The Role of Histone H3 Lysine 36 Methylation in Reprogramming of Fibroblasts and on Induced Pluripotent Stem Cell Generation

by

Burcu Özçimen

A Dissertation Submitted to the
Graduate School of Science and Engineering
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy
in
Molecular Biology and Genetics

KOÇ
UNIVERSITY

August,10, 2018

The Role of Histone H3 Lysine 36 Methylation in Reprogramming of Fibroblasts and on Induced Pluripotent Stem Cell Generation

Koç University

Graduate School of Science and Engineering

This is to certify that I have examined this copy of a doctoral dissertation by

Burcu Özçimen

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:

Assoc. Prof. Tamer Önder (Advisor)……

Assoc. Prof. Nathan Lack…….

Assoc. Prof. Tolga Emre……..

Prof. Nesrin Özören……

Assoc. Prof. Müjdat Zeybel……..

Date: 10.08.2018

ABSTRACT

Embryonic stem cells (ESCs) have the capacity to form all types of somatic cells in the body. Introduction of four transcription factors *OCT4*, SOX2, *KLF4* and *cMYC* (OSKM) into somatic cells induces reprogramming to generate ESC-like cells. Reprogramming is a step-wise process which requires several molecular changes. One of these is the remodeling of global chromatin landscape of somatic cells to gain pluripotency. However, there remain many unanswered questions regarding the influence of chromatin regulators on the gene expression network of somatic cells and how they affect reprogramming.

SETD2 is the only chromatin modifier responsible for generating the H3 Lysine 36 (H3K36) trimethylation, a histone mark that is associated with active transcription. The role of H3K36 methylation in reprogramming remains unknown. Therefore, we investigated the role of SETD2 in reprogramming via shRNA-mediated knock-down or gRNA-mediated knock-out strategies in human fibroblasts. We also performed RNA sequencing on SETD2-depleted cells before and after OSKM-induction to understand how loss of H3K36 methylation affects gene expression networks.

Our results demonstrate that, global H3K36me3 decreases upon SETD2 depletion and this phenotype enhances OSKM-mediated reprogramming of fibroblasts. Additionally, SETD2 depletion is sufficient to replace KLF4 and cMYC in OS-mediated reprogramming. Downregulation of SETD2 is effective in the early stages of reprogramming and leads to a decrease in somatic specific gene expression levels. This phenotype is also observed upon overexpression of a histone H3 mutant that dominantly blocks H3K36 methylation. We suggest that reduction of H3K36me3 active marks upon SETD2 depletion contributes to OSKM-mediated suppression of fibroblast specific genes.

Taken together these results indicate that SETD2 is a roadblock for reprogramming and reduction of its expression has a positive effect on reprogramming through suppression of somatic specific genes.

ÖZETÇE

Embriyonik kök hücreler (EKH) vücutta bulunan bütün somatik hücreleri oluşturabilme kapasitesine sahiptir. *OCT4*, *SOX2*, *KLF4* ve *cMYC* (OSKM) transkripsiyon faktörlerinin farklılaşmış hücrelere sunulması, yeniden programlamayı uyarıp EKH-benzeri uyarılmış plüripotent kök (uPKH) hücrelerin oluşmasını sağlamaktadır. Yeniden programlama, birden çok aşamalı bir süreçtir ve her aşamasında belirli moleküler değişikliklerin gerçekleşmesi gerekmektedir. Bununla birlikte, kromatin düzenleyicilerinin, somatik hücrelerin gen ifadeleri üzerindeki etkileri ve yeniden programlamayı nasıl etkilediği konusunda birçok cevaplanmamış soru bulunmaktadır.

SETD2, aktif transkripsiyon ile ilişkilendirilen Histone H3 lizin 36 (H3K36) kalıntısının tri-metilasyonundan sorumlu olduğu bilinen tek kromatin düzenleyicidir. H3K36 metilasyonunun yeniden programlama üzerindeki etkisi halen bilinmemektedir. Bu sebeple SETD2'nin insan fibroblast hücrelerinin yeniden programlanması üzerindeki rolünü, shRNA-tabanlı gen susturulması ya da gRNA-tabanlı gen çıkarılması stratejileri ile inceledik. Buna ek olarak, H3K36 metilasyonu kaybının gen ifadesi ağını nasıl etkilediğini anlamak amacı ile, OSKM-sunumu öncesi ve sonrası elde edilen SETD2 ifadesi azaltılmış hücrelere RNA dizileme analizi uyguladık.

Elde ettiğimiz sonuçlara göre, SETD2 ifadesin baskılanması, genel H3K36me3 ifadesinde azalmaya yol açmakta ve bu soytür, fibroblast hücrelerinin OSKM-tabanlı yeniden programlaması arttırmaktadır. Ek olarak, SETD2 ifadesindeki baskılanma, OS-tabanlı yeniden programlama sürecinde *KLF4* ve *cMYC* yerine geçmek için yeterlidir. SETD2 ifadesindeki azalış, yeniden programlamanın erken evrelerinde etkili iken, RNA dizileme ile de anlaşıldığı üzere, bu durum somatik gen ifadelerinin azalmasına yol açmaktadır. Bu soytür, H3K36 metilasyonunu baskın olarak engelleyen histone H3 mutantının aşırı ifadesi sonucunda da görülmektedir. Bu sebeple, SETD2'nin baskılanması sonucunda görülen H3K36me3 aktif işaretindeki azalışın, OSKM faktörlerinin fibroblast-özgü ifadeler üzerindeki baskılayıcı etkisine katkıda bulunduğunu öne sürmekteyiz.

Elde ettiğimiz sonuçlar göz önüne alındığında, SETD2'nin yeniden programlama için bir engel olduğu ve ifadesindeki azalışın somatik gen ifadelerini baskılaması yolu ile yeniden programlamayı olumlu etkilediği görülmektedir.

ACKNOWLEDGEMENTS

It is a genuine pleasure to express my deep sense of thanks and gratitude to my advisor, Dr. Tamer ÖNDER for his generous guidance and support. He made it possible for me to work on my dream topic. It was an honor to work with him and be one of his first students. This thesis would not have been possible without his encouragement and enthusiasm.

I would like to express my great appreciation to my thesis committee members Dr. Nathan Lack and Dr. Tolga Emre for their support, suggestions, and guidance. I am also grateful to Dr. Müjdat Zeybel and Professor Nesrin Özören for their invaluable comments on my thesis and agreeing to become a member of my thesis defense.

I am also grateful to Deniz Uğurlu Çimen for being not only a colleague but also a good friend. We shared a lot of things together and I thank her for all memories we have experienced for the past five years. I would like to thank Hilal Saraç and Filiz Şenbabaoğlu (Hiliz/Filal) for their beautiful friendships. I wish to present my special thanks to Fatma Özgün and Fidan Şeker; we shared a life together and I feel lucky to meet them.

I am indebted to many of my friends especially Önder Lab members for their support and contributions. I also want to thank all members of KUSOM and KUTTAM for creating a friendly environment. It was a pleasure for me to work with them.

Last but foremost, I would like to express my gratitude to my father Şerafet Özçimen, my mother Nurdane Özçimen, and my sister Büşra Özçimen for their endless love and unwavering support. Without them, I would not have come this far.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

ESC: Embryonic stem cell

ICM: Inner cell mass

hESC: Human embryonic stem cell

PCG: Primordial germ cell

SCNT: Somatic cell nuclear transfer

iPSC: Induced pluripotent stem cell

DNMT: DNA (cytosine-5)-methyltransferase

HAT: Histone acetyltransferase

HDAC: Histone deacetylase

SETD2: SETD Domain Containing 2

ccRCC: Clear cell renal cell carcinaoma

MET: mesenchymal-to-epithelial transition

PRC: Polycomb repressive complex

IVF: *in vitro* fertilization

HSC: Hematopoietic stem cell

DSB: Double-strand break

ENCODE: Encylopedia of DNA Elements

IGV: Integrative Genomics Viever

RT qPCR: Real time quantitative polymerase chain reaction

Ct: threshold cycle

NHDF: Normal human dermal fibroblast

GO: Gene Ontology

1. INTRODUCTION

1.1. Developmental potency

A Fertilized egg, also called a zygote, and blastomeres of the morula stage embryo have the ability to differentiate into all embryonic and extraembryonic tissues. These cells are called totipotent, which means ‘whole’ (totus) and ‘powerful and strong’ (potentem) in Latin. The inner cell mass of the developing embryo and the embryonic stem cells which are derived from them are able to differentiate into only embryonic tissues and are thus called pluripotent. Later during development, potency is further restricted. Tissue stem cells which can give rise to only certain types of cells, based on which tissue they reside, are called multipotent or unipotent depending on the variety of cells that they can differentiate into [1]. Finally, cells terminally differentiate into somatic cells, in some cases they become post-mitotic such as neurons.

1.2. Embryonic stem cells

Embryonic stem cells (ESCs) are obtained from the inner cell mass (ICM) of the blastocyst. Their ability to differentiate into the three germ layers (endoderm, ectoderm and mesoderm) and generate all types of cells in the body is called pluripotency. These cells are able to silence both endogenous and exogenous proviral sequences to maintain their genomic stability [2]. ESCs also have self-renewal capacity, meaning they can be propagated in cell culture indefinitely [3]. They express pluripotency-related transcription factors like OCT4, SOX2, and NANOG in addition to cell surface markers such as TRA-1-60, TRA-1-81 and SSEA4. They are positive for alkaline phosphatase, have high telomerase activity and exhibit shortened doubling time due to a reduction of the G1 phase of the cell cycle. ESCs are able to contribute to developing blastocysts resulting in chimeras and can form teratomas when injected into immunocompromised mice, which are two key evidences of their pluripotency *in vivo* [4]. *In vitro*, their ability to differentiate in the three germ layers can also be assessed by their ability to form embryoid bodies [5].

1.2.1. History of Embryonic Stem Cells

Researches on mammalian pluripotent stem cells began in the late 1960s with the studies of spontaneous testicular teratocarcinomas in inbred strains of male mice. These tumors consisted of multiple embryonic tissues and could be propagated by transplantation [6]. Mouse teratocarcinomas are originated from undifferentiated primordial germ cells (PGCs). PGCs differentiate into gametes and their migration to gonads are under the control of *steel* gene which encodes for a peptide growth factor. Heterozygous *steel* mutations cause tumor formation during PCG migration and differentiation [7]. When these tumors are transformed into cells lines and transplanted into the mouse's uterus, they can generate teratoma. This finding demonstrated that teratocarcinomas comprise pluripotent cells that have the capacity to form all types of cells in the body [8, 9].

In 1981 Evans and Kaufman were able to culture pluripotent stem cells obtained directly from mouse blastocysts. The inner cell mass of the blastocysts was isolated and cocultured with mitomycinC-treated STO feeder cells [10]. Obtained colonies showed embryonal carcinoma cell properties and maintained their undifferentiated state in culture with expression of pluripotency-related surface antigen SSEA-1. They were able to form teratomas when injected into mice [11]. Three years later male karyotype-bearing mouse ESC lines were injected into female blastocysts and the contribution of male cells to germ lines of chimeric offspring mice has been shown [12], demonstrating their capacity to differentiate when injected into mice.

The first isolation and *in vitro* cultivation of human ESCs (hESCs) was achieved in 1998 by Thomson and his group. They used human embryos obtained via in vitro fertilization (IVF), which were donated to be used for research purposes. These embryos were cultivated until they formed blastocysts. They isolated ICMs of these blastocysts to obtain hESC lines. Like the other ESCs obtained from nonhuman primates, these hESC lines expressed pluripotency-related markers such as SSEA-4, TRA-1-60, TRA-1-81 and alkaline phosphatase. For ethical reasons, they were not able to use hESCs to test the contribution of these cells into blastocysts and formation of chimeras. However, when hESCs were injected into SCID mice, they formed teratomas comprising the differentiated cell types from all the three germ layers [13].

hESCs are powerful tools to study genetic and molecular mechanisms acting on human development and disease-modeling. These cells are especially important for cell-based therapies and drug discoveries, allowing the investigation of the molecular responses of diverse cell types to many treatment approaches. These pluripotent cells are also important for regenerative medicine; because of their differentiation capacity, hESCs can be used as a source of tissue to treat human diseases and replace damaged or diseased cells [13]. In selected cases, human embryonic stem cells have begun to be used in clinical applications. For example, in 2015 a clinical trial started to treat age-related macular degeneration and Stargart's macular dystrophy patients with transplantation of hESC-derived retinal pigment epithelium cells. Cells were engrafted successfully and 72% of the patients developed increasing retinal pigment epithelium. However, teratoma forming capacity and likelihood of transformation into undesirable cell types, or immune rejection possibility of embryonic stem cells raise concerns about their use in treatment of diseases [14].

1.3. History of Cellular Reprogramming

Somatic cell nuclear transfer (SCNT) is a technique used to microinject nucleus of a fully differentiated cell into an enucleated egg. This technique was first developed in 1938 by Hans Spemann who transplanted nucleus from an salamander embryo to an enucleated egg [15]. In early 1950s, Briggs and King used SCNT technique in their experiments on *Rana pipiens* development. They obtained the nuclei of a late-stage embryo and transplanted them into enucleated oocyte, which was later referred to as 'cloning'. The resulting cells were able to differentiate normally and formed fertile tadpoles, those were later fixed to examine their gastrula stage of development [16].

10 years after Briggs and King's discovery, Sir John Gurdon reported the first cellular reprogramming by generating tadpoles (*Xenopus laevis*) from the nucleus of intestinal epithelium cells transplanted into enucleated oocyte. This breakthrough work destroyed the dogma that terminal differentiation is an irreversible phenomenon [17].

The first mammalian live offspring clone was created in 1996 via microinjection of a nucleus from an adult cell of Finn Dorset sheep into enucleated developing egg cell taken from Scottish Blackface sheep. The hybrid cell was stimulated and activated by electroporation and the resulting embryo was implanted into the uterus of a surrogate mother. The generated offspring was named as 'Dolly' and this work proved that, a differentiated mammalian nucleus can be reprogrammed via factors present in the ooctye [18, 19]. In 2001, the results of another study on hESCs have been announced. In this study, they were able to fuse thymocytes with mouse ESCs and generate hybrid cells with pluripotent characteristics; which indicated that ESCs possess certain factors in their cytoplasm, those are capable of reprogram somatic cells [20].

1.4. Discovery of Induced Pluripotent Stem Cells

To identify pluripotency-related genes that can reprogram somatic cells, Takahashi and Yamanaka performed a genetic screen with 24 transcription factors, those are highly expressed in ESCs and important for the maintenance of stem cell identity. These transcription factors were coexpressed in combinations from retroviral vectors in mouse fibroblasts, and ESC-like pluripotent cells were generated. After one-by-one withdrawal of these factors from the initial reprogramming cocktail, Takahashi and Yamanaka discovered that four transcription factors are sufficient to reprogram mouse fibroblast cells to an undifferentiated pluripotent stem cell state. These four transcriptions factors were *OCT4*, *SOX2*, *KLF4* and *cMYC* (also called OSKM or Yamanaka Factors). The resulting cells displayed characteristic features of ESCs, such as endogenous pluripotency-related gene expressions, increased proliferation and self-renewal, and finally ability to form teratomas [21]. These reprogrammed cells were called induced pluripotent stem cells (iPSCs). One year later, reprogramming of human fibroblast cells to an ESC-like state with the same defined four transcription factors was reported [22]. In 2012, Shinya Yamanaka and John Gurdon were awarded the Nobel Prize in Physiology or Medicine for their contribution to cellular reprogramming.

Somatic cells of several species have been used to generate iPSCs including mouse, human, rat, and rhesus monkey. Moreover, different cells types like keratinocytes, neurons, liver cells, melanocytes, pancreatic β islet cells, and lymphocytes were shown to be reprogrammed into their undifferentiated pluripotent state [1].

1.4.1. iPSC Generation Techniques

1.4.1.1. Reprogramming Factor Selection

Takahashi and Yamanaka were able to convert mouse and human fibroblasts into iPSCs with OSKM factors. Despite these four transcription factors being sufficient to reprogram fibroblasts into pluripotent cells, iPSC generation efficiency was quite low, which led scientists to seek additional factors, possibly more capable of converting somatic cells into pluripotent stem cells. First, alternate transcription factors with similar functions were used to replace Yamanaka factors such as KLF2 rather than KLF4 or L-MYC and MYCN rather than cMYC [23, 24]. In addition to this, somatic cells were reprogrammed with other pluripotency-related factors such as NANOG and LIN28 together with OCT4 and SOX2 [25].

Requirement for the transcription factors are also varying according to the somatic cell type. For example, cMYC-expressing fibroblast cells do not require ectopic expression of cMYC to be reprogrammed. This factor is usually added to the cocktail only to increase the efficiency [24]. Another example is neural progenitor cell reprogramming, which does not require SOX2 and cMYC, since starting cells already express these factors [26].

1.4.1.2. Delivery System Selection

Reprogramming of fibroblast cells into iPSCs was initially performed via retroviral transduction of selected transcription factors. Retroviruses cause transgenes to be randomly integrated into the host genome and generate mutations, which is incompatible with their use in regenerative medicine. Removal of sequences required for genomic integration from the viral systems may be a solution but, transient expressions of these transgenes cause a decrease in the efficiency iPSC generation [27]. Concerns about the use of viral systems to deliver transgenes prompted scientists to try safer ways for factor delivery and iPSC generation.

Using non-integrating vectors to generate iPSCs is excepted to be safer than stable transduction of genome-integrating viruses. For example, adenoviruses are used to deliver transgenes to reprogram mouse and human cells. Transfection of these viruses

into the host cell provides transient expressions of transgenes without genomic integration [28, 29]. Another example for non-integrating system is to use Sendai viruses. The genomic content of Sendai viruses is composed of single-stranded RNA and delivery of transgenes with these viruses does not result in integration to DNA-based genomes. The disadvantage of the Sendai virus is that, they are diluted in each cell division. The decrease in the copy number of viruses in the host cell is resulted in reduced levels of transgene expressions [30].

Sequential transfection of Yamanaka factor-containing plasmids into somatic cells without viral applications is another type of gene delivery system to generate iPSCs [31]. One other integration-free transgene delivery system is episomal transfection. Episomal plasmids are generated from segments of Epstein-Barr virus containing reprogramming factors and they are introduced to the cells by electroporation. The EBNA (Epstein-Barr nuclear antigen-1) sequence provides extrachromosomal replication to the plasmid in each cell division impeding dilution of the plasmid during reprogramming. Still, efficiency of iPSC generation is quite low compared to integrative systems [32].

Supercoiled DNA structures called minicircle vectors, which lack bacterial sequences, are also used for pluripotency-related gene transfer. Transfection efficiency of minicircles is high and they have the ability to express the transgenes in a long period of time in the host cell without genomic integration. These plasmid-like DNA structures are introduced to the cells by nucleofection [33].

Above mentioned techniques use DNA to deliver the genomic information. There are other ways to reprogram somatic cells into iPSCs which do not require this type of nucleic acid. Purified pluripotency-related protein delivery is one of the DNA-free techniques that includes direct introduction of recombinant proteins to the cell. This technique requires sequential transduction of the proteins [34]. Another way is to use synthetic mRNAs obtained via *in vitro* transcription. Contrary to expectations, efficiency is quite high and iPSC generation can also be boosted with the addition of LIN28 RNA to the cocktail [35].

Another important technique for cellular reprogramming is to use small molecule inhibitors. Combinations of small molecules inhibiting certain barriers for reprogramming were first used to generate mouse chemically induced pluripotent stem cells (CiPSCs). This combination of chemicals consisted of valproic acid (VPA), CHIR99021 (a GSK3-beta inhibitor), 616452 (TGF-beta receptor inhibitor) tranylcypromine (LSD1 inhibitor), forskolin (adenylyl cyclase activator) and DZNep (general histone methyl-transferase inhibitor) [36]. However, to date, a chemical combination that can reprogram human somatic cells has not been identified.

Above mentioned discoveries were the consequence of the curiosity to better understand the cellular reprogramming and to find out molecular mechanisms those have role in the iPSC generation. Generation of pluripotent stem cells upon reprogramming of fully differentiated cells is accomplished by reconstitution of chromatin structure that regulates the usage of genomic information [37]. The regulation of gene expressions is under the control of epigenetic mechanisms and this control mechanism ensures differential expressions those distinguish a somatic cell from a pluripotent one [38]. The mechanism by which different types of cells with the same genome have differential expressions and so phenotype is called “epigenetics” [39].

1.5. Epigenetics

The term "Epigenetics'' was first defined in 1942 by Waddington as programmed changes of phenotypes arising from genotypes [40]. Recently, epigenetics is defined as, cellular information that can be inherited independently from genomic changes. In other words, epigenetic is transferable information of reversible changes seen on phenotype without changing the genetic information. Epigenetic patterns that control the gene expression are inherited to daughter cells independent from DNA sequence [41].

1.5.1. DNA methylation

DNA methylation is one of the main epigenetic marks and is generated by the addition of methyl (CH3) group to Cytosine residues [42]. In humans, 70-80% of the CpG dinucleotides are regulated by DNA methylation [43]. Genes those undergo genomic imprinting or inactivation of X chromosome during development are inactivated through CpG methylation [44]. Methylation of CpG islands on promoters and regulatory sequences blocks the transcription factor-binding to those regions [45]. In addition to this, DNA methylation is found to be low in undifferentiated cells and increases throughout differentiation and aging [46].

DNA (cytosine-5)-methyltransferase 1 (DNMT1) is responsible for sustaining the DNA methylation and prefers adding methyl groups to the hemimethylated DNA. DNMT3A and DNMT3B instead, are in charge of *de novo* methylation and they maintain inherited DNA methylation during replication [47]. Methylated CpG dinucleotides are recognized by certain DNA binding proteins such as MECP2 and MBD2, those have repressive properties on gene expression [48-50].

DNA methylation is shown to be a barrier for cellular reprogramming. Decreasing global DNA methylation via using 5-Azacytidine, a chemical analog of Cytidine, which cannot be methylated, was reported to enhance reprogramming. siRNA-mediated inhibition of DNMT1 also improves fully-reprogrammed iPSC generation, which was also indicated by the same group [51]. Downregulation of AID deaminase, a protein which is responsible for conversion of Cytidine to Uracil, was also shown to increase

the reprogramming efficiency of mouse fibroblasts into iPSCs [52]. In addition, overexpression of DNA hydroxlase TET1 has been demonstrated to promote endogenous OCT4 activation by reducing DNA methylation and may even replace the exogenous OCT4 expression in OSKM-mediated cellular reprogramming [53]. All these studies have a common idea that, DNA methylation acts as a barrier for iPSC generation and removal of this epigenetic mark helps to improve cellular reprogramming.

1.5.2. Histone Modifications

Nucleosomes are spool-like protein structures with DNA wrapped around them. This structure ensures DNA packaging in eukaryotic cells. Nucleosomes regulate transcription by forming octamers with H2A, H2B, H3 and H4 Histone proteins [54, 55]. These Histone proteins are modified on their tail extensions, protruded from the octamer, with posttranscriptional changes such as acetylation, methylation, phosphorylation, and ubiquitination. Histone tail sections possess approximately 25-30% of histone mass and providing adequate extensions for protein binding while being sensitive to proteases [56-59]. In the context of reprogramming, acetylation and methylation of histones play particularly important roles and are therefore discussed in more detail below.

1.5.2.1. Histone Acetylation

Acetylation ensures nucleosomes to become less compact and more accessible to transcription factor-binding by neutralizing the positive charges of Lysine residues on histone amino tails [60-63]. Histone proteins are acetylated by histone acetyltransferases (HATs), those can be examined in two groups. TypeA HATs are stabilized in the nucleus and responsible for transcription activation, whereas TypeB HATs, are stabilized in the cytoplasm responsible for the acetylation of newly synthesized proteins before DNA replication [64]. Histone acetyltransferases and histone deacetylases (HDACs) regulates transcription by regulating acetylation of Lysine residues resided at histone tails [55, 65]. In many species, histones are acetylated at histone H3 Lysine 9, 14, 18, and 23 residues, whereas histone H4 proteins are also acetylated at Lysine 5, 8, 12, and 16 residues [56, 66, 67]. The steady

state histone acetylation is balanced by reverse activities of HATs and HDACs. In general, increase in histone acetylation is associated with transcriptional activation and decrease in this mark is associated with gene repression [68-70].

Several studies on the relationship between histone acetylation and cellular reprogramming indicated that, deacetylation of histone proteins impedes reactivation of pluripotent genes during reprogramming, which makes HDACs to be characterized as roadblocks for iPSC generation [71]. Therefore, several HDAC inhibitors are used to prevent histone deacetylation to facilitate activation pluripotency circuitry [72, 73]

1.5.2.2. Histone Methylation

Histone methylations have functions in several biological processes such as heterochromatin formation, X chromosome inactivation, genomic imprinting and silencing of homeotic genes [74-79]. Aberrant histone modifications can cause quite a few numbers of diseases especially cancers [80-84]. Methylation can be present at Lysine(K) and Arginine(R) residues residing at histone amino tails. Lysines those can be methylated are identified as H3K4, H3K9, H3K27, H3K36, H3K79 and H4-K20 [78-86]. Methylation of these residues is associated with gene activation/inactivation and DNA damage [85, 87]. Lysine residues can be mono-, di-, and tri-methylated. Methylated Lysine residues with different levels provide binding regions for chromatin modifiers such as histone methyltransferases, histone deacetylases, and chromatin remodelers resulting in different functions [85]. Euchromatic histone methylations have both activating and inhibiting effect on transcription. Histone H3K4, H3K36 and H3K79 tri-methylations for example, have activating role whereas H3K9 and H3K27 tri-methylations inhibit transcription via promoter inactivation or heterochromatin formation [79].

In previous studies, methylation was thought to be stable and irreversible [88, 89]. However, after the discovery of H3K4-specific demethylase LSD1, it is now known that histone methylation is a reversible and dynamic process [90]. Two years after LSD1 detection, a new histone demethylase, JHDM1 was identified. Unlike LSD1, JHDM1 was discovered to have a JmjC domain that is responsible for demethylation of H3K36me2 [91]. JmjC domain belongs to cupin superfamily of metalloenzymes and

around a hundred proteins possess this domain [92]. JMJD2 is one of the super families with JmjC domain-containing proteins and consisting of JMJD2A, JMJD2B, JMJD2C, and JMJD2D demethylases functioning as Lysine trimethyl-demethylases [93].

There are several reports stating that histone methylation regulates cellular reprogramming acting as enhancers or barriers. For example, SUV39H1/H2, EHMT1/2 and SETDB2 methyltransferases, those are responsible for methylation of histone H3K9 residue, are barriers for reprogramming and their inhibition improves iPSC generation by negatively regulating the repressive H3K9 methylation [94, 95]. Overexpression of KDM2A and KDM2B demethylases were also shown to enhance reprogramming via regulating H3K36 methylation [96]. Another barrier for reprogramming is DOT1L, which performs methylation of H3K79 residue, and its depletion is also enhances iPSC generation.

Considering the activities of histone modifiers and their influence on cellular reprogramming, we focused on SETD2 which is responsible for trimethylation of H3K36 residue [97]. The similarities in between H3K36me3 and H3K79me3 such as their presence on gene bodies and their association with active transcription [94, 98] let us to think that H3K36me3 may also regulate somatic program and its presence may be a barrier for cellular reprogramming. With the confidence that exogenous expression of KDM2B, an H3K36 demethylase, has been shown to have positive effect on iPSC formation [99], it reinforces the idea that SETD2 depletion may have similar effects on reprogramming.

1.6. SETD2

1.6.1. Identification of SETD2 as a methyltransferase

SETD2 protein was first identified in 1998 by its interaction with Huntingtin protein. Faber *et al.* used a yeast-two-hybrid system to find out interacting partners for Huntingtin. Using different fragments of Huntingtin protein as a bait and cDNA libraries as prey; they discovered three novel WW domain-containing proteins interacting with N-terminal of Huntingtin. One of these proteins was SETD2 and named as Huntingtin Yeast Partner B, or HYPB [100]. In two other studies of transcriptome analysis, several

transcripts were identified those were expressed in CD34+ hematopoietic stem/progenitor cells of umbilical cord, one of which was *HSPC069* and possessed the same DNA sequence with *SETD2* [101, 102]. Another broad study performed with cDNA libraries, obtained from adult human brain fractions, identified the cDNA clone KIAA1732 resulting 1325 amino acid-long protein which later turned out to be SETD2 [103].

Rega *et al.* was able to clone full-length SETD2 from HeLa cells with a cDNA of 2061 amino acids. Moreover, this sequence was mapped with FISH probes and found to be located at chromosome 3p21.3-p21-2 [104]. SETD2 protein was later found to have AWS-SET-PostSET domains which functions in Histone H3 Lysine 36 methylation. In this study the interaction between C-terminal 142 amino acids of SETD2 and hyperphosphorylated RNAPII was highlighted [105].

Ten years after its discovery, Edmunds *et al.* first clarified the functional role of this protein on histone methylation in mammals and the name of this protein was converted to SETD2. When they stimulated murine fibroblasts with EGF, H3K36me3 marks were localized on transcriptionally active sites and SETD2 was found to be responsible for only tri-methylation but not mono- and di-methylation of H3K36 residues at those regions. H3K36 tri-methylations were transcription elongation dependent, since using DRB transcription inhibitor decreased H3K36me3 levels on coding regions [106].

1.6.2. Functional domains of SETD2 protein

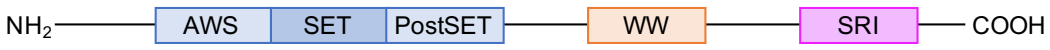


Figure 1. SETD2 protein with its functional domains.
Schematic representation of SETD2 protein with, catalytic AWS-SET-PostSET domain, WW protein-protein interaction domain, and SRI RNAPII binding domain.

1.6.2.1. AWS-SET-PostSET domain

The name of SET domain comes from ‘Suppressor of Variegation, Enhancer of zeste and Trithorax’ genes of *Drosophila melanogaster*. Human SET domain is an evolutionary conserved motif comprising 130 amino acids. SET domain is located in

between cysteine-rich AWS (associate with SET) and PostSET domains which are required for SET domain methyltransferase activity [74]. The catalytic SET domain of SETD2 uses *S*-adenosyl-*L*-methionine (SAM) as a substrate to transfer the third methyl group to lysine located at 36^{th} residue of Histone H3 [107].

1.6.2.2. WW domain

This domain is composed of hydrophobic, aromatic, and proline-rich 38 amino acids. Name of this domain comes from two highly conserved tryptophans (WW) resides 20-22 amino acids apart. WW domain have homology with SH2 and SH3 domains. They both recognize proline-rich proteins with Proline-Proline-x-Tyrosine (PPxY) motifs to mediate protein-protein interactions [108, 109]. WW domain of SETD2 interacts with proline-rich C–terminal segment of Huntingtin protein [100].

1.6.2.3. SRI domain

Rpb1 is the largest domain of RNAPII with C-terminal region composed of heptapeptide repeats. Phosphorylated serine amino acids at 2nd and 5th residues of heptapeptides are recognized by Set2 Rpb1 Interacting (SRI) domain of SETD2. Binding of SRI domain to elongating RNAPII recruits SETD2 to transcribed regions [110] which may be required for H3K36 trimethylation activity of this methyltransferase. Consistent with this notion deletion of SRI domain in yeast Set2 leads to complete loss of H3K36me3 [111].

1.6.3. Biological Functions of SETD2

1.6.3.1. Trimethylation of Histone H3 Lysine 36 Residue by SETD2

SETD2 is responsible for trimethylation of lysine 36 residue on histone H3. Evidence for this comes from several studies. Human SETD2 performs H3K36me3 in a complex including heterogeneous nuclear ribonucleoprotein L (HnRNP-L); since siRNA-mediated downregulation of either of these proteins decrease trimethylation levels but not mono- and di-methylation of H3K36 residue *in vivo* [112]. In another study, recombinant nucleosomes were used to characterize the substrate specificity of SETD2. Introducing lysine-to-alanine mutation at histone lysine residues, SETD2 found to be responsible for methylation of lysine at 36th residue but no other lysines of histone H3 protein, demonstrating its substrate specificity to H3K36 residue [97].

1.6.3.2. SETD2 Interaction Partners and Their Role in SETD2 Function

SETD2 protein was discovered when its interaction with Huntingtin protein was first identified. According to this study, N-terminus of Huntingtin protein was detected by 13 distinct proteins three of which possessing WW protein-protein interaction domain. HYPB, which was further named as SETD2, was one of these three WW domain-containing proteins and it had the ability to interact with both normal and mutant Huntingtin in lymphoblastoid cells [100]. In another study, SETD2 protein was found to contain polyproline (polyP) stretch at its C-terminus directly interacts with the following WW domain to form a closed conformation state. This inhibition of WW domain with polyp stretch impedes SETD2 binding to proline-rich region of Huntingtin [113].

Some of the PRC2 complex components recognize not SETD2 but its functional result, H3K36me3 marks, and this recognition is required for recruitment of PRC2 complex to certain transcribed regions. *PHF1* gene encodes for polycomb group PHD Finger Protein 1 (PHF1), a histone H3 Lysine27-specific methyltransferase complex (PRC2) component. PRC2 complex activity associated with its recruitment to the transcribed regions by H3K36me3 mark. Tudor domain of PHF1 recognizes H3K36me3 and PRC2 complex is localized at DNA double strand break sites via this interaction [114]. Another study showing that, N-terminal domains of two PRC2 complex components PHF1 and MTF2 bind to CpG clusters especially to the unmethylated ones where nucleosomes contain H3K36me3 marks [115].

IWS1/SPT6 protein partners are involved in interaction of SETD2 with RNAPII elongation complex. Binding of SETD2 to serine2-phosphorlylated CTD of RNAPII and its recruitment to the transcribed regions are ensured by IWS1. Like SETD2 depletion, downregulation of IWS1 protein causes increase in H3K36me3 levels at 5' ends, with substantial decrease in H3K36me3 levels at 3' ends of the transcribed genes [116].

SETD2 stability and its degradation are regulated by proteasome complexes where SPOP, a specific subunit of CUL3 ubiquitin E3 ligase complex, interacts with SETD2 and takes charge of polyubiquitination required for degradation [117].

1.6.3.3. Role of H3K36me3 in Transcription

SETD2 binds to serine-2 phosphorylated CTD of elongating RNAPII and is carried to coding regions of actively transcribed genes where it trimethylates histone H3K36 residues. Inhibition of RNAPII phosphorylation is resulted in displacement of H3K36me3 toward the 3' end of the genes demonstrating that SETD2 is associated with transcription [118]. Upon growth factor stimulation, H3K36me3 levels increases at 3' end of responsive genes concomitant with transcriptional activation [106]. In addition, DNase-I hypersensitive sites on three mammalian cell lines demonstrates that, localizations of serine2-phosphorylated RNAPII and H3K36me3 are highly correlated and their co-presence in certain regions assisted to predict actively transcribed genes [119].

Transcription is supported by and regulated with the accumulation of H3K27me1 marks on transcribed regions. H3K27me1 deposition on the intra- and intergenic regions is also copresent with SETD2-mediated H3K36me3 and they both are enriched in actively transcribed genes, which associates H3K36me3 with transcriptional regulation [120].

1.6.3.4. SETD2 Role in Preventing Spurious Transcription

H3K36me3 marks prevents spurious transcription sites by recruiting histone deacetylase complexes to the transcribed regions [121]. The association between H3K36 trimethylation and histone deacetylation was discovered in *S.cerevisiae* and has not been observed in mammalian gene regulation systems. Yet, inhibition of spurious transcription in mammals is regulated by another mechanism in which SETD2 is involved. 11% of the active genes initiate intragenic transcription upon SETD2 downregulation. SETD2 has role in loading of FACT (Facilitates Chromatin Transcription) complex components SPT16 and SSRP1 to the active sites. FACT is a histone chaperone and its recruitment to coding regions represses transcription via blocking cryptic promoters within the genes. Decrease in SETD2 expression levels causes reduction in nucleosome occupancy. SPT16 recruitment is also shown to require chromatin regions includes H3K36me3-modified nucleosomes [122].

SETD2 has role in DNA methylation which is also important for transcription regulation. DNMT3B binding to the genic portions of transcribed genes has been reported in

mouse stem cells. PWWP domain of DNMT3B recognizes SETD2-mediated H3K36me3 marks and is recruited to the transcribed regions [123]. DNTM3B-regulated DNA methylation protects active regions from undesirable RNAPII entry and initiation of cryptic transcription. This discovery also shows how SETD2 regulates and prevents spurious transcription in mammals [124].

1.6.3.5. Role of SETD2 in Splicing Machinery

One other function of SETD2 is to regulate splicing during transcription. Intron-containing genes are enriched for H3K36me3 relative to intronless genes. H3K36me3 levels at intron-containing genes are proportional to their transcription levels. Defects in splicing machinery lead to decreased recruitment of SETD2 and causes reduction of H3K36me3 levels. In addition, splicing activity increases with the increase in H3K36me3 [125]. One of the SETD2 complex component hnRNP-L is also a regulator for splicing machinery. Interaction of hnRNP-L with SETD2 regulates intron retention in human cells [126].

H3K36me3 is known to be enriched at alternatively spliced exonic regions. For example, *GLA* mutations in Fabry disease causes aberrant alternative splicing and those sites are enriched with H3K36me3 [127]. Another disease-related example is downregulation of *MECP2*, which is involved in mRNA splicing, initiates enrichment of H3K36me3 levels at spliced regions in mouse primary neurons [128]. In ICF1 syndrome, mutations in *DNMT3B* impair exon splicing in addition to effects on other transcription regulation mechanisms. H3K36me3 levels are seen to be decreased at splice-defected regions and redistributed over gene bodies [129].

1.6.3.6. SETD2 Role in Other Cellular Activities

There is not much information about SETD2 effect on immune response, but methylation of nonhistone proteins by SETD2 reveals that SETD2 has other functions besides histone H3K36 trimethylation. In the literature there are two notable examples of non-histone targets of SETD2. One of these nonhistone targets is STAT1, which plays an important role in IFNα-mediated immune response. Hepatocyte-specific deletion of SETD2 in mice causes elevation of Hepatitis B surface antigen (HBsAG) levels. This phenomenon is linked to SETD2-mediated methylation STAT1 on lysine

525. STAT1 methylation by SETD2 amplifies activation of antiviral cellular response [130].

Another important function of SETD2 is regulation of cytoskeletal remodeling. SETD2 is responsible for methylating lysine 40 residue of α-Tubulin protein during mitosis and cytokinesis. In addition to this, SETD2 deletion prevents αTubK40 methylation and leads to mitotic spindle and cytokinesis defects [131].

1.6.3.7. Role of SETD2 in Stem Cells and Development

SETD2 protein and its catalytic functions have important roles in stem cell maintenance and development. Embryonal carcinoma cells have open chromatin structures at pluripotency-related regions such as *OCT4*, *SOX2* and *NANOG* relative to more differentiated cell types. These genic regions have nucleosomes towards the 3' end of the gene bodies which are enriched with trimethylated Histone H3 at lysine 36 residues [98].

SETD2 expression level changes during developmental stages. For example, porcine sperms are shown to be rich in *SETD2* expression which decreases immediately after *in vitro* fertilization (IVF). 2-cell stage embryo again gains *SETD2* expression and peaks at 4-day stage of developmental period. Besides, embryos obtained by SCNT has *SETD2* expression throughout development [132]. Another study on porcine oocytes demonstrated that, Vitamin C have positive influence on meiotic maturation by decreasing H3K36me3 levels. When oocytes are treated with L-ascorbic acid 2-phospahte (AA2P), which is necessary for vitamin C synthesis, developmental competency of oocytes are increased by disappearance of H3K36me3 marks [133]. Germ cell-specific *SETD2* knockout mice model also points to the importance of *SETD2* expression for spermatogenesis and acrosome formation. Those mice are infertile due to loss of H3K36me3-related complications in sperm maturation [134].

Homozygous disruption of *SETD2* results in lethality during mouse development. *SETD2* knock-out mouse embryos lose H3K36me3 marks and die at E10.5-E11.5 days of embryonic stages. *SETD2* disruption causes vascular defects, which points to a role for SETD2 in angiogenesis [135]. *SETD2* expression induces primitive endoderm differentiation of mESCs. *SETD2* knockout mESCs are shown to have decreased *ERK* expression which results in downregulation of endoderm-related genes. In mESCs,

FGFR3 expression is regulated by SETD2-mediated H3K36me3 and *FGFR3* transgene expression rescues aberrant *ERK* expression and primitive endoderm differentiation [136].

SETD2 expression also affects hematopoietic stem cell (HSC) self-renewal and differentiation. Differentiation capacity of HSCs to multipotent progenitor cells decreases in conditional *SETD2* knock-out mice. [137]. In another mouse model containing SETD2 exon 6 knockout results in leukopenia, anemia, mild fibrosis in bone morrow and other HSC differentiation-defect related diseases [138].

Mammalian X chromosome inactivation is mainly regulated by the expressions of noncoding Xist and its antisense Tsix RNAs. Experiments on mouse male embryonic stem cells show that, H3K36me3 is recruited to Xist promoter on X chromosome causing suppression of Xist expression. Substitution of lysine 36 to methionine in histone H3 protein leads to decrease in H3K36me3 levels and removal of Xist repression. Moreover, Xist expression increases upon SETD2 depletion. These findings demonstrate that, X chromosome inactivation process is regulated by SETD2-mediated H3K36me3 [139].

1.6.3.8. SETD2 Role in Repair Mechanisms and Mutations at H3K36 residue

Chondroblastoma is a type of benign tumor affecting bones and originating from increase in number of immature cartilage cells affecting children and young adolescent. 90% of chondroblastoma types is determined to contain Lysine36-to-Methionine36 mutation in H3.3 variant of histone H3 protein. This mutation decreases the catalytic activity of SETD2 protein to methylate H3 proteins, thus causes decrease in global H3K36me3 levels and a concomitant increase in H3K27me3 levels. Augmentation of H3K27me3 levels by redistribution of PRC1 complex causes repression of certain genes related to mesenchymal differentiation [140]. Giant cell tumor of bone (GCTB) is another type of pediatric cancer effecting bones, bearing H3K36M mutation with 92% incidence. Therefore, H3K36M mutations are used as biomarker for diagnosis of bone tumors [141].

