive heterochromatin regions, which are known to be densely compact [129]. Consistent with the relaxed state, linker histones are only loosely bound to chromatin, allowing flexibility [130]. ES cells also contain widely spread active chromatin regions, marked with the activating histone acetylation and H3K4 methylation, as well as poor methylation of DNA [131-133].

Although the embryonic stem cell chromatin is more permissive, lineage-specific genes need to remain inactive to conserve pluripotency. Interestingly, many genes of developmental regulators are found to be marked with both the activating H3K4me3 and the repressive H3K27me3 marks [134]. These two marks are not found on the same protruding N-terminal histone tail, but are present on the same nucleosome [135]. Such bivalency is thought to repress developmental regulators, while still keeping them poised for activation [134].

During differentiation of embryonic stem cells, the open chromatin structure becomes more compact. Heterochromatin regions condense and become more abundant, as demonstrated by the significant increase in H3K9me2 marked regions, a modification associated with facultative heterochromatin [136]. The repressive H3K27me3 mark is also spread to repressed genes and intergenic regions in differentiated cell. Active genes specific to the somatic cell lineage are marked with H3K4me3 at the promoter, and

H3K4me1, H2K4me2 and H3K27ac at the enhancer regions [137-139]. These data indicate that genes are tightly regulated at the epigenetic level during development to ensure they are correctly expressed upon differentiation.

Reprogramming of differentiated cells to pluripotency is epigenetically challenging, because it requires dramatic reshaping of the epigenome. The four reprogramming factors transduced to cells are particularly responsible for the activation of endogenous pluripotency genes that are found in closed chromatin regions.

DNA packed with nucleosomes is less accessible to factors. Many transcription factors fail to bind to such tightly packed chromatin. However, reprogramming factors OCT4, SOX2 and KLF4 can bind such stretches of DNA in a cooperative manner. It is possible that binding of one of these factors help destabilize the nucleosome, aiding the binding of the remaining factors to sites that were previously inaccessible [140]. It is important to note that although cooperative binding is important, the OSK factors can nevertheless bind by themselves to their targets protected by closed chromatin [141], making them pioneer transcription factors.

Early in reprogramming, there are dramatic changes in both the enhancer and promoter regions. Gene expression changes firstly occur in regions marked with the activating H3K3me3 modification[142]. This is consistent with the idea that early in reprogramming, the major response of cells is achieved through the disruption of the lineage-specific gene expression.

Apart from genes that change their expression levels, many histone modifications also occur in other promoter regions. These are promoters of genes that will be required in later stages of reprogramming, such as *Lin28*, and they gain the H3K4me2 mark [142], which is associated with active or poised promoters and enhancers.

Normally, enhancers that are active are marked with H3K27ac, H3K4me1 and H3K4me2, but do not contain H3K4me3. When inactive, these enhancers remain unmarked, whereas in the poised state they contain H3K4me1 or me2 without the H3K27ac [143]. During reprogramming to pluripotency, enhancers have their histone modifications altered rapidly. About 60% of the enhancers of fibroblast-specific genes lose their active marks, and about 1000 of embryonic stem cell-specific enhancers become poised, as determined by the changes in their H3K4 methylation status [142]. Altogether, these global changes in the epigenome are crucial in the disruption of the differentiated cell gene expression profile, and establishment of the pluripotent stem cell state.

1.6.2 Epigenetics in Lineage Conversion

The best studied epigenetic changes during direct lineage conversion are in the transdifferentiation of mouse fibroblasts to functional neurons [70]. This protocol ectopically expressed *Ascl1*, *Brn2* and *Myt1l* in fibroblasts. Three years after the original protocol was published, the same group reported on the epigenetic mechanisms that are at play during this conversion [75]. According to their findings, ASCL1 is a pioneer transcription factor that can bind to chromatin irrespective of the state of compaction. This transcription factor binds to closed chromatin regions marked with H3K4me1, H3K27ac, and H3K9me3 marks. Its significance in this conversion is further demonstrated by the misdirection of BRN2 to its target sites without ASCL1.

In the same study, the binding regions of ASCL1 were shown to be preferentially in the closed state, whereas BRN2 and MYT1L primarily bind to regions marked with activating histone modifications. Interestingly, the binding of ASCL1 is dependent on the presence of the repressive mark; removal of the H3K9me3 mark by histone demethylase significantly decreases the efficiency of this lineage conversion [75]. This finding is supportive of the observation that certain cell types such as human keratinocytes and osteoblasts cannot be converted to induced neurons, because they do not contain the trivalent H3K4me1, H3K27ac, and H3K9me3 marks. On the other hand, cell types that

can be converted to functional neurons, such as mouse and human fibroblasts, and human myoblasts, do possess these trivalent marks [144].

1.6.3 Polycomb Repressive Complex

One of the most studied mediators of epigenetic modifications is the polycomb group (PcG). PcG is composed of two complexes. Polycomb repressive complex 2 (PRC) catalyzes the tri-methylation of lysine 27 residue of histone H3. Polycomb repressive complex 1 (PRC1) recognizes this modification, and catalyzes the mono-ubiquitylation of lysine 119 residue of histone H2A [145].

The main components of PRC2 complex are EZH2 (enhancer of zeste homologue 2), EED (embryonic ectoderm development) and SUZ12 (suppressor of zeste) [146]. EZH2 is a histone methyltransferase with a SET (su[*var*]3–9, enhancer of zeste, Trx) domain, which catalyzes the addition of methyl groups to lysine 27 of histone H3 [147]. EED has WD40 domains, which interact with H3K27me₃, allowing it to spread the mark. These three components are required for PRC2's activity *in vitro*, whereas their association with the histone chaperone protein RBBP4/7 is required *in vivo* [148].

1.6.4 Recruitment of and Transcriptional Repression by PRC2

In *Drosophila*, Polycomb Group proteins bind to DNA named Polycomb Response Elements (PREs). Specific PRE sequences have not been identified in mammalian cells. However, it is known that CpG rich sequences mediate PRC2 binding [149]. Some of the mechanisms proposed for PRC2 recruitment are association with interaction partners, transcription factors, and recruitment by non-coding RNAs [150]. The exact function of regulation of histone modifications by PRC complexes remains uncertain. H3K27 tri-methylation by PRC2 has been analyzed mostly in the context of PRC1 recruitment. Mono-ubiquitylation of H2AK119 by PRC1 component ubiquitin ligase RING1A and RING1B is thought to function in compaction of chromatin and thereby repression of transcription. Polycomb repressive complexes were shown to induce the condensation of

nucleosomes *in vitro* [151]. Other proposed mechanisms through which H2AK119ub1 might silence transcription are inhibition of H3K4 methylation, inhibition of RNAP II activity, and hindrance of H2A-H2B dimer displacement from transcription sites [150].

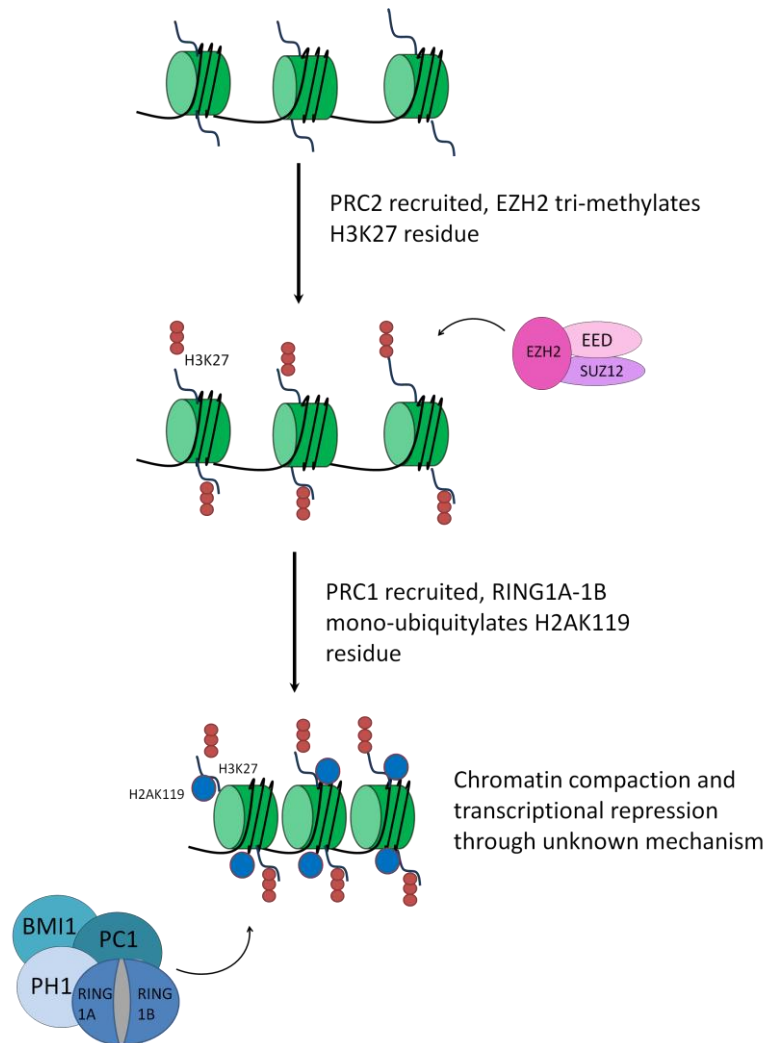


Figure 1.2 Polycomb Repressive Complexes

Polycomb repressive complex 2 (PRC2) is made up of EZH2, EED and SUZ12 components, where EZH2 methylates lysine 27 of histone H3. This modification recruits Polycomb repressive complex 1 (PRC1). Two of the PRC1 components are the ubiquitin ligases RING1A and RING1B, which mono-ubiquitylates lysine 119 of Histone H2A. This modification is believed to interrupt nucleosome dynamics, and hence inhibit transcription. Adapted from reference [152].

1.6.5 EZH2 in pluripotency, differentiation and reprogramming

In embryonic stem cells, at least 10% of genes are targeted by PRC2 [153]. Specifically, PRC2 targets *Hox* genes and other developmental regulators [154-156]. Notably, most PRC2 targeted genes also bare the H3K4me3 mark, which is associated with active transcription. This allows embryonic stem cells to have a flexible chromatin structure, which is considered to be important for keeping cells poised for differentiation [134].

For the maintenance of pluripotency, there are two models to explain the role of PRC2-mediated transcriptional repression. The first model claims that pluripotency is maintained through the repression of genes that induce differentiation, and the second states that pluripotency is lost when genes that are essential to maintain pluripotency are silenced [147]. Evidence against the first model came from studies that showed embryonic stem cells lacking a PRC2 component can still be maintained in an undifferentiated state [157]. This finding is supportive of the second model that suggests PRC2 is involved in the repression of pluripotency genes, which leads to differentiation. In support of this idea, it was shown that loss of PRC2 component SUZ12 is associated with inefficient silencing of *Nanog* and *Oct4*, two pluripotency factors [158].

PRC2 complex is highly important in development, illustrated by severe gastrulation defects when any one of the components is lost. *Ezh2* knockout embryos die on E8.5. *Ezh2*'s expression is upregulated following fertilization, and it is expressed during preimplantation [159]. EED first appears after E5.5. This implies that EZH2 may have PRC2 independent roles in preimplantation [160]. *Eed* knockout mice have gastrulation defects and fail to inactivate the X chromosome in extraembryonic cells. Loss of *Suz12* is lethal before E10.5 [161].

PRC2's role in normal differentiation was mostly studied in culture of mouse embryonic stem cells deficient in PRC2 components. When differentiation is induced in embryonic stem cells *in vitro*, *Eed* deficient cells were still able to contribute to all lineages in

chimeric embryos [157], *Suz12* deficient cells could not form the endodermal layer [162], *Ezh2* or *Eed* deficient cells showed impaired commitment to mesodermal lineages [163].

Considering the role of PRC2 in ES cell differentiation as described above, PRC2 loss is predicted to cause defects in differentiation. This is the case for adipogenic [164] and lymphogenic [165] lineages, where PRC2 inactivation prevents differentiation. However, PRC2 inactivation also induces differentiation toward myogenic lineage [166] and formation of epidermis [167].

During development, endodermal cells that differentiate into hepatic and pancreatic lineages are marked with specific histone marks. In these multipotent cells, the regulatory elements of pancreas-specific genes, such as the master regulator *Pdx1*, carry the activating H3K9 and H3K14 acetylation marks, as well as the repressive K3K27me3 mark [168]. In contrast, liver-specific genes of the undifferentiated endoderm cells, such as *Alb* and *Afp* have only low levels of three of these marks. Upon BMP4 signaling during development, histone acetyltransferase p300 is recruited to liver-specific genes and acetylates their regulatory regions [168]. Hence, hepatic specification is initiated. Meanwhile, PRC2 sustains the K27 tri-methylation mark of histone H3 in pancreas-specific genes, which inhibits differentiation toward pancreatic lineage [168].

H3K27 tri-methylation mediated by the PRC2 complex is important in reprogramming cells to pluripotency. This requirement was demonstrated by the inability of PRC2 depleted mouse embryonic stem cells to reprogram differentiated human B cells through the formation of heterokaryons [169]. It was later shown that knockdown of PRC2 components by shRNA also significantly decreases reprogramming efficiency of human fibroblasts into induced pluripotent stem cells by the Yamanaka factors [170]. These two studies reported similar effects for the inhibition of other PRC2 (EED, SUZ12) and PRC1 (BMI1, RING1) components.

1.6.6 DOT1L

DOT1L (disruptor of telomeric silencing like protein) is a histone methyl transferase that adds methyl groups to lysine 79 of histone H3. DOT1L is unique because it is the only known histone H3 K79 methyl transferase, and it does not have a SET (su[*var*]³⁻⁹, enhancer of zeste, Trx) domain that is common to histone lysine methyltransferases [169, 171]. Its catalytic activity lies in a domain that is conserved among class I methyl transferases [172].

1.6.7 DOT1L in pluripotency, differentiation and reprogramming

In mouse embryonic stem cells, the H3K79me₂ mark is commonly found with the activating H3K4me₃ mark near the transcription start site of actively transcribed genes [173]. However, DOT1L is dispensable for mouse embryonic stem cell maintenance. When *Dot1l* was depleted in ES cells, the H3K79me₂ mark was lost, but did not significantly affect cell proliferation or overall morphology [173]. Upon induction of differentiation, *Dot1l* knockdown cells arrested in the G2/M phase of the cell cycle, and had aneuploidy [173]. Although these ES cells had differentiation defects in culture, teratoma assays showed that both *Dot1l* expressing and deficient ES cells could differentiate into all three germ layers [173].

In normal mouse development, loss of *Dot1l* causes embryonic lethality at E9.5-E10.5 due to defects in growth and angiogenesis [174]. Moreover, *Dot1l* deficient mouse embryonic stem cells have less heterochromatic marks at centromeres and telomeres, causing a more relaxed chromatin state, and elongation of telomeres [174].

In line with the findings in embryonic stem cells, cardiac-specific knock out of *Dot1l* in mouse embryos leads to enlarged hearts, a condition which could be rescued by expressing *Dot1l* postnatally [174]. DOT1L is also involved in cartilage development, as shown by *Dot1l* knockdown resulting in chondrogenic defects [175].

The better studied role of DOT1L is in hematopoiesis, a process that occurs in both the embryo and the adult. DOT1L's role is best characterized in the case of leukemic transformation, which is caused by accumulating mutations in hematopoietic precursors and results in improperly differentiated blood cells [176]. One such mutation causes the fusion of the *MLL* gene with many different partners. In humans, these fusion proteins were shown to interact with DOT1L, and cause aberrant H3K79 methylation [177].

In the case of normal reprogramming that occurs during fertilization, it was shown that mouse oocyte has H3K79me2 and me3 marks, but these rapidly disappear following fertilization [178]. This indicates that the loss of H3K79 in development has some role in reprogramming to totipotency.

In artificial reprogramming, DOT1L inhibition with small molecules or genetic knockout was shown to increase the efficiency of conversion of mouse fibroblasts to induced pluripotent stem cells (iPSCs) about 3-4 fold [170]. Inhibition of DOT1L aids the loss of the H3K79me2 from fibroblast specific genes that are silenced upon reprogramming. This facilitates the transition from a mesenchymal to an epithelial state, which is also relevant for the case of induced hepatocytes.

1.7 Aim

In this study, we aim to investigate the chromatin modifications that regulate the conversion of mouse fibroblasts to induce hepatocytes. Better understanding of this lineage conversion might aid in the development of new techniques to generate these cells at higher efficiencies and in better qualities. Our specific aims are to identify which chromatin modifiers have a functional role in transdifferentiation, the downstream genes regulated by the identified chromatin modifiers, and whether modulation of chromatin modifier activity leads to changes in lineage conversion efficiency. With this purpose, we infected mouse fibroblasts with *Hnf4a* and *Foxa1* transcription factors, and treated them

with a number of small molecule inhibitors that were known to be important regulators of gene expression in previous studies.

Chapter 2

MATERIALS AND METHODS

2.1 Cell Culture

HEK 293T cells, mouse embryonic fibroblasts (MEFs) and mouse tail tip fibroblasts (TTFs) were cultured in DMEM (Gibco) supplemented with 10% FBS (Gibco) and 1% Penicillin-Streptomycin (Gibco), (D10 Medium), at 37°C with 5% CO₂. *Ezh2* fl/fl MEFs [163] were kindly provided by Stuart Orkin (Harvard Medical School) and *Dot1l* fl/fl TTFs [179] were a gift from. Scott Armstrong (Memorial Sloan Kettering Cancer Center). Induced hepatocytes were maintained in hepatomedium (hepatomedium: DMEM:F12 (Gibco) supplemented with 10% FBS, 1 µg/ml insulin (Sigma), 10⁻⁷ M dexamethasone (Sigma), 10 mM nicotinamide (Sigma), 2mM L-glutamine (Gibco), 50 µM β-mercaptoethanol (Gibco), 1% Penicillin-Streptomycin (Gibco), 20 ng/ml EGF (PeproTech) and 20 ng/ml HGF (PeproTech)). dH1f cells were cultured in alpha-MEM (Gibco) supplemented with 10% FBS, 1% Penicillin-Streptomycin (Gibco) and 1X Non-essential amino acids (Gibco) at 37°C with 5% CO₂. Puromycin selection was performed at 1 µg/ml for 3 days.

2.2 DNA Constructs

pGCDNsam-*Hnf4α*-IRES-GFP (Addgene plasmid # 33006) and pGCDNsam-*Foxa1*-IRES-GFP (Addgene plasmid # 33007) were purchased from Addgene [85]. shRNA constructs for sh*Ezh2* and shControl, and overexpression constructs pMIG-IRES-GFP and pMIG *Ezh2* were gifts from Dr. Stuart Orkin. MSCV-CreERT2 puro (Addgene plasmid #22776) and pBABE-puro SV40 LT (Addgene plasmid # 13970) were purchased from Addgene [180, 181].

2.3 Small Molecules

Small molecules used to test effect on fibroblast to hepatocyte conversion were added to the hepatomedium following the last infection with the transcription factors unless stated otherwise. These chemicals are listed in Table 2.1.

Table 2.1 List of Small Molecules Used

Small Molecule	Source	Concentration	Solvent	Target
5-azacytidine (5-AZAC)	Tocris Bioscience	5 μ M	DMSO	DNA methyltransferase inhibitor
BIX01294	Tocris Bioscience	1 μ M	DMSO	G9a inhibitor
Bromosporine	Tocris Bioscience	250nM	DMSO	Broad spectrum bromodomain inhibitor
CBP/Brd4	Dr. Udo Oppermann	1.25 μ M	DMSO	Bromodomain inhibitor
CHIR99021	Tocris Bioscience	10 μ M	DMSO	GSK-3 inhibitor
3-DZNep	Tocris Bioscience	50nM	water	EZH2 inhibitor
EPZ4777	Tocris Bioscience	3 μ M	DMSO	DOT1L inhibitor
Forskolin	Tocris Bioscience	10 μ M	DMSO	adenylate cyclase activator
Sodium Butyrate (NaB)	Sigma	0.5mM	water	HDAC inhibitor
RepSox (616452)	Tocris Bioscience	5 μ M	DMSO	TGF- β type I receptor inhibitor
SGC0946	Tocris Bioscience	0.1 μ M	DMSO	DOT1L inhibitor
Tranylcypromine	Tocris Bioscience	5 μ M	DMSO	Histone demethylase LSD1 inhibitor
TTNBP	Tocris Bioscience	1 μ M	DMSO	Analog of retinoic acid
Valproic Acid (VPA)	Tocris Bioscience	500 μ M	water	HDAC inhibitor

2.4 Isolation of primary mouse embryonic fibroblasts

Pregnant BALB/c mice were sacrificed 13.5 dpc by cervical dislocation. The uterine horns were dissected and dipped in 70% ethanol and then submerged in PBS until they could be transferred to the tissue culture hood. Once in the hood, the uterine horns were placed in sterile Petri dishes containing 2 ml PBS. Embryos were separated from their placenta and embryonic sac by sterile scissors and forceps. The head and internal organs were removed, and the remaining tissues were collected in a clean dish. The tissues were

minced with a sterile razor until they could be pipetted. The minced embryos were collected in a falcon tube, and incubated in trypsin (1 ml per embryo) for 15 min at 37°C. The cells were pipetted up and down every 5 min inside the hood. Trypsin was then inactivated by adding an equal volume of D10 medium. Cells were pelleted by centrifuging at 1500 rpm for 5 min. The supernatant was discarded and cells were resuspended in warm D10. Cells were plated onto 0.1% gelatin coated 10 cm dishes, two embryos per dish, and labeled as passage 0.