Another cancer type affected by H3K36me3 levels is pediatric glioblastoma (GBM). Histone H3.3 with G34-to-R or G34-to-V substitution mutations prevents SETD2

binding to histones and trimethylation of H3K36 residue. However, this mutation only changes redistribution of H3K36me3 and promotes tumorigenesis without changing global H3K36me3 levels [142]. Histone mutations affecting H3K36me3 levels and SETD2 catalytic activity have been shown in other and tumor types tissue formations such as adipose [143], and fibromatosis [144].

SETD2 also have influence on repair mechanisms. Proteins involved in double strand break (DSB) repair mechanisms are recruited to H3K36me3-enriched active regions [145]. Mutations in SETD2 or loss of H3K36me3 contribute to tumor progression via creating problems in repair system. As an example, depletion of SETD2 interfere with homology-directed repair and prevents homologous recombination at double-strand break (DSB) regions [146]. Decreased H3K36me3 levels has been shown in patients with skin lesions resultant of arsenic exposure. The diverse effect of H3K36me3 on mismatch repair system (MMR) and following promoter hypermethylation leads to accumulation of DNA damage [147].

SETD2-facilitated phosphorylation and activation of p53 protein induces apoptosis after DNA damage response. ATM is a downstream effector of p53 and activated by the presence of DNA DSBs. SETD2 protein also recruits RAD51 which plays role in homologous recombination at DSB regions. SETD2 loss or mutations lead to chromosome instability as a result of deficiency in DSB-repair system [148]. Histone H3K36M mutations also inhibits H3K36me3 and so RAD51 recruitment to DSB sites, which prevents homology-directed repair [149].

Acetylation of H4K16 residue is stimulated by H3K36me3 at DSB sites of active DNA regions. H4K16ac decides which repair mechanism is to be used. Both marks are elevated at DSB regions after cells are treated with DNA damaging agents [150]. SETD2 loss leads to chemoresistance to DNA-damaging agents via inhibiting apoptosis. Mutations are accumulated in the cells since proteins involved in DNA damage response cannot be recruited to DSB sites by demolished H3K36me3 marks [151].

PALB2 is a partner of BRCA2 protein and functioning in restoration of broken DNA structures and protection against DNA damage. Depletion of MRG15 transcription

factor, an interaction partner of PALB2, impairs PALB2 loading to the chromatin. MRG15 is known to recognize and bind to trimethylated H3K36 histones and recruited to actively transcribed regions together with PALB2. Therefore, protection of active genic loci from genotoxic stress is ensured by H3K36me3-MRG15-PALB2 interaction [152, 153].

During G1 and early S phase of mitosis human MutSalpha is recruited to newly replicated DNA via its PWWP domain interaction with H3K36me3 marks. SETD2-mediated H3K36me3 protects DNA from mismatch mutations and loss of either SETD2 or H3K36me3 induces microsatellite instability [154]. SETD2 mutations leads to abolishment of H3K36me3 marks which is required for MutSalpha recruitment. Frequency of microsatellite instability increases as a result of defects in the mismatch-repair system [155].

1.6.3.9. The Relationship Between SETD2 Activity and Cancer

Renal cell carcinoma (RCC) comprises 90-95% of all kidney-related cancers [156]. RCC is characterized by mutations or deletions in short arm of chromosome 3 including genes *VHL*, *PBRM1*, *BAP1* and *SETD2*. The most common subtype of RCC is clear cell renal cell carcinoma (ccRCC) and associate with inactivation of von Hippel-Lindau (VHL), tumor suppressor gene. *VHL* gene which is located at chromosome 3, in close proximity with *SETD2* gene. Large deletions at p arm of chromosome 3 affects both genes and cause ccRCC [157].

Sarcomatoid feature of ccRCC is associated with poor prognosis in which SETD2 contains somatic single nucleotide variants [158]. Loss-of-function mutations SETD2 lead to H3K36me3-negative tumors, those have 2-fold mortality risk than SETD2-intact tumor types, which shows SETD2 association with survival rate of ccRCC patients [159]. Non-promoter regions are DNA-hypomethylated due to SETD2 loss, which correlates with H3K36me3-directed DNMT3A recruitment [160]. Tumors bearing wild type *SETD2* preserves genome integrity, but *SETD2*-negative tumors have increased DNA damage, explaining importance of *SETD2* in genome stability and DNA repair mechanisms [161]. In non-metastatic ccRCC, inefficient *SETD2* expression and decreased level of H3K36me3 lead to large tumor size [162]. About 500 different

SETD2 mutations accounting for 12% of all incidences have been characterized in ccRCC, most of them are missense and frameshift mutations inhibiting *SETD2* activity [163].

Renal cell carcinoma is not the only tumor type consisting *SETD2* inactivation. *SETD2* is also mutated in most of the breast cancer patients. When breast cancer tumors are compared with noncancerous breast tissue in terms of gene expressions, *SETD2* expression is indicated to be low in tumors. In this case, it is clear that low levels of SETD2 is not sufficient to perform its tumor suppressor functionality [164]. Several other studies demonstrate that, *SETD2* mutations and expression profiles contribute to tumorigenesis in breast cancers [165-167]. Truncating mutations in *SETD2* gene are seen in high grade pediatric glioma tumor types with a 15% incidence. *SETD2* mutations classifies glioma patients, where *SETD2* is mutated in young patients and H3.3G34V/R mutations are seen in IDH mutant glioma types [168]. In diffuse cerebellar glioma type, *SETD2* truncating mutations and H3.3K36M mutations are also mutually exclusive [169]. *SETD2* alterations are considered as driver mutations for high grade gliomas [170]. Based on this set of evidence, SETD2 is thought to act as a tumor suppressor.

1.7. Phases of Reprogramming and Chromatin Factors Regulating iPSC Generation

Introduction of Yamanaka factors are sufficient to convert somatic cells into iPSCs. However, low efficiency of iPSC generation indicates that there are barriers blocking the cells from losing their lineage identity and become fully reprogrammed. Studies demonstrated that, inhibition of certain genes removes these obstacles during reprogramming and results in more efficient iPSC generation. Cellular reprogramming is a stepwise mechanism composed of three main stages (Figure 2). In each stage cells need to undergo certain changes to reach the fully-reprogrammed stage by reactivation of core pluripotency circuitry and barriers affecting to each stage differently.

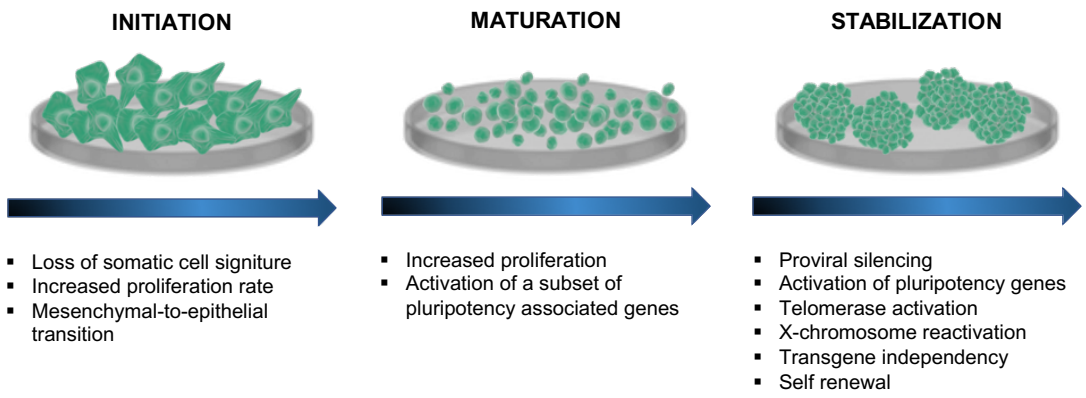


Figure 2. Illustration of somatic cell reprogramming stages.
Key molecular events for each cellular reprogramming stages are listed with bullet points.

Reprogramming of fibroblast cells first results in downregulation of somatic gene expressions and starts a mesenchymal to epithelial transition (MET) by activating E-cadherin and BMP receptor signaling [171-173]. Following MET, pluripotency-related early response genes such as *SSEA1*, alkaline phosphatase, *FBXO15* are activated [171, 174]. Core transcription factors for pluripotency like OCT4, SOX2 and NANOG start to be expressed in the late period. Expression of these factors is the main indicator of fully reprogrammed cells. In addition, reprogrammed cells become independent from transgene expressions and silenced [175].

Reprogramming of somatic cells into pluripotent stem cells requires a global change in transcriptional networks. Diverging from lineage identity and acquisition of pluripotency-related marker expressions are achieved by epigenetic regulations that affect gene expressions patterns. Somatic chromosome structure undergoes hypomethylation to provide accession of transcription factors to pluripotency-related gene promoters. Such promoters are regulated by bivalent histone methylation marks as H3K4me3/H3K27me3 [51, 176]. Epigenetic changes during differentiation process of ESCs to lineage-specific conversion is reverted, euchromatic regions gain DNA methylation while H3K9 trimethylated heterochromatin structures are targeted by demethylases [177].

Several genes show variable expression patterns in the initial stages of OSKM-induced reprogramming. These genes can also vary due to starting cell type since stochastic gene expressions are seen in different cells. Those genes are related to MET

proliferation and metabolism. ESC-like morphological changes also appear in the initial stage and cells those possess such changes are prone to be reprogrammed [173, 178, 179]. Initial-to-intermediate phase transition have a gate where cells are subjected to bottleneck effect. Only the cells who pass the requirements of reprogramming at initial stage can continue to the intermediate stage where others are stuck in a partially-reprogrammed stage. The late stage is interpreted as deterministic or hierarchical phase where *NANOG* starts to be expressed and pluripotency circuitry is activated. DNA methylation and X chromosome reactivation events also occur in the deterministic late phase.

Studies on fully/partially reprogrammed cells in terms of their expression patterns relative to differentiated cells help to understand insight of iPSC generation. Activating histone modifications such as H3K9ac, H3K27ac, H4K5ac, H3K4me3, and H3K36me3 appeared to be higher in fully reprogrammed cells. On the other hand, the deposition levels of these marks were shown to be similar in partially reprogrammed cells and fibroblasts. Heterochromatin conversion of actively transcribed regions is observed in early time points of reprogramming which is followed by activation of repressed genes especially pluripotency-related ones [180].

Distribution of histone modifications start to change immediately after OSKM induction. Promoters and enhancers of actively transcribed lineage-specific genes lose H3K4me3 marks, and H3K4me2 is deposited at the promoter regions of pluripotency genes where OCT4 and SOX2 binding sites are present [181]. Repressed promoters are gradually depleted of H3K27me3 and are hypomethylated to become fully accessible for transcription factor binding at late stage [179].

Changes in the gene expressions during reprogramming process are controlled by epigenetic modifiers (Table 1 and Table 2). Mammalian Trithorax group protein WDR5 for example, is an effector protein for H3K4 methylation and its interaction with OCT4 provides their translocation to promoter regions of genes those related to pluripotency and self-renewal [182]. Removal of repressive marks on early activated pluripotency gene promoters is achieved by KDM6A protein, which interacts with OSK transcription factors and act as a H3K27 demethylase. Depletion of KDM6A inhibits reprogramming due to aberrant H3K27me3 distribution [183]. Like histone demethylases, DNA

demethylases also have roles in reprogramming. TET1 and TET2 proteins are responsible for the conversion of 5-methylcytosines into 5-hydroxymethylcytosines which eventually lose a methyl group and become cytosines in the DNA structure. TET1 and TET2 are OCT4-regulated enzymes and upon OSKM induction, they regulate expressions of pluripotency-related genes such as *NANOG*, *ESRRB* and *OCT4* [184].

KDM2A and KDM2B are two demethylases whose expression decreases global H3K36me2 levels. Removal of H3K36me2 marks on promoter regions of early-activating genes by KDM2B, increases OCT4-GFP positive cell numbers in mouse fibroblast reprogramming [99]. The positive effect of vitamin C supplementation on reprogramming has been demonstrated. KDM2A and KDM2B act as downstream effectors of vitamin C and enhance reprogramming by regulating H3K36 methylation [185].

Other enhancers of reprogramming process are H3K9 and H3K27 methylation-related Polycomb repressive complex components. EHMT1 and SETDB1 are responsible for H3K9 methylation and their catalytic activities silence somatic gene expressions. PRC2 complex components such as EZH2, EED and SUZ12 catalyze H3K27 methylation and act similarly to suppress lineage-specific genes in early stages of reprogramming. As expected, downregulation of these genes reduces iPSC generation. DOT1L on the other hand, is the only methyltransferase performing all H3K79 methylation which is correlated with active transcription. Depletion of DOT1L reduces global H3K79me2 levels and results in efficient suppression of lineage-specific regulators of fibroblasts in initial stages of reprogramming. This phenomenon allows a proper MET which is required for iPSC generation [94].

F-BOX E3 ligase (FBXW7) is a type of E3 ubiquitin ligase with specific targets for ubiquitin conjugation. This protein regulates ubiquitination of certain proteins for degradation including cMYC, transcription factor used for cellular reprogramming. siRNA-mediated downregulation of FBXW7 in mouse fibroblasts increases alkaline phosphatase positive colonies with expressing endogenous *OCT4* and *NANOG* [186].

BMP4 and vitamin C were shown to regulate the H3K9me3 levels during mouse cell reprogramming. Considering the negative effect of BMP4 induction and positive effect of vitamin C supplementation on iPSC generation, H3K9me3 was thought to be a barrier for reprogramming. As suspected, downregulation of SUV39H1, SUV39H2, EHMT2 and SETDB1, regulators of H3K9 methylation, improved reprogramming of pre-iPSCs those were arrested at intermediate stage [187].

In the light of all these studies on cellular reprogramming and epigenetic regulations during developmental processes, we hypothesized that, SETD2 may be a barrier for iPSC generation and depletion of this protein may positively affect reprogramming process. Several experiments that we performed demonstrated that, downregulating SETD2 in initial stages facilitates reprogramming of fibroblast cells into iPSCs. Prior to our work, there is no other study in the literature identifying the role of SETD2 in cellular reprogramming or describing their relation. In addition to this, using mutant histones to regulate iPSC generation was another study that we performed, which has no similar work has been done in the reprogramming field.

Table 1. Histone methyltransferases, those regulating cellular reprogramming				
Modifier	**Activity**	**Target**	**Action**	**Reference**
WDR5	methyltransferase	H3K4	enhancer	[182]
KMT2A/ MLL1	methyltransferase	H3K4	enhancer	[188]
KMT2D/ MLL2	methyltransferase	H3K4	enhancer	[189]
KMT7/ SETD7	methyltransferase	H3K4	barrier	[190]
KMT3E/ SMYD3	methyltransferase	H3K4	enhancer	[191]
KMT1A/ SUV39H1	methyltransferase	H3K9	barrier	[94, 187]
KMT1B/ SUV39H2	methyltransferase	H3K9	barrier	[94, 187]
KMT1D/ EHMT1	methyltransferase	H3K9	barrier	[94]
KMT1C/ EHMT2	methyltransferase	H3K9	barrier	[94, 187]
KMT1E/ SETDB1	methyltransferase	H3K9	barrier	[95, 187]
KMT1F/ SETDB2	methyltransferase	H3K9	enhancer	[191]
KMT6A/ EZH2	methyltransferase	H3K27	enhancer	[94, 192]
KMT3B/ NSD1	methyltransferase	H3K36	enhancer	[193]
KMT3G/ NSD2	methyltransferase	H3K36	enhancer	[191]
KMT4/ DOT1L	methyltransferase	H3K79	barrier	[94]

Table 2. Histone demethylases, those regulating cellular reprogramming				
Modifier	**Activity**	**Target**	**Action**	**Reference**
KDM1A/ LSD1	demethylase	H3K4	barrier	[194]
KDM5A/ JARID1A	demethylase	H3K4	enhancer	[195]
KDM5B/ JARID1B	demethylase	H3K4	barrier	[196]
KDM3A/ JMJD1A	demethylase	H3K9	enhancer	[197]
KDM3B/ JMJD1B	demethylase	H3K9	enhancer	[197]
KDM4B/ JMJD2B	demethylase	H3K9	enhancer	[197, 198]
KDM4C/ JMJD2C	demethylase	H3K9	enhancer	[197]
KDM1B/ LSD2	demethylase	H3K9	barrier	[199]
KDM6A/ UTX	demethylase	H3K27	enhancer	[183, 195, 200]
KDM6B/ JMJD3	demethylase	H3K27	enhancer	[195]
KDM2A/ JHDM1A	demethylase	H3K36	enhancer	[99, 185]
KDM2B/ JHDM1B	demethylase	H3K36	enhancer	[99, 185]

2. METHODS

2.1. Bacterial Experiments

2.1.1. Competent bacteria preparation

Competent bacteria preparation protocol was adapted from Hanahan *et al*. [201]. RF1 (Table 3) and RF2 (Table 4) buffers were prepared with given reagents in required amounts. Single bacteria colony was picked and cultured overnight in 2ml of antibiotic-free LB broth. 0.5 ml of the culture was diluted into 100 ml of LB broth and incubated for 3 hrs. Using LB broth as control, optic density (OD) of the of the culture was measured at 600 nm until it was in between 0.4 and 0.6. When the required OD was achieved, culture was incubated on ice for 15 min and then centrifuged at 3000 rpm for 15 min. Bacterial pellet was resuspended with 33 ml of RF1 buffer and incubated on ice for 15 min and then centrifuged as before. Bacterial pellet was resuspended with 8 ml of RF2 buffer and 200µl of the bacteria was aliquoted into 1.5 ml centrifuge tubes on ice and snap frozen in liquid nitrogen.

Table 3. RF1 Buffer preparation		
	Amounts	Final conc.
RbCl	12 g	100 mM
MnCl2.4H2O	9.9 g	50 mM
CH3CO2K	30 ml from 1M stock	30 mM
CaCl2	1.1 g	10 mM
Glycerol	150 g	15% (w/v)
dH2O		up to 1 L
pH was adjusted to 5.8 with 0.2 M CH3COOH		

Table 4. RF2 Buffer preparation		
	Amounts	Final conc.
MOPS	20 ml of 0.5 M stock	10 mM
RbCl2	1.2 g	10 mM
CaCl2	8.25 g	75 mM
Glycerol	150 g	15% (w/v)
dH2O		up to 1 L
pH was adjusted to 6.8 with NaOH		

2.1.2. Bacterial Transformation

Plasmids were amplified and isolated with NucleoBond midi plasmid DNA purification kit (MN 740410.10). 50 ng of plasmid was mixed with 50 µl of competent Stabl3 bacteria and mixture was heat shocked at 37ºC for 30 sec and then placed immediately on ice. 200 µl of LB broth was added 5 min after heat shock and bacteria was shaken at 37ºC for 15 min at 225 rpm. 100 µl of bacteria was spread onto agar plates containing antibiotic (ampicillin or kanamycin) and incubated for 16 h at 37ºC. Single colonies were picked and inoculated into 200 ml of LB broth and then incubated for 16 h at 37ºC by shaking at 225 rpm.

Plasmid DNA was extracted using NucleoBond midi plasmid DNA purification kit according to manufacturer's instructions. Concentration of purified plasmid DNA was determined with Nanodrop. 1000 ng of plasmid was digested with suitable endonucleases for 1 h and run on 1.5% agarose gel for diagnostic purposes. When expected bands were obtained, plasmids were used for further experiments.

2.1.3. Cloning experiments

2.1.3.1. Cloning of SETD2 shRNAs into pSMP vector

shRNA oligo sequences (Table 5) were designed using RNAi Codex tool: (http://cancan.cshl.edu/cgi-bin/Codex/Codex.cgi).

Table 5. shRNA oligo sequences

shRNA oligos	Sequences
shCntrl	5'_TGCTGTTGACAGTGAGCGCCCGCCTGAAGTCTCTGATTAATAGTGAAGCCA CAGATGTATTAATCAGAGACTTCAGGCGGTTGCCTACTGCCTCGGA_3'
shSETD2-1	5'_TGCTGTTGACAGTGAGCGCGCTGACTCACGGTGTTATGAATAGTGAAGCCA CAGATGTATTCATAACACCGTGAGTCAGCTTGCCTACTGCCTCGGA_3'
shSETD2-2	5'_TGCTGTTGACAGTGAGCGCGCCATGTTACTTTGATCTTATTAGTGAAGCCA CAGATGTAATAAGATCAAAGTAACATGGCATGCCTACTGCCTCGGA_3'
shSETD2-3	5'_TGCTGTTGACAGTGAGCGCGGAATTAGACTCTTTATCTAATAGTGAAGCCA CAGATGTATTAGATAAAGAGTCTAATTCCTTGCCTACTGCCTCGGA_3'

Designed shRNA oligos were synthesized by Macrogen (http://www.macrogen.com/en/main/index.php) and then amplified with primers (Table 6) compatible with restriction ends of pSMP vector.

Table 6. Primers for pSMP shRNA cloning	
Forward	5'-GATGGCTGCTCGAGAAGGTATATTGCTGTTGACAGTGAGCG-3'
Reverse	5'-GTCTAGAGGAATTCCGAGGCAGTAGGCA-3'

PCR reaction to amplify shRNA oligos was prepared with following reagents (Table 7):

Table 7. PCR reaction mix for shRNA oligo amplification			
	Stock conc.	Final Conc.	1x reaction volume
DreamTaq	2X	1X	25 µl
Betaine	5 M	1 M	10 µl
dNTPs	2.5 mM	50 µM	1 µl
Primer mix	25 µM	0.5 µM	1 µl
Template oligo	10 µM	0.1 µM	0.5 µL
dH_2O			12.5 µl

Amplified PCR product was run on 1.5% agarose gel containing 1% EthBr (v/v) then desired band (132 bp) was cut from the gel and DNA was purified from the gel with NucleoSpin PCR clean-up Gel extraction kit (MN 740609.50) as described in manufacturer's instructions. Obtained DNA was digested with EcoRI and XhoI for 3 hours. 5 µg of pSMP backbone was also digested simultaneously with same enzymes. After 3 hours of enzymatic digestion of pSMP backbone, 1µl of Antarctic phosphatase (NEB M0289S) was added into the reaction to prevent self-ligation and they were incubated for 15 min at 37°C then inactivated at 70°C for 30 min. Digested backbone was loaded onto 1.5% agarose gel and desired band (6699 bp) was obtained then purified with NucleoSpin PCR clean-up Gel extraction kit. Digested oligo fragment was also purified with using PCR purification protocol of the same kit. Concentrations of both oligo and backbone was measure with Nanodrop.

Digested shRNA oligo and pSMP backbone was ligated as 1:3 molarity ratio with Quick ligation (NEB 2200S) protocol according to manufacturer's instructions. Ligation mix (Table 8) was incubated at room temperature (RT) for 10 min and 5 µl of ligated vector was transformed into Stabl3 competent bacteria. Transformation was performed with 50 µl of bacteria and 5µl of ligation product. Single colonies were picked, and plasmid DNA was obtained with mini prep. Clones were verified by digesting with EcoRI/NdeI, and Sanger sequencing with MSCV primer (Figure 3).

Table 8. Ligation mix for shRNA cloning	
Reagents	20 µl reaction volume
Quick Ligase Reaction Buffer (2X)	10 µll
Backbone (~6500 bp)	100 ng
Insert DNA (~120 bp)	4.6 ng
dH_2O	Up to 20 µl

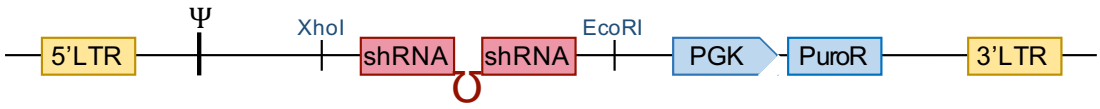


Figure 3. Schematic representation of pSMP plasmid with shRNA insert.
shRNA oligos were amplified and cloned into pSMP backbone in between EcoRI and XhoI cut sites.

2.1.3.2. Cloning of SETD2 gRNAs into V1 pLentiCrispr vector

gRNA sequences were designed using Crispr design tool (http://crispr.mit.edu/). Compatible sequences, creating BsmBI cut sites, were added to the both ends of the oligos (Table 9).

Table 9. gRNA oligo sequences		
gRNA oligos		Sequences
gCntrl	top	5'-CACCTGAACCGCATCGAGCTGAA-3'
	bottom	5'-AAACTTCAGCTCGATGCGGTTCA-3'
gSETD2-1	top	5'-CACCGACATCTACTGGTAAGGGTAC-3'
	bottom	5'-AAACGTACCCTTACCAGTAGATGTC-3'
gSETD2-2	top	5'-CACCGATCTAACTGCTACATCTAC-3'
	bottom	5'-AAACGTAGATGTAGCAGTTAGATC-3'
gSETD2-3	top	5'-CACCGAACTGCTACATCTACTGGTA-3'
	bottom	5'-AAACTACCAGTAGATGTAGCAGTTC-3'

Designed oligos were synthesized by Macrogen. gRNA oligo cloning into pLentiCRISPR v1 vector was performed according to the Crispr cloning protocol defined in (https://www.addgene.org/crispr/zhang/). First, complementary gRNA oligos were phosphorylated with T4 PNK (NEB M0201S) and then annealed in a thermocycler at 37ºC for 30 min and ramping down the temperature from 95ºC to 25ºC, with 5ºC/min intervals (Table 10).

Table 10. gRNA oligo pairing mix		
Reagents	Conc.	1X reaction
Oligo 1	100 µM	1 µl
Oligo 2	100 µM	1 µl
10X T4 Ligation Buffer		1 µl
T4 PNK		0.5 µl
dH_2O		6.5 µl

1 µg of pLentiCRISPR v1 (Addgene 49535, not available in Addgene anymore) was digested with BsmBI endonuclease for 30 min. Digested plasmid was purified with MN PCR Clean-up gel extraction kit. 1 µl of (1/200 diluted with dH_2O) annealed oligo insert was ligated into 50 ng of pLentiCRISPR v1 vector, in molar ratios, with Quick Ligase (NEB M2200S). Ligation mixture was incubated at 37ºC for 30 min and transformed into Stabl3 competent bacteria. Same procedure with shRNA cloning was performed for the subsequent steps. Clones were verified by Sanger sequencing using MSCV primer (Figure 4).

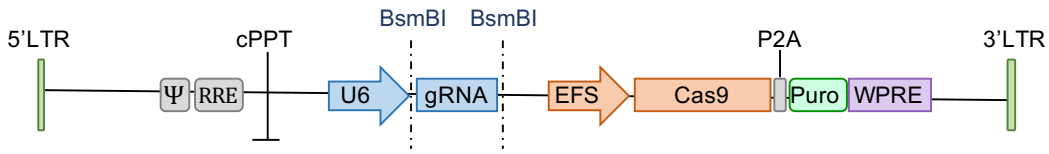


Figure 4. Schematic representation of pLentiGuide v1 plasmid with gRNA insert.

2.1.3.3. Cloning H3.3 mutants into pWZL vector

pWZL-Blast-GFP plasmid (Figure 5) was used as backbone to clone histone H3.3 cDNA sequence. First of all, GFP sequence was removed from the plasmid digesting with EcoRI endonuclease at 37ºC for 3 hours. Linear backbone (5100 bp) was cut from 1.5% agarose gel and purified with NucleoSpin Gel and PCR clean-up kit (MN 740952.50). Linear backbone was self-ligated with T4 DNA ligase (NEB M0202) at 37ºC for 1 hour, then enzyme was inactivated at 65ºC for 15 min and chilled on ice. Ligated backbone was transformed into Stabl3 bacteria. Single colonies were picked and inoculated into 5 ml of LB broth and then incubated for 16 h at 37ºC by shaking at 225 rpm for plasmid isolation with Nucleospin plasmid DNA purification kit (MN 740588.50).

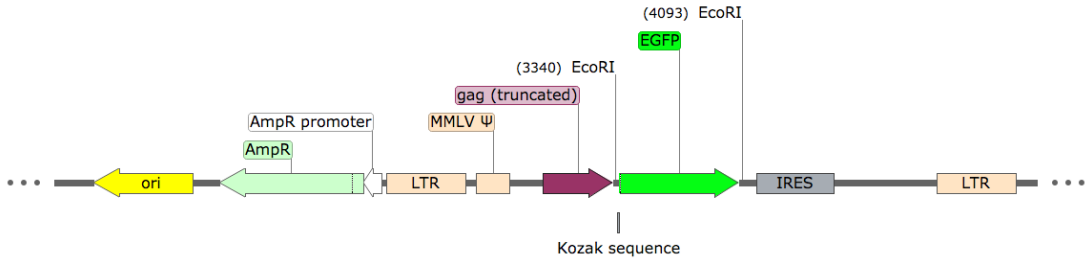


Figure 5. pWZL-Blast-GFP plasmid sequence (Addgene 12269).

Purified backbone was sent for sequencing with IRES-rev primer to confirm the removal of GFP sequence (Figure 6). pWZL-Blast backbone was digested with XhoI endonuclease and treated with Antartcic phospahatase and stored at -20ºC until histone H3.3 cloning.

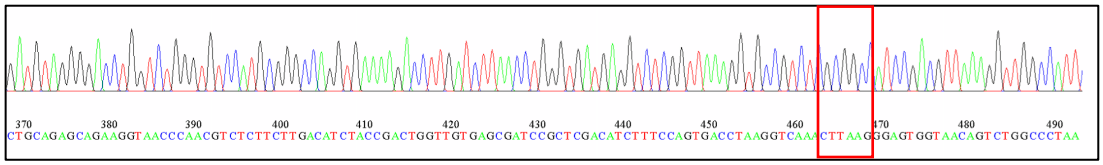


Figure 6**. pWZL-Blast backbone sequence without GFP sequence.**
Red box shows EcoRI recognition sequence (GAATTC) where two EcoRI sites are fused after the removal GFP sequence.

pcDNA-FLAG-H3.3-polyA plasmid (Figure 7) was used to obtain H3.3 insert sequence. Lysine amino acid sequence (AAG) at 36th residue was substituted to methionine (ATG) sequence by Q5 site directed mutagenesis kit (NEB E0554) with suitable primers (Table 11). PCR was performed with described as (Table 12) with defined reaction steps (Table 13).

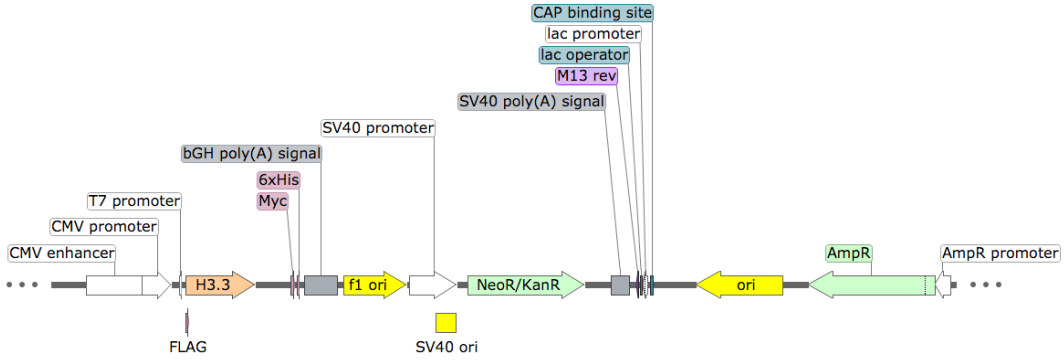


Figure 7. pcDNA-Flag-H3.3-polyA plasmid map (Addgene 59780)

Table 11. Primers used for Q5 site directed mutagenesis	
Primers	Sequences
H3.3K36M fwd	5'-CGGCGGGGTGatgAAGCCTCATC-3'
H3.3K36M rev	5'-GTAGAGGGAGCGCTTTTCC-3'

Table 12. Exponential amplification PCR		
Reagents	Conc.	1X reaction mix
Q5 2x master mix	2X	12.5 µl
primer mix (Table 11)	10 µM	1.25 µl
20 ng/µl plasmid	20 ng/µl	1 µl
dH20		9 µl

Table 13. PCR reaction steps			
	Temp.	Time	Cycle #
Initial Denaturation	98ºC	30 sec	1 cycle
Annealing	98ºC	30 sec	30 cycles
	67ºC	2 min 45 sec	
	72ºC	30 sec	
Final Extension	72ºC	5 min	1 cycle
Hold	4ºC	∞	

After lysine-to-methionine substitution at 36th residue of H3.3 with PCR amplification, linear plasmid was ligated with KDL mix (Table 14) for 5 min at RT and then chilled on ice. 5 µl of ligated plasmid was transformed into 50 µl of Stabl3 bacteria. Single colonies were picked and inoculated into 5 ml of LB broth and then incubated for 16 h at 37ºC by shaking at 225 rpm for plasmid isolation with Nucleospin plasmid DNA purification kit.

Table 14. KDL reaction mix	
Reagents	1X reaction mix
PCR product	1 µl
10x KDL enzyme mix	1 µl
2x KDL KDL reaction buffer	5 µl
dH_2O	Up to 10 µl

Purified backbone was sent for sequencing with primer (5'-GGATGGCACACAGGTTGGTA-3') compatible with H3.3 sequence, to confirm the mutation (Figure 8).

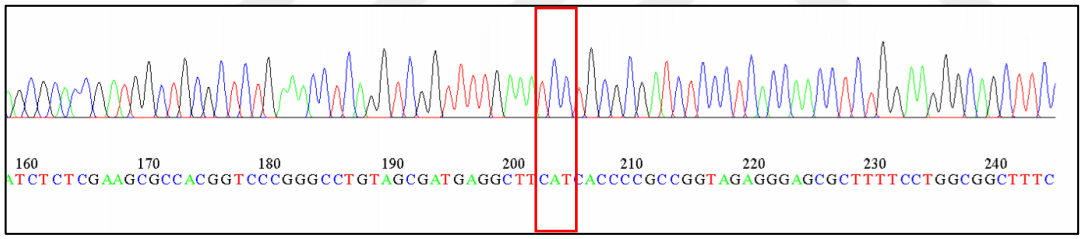


Figure 8. H3.3 sequence with lysine-to-methionine mutation at 36th residue

Since there were no compatible cut sites for cloning H3.3 insert into pWZL-Blast backbone, Additional XhoI cut site was added with at the ends of H3.3-Flag sequence in pcDNA plasmid via PCR amplification (Table 15) with designed primers (Table 16).

Table 15. PCR reaction for XhoI cut site insertion		
Reagents	Conc.	1x reaction mix
pcDNA-Flag-H3.3-PolyA	200 ng	
Phusion HF Reaction buffer	5x	10 µl
dNTP mix	10 µM	1 µl
Primer mix	10 µM	2.5 µl
Phusion High Fidelity DNA polymerase		0.5 µl
dH_2O		Up to 50 µl

Table 16. XhoI cut site inserting primers	
Primers	Sequences
XhoI-start-H3.3 fwd	5'-AGCTA**CTCGAG**ACCATGGACTACAAAGACGATGACG-3'
XhoI-stop-H3.3 rev	5'-TAGCT**CTCGAG**TTAAGCTCTCTCTCCCCGTATC-3'

Amplified PCR product was purified with NucleoSpin PCR clean-up Gel extraction kit. pWZL-Blast plasmid and pcDNA-Flag-H3.3 products were digested with XhoI endonuclease at 37ºC for 3 hours then digested products were obtained with agarose gel purification. Flag-H3.3 wild type and Flag-H3.3K36M inserts were cloned into pWZL-Blast backbone with previously explained ligation protocol and generated pWZL-Flag-H3.3 plasmids (Figure 9) were transformed into Stabl3 bacteria to propagate plasmids for further experiments.

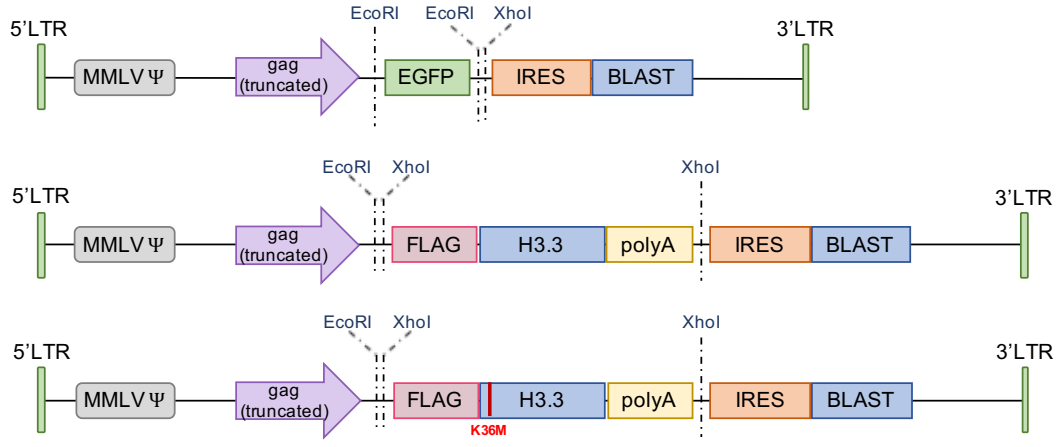


Figure 9. Schematic representation of pWZL-GFP plasmid with FLAG-H3.3-polyA insert.

2.2. Cell culture experiments

2.2.1. Mitomycin inactivated mouse embryo fibroblast (MEF) preparation

MEF feeder layer preparation was optimized from Jozefczuk *et al*. [202]. Balb/C mouse embryos were isolated at E13.5 of embryonic development and washed with 1X PBS. Head and visceral tissues were removed. The remaining tissues were minced using razor blades and incubated with 0.25% trypsin at 37ºC for 10 min. Trypsin was inactivated with FBS-containing D10 medium (DMEM 1X, 10% FBS, 1% Penicillin/Streptomycin) and cells were dissociated by pipetting. Cells were then transferred onto 0.1% gelatin-coated (Sigma G1890-5006) 10 cm plates and incubated with D10 medium at 37ºC under 5% CO_2. Medium was changed daily, and cells were passaged when monolayer cells were reached to 80% confluency.

MEF inactivation was performed with mitomycin C (Millipore 475820). Cells were propagated on gelatin-coated 15 cm plates. MitoC was diluted with 1X PBS to 1 mg/ml. Cells were washed with 1X PBS and then treated with 20ml D10 medium containing 10 µg/ml of mitoC for 2 h at 37ºC. mitoC-treated MEFs were washed twice with 1x PBS and then trypsinized to single cells. Cells were counted with hemocytometer, aliquoted to cryovials in freezing medium (50% FBS, 40% D10 medium, 10% DMSO).

2.2.2. Packaging of retro- and lentiviral plasmids

293T cells were cultured in D10 medium at 37°C with 5% CO_2 Virus packaging for retro- and lentiviral plasmids were performed as explained in Onder *et al*. [203]. $2.5x10^6$ 293T cells were seeded onto 10cm plates for plasmid transfection. Required packaging plasmids were prepared in the following concentrations (Table 17).

Table 17. Viral packaging mix		
	Retroviral	Lentiviral
Plasmid name	2500 ng	2500 ng
pCMV-VSV-G (Addgene 8454)	225 ng	225 ng
pUMVC (Addgene 8449)	2250 ng	-
pCMV-dR8.2 dvpr (Addgene 8455)	-	2250 ng
DMEM 1X	up tp 400 µl	up to 400 µl

15 µl of FuGENE transfection reagent (Promega E2691) was placed into 180 µl of DMEM 1X and plasmid mixture was added onto FuGENE. After 30 min incubation, transfection solution was dropwise applied to 293T cells. Medium was changed with fresh D10 after 12 hours. Medium containing viruses were collected at 48h and 72h after transfection and filtered through 0.45 µm filters to remove 293T cells and prevent cross contamination. Viral supernatants were aliquoted in small volumes and stored at -80ºC.

Concentrated virus preparation was performed using polyethylene glycol, PEG 8000 (Sigma 1546605). 1/5 volume of 50% PEG 8000 (dissolved in 1X PBS) was added onto the filtered viral supernatants and incubated for 2 days at +4ºC. Viruses were then precipitated by centrifugation at 2500 rpm for 20min and medium was removed. Pellet was dissolved in 1X PBS with 1/100 volume of initial volume, then aliquot and stored -80ºC.

2.2.3. Viral infection

Infection procedures of retro- and lentiviruses were same and processed as explained in Onder *et al*. [203]. $1x10^5$ dH1f cells were seeded into 1 well of 6-well plate. Next day, cells were incubated with 750 µl of unconcentrated virus, 250 µl of D10 medium, and 1 µl of P/S (8 mg/ml). Viral infection was performed sequentially for two days. After infection, cells were treated with 1µl/2ml of puromycin (Gold Biotechnology P-600-100) or other required antibiotics such as 7 µg/ml of blasticidin and uninfected cells were eliminated in 3-day selection. Cells were propagated and used for further experiments.

2.2.4. Culture of dH1f and iPSC cells

dH1f cell culture was performed as previously described [204]. $1x10^6$ dH1f cells were thawed in 37°C for 3 min and DMSO-containing freezing medium was removed by

centrifugation at 1200 rpm for 5 min. Cells were dissolved in D10 medium and placed into 10 cm plates and incubated for 2 days at 37°C under 5% CO_2 pressure. Medium was changed every 2 days and when cells were become 80% confluent, detached from the plate by trypsinization. After 3 min incubation at 37°C, trypsin was inactivated with FBS-supplemented medium and cells were counted with hemocytometer. Cells were seeded according to need for the experiment.

IPSC passaging was performed either with Collagenase IV (Gibco 17104-019) or ReLeSR (STEMCELL 05873). Colonies were washed with DMEM F12 base medium for two times and incubated with ColIV (0.05 g/50 ml DMEM F12) for 5 min at 37°C. The ColIV was removed and colonies were detached from the plate surface by scraping in DMEM F12 medium. Collected colonies were centrifuged and split and plated into mitoMEF feeder layer-seeded plates then were cultured with hES medium (385 ml DMEM F12, 100 ml KOSR, 5 ml Penicillin/Streptomycin, 5 ml L-Glutamine, 5 ml NEAA, 500 µl 10µg/ml FGF2, 500 µl 2-mercaptoethanol) medium. For passaging with ReLeSR was performed with a different procedure. Cells were first washed with MEM F12 base medium for two times and incubated with ReLeSR for 1 min at RT. Then ReLeSR was removed and cells were incubated at 37°C for 6 min. hES medium was added to the plate and cells were detached by pipetting up and down. Single-cell suspension was then split into mitoMEF feeder layer-seeded plates.

2.2.5. Reprogramming of fibroblasts into iPSCs

Reprogramming of fibroblast cells was performed according to a previously published protocol Park *et al*. [204]. Briefly, $5x10^4$ fibroblast cells were seeded into 1 well of 12-well plate and were infected with 250 µl of pSIN4-EF2-O2S (Addgene 21162) and 250µl of pSIN4-CMV-K2M (Addgene 21164) unconcentrated viruses, 200 µl of D10 medium and 0.7 µl of P/S (8 mg/ml). 12 hours after incubation, transduction medium was removed and replaced with fresh D10 medium. Medium change with D10 was performed at 1^{st}, 3^{rd}, and 5^{th} days of reprogramming. On day 6, cells were detached from the plate by trypsinization and transferred onto mitoMEF seeded plates. Medium was switched to into hES medium and changed in every two days until 14^{th} day of reprogramming, then changed daily until day 21.

iPSC colonies obtained by fibroblast reprogramming with defined factors were quantified with Tra-1-60 staining. Cells were washed with 1X PBS for 5 times for 2 min in every was and fixed with 4% PFA for 20 min. After fixation cells were washed with 1X PBS. Antibody solution was prepared with 1X PBS supplemented with 3% Triton X-100 and 3% FBS. Fixed cells were treated with 1:250 diluted biotin-conjugated Tra-1-60 antibody (BioLegend 330604) and incubated overnight at +4ºC. Next day antibody solution was removed, and cells were washed with 1X PBS then incubated with 1:500 diluted HRP-conjugated Streptavidin (BioLegend 405210) for 2 hours. After washing steps to remove unbound streptavidin cells were treated for 20 min with DAB solution (Vector SK-4100) following manufacturer's instructions. Plates were scanned and number of Tra-1-60 positive colonies were quantified using ImageJ.

2.2.5.1. T7 endonuclease assay

gCntrl and gSETD2 virus-infected fibroblasts were subjected to T7 endonuclease assay to verify in/del mutations introduced at targeted regions. Cells were infected with gRNA lentiviruses as explained in viral infection procedure. Puromycin-selected cells were then cultured and propagated at least for 5 days. Cells were collected by trypsinization and centrifugation. Genomic DNA isolation was performed with NucleoSpin Tissue kit using manufacturer's instructions. Concentration of isolated genomic DNA was measure with Nanodrop. gRNA-targeted region was amplified with PCR. 50 µl of PCR reaction was prepared including designated reagents and amounts (Table 18):

Table 18. PCR amplification mix	
Reagents	50 µl reaction mix
2X DreamTaq (Thermo Scientific K1081)	25 µl
5X Betaine (Sigma B0300-1VL)	10 µl
Primer mix (2.5 µM) (Table 19)	5 µl
Template DNA	250 ng
dH_2O	up to 50 µl

Table 19. Primer Sequences for PCR amplification of gSETD2 target site	
Primers	Sequences
T7 gSETD2 fwd	5'-CATGCAGCACCATTACCAGC-3'
T7 gSETD2 rev	5'-TCGTGAAGGTTTCCCTAGATCC-3'

5 µl of the PCR product was run on the 1.5% agarose gel to verify the amplification and the rest was subjected to PCR cleanup procedure using Nucleospin Gel and PCR clean-up (MN 740609.50). Concentration of purified PCR product was measured with Nanodrop and 400 ng of DNA was subjected to T7 endonuclease assay.

First of all, 400 ng of PCR product was mixed with 2 µl of 10X NEBuffer2 and dH_2O up to 19 µl, hybridized by heating up to 95ºC and then cooling down to 4ºC with 0.1ºC/sec ramping rate. Reannealed heteroduplex PCR product was digested with 1 µl of T7 endonuclease (NEB M0302S) at 37ºC for 2 hours. Reaction was stopped with 1.5 µl of 0.25M EDTA and digested DNA fragments were run on 1.5% agarose gel.

2.2.5.2. RNA isolation and cDNA synthesis

RNA isolation was performed using MN NucleoSpin RNA isolation kit according to manufacturer's instructions. Cells were lysed with lysis buffer supplemented with β-

mercaptoethanol. Concentration of eluted RNA was measured by Nanodrop. 1000 ng of isolated RNA sample was converted into cDNA. cDNA synthesis mixture was put into a RNase-free PCR tube with calculated amounts (Table 20).

Table 20. cDNA synthesis mix		
Reagents	Cat. no	1X reaction
RNA		1000 ng
2 mM dNTP	Thermo Scientific R0192	2.5 µl
Random Hexamer	Invitrogen N8080127	1 µl
dH_2O		Up to 16.5 µl

cDNA melting was performed with PCR machine at 65°C for 5 min then mixture was immediately chilled on ice. Then RT mix was prepared (Table 21) and mixed with cDNA synthesis mixture.

Table 21. Reverse Transcriptase mix		
Reagents	Cat. no	1X reaction
5X 1st Strand Buffer	Invitrogen Y02321	5 µl
0.1M DTT	Invitrogen Y00147	2 µl
RNasin	Promega N251B	0.5 µl

Total mixture was incubated at RT for 10 min. 1 µl of M-MLV RT reverse transcriptase (Invitrogen 28025-013) was added and mixture was incubated at 37°C for 1 h and the 70°C for 15 min at PCR machine. After incubation, 75 µl dH_2O was added for ¼ dilution.

2.2.5.3. Real Time Quantitative PCR

Relative mRNA levels were quantified with quantitative real time PCR. cDNAs converted from total RNA were used as template. Reaction mix was prepared with the reactions as defined (Table 22):

Table 22. RT-PCR mix		
Reagents	Stock conc.	1X reaction
2X SYBR Green	2X	10 µl
Primer mix (Table 23)	2.5 µM	2 µl
cDNA	Converted from 1 µg RNA	2 µl
dH_2O		6 µl

Beta-Actin was used as internal control to normalize cDNA amounts between test and control samples. Obtained data was analyzed with ΔΔCT method where Ct stands for threshold cycle where fluorescent first reaches a predetermined level [205]. qPCR reaction was run on LightCycler (Roche 480).

Table 23. Primer sequences used in RT-PCR experiments		
Primers	**Direction**	**Sequences**
SETD2	Fwd	5'-GGAAAACACAACAACTGAACGAG-3'
	Rev	5'-TGAGTTTGCTGGTCTGGGTCT-3'
OCT4	Fwd	5'-AGCGAACCAGTATCGAGAAC-3'
	Rev	5'-TTACAGAACCACACTCGGAC-3'
SOX2	Fwd	5'-AGCTACAGCATGATGCAGGA-3'
	Rev	5'-GGTCATGGAGTTGTACTGCA-3'
NANOG	Fwd	5'-TGATTTGTGGGCCTGAAGAAA-3'
	Rev	5'-TGGTGGTAGGAAGAGTAAAG-3'
LIN28A	Fwd	5'-TGCGGGCATCTGTAAGTGG-3'
	Rev	5'-GGAACCCTTCCATGTGCAG-3'
LIN28B	Fwd	5'-TACCAAACCTCTCCCACCAG-3'
	Rev	5'-ACTATGGCCCCTAGCCAGTT-3'
KLF4	Fwd	5'-TCTCAAGGCACACCTGCGAA-3'
	Rev	5'-TAGTGCCTGGTCAGTTCATC-3'
cMYC	Fwd	5'-ACTATGACCTCGACTACGACTC-3'
	Rev	5'-TCGTCGCAGTAGAAATACGGC-3'
TBX2	Fwd	5'-CGTCACTGCCTACCAGAATGA-3'
	Rev	5'-TACAAGCGTAGAGACGGCAG-3'
SMAD3	Fwd	5'-GAAACTCAAGAAGACGGGGC-3'
	Rev	5'-CGGCAGTAGATGACATGAGGG-3'

NID2	Fwd	5'-CGCTTACGAGGAGGTCAAGC-3'
	Rev	5'-AGGCCGTTGGCAGGATAAAG-3'
MYOCD	Fwd	5'-GGAAACCCTTGTCCCAAAAGC-3'
	Rev	5'-AGGATGGAGGCTGTGTTACCT-3'
TGFBR3	Fwd	5'-GTGGATCTCCGCTGGACACA-3'
	Rev	5'-TCCAACAGTGCGTCTCGTC-3'

2.2.5.4. Western Blot experiments

Western blot experiments were optimized from protocol described by Mahmood *et al*. [206]. Fibroblast cells were collected after virus infection and selections and pellet was lysed according to protein sample needed. Histone acid extraction was performed with TEB buffer described in Abcam protocols website (https://www.abcam.com/protocols/histone-extraction-protocol-for-western-blot).

Fibroblast cells were trypsinized and centrifuged at 1200 rpm for 5 min to remove medium. Cell pellet was resuspended with TEB buffer composed of 0.5% Triton X-100 (v/v), 2 mM PMSF, and 0.02% (w/v) NaN_3. Cells were lysed on ice for 10 min by shaking and then centrifuged at 6500 g for 10 min. Nuclei-containing pellet was washed with TEB buffer for a second time with same centrifugation process and then resuspended in 0.2N HCl. Nuclei lysis was performed overnight in acid and then centrifuged again to obtain histone proteins. Supernatant was collected, and acid was neutralized with 1/10 volume of 2M NaOH. Histone proteins were stored at -20ºC as a short-term storage until western blot experiments. Whole protein was obtained with lysis buffer prepared with defined reagents and amounts (Table 24);

Table 24. Whole cell lysis buffer		
	Stock conc.	Final conc.
TRIS	1 M	50 mM
NaCl	5 M	250 mM
EDTA	25 mM	5 mM
NaF	1.2 M	50 mM
NP40	10%	1%
NaN3	20%	0.02%
PMSF	100 mM	1 mM
Protease Inhibitor	10X	1X
dH2O		Required volume

Cell pellets were lysed on ice by shaking for 30 min and centrifuged at 13000g for 10 min to precipitate cell debris. Protein concentration in the supernatant was quantified by BCA kit (Thermo Scientific 23225). Nuclear/cytosolic fractionation method was used to enrich nuclear proteins those are expressed in low amounts. First, cell pellets were resuspended with cytosolic lysis buffer (Table 25) and incubated on ice for 15 min and then centrifuged at 3000g for 3 min. Cytosolic protein-containing supernatant was placed into another tube and pellet was resuspended with nuclear lysis buffer (Table 26). Nuclei sample was sonicated twice at 40 amplitude for 10 sec using QSonica Q700 sonicator and then centrifuged at 15000g for 5 min. Supernatant containing nuclear proteins was placed into another tube and samples were stored at -80ºC for further experiments.

Table 25. Cytosolic lysis buffer		
	Stock conc.	Final conc.
HEPES pH 7.9	10 mM	500 µM
KCl	10 mM	500 mM
EDTA	0.1 M	25 mM
IGEPAL		0.4%
Protease Inhibitor	10X	1X µl
dH2O		Required volume

Table 26. Nuclear lysis buffer		
	Stock conc.	Final conc.
HEPES pH 7.9	20 mM	500 mM
NaCl	400 mM	5 mM
EDTA	1 M	25 mM
Glycerol	10%	1%
Protease Inhibitors	10X	1X
dH2O		Required volume

Concentrations of proteins were calculated with BCA protein assay kit (Pierce 23227) using BSA as standard according to manufacturer's instructions. Color detection was quantified by plate reader at 562 nm. Protein concentration was calculated by using absorbance value of standards. Loading buffer for sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was prepared by mixing. 900 µl of 4X Laemmli sample buffer (Bio-Rad 161-0747) with 100 µl of 2mercaptoethanol (Bio-Rad 161-0710). Protein concentrations were equalized with 1X loading dye and dH_2O, then

incubated at 95ºC for 10 min. 10X TGS (Tris/Glycine/SDS) buffer (BioRad 161-0772) was diluted to 1X and used as running buffer.

Protein samples were loaded onto polyacrylamide gels (Mini-Protean 4568084) and run at 25 mA for 45 min then transferred onto methanol-activated PVDF membrane via semi-dry transfer using Trans-Blot Turbo transfer system (BioRad 1704150). Membrane was blocked with nonfat dry milk (Bio-Rad 1706404) by shaking for 1h at RT. Membrane was incubated with 1:1000 diluted 1º antibody overnight at +4ºC. Membranes were washed several times with TBS-T (6.05 g Tris, 8.76 g NaCl in up to 1L dH_2O with pH 7.6, supplemented with 0.1% Tween20) and then incubated with 1:10000 diluted infrared-conjugated 2º antibody for 1h at RT in dark (Table 27). After several washing steps, infrared signal coming from 1º antibody-bound 2º antibody by Li-cor Odyssey Fc imaging system.

Table 27. Antibodies used in western blot experiments

Antibodies	**Cat. no**	**Source**	**Working con.**
Histone H3 (D1H2) XP	Cell signaling 4499	Rabbit	1:1000
Tri-Methyl-Histone H3 (Lys36) (D5A7)	Cell Signaling 4909	Rabbit	1:1000
Di-Methyl-Histone H3 (Lys36) (C75H12)	Cell Signaling 2901	Rabbit	1:1000
Tri-Methyl-Histone H3 (Lys4) (C42D8)	Cell Signaling 9751	Rabbit	1:1000
Tri-Methyl-Histone H3 (Lys9) (D4W1U)	Cell Signaling 13969	Rabbit	1:1000
Tri-Methyl-Histone H3 (Lys27) (C36B11)	Cell Signaling 9733	Rabbit	1:1000
Tri-Methyl-Histone H3 (Lys79)	Cell Signaling 4260	Rabbit	1:1000
SETD2 polyclonal antibody	ABclonal A3194	Rabbit	1:1000
Alpha-Tubulin	Sigma T9026	Mouse	1:1000
IRDye 800CW anti-Rabbit IgG	Li-Cor 925-32211	Goat	1:10000
IRDye 800CW anti-Mouse IgG	Li-Cor 925-32211	Goat	1:10000

2.2.5.5. Flow Cytometry analysis

Tra-1-60-PE staining was performed for post-OSKM samples on 7th day of reprogramming. Cells were detached from the plate with trypsinization and centrifuged at 5000g for 10 min. Washing steps were performed for 3 times with 1X PBS supplemented with 1% FBS. Cells were resuspended in 100µl of washing buffer including 0.15 µl of PE anti-human Tra-1-60-& (BioLegend 330610) antibody. Cells were incubated for 1h in dark on ice. Antibody was diluted with wash buffer and cells were centrifuged as before. After several washing steps to remove unbound antibody,

cells were resuspended in 400 µl of washing buffer and subjected to flow cytometry analysis using a BD Accuri C6 flow cytometer. Uninfected dH1f cells were used to gate the negative area in FL1/FL2 channel where the former is used to detect FITC and the latter to detect PE. FITC channel was also taken into consideration to eliminate autofluorescence that may interfere with PE signal. At least $1x10^4$ events were recorded.

2.2.5.6. Karyotype analysis of iPSCs

$1x10^6$ mitoMEF cells were seeded into gelatin-coated 10cm plates and cultured with D10 medium until they attached to the plate surface. D10 medium was removed and cells were cultured with 10ml of hES medium. Conditioned medium (MCM) was collected in every 24 hours for 10 days and stored at -80ºC after filtration to eliminate cross-contamination. iPSC colonies were seeded into Matrigel-coated plates and cultured with MCM medium supplemented with 10µg/ml FGF2. iPSC colonies were propagated and then passaged with ReLeSR onto matrigel-coated cell culture flasks with MCM medium and subjected to karyotyping with Giemsa-banding at Koç University Hospital Department of Medical Genetics.

2.2.5.7. Immunofluorescent staining

Autoclaved cover slips were placed into the wells of 6-well plate and stabilized to the well with 200 µl of D10 medium. $2x10^5$ fibroblasts were seeded onto the cover slips and incubated at 37ºC under 5% CO_2 pressure until cells were attached to cover slips. Wash Buffer was prepared with 1X PBS containing 0.2% Triton X-100. Cells were fixed with 4% PFA for 30 min at RT. PFA was removed by washing steps. Cells were permeabilized with Perm buffer (1X PBS containing 0.5% Triton X-100) for 10 min at RT. Cells were washed to remove Perm buffer and blocked for 1 hour at RT with Blocking buffer (1X PBC containing 0.2% Triton X-100, 3% BSA and 5% goat/donkey serum depending on 2º antibody source) (Table 28).

iPSC characterization for the expression of pluripotency-related proteins was performed with immunofluorescent staining. IPSCs were grown on coverslips with mitoMEF. OCT4, NANOG, and SSEA4 protein levels were determined in control iPSC cell line whose characterization procedures has already been completed. pluripotency-related protein expressions of iPSCs obtained from reprogramming of shRNA-infected fibroblasts were compared with expressions of control iPSCs.

1º antibody was diluted 1/100 with blocking buffer and coverslips were closed onto 15 µl of antibody. 1º antibody incubation was performed overnight at +4ºC. Next day, coverslips were placed onto wells of 12-well plate and unbound antibody was removed by washing steps for 3 times. 2º antibody conjugated with fluorophore was diluted 1/500 with blocking buffer and coverslips were closed onto 15 µl of antibody. 2º

antibody incubation was performed at RT for 1 hour in dark. Coverslips were placed onto wells of 12-well plate again and unbound antibody was removed by washing steps for 3 times. 5 µl of mounting medium with DAPI (Vector H-1500) was placed onto glass slide and cover slip was closed on the medium. Edges of the cover slip was covered with transparent nail polish and slides were stored +4ºC in dark until microscopy imaging. Images were obtained with Nikon 90i confocal microscope.

Table 28. Antibodies used in immunofluorescent staining

Antibodies	Cat. no	Source	Working con.
SSEA4	BD560218	Mouse	1:100
NANOG	Abcam ab21624	Rabbit	1:100
OCT4	Abcam ab19857	Rabbit	1:100
SETD2	ABclonal A3194	Rabbit	1:100
SETD2	LSBioscience LS-C332416	Rabbit	1:100
Alexa Fluor 555 anti-Rabbit	Thermo Fischer A27039	Goat	1:100

2.2.5.8. Teratoma formation

IPSCs were grown on mitoMEFs in 10cm plates in hES medium. When colonies reached to 80% confluency cells were detached from the plate with ReLeSR and mixed at 1:1 ratio with Matrigel and kept on ice.

8-week old immunocompromised SCID mice were obtained from the animal facility of Koç University Research Center for Translational Medicine (KUTTAM) according to ethics committee approval of Koç University Local Ethics Committee for Animal Experiments (approval code: 10.12.2013 2013-4) 150 µl of iPSCs in matrigel were injected into the leg of mouse intramuscularly and animals were examined every week. 6-8 weeks later, mice were sacrificed and teratomas were removed, washed with 1X PBS and stored at +4ºC in formalin solution. Histological sectioning and Hematoxylin/eosin staining were performed at Koç University Hospital Department of Pathology. Images was obtained with Zeiss Axio Imager.

2.3. RNA sequencing and data quantification

Reprogramming of fibroblast cells was performed as described previously. Cells were infected with 2 independent shRNAs targeting SETD2 mRNA and a control luciferase targeting shRNA (shCntrl). From each shRNA-infected cell, $5x10^4$ fibroblast cells were

seeded into 3 wells of 12-well plates in two replicates. First replicate was subjected to RNA isolation 3 days after seeding and denoted as "pre-OSKM" samples.

Second replicate was subjected to OSKM-infection with 250 µl of O2S and 250 µl of K2M unconcentrated viruses, and RNAs were isolated at day 6 of reprogramming (designated as "post-OSKM" samples). RNA isolation was performed with Trizol and continued with direct-zol RNA miniprep kit (Zymogen R2050) according to manufacturer's instructions. Concentration of isolated RNA was measured with Nanodrop and 1 µg of RNA sample was stored at -80ºC for RNA sequencing experiment. Rest was subjected to cDNA synthesis and used for verification of RNA sequencing results by RT-PCR.

RNA samples were sent to Oxford university, Botnar Research Centre (UK). Library preparation and high-throughput RNA sequencing was performed by Dr. Martin Philpott in Prof. Udo Oppermann's laboratory. Library preparation was performed with NEBNext Ultra RNA library prep kit for Illumina (NEB E7770) according to manufacturer's protocol. Libraries were sequenced on NexSeq 500 platform (Illumina) with paired-end run of 2x75 reads, as explained in Cribbs *et al*. [207]. RNA sequencing data analysis was performed at Koç university School of Medicine by Fırat Uyulur, Tunç Morova and Kenan Sevinç. FASTQ data was aligned with TopHat2 Alignment tool. Differentially expressed genes between shCntrl and shSETD2 samples were quantified using DESeq2 tool (https://bio.tools/deseq2). To determine statistically significant differences in gene sets among control and SETD2 knock-down samples, Gene Set Enrichment Analysis (GSEA) was performed with GSEA software using Molecular Signature Database (MSigDB) [208].

2.4. Chromatin Immunoprecipitation-coupled qPCR

Chromatin Immunoprecipitation (ChIP) sequencing data comprising normal human dermal fibroblast, H1 hESC, and hiPSC cell lines were obtained from NCBI Encyclopedia of DNA Elements (ENCODE) Consortium, aligned on Human GRCh37/hg19 reference genome and visualized by using Integrative Genomics Viewer (IGV) program. Chromatin data, immunoprecipitated with either H3K36me3 or H3K27me3 antibodies, were used to design primers to detect epigenetic changes on

distinct genomic regions of reprogramming-related genes. Amplification of promoter, exonic and intronic regions were provided with primer-1, primer-2 and primer-3 in order (Table 29). We aimed to investigate the enrichment of H3K27me3 and/or H3K36me3 on these regions in our shRNA infected pre- and post-OSKM cells.

Table 29. Primer pairs designed for ChIP-qPCR analysis			
Primer pairs	Direction	Target region	Primer sequence
GREM1-1	Fwd	promoter	5'-TCATTTGACGTGGCTGATGGA-3'
	Rev		5'-GGTTCCTCAAAGAGGCCGTA-3'
SMAD3-1	Fwd	promoter	5'-TCCATCCCACCTCTCCTGAA-3'
	Rev		5'-GCTGCTCTGGATGGGTACAA-3'
SMAD3-2	Fwd	exon	5'-CAGCTCTGAGACCCAGGAAC-3'
	Rev		5'-ATAGGATCGTGCAGCAGGTG-3'
TGFBR3-1	Fwd	promoter	5'-TCCCCCATGCACACATTCAT-3'
	Rev		5'-CGGCTGGGTAATGCAGTCTC-3'
TGFBR3-3	Fwd	intron	5'-AGCGGTCACAGAATCCAAGG-3'
	Rev		5'-ACTCAGAAAGCCCAGAGAACA-3'
TBX2-1	Fwd	promoter	5'-TGCCGGAAGACAGAAAACGA-3'
	Rev		5'-GAAGCGCACGTCTCCTAAGT-3'
TBX2-3	Fwd	intron	5'-AAAGCACTCCTGGCTACTGC-3'
	Rev		5'-GCTCTGCCCCACTAGAACAA-3'
NANOG-1	Fwd	promoter	5'-CACACACACCCACACGAGAT-3'
	Rev		5'-TTGTCTATCCCTCCTCCCAGG-3'

ChIP-IT High Sensitivity Kit (Active Motif 53040) was used for ChIP experiments. First, uninfected fibroblasts were subjected to immunoprecipitation to determine primer efficiencies, as well as enrichment of aforementioned marks on target regions. $15x10^6$ cells were crosslinked with fixation buffer, supplemented with 37% formaldehyde (Fisher Scientific BP531-25). After several lysis processes, nucleus bearing specimen was sonicated with Bioruptor sonicator (Diagenode B01020001) to shear chromatin into small fragments.

1/6 of sheared chromatin was used for input preparation and the rest was divided into 3 for immunoprecipitation experiments with ChIP-grade H3K27me3 (Cell Signaling 4909) and H3K27me3 (Millipore 07-449) 1º antibodies. Chromatin-wrapped Histone proteins were pulled-down with antibodies recognizing methylation marks with agarose beads, then histones were removed from DNA by decrosslinking. Obtained DNA fragments and input DNA were subjected to quantitative PCR (qPCR) analysis to

identify enrichment of primer-targeted sequences, those were immunoprecipitated with antibodies. Percent enrichment scores were calculated with difference in threshold cycles (C_T) of input and immunoprecipitated DNA, regarding to their initial amounts.

After identification of H3K36me3- and H3K27me3-bearing regions in fibroblast cells, same procedure was performed for shCntrl and shSETD2-infected preOSKM and postOSKM samples to determine the redistribution of those histone methylation marks. Enrichments of repressing H3K27me3 and activating H3K36me3 marks on targeted regions were compared in between shCntrl and shSETD2 samples to correlate the changes gene expression profiles with their histone methylation status.

3. RESULTS

3.1. Inhibition of SETD2 expression using short hairpin RNAs increases reprogramming efficiency

To test the role of SETD2 in reprogramming, we cloned three independent shRNA sequences the retroviral pSMP vector. As a control we used an shRNA vector targeting the firefly luciferase gene (shCntrl). We first checked whether shRNAs designed to target SETD2 was sufficient to decrease its mRNA level when introduced into fibroblast cells. After transduction of the retroviral shRNA vectors, puromycin selection was applied to dH1f fibroblasts to eliminate uninfected cells. Three independent shRNAs reduced *SETD2* mRNA expression by at least 50 % compared to control shRNA. Efficiencies of shRNA viruses to knockdown SETD2 were checked for every virus batch. Knockdown levels from three biological replicates are shown in (Figure 10).

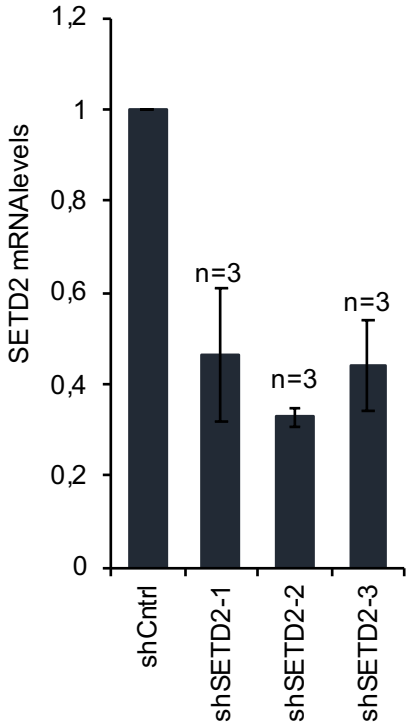


Figure 10. shRNAs are sufficient to reduce SETD2 mRNA levels.
SETD2 mRNA levels of shSETD2-infected cells normalized to shCntrl-infected group. Results are represented with 3 biological replicates each with 2 technical replicates; error bars, ± se.

Next, we sought to examine the effect of SETD2 depletion on histone methyl marks. As expected, depletion of SETD2 significantly decreased the overall cellular Histone H3 Lysine 36 tri-methyl (H3K36me3) levels. Interestingly, we observed a slight increase in the H3K36 di-methyl mark upon SETD2 knockdown, which may be due to the continuous activity H3K36 di-methyl transferases present in the cell. Global levels of other major chromatin marks such as H3K4me3, H3K9me3, H3K27me3 and H3K79me2 were not affected in SETD2-depleted cells compared to control cells (Figure 11). The near total depletion of H3K36 tri-methyl mark upon SETD2 knock-down is in line with previous literature [209] indicating that SETD2 is the only enzyme that catalyzes the conversion of di-methyl H3K36 to its tri-methylated state

Consistent with literature, this data demonstrates that methyltransferase activity of SETD2 is specific for trimethylation of histone H3 lysine 36 residue.

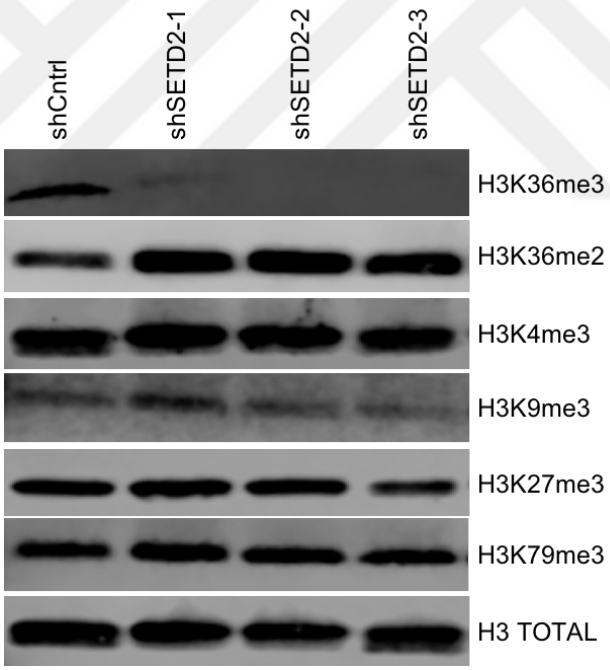


Figure 11. Downregulating SETD2 with shRNAs decreases global levels of H3K36me3 mark. Immunoblot analysis of major histone marks in control (shCntrl) and SETD2 shRNA expressing cells. Total Histone H3 was used as loading control for histone extracts.

SETD2 expressions were examined by immunoblot analysis. First of all, 293T cells were transfected with plasmid containing N-terminus truncated SETD2 in fusion with GFP protein and pBp-GFP plasmid for transfection control. Uninfected dH1f fibroblasts were examined to analyze endogenous SETD2 protein levels. Total proteins were extracted with whole cell lysis protocol. We were not able to detect endogenous SETD2 protein in dH1f cells and pBp-GFP transfected 293T cells in 60 µg total protein extracts. We detected a slight band at around 250 kDa size in SETD2-GFP transfected 293T cell protein extracts (Figure 12.a).

Then we performed nuclear/cytosolic fractionation to enrich SETD2 protein. Similar to whole cell extracts, we were not able to detect SETD2 band in nuclear fractions of uninfected dH1f cells and pBp-GFP transfected 293T cells. We detected a darker and a light band at around 250 kDa size in nuclear /cytocolic fractions of SETD2-GFP transfected 293T cells which we thought that they belonged to SETD2 protein. Tubulin and H3 total blots demonstrated that we were unable to fractionate cells sufficiently since we detected tubulin protein in nuclear fractions and H3 protein in cytosolic fractions (Figure 12.b).

These results demonstrate that, endogenous SETD2 expression is quite low to be detected with immunoblot analysis.

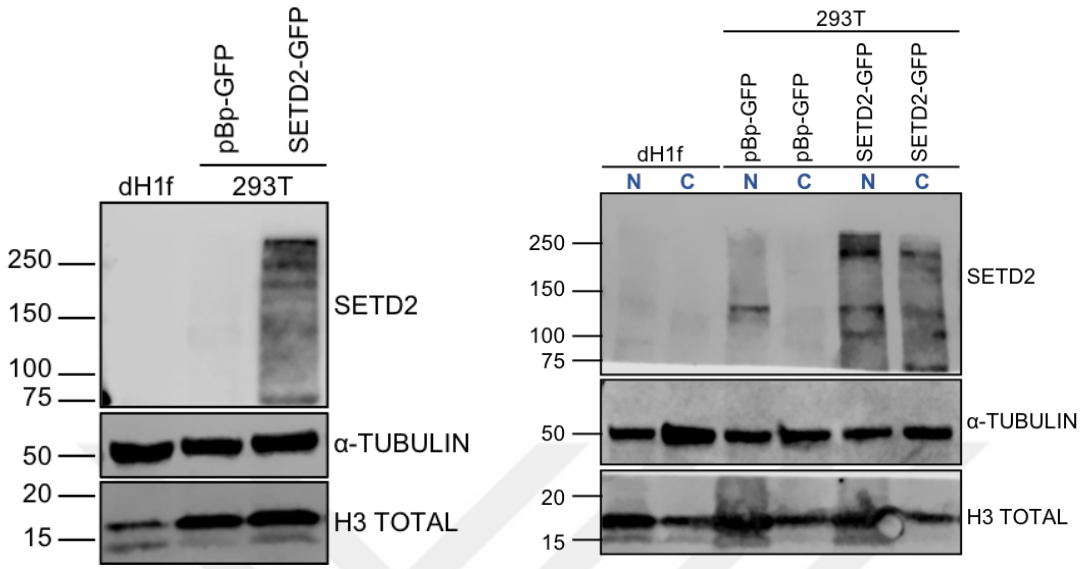


Figure 12**. Immunoblot analysis of SETD2 protein in SETD2-overexpressed fibroblasts.** Immunoblot analysis for SETD2 performed with 1:1000 diluted SETD2 1º antibody (ABClonal A3194). a. Blot for 1º antibody detecting SETD2 protein among whole cell proteins in uninfected dH1f and pBp-GFP or SETD2-GFP infected 293T cells. Alpha-tubulin and H3 total blots were used as loading control of total proteins b. Blot or 1º antibody detecting SETD2 proteins in nuclear/cytosolic fractions of uninfected dH1f and pBp-GFP or SETD2-GFP transfected 293T cells. Alpha-tubulin was used as loading control for cytosolic fractions and H3 total was for nuclear fractions. SETD2-GFP fusion protein weight was expected to be 260 kDa.

We then performed immunostaining on control and shSETD2-infected fibroblasts. Immunofluorescence (IF) signal in control fibroblasts was mostly nuclear with some cytoplasmic staining, which is consistent with recent literature showing that SETD2 has functions in both cellular compartments [131]. As a negative control, cells were stained with only fluorophore-conjugated 2º antibody which did not exhibit any IF signal. In two independent shSETD2 samples, we observed lower signals both in cytoplasm and in the nucleus suggesting that overall SETD2 protein levels were reduced upon knockdown (Figure 13).

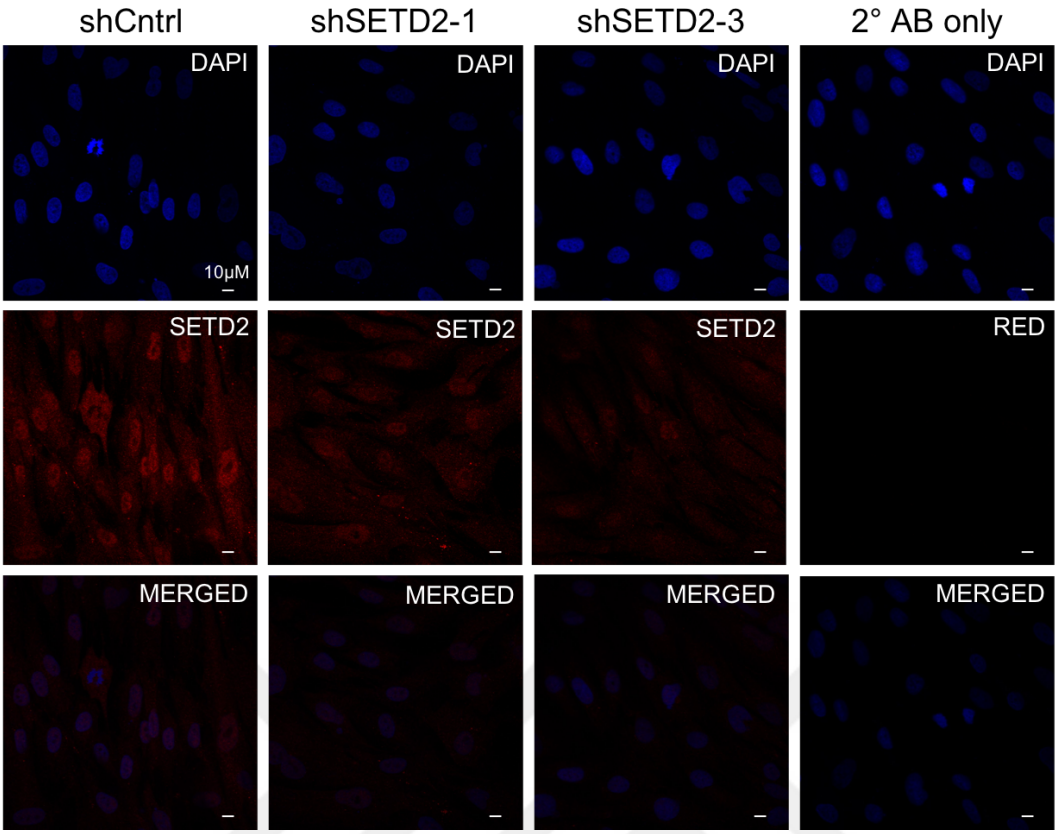


Figure 13. Immunofluorescence staining for SETD2 in control and shSETD2 cells.
Nuclei were stained with DAPI. 2º antibody only stained cells exhibited background signal of conjugated fluorophore at an undetectable level. Immunofluorescence staining was performed with SETD2 rabbit 1ºAb (ABClonal A3194) and Alexa Fluor 555 anti-rabbit 2ºAb. Images were taken with 60X magnification, scale bar 10 μm.

We next performed OSKM-mediated reprogramming of dH1f cells expressing control shRNA and three independent shRNAs against SETD2 (Figure 14).

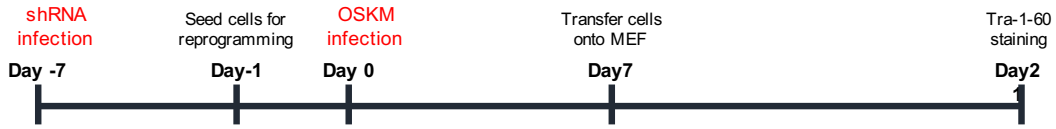
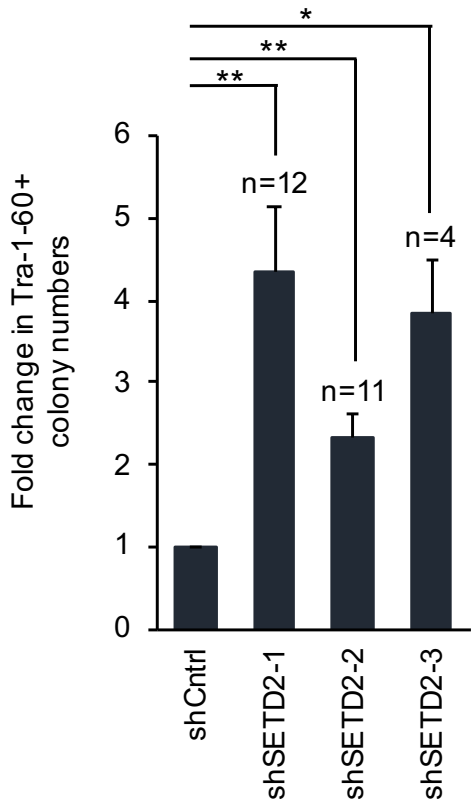


Figure 14. Timeline for reprogramming of shSETD2-infected fibroblasts.
shRNAs were introduced into cells with sequential infections and cells were selected with puromycin (2 µg/ml final conc.)

Number of IPSC colonies were quantified by immunostaining with antibody against Tra-1-60, a cell-surface marker highly specific for embryonic stem cells and fully reprogrammed cells [210]. Reprogramming for each shRNA were performed in at least 4 independent biological experiments. Quantification of Tra-1-60 positive colonies demonstrated that SETD2 knock-down significantly increased iPSC colony numbers resulting in 2-4 fold greater reprogramming efficiency (Figure 15.a). We also checked whether SETD2 downregulation enhances reprogramming of another somatic fibroblast cell line (NO wild type fibroblasts obtained from skin biopsy of a healthy female donor, cell lines were generated in our laboratory) and verified increase in iPSC colony numbers in shSETD2-infected cells compared to shCntrl samples (Figure 15.b).

These data show that, SETD2 is a barrier for reprogramming of fibroblast cells into induced pluripotent stem cells and its depletion with shRNAs enhances reprogramming.

a.



b.

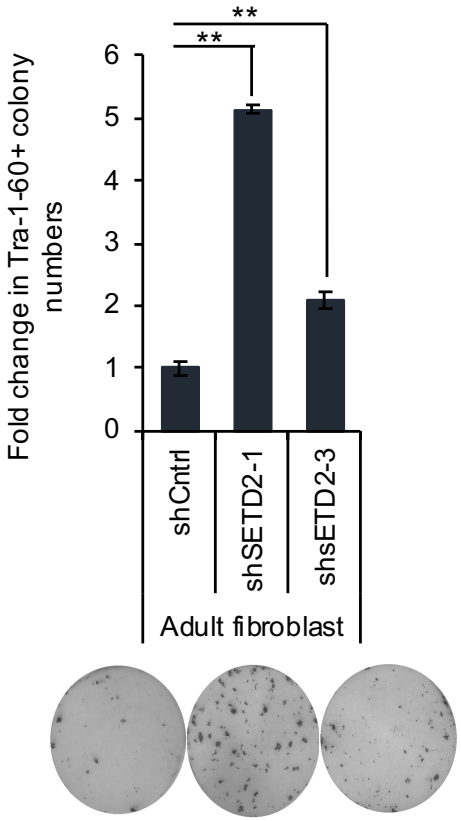


Figure 15. Inhibition of SETD2 enhances reprogramming of fibroblasts into iPSCs.
a. Fold change in the number of Tra-1-60+ colonies. Number of biological independent experiments are indicated with n, each with 3 technical replicates. P values were determined by a two-tailed Student t-test; * $P < 0.05$, ** $P < 0.01$; error bars, ± se. **b.** Fold change in the number of Tra-1-60+ colonies from reprogramming of control (shCntrl) SETD2-depleted NO wt fibroblasts indicated with 3 technical replicates. P values were determined by a two-tailed Student t-test; * $P < 0.05$, ** $P < 0.01$; error bars, ± sd.