2.5 Generation of *Ezh2* Knockout MEF and *Dot1l* Knockout TTF Cell Lines

P3 *Ezh2* fl/fl MEFs were infected with pBABE-puro SV40 LT and MSCV-CreERT2 puro vector, and selected with puromycin for 3 days. P3 *Dot1l* fl/fl MX-cre TTFs were infected with the MSCV-CreERT2 vector and selected with puromycin for 3 days. The surviving cells were then infected with pBABE-puro SV40 LT vector. To create knock-out cells, creER infected *Ezh2* fl/fl MEFs and *Dot1l* fl/fl TTFs were treated for 5 days with 0.5 μ M 4-OHT (Sigma).

2.6 Production of Viral Particles

HEK 293T cells were seeded at a density of 2.5×10^6 cells to 10 cm plates. The next day, in the afternoon, cells were transfected with 225 ng envelope protein (VSVG), 2250 ng Gag Pol (pUMVC for retroviruses and 8.2DeltaVPR for lentiviruses) and 2500 ng viral vector using 20 μ l FuGENE® 6 Transfection Reagent (Promega) in a total volume of 400 μ l serum-free DMEM according to the manufacturer's instructions. In the morning, the transfection medium was washed off and replaced with fresh 8 ml 293T media. 48 hours after transfection, first viral particles were collected and stored at 4°C. 293T media was refreshed and a second virus collection was performed at 72h. Viral supernatants from the two collections were combined and filtered through a 45 μ m low protein binding syringe filter to clear viral particles from cells and debris. All plastic materials that came in contact with the virus were bleached and trashed accordingly.

2.7 Concentration of Viruses

PEG-8000 (Sigma) dissolved in PBS was added to collected viral supernatant to a final concentration of 10% (w/v). The mixture was incubated at 4°C overnight. The next day, the mixture was centrifuged at 2500 rpm for 20 min at 4°C. The supernatant was discarded, leaving a small volume in the tube not to disrupt the pellet. The tube was centrifuged once more at 1200 rpm for 5 min to collect the remaining fluid in the tube. The remaining supernatant was discarded and the pellet containing the viral particles was resuspended in PBS, concentrating the virus 100 fold. The virus was stored in aliquots at -80°C.

2.8 Titering of Viruses

P1 BALB/c MEFs were seeded to 0.1% gelatin coated 6 well plates at a density of 200000 cells/well. The next day in the afternoon, cells were infected with 10, 1 or 0.1 µl of concentrated virus with 8 µg/ml protamine sulfate (Sigma-Aldrich) in the infection medium. In the morning, virus was washed away. The next day, 48 hours after infection, cells were analyzed by flow cytometry to quantify the percentage of GFP-positive cells. The titer of the virus (transducing units / ml) was calculated by multiplying this percentage by the total number of plated cells and then dividing it by the volume of virus used.

2.9 Flow Cytometry

Cells were harvested by trypsinization for the analysis of intracellular proteins, and washed once with PBS. Cells were centrifuged at 1500 rpm for 5 min and resuspended in FACS buffer (PBS with 1% FBS). BD Accuri C6 Flow Cytometer (BD Biosciences) was used to analyze the samples according to the manufacturer's instructions.

2.10 Generation of induced hepatocytes

P1 mouse embryonic fibroblasts (MEFs) were seeded onto 0.1% gelatin coated 12 well plates at a density of 5×10^3 cell per well. The next day in the afternoon, cells were

infected with concentrated retroviruses carrying mouse *Hnf4a* and *Foxa1* transcription factors at MOI 10. In the morning, medium was changed. Infection was repeated in the afternoon, after a minimum of 8 hours. The day after, the medium was replaced with growth-factor free hepatomedium. After one week of culture of cells in hepatomedium, cells were replated on to type I collagen coated plates, and grown in hepatomedium containing HGF and EGF.

2.11 Histone Extraction

A modified version of Abcam's histone acid extraction protocol was used. Cells were harvested and washed once with ice-cold PBS. Cells were then resuspended in Triton Extraction Buffer (TEB: PBS with containing 0.5% Triton X 100 (v/v), 2 mM phenylmethylsulfonyl fluoride (PMSF) added fresh, 0.02% (w/v) NaN₃) at a density of 10⁷ cells per ml, and lysed on ice for 10 minutes with gentle stirring. Following lysis, cells were centrifuged at 6500 x g for 10 minutes at 4°C to pellet the nuclei, and the supernatant was discarded. The nuclei were washed in half the volume of TEB, and resuspended in 0.2 N HCl at a density of 4x10⁷ nuclei per ml. Histones were acid extracted overnight at 4°C. The next day, samples were centrifuged at 6500 x g for 10 minutes at 4°C to pellet debris. The supernatant containing the histone proteins were collected, and mixed with 1/5 volume of 0.1M NaOH to neutralize the pH. Protein concentrations of samples were determined by Pierce BCA Protein Assay Kit (Life Technologies) according to the manufacturer's instructions. Samples were stored at -80°C.

2.12 Western Blotting

Histone proteins were mixed with 4X Loading Dye (9:1 mixed 4X Laemmli Sample Buffer (Bio-Rad), and β-mercaptoethanol (Bio-Rad)) to a final concentration of 1X, and boiled at 95°C for 15 min. 1-2 µg of histone extracts were run on 4-15% gradient Mini-PROTEAN® TGX™ Precast Protein Gels (Bio-Rad) using Tris/Glycine/SDS running buffer (Bio-Rad) in a Mini-PROTEAN Tetra cell electrophoresis chamber (Bio-Rad) at

20mA for 30-40 min. Proteins were transferred to PVDF membrane (Bio-Rad) by semi-dry transfer using the Trans-Blot® Turbo™ Transfer System (Bio-Rad) at 25V for 30 min. Tris/Glycine/SDS containing 20% ethanol was used as the transfer buffer. Following transfer, proteins were visualized by staining with Ponceau S Red solution for 10 min (Bio-Rad). The stain was washed away with TBS containing 0.1% Tween 20 (TBS-T). The membrane was blocked with 5% non-fat dry milk (w/v) (Bio-Rad) dissolved in TBS-T for an hour at RT. Blocked membrane was incubated with the primary antibody dissolved in antibody dilution solution (2% bovine serum albumin (BSA), 0.02% NaN₃ in TBS-T) overnight at 4°C with gentle shaking. The next morning, the unbound antibody was washed three times with TBS-T, 15 min each on shaker. Membrane was then incubated with HRP-conjugated secondary antibodies diluted in 5% milk in TBS-T for 1 h at RT. Following this incubation, the membrane was washed three times with TBS-T, 15 min each on shaker. The protein bands were visualized by chemiluminescent detection using Pierce ECL Western Blotting Substrate (Life Technologies) and CL-XPosure X-ray Film (Life Technologies) according to the manufacturer's instructions. All of the antibodies used and their dilutions are listed in Table 2.2.

Table 2.2 List of all antibodies used for Western Blotting

Antibody	Host	Brand	Catalog Number	Dilution
Histone H3	Rabbit	Cell Signalling	9715	1 : 1000
H3K27me3	Rabbit	Millipore	07-449	1 : 1000
H3K79me2	Rabbit	Abcam	ab3594	1 : 1000
rabbit IgG-HRP	Goat	Santa Cruz	sc-2054	1 : 2000

2.13 Gene Expression Analysis by Quantitative PCR (qPCR)

RNA was isolated from harvested cells by using NucleoSpin RNA isolation kit (Macherey-Nagel). To synthesize cDNA, a 16.5 µl reaction was set up with 200 ng -1 µg RNA, 2.5 µl 2 mM dNTP (Life Technologies) and 1 µl 50 µM hexanucleotide mix (Invitrogen) in a PCR tube. The mixture was incubated at 65°C for 5 min, and quickly chilled on ice. RT mix prepared from 5 µl 5X First Strand Buffer (Invitrogen), 2 µl 0.1M

DTT (Invitrogen) and 0.5 μ l RNasin (Promega) was added to the cooled sample, and incubated for 10 min at RT. 1 μ l of MMLV-RT enzyme (Invitrogen) was added and cDNA was synthesized at 37°C for 1 hour. The enzyme was then inactivated by incubating the samples at 70°C for 15 min. The synthesized cDNA was diluted to 100 μ l with dH₂O. A 20 μ l PCR reaction was set up in 96 well opaque plates (Roche) using 10 μ l of LightCycler 480 SYBR Green I Master Mix (Roche), 2 μ l of cDNA, 2 μ l of 2.5 mM forward and reverse primer mix, and 6 μ l of dH₂O. LightCycler 480 Instrument II (Roche) was used to determine the threshold cycle (Ct) values according to the manufacturer's instructions. See List of primers used for qPCR for a list of primers used in this study.

Table 2.3 List of primers used for qPCR

Gene	Species	Forward Primer	Reverse Primer
<i>Afp</i>	mouse	CCTGTGAACTCTGGTATCAGCCA	CTCAGAAAACCTGGTGATGCATAGC
<i>Alb</i>	mouse	TGTCGCCAAAGAGTTTAAAGCTG	TCTTAATCTGCTTCTCCTTCTCTGG
<i>Aat</i>	mouse	TCGGAGGCTGACATCCACAA	TCAGCTGCAGCTCACTGTCTGG
<i>Cdh1</i>	mouse	CAGGTCTCCTCATGGCTTTGC	CTTCCGAAAAGAAGGCTGTCC
<i>Vim</i>	mouse	CGTCCACACGCACCTACAG	GGGGGATGAGGAATAGAGGCT
<i>Pcolce</i>	mouse	GCCAGACCCCCAACTACAC	CCGTAATTGTCCAGATGCACTT
<i>Acta2</i>	mouse	GTCCCAGACATCAGGGAGTAA	TCGGATACTTCAGCGTCAGGA
<i>Col1a1</i>	mouse	GCTCCTCTTAGGGGCCACT	CCACGTCTCACCATTGGGG
<i>Col1a1</i>	mouse	GCTCCTCTTAGGGGCCACT	ATTGGGGACCCTTAGGCCAT
<i>Col1a1</i>	mouse	TAAGGGTCCCCAATGGTGAGA	GGGTCCCTCGACTCCTACAT
<i>Col1a1</i>	mouse	CTGGCGGTTCAAGGTCCAAT	TTCCAGGCAATCCACGAGC
<i>Fbln2</i>	mouse	GCGTACAAGGAGGGTTTGTG	AATGCAGTTGGGTGGTAGCTC
<i>S100a4</i>	mouse	TCCACAAATACTCAGGCAAAGAG	GCAGCTCCCTGGTCAGTAG
<i>Actb</i>	mouse	TTGCTGACAGGATGCAGAAGGAGA	ACTCCTGCTTGCTGATCCACATCT
<i>Ezh2</i>	mouse	AGTGACTTGGATTTTCCAGCAC	AATTCTGTTGTAAGGGCGACC
<i>Twist1</i>	mouse	GGACAAGCTGAGCAAGATTCA	CGGAGAAGGCGTAGCTGAG
<i>Snai1</i>	mouse	CACACGCTGCCTTGTGTCT	GGTCAGCAAAAGCACGGTT
<i>Snai2</i>	mouse	TGGTCAAGAAACATTTCAACGCC	GGTGAGGATCTCTGGTTTTGGTA
<i>Hnf1a</i>	mouse	GACCTGACCGAGTTGCCTAAT	GCGAAGTCTTCCCCATCGTC
<i>Zeb1</i>	mouse	GCTGGCAAGACAACGTGAAAG	GCCTCAGGATAAATGACGGC
<i>Zeb2</i>	mouse	ATTGCACATCAGACTTTGAGGAA	ATAATGGCCGTGTGCTTCG
<i>Pdx1</i>	mouse	CCCCAGTTTACAAGCTCGCT	CTCGGTTCCATTCGGGAAAGG
<i>Hnf6</i>	mouse	GGCAACGTGAGCGGTAGTTT	TTGCTGGGAGTTGTGAATGCT
<i>ACTB</i>	human	TGAAGTGTGACGTGGACATC	GGAGGAGCAATGATCTTGAT
<i>ALB</i>	human	TGCAACTCTTCGTGAAACCTATG	ACATCAACCTCTGGTCTCACC
<i>AAT</i>	human	TTCGTTGTCATGTCCTCAAACCTG	TGTGGTCAGGTAGTCTTCAAGG
<i>Hnf4a</i>	mouse	CCCCTAACCCAGAAATCCAA	GCTGGGGATTGTCATGCTTT
<i>Afp</i>	mouse	CTCTACGGCATTGAGGGTGA	AACAGCAACAAATGGGCACT
<i>Hnf1a</i>	mouse	TGGAAAACCACACACACCAG	AGAGGGCCATGCTCCATTT
<i>Alb</i>	mouse	TACGATGGCAAAGACCATCA	GTGTGAAACGGGTCTGGAGT

2.14 Periodic Acid-Schiff Staining

Cells attached to the plates were washed in PBS and fixed in 10% formalin for 20 min. Formalin was washed with PBS and dH₂O. Cells were incubated in periodic acid solution for 5 min at RT, and rinsed several times with dH₂O. Next, samples were incubated in Schiff's reagent for 15 min at RT. The solution was washed away with running tap water for 5 min. Samples were left in PBS.

2.15 Immunofluorescence

Cells grown on cover slips were washed once in PBS and fixed with 4% paraformaldehyde for 30 min at RT. PFA was washed with PBS and samples were blocked with PBS containing 20% serum and 0.2% Triton X-100. Cells were then incubated with the primary antibodies overnight at 4°C. The next day, the primary antibodies were washed three times with PBS containing 0.2% Triton X-100. The cells were then incubated with the secondary antibodies for 1 h at RT. After washing cells three times with PBS 0.2% Triton-X, cover slips were mounted onto glass slides with 50% glycerol in PBS containing 1:10000 Hoechst. Images of samples were taken with the Nikon Eclipse 90i Confocal Microscope. Table 2.4 lists all antibodies used for immunofluorescence microscopy.

Table 2.4 List of antibodies used for immunofluorescence microscopy

Antibody	Host	Brand	Catalog Number
E-cadherin	Mouse	BD Biosciences	#610081
Vimentin	Rabbit	Abcam	ab92547
Albumin	Rabbit	Abcam	ab135575
Rabbit IgG (H+L), Alexa Fluor 488 conjugate	Goat	Life Technologies	A-11008
Mouse IgG (H+L), Texas Red-X conjugate	Goat	Life Technologies	T-6390

2.16 Immunostaining

Cells on plates were fixed with 4% paraformaldehyde for 30 min at RT, washed with PBS, and incubated with biotin conjugated anti-E-cadherin (eBioscience #13-3249-82) primary antibody diluted in staining buffer (PBS, 3% FBS, 0.3% Triton X-100) overnight at 4°C on orbital shaker. The next day, primary antibody was washed away with 5 rounds of 2 min PBS washes. Samples were incubated with streptavidin-HRP (BioLegend #405210) diluted 1:500 in staining buffer 2 h at RT. Antibody was washed away with 5 PBS washes. Stain was developed with DAB kit (Vector Labs) (Development solution: 2 drops of buffer, 4 drops DAB, 2 drops sodium hydroxide and 2 drops nickel solution added sequentially to 5 ml distilled water and mixed after each addition). The stain was developed within 15 min, and the development solution was washed away with 2 PBS washes. Cells positive for staining were quantified with ImageJ software.

2.17 Chromatin Immunoprecipitation (ChIP)

Cells were harvested by trypsinization and collected in a suspension of 10^6 cells/ml. For crosslinking of proteins to DNA, 37% formaldehyde was added to the cells to a final concentration of 1%. Cells were rocked at room temperature for 10 min. This reaction was stopped by adding 2.5 M glycine to a final concentration of 0.125 M, and rocking cells at RT for 5 min. Cells were pelleted by centrifugation at 1500 rpm at 4°C, and washed twice with ice-cold PBS, using at least half of the crosslinking volume. After the washing steps, the cells were pelleted and the supernatant was removed. At this point, pellets could be stored at -80°C. Cells were resuspended in ChIP SDS lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl, pH 8.1) with protease inhibitor cocktail (Roche) at a density of 2×10^6 cells/ 100 μ l and aliquoted 100 μ l per 0.5 ml Bioruptor microtubes (Diagenode). Samples were sonicated in Bioruptor Plus sonicator (Diagenode) in a water bath cooled to 4°C. Sonication was performed using high power level, 30 second on-30 second off duty cycles for three rounds of 5 min. Sample was briefly vortexed and spun down after each round. Following sonication, the sample was centrifuged at top speed for 10 min at 4°C. The supernatant was transferred into a fresh tube, and the pellet containing

cell debris was discarded. The mixture was 1:10 using ChIP dilution buffer (0.01% SDS, 1.1% Triton-X 100, 1.2 mM EDTA, 16.7 mM Tris-HCl, pH 8.1 and 167 mM NaCl) with protease inhibitors. At this point, 10% of the sample was taken as the input control. NaCl was added to the input to a final concentration of 0.2 M. The input control was incubated 4 hours to overnight at 65°C. Input was then stored at -20°C until further use. The remaining sample was incubated with primary antibody overnight at 4°C in a rotator. 5 µg of antibody was used per 2×10^6 cells. Table 2.5 lists all antibodies used for ChIP.

Table 2.5 List of antibodies used in chromatin immunoprecipitation experiments

Antibody	Host	Brand	Catalog Number
H3K27me3	Rabbit	Millipore	07-449
Normal rabbit IgG		Santa Cruz	sc-2027

In the morning, protein G magnetic Dynabeads (Life Technologies) were washed 3 times in TE (1 mM EDTA, 10 mM Tris-HCl, pH 8.1). Beads were pelleted with the help of a magnetic rack. 50 µl of beads were added to the sample, and the mixture was rotated at 4°C for 1 hour. Following incubation with the beads, the unbound sample was discarded. The beads and the bound protein-chromatin complex were washed with low salt wash buffer (0.1% SDS, 1% Triton-X 100, 2 mM EDTA, 20 mM Tris-HCl, pH 8.1, 150 mM NaCl), high salt wash buffer (0.1% SDS, 1% Triton-X 100, 2 mM EDTA, 20 mM Tris-HCl, pH 8.1, 500 mM NaCl), LiCl wash buffer (0.25 M LiCl, 1% IGEPAL-CA 630, 1% deoxycholic acid (sodium salt), 1 mM EDTA, 10mM Tris 8.1) and twice with TE. The tube was rocked 5 min per wash at 4°C. Washed samples were eluted with 150 µl of freshly prepared elution buffer (1% SDS, 0.1M NaHCO₃), by rocking the beads 15 min at RT. Elution was repeated once more with fresh elution buffer. Eluants were combined and reverse crosslinked by incubation at 65°C for 4 hours to overnight with 0.2M NaCl added to the sample. 5 µl of 20 mg/ml Proteinase K was then added to cooled the sample and input, and incubated at 37°C for 2 hours. Proteinase K digested samples were purified using ChIP DNA Clean & Concentrator Kit (Zymo Research). Sample and input were eluted using 80 µl of elution buffer. 1 µl of RNase A (Life Technologies) was added to

each sample to digest any RNA present. Tubes were incubated at 37°C for 30 min. Following digestion, samples were stored at -20°C.