3.2. SETD2 knockdown and Dot1L inhibition have additive effects on reprogramming efficiency.

Previous work has shown that inhibition of the H3K79 methyltransferase DOT1L via shRNAs or small molecule inhibitors also increases iPSC generation [94]. As H3K79me2 levels were unaffected by SETD2 knockdown, we hypothesized that SETD2's effect on reprogramming is independent of DOT1L. To test this notion, we combined DOT1L inhibition with SETD2 knockdown in reprogramming. Control and three independent shSETD2-infected fibroblasts were treated with DMSO or a DOT1L inhibitor, (iDOT1L, EPZ004777) for the first 7 days of reprogramming. Experiments were repeated at least for 2 times for each shSETD2 samples. In line with previous studies, we observed approximately 3-fold increase in iPSC colony numbers of control cells with iDOT1L treatment. SETD2 knockdown had an additive effect with Dot1L inhibitor treatment and further boosted the reprogramming efficiency to nearly 5-fold (Figure 16 a. and b). These results suggest that these two chromatin modifiers act as two separate barriers to reprogramming.

According to reprogramming data, DOT1L inhibition have additive effect on reprogramming of fibroblasts into iPSCs when combined with SETD2 depletion.

a. b.

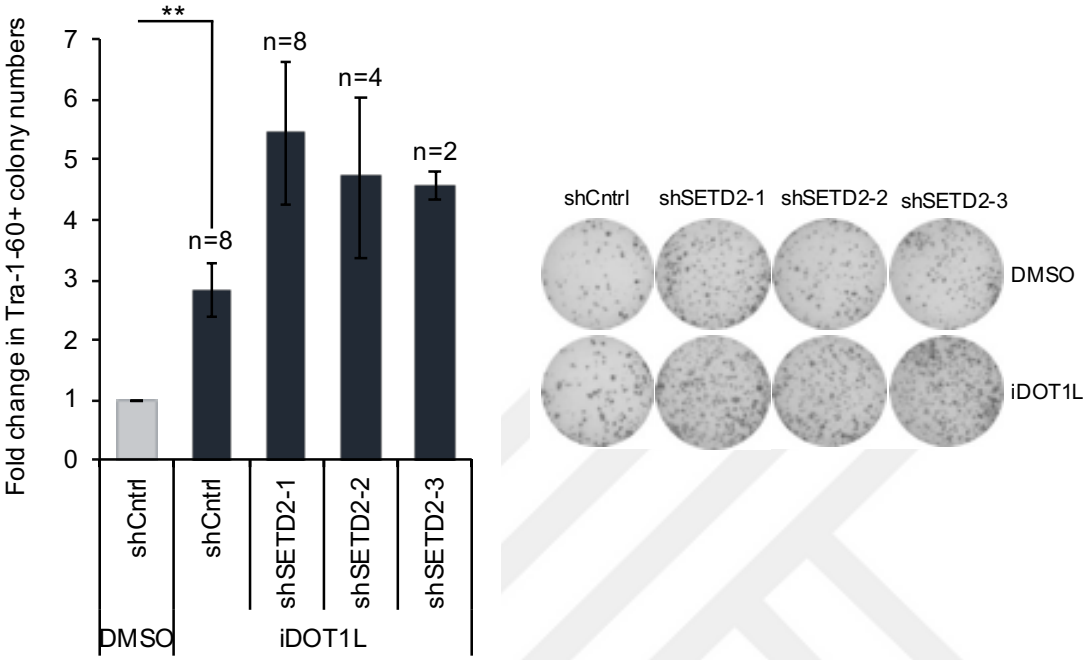


Figure 16. Inhibition of DOT1L has an additive effect on reprogramming when combined with SETD2 knock-down.
a. Fold change in Tra-1-60+ colonies in control and shSETD2 cells treated with iDOT1L compared to control cells treated with DMSO. Number of biological independent experiments are indicated with n, each with 3 technical replicates. *P* values were determined by a two-tailed Student t-test; * $P < 0.05$, ** $P < 0.01$; error bars, ± s.e. **b.** Representative well images from one of the reprogramming experiments quantified in a.

3.3. Inhibition of SETD2 expression enhances iPSC generation via regulating initial stages of reprogramming

Somatic mutations in SETD2 are associated with aggressive progression of cells in certain cancer types [211, 212]. To test whether SETD2 depletion increases cell proliferation, we calculated normalized Tra-1-60+ colony sizes obtained from 3 biological replicates of reprogramming experiments. If downregulation of SETD2 increases cell proliferation, then we would get bigger colonies in knockdown samples than in shCntrl group. However, iPSC colony sizes of SETD2-depleted cells did not show any significant difference compared to control cells (**Figure 17**.a). To directly asses if SETD2 knockdown has an effect on cell proliferation we performed cell counts of control and shSETD2 infected fibroblasts prior to and following OSKM transduction. We counted cells with hemocytometer and normalized the cell count of shCntrl cells into 100. In both settings, we did not detect any significant differences between the total cell numbers of SETD2-depleted and control cells (**Figure 17**. b and c).

These findings clarify that SETD2 depletion does not provide any proliferation advantage for fibroblasts during iPSC generation. The enhancement of reprogramming upon SETD2 downregulation should be related some other functions of SETD2 protein.

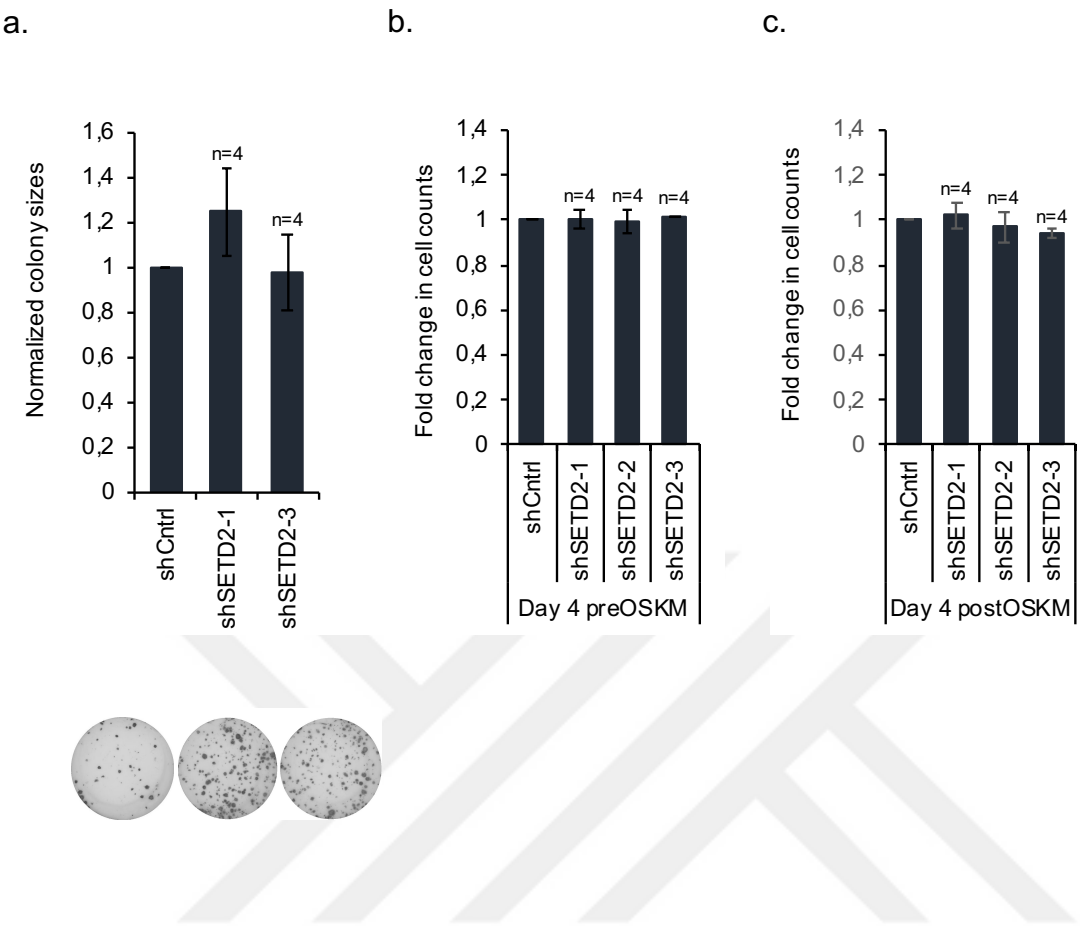


Figure 17. Downregulation of SETD2 with shRNAs has no influence on proliferation rates of reprogramming cells.

a. Normalized colony sizes, those calculated with the proportion of colony numbers to Tra-1-60+ total area of reprogrammed cells. Number of biological independent experiments are indicated with n, each with 3 technical replicates. **b. and c.** Cell count of shSETD2-infected fibroblasts from preOSKM and postOSKM samples compared to shCntrl group whose cell counts are normalized to 100, each represented with 3 technical replicates. *P* values were determined by a two-tailed Student t-test; * $P < 0.05$, ** $P < 0.01$; error bars, ± s.e.

After observing that downregulation of SETD2 does not confer a proliferative advantage to fibroblasts, we next investigated potential cellular mechanisms underlying the increase in reprogramming efficiency of SETD2-depleted fibroblasts. Based on the knowledge that DOT1L inhibition is effective in initial stages of reprogramming [94], we analyzed if SETD2-depletion has similar effects in initial stages of reprogramming. First, we assessed the percentage of Tra-1-60 positive cells emerging at an early time point in reprogramming and performed flow cytometry with a fluorophore conjugated Tra-1-60 antibody 7 days post-OSKM transduction. SETD2 knockdown cells yielded 2-fold greater number of Tra-1-60+ cells at this early time point compared to control cells. These results suggest that SETD2 is an early barrier to reprogramming similar to DOT1L. These results suggest that, *SETD2 is an early barrier to* reprogramming similar to DOT1L (Figure 18).

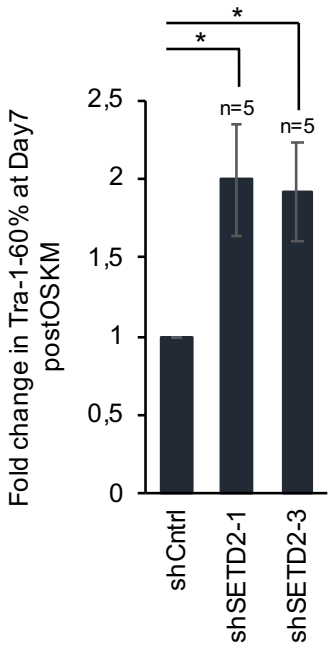


Figure 18. The positive effect of SETD2 downregulation on reprogramming process is apparent at initial stages.
Fold change in Tra-1-60+ cell numbers of control (shCntrl) ad shSETD2 expressing cells, counted with flow cytometry analysis of Tra-1-60-PE stained reprogramming cells, 7 days after OSKM-induction. Number of biological independent experiments are indicated with n, each with 3 technical replicates. *P* values were determined by a two-tailed Student t-test; * $P < 0.05$, ** $P < 0.01$; error bars, ± s.e.

To identify the time-window during reprogramming in which SETD2 depletion is critical we cloned an inducible shRNA against SETD2 using the doxycycline-inducible

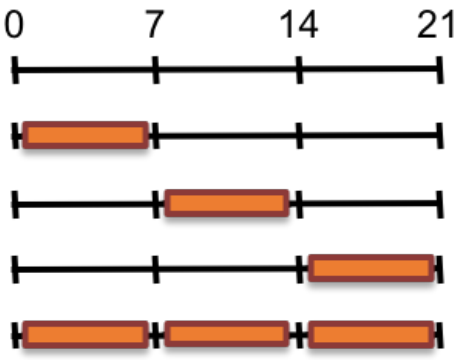
lentiviral backbone LT3GEPIR [213]. This vector allows for the expression of shRNAs in desired time windows during iPSC generation (Figure 19.a). We first checked SETD2 mRNA levels upon doxycycline treatment of fibroblasts and detected SETD2 downregulation in LT3GEPIR-shSETD2-infected cells but not in shCntrl samples (Figure 19.b). Then we supplemented media with doxycycline during reprogramming and downregulated SETD2 in several time intervals. We observed that SETD2 depletion enhanced reprogramming only when it occurred in initial stage (Days 1-7) and this positive effect was diminished when SETD2 was depleted for the entire reprogramming period (Days 1-21) (Figure 19.c).

We did not see any enhancement in reprogramming when we deplete SETD2 at intermediate stage, where H3K36me3 deposition seems to be not present already. Moreover, downregulation of SETD2 at late stage had no effect on reprogramming efficiency, since proper suppression has not been proceeded in initial stage and cells could not meet the requirements to reach the late stage.

The most important finding was when we force shSETD2 throughout reprogramming, the positive influence of that we saw at initial stage was neutralized by shSETD2 expressions at intermediate and late stages. This discovery states that *SETD2 expression was required at late stage most likely to regulate pluripotency gene expressions and somehow, we observe enhancement in iPSC generation even we stably express shSETD2 with under the control of constitutively active MSCV LTR promoter*.

.

a.



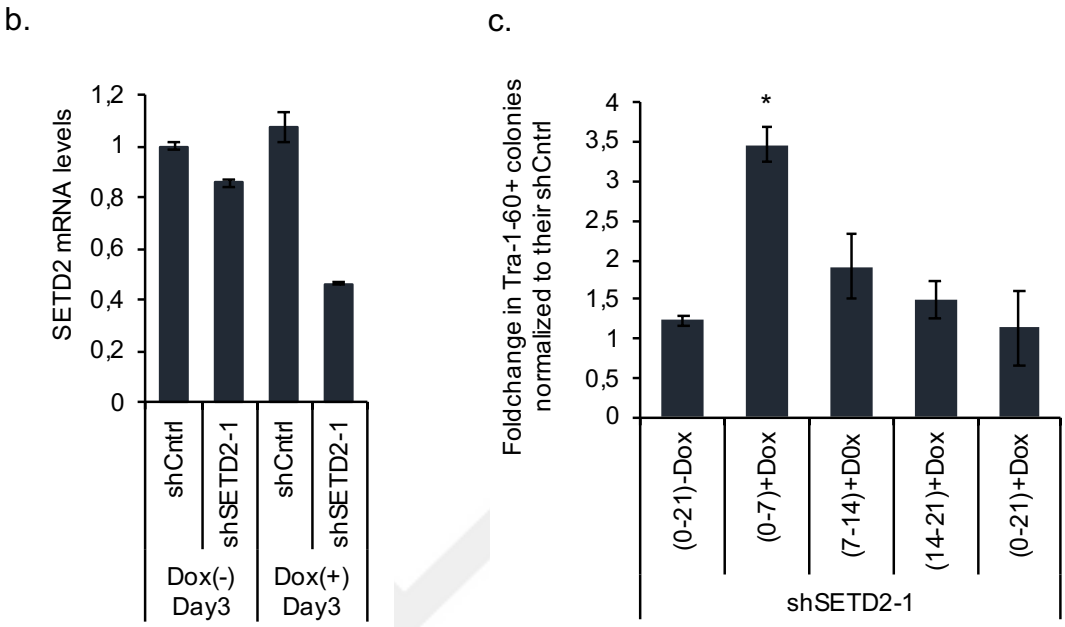


Figure 19. SETD2 depletion in specific time windows during reprogramming.
a. Orange bars designates the time intervals of doxycycline treatment to induce the expression of shRNAs. b. SETD2 mRNA levels in fibroblasts infected with doxycycline-inducible shSETD2-expressing viruses, in the absence or presence of Doxycyline (Dox) normalized to shCntrl untreated samples. Results are represented with 2 technical replicates; error bars, ± sd. c. Fold change in Tra-1-60+ colony numbers upon Dox addition for the indicated time windows. Colony numbers for each time window were normalized to shCntrl cells treated for the same duration with Dox. Data is represented with 3 technical replicates. P values were determined by a two-tailed Student t-test; * $P < 0.05$, ** $P < 0.01$; error bars, ± sd.

The fact that SETD2 downregulation by Dox addition throughout the entire duration of reprogramming had no effect on efficiency, prompted us to investigate whether retroviral shRNA mediated SETD2 downregulation levels fluctuates during reprogramming. It has long been recognized that pluripotent stem cells have the property of repressing viral promoters by inducing heterochromatin at viral integration sites [214]. Therefore, we wanted to check *SETD2* mRNA levels at different time points of the reprogramming process. RT-PCR results demonstrated that, downregulation of

SETD2 was achieved in the initial stage, but the level of knockdown was diminished starting from intermediate stage. This outcome indicated that, shRNAs used to downregulate SETD2 were no longer being expressed after initial stage due to proviral silencing, and that the enhancement in iPSC generation we observed was due to SETD2 knockdown mainly in first 7 days of reprogramming (Figure 20).

This finding confirmed proviral silencing feature of iPSCs which we also observed similar results from previous experiments and determined suppression of proviral particles via demonstrating recovery of SETD2 mRNA levels during reprogramming.

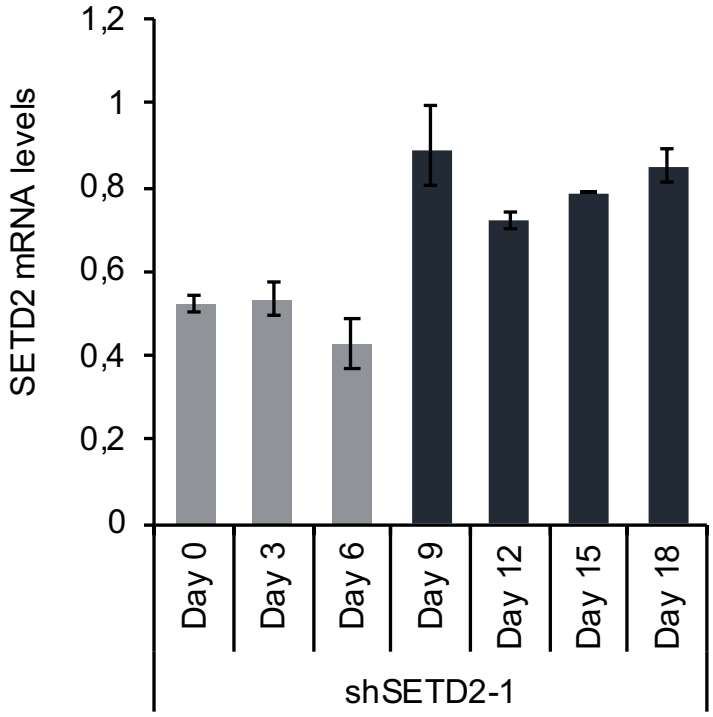


Figure 20. SETD2 expressions in shSETD2 cells returns to its normal levels after initial stage of reprogramming.
RT-PCR analysis of relative SETD2 expressions in shSETD2-infected cells at the indicated days after OSKM transduction. Expression values were normalized shCntrl samples on the same days. Data is represented with 2 technical replicates. *P* values were determined by a two-tailed Student t-test; * P < 0.05, ** P < 0.01; error bars, ± s.d.

3.4. SETD2 knock-out phenotype enhances reprogramming of fibroblasts into iPSCs

After establishing the positive effect of SETD2 downregulation on reprogramming, the we asked whether knockout of SETD2 would also increases iPSC generation. To test this notion, we designed 3 independent gRNAs targeting second exon sequence of SETD2 gene to knockout SETD2 via Crispr-Cas9 technology (Figure 21).

Figure 21. gRNAs and their target sequences in second exon of SETD2 gene.
Illustration of 3 independent gRNAs specific for SETD2 and their target regions in the second exon of the gene. Navy bars indicate gRNAs, complementary to target genome sequences and blue bars show PAM sequences (NGG), which is not present gRNA sequence but required to be adjacent to gRNA target sequence for Cas9 recognition.

We introduced these gRNAs to dH1f cells using a lentiviral vector that also expresses Cas9 (lenticrisprv1 vector) and subjected these cells for T7 endonuclease assay to check for genome editing at targeted regions. We amplified gRNA-targeted regions with a pair of primers and performed heteroduplex formation. Enzymatic digestion of gRNA-targeted region with T7 endonuclease gave clear digested bands in the presence of T7 enzyme which were not digested in negative control groups. Amplified DNA region from gCntrl-infected cells was not digested with T7 enzyme, demonstrating that no mutation was present in gCntrl infected cell populations (Figure 22).

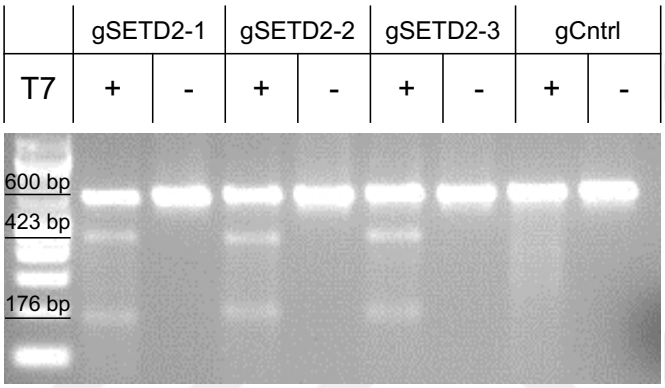


Figure 22. Agarose gel image demonstrating T7 endonuclease-digested gRNA target regions. T7-recognised mismatches on heteroduplex DNA are digested with the enzyme resulting in 2 smaller bands (423 bp and 176 bp) than amplified SETD2 DNA region (600 bp) comprising gRNA target sequence.

Next, we examined the functional loss of SETD2 activity upon gSETD2 infection with western blotting. We extracted histone proteins from control (non-targeting) and SETD2-gRNA introduced fibroblasts and performed western blot with anti-H3K36me3 antibody. We observed partial loss of H3K36me3 marks in gSETD2-infected fibroblasts compared to gCntrl-infected cells. Total H3 levels was used to confirm equal loading (Figure 23).

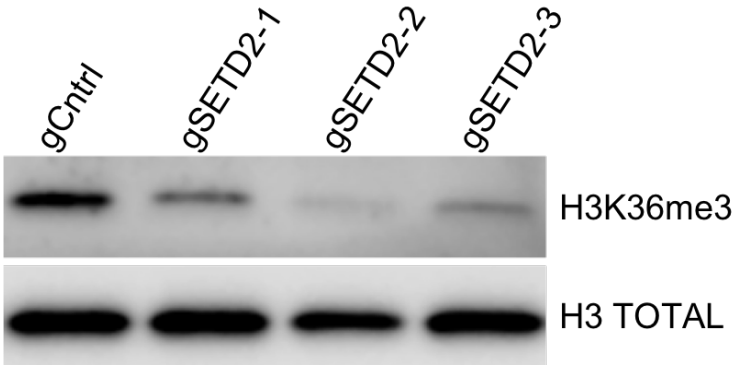


Figure 23. Knock-out phenotype in fibroblasts decreases global levels of H3K36me3 mark. Immunoblot analysis of H3K36me3 mark in control (gCntrl) and SETD2 gRNA expressing cells. Total Histone H3 was used as loading control for histone extracts.

We also performed immunostaining to gSETD2-infected fibroblasts to check SETD2 expressions at protein level. Even though, there was not a necessity to see a decrease in SETD2 levels when mutations were introduced, luckily, we gathered lower signals from all three gSETD2 samples than obtained from gCntrl-infected cells. We also stained cell only with fluorophore-conjugated 2º antibody to eliminate background signals (Figure 24).

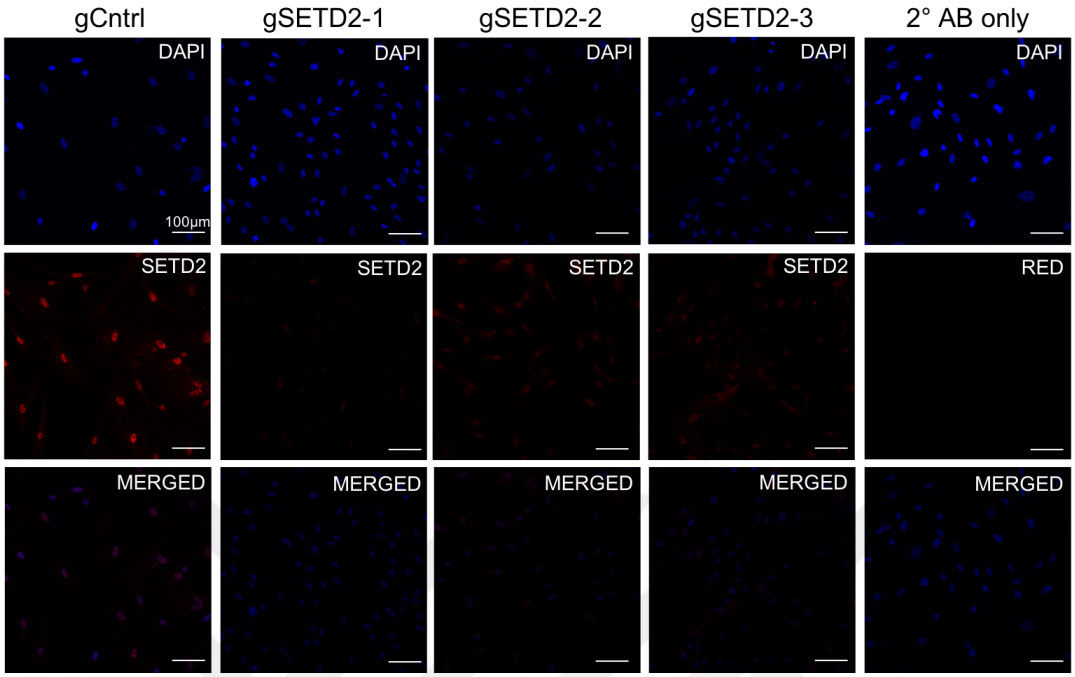


Figure 24. SETD2 proteins exhibit low expression in gSETD2 cells compared to gCntrl group. Immunofluorescence staining for SETD2 protein of gSETD2 and gCntrl infected fibroblasts. Nuclei are stained with DAPI. 2º antibody only stained cells demonstrate background signal of conjugated fluorophore at an undetectable level. Immunofluorescence staining was performed with SETD2 rabbit 1ºAB (LSBioscience 342775) and Alexa Fluor 555 anti-rabbit 2ºAB. Images were taken with 20X magnification, scale bar 100µm.

We performed OSKM-reprogramming of dH1f cells, those were initially subjected to 3 independent gSETD2 infections and puromycin selection. We carried out reprogramming experiments in 3 biological each composed of 3 technical replicates. Colony counts obtained by Tra-1-60 staining of reprogramming experiments demonstrated that, iPSC colony numbers increased in gSETD2 samples compared to gCntrl cells similar to downregulation effect of SETD2 on reprograming (Figure 25).

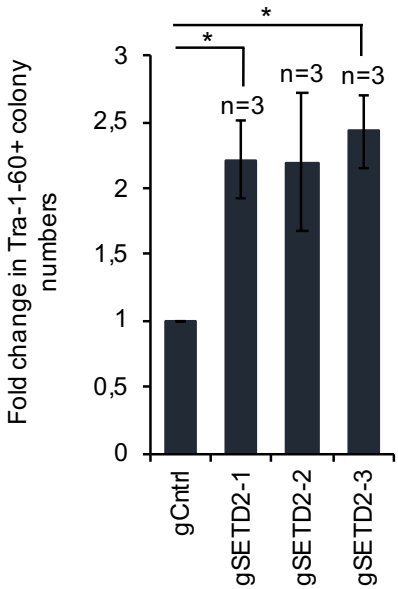


Figure 25. SETD2 knock-out phenotype enhances reprogramming of fibroblasts into iPSCs.
Fold change in Tra-1-60+ colony of numbers three independent reprogramming of gCntrl and gSETD2 fibroblasts. Number of biological independent experiments are indicated with n, each with 3 technical replicates. *P* values were determined by a two-tailed Student t-test; * P < 0.05, ** P < 0.01; error bars, ± s.e.

Repeated reprogramming experiments indicated that, SETD2 knock-out fibroblasts were sufficient to be reprogrammed with traditional OSKM-mediated reprogramming method, and even have increased iPSC colony numbers compared to gCntrl cells.

We also tested additional effect of DOT1L inhibition in combination with SETD2 knock-out. We performed experiments with 3 independent gSETD2-infected fibroblasts and treated these cells with DOT1L inhibitor, iDOT1L in the first 7 days of reprogramming. Experiments were repeated for 2 times for each gSETD2 samples. As expected, iDOT1L treatment in control fibroblasts resulted in a three-fold increase reprogramming (Figure 26.a). However, DOT1L inhibition seemed to have no additive effect on reprogramming when combined with SETD2 knock-out phenotype (Figure 26.b).

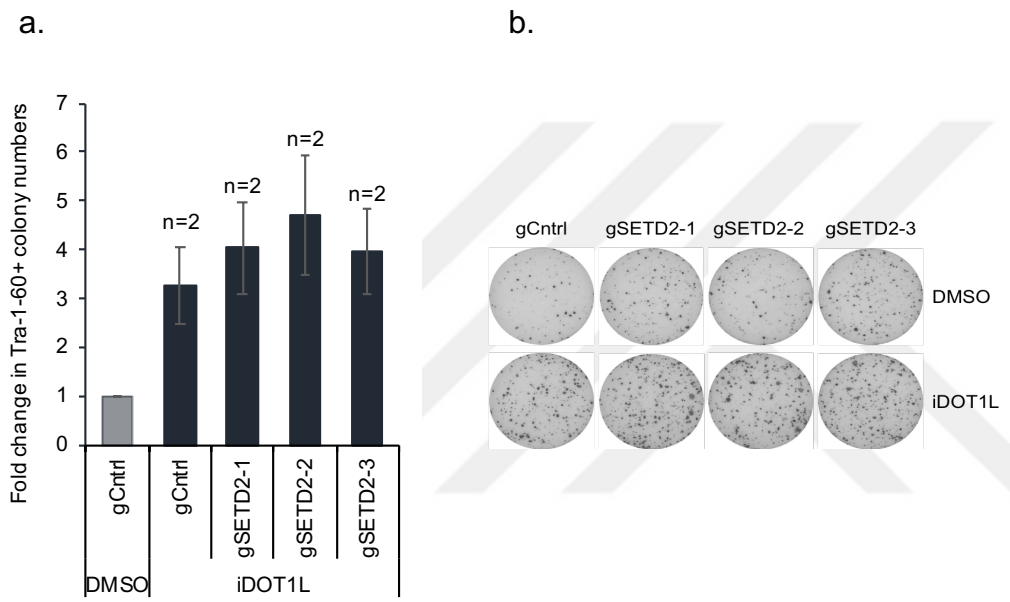


Figure 26. DOT1L inhibition has additive effect on reprogramming when combined with SETD2 knock-out phenotype.
a. Fold change in Tra-1-60+ colonies in control and gSETD2 cells treated with iDOT1L compared to control cells with DMSO. Number of biological independent experiments are indicated with n, each with 3 technical replicates; error bars, ± s.e. **b.** Representative well images from one of the reprogramming experiments quantified in a.

To investigate the proportion iPSC colonies with a complete knockout of SETD2, we picked 33 of reprogrammed colonies from gSETD2-2 and gSETD2-3 infected fibroblasts and examined the total H3K36me3 levels by immunoblotting. This analysis revealed that, 2 colonies (6% of all) had SETD2 double knock-out phenotype with complete loss of H3K36me3 marks and 4 colonies (12% of all) had SETD2 heterozygote knock-out phenotype with partial loss of H3K36me3 marks. Blot showing total H3 levels were used as loading control (Figure 27).

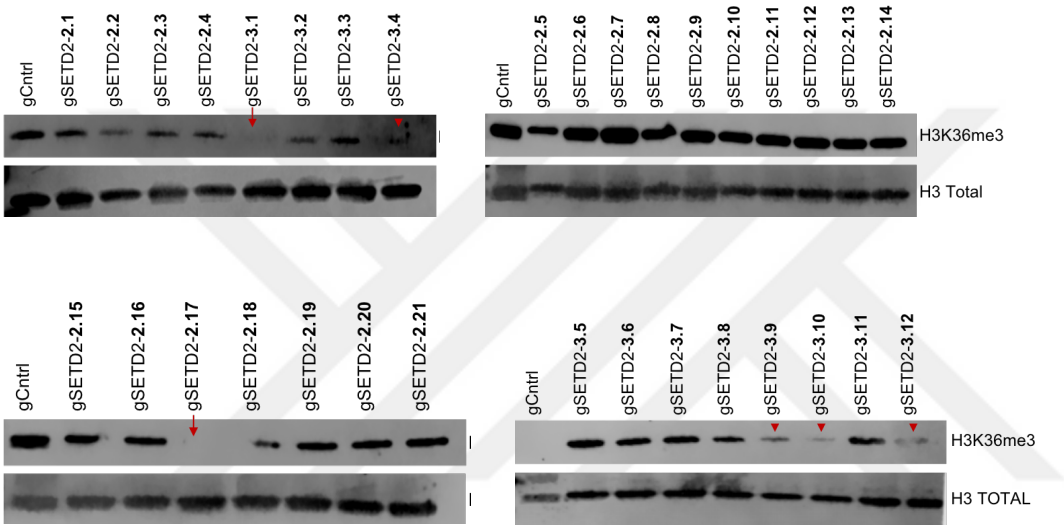


Figure 27. Knock-out phenotype in iPSCs decreases global levels of H3K36me3 mark.
Immunoblot analysis of H3K36me3 mark in control (gCntrl) and iPSC colonies generated via reprogramming of gSETD2-infected fibroblasts. Total Histone H3 was used as loading control for histone extracts. Red arrows showing SETD2 double knockout and arrow heads shows heterozygous knockout iPSC cells for SETD2.

3.5. Decrease in H3K36me3 levels have positive influence on reprogramming

Lysine-to-Methionine mutation in histone H3.3 variant at residue 36 has been shown to reduce global H3K36me3 levels [215]. This dominantly acting mutation is seen in several cancer types including chondroblastoma [216] and sarcoma [140]. In the light of this knowledge, we asked whether decrease in H3K36me3 levels with introduction of H3.3K36M mutated histones can mimic the phenotype of SETD2 downregulation on reprogramming process. We first introduced lysine-to-methionine mutation at K36 residue of histone H3.3 construct via site directed mutagenesis and confirmed substitution with Sanger sequencing (Figure 28). Then we cloned mutated H3.3 sequence together with FLAG tag into a mammalian expressing retroviral plasmid (pWZL-H3.3K36M-Blast).

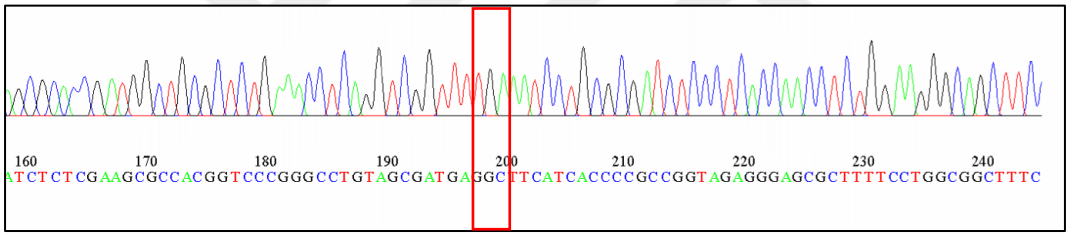
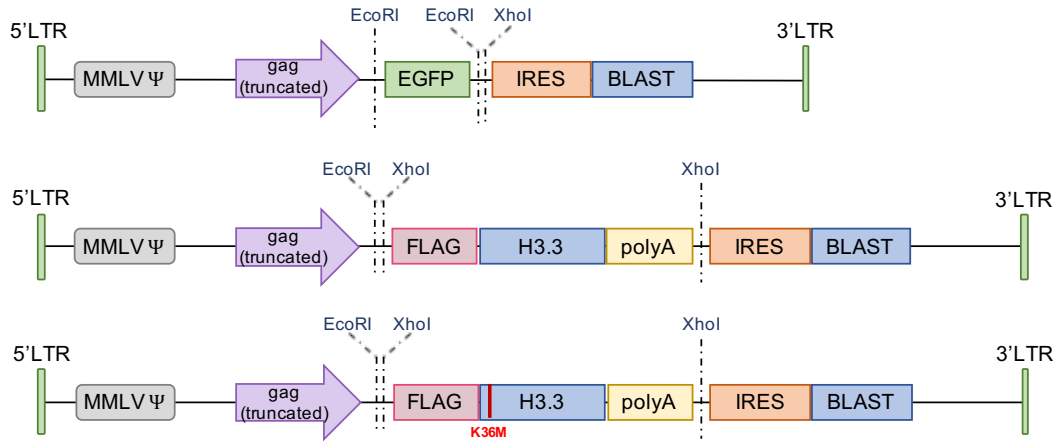


Figure 28. H3.3 sequencing result showing lysine-to-methionine mutation at 36th residue. Lysine-coding AAG sequence was substituted with ATG sequence which codes for methionine. Substitution was performed with site-directed mutagenesis.

We introduced H3.3K36M mutant histone and its control wild type H3.3 to dH1f cells and selected the cells with blasticidin (Figure 29.a). We first checked H3K36me3 levels to confirm global reduction of this mark upon H3.3K36M expression. Western blot results demonstrated that, overexpression of H3.3K36M mutated histones decreased global H3K36me3 and H3K36me2 levels compared to control group. We detected overexpression of the mutant histones with an anti-FLAG antibody, which was not present in GFP control cells (Figure 29.b).

a.



b.

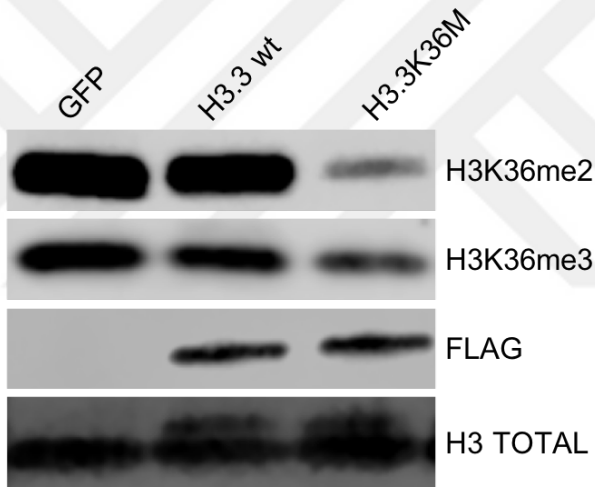


Figure 29. Overexpression of H3.3K36M mutant histones decreases global levels of H3K36me2/3 marks.
a. illustration of pWZL plasmid with GFP, FLAG-H3.3 and FLAG-H3.3K36M inserts. **b.** Immunoblot analysis of H3K36me2/3 marks in pWZL-GFP, pWZL-H3.3 wt and pWZL-H3.3K36M transduced cells. FLAG Immunoblot confirms the overexpression of the H3.3 constructs and total Histone H3 was used as loading control for histone extracts.

After verification of decrease in H3K36me3 levels, we subjected histone H3.3-overexpressed dH1f cells to OSKM-reprogramming. We carried out reprogramming experiments in 3 biological replicates, each composed of 3 technical replicates. Colony counts obtained by Tra-1-60 staining of reprogramming experiments demonstrated that, iPSC colony numbers increased in H3.3K36M-overexpressed cells compared to H3.3 wild type control group, demonstrating the influence of H3K36 methylation on reprogramming (Figure 30).

a. b.

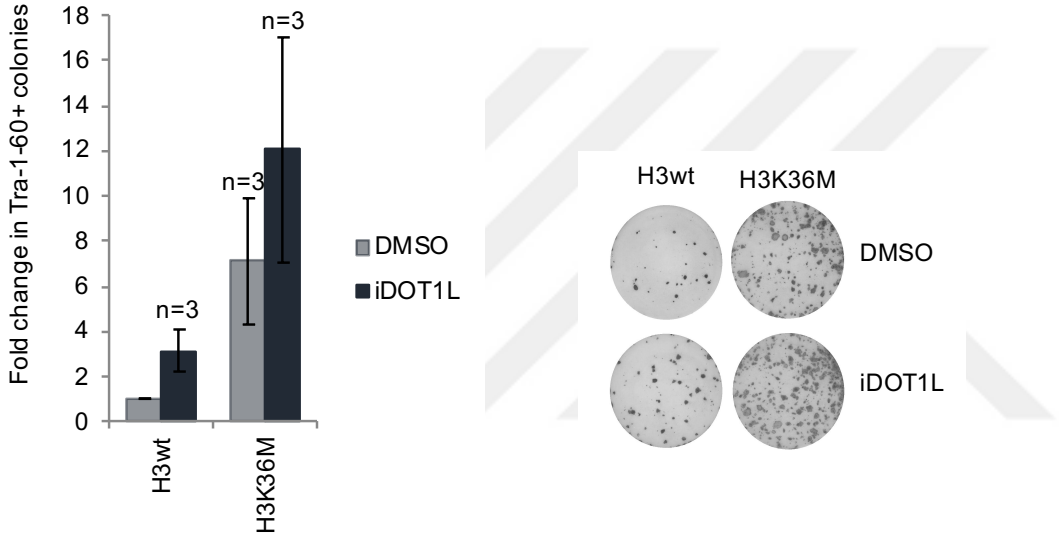


Figure 30. Overexpression of H3.3K36M mutant histones enhances reprogramming of fibroblast into iPSCs.
a. Fold change in Tra-1-60 positive colonies in control H3.3 wt and H3.3K36M cells treated with either DMSO or Dot1L. Number of biological independent experiments are indicated with n, each with 3 technical replicates; error bars, ± s.e. *P* values were determined by a two-tailed Student t-test; * P < 0.05, ** P < 0.01; error bars, ± s.e. **b.** Representative well images from one of the reprogramming experiments quantified in a.

This result demonstrated that, decrease in H3K36me3 levels without intervening other cellular activities of SETD2 protein was sufficient to enhance reprogramming of fibroblast cells into iPSCS. Positive influence of H3K36M mutation on reprogramming process, is a novel work that will shed a light on literature.

3.6. iPSCs derived from SETD2-depleted fibroblasts show characteristic features of pluripotency

Tra-1-60 staining is one of the most commonly used method for iPSC characterization. However, it is not enough to be sure that cells are fully reprogrammed. iPSCs need to express certain pluripotency markers and differentiate into embryonic layers which are characteristic features for embryonic stem cells. Therefore, we picked iPSC colonies reprogrammed from shCntrl and shSETD2-infected fibroblasts and propagated them to generate cell lines. We also used one other control iPSC cell line (NO wt iPSC-2 cell line generated in our lab via reprogramming of NO wt fibroblasts with episomal transfection) for comparison, whose characterization has been performed and required expressions have been displayed (data not shown).

First, we checked whether shSETD2 iPSC cell line expressing SETD2 in normal levels as a consequence of proviral silencing of shRNA sequence. We performed immunofluorescent staining to Cntrl, shCntrl and shSETD2 iPSC cell lines with 1º antibody specific for SETD2 protein. SETD2 protein levels in shSETD2 iPSC cell line were exhibited a comparable level with Cntrl and shCntrl iPSC cell lines. We also stained cell only with fluorophore-conjugated 2º antibody to eliminate background signals (Figure 31).

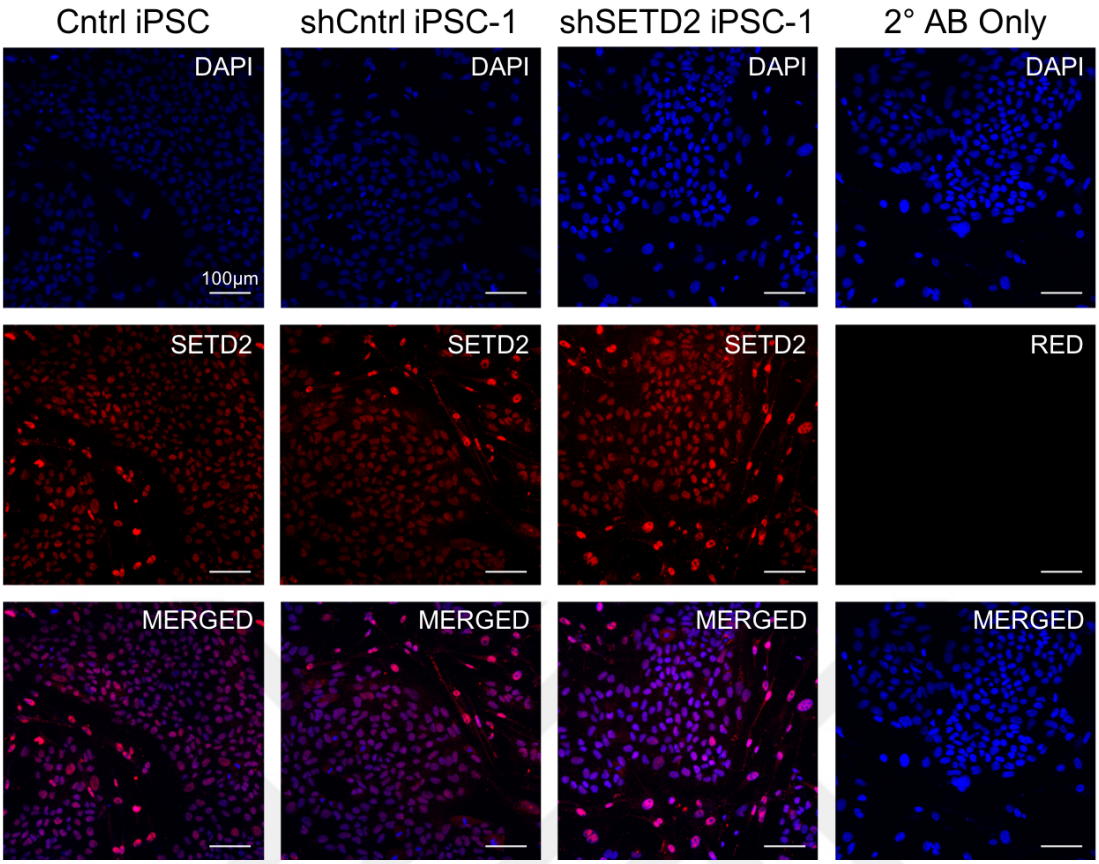


Figure 31. shRNA effect on SETD2 expression is suppressed in shSETD2 iPSCs due to proviral silencing.
Immunofluorescence staining for SETD2 protein of Cntrl, shCntrl and shSETD2 iPSC cell lines. Nuclei are stained with DAPI. 2º antibody only stained cells demonstrate background signal of conjugated fluorophore at an undetectable level. Immunofluorescence staining was performed with SETD2 rabbit 1ºAB (LSBioscience 342775) and Alexa Fluor 555 anti-rabbit 2ºAB. Images were taken with 20X magnification, scale bar 100µm.

Then, we examined mRNA expressions of pluripotency markers together with KLF4 and cMYC as control. We demonstrated that all iPSC cell lines including shSETD2-1 iPSC was expressing OCT4, SOX2, NANOG, LIN28A, and LIN28B which we did not observe expressions of these genes in fibroblast cells. KLF4 and cMYC expressions were detected in all cells and βACTIN was used as PCR control (Figure 32).

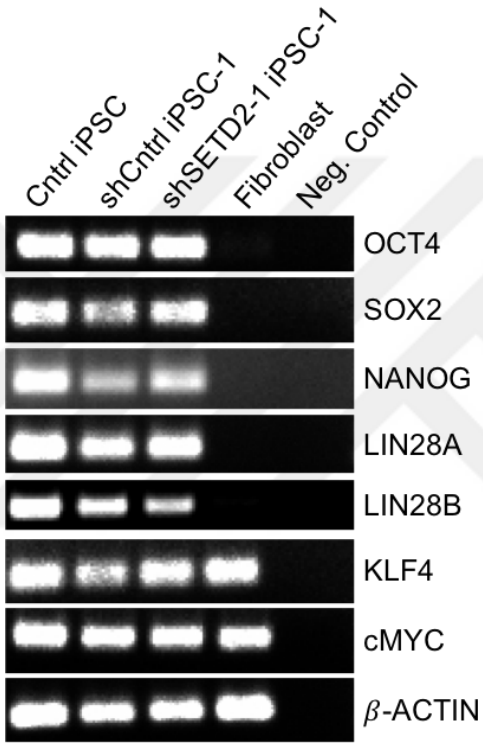


Figure 32. iPSCs derived from shSETD2-infected fibroblast reprogramming expresses pluripotency markers.
Agarose gel image of amplified cDNAs from Cntrl, shCntrl and shSETD2 iPSC cell lines for showing pluripotency markers (OCT4, SOX2, NANOG, LIN28A, and LIN28B). BetaACTIN is used as a control. No template control is Neg. Control.

After demonstrating that shSETD2 iPSCs express pluripotency-related genes at mRNA level, we tested whether these mRNAs are translated to proteins in similar levels. To test this notion, we stained iPSC colonies with 1º antibodies specific for OCT4 and NANOG transcription factors, and SSEA4 cell surface marker. As expected, we observed OCT4 and NANOG localization in the nucleus and SSEA4 in cell surface of all iPSC cell lines (Figure 33).

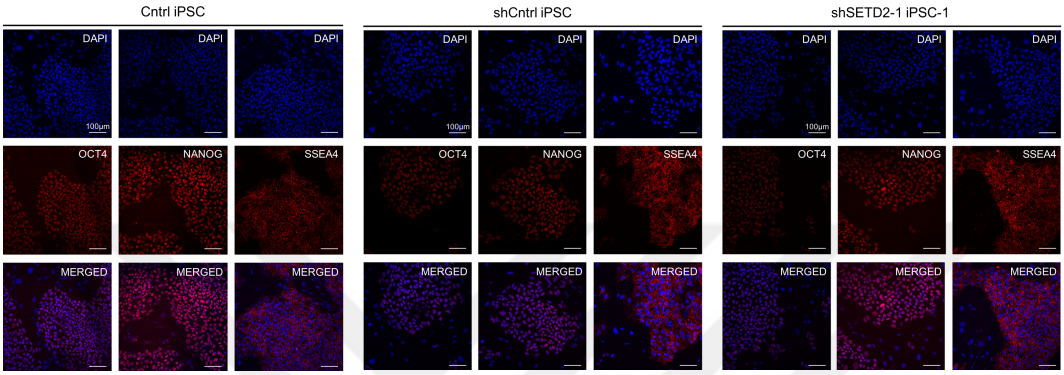


Figure 33. iPSCs derived from shSETD2-infected fibroblasts express pluripotency-related proteins.
Immunofluorescence staining for OCT4, NANOG and SSEA4 proteins demonstrates the presence of their expressions at the protein level in Cntrl, shCntrl and shSETD2 iPSC cell lines. Nuclei are stained with DAPI. DAPI stained and marker unstained cells are mitoMEFs used as feeder layer for iPSCs and act as negative control for pluripotency marker expression. Images were taken at 20X magnification, scale bar 100µm.

To test whether iPSCs generated from SETD2 depleted fibroblasts can attain pluripotency, we performed teratoma formation assay with shCntrl and shSETD2 iPSC cell lines by injecting them intramuscularly into immunocompromised SCID mice. After 6-8 weeks of incubation in mice, all iPSC lines generated cystic teratomas. Tissue sections were prepared from paraffin-embedded tumors and immunohistochemical staining was performed with hematoxylin/eosin. By histological examination of the tissue sections we were able to identify cell structures belonging to all three germ layers, such as ectoderm (neural rosettes), mesoderm (cartilage), and endoderm (glandular epithelium) (Figure 34). This data indicates that, iPSCs generated from reprogramming of shSETD2-infected fibroblasts are able to differentiate into different types of cells representing one of three-germ layers which is similar to control cells.

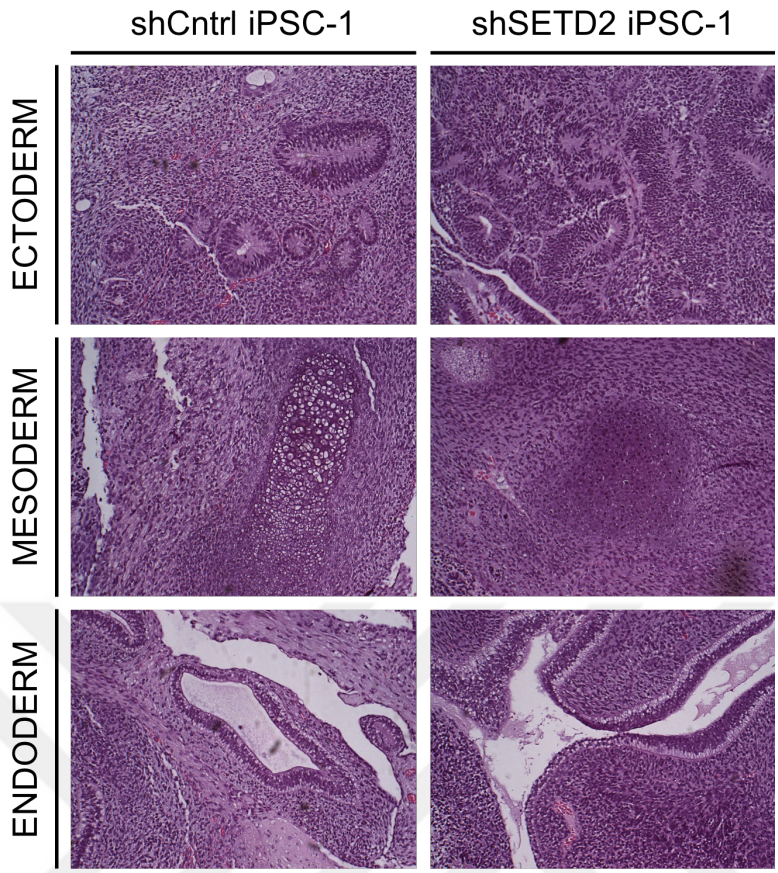


Figure 34. iPSCs derived from shSETD2-infected fibroblasts can differentiate into cells representing all 3-germ layers.
Injection of shCntrl and shSETD2 iPSC cell lines into SCID mice generated teratomas. Hematoxylin/eosin staining of teratomas indicates the presence of differentiated cells comprising characteristics of endoderm, mesoderm, and ectoderm.

These characterization experiments identified the pluripotency and stemness capacity of iPSCs generated from reprogramming of shSETD2-infected fibroblast cells.

3.7. SETD2 inhibition have global effect on gene expressions during reprogramming

Inducible shRNA experiments and increase in the emergence of Tra-1-60 positive cells on day 7 of reprogramming suggested that, downregulation of SETD2 provides advantage in an early time point of reprogramming process. To understand the molecular events occurring upon SETD2 knock-down, we collected RNA samples from control and two independent shSETD2-infected fibroblasts before and 6 days after OSKM-transduction. The latter group coincided with the last day of initial stage. We subjected rest of the cells to verify the increase in iPSC generation in shSETD2-infected cells compared to shCntrl ones (Figure 35).

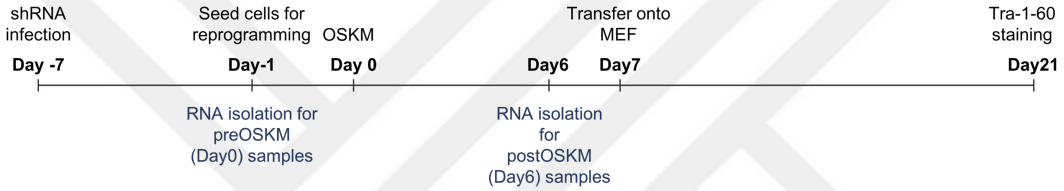


Figure 35. Timeline showing RNA sequencing sample collection points during reprogramming of shRNA infected fibroblasts.
RNA samples were collected from shCtnrl and two independent shSETD2-infected fibroblasts as preOSKM specimens and 6 days after OSKM induction as postOSKM specimens (n=3 for technical replicates).

Library preparations and RNA sequencing was performed in Prof. Udo Oppermann's group at Oxford University, Botnar research Centre and raw data was analyzed in collaboration with Firat Uyulur, Tunc Morova and Kenan Sevinc from Koç University School of Medicine. We first generated cluster dendrogram to ascertain the correlation between technical replicates of same conditions. We identified that, all replicates of same the conditions were clustered together except one set of shSETD2-1 and shSETD2-3 preOSKM samples (Figure 36). Those samples were sequenced in a separate sequencing run, pointing to a batch effect. Therefore, we excluded those two samples and proceeded the following analysis with rest of the replicates.

a.

b.

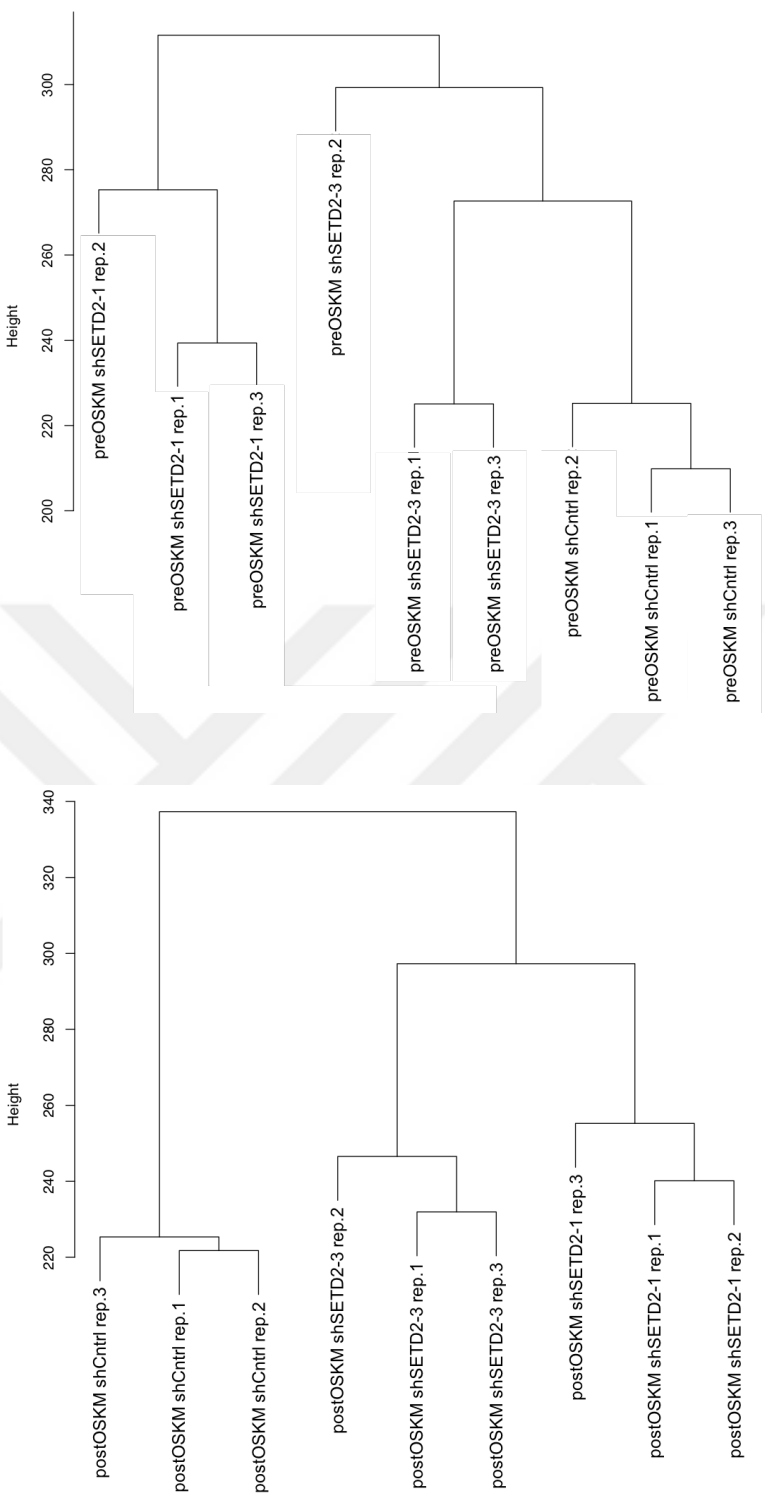


Figure 36. Cluster analysis of RNA sequencing samples.
Samples were clustered according to similarities in their gene expression profiles. a. Clustering of preOSKM samples revealed the divergence of two shSETD2-infected replicas, one belongs to shSETD2-1 and the other to shSETD2-3 samples. b. All triple replicates were clustered together in postOSKM samples.

Differentially-expressed genes among samples was identified using DESeq2 tool. We determined the common downregulated and upregulated genes in two independent shSETD2 pre- and post-OSKM samples with the following cutoffs: p-value ≤0.001 and log2 fold change ≥0.5 (Figure 37). According to this analysis, SETD2 knockdown in fibroblasts resulted in 55 downregulated and 54 upregulated genes. Upon OSKM expression, the number of differentially expressed genes increased, with 274 being downregulated and 219 being upregulated.

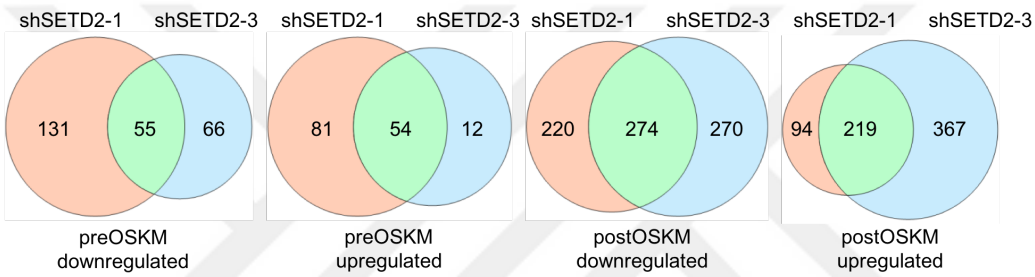


Figure 37. Gene numbers commonly up- or down-regulated among shSETD2 cells of preOSKM and postOSKM samples.
Intersections of venn diagrams shows the number of genes downregulated or upregulated in both shSETD2-infected cell populations compared to shCntrl group (p-value ≤0.001 and log2 fold change ≥0.5).

We also generated heatmaps of gene expression profiles to show similarities in between technical replicates of same samples. Heatmaps were constructed for genes up- or downregulated in both sets of shRNA infected cells compared to shCntrl cells. Differentially-expressed genes were again restrained at the p-value ≤0.001 and log2 fold change ≥0.5 (Figure 38).

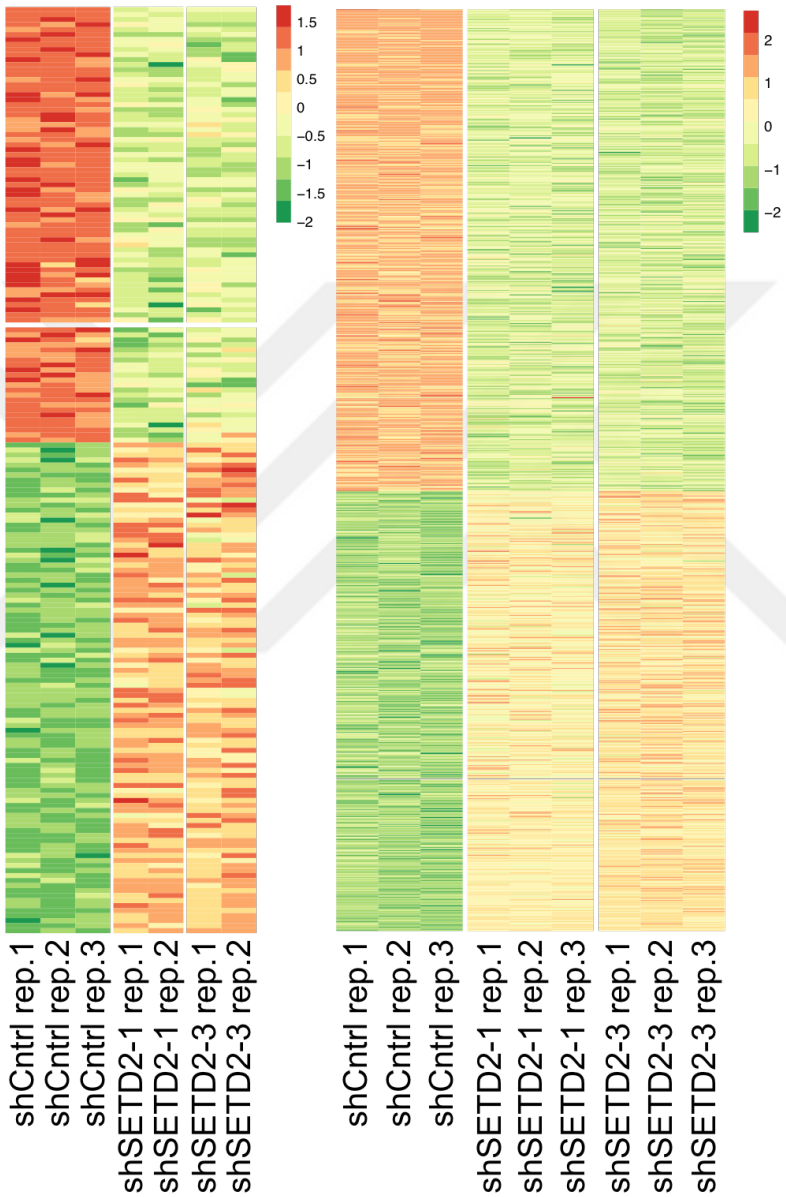
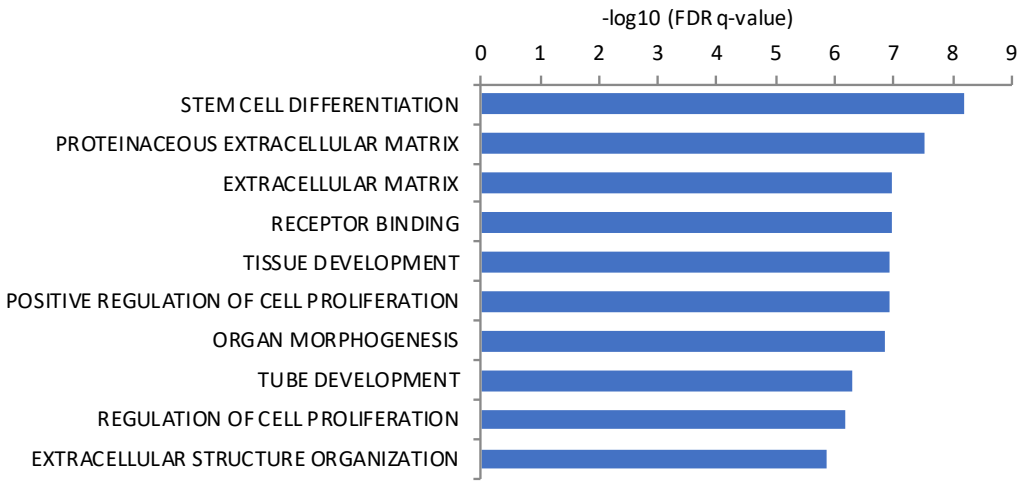


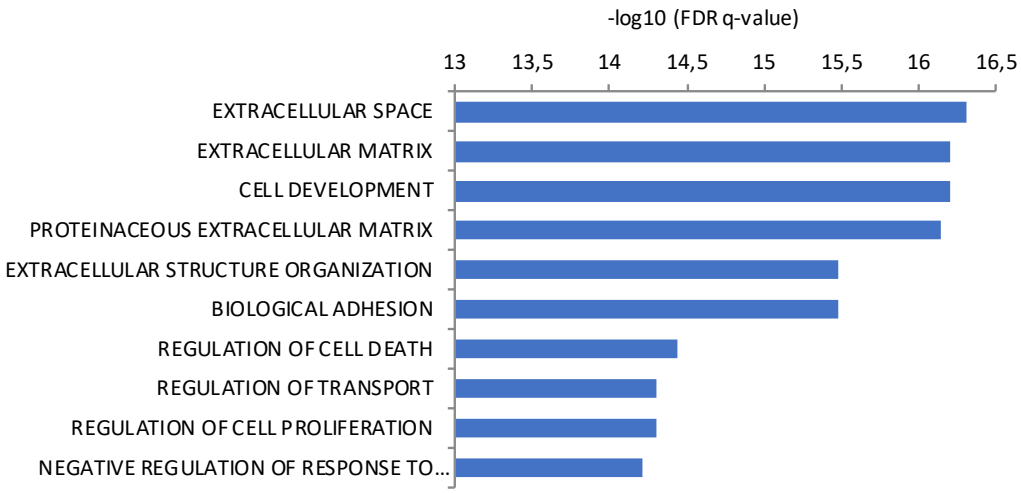
Figure 38. Heatmap showing differentially expressed genes in shCntrl and shSETD2 cells of preOSKM and postOSKM samples.
Color code represents expression levels relative to normalized expression levels (p-value ≤0.001 and log2 fold change ≥0.5). Green bars correspond to downregulated and orange bars to upregulated genes.

We performed pathway analysis for genes those differentially expressed in shSETD2 cells compared to their control cells in preOSKM and postOSKM samples. Using Gene Ontology (GO) analysis, we analyzed genes for their functions in cellular processes, activity regions and biological roles to understand their influence on reprogramming. GO analysis for downregulated genes in preOSKM samples demonstrated that, SETD2 depletion have negative influence on stem cell differentiation and several developmental processes (Figure 39.a). Similarly, downregulated genes in postOSKM samples were identified as regulating cellular development (Figure 39.b). Genes those upregulated upon SETD2 depletion in preOSKM and postOSKM samples were identified as they have role in external or internal stimuli and defense mechanisms against infections (Figure 39. c and d).

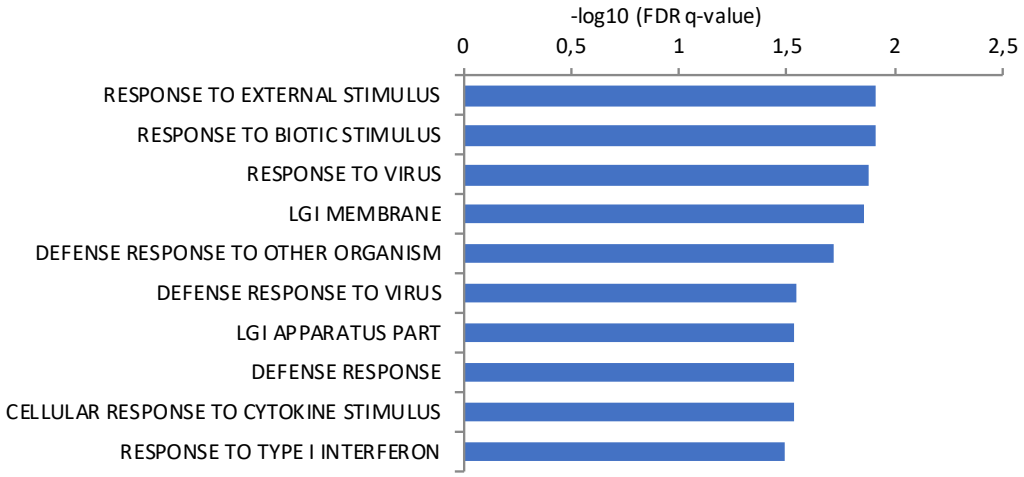
a.



b.



c.



d.

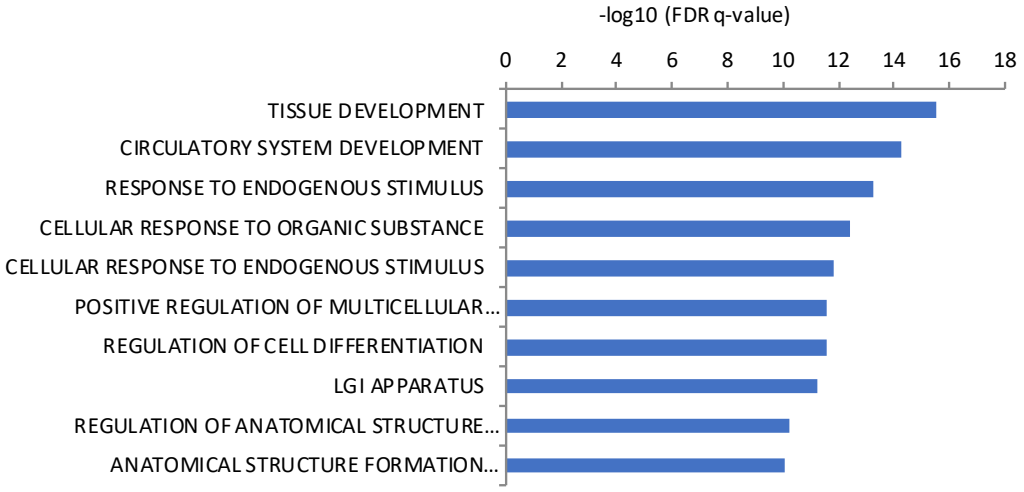


Figure 39. Gene ontology analysis for downregulated and upregulated genes in pre- and postOSKM samples.
Pathway analysis for downregulated genes in **a.** preOSKM **b.** postOSKM samples and upregulated genes in **c.** preOSKM and **d.** postOSKM samples (p-value ≤0.001 and log2 fold change ≥0.5).

Based on the fact that many differentiation and fibroblast function-related GO categories were enriched among the common downregulated genes, we hypothesized that SETD2 inhibition facilitates reprogramming by decreasing the transcriptional activity of fibroblast/mesenchyme specific genes. To test this hypothesis, a fibroblast-related gene set (307 genes) and a pluripotency gene set (308 genes) was constructed based on previously published gene expression profiling of human fibroblasts and their iPSC derivatives (GSE55679) (Figure 40).

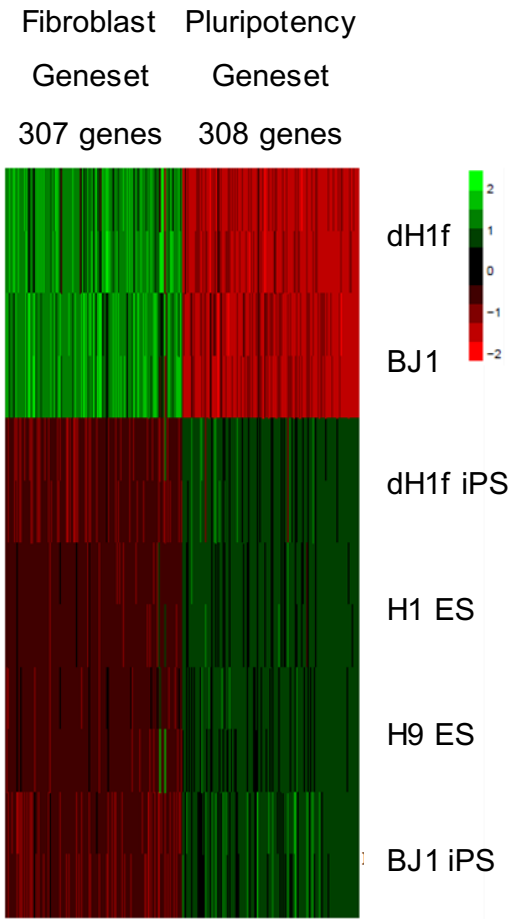


Figure 40. Heatmap showing differentially expressed genes fibroblast and pluripotent stem cells lines.
Heatmap was generated with microarray data of 2 fibroblast and 4 pluripotent stem cell lines with human origin. Color code represents expression levels relative to normalized expression levels. Red bars correspond to downregulated and green bars to upregulated genes.

DESeq2 output of RNA sequencing data was analyzed for gene set enrichment analysis (GSEA). RNA sequencing data was first ranked according to log2 fold change and then compared with fibroblast and pluripotency gene sets. Fibroblast-related gene set was highly negatively enriched in both shSETD2 samples of preOSKM condition (Normalized enrichment score (NES) = -1.61, FDR q-val = 0 for shSETD2-1 and, NES = -1.98 FDR q-val = 0 for shSETD2-3), while pluripotency gene set was not enriched in neither. These results suggest that SETD2 knockdown in fibroblasts results in an overall downregulation of fibroblast-related gene sets without affecting pluripotency genes.

When we examined shSETD2 cells of postOSKM samples it was clear that, fibroblast gene set was further downregulated in both (NES = -2.53 FDR q-val = 0 for shSETD2-1 and, NES = -2.54 FDR q-val = 0 for shSETD2-3), suggesting that depletion of SETD2 results in a more extensive repression of the starting fibroblast gene expression program. In contrast, at this time-point pluripotency gene set was positively enriched in both SETD2-depleted samples (NES = 2.99 FDR q-val = 0 for shSETD2-1 and, NES = 2.96 FDR q-val = 0 for shSETD2-3). This result is consistent with our earlier observations that upon OSKM expression, reprogramming is accelerated in SETD2 knockdown cells (Figure 41).

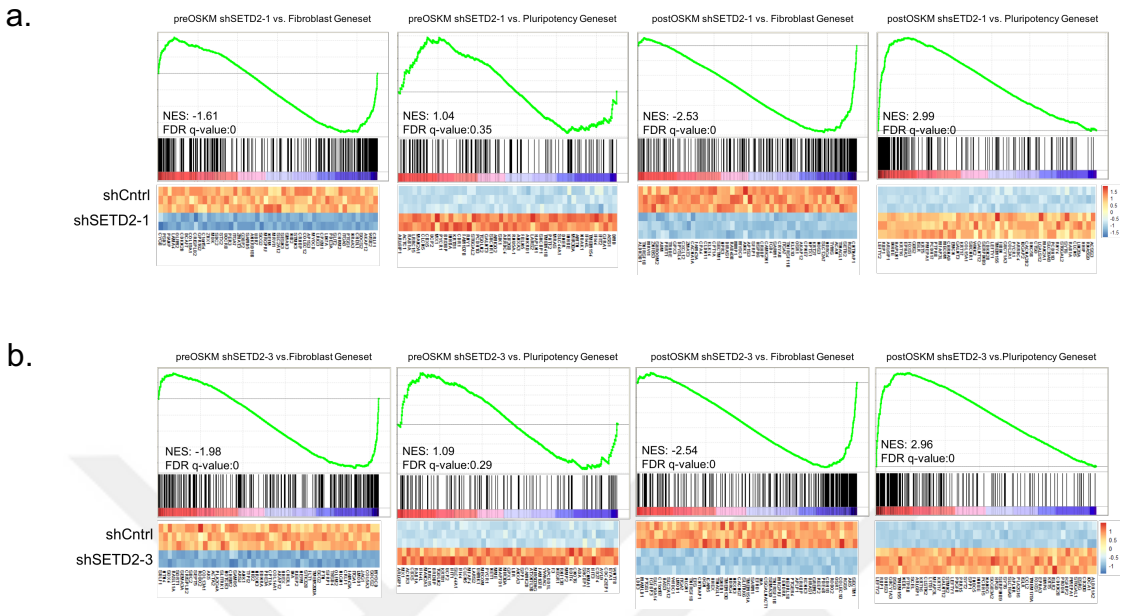
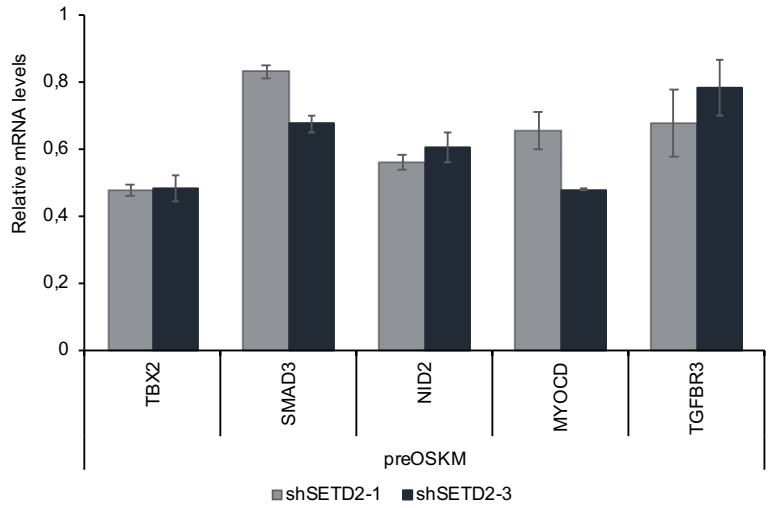


Figure 41. GSEA analysis for a. shSETD2-1 and b. shSETD2-3 samples showing enrichment profiles in fibroblast and pluripotency gene sets.
Heatmaps below the enrichment graphs show the relative expression of the top 50 genes belonging to the negatively or positively (according to enrichment profile) enriched gene sets queried.

We verified the relative expression levels selected genes from the RNA sequencing data by quantitative RT-PCR. We collected RNA samples from shCntrl- and shSETD2-infected cells in preOSKM and postOSKM conditions. We selected *TBX2*, *SMAD3*, *NID2*, *MYOCD* and *TGFBR3* as representative common downregulated genes in both shSETD2-1 and shSETD2-3 samples and performed RT-PCR experiments for 2 different biological replicates obtaining an average relative gene expression level. All selected genes had reduced mRNA levels in both shSETD2 samples compared to control cells. The extent of downregulation was variable but was apparent in both pre- and post-OSKM conditions. These results confirmed our RNA-sequencing analysis and demonstrated that depletion of SETD2 downregulates this specific set of fibroblast-related genes (Figure 42).

a.



b.

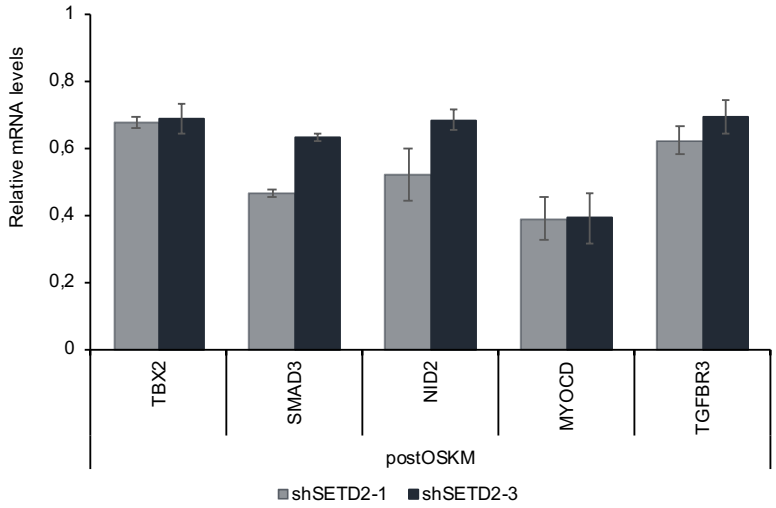


Figure 42. Expression levels of selected genes from RNA-sequencing analysis of shSETD2 cells. mRNA expressions of both shSETD1 and shSETD2-3 infected cells of **a.** preOSKM and **b.** postOSKM samples were normalized to their shCntrl group. BetaACTIN was used as internal control. Number of biological independent experiments are indicated with n, each with 2 technical replicates; error bars, ± se.

This finding supports that SETD2 directly regulates SMAD3 and MYOCD and their downregulation during reprogramming may be one of the reasons for improvement in iPSC generation. Over-representation of pluripotency related LIN28A/B, LEFTY1/2, NANOG, DPPA4 and SALL4 genes [217] in SETD2 downregulated postOSKM samples were a strong indicator of positive influence of SETD2 depletion on

reprogramming process. As a result of all these studies, we clarified the barrier effect of SETD2 and H3K36me3 mark on iPSC generation via explaining the molecular mechanisms underlying changes in initial stages of reprogramming.

3.8. SETD2 knock-down changes distribution of histone marks

Having shown that depletion of SETD2 results in repression of fibroblast-related genes, we next investigated how this perturbation affects the chromatin structure of target genes. To this end, we examined H3K36me3 and H3K27me3 levels at select genes by chromatin immunoprecipitation (ChIP).

We illustrated distributions of H3K27me3 and H3K36me3 marks at regions, those will be amplified with selected primer sets. Using IGV viewer, we marked amplification sites from the ChIP data obtained from in NHDF, H1 hESC, and hiPSC cell lines. ChIP data with pink peaks demonstrates H3K36me3-enriched regions for each cell line and green peaks show H3K27me3-enriched sites in the selected regions. Red peaks under each figure are RNA sequencing data representations for human iPSC cell line and normal human dermal fibroblasts, identifying expressed regions. Primer sets were designed to amplify the regions pointed out with vertical red lines (Figure 43).