Chapter 3

RESULTS

3.1 Primary screen identifies Ezh2 and Dot1L inhibitors as hits

To understand epigenetic modifications that are important in fibroblast to hepatocyte transdifferentiation, we infected mouse embryonic fibroblasts with *Hnf4a* and *Foxa1* vectors, and treated cells with various small molecules (see Table 2.1 for a complete list of chemicals used). These molecules were chosen based on studies that have previously shown their potential to influence gene expression and reprogramming to a pluripotent state [170, 182]. On day 7, relative gene expression levels of the hepatocyte cell marker albumin were determined (Figure 3.1a). With the current published protocol [85], the control DMSO treated cells could only reach 2.1% of the liver albumin mRNA expression levels (Figure 3.1b). Of the chemicals tested, DOT1L inhibitors SGC0946 and EPZ4777, adenylyl cyclase activator forskolin, and EZH2 inhibitor DZNep increased the albumin gene expression level 2-3 fold compared to DMSO treated control cells (Figure 3.1b). Both of the histone deacetylase inhibitors tested, sodium butyrate (NaB) and valproic acid (VPA), decreased the albumin mRNA levels. To assess the effect of the small molecules without the transcription factors, fibroblasts were also treated by chemicals alone for 7 days. The chemicals alone did not dramatically change albumin mRNA levels compared to liver tissue (Figure 3.1c).

We verified these initial results by repeating some of the more significant results in triplicate. Of these, DOT1L inhibitors SGC0946 and EPZ4777, and EZH2 inhibitor DZNep consistently performed better than DMSO treated controls (Figure 3.2). The expression of hepatocyte-specific genes albumin (*Alb*) and alpha-1 antitrypsin (*Aat*), and epithelial marker e-cadherin (*Cdh1*) were about two fold higher in EPZ4777 treated cells. Similarly, DZNep treated iHeps had about three fold higher expression of *Alb* and *Cdh1* compared to DMSO treated controls. Conversely, cytidine analog and DNA

methyltransferase inhibitor 5-AZAC, and histone deacetylase inhibitor valproic acid treated samples had less expression of hepatocyte markers, as in the primary screen.

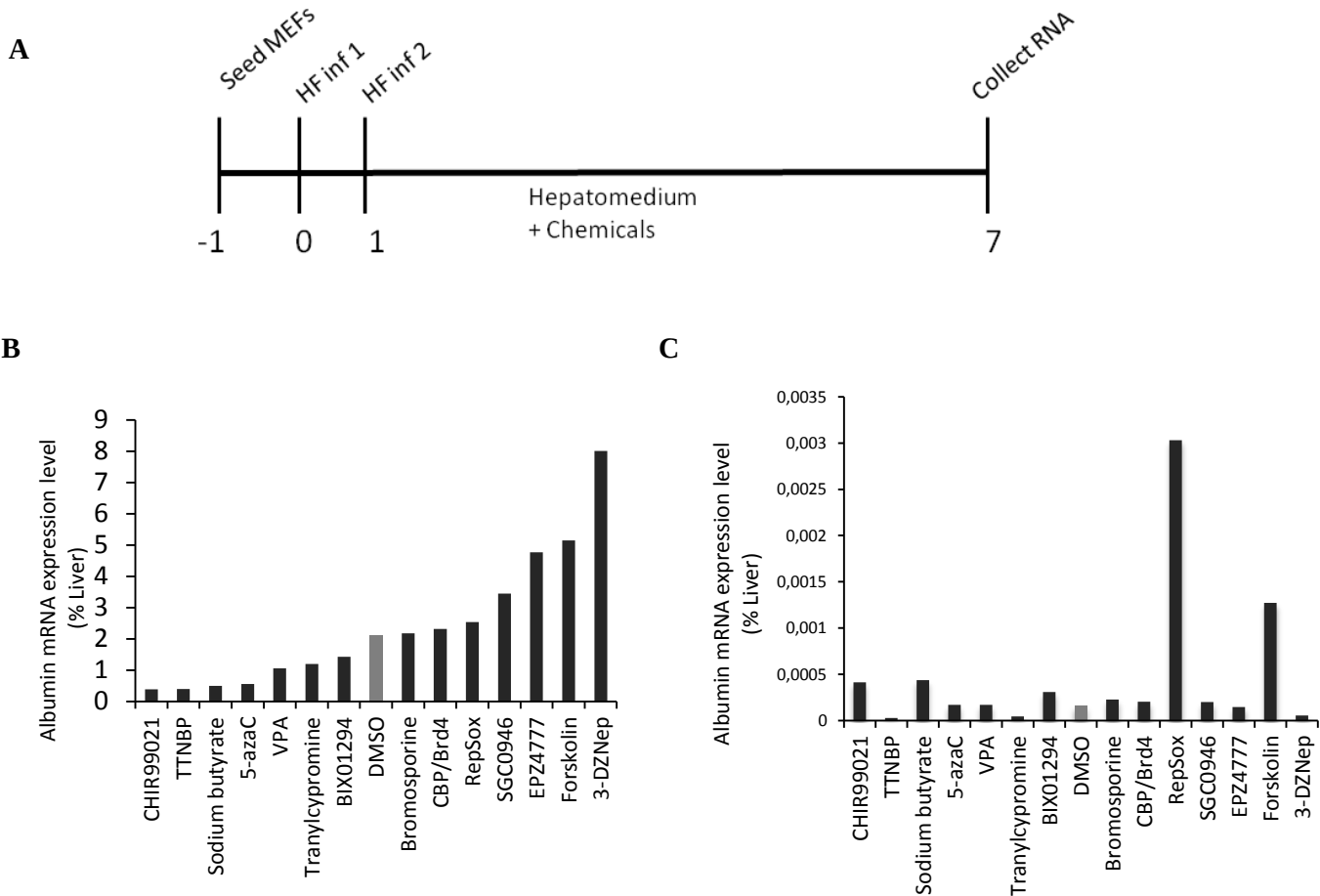


Figure 3.1 Primary screen for small molecule inhibitors
(A) Experimental setup. MEFs were infected with *Hnf4a* and *Foxa1* vectors, and cells were maintained in hepatomedium with the indicated small molecules. RNA was collected on day 7. (B) Relative albumin mRNA expression level of day 7 iHeps treated with different small molecules. mRNA levels are expressed as % liver. (C) Albumin mRNA levels of uninfected mouse embryonic fibroblasts treated with small molecules for 7 days without infection with any transcription factor vectors. mRNA levels are expressed as % liver.

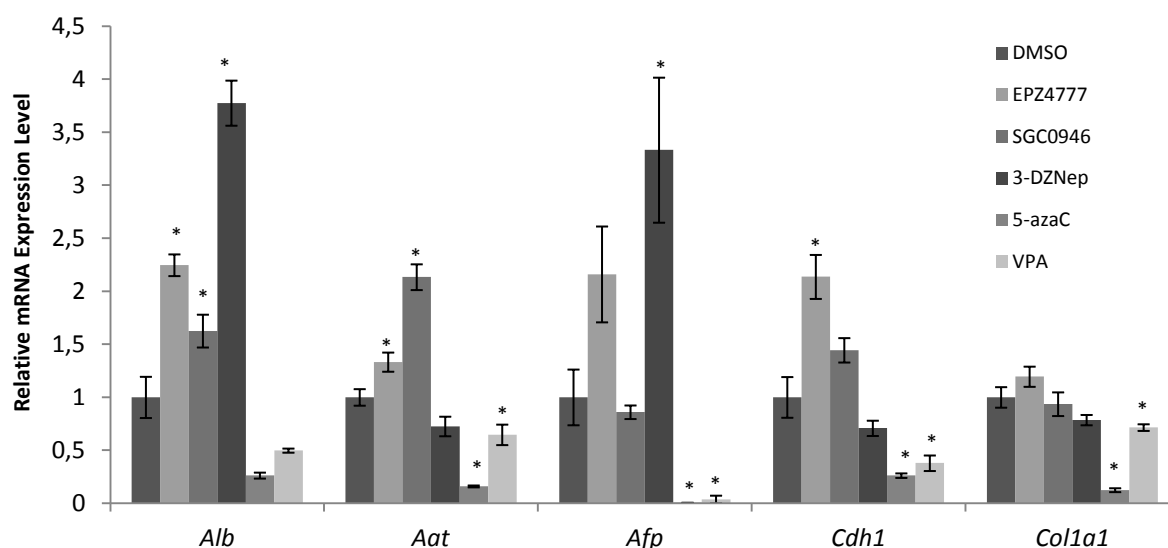


Figure 3.2 Verification experiments for small molecules

Relative mRNA expression levels of specified genes of day 7 iHeps treated with the indicated small molecules. * $p < 0.01$ compared to DMSO treated control iHeps ($n=3$, technical replicates; error bars, $\pm$ s.d.).

3.2 Post-infection chemical treatment is slightly more effective than pre-treatment

We next investigated whether pre-treatment of cells with the small molecules would be more effective than post-infection treatment. We treated MEFs with two of the most effective compounds, EPZ4777 and DZNep, alone or in combination, prior to or following infection with the transcription factor vectors. The gene expression levels of the resulting iHeps were determined on day 7. EPZ4777 treatment following infection led to 4-fold higher expression level of *Alb* and 3-fold higher expression level of *Aat* in the resulting iHeps, compared to pre-treated controls (Figure 3.3). Compared to pre-treatment, DZNep treatment after infection resulted in a 2.5-fold decrease in *Aat* and a 10-fold decrease *Afp* mRNA levels (Figure 3.3). The decrease in *Afp* is in fact the desired outcome, since mature hepatocytes do not express *Afp*. The combination of DZNep and EPZ4777 worked better in the pre-treatment, with *Aat* and *Cdh1* levels being 2.5-fold and 2 fold higher than post-infection treatment, respectively. Double treatment

had trends similar to the DZNep treatment in terms of marker gene expression, suggesting DZNep is the dominant player in this conversion.

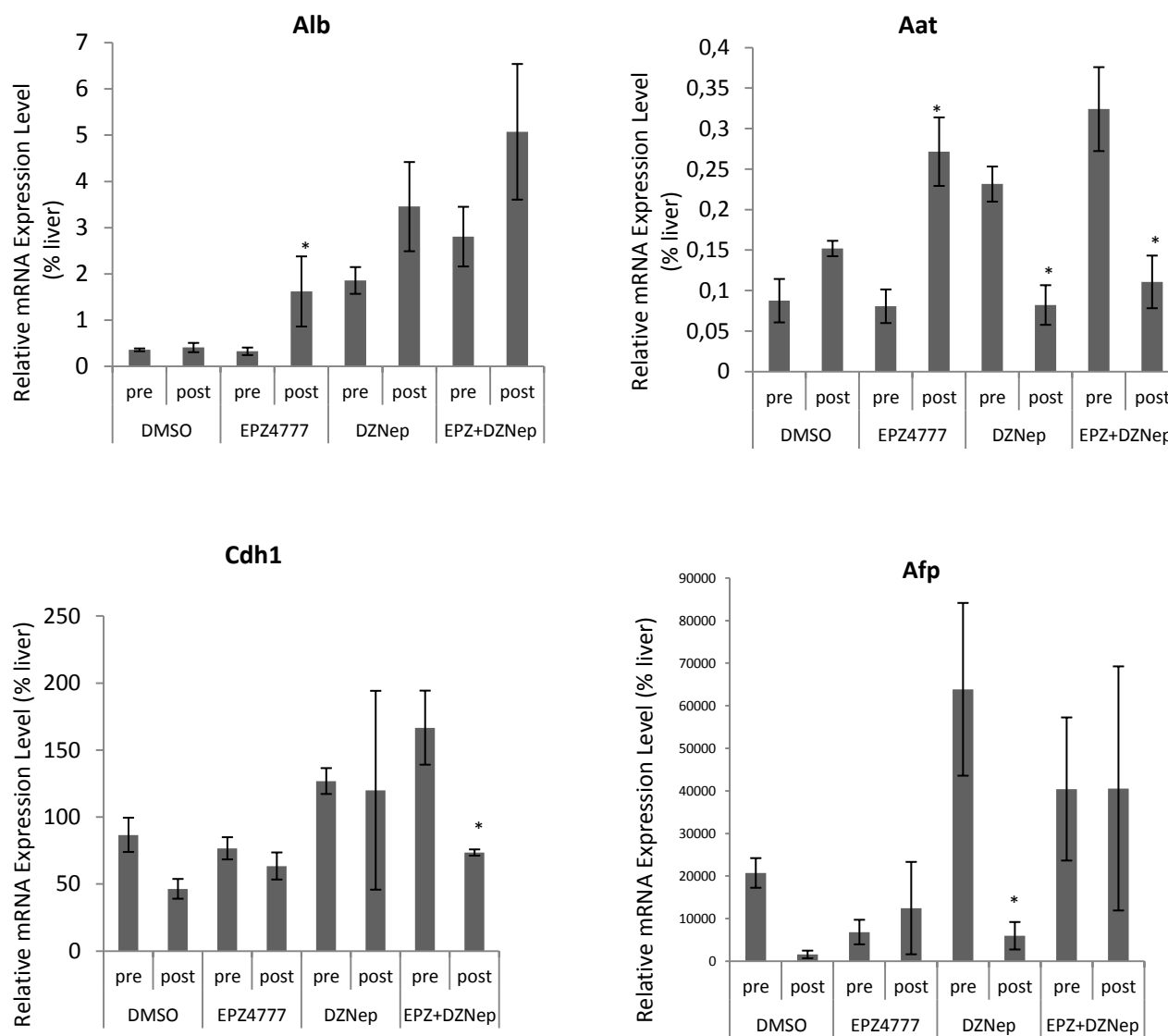
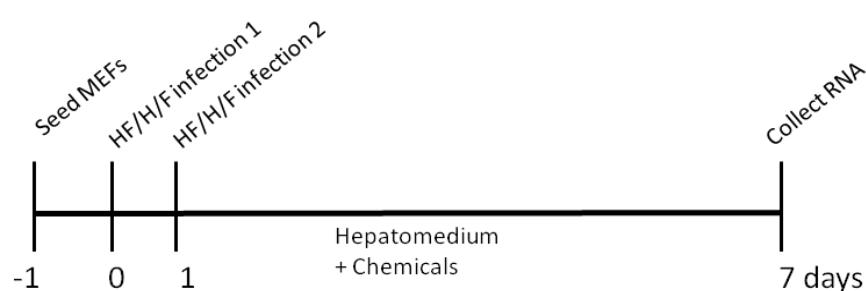


Figure 3.3 Pre-infection vs. post-infection treatment of cells with small molecules. Mouse embryonic fibroblasts were treated with the indicated small molecules for 7 days either before (pre) or after (post) they are infected with *Hnf4a* and *Foxa1* viral vectors. Data were normalized to beta-actin mRNA levels. * $p < 0.01$ compared to pre-treatment control day 7 iHeps (n=3, technical replicates; error bars, $\pm$ s.d.).

3.3 EZH2 and DOT1L inhibitors cannot replace the two transcription factors

To better assess how well EZH2 and DOT1L inhibitors perform in fibroblast to induced hepatocyte conversion, we tested if these two molecules, alone or in combination, can replace any of the two transcription factors (Figure 3.4a). Fibroblasts were infected with *Hnf4a*, *Foxa1*, or *Hnf4a* and *Foxa1* vectors, and treated with EPZ4777, DZNep, or both for 7 days. On day 7, gene expression levels of hepatocyte markers were determined. Chemical treatment of fibroblasts, infected with *Hnf4a* vector only, cannot compensate for FOXA1. These cells only have a maximum of 4% mRNA expression of hepatocyte marker genes compared to the double infected, DMSO treated control iHeps (Figure 3.4b). *Foxa1* infected, EPZ4777 and DZNep treated fibroblasts perform slightly better, reaching ~ 10% of the two factor (HF) infected iHeps' *Alb* and *Aat* mRNA levels (Figure 3.4b). EPZ4777, DZNep, or double treated, and *Foxa1* infected cells reached about 12% of the *Aat* levels of HF infected cells. Infection with *Foxa1* vector, together with DZNep treatment led to an E-cadherin induction of 80%, compared to HF infected controls. Interestingly, relatively high *Cdh1* expression levels were achieved for most samples. HNF4 α alone samples had 40%, and FOXA1 alone samples had 30% of *Cdh1* mRNA levels compared to HF infected controls (Figure 3.4b, lower graph).

A



B

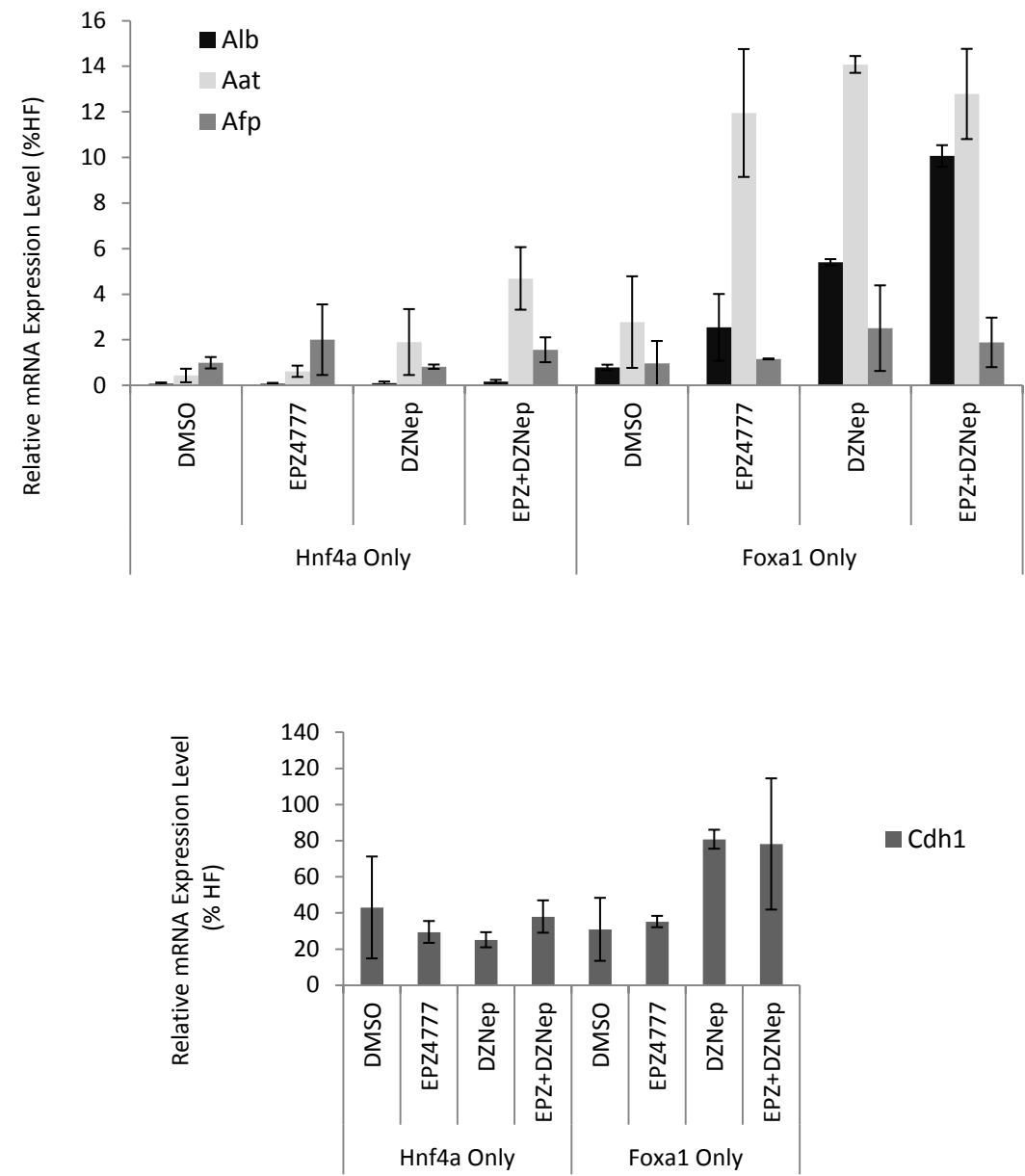


Figure 3.4 EPZ4777 and DZNep cannot replace the two transcription factors
(A) Mouse fibroblasts were infected with the indicated transcription factors and treated with the indicated small molecules for 7 days. (B) Gene expression was analyzed on day 7. Data were normalized to beta-actin mRNA levels.

3.4 Investigating EZH2's effect on fibroblast to hepatocyte conversion

3.4.1 DZNep treatment reduces H3K27 trimethylation

To determine the effect of DZNep treatment on the global levels of H3K27 trimethylation, we treated MEFs with the effective concentration of DZNep used in the primary screen for 5 days. Interestingly, 50nM DZNep did not lead to a significant decrease in the H3K27me3 levels (Figure 3.5). There was only a slight decrease in trimethylation levels, as indicated by the small difference between the loading controls (total histone H3 levels). To ensure that the small molecule was working, we tested the effect of increasing concentrations of DZNep. Using higher concentrations of 250nM and 500nM indeed reduced the K27me3 levels.