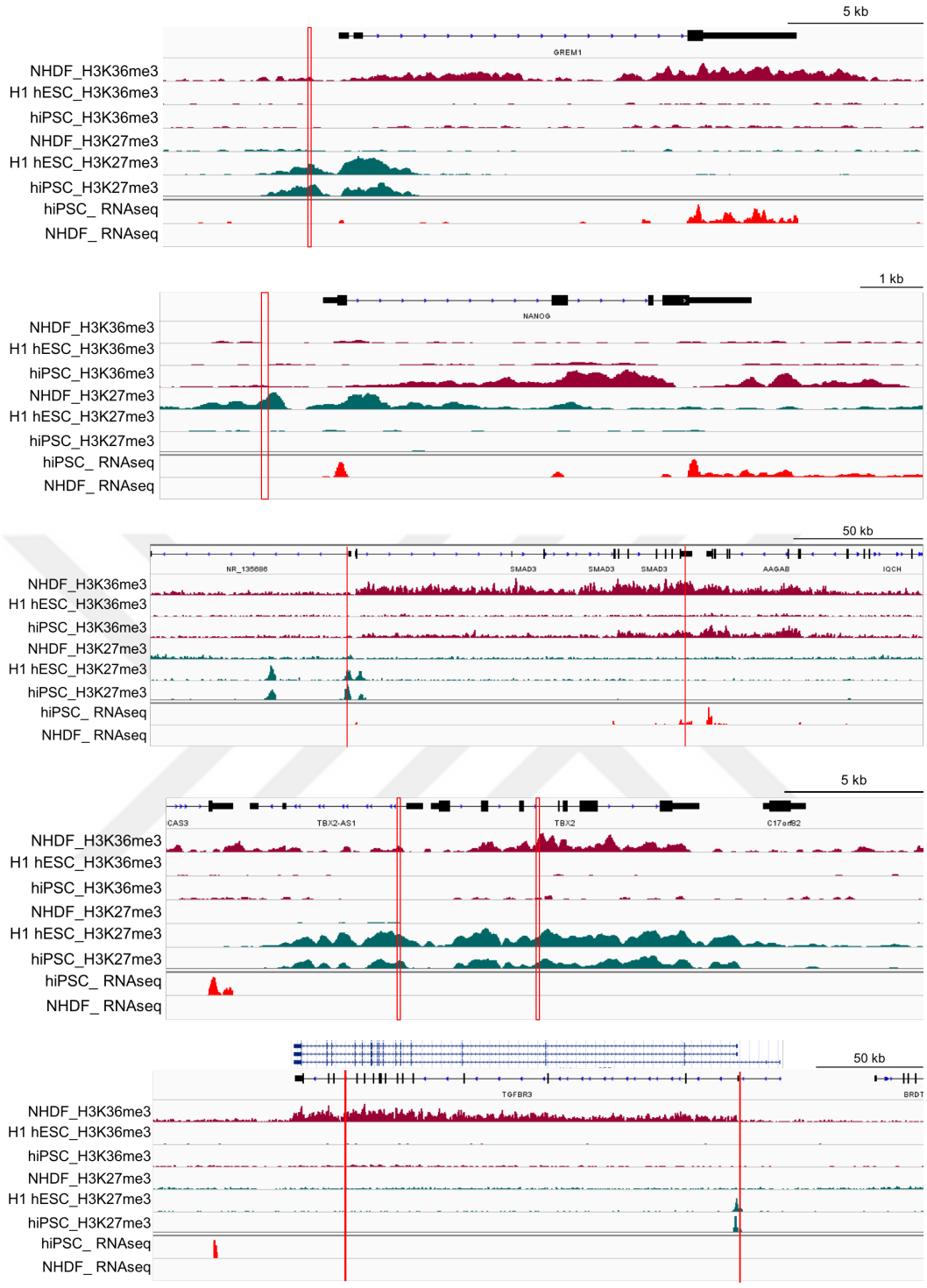


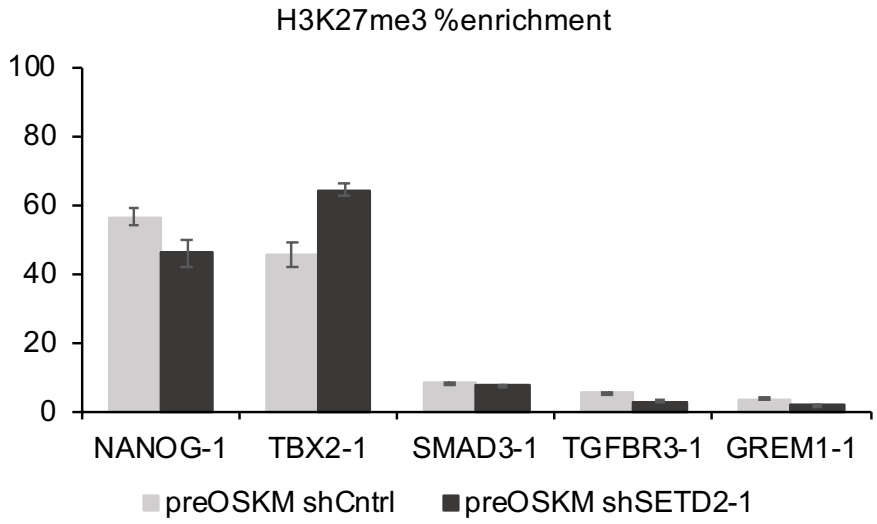
Figure 43. Illustration of ChIP-seq and RNAseq data from fibroblast and pluripotent stem cell lines.
IGV images for ChIP sequencing of cell lines obtained with H3K36me3 and H3K27me3 antibody pull downs. Vertical red lines demonstrate amplification sites with primer pairs designed for ChIP-qPCR analysis.

Primer sets were designed according to RNA sequencing data to detect H3K27me3 distributions on downregulated genes and their regulatory sequences. First, uninfected fibroblast cells were subjected to ChIP with H3K27me3 1º antibody to obtain DNA regions enriched in those epigenetic marks. Immunoprecipitated Histone-DNA structures were decross-linked to obtain pull down DNA fragments. qPCR analysis was performed with primers amplifying regions aforementioned in (Figure 44).

We analyzed H3K27me3 enrichments at promoter regions of upregulated *NANOG* and downregulated several genes upon SETD2 knock-down, in preOSKM samples. When *NANOG* promoter was considered, we detected a decrease in repressive H3K27me3 deposition in shSETD2 cells compared to control cells. In addition to that, H3K27me3 enrichment was quite high at *TBX2* promoter in shSETD2 samples compared to control cells. We could not detect differential enrichment for H3K27me3 mark at promoter regions of other downregulated genes among control and shSETD2 samples (Figure 44.a).

Ultimately, we analyzed H3K27me3 enrichments at promoter regions of pluripotency-related *NANOG* and genes differentially expressed in fibroblast cells in postOSKM samples. As expected, we saw a drastic decrease in deposition of repressive H2K27me3 mark at *NANOG* promoter of shSETD2 cells compared to control ones. We also detected an enrichment for H3K27me3 at *TBX2* promoter region in shSETD2 samples compared to control cells. However, other selected genes (*SMAD3*, *TGFBR3*, *GREM1*), those were known to be downregulated upon SETD2 depletion, showed no enrichment of H3K27me3 at their promoter regions in shSETD2 samples (Figure 44.b).

a.



b.

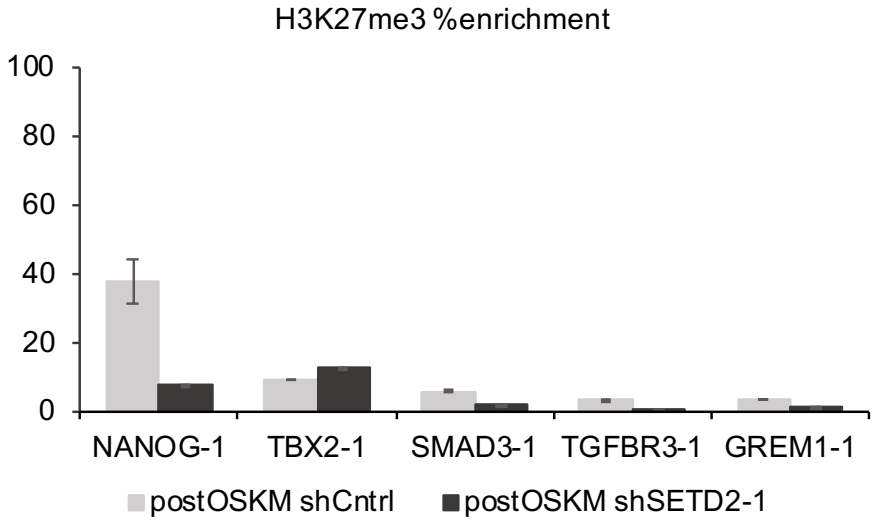


Figure 44. H3K27me3 enrichments in DNA regions of control and shSETD2 cell of pre- and postOSKM samples.
ChIP analysis coupled with qPCR showing H3K27me3 deposition levels on promoter regions of selected genes in control and shSETD2 cells of **a.** pre-OSKM and **b.** postOSKM samples. Data is represented with 2 technical replicates; error bars, ± se.

3.9. SETD2 downregulation eliminates requirement of KLF4 and cMYC for reprogramming

We have clarified that downregulation of SETD2 in initial stages have positive influence on reprogramming process. Especially, RNA sequencing data unveiled under-represented genes in shSETD2 cells are related to fibroblast identity and required to be suppressed efficiently for cells to pass the bottle neck through intermediate state. We already know that introduction of OSKM transcription factors is sufficient to reprogram fibroblast into their undifferentiated pluripotent stem cell state, and their ectopic expression force cells to suppress genes associated with fibroblast identity [210]. Therefore, we asked whether downregulation of SETD2 may compensate some of the functions of OSKM factors. So, we introduced shSETD2 to fibroblasts and subjected these cells with OS-reprogramming excluding *KLF4* and *cMYC* transgenes. We conducted OS-reprogramming experiments in 3 biological replicates and all obtained results demonstrated that iPSC colony numbers were higher in shSETD2-reprogramming compared to shCntrl ones. These results helped us to explain the role of SETD2 in reprogramming since fibroblasts were sufficient to be converted into iPSCs with SETD2 downregulation in the absence of KM transcription factors (Figure 45).

These data that we obtained demonstrated that, SETD2 depletion intercepted the requirement of KLF4 and cMYC via supplementing some of their functions.

a. b.

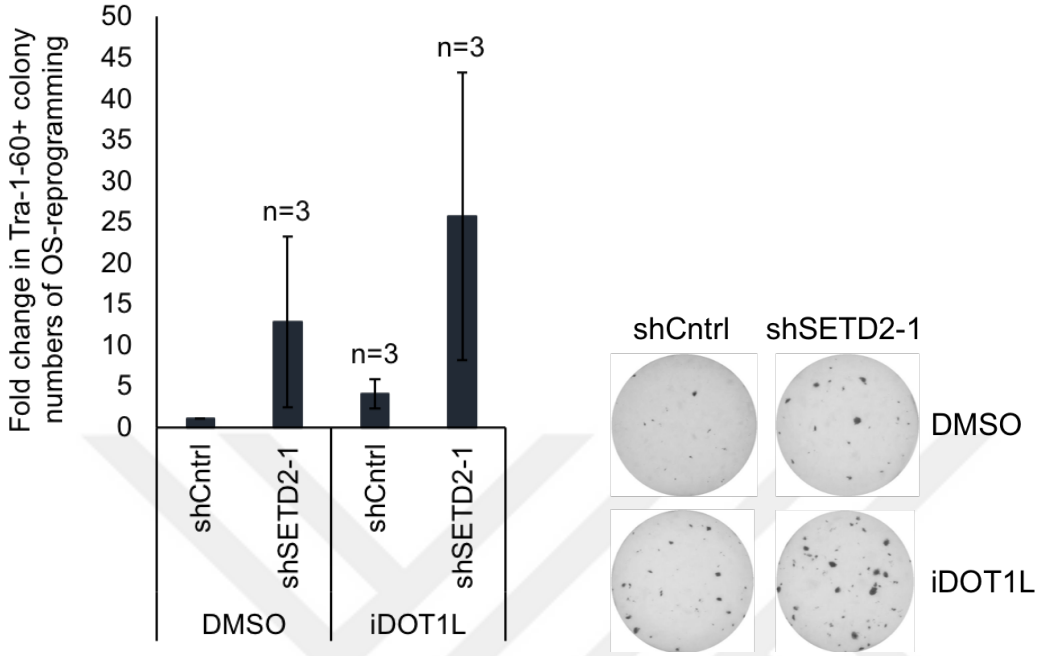


Figure 45. Downregulation of SETD2 is sufficient to induce reprogramming with OS transcription factors.

a. Fold change inthe number of Tra-1-60 positive iPSC clonies of OS-mediated reprogramming of control and shSETD2 cells treated with either DMSO or iDot1L. Number of biological independent experiments are indicated with n, each with 3 for technical replicates. P values were determined by a two-tailed Student t-test; * $P < 0.05$, ** $P < 0.01$; error bars, ± se. **b.** The increment in Tra-1-60+ colony numbers in shSETD2-infected cells is more visible in representative well images both in DMSO and iDOT1L samples.

4. DISCUSSION

This project was carried out to investigate the role of one of the epigenetic regulators, SETD2, on reprogramming of fibroblasts into induced pluripotent stem cells. Several studies performed to identify the influence of epigenetics on reprogramming. During reprogramming, transcriptional network undergoes global changes to suppress somatic expressions and activate pluripotency-related gene expressions. Diverging from somatic identity thorough pluripotency is under the control of epigenetic mechanisms [218].

SETD2 is a multi-domain protein with several biological functions. The best-known function of this protein is its methyltransferase activity, where it is responsible for trimethylation of Histone H3 Lysine 36 residue. SETD2 catalyze H3K36me3 via its recruitment to H3K36me3 is associated with active transcription [105]. SETD2 binds to serine-2 phosphorylated CTD of elongating RNAPII and recruited to transcribed regions to catalyze trimethylation of histone H3K36 residue [118]. In the light of these findings, we hypothesized that catalytic function of SETD2 may be important for regulation of reprogramming.

There are several studies discussing the influence of histone H3 lysine36 methylation on development and reprogramming. For example, one study explains that H3K36 demethylase KDM2B removes H3K36me2 marks on the promoters of early activating genes and increases the percentage of OCT4-GFP positive cells in mouse reprogramming. The repression activity of KDM2B on senescence by downregulating Ink4a/Arf complex and thus increase in the cell proliferation is another contributor to the enhancement of reprogramming [99]. Another supporting study explaining that, supplementation of vitamin C activates its downstream effectors, KDM2A and KDM2B, and their activation enhances reprogramming via decreasing H3K36 methylation [185].

Histone demethylation role of KDM2B is counteracting with histone methyltransferase activity of SETD2 protein. Therefore, we considered the possibility that, positive effect of KDM2B overexpression on iPSC generation may be mimicked by suppression of SETD2. As we expected, we demonstrated the positive influence of SETD2

downregulation on reprogramming of fibroblast cells under the influence of H3K36me3 depletion affecting somatic gene expressions in an early stage.

KDM4B is H3K9/H3K36 demethylase whose role in reprogramming was studied. Overexpression of KDM2B in mouse embryonic fibroblasts was shown to reduce H3K9 and H3K36 methylation levels and support pluripotency of iPSCs generated by SCNT. It is not quite clear that, which of the epigenetic marks is regulating reprogramming, but this study also gives an idea about the involvement of epigenetic regulations and H3K36 methylation on reprogramming.

Another motivation for us to study the influence of H3K36 methylation on reprogramming was based on another supporting investigation that, DOT1L inhibition enhances reprogramming by assisting sufficient downregulation of lineage specific genes in initial stages. DOT1L is only known methyltransferase catalyzing mono- di- and tri-methylation of histone H3 lysine 79 residues. H3K79me3 is situated on genic sequences in increasing amount through 3' end and associated with active transcription [94]. In view of SETD2 being only trimethyltransferase for H3K36 residue, which is also associated with active transcription; we asked if these two epigenetic regulators act on similar processes to enhance suppression of somatic genes in initial stages of reprogramming. Similarities in between H3K79 and H3K36 methylations on gene expressions led us to identify the positive influence of SETD2 depletion on reprogramming via suppression of somatic cell program.

We extended out study to investigate the combinatorial effect of DOT1L inhibition and SETD2 depletion on reprogramming. Prior to RNA sequencing experiments to identify differentially-expressed genes under the influence of SETD2 depletion during reprogramming. We first checked if removal of two activating marks have different targets, thus enhance reprogramming through different pathways. Our findings demonstrated that, downregulation of SETD2 increased iPSC colony numbers compared to its DMSO-treated counterpart, when combined with DOT1L inhibition. Having the knowledge that DOT1L inhibition influences on somatic gene expressions in early stages, we thought the possibility that an efficient suppression of somatic program was accomplished by contribution of SETD2 depletion in early time point of reprogramming. Reprogramming requires suppression of somatic gene expressions to initiate mesenchymal-to-epithelial transition and cells those failed to succeed aforementioned requirements are stuck in the initial stage and become partially reprogrammed [210]. Perhaps, depletion of H3K36me3 steps in at this stage and completes the requirements of the partially-reprogrammed cells so that they can pass to the next stage (Figure 46).

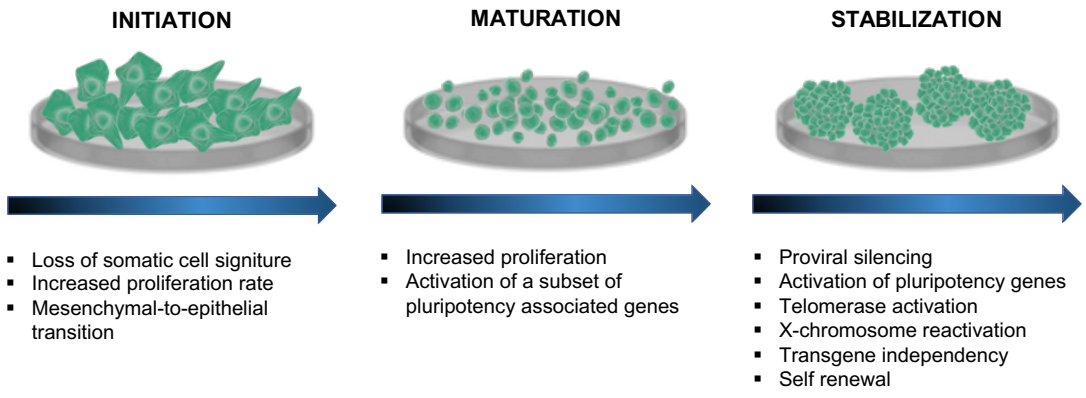


Figure 46**. Illustration of somatic cell reprogramming stages.**
Key molecular events for each cellular reprogramming stages are listed with bullet points.

In the light of our investigations, we decided to identify the genes those are under the control of SETD2 during reprogramming. Therefore, we investigated the positive influence of SETD2 depletion in different stages of reprogramming. Our findings clarified that, downregulation of SETD2 is important in initial stages of reprogramming possibly affecting somatic gene expressions. This investigation let us to analyze genes those expressions are affected by SETD2 depletion in initial stages of reprogramming. For this purpose, we performed RNA sequencing to SETD2-depleted cells prior to reprogramming and 6 days after OSKM induction. RNA sequencing helped us to investigate molecular mechanisms underlying the positive influence of SETD2 depletion on reprogramming. We investigated differentially expressed genes upon SETD2 depletion, whether they were already related to any kind of pluripotency, differentiation or developmental processes.

Prior to RNA sequencing, we first identified SETD2 depletion effect in every individual stages of reprogramming to decide which stage is most regulated via SETD2 downregulation. Reprogramming is regulated by epigenetic mechanisms at every stage. Changes in quantity and distribution of epigenetic marks during reprogramming allows cells to lose their somatic identity and conversion into an undifferentiated state. When we take H3K36me3 into consideration, this active mark is present in initial stage probably regulating somatic gene expressions and a sharp decline is seen in intermediate stage, then again, its presence become apparent in late stage to regulate pluripotency-related genes expressions [181, 219]. These findings correlate with our results, since we did not see any enhancement in reprogramming when we deplete SETD2 at intermediate stage, where H3K36me3 deposition seems to be not present already. Moreover, downregulation of SETD2 at late stage had no effect on reprogramming efficiency, since proper suppression has not been proceeded in initial stage and cells could not meet the requirements to reach the late stage (Figure 47).

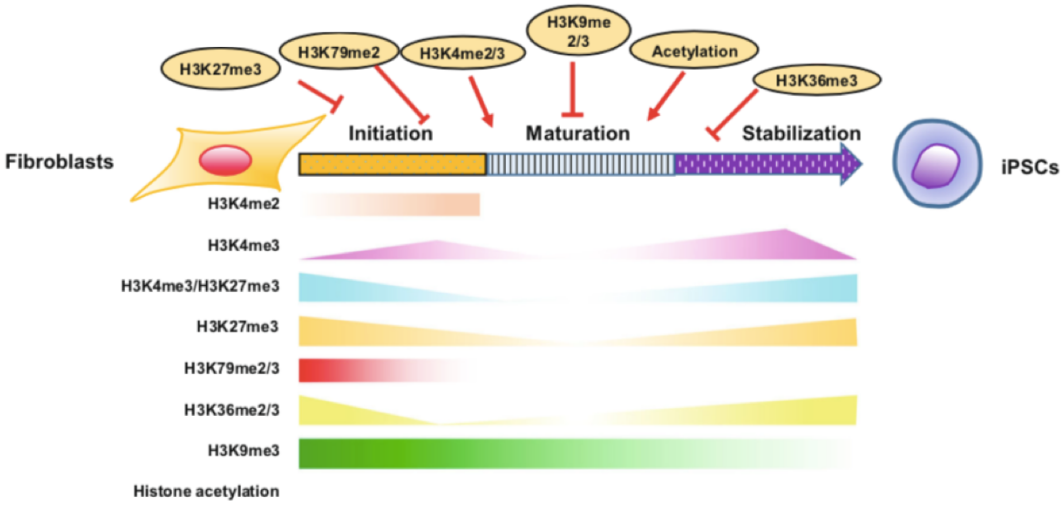


Figure 47. Change in histone modification levels throughout reprogramming.
(Adapted from Qin et al. [219]).

The most important finding was when we force shSETD2 throughout reprogramming, the positive influence of that we saw at initial stage was neutralized by shSETD2 expressions at intermediate and late stages. This discovery states that SETD2 expression was required at late stage most likely to regulate pluripotency gene expressions and somehow, we observe enhancement in iPSC generation even we stably express shSETD2.

Our findings let us to subject cells those are under the influence of SETD2 depletion prior to OSKM-induction and cells those belongs to initial stage of reprogramming, to RNA sequencing. According to RNA sequencing results, where we compared gene expression levels among control cell and SETD2-depleted cells, one of the under-represented gene in SETD2-depleted cells was TGFBR3. This glycoprotein is functioning in one of three TGF-β receptors, can be found as soluble or membrane-bound form inside the cell [220]. We observed that this receptor was downregulated in postOSKM samples as a result of SETD2 depletion. TGF-β is already known to be a barrier for reprogramming and its inhibition is displayed to substitute both SOX2 and cMYC in iPSC generation process [221]. One other study is also claiming that, eliminating TGF-β with its specific small molecule inhibitor can replace OCT4 via modulating downstream kinases during mouse reprogramming [222]. Therefore, we investigated the TGFBR3 function on reprogramming even it did not come up as an under-represented gene in preOSKM samples. Transcriptome analysis during

reprogramming of human fibroblasts is clarifying that, TGBBR3 has higher expression levels in early phases functioning in inhibition of mesenchymal-to-epithelial transition, then its expression falls in later phases [223]. Besides, TGFBR3 is known to be already inhibited with OCT4 and SOX2 expressions [172], we think that sufficient downregulation of TGFBR3 under the influence of SETD2 depletion positively affected iPSC generation.

SMAD3 was one of the commonly downregulated genes both in preOSKM and postOSKM samples. SMAD3 is a downstream protein of TGF-β signaling pathway and regulated differentiation of smooth muscle from mesenchymal stem cells (MSCs). Inhibiting SMAD3 with small molecules leads to decrease in MYOCD levels [224], which was also downregulated upon SETD2. Both proteins are required for smooth muscle differentiation and their downregulation even in preOSKM samples might have supported pluripotency by interfering with lineage identity.

Another approach claiming that, inhibition of TGF-β increases SMAD3 expression and this situation enhances iPSC generation. Overexpression of SMAD3 positively regulates reprogramming via interacting with core pluripotency markers but depletion of endogenous SMAD3 have no significant effect on iPSC generation [225]. One other study explains that, counteracting activity of SMAD3 with Polycomb repressive complex regulates OCT4 expression and is inhibiting differentiation of mouse ESCs [226]. This finding does not support our consideration that downregulation of SMAD3 may avoid muscle differentiation and help fibroblasts to lose their lineage identity. However, additional tests should be done to identify the effect of SMAD3 expression on reprogramming process.

MYOCD was also identified as commonly downregulated in SETD2-depleted cells both in pre- and post-OSKM samples. MYOCD is required for cardiomyocyte and smooth muscle differentiation from ESCs and this protein is also a downstream effector protein of TGF-β signaling [227]. Despite TGF-β not being in the confidence interval of downregulated gene list in preOSKM samples, SMAD3 and MYOCD were identified as under-represented in those cells. This finding supports that SETD2 directly regulates SMAD3 and MYOCD and their downregulation during reprogramming may be one of the reasons for improvement in iPSC generation.

Another downregulated gene both in pre- and post-OSKM samples was NID2, mainly functioning in extracellular matrix remodeling, invasion and migration. It has been demonstrated that NID2 is one of the genes upregulated during BMP4-mediated trophoblast differentiation of hESCs [228]. Considering our results, downregulation of NID2 via shSETD2 introduction to cells might have negatively affected fibroblasts inclining to trophoblast differentiation during reprogramming. As a future direction, it will be a good approach to test the effect of NID2 depletion on reprogramming of fibroblasts into iPSCs.

TBX2 is one of the T-box transcription factor family protein functioning as a repressor of transcription and leading to differentiation to certain tissue types during development. While TBX family proteins are generally involved in limb bud formation, TBX2 is known to be also important for digit development [229]. Induction of TBX2 expression is also seen during differentiation of iPSCs into cardiac lineage [230]. As can be understood from the results of the studies done, TBX2 is an indicator that plays role in differentiation process and its expression is not present in pluripotent stem cells. Under-representation of TBX-2 both in pre- and post-OSKM samples upon SETD2 downregulation might prevent fibroblasts to transdifferentiate into other cell types during reprogramming. We think that TBX2 downregulation upon shSETD2-infection might be effective in loss of fibroblast identity and facilitation of mesenchymal-to-epithelial transition.

Since H3K36me3 is being an active mark and SETD2 downregulation will cause decrease in expression of actively-expressed genes in fibroblasts, we considered over-represented genes were as a consequence of indirect effect of SETD2 downregulation. We did not come across any genes that we thought were directly related to pluripotency or facilitation of pluripotency in preOSKM samples. At least, over-representation of pluripotency related LIN28A/B, LEFTY1/2, NANOG, DPPA4 and SALL4 genes [217] in SETD2 downregulated postOSKM samples were a strong indicator of positive influence of SETD2 depletion on reprogramming process.

We had available fibroblast and pluripotency gene sets those were generated by identification of differentially-expressed genes obtained by comparing microarray data from 2 different fibroblast and 4 different pluripotent stem cell lines. In the light of these

gene sets, we procured a list of genes those most probably regulate fibroblast identity and pluripotency. We compared our RNA sequencing data with these gene sets and designated common genes those were under the influence of SETD2 downregulation.

We performed comparison analysis among gene sets and RNA sequencing data obtained with RNAs collected at 2 distinct time points of reprogramming. The former data included differentially-expressed genes upon SETD2 downregulation in fibroblast cells prior to OSKM-induction and the latter was obtained from RNA pool of shSETD2-infected cells those were subjected to OSKM-mediated reprogramming and collected at the end of initial stage.

When we analyzed top 50 negatively-enriched genes in fibroblast gene set, we came up with MYOCD and NID2 in both shSETD2 samples and SMAD3 for shSETD2-1. We believe that, genes those negatively-enriched in fibroblast gene set for both pre- and post-OSKM samples facilitated reprogramming process under the influence of SETD2 downregulation. Therefore, our claim is that decrease in SETD2 expression resulted in loss of H3K36me3 active marks on lineage-specific genes in early stages there so, helped to lose fibroblast identity more efficiently to induce MET transition. As a result of all these studies, we clarified the barrier effect of SETD2 and H3K36me3 mark on iPSC generation via explaining the molecular mechanisms underlying changes in initial stages of reprogramming.

We demonstrated that SETD2 depletion downregulates several genes related to fibroblast identity those are mainly suppressed by induction of Yamanaka factors. The increment in iPSC generation upon SETD2 knock-down can be explained as, suppression of lineage-specific genes those should be silenced in initial stages of reprogramming by induction of OSKM factors were supported by loss of H3K36me3 active marks on those genes. Therefore, we asked if SETD2 depletion may replace some of the functions of these 4 transcription factors on iPSC generation.

Yamanaka factors are sufficient to induce reprogramming. However, some those defined factors are not mandatory for iPSC generation. Depletion of several genes has been shown to replace one or more OSKM factors. For example, alternate transcription factors such as KLF2 and L-MYC and cMYC have been replaced with KLF4 and MYCN

to facilitate reprogramming [23, 24]. We also mentioned before that NANOG and LIN28 replaces KLF4 and cMYC [26].

There are several experiments identifying some other genes those sufficient to induce iPSC generation in the absence of one or more Yamanaka factors. Inhibition of TGF-β signaling has been shown to abolish the requirement of SOX2 and cMYC in mouse reprogramming [221]. Estrogen-related receptor b (ESRRB) is one of the NANOG responsive gene, whose overexpression can reprogram NANOG$^{-/-}$ mouse fibroblast cells and sufficient to sustain reprogramming by replacing KLF4 [231]. TET1 is also claimed to replace OCT4 in mouse reprogramming via demethylation of endogenous OCT4 DNA and induce its activation [53]. Inhibition of DOT1L methyltransferase which was reported to enhancing iPSC generation is one other replacer for KLF4 and cMYC [94].

In the light of studies performed to identify genes those have the ability to replace OSKM factors, we wasted to check the ability of SETD2 depletion to replace one or more of Yamanaka's factors for iPSC generation. Therefore, we subjected shSETD2-infected fibroblasts to OS reprogramming in the absence of KLF4 and cMYC transgene expressions. We obtained very small numbers of iPSC colonies of shCntrl cells, where we observed SETD2 depletion increased iPSC colony numbers in OS reprogramming compared to shCntrl group. DOT1L inhibition increased iPSC colony numbers of OS-mediated reprogramming both in shCntrl- and shSETD2-infected cells showing its additive effect when coupled with SETD2 downregulation. These data we obtained demonstrated that, SETD2 depletion intercepted the requirement of KLF4 and cMYC via supplementing some of their functions.

H3K36me3 being an active mark, regulating thousands of genes and SETD2 downregulation therefore, is affecting gene expressions globally in the genome. Using shRNAs decrease mRNA levels to a level where the remainder may be enough for cells to survive and sustain cellular functions. SETD2 is vital for mouse development and mouse embryos comprising homozygous knock-out deletion of SETD2 dies at E10.5-E11.5 days of embryonic stages [135].

We have previously shown the positive influence of SETD2 downregulation on iPSC generation and wanted to explore the effect of total loss of SETD2 on reprogramming. Using Crispr-Cas9 system, we generated SETD2 knock-out fibroblasts with decrease H3K36me3 phenotype and subjected them to OSKM-mediated reprogramming. Repeated reprogramming experiments indicated that, SETD2 knock-out fibroblasts were sufficient to be reprogrammed and even increased iPSC colony numbers compared to shCntrl cells. We also observed additive effect of DOT1L inhibition on iPSC generation when SETD2 knock-out cells were treated with small molecule inhibitor specific for DOT1L. iPSC colonies obtained from reprogramming of SETD2 knock-out fibroblasts were expressing Tra-1-60 cell surface marker, however further experiments should be performed examine pluripotency and stemness capacities of those obtained iPSC cells.

SETD2 protein has several functions rather than being methyltransferase for histone H3K36me3. Thus, the increment in iPSC generation may not be related to only decrease in H3K36e3 levels upon SETD2 knock-down. Therefore, we approached our experiments with a different perspective to support our hypothesis that H3K36me3 active marks being important for fibroblast-specific gene expressions and its removal facilitates suppression of target genes during reprogramming. Thus, we targeted histone proteins without interfering with SETD2 expression.

H3.3K36M mutation has been known to prevent methylation of histone H3 at lysine 36 residue [169]. Therefore, we generated lysine-to-methionine mutations at 36th residue of histone H3.3 gene and overexpressed this construct in fibroblast cells. We saw a clear enhancement in reprogramming for H3.3K36M-infected cells compared to wild type H3.3 overexpressed control group. Moreover, DOT1L inhibition also increased iPSC colony numbers when combined with H3.3K36M overexpression. This result demonstrated that, decrease in H3K36me3 levels without intervening other cellular activities of SETD2 protein was sufficient to enhance reprogramming of fibroblast cells into iPSCS.

So far, there has been no indication issued in literature explaining effects of histone mutations on reprogramming. This study, in which we demonstrated the positive

influence of H3K36M mutation on reprogramming process, is a novel work that will shed a light on literature.

As mentioned before, SETD2 protein have different functions in several biological processes. Most of them are related to its catalytic function serving as recruitment site for the writer proteins those have PWWP domain [232]. H3K36me3 recruits tumor suppressor gene ZMYND11 to actively transcribed regions, to regulate transcription by suppressing elongation of RNAPII which is an important in cell cycle regulation [233]. PRC2 complex is also recruited to double strand break regions via interacting with H3K36me3 mark [114]. DNMT3B is another reader protein of H3K36me3 where it is recruited to transcribed regions to prevent cryptic transcription by methylating DNA at those regions [124]. Studies on DNMT3B demonstrated that this protein is dispensable for cellular reprogramming [234]. One other function of SETD2-mediated H3K36me3 is regulating splicing machinery. Exonic DNA regions, whose products are determined to be included to the mature mRNA, are enriched in H3K36me3 marks where change in the deposition of H3K36me3 affects splicing rates [126].

SETD2 protein itself is recruited to actively transcribed regions via directly interacting with IWS1 protein which provides interaction between SETD2 and elongating RNAPII protein during transcription [116]. Expression of SETD2 protein is tightly regulated by SPOP, which is a specific subunit of CUL3 ubiquitin E3 ligase complex. Direct binding of SPOP to SETD2 induced polyubiquitination of this protein further destined to degradation [117]. From all these functional activities, we are not sure if any of them have influence on reprogramming. Further studies should be done to identify their roles in iPSC generation.

5. APPENDIX

Table 30. preOSKM downregulated genes in shSETD2-1		
preOSKM shSETD2-1		
log2FoldChange	padj	gene name
-0,501263613	1,53743E-05	AR
-0,502003176	4,08729E-42	CAV1
-0,502901988	1,80575E-39	IGFBP4
-0,503820696	4,30529E-07	SCARA3
-0,508982254	0,009107854	HNRNPCP2
-0,510141833	0,006411414	ISYNA1
-0,513038133	0,000211605	JADE2
-0,514230328	0,000960281	MICALL2
-0,51444773	2,37654E-05	MZT1
-0,51488373	2,47166E-14	PCDH18
-0,516309981	6,37145E-11	SCUBE3
-0,517987584	6,1836E-18	CFL2
-0,519803422	4,21006E-20	ANLN
-0,523521148	0,009508563	MYRF
-0,527031724	0,005335612	RPL4P4
-0,528577759	0,005897618	RPLP0P6
-0,533730084	2,98433E-05	KALRN
-0,534946161	0,003007543	XKR8
-0,535465171	3,78842E-29	TMED10
-0,536860307	0,000271697	TBX2
-0,537369356	6,38207E-06	PLAT
-0,539941617	2,98729E-05	GEM
-0,54170531	6,58174E-09	OXR1
-0,542370756	0,009779858	AC009093.1
-0,543382915	3,77969E-07	DCHS1
-0,543552011	6,59082E-05	ADAMTS7
-0,544085937	0,009367194	LSAMP
-0,54433954	0,001997547	TEF
-0,551605982	0,007818441	ZNF607
-0,554574985	0,005204215	VPS37D
-0,555042199	5,87837E-09	SOCS5
-0,556164789	0,003179454	CNNM4
-0,558284061	0,000299585	HCFC2
-0,561301834	2,08922E-06	OLFML2B
-0,563354006	0,007818441	PCBP3
-0,564489255	0,001085671	SOBP
-0,568187981	0,000250858	PLA2G4A

-0,569841509	0,003490048	FICD
-0,577689689	0,00620928	RAB40B
-0,578747867	0,005724399	AC080112.2
-0,579738449	0,002658538	NDRG4
-0,582985461	6,98512E-05	PALM
-0,583142997	0,00571705	NCALD
-0,584485104	4,29103E-15	DICER1
-0,588628093	0,005215059	TNFSF12
-0,591153756	3,38698E-05	STK32B
-0,591370064	3,72676E-15	SEMA3A
-0,593044137	2,73148E-11	BTBD1
-0,593151205	0,00069888	ZNF853
-0,593588893	2,42694E-23	RGS4
-0,593762187	0,000128658	GAMT
-0,593980046	0,00415815	B4GALNT1
-0,59494754	1,46388E-29	SFRP1
-0,596287832	3,87751E-07	THSD4
-0,596479931	5,30445E-05	ITGA6
-0,598077392	3,77358E-06	LDB2
-0,598184127	3,59773E-07	CPNE2
-0,600564796	5,84466E-05	DRP2
-0,60120122	1,3511E-10	AK5
-0,603498773	1,37695E-10	H1F0
-0,606998345	0,000369019	APOE
-0,61013714	9,67171E-08	RNF152
-0,610822654	0,002266113	RPL7P9
-0,61100938	0,001615841	MAN1C1
-0,612786047	0,002381349	KIF26A
-0,618011264	0,002671218	CTXN1
-0,620692263	2,03164E-08	SVIL
-0,621281627	0,001125952	PRTFDC1
-0,621771805	4,24961E-17	NOTCH3
-0,622285204	1,12787E-23	MXRA5
-0,622553605	0,002233117	DENND2A
-0,623438571	0,001168932	AC134312.5
-0,623955226	0,002446135	MARCH9
-0,624214531	6,3713E-08	HSPB6
-0,624346111	2,80686E-35	ACTG2
-0,625081681	0,000186201	TNFRSF19
-0,626680876	1,09887E-24	KITLG
-0,627165619	1,61285E-29	MMP1
-0,627198679	0,001257545	SHC4

-0,627284041	1,10571E-07	PGM2L1
-0,629493887	1,81131E-08	CPT1A
-0,632309849	3,78227E-15	TBC1D9
-0,633201401	0,001874821	TCF7
-0,634708921	0,00050633	GBP2
-0,638432464	0,00016508	GAS1
-0,640904786	0,000741003	PSMC1P1
-0,641875569	3,98376E-05	ZBTB10
-0,64213919	5,71242E-09	ANGPTL2
-0,64216964	0,001541994	GABRB1
-0,642275809	2,38957E-06	BCAM
-0,645031183	1,86236E-06	PLTP
-0,647719048	0,000250597	COL13A1
-0,649658457	5,281E-15	NETO2
-0,653420408	0,001257545	PRRT2
-0,653587226	1,00089E-09	PAMR1
-0,654134639	0,000413021	APLP1
-0,654618229	0,000793893	CA11
-0,658193298	0,000444897	CCDC136
-0,659037577	1,144E-21	SPON2
-0,659447296	0,000462218	GAL3ST4
-0,664058853	0,000258221	TGFBR3
-0,665945711	0,00015963	STC1
-0,668291418	1,98515E-15	BLOC1S6
-0,668436395	1,20015E-06	AC022149.1
-0,674403397	6,70104E-08	PLEKHG4
-0,676176869	0,000617605	OIP5
-0,67724294	1,02786E-43	FHL1
-0,678216911	4,91108E-06	SYNJ1
-0,678237367	4,27213E-12	EPHA5
-0,679087724	1,85904E-06	PREX1
-0,680179903	5,43429E-08	DENND3
-0,680352546	8,18099E-21	COL5A3
-0,681685835	4,91068E-10	SUMF1
-0,683681006	7,83742E-36	ARSJ
-0,687210115	3,66179E-08	LAMA5
-0,688443781	9,0061E-11	THBS3
-0,695050323	6,57367E-13	TM4SF1
-0,698154752	0,000234566	A2M
-0,702904899	0,000397948	FENDRR
-0,703738171	3,87489E-10	NCEH1
-0,708282754	2,48491E-18	CNTNAP1

-0,709895137	1,49178E-05	BHLHE41
-0,710307728	8,96647E-06	OLFML2A
-0,714506025	9,40758E-07	FGD4
-0,717874357	2,43374E-08	CLU
-0,720794797	6,37145E-11	SAMD8
-0,724343517	0,000106336	CYGB
-0,725446005	2,1858E-19	PSD3
-0,730085578	0,000196017	PIANP
-0,734480671	2,58561E-40	LAMA4
-0,737343617	2,94475E-12	OLFM2
-0,739384308	8,84112E-05	CDKN2D
-0,74350432	1,19708E-29	DCAF7
-0,747905528	4,2459E-16	COL18A1
-0,748173725	9,69153E-09	CRISPLD2
-0,754171333	0,000109377	GPR162
-0,75595613	8,32836E-05	C1orf54
-0,77329133	1,64705E-16	ZHX1
-0,774270878	5,01602E-25	MDK
-0,779958468	2,55075E-41	SDC2
-0,783292981	8,15099E-19	STX2
-0,784741833	7,81785E-23	UQCRB
-0,790087885	7,02594E-06	PER3
-0,791627126	2,18433E-07	ITGA7
-0,793638532	1,15269E-10	ZMAT3
-0,801055617	4,92153E-48	G3BP2
-0,81030113	7,28397E-07	SAMD5
-0,810752621	2,25381E-06	ANKRD18B
-0,823694631	5,14087E-13	UBR7
-0,826404411	1,52016E-10	SCG2
-0,827885086	1,55624E-07	MRGPRF
-0,832759351	2,96269E-10	NEFM
-0,834033142	8,59977E-15	MRVI1
-0,841371208	1,47602E-06	SLC2A12
-0,84735068	4,03555E-11	SESN3
-0,848356971	4,65507E-42	SMAD3
-0,857739235	5,9252E-57	NID2
-0,863909534	1,54462E-08	STMN3
-0,864497791	4,21821E-07	CRHBP
-0,864826075	8,7974E-24	COLEC12
-0,884269504	2,12236E-06	AP1S3
-0,884685919	1,03785E-25	MYOCD
-0,893610827	1,5285E-17	LRIG3