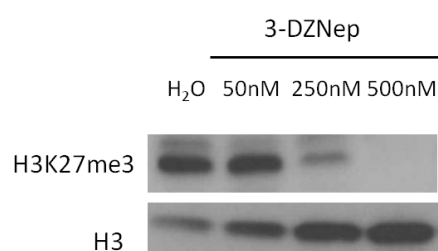


Figure 3.5 DZNep treatment reduces global H3K27 trimethylation levels

Mouse embryonic fibroblasts were treated with varying concentrations of DZNep for 5 days. H3K27me3 levels were determined from histone extracted samples.

3.4.2 *Ezh2* knockdown verifies chemical treatment data

To verify the results obtained by DZNEP treatment, shRNAs were used to knockdown *Ezh2* (Figure 3.6a). MEFs were infected with either sh*Ezh2* or a scrambled shRNA as control, and treated with puromycin for 3 days to select for the cells that were infected with the vector. At the end of the selection, *Ezh2* mRNA levels were tested in the two samples. A knockdown efficiency of 60% was achieved (Figure 3.6b). The remaining cells were reseeded, and infected with *Hnf4a* and *Foxa1* vectors, and maintained in hepatomedium for 7 days. On day 7, gene expression levels of each sample were

determined by qPCR. Compared to control shRNA infected iHeps, *shEzh2* infected iHeps had 3-fold higher expression of hepatocyte marker albumin and 1.5-fold higher expression of epithelial marker e-cadherin, whereas alpha-fetoprotein, a marker for embryonic hepatocytes, had 5-fold lower expression (Figure 3.6b).

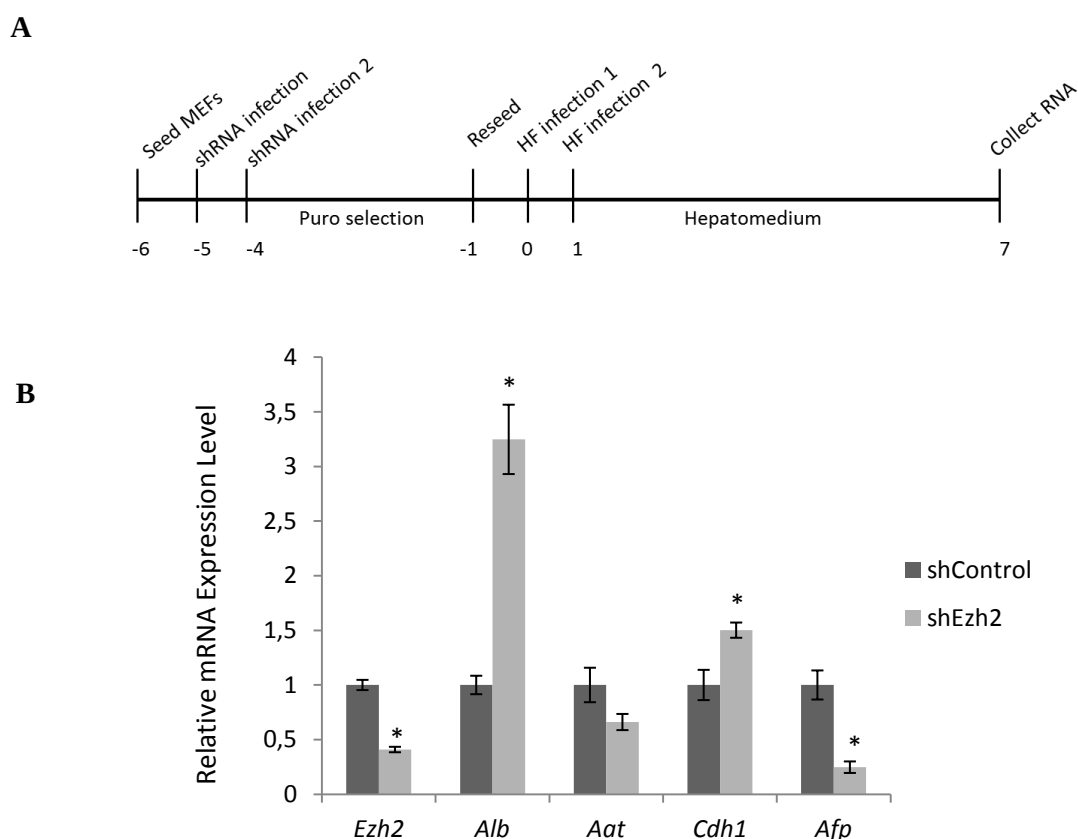


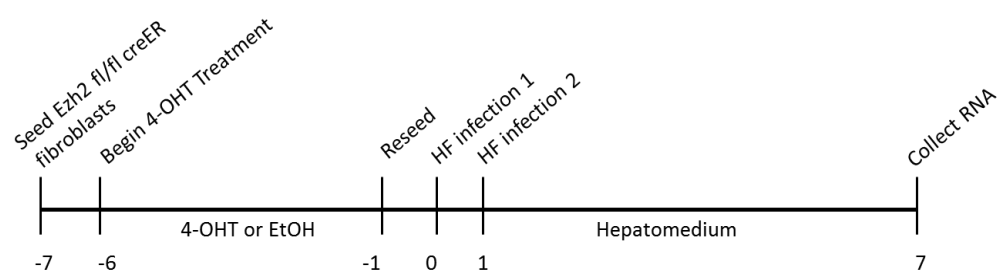
Figure 3.6 iHeps derived from *Ezh2* knock down MEFs show increased expression of mature hepatocyte and epithelial markers.

(A) MEFs infected twice with *shEzh2* or *shControl* viral constructs were selected with puromycin, and infected with HF vectors. (B) Gene expression levels of *Ezh2* and several hepatocyte markers were determined on day 7. Data were normalized to endogenous actin mRNA levels. * $p < 0.01$ compared to control shRNA-expressing iHeps ($n=3$, technical replicates; error bars, $\pm$ s.d.).

3.4.3 *Ezh2* knockout verifies chemical treatment data

We next tested whether *Ezh2* knockout could further improve fibroblast to iHep conversion. We used immortalized *Ezh2* fl/fl cre-ER MEFs treated with active the metabolite of tamoxifen (4-OHT) or the vehicle ethanol to create *Ezh2* knockout or wild type cells (Figure 3.7a). At this point, the H3K27 trimethylation levels were reduced in 4-OHT treated MEFs, as shown by H3K27me3 Western blot (Figure 3.7b). Induced pluripotent stem cells derived from *Ezh2* wild type and knockout MEFs were used as controls in the blot. The *Ezh2* wild type and knockout fibroblasts were then infected with *Hnf4a* and *Foxa1* to induce lineage conversion. On day 7 following infection, gene expression levels of the generated cells were examined. iHeps derived from EZH2 depleted cells exhibited increased expression of hepatocyte marker genes (Figure 3.7c). Specifically, *Alb* mRNA level increased 5-fold, *Aat* mRNA level increased 2-fold, *Hnf1a* mRNA level increased 2.5-fold, *Cdh1* mRNA level increased 4-fold in *Ezh2* knockout iHeps, compared to wild type controls.

A



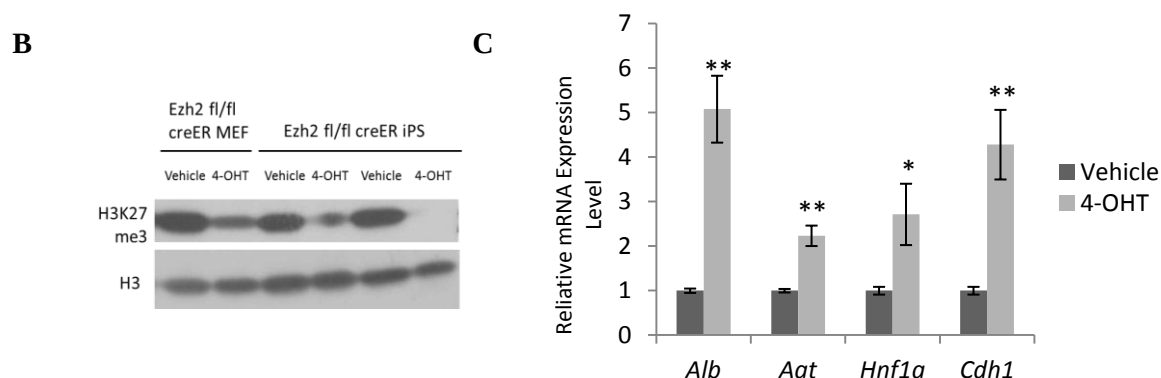


Figure 3.7 iHeps derived from *Ezh2* knockout MEFs show increased expression of mature hepatocyte and epithelial markers.

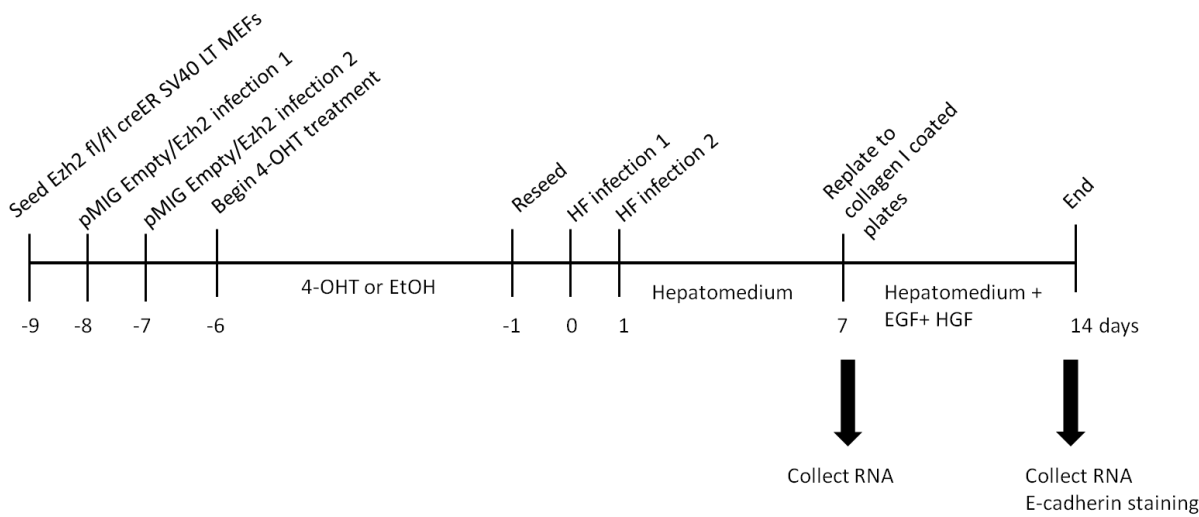
(A) Experimental setup. For knockout, MEFs were derived from *Ezh2* fl/fl mice. MEFs were first infected with a MSCV-CreERT2 vector. Vehicle (ethanol) or 4-OHT was added to cultures for 5 days. (B) 4-OHT treatment reduced H3K27me3 levels. Mouse iPS cells derived from *Ezh2* WT or KO cells were used as control. (C) Relative gene expression levels of key hepatocyte and epithelial markers. Expression levels were normalized to endogenous actin gene expression level. * $p < 0.05$ and ** $p < 0.01$ compared to control vehicle (ethanol) treated samples ($n=3$, technical replicates; error bars, $\pm$ s.d).

3.4.4 *Ezh2* overexpression reverses the knockout effect on fibroblast to hepatocyte conversion

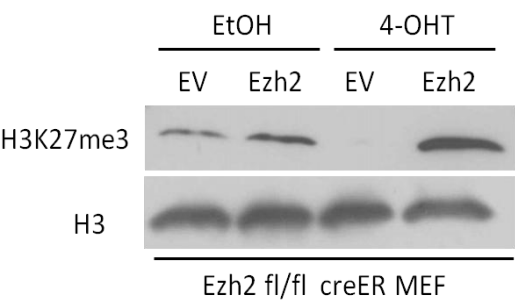
To evaluate if *Ezh2* overexpression could reverse the effects of *Ezh2* knockout, we overexpressed *Ezh2* in *Ezh2* wild type and knockout fibroblasts, and reprogrammed them to induced hepatocytes (Figure 3.8a). These cells were generated by first infecting the *Ezh2* fl/fl cre-ER MEFs with *Ezh2* overexpression or control *Gfp* expression vector, and then treated with 4-OHT or the vehicle ethanol to create *Ezh2* knockout or wild type cells. The H3K27me3 status of each sample was determined by Western blotting prior to reprogramming. 4-OHT treatment depleted trimethylation from H3K27 residues, while *Ezh2* overexpression reversed that phenotype (Figure 3.8b). The cells were then reseeded, and infected with *Hnf4a* and *Foxa1* vectors to induce conversion to iHeps. Cells were maintained in culture for 14 days in hepatomedium. Hepatocyte and epithelial marker

gene expression levels of the resulting cells were analyzed on day 7 and day 14. Similar to the results of the previous knockdown experiment, at the end of day 7, *Ezh2* knockout, *Gfp* overexpressed iHeps (4-OHT + EV) had significantly higher levels of *Alb*, *Aat*, and *Cdh1* compared to *Ezh2* wild type, *Gfp* overexpressed iHeps (EtOH + EV) (Figure 3.8b). By day 14, there was little difference in marker gene expression between these two samples, and was not significant (Figure 3.8d). The overexpression of *Ezh2* in *Ezh2* wild type cells (EtOH + EZH2) did not change the marker gene expression levels compared to *Ezh2* wild type, *Gfp* overexpressed cells, as assessed both on day 7 and day 14. One exception to this was *Afp* mRNA levels at the end of day 7, with an increase of about 25-fold in *Ezh2* overexpressed samples (Figure 3.8c). *Ezh2* overexpression in *Ezh2* knockout cells reversed the effect of *Ezh2* depletion, shown by the 3-fold decrease in *Alb* expression in *Ezh2* overexpressed cells compared to controls on day 7 (Figure 3.8c). The decrease was less prominent, but still significant for *Aat* and *Cdh1* mRNA levels, with a 1.8- and 2-fold decrease, respectively. Interestingly, this effect was much less apparent by day 14 (Figure 3.8d), with only 1.2 fold difference remaining in *Alb* mRNA levels among the two samples.

A



B



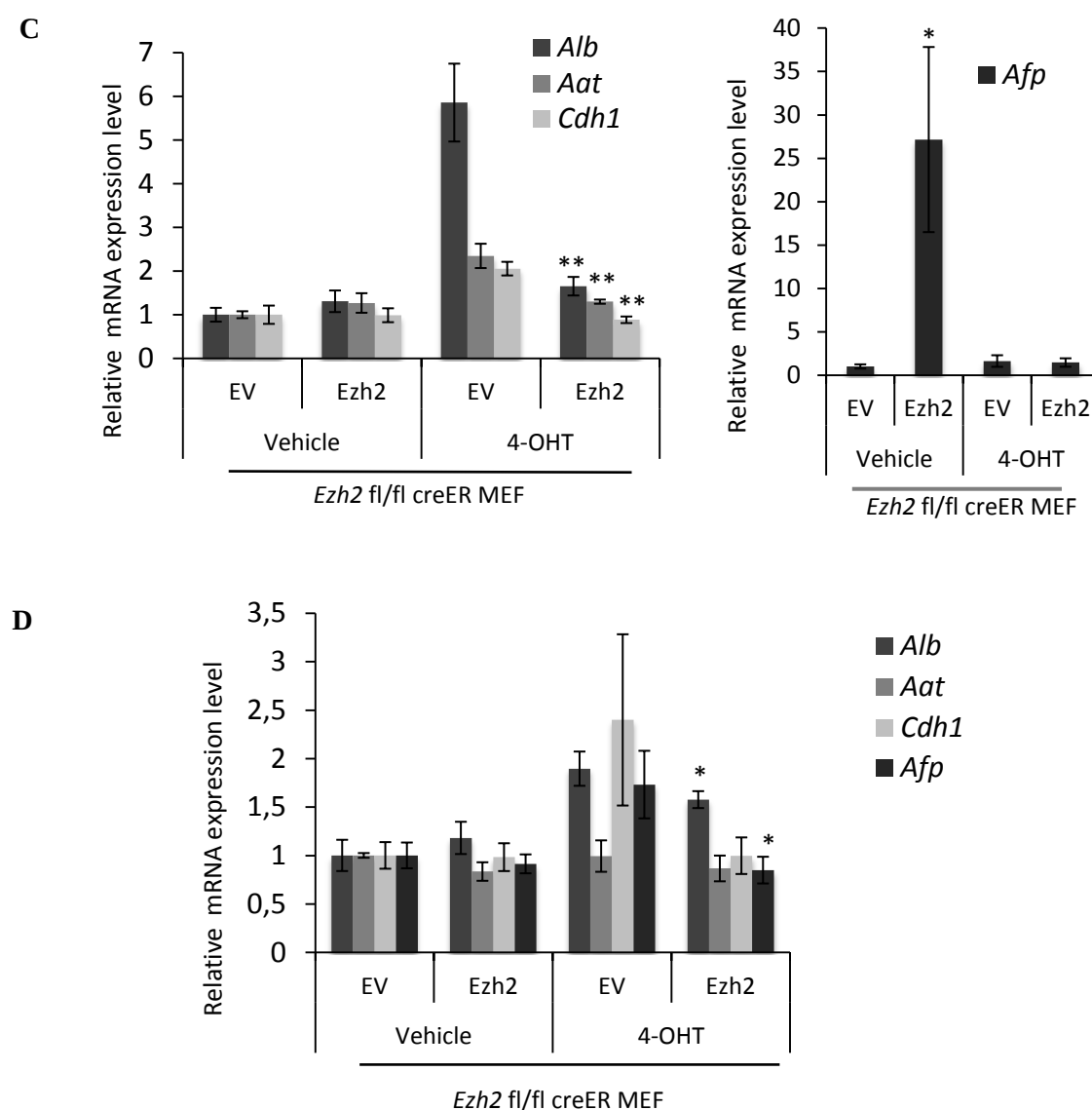


Figure 3.8 Ezh2 overexpression reverses the effect of Ezh2 knockout

(A) Experimental setup. MEFs derived from *Ezh2* fl/fl mice and infected with a MSCV-CreERT2 vector were infected with pMIG *Ezh2* or pMIG IRES GFP vector. Vehicle (ethanol) or 4-OHT was added to cultures for 5 days. (B) 4-OHT treatment reduced, whereas *Ezh2* overexpression restored H3K27me3 levels. (C,D) Relative gene expression levels of key hepatocyte and epithelial markers, normalized to endogenous actin mRNA levels at day 7 (C) and day 14 (D). * $p < 0.05$ and ** $p < 0.01$ compared to control empty vector infected samples within same genetic background ($n=3$, technical replicates; error bars, $\pm$ s.d.).

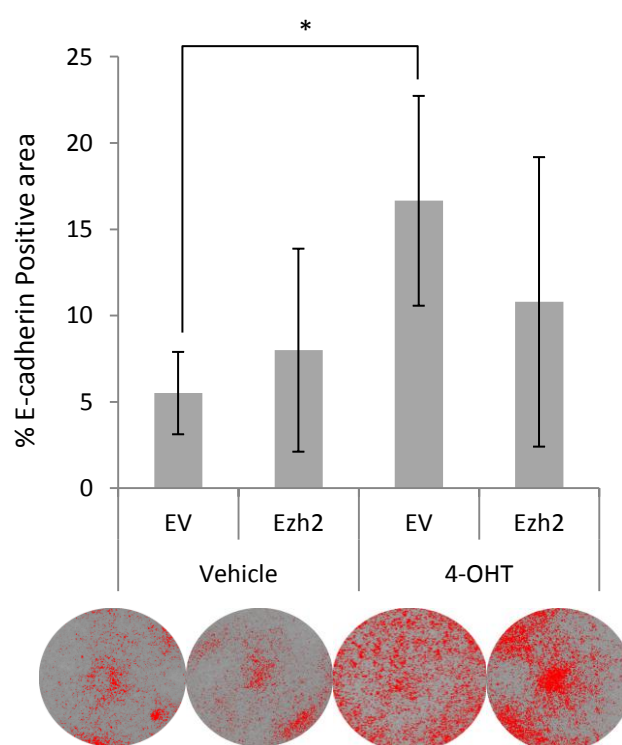


Figure 3.9 *Ezh2* knockout MEFs generates more E-cadherin positive iHep colonies

MEFs were derived from *Ezh2* fl/fl mice were infected with a MSCV-CreERT2 vector, and vehicle (ethanol) or 4-OHT was added to cultures for 5 days. Cells were stained for E-cadherin antibody on day 14. Representative images of wells with colonies marked in red by ImageJ are shown. *p<0.05 compared to control vehicle (ethanol) treated, empty vector infected samples (n=3, technical replicates; error bars, ± s.d.).

We next investigated whether the increase in marker gene expressions correlates with the number of induced hepatocyte cells. We analyzed E-cadherin expression at the protein level for *Ezh2* wild type and knockout iHeps with or without exogenous *Ezh2* overexpression, same as in the experimental setup described in Figure 3.8a. In control empty vector (EV) infected samples, *Ezh2* knock-out wells had three-fold more E-cadherin positive cells than *Ezh2* wild type wells (Figure 3.9). This verified that *Ezh2* knockout results in significantly more E-cadherin positive cells. *Ezh2* overexpression

partially reversed this phenotype; EV infected, *Ezh2* wild type cells and *Ezh2* overexpressed, *Ezh2* knockout iHeps were not significantly different. Moreover, in the *Ezh2* knockout samples, the *Ezh2* overexpression and empty vector overexpression were not significantly different.

3.4.5 H3K27 residues are less tri-methylated on hepatocyte-specific gene promoters in iHeps

To investigate the cause of the increase in hepatocyte-specific genes in *Ezh2*-inhibited iHeps, the presence of the H3K27me3 mark on promoters of hepatocyte specific genes were determined by chromatin immunoprecipitation (ChIP). For this experiment, uninfected fibroblasts and *Hnf4a* and *Foxa1* infected day 7 were compared. DNA-bound proteins were crosslinked to the chromatin with formaldehyde, and DNA was fragmented by sonication. H3K27me3 marked DNA stretches were pulled down with an H3K27me3 antibody and magnetic beads with high affinity to the antibody. The DNA pieces eluted from the beads were reverse crosslinked from the protein. Control ChIPs were performed with normal IgG. The pulled down DNA was analyzed for the presence of the promoter regions of hepatocyte-specific genes by qPCR. The results demonstrate that the induction lineage conversion results in loss of H3K27 methylation from the endogenous promoters of *Alb*, *Hnf1a* and *Hnf4a* by 6-, 5- and 2.5- fold, respectively (Figure 3.10). This finding is consistent with the notion that lineage conversion is accompanied by the de-repression of PRC2-silenced hepatocyte specific promoters in MEFs. Furthermore, this finding suggests that inhibition of PRC2 either through small molecule inhibitor DZNep or through loss of *Ezh2* expression might facilitate the expression of hepatic genes associated with lineage conversion, by contributing to the de-repression of hepatocyte specific promoters.

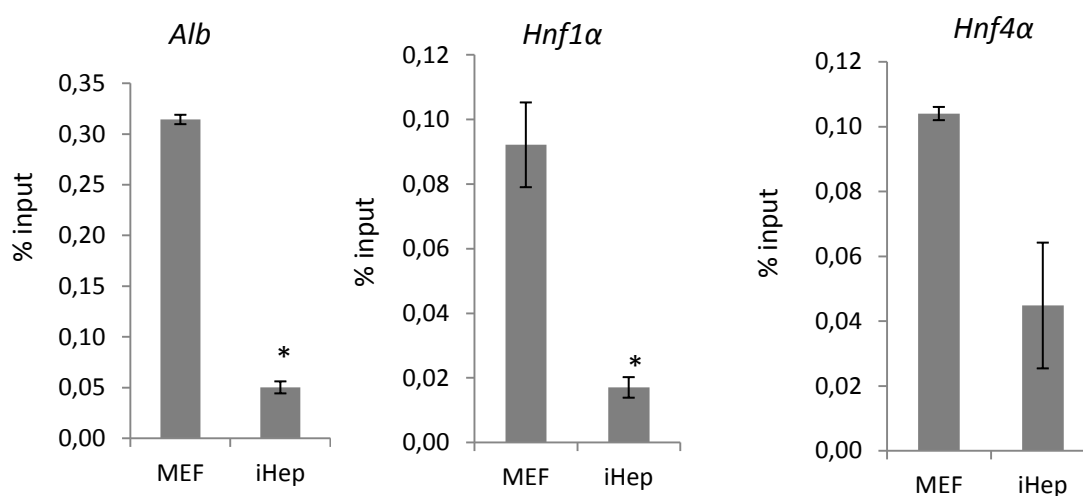


Figure 3.10 H3K27 tri-methylation is reduced on hepatocytes specific gene promoters in iHeps.

Day 7 iHep and uninfected fibroblast chromatin was pulled down with H3K27me3 antibody. The presence of individual genes was determined by qPCR. Pull downs with normal IgG were performed as control. * $p < 0.05$ compared to uninfected fibroblasts.

3.5 Investigating DOT1L's effect on fibroblast to hepatocyte conversion

3.5.1 EPZ4777 treatment reduces H3K79 di-methylation

To ensure that the small molecule inhibitor EPZ4777 treatment is inhibiting DOT1L, the global H3K79 methylation levels were assessed by Western blotting. Five day treatment of mouse fibroblasts with EPZ4777 reduced, although did not completely abolish the di-methylated H3K79 residues (Figure 3.11). iPS cells derived from *Dot1l* fl/fl MEFs treated with vehicle (ethanol) or tamoxifen were used as controls.

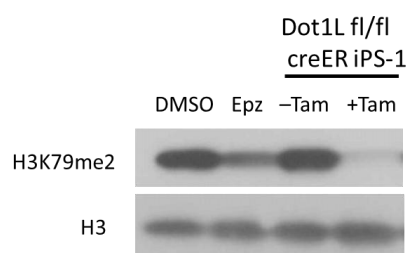


Figure 3.11 EPZ4777 reduces global H3K79me2 levels

Mouse embryonic fibroblasts were treated with vehicle (DMSO) or EPZ4777 for 5 days. H3K79me2 levels were determined by Western blotting.

3.5.2 *Dot1l* knockout verifies chemical treatment data

To investigate whether genetic knockout of *Dot1l* can verify the chemical treatment data, *Dot1l* fl/fl cre-ERT2 tail tip fibroblasts (TTFs) were treated with the active metabolite of tamoxifen (4-OHT) or the vehicle (EtOH) for 5 days, and then infected with *Hnf4a* and *Foxa1* transcription factors to induce fibroblast to hepatocyte conversion (Figure 3.12a). Prior to the start of reprogramming, the H3K79me2 status of each sample was determined by Western blotting. iPS cells derived from *Dot1l* fl/fl MEFs treated with vehicle (ethanol) or tamoxifen were used as controls. 4-OHT treated samples had decreased levels of H3K79 di-methylation, as expected (Figure 3.12b). 7 days after infection with the transcription factors, the gene expression levels of hepatocyte and epithelial cell markers were determined by qPCR. 4-OHT treated samples had 2-fold increased *Alb*, 3-fold increased *Aat*, 2-fold increased *Afp*, and 3.5-fold increased *Cdh1* mRNA levels compared to vehicle treated controls (Figure 3.12c). *Hnf1a* expression levels were not significantly different.

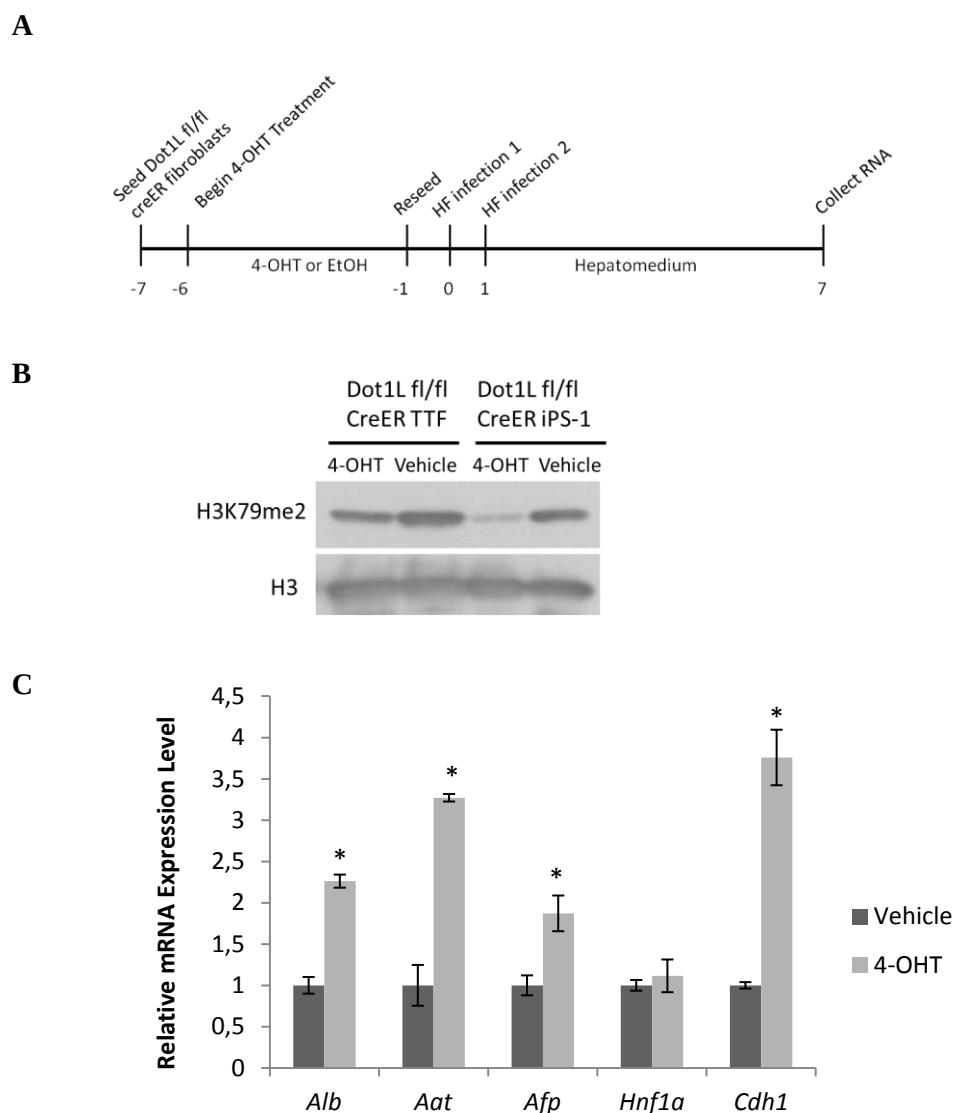


Figure 3.12 iHeps derived from *Dot1l* knockout TTFs more robustly express hepatocyte markers.

(A) Experimental setup. TTFs were derived from *Dot1l* fl/fl mice and infected with a MSCV-CreERT2 vector. Vehicle (ethanol) or 4-OHT was added to cultures for 5 days. (B) 4-OHT treated TTFs have reduced H3K79me2 levels. iPS cells derived from *Dot1l* fl/fl MEFs treated with vehicle (ethanol) or tamoxifen were used as controls. (C) *Dot1l* knockout and wild type TTFs were converted to iHeps, and marker gene expressions were analyzed on day 7. Expression levels were normalized to endogenous actin gene expression level. * $p < 0.01$ compared to control vehicle (ethanol) treated samples ($n=3$, technical replicates; error bars, $\pm$ s.d.).

3.5.3 DOT1L inhibition does not reduce the expression of fibroblast-specific genes

Removal of the activating K79me2 mark from H3 histones may help silencing of fibroblast specific genes, thereby facilitating lineage conversion. We tested this idea by determining the gene expression levels of fibroblast specific genes in iHeps. Mouse fibroblasts infected with *Hnf4a* and *Foxa1* vectors, and treated with DMSO or EPZ4777 for 7 days were maintained in culture for three weeks. mRNA levels of genes enriched in fibroblasts were determined on day 7, 14 and 21. EPZ4777 treated iHeps did not significantly differ from DMSO treated controls in terms of vimentin (*Vim*), collagen (*Col1a1*), alpha-actin-2 (*Acta2*), fibulin-2 (*Fbln2*) and S100 calcium-binding protein A4 (*S100a4*) mRNA levels for all time points tested (Figure 3.13). Only procollagen C-endopeptidase enhancer (*Pcolce*) mRNA levels of day 7 EPZ4777 treated iHeps were 1.75-fold less compared to DMSO treated controls. However, this difference disappeared on day 14 and 21. Notably, most of the fibroblast marker gene expression levels increased in both DMSO and EPZ4777 treated iHeps in later weeks.

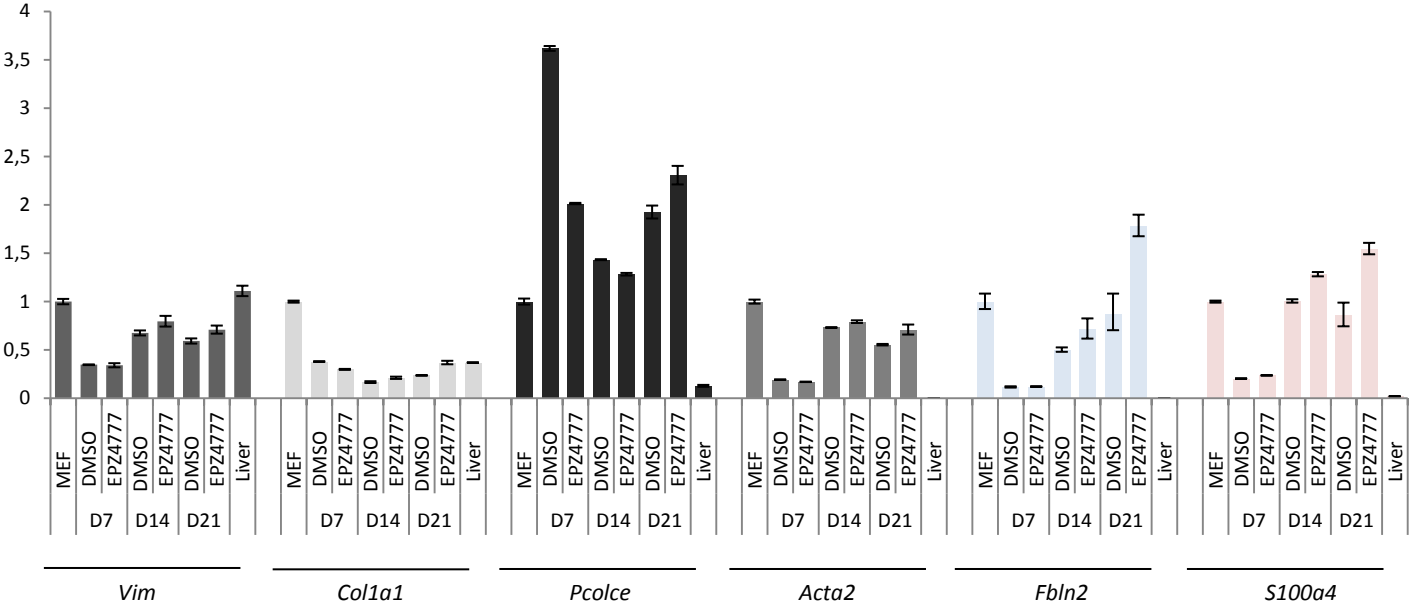


Figure 3.13 Fibroblast specific genes are expressed in similar levels in DOT1L inhibited and control iHeps

MEFs infected with *Hnf4a* and *Foxa1* viral vectors were maintained in hepatomedium and treated with the indicated small molecules for 7 days. Gene expression levels are determined on day 7, 14 and 21. Expression levels were normalized to endogenous actin gene expression level.

3.5.4 DOT1L inhibition does not reduce the aberrant other lineage-specific gene expression

One other possible mechanism through which DOT1L inhibition facilitates *Hnf4a* and *Foxa1* induced changes could be the repression of other lineage-specific genes. It was shown that iHeps aberrantly express the intestinal marker *Cdx2* [108]. Removal of the activating H3K79me2 mark from the regulatory regions of this gene could be aiding the transdifferentiation process. To test this idea, mouse fibroblasts were infected with *Hnf4a* and *Foxa1* viral vectors, and treated for 7 days with EPZ4777 or DMSO as control. Cells were maintained 14 days in culture. The levels of *Cdx2* mRNA were determined on days 7 and 14. However, the *Cdx2* levels of EPZ4777 treated iHeps were similar to DMSO treated iHeps, on both day 7 and day 14 (Figure 3.14).

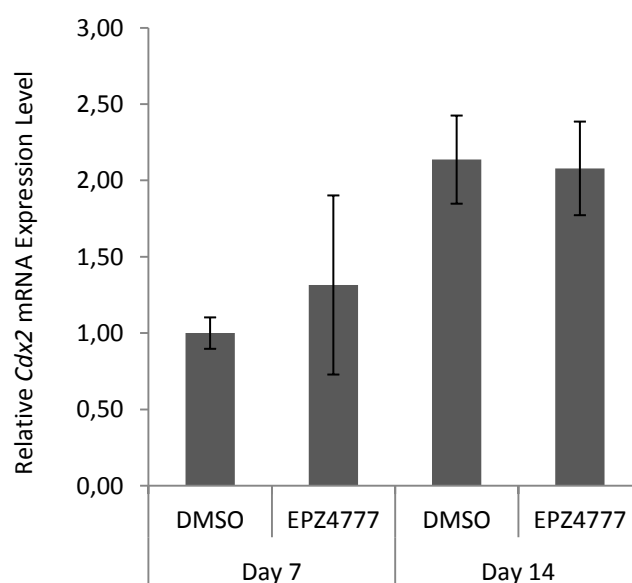


Figure 3.14 DOT1L inhibition does not reduce the expression of the intestinal marker *Cdx2*

MEFs infected with *Hnf4a* and *Foxa1* viral vectors were maintained in hepatomedium and treated with the indicated small molecules for 7 days. Gene expression levels are determined on day 7 and 14. Expression levels were normalized to endogenous actin gene expression level. (n=3, technical replicates; error bars, $\pm$ s.d.).

3.5.5 *Dot1l* knockout does not dramatically affect the expression of EMT Genes

One possible mechanism through which EPZ4777 treatment might act is suppression of epithelial–mesenchymal transition inducers. To test this, we examined the gene expression levels of direct or indirect EMT inducers. In contrast to our expectations, we observed an increase in the expression of *Snai1*, a repressor of E-cadherin (Figure 3.15). The expression of other E-cadherin repressors were similar in *Dot1l* knockout iHeps compared to wild type, or only slight lower.

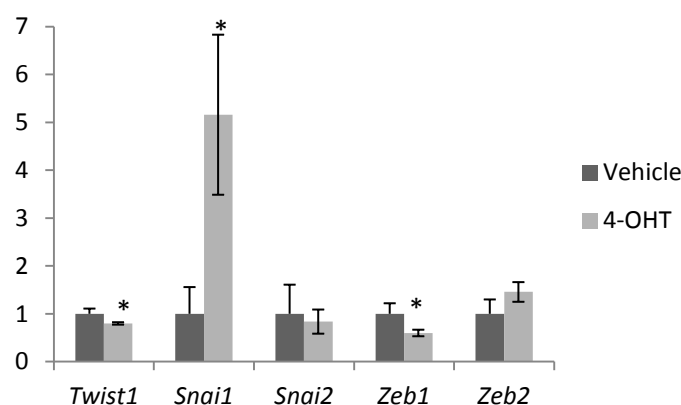


Figure 3.15 *Dot1l* knockout and wild type iHeps have similar EMT gene expression levels

Gene expression levels of epithelial-mesenchymal transition (EMT) inducers of day 7 *Dot1l* WT and KO iHeps were determined. TTFs were derived from *Dot1l* fl/fl mice and infected with a MSCV-CreERT2 vector were treated with vehicle (ethanol) or 4-OHT for 5 days. Expression levels were normalized to endogenous actin gene expression level. * $p < 0.05$ compared to control vehicle (ethanol) treated samples ($n=3$, technical replicates; error bars, $\pm$ s.d.).

Chapter 4

DISCUSSION

With this study, we showed that inhibition of histone methyltransferase EZH2 and DOT1L result in increased expression of hepatic genes associated with lineage conversion carried out by *Hnf4a* and *Foxa1*. Inhibition by small molecules is not sufficient to replace any of the transcription factors.