-0,901259887	7,1663E-07	CPXM1
-0,909474077	2,01872E-09	LIN7A
-0,911461497	4,70512E-16	PDE3A
-0,933414961	3,23986E-20	PLPPR4
-0,943330891	1,06358E-24	SLIT2
-0,946598622	2,18637E-07	CHRM2
-0,95433222	4,24736E-11	ITGA8
-0,978171957	9,2119E-30	PTGS1
-0,986447086	3,53102E-09	NR4A3
-0,991359123	3,31401E-08	CHI3L1
-1,000059462	2,0733E-09	SOCS2
-1,015545725	4,7772E-101	AKAP12
-1,017814923	6,55703E-51	WSB2
-1,024942283	3,46572E-10	CRLF1
-1,068643564	1,49484E-20	TRHDE
-1,145819077	1,03877E-16	VAT1L
-1,147191946	2,69957E-29	TFPI2
-1,23032706	9,15446E-21	PTN
-1,367537613	8,83901E-16	PTGER2
-1,41062926	5,49244E-87	SETD2
-1,417764925	4,32386E-18	HGF
-1,713330482	6,55703E-51	TYMS
-2,130320169	2,9315E-45	NPTX1

Table 31. preOSKM downregulated genes in shSETD2-3		
preOSKM shSETD2-3		
log2FoldChange	padj	gene name
-0,500509638	0,00238081	PRKAR2B
-0,501106692	0,000194411	BCL6
-0,503759418	0,007204212	CACNG4
-0,507254508	1,64802E-24	LTBP1
-0,510398502	0,009831123	FNDC10
-0,516258254	1,83624E-16	LAMA4
-0,516517243	0,00454731	HAPLN3
-0,516560332	5,55433E-06	SLC7A2
-0,517636888	1,12615E-12	LMO7
-0,520180196	2,23089E-13	EPS8
-0,520487599	0,00744731	C11orf87
-0,522411188	6,84517E-09	SLC1A4
-0,522660106	0,005640961	PRSS35
-0,523356594	1,24981E-05	PAMR1
-0,524895979	2,34801E-05	DMPK
-0,52569301	0,004414211	CXCL1
-0,526134917	2,75547E-30	STC2
-0,527831907	4,56911E-12	USP53
-0,530281433	1,52117E-14	KITLG
-0,53548092	0,00505148	SHC4
-0,538392477	0,007969642	PODXL2
-0,539711388	0,00744731	FENDRR
-0,544671972	1,39295E-06	ANKH
-0,548803856	0,000566793	ELN
-0,549833293	4,33268E-06	LRIG3
-0,550327238	0,000293837	PDGFRL
-0,551803242	8,68365E-06	FZD8
-0,553414978	1,1842E-08	H1F0
-0,553758353	0,005780342	DIRAS3
-0,556442998	0,000644043	IL1RAP
-0,557822638	6,4775E-24	P3H2
-0,558124024	1,20522E-07	NPR3
-0,56161163	3,28071E-06	GNG11
-0,56242999	0,004010848	FBXL16
-0,56270445	0,000973252	EFNB3
-0,564116336	0,000325272	ADAMTS14
-0,566218699	1,3385E-13	ZNF469
-0,566308	2,90919E-11	NETO2
-0,570026101	1,90762E-28	FHL1

-0,585805386	3,014E-18	RGS4
-0,588690417	0,000242981	MBNL1-AS1
-0,593776987	0,001000733	APLP1
-0,594540971	0,001858739	SLC2A12
-0,601434634	0,000218614	NAP1L3
-0,601893301	0,000949368	PTGER2
-0,602800442	8,79517E-32	UACA
-0,606925068	5,03455E-06	SYNGR2
-0,610734604	2,38543E-05	ADAMTS5
-0,610974825	4,49583E-06	BCYRN1
-0,612829823	0,001483558	B4GALNT1
-0,613195145	1,70213E-05	MAMLD1
-0,62024101	1,51624E-20	SMAD3
-0,624098468	0,000258673	HGF
-0,628127867	0,000516849	A2M
-0,630215299	2,89979E-11	SEMA5A
-0,631625025	8,11034E-06	VAT1L
-0,632508293	0,000723389	ICAM5
-0,634760325	6,24998E-07	FAT3
-0,636343936	3,17316E-38	IGFBP5
-0,637693531	4,78126E-06	TBX2
-0,637930888	5,95226E-05	LIN7A
-0,651551311	5,01043E-29	SULF1
-0,659687956	1,88656E-20	NTN4
-0,663690541	2,16841E-09	SOX4
-0,667770468	8,10429E-05	RASL11A
-0,669113668	9,70464E-16	SORT1
-0,669440164	4,61862E-18	SEMA3A
-0,671222721	6,67488E-10	CRISPLD2
-0,677376806	3,46745E-35	SDC2
-0,679136727	1,05775E-05	SEPT5
-0,68077853	0,000178217	ROBO2
-0,683388334	4,93625E-05	COL13A1
-0,684644343	1,27306E-13	AK5
-0,685567042	7,51583E-16	NECTIN3
-0,68804611	1,00462E-27	ACTA2
-0,692661681	6,06281E-08	PLA2G4A
-0,700481199	1,5625E-30	CEMIP
-0,707327487	7,68994E-05	SLITRK1
-0,713110519	2,18812E-21	NOTCH3
-0,719136541	9,66802E-05	CHRM2
-0,740820476	6,06289E-06	SAMD5

-0,756459521	2,39479E-36	ARSJ
-0,759286341	2,36471E-05	ANK1
-0,760044683	8,28755E-12	TFPI2
-0,761192935	8,32235E-45	NID2
-0,782818671	9,16229E-08	NRXN3
-0,796835401	8,09679E-17	EDNRA
-0,79830417	3,77072E-13	PDE3A
-0,801677969	6,272E-13	CPT1A
-0,813950813	1,1259E-24	COL14A1
-0,815666085	8,57642E-58	AKAP12
-0,822795783	1,80591E-08	HHEX
-0,840249803	2,28559E-06	CHI3L1
-0,846964834	1,52884E-07	INA
-0,852121385	1,62418E-07	REEP2
-0,862290142	1,36081E-10	NEFM
-0,864320877	1,38075E-10	STK32B
-0,880240979	2,83545E-14	FLT1
-0,915409791	3,5883E-19	TMEM200A
-0,922511163	2,59132E-13	SCG2
-0,922770668	7,20528E-13	PTN
-0,940762888	3,90271E-57	SFRP1
-0,944356608	3,04527E-18	THSD4
-0,945971004	1,75905E-17	OLFM2
-0,954701663	9,20791E-19	NLGN1
-0,963030548	7,76047E-10	CRLF1
-0,979391636	6,93751E-24	FAM43A
-0,995860405	1,32283E-49	ITGA1
-1,059236962	1,03949E-30	LMOD1
-1,077533741	2,61872E-46	JAG1
-1,08156669	3,67884E-54	COL5A3
-1,140787666	4,9111E-13	SOCS2
-1,142778335	3,60975E-43	MYOCD
-1,157252139	3,18168E-27	MRVI1
-1,271690669	4,98074E-39	EPHA5
-1,393625067	3,823E-43	PLPPR4
-1,402851259	7,80027E-22	FGD4
-1,465515561	2,90148E-20	NR4A3
-1,494734713	1,28288E-26	ITGA8
-1,598447061	3,4288E-102	SETD2
-1,940170434	5,31925E-42	NPTX1

Table 32. preOSKM upregulated genes in shSETD2-1		
preOSKM shSETD2-1		
log2FoldChange	padj	gene name
1,94344775	1,5191E-132	ARL6IP1
1,584183121	2,61564E-41	ACTC1
1,221755614	1,4278E-126	RAB1A
1,191749419	8,55183E-24	LRRC15
1,143970797	6,36295E-46	MAN2A1
1,129233822	1,59473E-19	CLCN5
1,097395547	1,12102E-80	CBX5
1,05536625	4,59277E-09	UCP2
1,052246361	1,31344E-09	MX1
1,041036951	1,44181E-42	PYCR1
1,035811073	2,97788E-10	BRI3BP
1,016522742	3,67355E-82	INHBA
1,009490027	2,80854E-14	ADGRG6
1,005776898	3,762E-08	KRT81
0,999591424	1,08032E-08	CDR1
0,992603666	3,10666E-23	FAM107B
0,982494295	2,0733E-09	RNY1
0,972768617	9,49348E-19	ST6GAL2
0,954578735	1,25407E-19	PCDH10
0,931058067	1,87023E-63	TGOLN2
0,928993594	2,01872E-09	GALNT3
0,926607153	3,83521E-07	SUCNR1
0,910580401	1,03659E-37	AMIGO2
0,910086397	7,49056E-10	SIK1
0,889201031	6,39311E-08	ARRDC4
0,877216238	4,07758E-12	KIF26B
0,870030716	9,99319E-08	RNU5A-1
0,866011072	5,02754E-09	ANXA3
0,848549238	2,81983E-27	FBLN5
0,833025502	1,63596E-20	ANKRD1
0,832752506	2,95068E-10	LARS2
0,806607346	1,80462E-09	HAPLN1
0,801645169	3,64929E-33	HACD3
0,800357463	5,15234E-15	SHCBP1
0,79091052	2,16827E-11	TMEM41B
0,7882641	9,00954E-22	PHTF2
0,782511975	5,09504E-15	MRAS

0,775230764	4,5651E-12	SLC44A1
0,762106556	2,39434E-05	DNER
0,759050257	1,15907E-07	INHBE
0,755720715	4,18308E-05	MIR30A
0,755438174	8,60084E-05	AMTN
0,749356685	2,06754E-20	NUAK1
0,748641691	9,27796E-05	IFI44L
0,746042355	3,43591E-06	TMEM154
0,743909619	0,00011205	IFI44
0,734541965	0,000165582	IFIT2
0,731763606	6,91919E-22	GFRA1
0,728951725	3,60826E-08	ASS1
0,724041067	1,52936E-11	DMD
0,718797106	5,83199E-24	MFAP4
0,714295284	2,00584E-18	UHMK1
0,687608468	2,06531E-05	CSRNP1
0,68568965	2,08922E-06	HAS2
0,681650612	9,94653E-09	TBC1D2
0,67727521	3,97652E-14	CHN1
0,674289792	2,648E-06	SPOPL
0,668557758	0,000376879	EDN1
0,666570078	8,63253E-53	SLC7A5
0,664580028	2,13281E-06	MINDY3
0,660530148	7,69271E-08	DACT1
0,660513209	9,03361E-17	ATXN7L3B
0,658519317	5,45651E-34	MAP7D1
0,658249881	2,80297E-06	GDF6
0,656733822	3,82117E-18	CDC42EP1
0,649491081	5,64454E-10	RAB12
0,649395097	0,00022759	EIF5A2
0,643967903	9,28584E-07	MLX
0,643820551	2,72214E-05	SPP1
0,633613862	0,000748548	COL4A4
0,632894978	1,35675E-25	SEMA3C
0,632386112	2,959E-05	LHX8
0,624091792	3,8613E-05	UBE2D1
0,620168914	8,0272E-10	KCTD5
0,619260877	3,14998E-05	RNA5SP149
0,618165188	3,34412E-05	C21orf91
0,616414125	9,30868E-07	MIGA1
0,6098769	0,000109966	IFIT3
0,608951689	0,002410239	PCSK9

0,608317048	0,00050633	PMEPA1
0,606566581	6,72364E-20	SNX3
0,604854933	1,22159E-08	IFI6
0,600357639	1,03606E-05	PLEKHA8
0,6003524	2,49411E-12	OGFRL1
0,599354727	2,0507E-09	SUDS3
0,594554369	1,30068E-12	UNC5B
0,594227451	5,01344E-05	FAM83G
0,592548524	3,18378E-07	ASF1A
0,592098467	1,05231E-05	SPSB1
0,590908158	0,004636453	NCAM1
0,590531565	1,93289E-22	SLC20A1
0,590499363	1,01487E-13	RFLNB
0,584566452	0,000127324	CTH
0,574916716	1,31169E-10	CLCN3
0,5686908	0,002170443	RDH10
0,565988855	2,82739E-65	MMP2
0,564830632	0,00032839	ZBED3
0,564485803	8,40526E-06	DCLK2
0,560991616	7,96453E-10	LBR
0,558684905	1,33375E-07	ANK2
0,555939061	0,0098946	AC010618.3
0,553689187	2,67681E-28	SLC39A14
0,553648127	0,002050728	KCTD12
0,547620605	0,001097968	IFIT1
0,541812596	3,19711E-06	AVL9
0,540774439	1,6081E-05	PKP2
0,538963643	3,66055E-52	MALAT1
0,53669146	1,15394E-08	CACUL1
0,536666016	3,62262E-20	MME
0,536557742	0,005086307	SNORA73B
0,536240516	0,000121566	DEDD
0,536039949	2,3821E-05	GCLC
0,534611869	5,75861E-15	NNMT
0,533028904	0,000170892	STX3
0,532058794	4,91068E-10	IER3
0,531719859	8,1086E-06	TUFT1
0,531677577	1,23205E-07	NOL4L
0,531056572	0,009131951	SAMD11
0,531043566	5,80662E-31	VCAN
0,5306743	2,57998E-07	SLC30A1
0,529207423	7,49506E-06	NRG1

0,523548978	0,009367194	NOX4
0,521625912	1,49442E-09	RUSC2
0,518542376	9,9171E-12	PCK2
0,516858343	6,3888E-09	NUS1
0,512861882	0,003507091	PDP2
0,511927954	6,72364E-20	CDV3
0,511911082	5,80272E-06	F3
0,511322131	0,000409953	RNU2-59P
0,507078488	3,84155E-05	GADD45B
0,506425403	0,001736816	SULT1E1
0,504917559	1,24411E-08	REEP3
0,502026647	6,65035E-26	GLS
0,501818783	0,002297459	OAS3
0,501284111	0,002699883	SLC38A5

Table 33. preOSKM upregulated genes in shSETD2-3		
preOSKM shSETD2-3		
log2FoldChange	padj	gene name
1,996017048	1,4213E-142	ARL6IP1
1,250821856	6,52622E-16	MX1
1,15221425	2,07673E-17	ACKR3
1,109711748	7,05484E-64	CBX5
1,061058302	2,50302E-87	RAB1A
0,981575956	9,13954E-09	IFI44L
0,976989502	1,01059E-33	MAN2A1
0,974762403	1,17597E-16	LRRC15
0,951585552	1,40839E-09	BRI3BP
0,935960514	5,87725E-56	TGOLN2
0,92465475	9,52153E-08	KRT81
0,915041909	1,45126E-07	UCP2
0,914960292	2,05947E-09	IFIT1
0,885862569	7,41466E-17	SLC44A1
0,884423681	2,81635E-07	IFI44
0,859385805	1,25422E-10	CLCN5
0,858493291	2,60729E-31	MFAP4
0,836107562	4,44785E-14	LARS2
0,823210818	2,85107E-21	UHMK1
0,803391372	1,23411E-08	MECOM
0,789336114	9,0485E-21	PYCR1
0,777245627	4,78424E-14	NET1
0,773733343	2,93873E-35	MME
0,772852579	2,66099E-41	MAP7D1
0,762317619	8,11034E-06	MIR30A
0,738044196	8,27942E-20	GPX1
0,735457213	1,83624E-16	LBR
0,727347786	1,18265E-06	ANXA3
0,723731928	8,92827E-07	SPP1
0,704848521	1,40866E-18	RFLNB
0,700859119	2,58E-25	CAMK2D
0,696646729	4,13133E-24	HACD3
0,695942244	3,04478E-17	ATXN7L3B
0,693845669	6,7303E-08	MIR503HG
0,688612792	8,62898E-11	FAM107B
0,685985207	3,26315E-07	MINDY3
0,683926344	6,68564E-36	YWHAG
0,683372947	3,08438E-12	CRYBG1
0,681927263	2,33477E-17	APCDD1L

0,681606156	2,12096E-10	F3
0,680838943	2,60586E-10	NEGR1
0,679715821	4,82945E-06	RNF138
0,676962121	1,0771E-65	MMP2
0,676208305	3,25447E-09	DDIT4
0,672027748	1,18265E-06	ADGRG6
0,671126511	1,11433E-10	CNOT6
0,659059965	7,77091E-07	ADA
0,651328929	7,54662E-10	SHCBP1
0,650816644	4,37951E-30	KCNMA1
0,647236353	8,17828E-05	IFIT3
0,644137548	0,000173505	CD274
0,635623708	4,34267E-09	AVL9
0,624716833	9,61319E-15	CDC42EP1
0,624288825	0,000860814	EYA1
0,618199256	0,001372932	WNT7B
0,609774153	5,28119E-06	RBBP9
0,600510904	3,60687E-05	LHX8
0,600186052	2,94802E-11	NUS1
0,596797177	3,40651E-15	EIF1AX
0,595666477	4,04978E-07	ETS2
0,592311418	2,06906E-09	SUDS3
0,592128571	2,05192E-11	FBLN5
0,580463358	1,62168E-11	MT1E
0,576041077	7,77325E-23	IPO7
0,5663077	2,40738E-13	ECT2
0,566095453	3,09561E-06	TBC1D2
0,564067175	0,004786718	ABCC3
0,563264258	0,004010848	BCL2L11
0,561566437	1,61381E-14	AGRN
0,56126003	0,004130862	SYTL5
0,559917116	1,40184E-09	NAV2
0,559780748	1,61005E-15	CXCL12
0,559008247	5,13929E-11	TMEM181
0,558506754	4,31984E-12	GFRA1
0,554703651	2,89979E-11	TRIM25
0,550423877	0,000535381	INHBE
0,550274188	1,69479E-06	ASF1A
0,549526994	0,000239054	CASC10
0,54858937	0,004397074	TRIM6
0,545565228	1,25209E-20	ASPH
0,545525613	0,001483558	TMEM154

0,545138205	8,80194E-06	TOX
0,543837445	8,34654E-07	ST6GAL2
0,543362007	0,000980017	B3GNT5
0,542254283	6,34466E-17	SEMA3C
0,541551201	4,39622E-18	NRP2
0,53947069	0,002646881	LFNG
0,53614546	4,35389E-12	TMPO
0,53529813	0,00025397	SPOPL
0,534599729	0,006782888	DLEU2
0,533511806	3,22808E-07	FIBIN
0,524425156	0,007456627	BCHE
0,523629891	1,31354E-06	RAB12
0,523060412	9,19403E-06	NFIB
0,521168105	6,51169E-08	TCF4
0,520118976	0,007863644	SKIDA1
0,51692717	1,53581E-07	TAF10
0,516763051	0,009886773	C20orf96
0,513968345	0,000178192	MLX
0,511747316	3,07436E-09	REEP3
0,510267881	0,000429588	HCFC1R1
0,508547416	5,47727E-08	CLCN3
0,504615401	4,39105E-08	PHTF2
0,501914004	0,008754468	HLX
0,500690061	4,63506E-06	FOPNL

Table 34. postOSKM downregulated genes in shSETD2-1		
postOSKM shSETD2-1		
log2FoldChange	padj	gene name
-0,500254237	2,52386E-05	CMTR2
-0,501699991	0,001037914	L1CAM
-0,501789235	2,31988E-11	LONP2
-0,502125804	5,25905E-05	NOV
-0,502134613	1,16204E-19	MYLK
-0,502971358	7,70803E-07	RASSF8
-0,503879645	0,001298206	MARCH4
-0,504427318	1,73966E-05	APOE
-0,50454055	1,33892E-14	PTGIS
-0,505429575	6,78447E-11	GPNMB
-0,50640203	3,908E-24	VASN
-0,506760515	5,70288E-05	FGF5
-0,50705077	1,24213E-07	RECK
-0,507488051	1,71232E-09	NECTIN3
-0,507526238	1,57164E-18	KITLG
-0,508943753	1,24384E-13	DICER1
-0,509533727	8,76478E-05	HRH1
-0,510240285	0,000991095	FAM46B
-0,513425414	1,04187E-05	GAS2L1
-0,514485253	1,11888E-06	NLRP1
-0,514521776	0,000765187	AFAP1L1
-0,515213308	1,24337E-08	SLC8A1
-0,516119517	1,42887E-06	PRG4
-0,517176675	0,000362773	GRB14
-0,517648036	0,005510756	ZNF470
-0,51797276	4,51771E-05	ATP11A
-0,519636301	1,9061E-10	FLT1
-0,520347329	0,000270727	LAMA1
-0,520579487	1,77852E-05	GALNS
-0,520589182	5,6674E-05	FAM167A
-0,521420772	0,001002116	ADAM33
-0,523003825	0,003379941	CACNA1C
-0,523152062	5,77309E-05	STMN3
-0,525168246	0,001246674	TLL1
-0,526135043	1,33189E-37	ITGA1
-0,526944254	1,41446E-10	OLFML2A
-0,527002163	2,52122E-08	CYTL1
-0,527532073	0,006264082	HSPB7
-0,528216751	0,000273709	PACSIN3

-0,529056159	5,88272E-05	DPP7
-0,530371525	0,000291054	MTRNR2L1
-0,531822487	2,37169E-17	TBC1D2B
-0,532278327	2,6243E-06	NAGA
-0,53246953	0,00999639	LINC00116
-0,533118731	0,001075204	PGM2L1
-0,533424889	0,000332297	SYNGR1
-0,53524439	2,77402E-06	ABLIM3
-0,535716504	5,6717E-13	ITPK1
-0,535737699	3,01352E-14	KIF13A
-0,535826606	2,41465E-07	STK32B
-0,536631904	2,1824E-73	FBN2
-0,538589386	6,74163E-09	NABP1
-0,538876556	0,006287851	IL17D
-0,538888671	5,92158E-20	HMGA1
-0,539001486	0,007187259	SLC22A18
-0,540442463	0,001273504	NBL1
-0,5441147	1,13577E-12	TGFBR3
-0,544207319	0,000834692	FOXRED2
-0,54481501	5,7645E-13	SLC20A2
-0,545125874	7,28984E-13	PDE5A
-0,546037768	0,004190785	EML2
-0,546314742	1,76338E-05	ZDHHC20
-0,549014482	2,21825E-42	MAP1A
-0,549482334	0,001896278	PPM1M
-0,549702837	0,005070216	ZNF135
-0,549999675	0,001961313	HPSE
-0,55043186	6,33544E-05	CPAMD8
-0,550673875	5,04692E-48	COL4A1
-0,551764273	1,98186E-08	LTBP4
-0,552366675	0,000448275	SEMA3F
-0,553434441	1,36115E-13	FAT4
-0,55363484	9,24698E-05	PPARA
-0,553654058	4,11945E-05	ZNF528
-0,553991359	4,99323E-07	CSF1
-0,554026164	0,002964833	XKR8
-0,554217174	0,005599333	MROH1
-0,555131179	3,88643E-17	LAMA5
-0,556819462	0,00717711	DRP2
-0,557035214	4,37038E-15	LOXL3
-0,558337207	2,71285E-13	TSPAN4
-0,558479095	4,19514E-18	MEF2A

-0,559825051	4,4114E-12	LRRC58
-0,559965668	6,45465E-17	FAM3C
-0,560692955	2,8171E-26	CD248
-0,560739194	0,003693017	ZNF33B
-0,561577602	0,003180033	PPM1H
-0,561894113	0,001581142	CXCL8
-0,562506796	4,92493E-06	AK3
-0,562934713	0,001153644	CD83
-0,563213482	7,37025E-15	GRK5
-0,563269812	1,02565E-33	TMED10
-0,563967824	0,000640316	EFNB3
-0,564262579	0,002367339	DNAJC30
-0,565663169	0,005867047	ABHD11
-0,566141019	0,002465754	TMEM132D
-0,566295029	0,004535097	ERVK13-1
-0,567151441	0,007425034	FBXW9
-0,567573026	3,19753E-11	SYT11
-0,567779526	7,01906E-05	SYTL2
-0,567842098	0,007406085	TBC1D8
-0,568031645	0,00449634	PIK3CD
-0,568287132	1,89915E-18	PTX3
-0,569218098	0,000248026	APBA1
-0,569473621	3,89919E-08	RMND5A
-0,570124877	0,008064759	GGT5
-0,570674509	0,007414692	MOK
-0,571674778	2,81029E-13	DMPK
-0,571819704	0,001321604	ALDH4A1
-0,571943927	1,71134E-15	MICAL1
-0,572085905	0,000576363	SYNE3
-0,573350853	0,000703151	MAP3K5
-0,573583951	0,001158549	MEGF6
-0,574169753	0,007425034	THSD7A
-0,574608846	6,08528E-05	UPK1B
-0,574984117	1,30384E-30	COL5A3
-0,575151534	0,005617398	ZC4H2
-0,575163656	8,59658E-20	ABHD2
-0,575344395	0,004931696	MED11
-0,575415982	0,000107694	KIAA1644
-0,576824787	3,09264E-07	C11orf57
-0,577480502	5,52729E-22	EGFR
-0,579910273	0,000326674	LRRK1
-0,580202317	0,000138802	MGAT3

-0,580562328	1,97035E-08	RIN1
-0,581217663	1,4275E-23	ENDOD1
-0,582932616	0,002248032	CHCHD10
-0,583371903	0,000226253	MICB
-0,584160009	2,10928E-05	ADGRE5
-0,585063738	0,006331486	EEF1A2
-0,586150365	1,18357E-26	TMEM123
-0,586541302	3,13691E-07	CPEB2
-0,587170422	0,001403392	BARX1
-0,587477055	0,003914526	AL137003.2
-0,587529426	0,00305581	ZNF667-AS1
-0,587688138	3,31772E-16	SNTB2
-0,589327608	0,005325746	BHLHE41
-0,590660641	0,007885521	CEP19
-0,590879798	7,80831E-08	EPB41L1
-0,592060693	1,50907E-32	CYBRD1
-0,592434813	3,09678E-20	NACC2
-0,592889752	0,005510756	ERFE
-0,593022915	3,60982E-22	CITED2
-0,593149705	8,43929E-14	PDE3A
-0,593510894	2,88759E-13	ADAMTS7
-0,595949445	1,26155E-49	ATP2B4
-0,597083541	1,98238E-06	DPP4
-0,597591035	1,42995E-05	TMEM159
-0,599628653	4,79006E-12	CA12
-0,599843729	7,07034E-10	SYNPO2
-0,601270793	0,000707461	HCN2
-0,601889943	7,27763E-05	GPRC5A
-0,602399679	2,06857E-09	THSD4
-0,603496241	1,10348E-25	ITGBL1
-0,604094819	2,02648E-19	OLFML2B
-0,604495127	0,002112686	RASA4
-0,605869468	5,0992E-69	CAV1
-0,605923514	2,92653E-21	MGLL
-0,606052776	5,40671E-23	FBLN1
-0,606250607	0,00801051	FBXO2
-0,608727354	8,4005E-08	MATN2
-0,60980944	0,000219838	NHSL2
-0,611108242	0,000281478	PPARG
-0,611862216	0,003668243	ARHGAP6
-0,612520644	0,005705071	TGFBR3L
-0,614435068	0,007712574	ZNF497

-0,614710902	9,26696E-24	CNN1
-0,615571867	7,15455E-05	MTURN
-0,618456519	0,00119254	SNED1
-0,618522766	2,57887E-10	NPDC1
-0,619145704	9,30627E-05	DHRS1
-0,619273952	1,47898E-30	ABCC1
-0,620246573	5,26641E-09	TCTA
-0,62118737	0,005183936	AADAC
-0,62148926	6,19082E-05	FAM212B
-0,621622917	0,001947911	GALNT16
-0,621719341	0,007449628	ZNF471
-0,62265629	1,63955E-06	SAMD5
-0,623577735	0,002491187	KCNK2
-0,625034385	0,003845925	ZNF558
-0,625142022	0,000151495	WFDC1
-0,626258409	8,25963E-37	UACA
-0,626826887	6,18547E-08	IRS2
-0,629269423	0,002164182	SPATA13
-0,62934082	7,42808E-08	GNG11
-0,629859136	3,19208E-09	THBS3
-0,630087088	7,16143E-06	TUB
-0,630332408	0,000226253	MAPK11
-0,630870194	5,94576E-06	CCBE1
-0,633374431	0,001961313	KIFC2
-0,63581735	0,007736236	ADRB2
-0,635923211	0,009110582	PKP3
-0,636378467	0,003998211	ASPHD1
-0,63692463	1,69597E-05	MAMLD1
-0,638616951	3,16301E-12	PCGF5
-0,639847175	9,6818E-05	NPTXR
-0,640216732	4,56566E-07	ANPEP
-0,64075759	2,2285E-17	ATP8B1
-0,640803344	0,008361713	EBI3
-0,64226257	4,59761E-08	FOSL1
-0,643854358	0,008762873	TMCC1-AS1
-0,644011876	2,47436E-27	NEK7
-0,644158286	0,002720101	GLIS1
-0,648139069	0,004147842	BDKRB1
-0,648240622	0,00778165	SH3TC2
-0,648484412	0,000122928	SIGIRR
-0,648645743	0,002747251	KIRREL3
-0,650042157	3,64155E-21	SDC3

-0,650975671	0,001404462	WNT2B
-0,651469524	5,91493E-12	USP53
-0,652307695	0,001337867	FAS
-0,652847365	6,58504E-06	GREM1
-0,653139763	0,004047547	ZFHX2
-0,65392932	0,007677614	AGAP2-AS1
-0,654296733	7,18759E-19	PLXNA3
-0,654863241	2,55226E-10	HPCAL1
-0,655096913	0,008516513	DBNDD1
-0,655542451	0,002371705	AC091940.1
-0,658099686	0,003599694	SPINT1
-0,658774761	3,51351E-28	ABI3BP
-0,659402428	4,19028E-06	EFHD1
-0,660227847	2,83968E-19	LRP5
-0,660260162	3,87231E-05	ASIC1
-0,660831714	6,42782E-30	C12orf75
-0,661383929	0,005075223	VAMP1
-0,662032336	3,08994E-38	G3BP2
-0,662453086	0,003509892	ZNF415
-0,66266425	0,006759844	VWF
-0,663161365	7,40711E-12	C1orf21
-0,663499676	0,000186427	CREBRF
-0,66535728	0,004567639	CST6
-0,667587455	1,41663E-07	HCFC2
-0,667760275	0,005970074	P2RX5
-0,668193614	0,007744697	CPM
-0,668756272	3,81818E-05	TEF
-0,668787719	0,00063292	HAS3
-0,668976439	1,38206E-58	C1orf198
-0,66919966	0,000280773	CXCL6
-0,671767937	3,14178E-12	SRGN
-0,673792736	0,005990874	CNIH2
-0,674154371	0,001049534	LINC00342
-0,674665354	1,21661E-12	CLIP3
-0,675646273	0,006204562	ANGPTL4
-0,677345774	0,005679276	TMEM51
-0,677356339	1,47127E-06	CMBL
-0,67762461	0,001178902	MMP19
-0,677711794	1,4611E-07	HSPB8
-0,678980052	0,003998211	TMEM121
-0,679299475	2,34327E-43	NR2F2
-0,681664585	0,003661213	CPLX1

-0,681847775	5,086E-11	SEMA3D
-0,68195144	9,25403E-17	FLOT2
-0,684462674	1,33263E-21	COLEC12
-0,684500477	8,01057E-10	MEIS3
-0,685382232	0,00111001	TMEM59L
-0,685932967	2,65933E-16	BLOC1S6
-0,686547286	4,94714E-09	HAPLN3
-0,687118898	1,41886E-48	ITGA3
-0,687837305	0,001602519	IL6R
-0,688045382	1,10663E-13	ADCY3
-0,688297714	0,002885305	SHC4
-0,689322539	0,002593068	SLC16A14
-0,689795253	1,24698E-19	PSD3
-0,690636334	6,08154E-12	ZHX1
-0,690724793	2,23928E-32	DCAF7
-0,691105063	0,00108318	CSDC2
-0,696416167	1,01467E-05	APLP1
-0,69658396	0,004888289	NPW
-0,697449946	0,001404499	VAT1L
-0,698575948	0,00060393	SLC18A2
-0,702011966	0,002238641	TRPC6
-0,703521571	6,86382E-33	SPON2
-0,705587073	0,000346799	CLEC11A
-0,706172942	0,003180033	VSTM2L
-0,706183165	0,003032391	COL9A2
-0,706403741	0,004598981	MYEOV
-0,706808984	0,001281197	ZNF132
-0,709352492	0,003257627	AC021016.2
-0,717301822	1,63091E-30	EFEMP2
-0,719586733	0,000988356	HAGHL
-0,7222094	3,81173E-05	PALMD
-0,722588885	3,5262E-23	CD9
-0,725331429	0,003439484	AC090877.2
-0,727065376	1,70225E-52	ASAP1
-0,727484649	0,002267787	HOPX
-0,727631964	0,001542464	PLXNB3
-0,728128608	0,000195617	ZNF702P
-0,7299563	0,000372157	ARHGEF4
-0,730248572	2,67183E-24	OLFM2
-0,730672254	0,000552236	CACNG8
-0,732104898	0,002259071	AC009093.2
-0,733605681	4,25198E-39	ATP13A3

-0,733618277	2,93649E-88	IGFBP4
-0,733653257	4,92424E-08	FGF14
-0,735591255	6,50526E-09	UBR7
-0,735757795	0,000365703	CCDC136
-0,736836173	5,61233E-10	DNASE1L1
-0,737838161	0,002741605	MYPN
-0,738213937	0,000400474	KLHL41
-0,738468596	1,39898E-12	SEZ6L2
-0,738719813	9,43401E-13	SOCS5
-0,740966278	8,34436E-30	LMOD1
-0,74169125	3,73362E-95	CRIM1
-0,741897113	0,002402083	GNGT2
-0,742833793	4,10698E-08	SNTA1
-0,744619472	0,001201757	TRBC2
-0,744673191	0,000197578	SLC29A2
-0,745411268	2,90582E-10	SLC7A2
-0,747863938	0,001315498	CD52
-0,749049078	2,81444E-17	SLC9A3R2
-0,74992333	1,02533E-39	SETD7
-0,750453748	4,58306E-07	GAMT
-0,751578194	1,10974E-28	UQCRB
-0,752567807	3,54753E-06	ICAM5
-0,753350378	0,00023844	NCALD
-0,754238725	9,52926E-05	FICD
-0,754710893	1,0113E-07	KCNK3
-0,757238616	0,000647319	CHI3L1
-0,757518578	4,59761E-08	MAMDC2
-0,759336721	0,001892246	EPHB6
-0,761986134	3,90334E-79	ACTG2
-0,762844465	1,06616E-11	PTN
-0,763150136	0,001846274	RDH5
-0,763371525	1,16279E-37	SMAD3
-0,763551647	1,33008E-06	ABI3
-0,76545618	6,03604E-07	JADE2
-0,765860633	0,001437878	ADIRF-AS1
-0,766182342	9,799E-05	SBSN
-0,768251168	5,52677E-07	GSTM2
-0,76846689	2,17281E-32	LPAR1
-0,770073307	0,000146296	MAN1C1
-0,77345348	1,2734E-91	NID2
-0,77352107	1,96119E-06	UNC93B1
-0,774859927	4,13845E-06	SLC43A2

-0,776552401	0,00033656	CACNG7
-0,777693514	5,49123E-07	GPER1
-0,781550787	1,29782E-07	ASPN
-0,782715287	2,78548E-06	RCAN2
-0,782721012	2,71036E-06	SLC9A3-AS1
-0,783033241	6,93949E-27	CFL2
-0,784697464	0,001246132	GAL3ST4
-0,785208112	3,2235E-08	SGSH
-0,786730609	0,000240062	VSIG10L
-0,78831484	1,01375E-18	PRKAG2
-0,788838863	0,001174517	COL17A1
-0,789712973	2,38351E-08	SERPINB2
-0,790925547	2,42444E-13	RALGPS2
-0,792095701	2,12695E-16	SLIT2
-0,792690848	0,000294055	SQOR
-0,794372273	3,34031E-07	PODN
-0,795612001	2,80485E-05	COMP
-0,795993196	0,000270364	CPXM1
-0,796358559	4,58906E-05	NEURL1B
-0,798055871	0,000207594	THBD
-0,80156826	1,02643E-19	SMPD1
-0,803906257	7,12267E-06	PTGER4
-0,806270547	1,33779E-05	KCNJ12
-0,806316897	1,5808E-15	SCG2
-0,806970752	1,65521E-05	CDH18
-0,812230888	2,32026E-49	ARSJ
-0,813124404	0,000756687	RASD2
-0,814965437	1,45523E-06	MBNL1-AS1
-0,817958146	8,76478E-05	RASL11A
-0,822670289	1,57351E-81	MYO1B
-0,823099961	0,000631141	FENDRR
-0,827103359	0,000162613	C11orf87
-0,827289355	3,00259E-05	LDB3
-0,827825375	8,22764E-09	EFS
-0,830583612	7,24335E-18	SAMD8
-0,832895506	1,12338E-08	PPL
-0,836189821	2,69029E-10	CLEC3B
-0,836690986	0,000377955	ACSS1
-0,837537926	2,58991E-25	BTBD1
-0,837941098	9,1195E-142	THBS2
-0,841050561	8,94221E-05	IRAK2
-0,841691734	2,74599E-11	FMN2

-0,841942168	0,000197581	AOX1
-0,84195982	2,9719E-39	ADAMTS2
-0,84588688	0,000302997	AMPD3
-0,847580534	8,33428E-07	PDE1C
-0,849848163	1,87766E-14	ANGPTL2
-0,857660669	1,42685E-10	HHEX
-0,857944507	1,16459E-57	RGS4
-0,859642181	6,86708E-05	SAMD14
-0,859801721	3,37187E-05	NPAS1
-0,861202917	2,8539E-07	RAB3IL1
-0,864416534	1,08167E-10	INA
-0,864461846	0,000283493	MASP1
-0,864898329	4,81812E-31	TP53I11
-0,86690268	1,46772E-10	SERINC2
-0,869458792	3,64419E-13	METRN
-0,870157495	0,000176158	MARCH9
-0,871665614	0,000110607	PIANP
-0,875233924	9,28847E-47	PLPPR4
-0,878621853	3,95739E-07	SUSD2
-0,87970307	9,75691E-16	AK5
-0,887338343	3,79611E-19	CRISPLD2
-0,890705726	1,59352E-41	SCUBE3
-0,894406187	1,07062E-16	SGIP1
-0,894941634	1,03355E-81	MCAM
-0,896281015	4,45249E-22	LRIG3
-0,898500846	1,69527E-23	AHRR
-0,90237263	2,69247E-05	PTGER2
-0,905141505	1,57515E-09	HGF
-0,914862538	4,25874E-07	RAMP1
-0,926902373	8,55289E-26	STX2
-0,927150222	6,65963E-17	NCEH1
-0,929517847	3,70514E-15	BCAM
-0,931106141	5,01352E-05	CHRDL1
-0,931174651	3,82124E-09	CSGALNACT1
-0,931233655	2,59716E-06	TMEM130
-0,931827086	4,77021E-13	ICAM1
-0,933565038	4,85202E-10	DENND3
-0,934246164	2,49489E-12	NDN
-0,942543339	8,72333E-18	COBLL1
-0,943175445	4,11945E-05	MAPK8IP2
-0,943352356	3,10426E-05	OGN
-0,946563274	2,76981E-05	TMEM151A

-0,948165322	2,8996E-07	RBM24
-0,953431054	8,57865E-19	SUMF1
-0,953534162	3,78733E-49	WSB2
-0,955353184	1,06204E-06	ELOVL2
-0,958437702	1,92609E-25	OXTR
-0,959416492	4,85431E-07	PER3
-0,959420616	3,35604E-58	PLAT
-0,960360196	1,30446E-05	PAX8
-0,962157542	9,801E-139	FHL1
-0,963812453	8,26521E-06	FBXO41
-0,965383477	1,40742E-12	SYNJ1
-0,96767438	1,79646E-22	ITGA8
-0,97748232	2,76179E-05	SH2D5
-0,983146902	1,6027E-117	TNXB
-0,983301758	1,94398E-05	B4GALNT1
-0,98680675	1,65947E-10	STC1
-0,996854646	2,2285E-17	GAS1
-1,001184039	2,59071E-11	ALDH3B1
-1,007968684	1,41881E-05	TNFRSF1B
-1,011735683	2,65882E-44	MRVI1
-1,013284689	3,81746E-08	ZNF853
-1,019469876	1,08725E-10	SHROOM2
-1,036867376	6,39921E-13	ADM
-1,038293551	6,06114E-09	PRRT1
-1,040803408	1,25961E-15	TRHDE
-1,04295704	7,30753E-61	PLPP3
-1,044957398	6,15183E-08	GPX3
-1,04773361	8,38029E-14	DPYSL4
-1,050013116	3,33399E-44	CXCL12
-1,050461403	2,39427E-20	ZMAT3
-1,052434838	6,35244E-07	CACNA1A
-1,060885562	1,98518E-34	FAM43A
-1,068147159	9,89315E-76	CPA4
-1,073120032	9,07705E-28	CLDN11
-1,0735138	4,5317E-18	KLK14
-1,078171633	6,29684E-44	CPT1A
-1,08414169	5,35918E-07	SECTM1
-1,091349069	2,16109E-29	NPR3
-1,093074875	1,22678E-10	DIRAS1
-1,097135649	6,79502E-08	RAB40B
-1,10544466	1,23505E-47	MMP1
-1,116926657	3,48785E-31	MYOCD