4.1 Possible mechanisms of Ezh2 inhibition on transdifferentiation

The histone mark maintained by EZH2 of the PRC2 complex is the tri-methylation of lysine 27 of histone H3. This modified residue is recognized by the PRC1 complex, and leads to the mono-ubiquitylation of H2AK119, which compacts the nucleosome arrays to inhibit transcription. Hepatocyte specific genes, especially their promoter regions are H3K27 methylated in mouse fibroblasts, whereas these marks are absent in induced hepatocytes (Figure 3.10). Treatment of cells with an EZH2 inhibitor helps remove this mark. Relaxing of chromatin may be aiding the conversion process, despite not causing a long lasting effect. This is observed by the experiment where *Ezh2* wild type and knockout MEFs are converted to iHeps; although *Ezh2* deficient iHeps seem to be better expressing hepatocyte markers at day 7 (Figure 3.8c), this effect is largely lost by the end of day 14 (Figure 3.8d). The reason for the loss of this effect could be due to FOXA1 being a pioneer transcription factor that can bind to compact chromatin. Since it can already bind compact chromatin on its own, the relaxing chromatin would only speed up the process. Eventually, the untreated samples would also express hepatocyte markers

Another important aspect to consider is the possible H3K27me3 independent functions of EZH2 in direct lineage conversion. EZH2 methylates proteins other than histones to regulate their function, such as STAT3 [183]. Methylation of STAT3 leads to increased tyrosine phosphorylation and activates STAT3. A number of groups have also reported

that EZH2 acts as a transcription activator. For example, EZH2 interacts with β -catenin to increase its accumulation in the nucleuse and facilitate its activation [184]. Moreover, EZH2 interaction with β -catenin, together with estrogen receptor α , regulates the expression of cyclin D1 in breast cancer in a methylation independent manner [185]. It is important to be aware of such H3K27me3 independent functions of EZH2, since they might also play a role in fibroblast to hepatocyte conversion. One set of experiments to address this question could be to express a mutant *Ezh2* that cannot be recruited to histones in an *Ezh2* knockout background, and observe whether the effects of the knockout are reversed. The reversal of increase in hepatocyte marker expression would implicate the methylation independent roles of EZH2 in the conversion process.

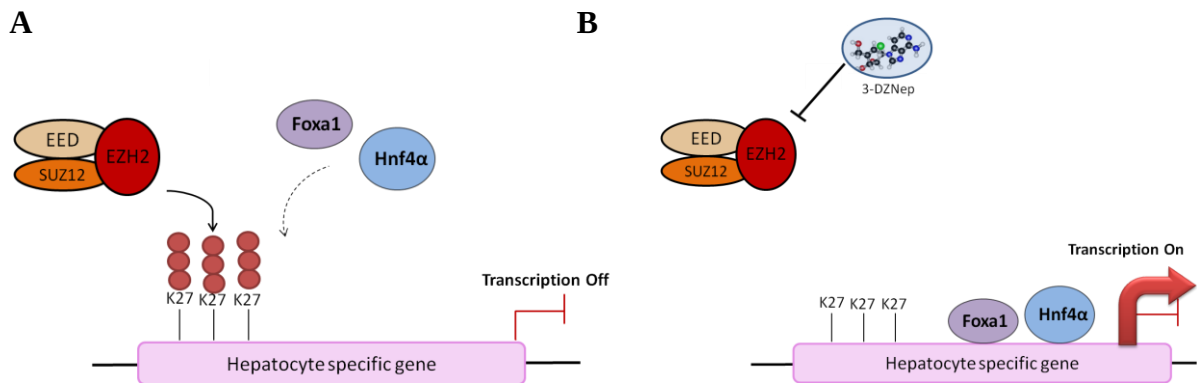


Figure 4.1 Model of Ezh2's role in conversion of mouse fibroblasts to iHeps

(A) In fibroblasts, regulatory regions of hepatocyte-specific genes H3K27 tri-methylated. When *Hnf4 α* and *Foxa1* are exogenously expressed, the transcription factors can weakly bind to and initiate the transcription of these hepatocyte-specific genes. (B) Upon inhibition of EZH2 with 3-DZNep, the PRC2 complex is inactivated, and H3K27me3 marks are lost. This enables stronger binding of HNF4 α and FOXA1 to their targets.

4.2 Possible mechanisms of Dot1L inhibition on transdifferentiation

DOT1L maintains the activating H3K79 methylation mark. One possible role of inhibition of DOT1L during fibroblast to hepatocyte conversion could be that it allows better silencing of fibroblast specific genes. However, this turned out not to be the case (Figure 3.13). We tested this by looking at the gene expression levels of fibroblast specific genes that were expressed in low levels in primary hepatocytes, but were aberrantly expressed in iHeps, as shown in a microarray by Sekiya and Suzuki [85]. We compared EPZ4777 treated vs. untreated samples, and found that DOT1L inhibition did not significantly alter their levels (Figure 3.13). Therefore, DOT1L inhibition's effect was not through downregulation of fibroblast specific genes.

One shortcoming of induced hepatocytes is the irregular expression of the intestinal cell marker *Cdx2*, which is absent from fibroblasts and hepatocytes [108]. Previous studies have shown that the knockdown of *Cdx2* by shRNAs can increase the hepatocyte marker gene expression [108]. We tested whether inhibition of DOT1L could be removing the activating di-methylation mark from H3K79 in the regulatory regions of these genes, and facilitating transdifferentiation. However, we did not observe a difference in *Cdx2* expression between EPZ4777 treated and DMSO treated control iHeps. This suggests there is some other mechanism through which DOT1L inhibition operates.

Fibroblast to induced hepatocyte conversion is a mesenchymal to epithelial transition. We therefore explored whether reduction of the activating H3K79 methylation marks globally reduced the activating marks of EMT inducers. The direct repressor of E-cadherin *Zeb1*, and the indirect repressor of E-cadherin *Twist1* appear only slightly downregulated, whereas the *Snai2* and *Zeb2* levels remain unchanged (Figure 3.15). Unexpectedly, the levels of *Snai1*, another direct repressor that binds to the E-cadherin promoter, is higher in *Dot1l* knockout iHeps.

None of the hypotheses tested regarding the mechanism of DOT1L inhibition in reprogramming of mouse fibroblasts to induced hepatocytes were supported by our findings. An alternative potential mechanism could be that inhibition of DOT1L removes activating marks from an inhibitor of the hepatic lineage, thus aiding in transdifferentiation. For example, it is known that transcription factor hematopoietically expressed homeobox (*Hhex*), induces hepatic specification through the repression of eomesodermin (*Eomes*) expression, a regulator of definitive endoderm [186]. Therefore, *EOMES* is an inhibitor of hepatic lineage specification. Removal of H3K79me2 mark from this gene may be facilitating fibroblast to iHep conversion. To test this potential, gene expression level of *Eomes* in DMSO and EPZ4777 treated day 7 iHeps could be compared. This approach could also be extended to other known regulators of hepatic fate.

4.3 Testing Chromatin Modifiers for the Regulation of Cell Conversion in Other Contexts

Since polycomb repressive complexes are known to be important regulators of many aspects of normal development, it is plausible to think that they may also regulate cell conversions in other contexts. To test this, we utilized the conversion of mouse fibroblasts to induced macrophages through the forced expression of the transcription factors PU.1 and C/EBP α [69]. Mouse fibroblasts were infected with the two transcription factors and treated with EPZ4777 or DZNep for 7 days. On day 7, the efficiency of conversion was determined by quantifying the percentage of cells that are positive for the macrophage marker MAC1. We found that EPZ4777 moderately increases, and DZNep decreases the percentage of MAC1⁺ cells. (See Appendix C, Figure C.1). Therefore, it is possible that the same chromatin modifications regulate different types of lineage conversions in a different manner. From this set of experiments, repressing H3K27me3 appears to be a required mark for conversion of fibroblasts to induced macrophages, and inhibition of the activating H3K79me2 mark is a facilitator of the same conversion.

4.4 Future Directions

This study mostly relies on the expression levels of hepatocyte and fibroblast markers to assess the effects of small molecule inhibitors of chromatin modifiers. A more indicative assay would be to look at a marker at the protein level. In their paper, Sekiya and Suzuki estimate the efficiency of lineage conversion by % E-cadherin-positive cell clusters. With this method, the efficiency of conversion is about 0.1% [85]. The choice of this marker is most likely because of its easier detection on the cell surface. However, this approach relies on an epithelial cell marker, rather than a hepatocyte marker. Therefore, a better indicator could be albumin, or transcription factor HNF1 α .

While gene expression profiles and marker protein expression are important indicators, functional data on the performance of the generated iHeps would be more informative. Albumin secretion, LDL uptake, indocyanine green uptake and release, and drug metabolism of the generated hepatocytes would be some of the key comparisons to normal hepatocyte function. Ultimately, one of the more stringent tests of functionality would be to assess the functional reconstitution capabilities of iHeps by injecting them in to the livers of Fah deficient mice. However, since some of the initial advantages of EZH2 inhibition on hepatocyte marker gene expression is compensated by untreated controls by the end of the second week, functional performance may not be significantly different.

Another possible test of the quality of the induced hepatocytes could be through the analysis of the gene regulatory networks (GRN) of iHeps. RNA-seq data of iHep cells treated with EPZ4777 or DZNep could be used to extract the active GRNs using CellNet, the tool that originally analyzed iHep GRNs to find they are not very similar to mature hepatocytes [108]. This could determine the performance of the engineered cells in a more systematic way.

A deeper look at the dynamics of EZH2 inhibition in transdifferentiation would be through comparison of HNF4 α and FOXA1 binding to their target genes in *Ezh2* wild type and knockout cells. The effect of EZH2 inhibition is best observed on day 7 of reprogramming. Chromatin immunoprecipitation of reprogramming *Ezh2* wild type and knockout fibroblasts with HNF4 α and FOXA1 antibodies, and looking at their binding sites would be informative. Comparison of day 3 and day 7 of the conversion would also reveal the time of the removal of the tri-methylation mark.

In summary, this study investigates the epigenetic changes during mouse fibroblast to hepatocyte conversion. Our results reveal removal of the inhibitory H3K27 tri-methylation mark by inhibition of EZH2 and removal of the activating H3K79 methylation mark by inhibition of DOT1L results in increased expression of hepatic genes associated with lineage conversion.

APPENDIX

Appendix A: Verification Experiments for Induced Hepatocytes

To verify that the conversion protocol of mouse fibroblasts to induced hepatocytes were working as described, the initial population of fibroblasts and the generated induced hepatocytes were analyzed. Only 14 days following infection with the transcription factors, the cells gained epithelial morphology (Figure A.1).

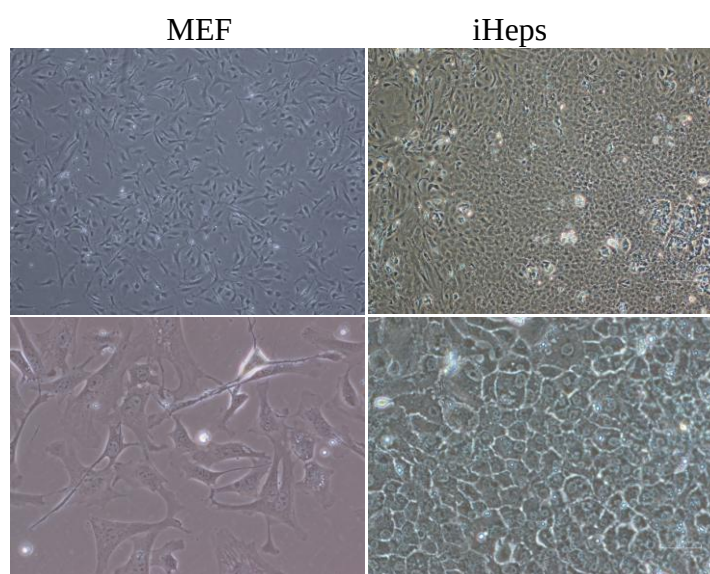


Figure A.1 Bright field microscopy images of fibroblasts and induced hepatocytes.

The lower panels show the same cells with higher magnification.

The cells were also characterized for their expression of vimentin and E-cadherin by immunofluorescence microscopy. Uninfected fibroblasts highly express the intermediate filament vimentin, but lack the epithelial cell marker E-cadherin (Figure A.2). In day 7 iHeps, the level of vimentin is reduced, and cells are positive for E-cadherin and albumin (Figure A.3). iHeps are also able to store glycogen, as assessed by periodic acid–Schiff (PAS) staining (Figure A.4).

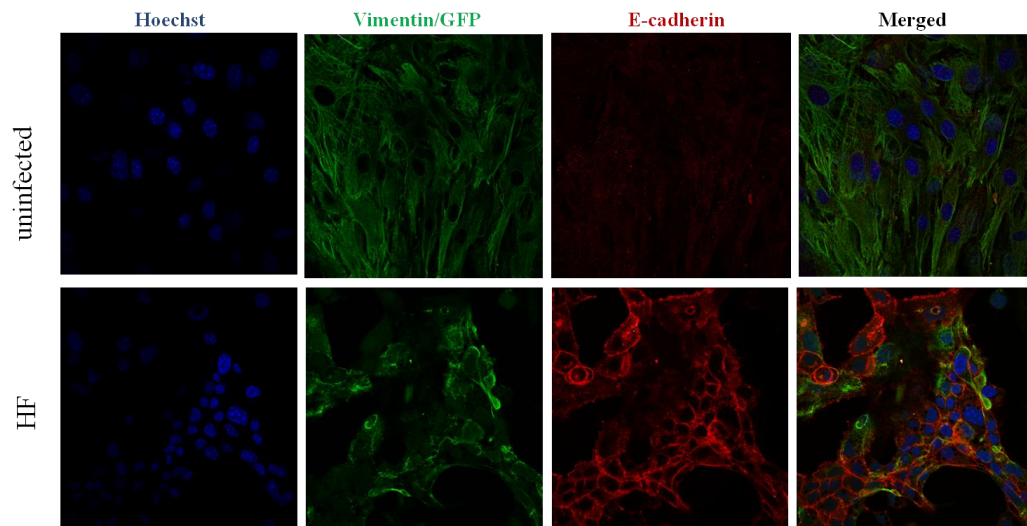


Figure A.2 Induced hepatocytes express E-cadherin at the protein level. Uninfected fibroblasts (top panel) and day 7 *Hnf4α* and *Foxa1* expressing iHeps (bottom panel) were stained for vimentin (green) and E-cadherin (red).

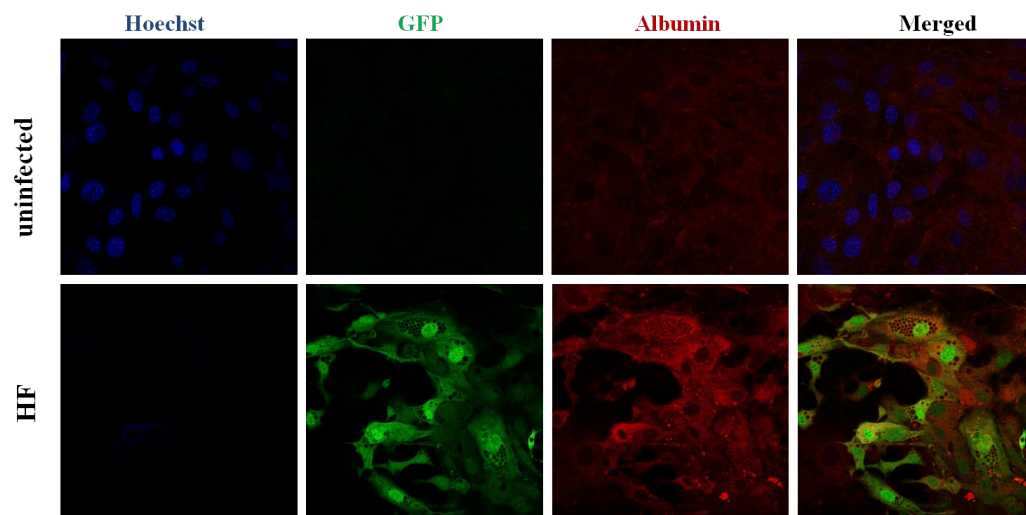


Figure A.3 Induced hepatocytes express albumin, a key hepatocyte marker. Uninfected fibroblasts (top panel) and day 7 *Hnf4α* and *Foxa1* expressing iHeps (bottom panel) were stained for and Albumin (red).

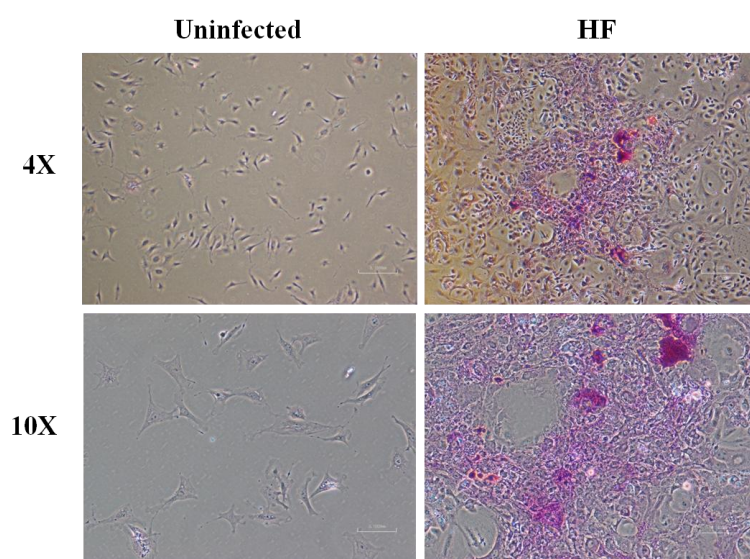


Figure A.4 Periodic acid-Schiff staining of MEFs and D14 iHeps.

Appendix B: Human Induced Hepatocytes

To determine whether mouse HNF4 α and FOXA1 transcription factor can be used to reprogram human fibroblasts into induced hepatocytes (hiHeps), fibroblasts differentiated from the human embryonic stem cell line H1 (dH1f) were infected with m*Hnf4 α* and m*Foxa1* viral vectors. On day 7 after infection, human albumin mRNA expression level was determined in the samples. Albumin gene expression levels were found to be 80-fold higher in mouse HF infected dH1fs compared to uninfected control dH1fs (Figure B.1). Treatment of cells with EPZ4777 and DZNep augmented this effect by 3-fold and 7.5-fold, respectively.

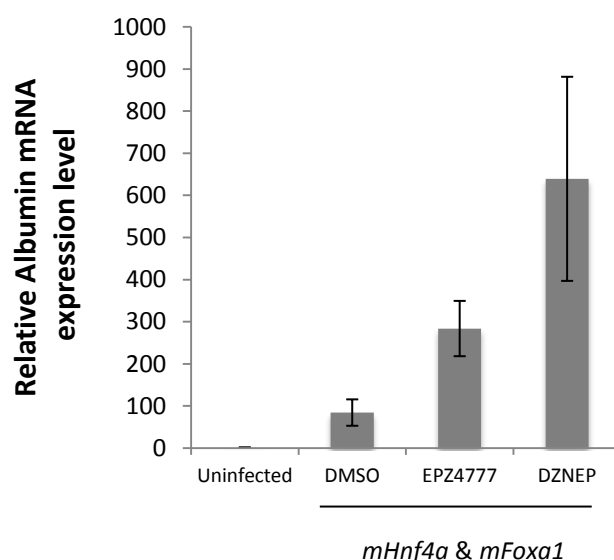


Figure B.1 dH1f lineage conversion using mouse transcription factors

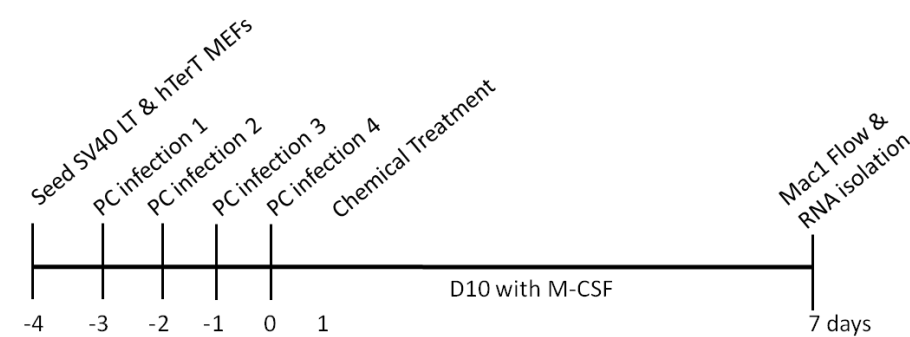
dH1f cells were infected with mouse *Hnf4 α* and *Foxa1* viral vectors, and maintained in hepatomedium containing the specified chemicals. Albumin mRNA levels were determined on day 7. Expression levels were normalized to endogenous actin gene expression level. * $p < 0.01$ compared to control uninfected fibroblasts (n=3; error bars, $\pm$ s.d.).

Appendix C: Induced Macrophages

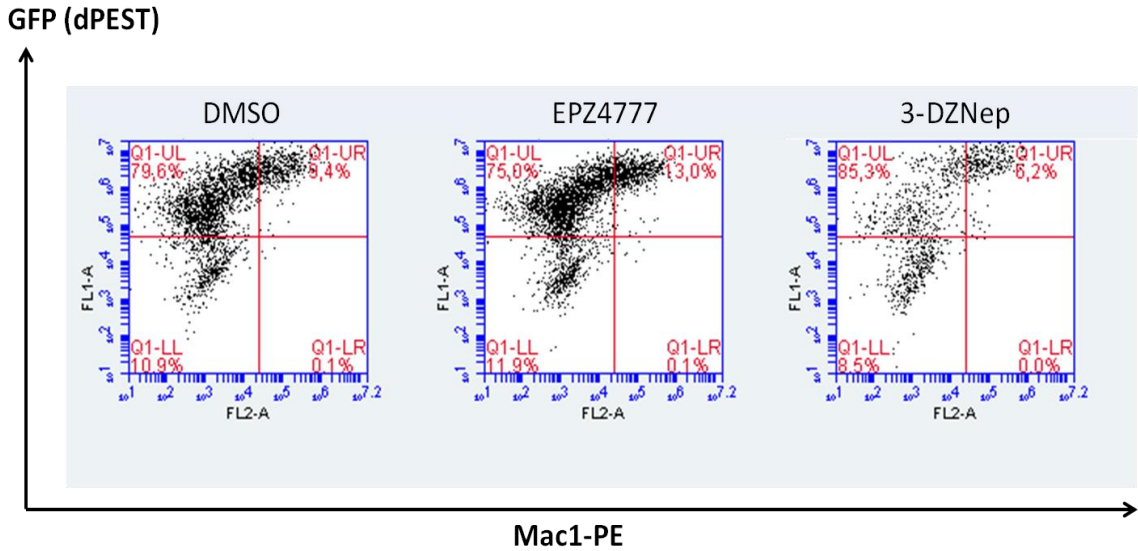
To investigate the roles of EZH2 and DOT1L inhibition in reprogramming in other contexts, we utilized the conversion of mouse fibroblasts to induced macrophages (iMacs). In this protocol, the expression of lineage-specific transcription factors *Pu.1* and *Cebpa* in mouse fibroblasts, and maintenance of cells in medium containing macrophage colony stimulating factor (M-CSF) generates macrophage-like cells that are positive for the pan-hematopoietic marker CD45, and macrophage marker MAC1 [69].

In this study, mouse embryonic fibroblasts were resistant to reprogramming by this protocol. One possible reason for this is the slow rate of cell proliferation. To facilitate reprogramming, MEFs were infected with SV40 large T antigen and human telomerase hTerT vectors. The immortalized MEFs were infected with *Pu.1delta PEST* and *Cebpa* retroviral vectors, and treated with EPZ4777 and DZNep at the same effective concentrations used for the conversion of fibroblasts to induced hepatocytes. At day 7, the cells were analyzed for their expression of the macrophage marker MAC1 by flow cytometry (Figure C.1Figure C.). DMSO treated control iMacs were 9.6% MAC1-positive, whereas EPZ4777 treated samples were 13% positive, showing a minor, but significant increase in macrophage-like cells ($p < 0.5$) (Figure C.1b and c). DZNep treatment hindered the production of macrophage-like cells, as indicated by the presence of only 6.3% MAC1-positive cells in the sample. Gene expression levels of *Mac1* and macrophage-specific gene *Csf1r* was analyzed in day 7 induced macrophages. *Mac1* expression was not statistically different among the three samples. DZNep treated samples had 50% less *Csf1r* mRNA levels compared to DMSO treated controls.

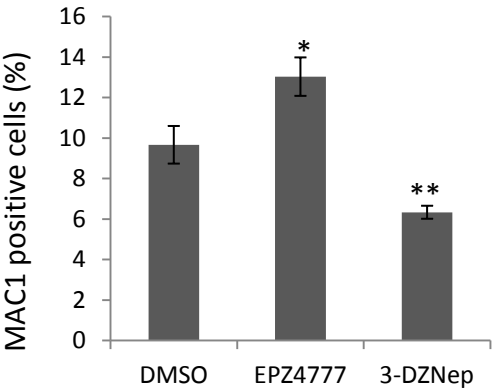
A



B



C



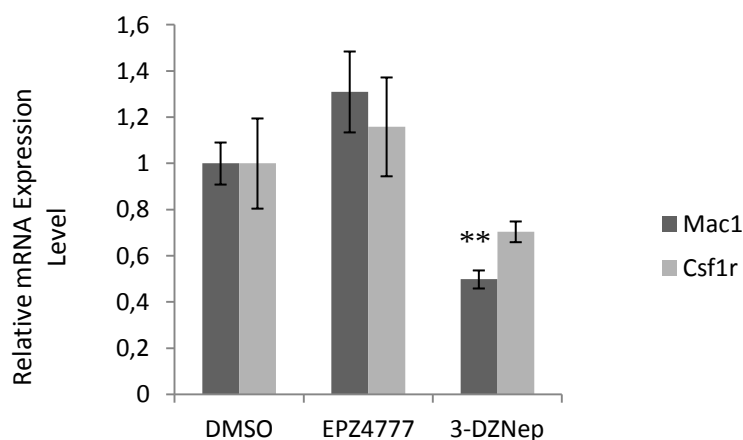
D

Figure C.1 DOT1L inhibition slightly increases the efficiency of fibroblast to macrophage transdifferentiation

(A) Immortalized mouse embryonic fibroblasts were infected with *Pu.1delta PEST* and *Cebpa* vectors, and maintained in M-CSF containing medium containing the indicated chemicals for 7 days. (B) On day 7, cells were analyzed for their expression of MAC1 by flow cytometry. (C) % MAC1-positive cells as analyzed by flow cytometry. * $p < 0.05$ and ** $p < 0.1$ compared to control DMSO treated samples ($n=3$; error bars, $\pm$ s.d.). (D) mRNA expression levels of day 7 induced macrophages. Data were normalized to endogenous actin mRNA levels. * $p < 0.1$ compared to control DMSO treated samples ($n=3$, technical replicates; error bars, $\pm$ s.d.).

Appendix D: Supplemental Methods for Induced Macrophages

Cell Culture

Induced macrophages were maintained in D10 containing 10 ng/ml macrophage colony stimulating factor (M-CSF) (PeproTech). G418 selection was performed at 400 ug/ml for one week.

DNA Constructs

MSCV PU.1 delta PEST IRES GFP construct was a kind gift from Dr. Thomas Graf [69]. pBABE-puro-hTERT (Addgene Plasmid #1771) [187] and WZLneo-C/EBP α (Addgene Plasmid #34567)[164] were purchased from Addgene.

Titering of CEBPA retroviral construct

P1 BALB/c MEFs were seeded to 0.1% gelatin coated 6 well plates at a density of 100000 cells/well. The next day in the afternoon, cells were infected with 10, 1 or 0.1 μ l of concentrated virus with 8 μ g/ml protamine sulfate (Sigma-Aldrich) in the infection medium. In the morning, virus was washed away. The next day, G418 was added to the culture medium. Cells were selected for one week, or until cells in the uninfected wells died completely. Cells were stained with crystal violet to calculate the titer.

Crystal Violet Staining

Cells were fixed with 4% paraformaldehyde for 5 min at RT, and stained with 0.05% crystal violet for 10 min with gentle shaking. The stain was washed away with tap water, and samples were dried by inverting the plates.

Generation of immortalized BALB/c MEFs

Fibroblasts were infected with pBABE-puro SV40 LT and pBABE-puro-hTERT constructs. The cells that received these plasmids were selected based on their ability to proliferate for extended periods of time in culture.

Generation of Induced Macrophages

Immortalized BALB/c MEFs were seeded onto 0.1% gelatin coated 12 well plates at a density of 50 000 cells/well. The next day, cells were infected with the retroviral vectors *Pu.1* and *Cebpa* at MOI:10 with 8 µg/ml protamine sulfate (Sigma-Aldrich). The virus was washed off and fresh media was added to the wells the next morning. Infection was repeated three times in the next three days. The day after the final infection, cells were switched to D10 medium containing M-CSF. Cells were analyzed on day 7.

Flow Cytometry

Cells were harvested and washed once with PBS. Cells were centrifuged at 1500 rpm for 5 min and resuspended in FACS buffer (PBS with 1% FBS) containing PE conjugated Mac1 antibody (BioLegend #101207) for 30 min on ice in the dark. The unbound antibody was washed with FACS buffer, in three washes. The cells were resuspended in fresh FACS buffer. Samples were analyzed using BD Accuri C6 Flow Cytometer (BD Biosciences) according to the manufacturer's instructions.

qPCR Primers

The qPCR primers used are listed in Table D.1.

Table D.1 List of qPCR primers used in induced macrophage experiments

Gene	Species	Forward Primer	Reverse Primer
Mac1	mouse	ATGGACGCTGATGGCAATACC	TCCCCATTCACGTCTCCCA
Csf1r	mouse	TGTCATCGAGCCTAGTGGC	CGGGAGATTCAGGGTCCAAG

Bibliography

1. Gurdon, J.B., T.R. Elsdale, and M. Fischberg, *Sexually mature individuals of Xenopus laevis from the transplantation of single somatic nuclei*. Nature, 1958. **182**(4627): p. 64-5.
2. Gurdon, J.B., *Transplanted nuclei and cell differentiation*. Sci Am, 1968. **219**(6): p. 24-35.
3. Gurdon, J.B., *Changes in somatic cell nuclei inserted into growing and maturing amphibian oocytes*. J Embryol Exp Morphol, 1968. **20**(3): p. 401-14.
4. Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors*. Cell, 2006. **126**(4): p. 663-76.
5. Yamanaka, S. and H.M. Blau, *Nuclear reprogramming to a pluripotent state by three approaches*. Nature, 2010. **465**(7299): p. 704-12.
6. Wheeler, G.N. and A.W. Brandli, *Simple vertebrate models for chemical genetics and drug discovery screens: lessons from zebrafish and Xenopus*. Dev Dyn, 2009. **238**(6): p. 1287-308.
7. Gurdon, J.B., *From nuclear transfer to nuclear reprogramming: the reversal of cell differentiation*. Annu Rev Cell Dev Biol, 2006. **22**: p. 1-22.
8. Spemann, H., *Die Entwicklung seitlicher und dorso-ventraler Keimhälften bei verzögerter Kernversorgung*. Z Wiss Zool, 1928. **132**: p. 105-134.
9. Briggs, R. and T.J. King, *Transplantation of Living Nuclei From Blastula Cells into Enucleated Frogs' Eggs*. Proc Natl Acad Sci U S A, 1952. **38**(5): p. 455-63.
10. King, T.J. and R. Briggs, *Changes in the nuclei of differentiating gastrula cells, as demonstrated by nuclear transplantation*. Proc Natl Acad Sci U S A, 1955. **41**(5): p. 321-5.
11. Gurdon, J.B., *The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles*. J Embryol Exp Morphol, 1962. **10**: p. 622-40.
12. Gurdon, J.B., *Adult frogs derived from the nuclei of single somatic cells*. Dev Biol, 1962. **4**: p. 256-73.
13. DiBerardino, M.A. and N.J. Hoffner, *Gene reactivation in erythrocytes: nuclear transplantation in oocytes and eggs of Rana*. Science, 1983. **219**(4586): p. 862-4.
14. Wilmut, I., et al., *Viable offspring derived from fetal and adult mammalian cells*. Cloning Stem Cells, 2007. **9**(1): p. 3-7.
15. Wakayama, T., et al., *Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei*. Nature, 1998. **394**(6691): p. 369-74.
16. Hochedlinger, K. and R. Jaenisch, *Monoclonal mice generated by nuclear transfer from mature B and T donor cells*. Nature, 2002. **415**(6875): p. 1035-8.
17. Egli, D., et al., *Developmental reprogramming after chromosome transfer into mitotic mouse zygotes*. Nature, 2007. **447**(7145): p. 679-85.
18. Colman, A., *Somatic cell nuclear transfer in mammals: progress and applications*. Cloning, 1999. **1**(4): p. 185-200.
19. Thuan, N.V., S. Kishigami, and T. Wakayama, *How to improve the success rate of mouse cloning technology*. J Reprod Dev, 2010. **56**(1): p. 20-30.
20. Wakayama, S., et al., *Production of healthy cloned mice from bodies frozen at -20 degrees C for 16 years*. Proc Natl Acad Sci U S A, 2008. **105**(45): p. 17318-22.
21. Yang, X., et al., *Nuclear reprogramming of cloned embryos and its implications for therapeutic cloning*. Nat Genet, 2007. **39**(3): p. 295-302.
22. Byrne, J.A., et al., *Producing primate embryonic stem cells by somatic cell nuclear transfer*. Nature, 2007. **450**(7169): p. 497-502.

23. Noggle, S., et al., *Human oocytes reprogram somatic cells to a pluripotent state*. Nature, 2011. **478**(7367): p. 70-5.
24. Simonsson, S. and J. Gurdon, *DNA demethylation is necessary for the epigenetic reprogramming of somatic cell nuclei*. Nat Cell Biol, 2004. **6**(10): p. 984-90.
25. Weiss, M.C. and M. Chaplain, *Expression of differentiated functions in hepatoma cell hybrids: reappearance of tyrosine aminotransferase inducibility after the loss of chromosomes*. Proc Natl Acad Sci U S A, 1971. **68**(12): p. 3026-30.
26. Davidson, R.L., B. Ephrussi, and K. Yamamoto, *Regulation of pigment synthesis in mammalian cells, as studied by somatic hybridization*. Proc Natl Acad Sci U S A, 1966. **56**(5): p. 1437-40.
27. Harris, H., et al., *Suppression of malignancy by cell fusion*. Nature, 1969. **223**(5204): p. 363-8.
28. Blau, H.M., C.P. Chiu, and C. Webster, *Cytoplasmic activation of human nuclear genes in stable heterocaryons*. Cell, 1983. **32**(4): p. 1171-80.
29. Blau, H.M., et al., *Plasticity of the differentiated state*. Science, 1985. **230**(4727): p. 758-66.
30. Pavlath, G.K. and H.M. Blau, *Expression of muscle genes in heterokaryons depends on gene dosage*. J Cell Biol, 1986. **102**(1): p. 124-30.
31. Tada, M., et al., *Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells*. EMBO J, 1997. **16**(21): p. 6510-20.
32. Tada, M., et al., *Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells*. Curr Biol, 2001. **11**(19): p. 1553-8.
33. Kimura, H., et al., *Histone code modifications on pluripotent nuclei of reprogrammed somatic cells*. Mol Cell Biol, 2004. **24**(13): p. 5710-20.
34. Pereira, C.F., et al., *Heterokaryon-based reprogramming of human B lymphocytes for pluripotency requires Oct4 but not Sox2*. PLoS Genet, 2008. **4**(9): p. e1000170.
35. Bhutani, N., et al., *Reprogramming towards pluripotency requires AID-dependent DNA demethylation*. Nature, 2010. **463**(7284): p. 1042-7.
36. Tokuzawa, Y., et al., *Fbx15 is a novel target of Oct3/4 but is dispensable for embryonic stem cell self-renewal and mouse development*. Mol Cell Biol, 2003. **23**(8): p. 2699-708.
37. Wernig, M., et al., *In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state*. Nature, 2007. **448**(7151): p. 318-24.
38. Okita, K., T. Ichisaka, and S. Yamanaka, *Generation of germline-competent induced pluripotent stem cells*. Nature, 2007. **448**(7151): p. 313-7.
39. Takahashi, K., et al., *Induction of pluripotent stem cells from adult human fibroblasts by defined factors*. Cell, 2007. **131**(5): p. 861-72.
40. Yamanaka, S., *A fresh look at iPS cells*. Cell, 2009. **137**(1): p. 13-7.
41. Wernig, M., et al., *A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types*. Nat Biotechnol, 2008. **26**(8): p. 916-24.
42. Yamanaka, S., *Elite and stochastic models for induced pluripotent stem cell generation*. Nature, 2009. **460**(7251): p. 49-52.
43. Yamada, T. and D.S. McDevitt, *Conversion of iris epithelial cells as a model of differentiation control*. Differentiation, 1984. **27**(1): p. 1-12.
44. Tsonis, P.A., et al., *A newt's eye view of lens regeneration*. Int J Dev Biol, 2004. **48**(8-9): p. 975-80.
45. Eilken, H.M., S. Nishikawa, and T. Schroeder, *Continuous single-cell imaging of blood generation from haemogenic endothelium*. Nature, 2009. **457**(7231): p. 896-900.

46. Slack, J.M., *Metaplasia and transdifferentiation: from pure biology to the clinic*. Nat Rev Mol Cell Biol, 2007. **8**(5): p. 369-78.
47. Yang, J. and R.A. Weinberg, *Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis*. Dev Cell, 2008. **14**(6): p. 818-29.
48. Hadorn, E., *Transdetermination in cells*. Sci Am, 1968. **219**(5): p. 110-4 passim.
49. Halder, G., P. Callaerts, and W.J. Gehring, *Induction of ectopic eyes by targeted expression of the eyeless gene in Drosophila*. Science, 1995. **267**(5205): p. 1788-92.
50. Schneuwly, S., R. Klemenz, and W.J. Gehring, *Redesigning the body plan of Drosophila by ectopic expression of the homoeotic gene Antennapedia*. Nature, 1987. **325**(6107): p. 816-8.
51. Taylor, S.M. and P.A. Jones, *Multiple new phenotypes induced in 10T1/2 and 3T3 cells treated with 5-azacytidine*. Cell, 1979. **17**(4): p. 771-9.
52. Jones, P.A. and S.M. Taylor, *Cellular differentiation, cytidine analogs and DNA methylation*. Cell, 1980. **20**(1): p. 85-93.
53. Lassar, A.B., B.M. Paterson, and H. Weintraub, *Transfection of a DNA locus that mediates the conversion of 10T1/2 fibroblasts to myoblasts*. Cell, 1986. **47**(5): p. 649-56.
54. Davis, R.L., H. Weintraub, and A.B. Lassar, *Expression of a single transfected cDNA converts fibroblasts to myoblasts*. Cell, 1987. **51**(6): p. 987-1000.
55. Kulesa, H., J. Frampton, and T. Graf, *GATA-1 reprograms avian myelomonocytic cell lines into eosinophils, thromboblats, and erythroblats*. Genes Dev, 1995. **9**(10): p. 1250-62.
56. Visvader, J.E., et al., *GATA-1 but not SCL induces megakaryocytic differentiation in an early myeloid line*. EMBO J, 1992. **11**(12): p. 4557-64.
57. Nerlov, C. and T. Graf, *PU.1 induces myeloid lineage commitment in multipotent hematopoietic progenitors*. Genes Dev, 1998. **12**(15): p. 2403-12.
58. Xie, H., et al., *Stepwise reprogramming of B cells into macrophages*. Cell, 2004. **117**(5): p. 663-76.
59. Laiosa, C.V., et al., *Reprogramming of committed T cell progenitors to macrophages and dendritic cells by C/EBP alpha and PU.1 transcription factors*. Immunity, 2006. **25**(5): p. 731-44.
60. Zhang, P., et al., *Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha*. Immunity, 2004. **21**(6): p. 853-63.
61. Zhou, Q., et al., *In vivo reprogramming of adult pancreatic exocrine cells to beta-cells*. Nature, 2008. **455**(7213): p. 627-32.
62. Jonsson, J., et al., *Insulin-promoter-factor 1 is required for pancreas development in mice*. Nature, 1994. **371**(6498): p. 606-9.
63. Ahlgren, U., et al., *beta-cell-specific inactivation of the mouse Ipf1/Pdx1 gene results in loss of the beta-cell phenotype and maturity onset diabetes*. Genes Dev, 1998. **12**(12): p. 1763-8.
64. Gu, G., J. Dubauskaite, and D.A. Melton, *Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors*. Development, 2002. **129**(10): p. 2447-57.
65. Zhang, C., et al., *MafA is a key regulator of glucose-stimulated insulin secretion*. Mol Cell Biol, 2005. **25**(12): p. 4969-76.
66. Seale, P., et al., *PRDM16 controls a brown fat/skeletal muscle switch*. Nature, 2008. **454**(7207): p. 961-7.
67. Kajimura, S., et al., *Initiation of myoblast to brown fat switch by a PRDM16-C/EBP-beta transcriptional complex*. Nature, 2009. **460**(7259): p. 1154-8.

68. Ieda, M., et al., *Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors*. Cell, 2010. **142**(3): p. 375-86.
69. Feng, R., et al., *PU.1 and C/EBPalpha/beta convert fibroblasts into macrophage-like cells*. Proc Natl Acad Sci U S A, 2008. **105**(16): p. 6057-62.
70. Vierbuchen, T., et al., *Direct conversion of fibroblasts to functional neurons by defined factors*. Nature, 2010. **463**(7284): p. 1035-41.
71. Nieto, M., et al., *Neural bHLH genes control the neuronal versus glial fate decision in cortical progenitors*. Neuron, 2001. **29**(2): p. 401-13.
72. Leone, D.P., et al., *The determination of projection neuron identity in the developing cerebral cortex*. Curr Opin Neurobiol, 2008. **18**(1): p. 28-35.
73. Cahoy, J.D., et al., *A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function*. J Neurosci, 2008. **28**(1): p. 264-78.
74. Weiner, J.A. and J. Chun, *Png-1, a nervous system-specific zinc finger gene, identifies regions containing postmitotic neurons during mammalian embryonic development*. J Comp Neurol, 1997. **381**(2): p. 130-42.
75. Wapinski, O.L., et al., *Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons*. Cell, 2013. **155**(3): p. 621-35.
76. Chanda, S., et al., *Generation of induced neuronal cells by the single reprogramming factor ASCL1*. Stem Cell Reports, 2014. **3**(2): p. 282-96.
77. Pang, Z.P., et al., *Induction of human neuronal cells by defined transcription factors*. Nature, 2011. **476**(7359): p. 220-3.
78. Yoo, A.S., et al., *MicroRNA-mediated conversion of human fibroblasts to neurons*. Nature, 2011. **476**(7359): p. 228-31.
79. Pfisterer, U., et al., *Direct conversion of human fibroblasts to dopaminergic neurons*. Proc Natl Acad Sci U S A, 2011. **108**(25): p. 10343-8.
80. Xu, J., Y. Du, and H. Deng, *Direct lineage reprogramming: strategies, mechanisms, and applications*. Cell Stem Cell, 2015. **16**(2): p. 119-34.
81. Szabo, E., et al., *Direct conversion of human fibroblasts to multilineage blood progenitors*. Nature, 2010. **468**(7323): p. 521-6.
82. Kim, H.S., et al., *Direct lineage reprogramming of mouse fibroblasts to functional midbrain dopaminergic neuronal progenitors*. Stem Cell Res, 2014. **12**(1): p. 60-8.
83. Caiazzo, M., et al., *Direct generation of functional dopaminergic neurons from mouse and human fibroblasts*. Nature, 2011. **476**(7359): p. 224-7.
84. Son, E.Y., et al., *Conversion of mouse and human fibroblasts into functional spinal motor neurons*. Cell Stem Cell, 2011. **9**(3): p. 205-18.
85. Sekiya, S. and A. Suzuki, *Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors*. Nature, 2011. **475**(7356): p. 390-393.
86. Ambasudhan, R., et al., *Direct reprogramming of adult human fibroblasts to functional neurons under defined conditions*. Cell Stem Cell, 2011. **9**(2): p. 113-8.
87. Ring, K.L., et al., *Direct reprogramming of mouse and human fibroblasts into multipotent neural stem cells with a single factor*. Cell Stem Cell, 2012. **11**(1): p. 100-9.
88. Buganim, Y., et al., *Direct reprogramming of fibroblasts into embryonic Sertoli-like cells by defined factors*. Cell Stem Cell, 2012. **11**(3): p. 373-86.
89. Han, D.W., et al., *Direct reprogramming of fibroblasts into neural stem cells by defined factors*. Cell Stem Cell, 2012. **10**(4): p. 465-72.

90. Yu, B., et al., *Reprogramming fibroblasts into bipotential hepatic stem cells by defined factors*. Cell Stem Cell, 2013. **13**(3): p. 328-40.
91. Pereira, C.F., et al., *Induction of a hemogenic program in mouse fibroblasts*. Cell Stem Cell, 2013. **13**(2): p. 205-18.
92. Li, W., et al., *In vivo reprogramming of pancreatic acinar cells to three islet endocrine subtypes*. Elife, 2014. **3**: p. e01846.
93. Bredenkamp, N., et al., *An organized and functional thymus generated from FOXP1-reprogrammed fibroblasts*. Nat Cell Biol, 2014. **16**(9): p. 902-8.
94. Riddell, J., et al., *Reprogramming committed murine blood cells to induced hematopoietic stem cells with defined factors*. Cell, 2014. **157**(3): p. 549-64.
95. Sandler, V.M., et al., *Reprogramming human endothelial cells to haematopoietic cells requires vascular induction*. Nature, 2014. **511**(7509): p. 312-8.
96. Huang, P., et al., *Direct reprogramming of human fibroblasts to functional and expandable hepatocytes*. Cell Stem Cell, 2014. **14**(3): p. 370-84.
97. Kim, Y.J., et al., *Generation of multipotent induced neural crest by direct reprogramming of human postnatal fibroblasts with a single transcription factor*. Cell Stem Cell, 2014. **15**(4): p. 497-506.
98. Ouyang, H., et al., *WNT7A and PAX6 define corneal epithelium homeostasis and pathogenesis*. Nature, 2014. **511**(7509): p. 358-61.
99. Schafer, B.W., et al., *Effect of cell history on response to helix-loop-helix family of myogenic regulators*. Nature, 1990. **344**(6265): p. 454-8.
100. Thayer, M.J. and H. Weintraub, *Activation and repression of myogenesis in somatic cell hybrids: evidence for trans-negative regulation of MyoD in primary fibroblasts*. Cell, 1990. **63**(1): p. 23-32.
101. Sancho-Martinez, I., E. Nivet, and J.C. Izpisua Belmonte, *The labyrinth of nuclear reprogramming*. J Mol Cell Biol, 2011. **3**(6): p. 327-9.
102. Panopoulos, A.D., S. Ruiz, and J.C. Izpisua Belmonte, *iPSCs: induced back to controversy*. Cell Stem Cell, 2011. **8**(4): p. 347-8.
103. Marro, S., et al., *Direct lineage conversion of terminally differentiated hepatocytes to functional neurons*. Cell Stem Cell, 2011. **9**(4): p. 374-82.
104. Ruiz, S., et al., *A high proliferation rate is required for cell reprogramming and maintenance of human embryonic stem cell identity*. Curr Biol, 2011. **21**(1): p. 45-52.
105. Zaret, K.S., *Regulatory phases of early liver development: paradigms of organogenesis*. Nat Rev Genet, 2002. **3**(7): p. 499-512.
106. Zaret, K.S. and M. Grompe, *Generation and regeneration of cells of the liver and pancreas*. Science, 2008. **322**(5907): p. 1490-4.
107. Overturf, K., et al., *Hepatocytes corrected by gene therapy are selected in vivo in a murine model of hereditary tyrosinaemia type I*. Nat Genet, 1996. **12**(3): p. 266-73.
108. Morris, S.A., et al., *Dissecting engineered cell types and enhancing cell fate conversion via CellNet*. Cell, 2014. **158**(4): p. 889-902.
109. Huang, P., et al., *Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors*. Nature, 2011. **475**(7356): p. 386-9.
110. Lai, E., et al., *HNF-3A, a hepatocyte-enriched transcription factor of novel structure is regulated transcriptionally*. Genes Dev, 1990. **4**(8): p. 1427-36.
111. Lai, E., et al., *Hepatocyte nuclear factor 3 alpha belongs to a gene family in mammals that is homologous to the Drosophila homeotic gene fork head*. Genes Dev, 1991. **5**(3): p. 416-27.

112. Sasaki, H. and B.L. Hogan, *Differential expression of multiple fork head related genes during gastrulation and axial pattern formation in the mouse embryo*. Development, 1993. **118**(1): p. 47-59.
113. Ang, S.L., et al., *The formation and maintenance of the definitive endoderm lineage in the mouse: involvement of HNF3/forkhead proteins*. Development, 1993. **119**(4): p. 1301-15.
114. Monaghan, A.P., et al., *Postimplantation expression patterns indicate a role for the mouse forkhead/HNF-3 alpha, beta and gamma genes in determination of the definitive endoderm, chordamesoderm and neuroectoderm*. Development, 1993. **119**(3): p. 567-78.
115. Ruiz i Altaba, A., et al., *Sequential expression of HNF-3 beta and HNF-3 alpha by embryonic organizing centers: the dorsal lip/node, notochord and floor plate*. Mech Dev, 1993. **44**(2-3): p. 91-108.
116. Kaestner, K.H., et al., *Inactivation of the winged helix transcription factor HNF3alpha affects glucose homeostasis and islet glucagon gene expression in vivo*. Genes Dev, 1999. **13**(4): p. 495-504.
117. Ang, S.L. and J. Rossant, *HNF-3 beta is essential for node and notochord formation in mouse development*. Cell, 1994. **78**(4): p. 561-74.
118. Weinstein, D.C., et al., *The winged-helix transcription factor HNF-3 beta is required for notochord development in the mouse embryo*. Cell, 1994. **78**(4): p. 575-88.
119. Kaestner, K.H., H. Hiemisch, and G. Schutz, *Targeted disruption of the gene encoding hepatocyte nuclear factor 3gamma results in reduced transcription of hepatocyte-specific genes*. Mol Cell Biol, 1998. **18**(7): p. 4245-51.
120. Chen, W.S., et al., *Disruption of the HNF-4 gene, expressed in visceral endoderm, leads to cell death in embryonic ectoderm and impaired gastrulation of mouse embryos*. Genes Dev, 1994. **8**(20): p. 2466-77.
121. Duncan, S.A., A. Nagy, and W. Chan, *Murine gastrulation requires HNF-4 regulated gene expression in the visceral endoderm: tetraploid rescue of Hnf-4(-/-) embryos*. Development, 1997. **124**(2): p. 279-87.
122. Li, J., G. Ning, and S.A. Duncan, *Mammalian hepatocyte differentiation requires the transcription factor HNF-4alpha*. Genes Dev, 2000. **14**(4): p. 464-74.
123. Du, Y., et al., *Human hepatocytes with drug metabolic function induced from fibroblasts by lineage reprogramming*. Cell Stem Cell, 2014. **14**(3): p. 394-403.
124. Kouzarides, T., *Chromatin modifications and their function*. Cell, 2007. **128**(4): p. 693-705.
125. Barski, A., et al., *High-resolution profiling of histone methylations in the human genome*. Cell, 2007. **129**(4): p. 823-37.
126. Koch, C.M., et al., *The landscape of histone modifications across 1% of the human genome in five human cell lines*. Genome Res, 2007. **17**(6): p. 691-707.
127. *An integrated encyclopedia of DNA elements in the human genome*. Nature, 2012. **489**(7414): p. 57-74.
128. Ho, L. and G.R. Crabtree, *Chromatin remodelling during development*. Nature, 2010. **463**(7280): p. 474-84.
129. Park, S.H., et al., *Ultrastructure of human embryonic stem cells and spontaneous and retinoic acid-induced differentiating cells*. Ultrastruct Pathol, 2004. **28**(4): p. 229-38.
130. Meshorer, E., et al., *Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells*. Dev Cell, 2006. **10**(1): p. 105-16.
131. Azuara, V., et al., *Chromatin signatures of pluripotent cell lines*. Nat Cell Biol, 2006. **8**(5): p. 532-8.

132. Guenther, M.G., et al., *A chromatin landmark and transcription initiation at most promoters in human cells*. Cell, 2007. **130**(1): p. 77-88.
133. Meissner, A., et al., *Genome-scale DNA methylation maps of pluripotent and differentiated cells*. Nature, 2008. **454**(7205): p. 766-70.
134. Bernstein, B.E., et al., *A bivalent chromatin structure marks key developmental genes in embryonic stem cells*. Cell, 2006. **125**(2): p. 315-26.
135. Voigt, P., et al., *Asymmetrically modified nucleosomes*. Cell, 2012. **151**(1): p. 181-93.
136. Wen, B., et al., *Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells*. Nat Genet, 2009. **41**(2): p. 246-50.
137. Creyghton, M.P., et al., *Histone H3K27ac separates active from poised enhancers and predicts developmental state*. Proc Natl Acad Sci U S A, 2010. **107**(50): p. 21931-6.
138. Rada-Iglesias, A., et al., *A unique chromatin signature uncovers early developmental enhancers in humans*. Nature, 2011. **470**(7333): p. 279-83.
139. Xie, W., et al., *Epigenomic analysis of multilineage differentiation of human embryonic stem cells*. Cell, 2013. **153**(5): p. 1134-48.
140. Adams, C.C. and J.L. Workman, *Binding of disparate transcriptional activators to nucleosomal DNA is inherently cooperative*. Mol Cell Biol, 1995. **15**(3): p. 1405-21.
141. Soufi, A., G. Donahue, and K.S. Zaret, *Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome*. Cell, 2012. **151**(5): p. 994-1004.
142. Koche, R.P., et al., *Reprogramming factor expression initiates widespread targeted chromatin remodeling*. Cell Stem Cell, 2011. **8**(1): p. 96-105.
143. Papp, B. and K. Plath, *Epigenetics of reprogramming to induced pluripotency*. Cell, 2013. **152**(6): p. 1324-43.
144. Hysolli, E. and I.H. Park, *Trivalent chromatin marks the way in*. Cell Stem Cell, 2013. **13**(5): p. 510-2.
145. Campos, E.I. and D. Reinberg, *Histones: annotating chromatin*. Annu Rev Genet, 2009. **43**: p. 559-99.
146. Cao, R., et al., *Role of histone H3 lysine 27 methylation in Polycomb-group silencing*. Science, 2002. **298**(5595): p. 1039-43.
147. Margueron, R. and D. Reinberg, *The Polycomb complex PRC2 and its mark in life*. Nature, 2011. **469**(7330): p. 343-9.
148. Kuzmichev, A., et al., *Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein*. Genes Dev, 2002. **16**(22): p. 2893-905.
149. Mendenhall, E.M., et al., *GC-rich sequence elements recruit PRC2 in mammalian ES cells*. PLoS Genet, 2010. **6**(12): p. e1001244.
150. Di Croce, L. and K. Helin, *Transcriptional regulation by Polycomb group proteins*. Nat Struct Mol Biol, 2013. **20**(10): p. 1147-55.
151. Francis, N.J., R.E. Kingston, and C.L. Woodcock, *Chromatin compaction by a polycomb group protein complex*. Science, 2004. **306**(5701): p. 1574-7.
152. Spivakov, M. and A.G. Fisher, *Epigenetic signatures of stem-cell identity*. Nat Rev Genet, 2007. **8**(4): p. 263-71.
153. Mohn, F., et al., *Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors*. Mol Cell, 2008. **30**(6): p. 755-66.
154. Lee, T.I., et al., *Control of developmental regulators by Polycomb in human embryonic stem cells*. Cell, 2006. **125**(2): p. 301-13.

155. Boyer, L.A., et al., *Polycomb complexes repress developmental regulators in murine embryonic stem cells*. Nature, 2006. **441**(7091): p. 349-53.
156. Bracken, A.P., et al., *Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions*. Genes Dev, 2006. **20**(9): p. 1123-36.
157. Chamberlain, S.J., D. Yee, and T. Magnuson, *Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency*. Stem Cells, 2008. **26**(6): p. 1496-505.
158. Walker, E., et al., *Polycomb-like 2 associates with PRC2 and regulates transcriptional networks during mouse embryonic stem cell self-renewal and differentiation*. Cell Stem Cell, 2010. **6**(2): p. 153-66.
159. O'Carroll, D., et al., *The polycomb-group gene Ezh2 is required for early mouse development*. Mol Cell Biol, 2001. **21**(13): p. 4330-6.
160. Shumacher, A., C. Faust, and T. Magnuson, *Positional cloning of a global regulator of anterior-posterior patterning in mice*. Nature, 1996. **383**(6597): p. 250-3.
161. Pasini, D., et al., *Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity*. EMBO J, 2004. **23**(20): p. 4061-71.
162. Pasini, D., et al., *The polycomb group protein Suz12 is required for embryonic stem cell differentiation*. Mol Cell Biol, 2007. **27**(10): p. 3769-79.
163. Shen, X., et al., *EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency*. Mol Cell, 2008. **32**(4): p. 491-502.
164. Wang, L., et al., *Histone H3K27 methyltransferase Ezh2 represses Wnt genes to facilitate adipogenesis*. Proc Natl Acad Sci U S A, 2010. **107**(16): p. 7317-22.
165. Su, I.H., et al., *Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement*. Nat Immunol, 2003. **4**(2): p. 124-31.
166. Caretti, G., et al., *The Polycomb Ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation*. Genes Dev, 2004. **18**(21): p. 2627-38.
167. Ezhkova, E., et al., *Ezh2 orchestrates gene expression for the stepwise differentiation of tissue-specific stem cells*. Cell, 2009. **136**(6): p. 1122-35.
168. Xu, C.R., et al., *Chromatin "prepattern" and histone modifiers in a fate choice for liver and pancreas*. Science, 2011. **332**(6032): p. 963-6.
169. Pereira, C.F., et al., *ESCs require PRC2 to direct the successful reprogramming of differentiated cells toward pluripotency*. Cell Stem Cell, 2010. **6**(6): p. 547-56.
170. Onder, T.T., et al., *Chromatin-modifying enzymes as modulators of reprogramming*. Nature, 2012. **483**(7391): p. 598-602.
171. Feng, Q., et al., *Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain*. Curr Biol, 2002. **12**(12): p. 1052-8.
172. Cheng, X., R.E. Collins, and X. Zhang, *Structural and sequence motifs of protein (histone) methylation enzymes*. Annu Rev Biophys Biomol Struct, 2005. **34**: p. 267-94.
173. Barry, E.R., et al., *ES cell cycle progression and differentiation require the action of the histone methyltransferase Dot1L*. Stem Cells, 2009. **27**(7): p. 1538-47.
174. Jones, B., et al., *The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure*. PLoS Genet, 2008. **4**(9): p. e1000190.
175. Castano Betancourt, M.C., et al., *Genome-wide association and functional studies identify the DOT1L gene to be involved in cartilage thickness and hip osteoarthritis*. Proc Natl Acad Sci U S A, 2012. **109**(21): p. 8218-23.
176. Lin, Y.W. and P.D. Aplan, *Leukemic transformation*. Cancer Biol Ther, 2004. **3**(1): p. 13-20.

177. Okada, Y., et al., *hDOT1L links histone methylation to leukemogenesis*. Cell, 2005. **121**(2): p. 167-78.
178. Ooga, M., et al., *Changes in H3K79 methylation during preimplantation development in mice*. Biol Reprod, 2008. **78**(3): p. 413-24.
179. Bernt, K.M., et al., *MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L*. Cancer Cell, 2011. **20**(1): p. 66-78.
180. Zhao, J.J., et al., *Human mammary epithelial cell transformation through the activation of phosphatidylinositol 3-kinase*. Cancer Cell, 2003. **3**(5): p. 483-95.
181. Kumar, M.S., et al., *Dicer1 functions as a haploinsufficient tumor suppressor*. Genes Dev, 2009. **23**(23): p. 2700-4.
182. Hou, P., et al., *Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds*. Science, 2013. **341**(6146): p. 651-4.
183. Kim, E., et al., *Phosphorylation of EZH2 activates STAT3 signaling via STAT3 methylation and promotes tumorigenicity of glioblastoma stem-like cells*. Cancer Cell, 2013. **23**(6): p. 839-52.
184. Li, X., et al., *Targeted overexpression of EZH2 in the mammary gland disrupts ductal morphogenesis and causes epithelial hyperplasia*. Am J Pathol, 2009. **175**(3): p. 1246-54.
185. Shi, B., et al., *Integration of estrogen and Wnt signaling circuits by the polycomb group protein EZH2 in breast cancer cells*. Mol Cell Biol, 2007. **27**(14): p. 5105-19.
186. Watanabe, H., et al., *HHEX promotes hepatic-lineage specification through the negative regulation of eomesodermin*. PLoS One, 2014. **9**(3): p. e90791.
187. Counter, C.M., et al., *Dissociation among in vitro telomerase activity, telomere maintenance, and cellular immortalization*. Proc Natl Acad Sci U S A, 1998. **95**(25): p. 14723-8.

Vita

Duygu Payzın was born in İzmir, Turkey on November 11, 1989. She received her Bachelor of Science degree in Molecular Biology and Genetics from Bilkent University, Ankara in 2012. She is working towards her Master of Science degree in Molecular Biology and Genetics from Koç University, Istanbul. Her thesis is on understanding the epigenetic modifications during mouse fibroblast transdifferentiation to induced hepatocytes.