-1,121439171	4,81034E-08	ANK1
-1,148256901	8,13472E-08	AP1S3
-1,156006453	1,8657E-103	SFRP1
-1,156977876	1,50337E-30	HSPB6
-1,161859537	1,79577E-38	CRHBP
-1,163241065	1,97901E-10	CAMK2N1
-1,164065597	2,48996E-20	FGD4
-1,173410073	1,13574E-10	SARM1
-1,182836906	6,12994E-39	CRYAB
-1,183944596	1,26923E-46	ITGA7
-1,187758041	2,25231E-10	TNFRSF11B
-1,193427987	2,80522E-21	KLK10
-1,196640239	2,34413E-14	OSR1
-1,211821315	7,22318E-08	VTN
-1,258295448	2,2842E-235	AKAP12
-1,262207585	2,41621E-10	CPNE7
-1,276764109	3,15845E-52	NPTX1
-1,278346859	1,121E-09	KRT27
-1,302311071	1,98263E-10	NOS3
-1,324564042	2,39308E-11	SLC12A7
-1,334982001	5,07765E-21	NDRG4
-1,368340584	1,08341E-26	TYMS
-1,406527523	7,03353E-25	ACAN
-1,408379283	4,0418E-19	TINAGL1
-1,433781808	1,86812E-79	SETD2
-1,435502386	1,19707E-57	RGS5
-1,449673907	9,07023E-73	CNTNAP1
-1,461862582	3,72939E-48	TFPI2
-1,526077594	1,5847E-15	SLC2A12
-1,615347679	3,80234E-19	NR4A3
-1,625177016	3,24921E-14	CHRM2
-1,687634533	1,62528E-32	MRGPRF
-1,699931206	2,40322E-32	PTGS1
-1,759233201	1,15785E-41	EPHA5
-1,978796744	2,1799E-99	MYH11

Table 35. postOSKM downregulated genes in shSETD2-3		
postOSKM shSETD2-3		
log2FoldChange	padj	gene name
-0,50050253	7,63333E-14	PDE5A
-0,500610544	3,68608E-07	CRISPLD2
-0,50119243	5,51739E-06	ARMCX4
-0,502363725	0,000224357	ZNF154
-0,503029178	0,000127014	PIEZO2
-0,503175092	7,71619E-25	KDELR3
-0,504738056	1,91165E-05	MAN2A2
-0,505411276	0,001282009	CHST10
-0,505506409	1,68834E-07	CLDN11
-0,505710316	6,48423E-27	CYB5R3
-0,506386171	1,72462E-46	TAGLN
-0,506907198	0,000348718	MAP3K12
-0,508253558	1,1943E-24	TNKS2
-0,508964671	0,001953036	CXCL1
-0,509633909	0,004712727	ZNF641
-0,510095497	0,000451203	JADE2
-0,510364255	2,96288E-15	ANKH
-0,510903402	2,35086E-16	SORBS3
-0,510988169	7,79351E-32	MYO1C
-0,511632513	8,71303E-11	CNST
-0,512163968	0,00399605	BCL2
-0,512599574	0,001361526	RBM24
-0,514600546	1,10706E-06	ADCY9
-0,515486848	5,55624E-06	HOMER3
-0,516142387	2,61575E-11	ZNFX1
-0,516327155	0,0009608	FITM2
-0,516896061	8,5295E-16	ADCY6
-0,518525602	0,005248556	BEX1
-0,518693028	0,000248513	ACADSB
-0,520559331	8,31119E-07	MEGF9
-0,521092453	9,78687E-18	ROCK2
-0,521518646	2,0023E-105	IGFBP7
-0,52264954	2,51351E-58	ATP2B4
-0,522745236	0,000233903	SORD
-0,523096217	0,003126487	GPER1
-0,523365265	0,008874377	MROH1
-0,523879512	0,00072424	NLRC5
-0,524046181	4,06954E-07	TRPM4
-0,525248508	4,22303E-19	THBS1

-0,52755636	0,000248692	PITPNM3
-0,527558798	0,002196416	HCN2
-0,527786428	2,28674E-14	TGFBR3
-0,52880291	7,69088E-07	CRAT
-0,53030339	1,11485E-08	HSPB6
-0,530751583	0,003535324	PRDM16
-0,531420098	0,005211936	TRIM46
-0,531794817	8,05741E-20	NUCB2
-0,534430007	2,99365E-13	IRS1
-0,535556479	0,004313248	STK32C
-0,535684188	1,11398E-06	MARCH8
-0,537090231	4,34112E-05	DIS3L
-0,537132206	0,000772738	ALDH6A1
-0,539469601	5,92629E-05	DKK1
-0,539926266	0,003078404	MCRIP2
-0,539981411	0,006530524	SBSN
-0,540397875	0,002190152	MOB3C
-0,540899007	1,42322E-29	LUM
-0,541043668	4,82717E-19	KIF13A
-0,541116163	1,29345E-21	LITAF
-0,541367279	3,06057E-05	NBEA
-0,541439061	0,001391789	SLC4A8
-0,541702424	8,75407E-25	ZNF469
-0,541854128	1,59681E-10	C16orf45
-0,541924122	0,001265505	NRM
-0,541997623	0,000486591	SLC25A30
-0,542501889	4,08303E-05	IL1RAP
-0,543966574	4,68047E-06	RNF141
-0,544730594	0,00205798	TMEM106A
-0,545011056	5,17646E-13	FARP2
-0,545638747	0,001313267	PHF11
-0,546042291	0,000363359	PPP1R26
-0,546136621	2,10171E-07	NCEH1
-0,547358472	0,004622781	ELOVL2
-0,547844229	2,34378E-54	ITGA1
-0,54785615	0,003933251	NGF
-0,547901826	4,7088E-06	FAM213B
-0,548285936	1,32666E-09	COPZ2
-0,548502679	7,04762E-06	PDGFRL
-0,551046113	2,57552E-11	APOBEC3C
-0,55155542	3,03219E-33	VASN
-0,552368095	0,002017199	ELFN1

-0,552758785	1,08043E-05	LAMA1
-0,55286224	0,007770225	PARD6G
-0,553897452	4,49242E-43	NQO1
-0,554431809	5,21151E-08	SLIT2
-0,554637716	6,99604E-37	PCOLCE
-0,554667213	0,001353552	PPARG
-0,557664112	1,16721E-09	ZBTB47
-0,557773041	1,93712E-07	ICAM1
-0,558106496	0,002418302	WDR76
-0,558713951	5,71176E-20	TWSG1
-0,559014325	5,44861E-06	FOXL1
-0,559321081	0,004393767	TMCO4
-0,559797775	2,31401E-50	NID2
-0,560244262	8,76655E-11	CAT
-0,560892462	0,000455981	RAD51
-0,564530291	1,11992E-22	IDH2
-0,566666865	9,16809E-05	NACAD
-0,567199722	1,1943E-24	KITLG
-0,567531204	1,05478E-70	UACA
-0,567534992	0,007993901	BHLHE41
-0,567904817	0,000319428	RBFA
-0,568110054	0,000173131	SP110
-0,56843387	1,32823E-14	TCIRG1
-0,569335894	3,18243E-12	PCSK9
-0,569917084	2,84371E-36	EXT1
-0,570052602	1,11953E-11	CYBA
-0,570519192	0,003165341	KCNE4
-0,571324984	2,28777E-06	PLEKHA4
-0,571657906	0,007766747	GLIS1
-0,571672508	0,002023153	TRDMT1
-0,571795475	0,000150595	ALDH4A1
-0,571861528	0,000349413	SLC9A3-AS1
-0,572004483	0,004030313	ZNF667-AS1
-0,572451016	0,000880457	ARHGEF25
-0,572867675	8,4166E-25	CNN1
-0,573629233	0,000306141	MAP3K5
-0,57429812	0,002906622	PODXL2
-0,575372328	2,22056E-10	SMARCD3
-0,575756452	1,75067E-21	GPSM1
-0,576117957	4,87368E-05	JPH2
-0,577028951	3,28299E-07	DNASE1L1
-0,579447795	0,001759387	FUCA1

-0,580715749	8,71964E-06	SYTL2
-0,581238972	0,00059273	DBF4B
-0,582257376	6,89394E-67	CEMIP
-0,582931517	0,000109352	DHRS1
-0,583050605	3,22308E-13	MAP2K3
-0,583198995	4,53257E-06	PPFIBP2
-0,58380698	0,007259924	CYP4V2
-0,586796055	4,68589E-06	EBF3
-0,587793873	0,001393108	ZIK1
-0,588721907	7,47847E-15	NLGN2
-0,589608914	8,01956E-07	HACD1
-0,590972673	5,59565E-21	LOXL3
-0,591843316	9,51988E-05	RAC3
-0,592039253	0,002118108	ZNF793
-0,592686246	0,000114365	LRRK1
-0,593699985	4,14395E-08	B3GNT9
-0,594061901	6,76336E-16	HS6ST1
-0,595279367	3,00171E-11	SCARF2
-0,597970677	0,003691149	EIF2AK2
-0,598082708	2,59705E-09	ZNF503
-0,598654245	6,2925E-05	ABCB9
-0,599945765	3,07458E-20	TGFB1I1
-0,599985079	0,002795235	GK
-0,601537464	7,10668E-31	EFEMP2
-0,601816355	7,3556E-28	CAPN1
-0,602449952	5,79969E-08	FAM57A
-0,602493846	2,70568E-05	NLRX1
-0,603552891	0,001584195	COL13A1
-0,603751157	1,60352E-22	GPNMB
-0,605710156	8,46111E-23	TMEM119
-0,60714426	5,41713E-22	NACC2
-0,607312714	2,58278E-05	ADGRB2
-0,607878699	8,73294E-75	IGFBP4
-0,60797357	1,82906E-06	PPL
-0,608381085	2,59163E-20	NEXN
-0,60897541	5,81631E-21	ZBTB4
-0,609240899	2,58321E-12	ADCY3
-0,609396054	1,59037E-05	PACSIN3
-0,609592885	0,000344724	SLC1A1
-0,610157427	0,000415435	PPM1H
-0,610188147	4,13424E-05	PRUNE2
-0,61219067	3,94674E-09	DCHS1

-0,612305647	0,007072972	ICA1L
-0,612841516	0,009426774	MUC5AC
-0,616675443	8,45245E-05	BDNF
-0,617900118	0,001166313	LINC00116
-0,61843498	0,00626832	SHC2
-0,618543861	1,02192E-17	ERBB2
-0,62092119	0,009632775	BAALC
-0,622066819	0,000240811	SLC43A2
-0,62445512	1,06865E-08	STXBP6
-0,627204695	2,85096E-05	MSX2
-0,627538177	0,007325323	TMEM35A
-0,629302192	0,001745316	TMEM62
-0,633184425	3,97554E-09	ANPEP
-0,633222427	2,408E-20	ECM1
-0,633928567	3,01687E-11	THSD4
-0,634072137	5,85296E-05	TEF
-0,634544269	0,007609831	WSCD1
-0,634585685	4,44307E-09	MEIS3
-0,635009515	3,69689E-68	FAT1
-0,636780891	1,29099E-05	FOXF1
-0,637152993	3,45979E-52	CD248
-0,637783171	0,002508196	TAP1
-0,63880989	0,00017621	LGALS3BP
-0,639316072	3,64371E-27	MICAL1
-0,639872303	7,7284E-21	ARL2BP
-0,640552413	7,8798E-08	BLVRA
-0,640836589	0,003060458	THSD7A
-0,641114674	6,33141E-20	TSPAN4
-0,641668048	0,001085591	ITGA10
-0,642064303	4,28726E-09	FOSL1
-0,64232106	0,002201208	AC145124.1
-0,64401538	2,00227E-70	MAP1A
-0,644104794	1,54755E-21	RASSF8
-0,64468748	1,40652E-36	SMAD3
-0,645080842	0,001770751	CBR3
-0,645229522	2,77697E-27	SCUBE3
-0,646525343	3,89623E-10	BRIP1
-0,646592198	3,97874E-21	PLXNA3
-0,646694101	3,28528E-52	EMILIN1
-0,646952335	5,98434E-13	SYNPO2
-0,647671955	9,92179E-28	SORT1
-0,648558671	0,008283206	ANKRD36B

-0,64881789	0,002628809	ZFHX2
-0,649534195	0,004727136	DPF3
-0,65001944	1,46144E-05	ADAM33
-0,651960989	1,10569E-06	VIT
-0,652802978	7,36351E-09	ACOX3
-0,654186934	7,98726E-06	ATP6V0E2
-0,655954982	0,001416414	IL17D
-0,65649068	0,001835414	ADAMTSL5
-0,656515179	9,15925E-31	RCN3
-0,656751225	1,51308E-06	PDE4A
-0,658440741	2,06051E-06	ZSCAN18
-0,659756589	6,16147E-19	FLOT2
-0,659991725	1,36356E-10	NAGA
-0,661527928	0,001437011	ARHGAP6
-0,661528948	1,85638E-13	C1orf21
-0,662596078	0,000820947	PRKAR2B
-0,66408026	0,001365302	MAST3
-0,664286691	7,14415E-38	C12orf75
-0,665078114	0,007768011	DENND2C
-0,666583617	6,55708E-11	PAX9
-0,667924896	0,006434366	ARHGAP22
-0,669164211	1,90899E-46	CYBRD1
-0,669628846	1,778E-144	FBN2
-0,669783189	1,07242E-05	MTURN
-0,67251322	3,32466E-98	FHL1
-0,673466985	3,16765E-05	TMEM25
-0,674490362	2,39981E-14	CLIP3
-0,675269288	6,11379E-23	AFF3
-0,677145078	0,000226952	GALNT16
-0,677203268	1,11953E-11	FAM46A
-0,677306033	2,44579E-22	ADAMTS7
-0,678104273	1,61128E-07	PPARA
-0,679052424	2,16159E-24	ITGBL1
-0,679131785	8,62921E-05	TDRD7
-0,679818191	2,78648E-91	MYL9
-0,680247015	0,000306621	RAB36
-0,682089458	0,000172524	FAS
-0,682316189	7,56103E-10	FOXF2
-0,68234818	0,00011308	LIN7A
-0,684072081	0,000838449	STAT1
-0,685238578	0,000558709	KCNK2
-0,686214241	7,68431E-06	FOXRED2

-0,686672494	6,52991E-05	SLC22A18
-0,68969963	1,85282E-06	SERPINB2
-0,691311674	0,007391386	GDF5
-0,691803224	0,006325166	TGFBR3L
-0,692324293	0,001113051	CSDC2
-0,693409934	1,1651E-07	SLC2A6
-0,69567023	0,007892172	MASP1
-0,696356734	1,38222E-06	ATP7B
-0,697511617	0,003758721	FBXO41
-0,698532874	0,000566127	SELENBP1
-0,69917195	0,001389138	CERS4
-0,699522964	0,000425449	TBC1D8
-0,701225951	1,30063E-27	SPOCD1
-0,701466495	6,65443E-87	COL14A1
-0,702174321	0,003448304	OPLAH
-0,7022693	0,009540109	PPP1R3A
-0,703205484	1,06468E-22	MMP1
-0,70333742	0,000489837	RAB40B
-0,704725443	0,007239871	NR5A2
-0,705321443	0,006944836	SH2D5
-0,707255181	1,31896E-07	RILPL2
-0,709822	0,005950196	AC008770.1
-0,710494025	2,19228E-09	L1CAM
-0,710915552	0,008963832	LINC00854
-0,711039177	7,00792E-06	RAB3IL1
-0,711968671	5,08426E-05	SAMD14
-0,714361878	0,006454568	MARVELD2
-0,714463011	9,85968E-09	SGSH
-0,714675323	0,003732913	MUC1
-0,716292447	0,006030507	PTCH2
-0,718252243	1,25235E-27	SNTB2
-0,718341861	3,7639E-16	SRGN
-0,718482737	0,006113019	SH3TC2
-0,721128291	3,70592E-08	FGF14
-0,722964679	0,002078138	SCN1B
-0,723235053	5,37662E-17	TP53I11
-0,723536594	4,30462E-05	CADM1
-0,723892024	9,39678E-60	MCAM
-0,725736534	0,000212837	FSD1
-0,727457428	0,000731021	PAX8
-0,728927526	6,39406E-12	SLC22A17
-0,733601781	2,7613E-06	STARD9

-0,733768286	5,30259E-16	CPEB2
-0,735524677	2,96488E-98	ACTG2
-0,738928842	0,004167038	A1BG-AS1
-0,741804416	0,000585486	AGT
-0,742728179	1,66579E-58	COL5A3
-0,742969415	0,000592335	NR2F2-AS1
-0,744028041	2,16054E-19	COMMD7
-0,744595273	1,37277E-10	CDC14B
-0,745141821	2,71832E-05	BARX1
-0,747755356	3,61647E-08	ALDH3B1
-0,748440298	0,002489811	PCDHB10
-0,748493054	0,002866123	RGS11
-0,749768713	0,000156066	ZNF667
-0,75017794	0,004047723	AP001107.5
-0,750726846	0,005032544	DLX3
-0,75161017	6,25998E-35	PTX3
-0,752794527	0,004831095	EEF1A1P19
-0,753530812	0,00010965	SLC12A7
-0,753878479	0,000804795	HCN1
-0,75471502	0,000265741	SLC29A2
-0,757959155	0,003526611	MARCH9
-0,758233246	1,40803E-09	TUB
-0,758338111	8,47691E-09	PREX1
-0,758353715	9,31615E-10	SAMD5
-0,75835415	6,94943E-10	FAM111A
-0,758806886	2,00648E-49	PDGFRA
-0,759185635	0,00072453	VSIG10L
-0,760785001	0,002051683	AGAP2-AS1
-0,761126777	4,02467E-05	SLC43A1
-0,761236456	1,1438E-27	SYT11
-0,761698141	7,3538E-101	C1orf198
-0,762021779	0,000767607	PLXNB3
-0,766408683	6,24667E-05	APC2
-0,767013396	0,000804795	ENOX1
-0,767250455	0,002903634	PHC1P1
-0,76932311	1,18222E-19	USP53
-0,769962349	0,002481063	HTR2B
-0,773522444	4,83251E-05	KIAA1456
-0,774741936	9,35202E-31	SMPD1
-0,775902329	0,000197744	ZNF132
-0,776789452	3,31614E-05	EGFL7
-0,779035333	0,000149761	PSEN2

-0,779664093	8,88913E-65	MYLK
-0,781350662	0,00336742	EBI3
-0,781530138	7,68719E-10	NPTXR
-0,787187118	9,81592E-08	GSTM2
-0,787249325	2,49255E-47	FBLN1
-0,7888994	8,33444E-93	RGS4
-0,789192561	1,42828E-05	CAMK2N1
-0,791433529	0,001641024	PLN
-0,794406339	2,83975E-13	HAPLN3
-0,7945063	4,816E-28	FST
-0,795229805	8,23703E-06	PRR16
-0,798693398	0,000758875	ANKRD18B
-0,799346485	0,000429308	FUOM
-0,800589248	1,65645E-07	TLL1
-0,805495754	1,97818E-05	ZNF702P
-0,806206899	2,62965E-09	GRB14
-0,80749534	0,002277005	ITIH3
-0,808713596	1,98207E-19	ADAMTS5
-0,809315759	7,85977E-05	ZNF471
-0,809331372	8,80402E-14	SGIP1
-0,811132699	0,001362003	NYAP1
-0,813340577	2,2095E-16	TCF7L1
-0,815909305	8,73034E-11	LRP3
-0,817681767	0,000838449	CAMK1
-0,819300173	3,40687E-06	CYB561
-0,819474476	0,000252023	NUDT14
-0,819743836	1,81221E-56	SDC3
-0,820214958	1,3766E-141	CRIM1
-0,820292094	0,001745316	USP18
-0,823422441	5,63232E-37	FAT4
-0,824809074	1,9844E-170	ACTA2
-0,826112837	8,19269E-05	PTGER2
-0,829727708	0,000696549	MX2
-0,83289297	2,36244E-06	PER3
-0,832952269	3,55048E-07	ATOH8
-0,834806004	0,000297856	ZNF497
-0,837612924	8,28711E-93	TNXB
-0,840236292	0,000203645	MYH3
-0,841927401	0,000269641	B4GALNT1
-0,84369154	3,10347E-06	PKN3
-0,846961692	0,00125268	PLPPR3
-0,847635847	6,72212E-05	ZNF558

-0,852254719	0,001139698	FAM83H
-0,854286635	9,30488E-10	CHST15
-0,857088921	4,57172E-16	NOV
-0,861642971	6,89686E-31	PDE3A
-0,864834008	0,000719234	FBXL22
-0,86676423	1,31444E-14	GNG11
-0,867321476	7,1065E-06	MAN1C1
-0,8699881	7,25056E-10	GREM1
-0,870538755	2,49916E-13	DENND3
-0,872567692	1,5385E-10	SHROOM2
-0,87306993	3,97454E-15	PLPP1
-0,875420187	0,000642728	CTSF
-0,876836475	7,47335E-18	COBLL1
-0,882541495	8,33508E-08	MEGF6
-0,887695984	0,000294642	OASL
-0,888059159	1,05194E-09	RIMS1
-0,891549958	5,57562E-34	SHISA2
-0,891795894	1,75183E-09	FABP3
-0,8922792	3,22513E-10	MICB
-0,893918107	1,5994E-15	SEZ6L2
-0,898479408	2,35069E-28	SLC8A1
-0,900761211	5,05437E-10	UNC93B1
-0,901243939	0,000152207	SMIM10
-0,901763861	2,20858E-93	ABI3BP
-0,901887105	0,000174722	FAM196A
-0,902555552	8,1011E-08	PRRT1
-0,90820215	0,000421465	ACP5
-0,908831228	1,02956E-13	GPC3
-0,909163724	6,23806E-68	DCN
-0,915870007	8,62612E-09	GMPR
-0,91608817	5,59501E-28	RALGPS2
-0,916918975	0,000371608	ADRB2
-0,917177768	7,20492E-15	KLK10
-0,917792053	9,18687E-37	NPTX1
-0,918287623	2,33296E-10	APBA1
-0,920833627	1,59568E-15	METRN
-0,92280257	1,20232E-30	CRYAB
-0,926028641	1,65879E-15	BCAM
-0,926346736	3,33074E-08	MBNL1-AS1
-0,931652644	1,56287E-42	NECTIN3
-0,936313038	2,61491E-26	PLPPR2
-0,936505781	2,17575E-24	NPR3

-0,941846774	0,000207109	PYCARD
-0,942168404	1,8037E-11	SYNE3
-0,943875792	0,000166905	CCDC190
-0,94438081	4,73977E-25	PCGF5
-0,944760215	5,28916E-14	NDN
-0,945193915	6,14344E-96	MYO1B
-0,947894539	0,00010142	CEND1
-0,947940523	8,7156E-08	HSPB7
-0,948069883	1,6406E-11	DPYSL4
-0,961160166	2,67751E-05	FBXO6
-0,963579321	2,12062E-19	ANGPTL2
-0,968240203	0,000152088	AOC3
-0,971192483	6,27839E-07	TMEM59L
-0,974017544	5,48485E-05	FBXL16
-0,976055047	2,59296E-16	ADM
-0,981796433	2,59296E-16	FGD4
-0,986631178	1,59797E-64	OLFM2
-0,988973051	5,21514E-18	GAS1
-0,989391489	1,87776E-06	ASPHD1
-0,997356336	1,98378E-52	LPAR1
-0,997813053	1,10521E-39	DMPK
-0,9983101	1,62072E-21	PTN
-0,999179032	6,5686E-08	VAT1L
-1,0041442	3,27744E-05	CHRDL1
-1,007598359	8,43596E-20	DPP7
-1,008923157	6,25052E-08	ZNF470
-1,01309402	5,51038E-26	DSEL
-1,013265788	1,20051E-42	CPT1A
-1,020622984	4,52811E-11	RAMP1
-1,021683838	1,94812E-08	ZNF135
-1,021794148	5,39789E-05	SHANK1
-1,022125629	1,46109E-33	AHRR
-1,023392216	2,49991E-21	IRS2
-1,024318369	7,2016E-20	CLEC3B
-1,024755618	1,62386E-08	ZNF853
-1,025656354	3,04437E-16	FMN2
-1,026701258	4,3787E-05	GAL3ST4
-1,028125969	7,21346E-17	SNTA1
-1,029755424	5,11697E-07	NPAS1
-1,032187423	3,55274E-28	MYOCD
-1,03779592	1,00378E-15	KCNQ5
-1,038888201	4,21348E-06	PIANP

-1,039080426	6,85978E-13	RCAN2
-1,039317822	3,13889E-05	CHI3L1
-1,04284137	1,10127E-05	MAPK8IP2
-1,047220582	2,1205E-60	LRP5
-1,051651392	6,77591E-07	FNDC10
-1,052519973	8,42865E-68	PLPPR4
-1,060639304	1,89577E-86	ARSJ
-1,067291369	2,90509E-11	EFNB3
-1,067900481	1,21633E-07	CACNA1A
-1,06938776	4,15372E-17	SERINC2
-1,077672631	1,11745E-19	FZD8
-1,091401117	1,88212E-08	CPNE7
-1,094365749	3,25756E-08	GPX3
-1,095924583	4,11505E-09	CLEC11A
-1,096492734	3,53919E-27	SCG2
-1,101347341	1,10407E-29	ITGA8
-1,121215219	3,0193E-58	MRVI1
-1,127495998	5,2083E-08	HAGHL
-1,127631021	5,53875E-06	RASD2
-1,130051766	6,96382E-07	SHC4
-1,135397857	3,27277E-24	SLC7A2
-1,13893335	2,75351E-40	FAM43A
-1,162832767	7,25909E-11	ZFP28
-1,169074176	3,59276E-07	KCND3
-1,169504142	1,29398E-83	CITED2
-1,175897533	1,75161E-09	LDB3
-1,180675742	2,76628E-16	MAMLD1
-1,181451253	8,05375E-31	PTGS1
-1,18162943	7,17094E-10	FGFBP3
-1,187934755	1,56792E-08	SLC16A14
-1,188830863	6,12257E-08	C11orf87
-1,195721618	1,88525E-10	HAS3
-1,215157565	9,16609E-11	SLC2A12
-1,222327247	1,88543E-14	WFDC1
-1,249792366	9,08516E-76	LMOD1
-1,250793788	1,13316E-62	ITGA7
-1,251205571	3,42196E-24	NLGN1
-1,284242661	1,16157E-07	TNFRSF1B
-1,288361221	2,24415E-22	EFS
-1,295209625	9,3522E-65	CNTNAP1
-1,296308257	4,89907E-08	OGN
-1,311792949	1,28193E-18	ICAM5

-1,311977176	7,24335E-27	HHEX
-1,312862379	2,59296E-16	TINAGL1
-1,319062054	7,64252E-11	SNED1
-1,323339025	1,11953E-11	TMEM130
-1,323680452	3,8991E-23	NDRG4
-1,32529106	9,89914E-17	NRXN3
-1,342358481	7,48628E-10	CACNG7
-1,350248373	2,97375E-08	SLITRK6
-1,357801481	4,02513E-09	TMEM151A
-1,366788925	4,31415E-13	SARM1
-1,369482759	1,22643E-20	EFHD1
-1,376707101	5,99585E-24	INA
-1,378376264	5,49386E-22	CSGALNACT1
-1,380205958	6,34176E-41	SEMA3D
-1,396449402	6,79405E-13	TNFRSF11B
-1,414693181	2,15052E-27	MRGPRF
-1,422127346	3,93009E-09	FENDRR
-1,479136223	1,72391E-13	NEURL1B
-1,481883929	6,92707E-19	PTGER4
-1,486059278	6,76142E-22	APLP1
-1,486548027	1,79353E-63	CRHBP
-1,489207144	1,09373E-23	KCNK3
-1,515461269	6,2755E-110	PLPP3
-1,53970144	4,46868E-21	DIRAS1
-1,543199937	5,4663E-24	GREM2
-1,552749629	3,72392E-15	REEP2
-1,560472676	4,38725E-32	PDE4B
-1,593801948	9,56638E-16	CDH18
-1,605740834	2,33246E-15	ROBO2
-1,615967608	2,93938E-24	OSR1
-1,62962099	8,72272E-12	COLEC10
-1,659914571	1,59781E-77	RGS5
-1,674193733	1,3359E-47	AK5
-1,683863487	5,15849E-14	SECTM1
-1,684219024	1,2928E-121	SETD2
-1,697438198	3,22863E-13	VTN
-1,836025129	5,16103E-38	ACAN
-1,892058423	0	SFRP1
-1,922493056	7,60582E-65	PRG4
-1,938552295	1,46688E-17	CHRM2
-2,447093642	3,0775E-68	EPHA5
-2,612471987	1,06257E-45	NR4A3

-2,835942794	2,6475E-212	MYH11

Table 36. postOSKM upregulated genes in shSETD2-1		
postOSKM shSETD2-1		
log2FoldChange	padj	gene name
2,627345017	1,37598E-71	LEFTY2
1,862318744	1,26616E-27	RBP1
1,806350411	4,1025E-32	ARL6IP1
1,75989302	1,50295E-52	INHBE
1,667197462	3,802E-121	HAPLN1
1,552440218	3,1004E-173	KRT16
1,490545699	2,60099E-23	EPHA3
1,470419159	1,08148E-33	FOXS1
1,447147492	1,82947E-28	G0S2
1,398794136	8,43794E-31	ISLR
1,275262943	6,00527E-11	KRT5
1,250452412	3,08011E-24	PMEPA1
1,243763387	2,86445E-08	PCDH17
1,234677177	2,11786E-22	PTPRB
1,225271186	2,89075E-10	MFAP3L
1,219122358	1,96572E-15	MECOM
1,218226346	5,58997E-08	CHRNA9
1,207864679	4,51149E-10	EMCN
1,190830449	3,74086E-11	GALNT3
1,182728184	1,49257E-33	KRT17
1,167404701	1,11562E-29	COL15A1
1,162983946	8,0398E-11	SUCNR1
1,148957305	4,03631E-07	WNK2
1,131611853	7,63395E-10	GALNT12
1,118035853	5,95668E-23	SERPINB9
1,11444814	4,19629E-79	CDKN2B
1,106941517	8,00084E-07	LEFTY1
1,104586929	1,13282E-06	TMEM155
1,101832485	2,2299E-36	FIBIN
1,096893144	8,43632E-11	GUCY1A3
1,092693014	1,81601E-06	COL2A1
1,072411911	3,67725E-57	PYCR1
1,061596653	3,05371E-14	ARRDC4
1,050486475	9,79885E-09	IQGAP2
1,05006927	4,25162E-06	CACNA2D2
1,038007227	1,37157E-57	UNC5B
1,036338484	8,4511E-12	CTSO
1,030168989	4,79235E-06	LGALS12
1,024787771	4,74327E-39	MAN2A1

1,024684811	2,59465E-16	SLC39A8
1,01452677	3,41549E-22	PCDH10
1,014228939	9,6077E-06	RNY4
1,014175496	2,66708E-36	ST6GAL2
1,002556403	4,24401E-06	SYT6
1,001498894	2,32209E-83	RAB1A
0,997619256	1,50387E-21	CLCN5
0,997072318	1,79662E-05	LIN28A
0,99280309	1,16382E-15	HHIP
0,979530074	2,50822E-05	RNA5S9
0,974710041	9,65001E-07	ACSS3
0,970466198	3,27755E-05	IGDCC3
0,957776965	1,23757E-10	PDE4D
0,950992709	4,75509E-07	CLSTN2
0,936707889	2,00148E-08	MDFI
0,936512434	3,38638E-05	JAKMIP2
0,932043425	2,2299E-36	PCSK9
0,926000358	7,17318E-06	AP003119.3
0,921933003	6,91868E-05	ISLR2
0,917796939	2,40969E-07	LYPD1
0,908642454	3,09052E-08	SIX2
0,908226595	2,3368E-07	LIPG
0,902788631	1,14121E-07	L1TD1
0,901007621	2,20711E-05	SALL4
0,899709979	0,000143776	DSG3
0,899071115	3,83898E-05	ASXL3
0,898951015	9,98608E-10	KIF26B
0,898559534	1,30101E-05	EMB
0,891728546	1,2425E-18	ST6GAL1
0,888646251	1,2594E-146	VCAN
0,877010353	6,46589E-06	COL25A1
0,875643352	2,4727E-09	SEMA6A
0,87382642	1,29782E-07	BMP6
0,873722839	5,23864E-06	VASH2
0,863550655	5,68626E-05	FREM2
0,857637175	3,0422E-12	CD36
0,855869972	0,00011626	SOX21
0,851922592	1,05495E-05	PLPP4
0,85054191	4,94279E-25	DSG2
0,84738377	2,46662E-14	ACTC1
0,845180339	3,20871E-31	DDIT4
0,844168991	0,000273709	AL591845.1

0,839354927	2,76179E-05	ADGRV1
0,834593904	0,00024232	APLNR
0,833509011	5,08056E-17	LARS2
0,832749919	4,45197E-06	TLR3
0,828610735	0,000167216	DHRS9
0,825172352	2,2994E-10	SIK1
0,824841827	2,65169E-05	SAMD11
0,82423402	0,000619531	ENPP4
0,818995986	6,99807E-15	HAS2
0,816324985	5,25182E-12	CBX5
0,814834654	5,21709E-11	ANXA3
0,812814355	7,1638E-19	PRR5L
0,811317865	0,000627388	SNORD3A
0,808769639	5,90688E-05	CLGN
0,803739432	2,30789E-42	NNMT
0,80295646	2,06886E-08	CHD7
0,800739015	2,20711E-05	PLA2G16
0,798658319	2,52167E-05	TRIM71
0,794728692	5,64899E-06	BCHE
0,793309046	2,14909E-09	CSRNP1
0,793169679	2,1783E-15	TMEM41B
0,792071442	9,34363E-06	STMN2
0,790975217	0,000608532	RNU4-1
0,790969922	6,66056E-21	PCDH1
0,786054167	0,001225418	SNORA73B
0,784916447	2,60074E-16	SLC44A1
0,784610465	0,000479915	NANOG
0,781146769	3,82519E-06	FAM213A
0,77936268	5,34788E-42	HACD3
0,777774615	5,2917E-05	PTPRE
0,771992695	0,000158086	MT1X
0,768824903	0,000855064	RNU4-2
0,759662275	2,97874E-52	P4HA1
0,757963175	0,001104713	DNAJC12
0,757216215	1,8341E-07	CCND2
0,75643867	2,59305E-16	FABP5
0,745533846	0,001077709	CXADR
0,744949938	0,001327143	LINC01583
0,743975809	2,14601E-05	COL4A4
0,739636566	0,00139743	PLEKHG1
0,738670598	2,34015E-11	KRT14
0,728495773	0,001947911	AL590617.2

0,728022267	8,54044E-13	PLEKHA8
0,726743225	0,000194056	LIMCH1
0,724831963	0,003356288	DRGX
0,724563282	0,001941788	KDR
0,722009794	0,002753491	CFI
0,719802688	4,51403E-07	SCRG1
0,718754859	0,003654588	PCDH15
0,718583995	0,000283493	ZNF365
0,718457763	0,000714007	RCOR2
0,716548316	7,67931E-09	RUNX2
0,716445731	0,000218272	MCTP1
0,715700673	0,003964648	RMRP
0,712209239	0,001703746	AL513318.2
0,708581022	5,32436E-25	CHN1
0,708315277	0,003826633	TMEM56
0,704465626	6,50322E-10	F2RL1
0,704243041	2,83968E-19	SESN2
0,703537726	2,34327E-43	ROBO1
0,701392904	1,12947E-18	VLDLR
0,700858416	0,004516972	ITPKA
0,700271955	1,7266E-25	ACKR3
0,698590951	0,000446729	PRTG
0,697670575	1,77891E-42	BHLHE40
0,69572514	3,62523E-11	DACT1
0,694811297	5,04291E-13	SPSB1
0,690481872	0,00307991	PCYT1B
0,690061707	2,37248E-44	IDH1
0,688468879	0,000289401	TMEM170A
0,688404553	0,005021356	RNU5A-1
0,682767686	0,004562005	PRR15
0,682209592	2,1873E-23	ASS1
0,678950888	3,15808E-08	OSBPL6
0,678577857	1,76373E-05	SLC16A9
0,678385636	5,0658E-23	MIAT
0,677806555	8,43615E-05	TNFAIP6
0,674959984	0,000471553	GDF15
0,674397585	3,09988E-06	AC098864.1
0,67303608	1,16495E-06	HEY1
0,671058881	3,7919E-06	LHX8
0,668350859	1,72742E-06	ARL15
0,667575349	1,3555E-46	CKB
0,667281902	2,47075E-12	ULBP1

0,665909959	0,00474453	RNY3
0,66430675	0,006588681	SPARCL1
0,662526537	0,000573107	COL26A1
0,662226591	1,45228E-07	DPPA4
0,661700497	0,004196106	WISP1
0,660357339	0,000725752	NRARP
0,659579222	2,09303E-22	FAP
0,657605478	0,008516646	TTYH1
0,657586098	0,001788813	KIF21B
0,65525808	0,009420915	FAM124A
0,653351867	0,005998507	SSC4D
0,652059604	0,000356594	LIN28B
0,651333113	0,007894188	GADD45G
0,651073469	0,002741605	TSPAN11
0,649108195	4,7617E-05	CACHD1
0,647881448	0,00560114	NOX4
0,647846018	0,009888271	KRT81
0,64771784	0,000158071	RHOU
0,647058268	0,009195716	PDGFB
0,646642247	0,0001621	KIAA1211
0,646514654	0,005618756	TMEM100
0,644982869	0,00197996	CASC9
0,642777393	1,28282E-06	ZHX2
0,641844468	3,93589E-07	ATP10A
0,641732112	0,00801051	MBNL3
0,639481847	5,3593E-15	ROR1
0,638460843	4,50373E-19	MRAS
0,63736551	0,006204562	NEBL
0,637137657	7,75971E-24	TMPO
0,636392288	0,006826655	F11R
0,634012622	0,000372806	CBX2
0,633991209	0,009424949	RAP1GAP
0,631936272	0,00988295	ARHGEF26
0,630220351	0,004866525	ZNF674
0,629890312	0,00016753	KCNT2
0,627989026	1,07157E-10	STRA6
0,627213871	0,003054202	SOX21-AS1
0,626813358	0,007337482	SLCO3A1
0,625212842	4,7551E-22	SKIL
0,625144997	0,002488407	TSPAN12
0,6239441	5,64205E-07	UBE2D1
0,618941756	0,001689986	SEMA6B

0,618420987	2,64276E-05	RNF144A
0,618098124	0,000701388	AZIN2
0,615725279	4,98837E-29	MAP7D1
0,614863943	1,84047E-07	FAM107B
0,614158903	1,96116E-19	ANK2
0,613379044	1,83789E-25	PDGFC
0,612951045	0,000855242	KIF21A
0,611387859	0,0059732	GDF6
0,609463217	1,94146E-18	CDC42EP1
0,60846806	3,55826E-13	ITGA4
0,607032201	0,000321997	ZNF521
0,605446304	0,000628498	PCTP
0,604598729	0,001555206	MARS2
0,600861477	4,66785E-05	NHSL1
0,600563949	7,20681E-09	GPR1
0,600485491	3,30588E-07	POLR2K
0,596659214	0,004136742	MYO7B
0,595838329	9,15869E-14	JUP
0,593856231	5,75687E-11	RAB12
0,593628228	5,3312E-11	KCNG1
0,592807364	0,000434243	CDCA7
0,592036583	5,5489E-34	MFAP4
0,591308162	1,08088E-06	FRAS1
0,590096978	3,25739E-07	ASF1A
0,588233446	1,37023E-14	ADGRG6
0,587828702	9,97907E-12	ENC1
0,586726343	0,000626736	RNF125
0,585818953	0,005273205	COL21A1
0,584683251	1,08725E-10	NME1
0,584359841	2,61887E-06	RNASEH2B
0,582890987	3,63163E-38	PLOD2
0,579448947	1,14008E-27	IVNS1ABP
0,577316236	0,000149345	INHBA
0,576026207	0,008259396	DHRS3
0,576024557	0,000257263	CALCRL
0,575108127	1,31742E-06	MINDY3
0,571904003	2,65412E-06	ZNF385A
0,571809496	0,00119316	MATN3
0,570814715	0,000953462	GNG2
0,570544609	0,001069279	ARHGAP26
0,568189117	8,00764E-09	SUDS3
0,568130309	3,19117E-25	TGOLN2

0,567287377	4,20775E-06	GADD45B
0,564936078	3,86542E-16	CLCN3
0,563716898	0,000145153	SLC39A11
0,563573927	6,93045E-07	VASH1
0,561414199	0,001321604	CPS1
0,560434241	4,20822E-05	SPOPL
0,55822257	2,32708E-06	SNHG15
0,558111748	0,001607399	TMEM154
0,556688676	8,64013E-12	CHAC1
0,556639219	5,35995E-10	TSC22D3
0,556359532	5,039E-09	MID1IP1
0,55632072	3,0425E-09	DNAJC1
0,554231434	6,63923E-17	TRIB3
0,554045054	2,77031E-07	PTHLH
0,553703901	9,88031E-33	SLC2A1
0,553690225	3,85592E-09	TLN2
0,553081767	5,15526E-21	SLC20A1
0,552420499	2,58365E-05	PLXDC2
0,55141388	2,07173E-10	JUNB
0,550512133	8,61135E-58	NREP
0,550241275	2,46532E-24	RCC2
0,550133781	1,56084E-05	GUCY1B3
0,547819773	1,84265E-06	MIGA1
0,547636293	2,9609E-15	GPT2
0,546995547	7,64667E-05	EPB41
0,546245333	0,002195214	SPX
0,544977438	6,70344E-11	CFH
0,544511567	0,000932904	ZUFSP
0,54446497	0,000159056	CACNB1
0,543237317	1,05133E-12	ATXN7L3B
0,540922515	0,003900112	RPLP0P2
0,538962429	0,005600068	ZC3H12A
0,535189263	1,24157E-08	BAMBI
0,535060946	0,000446841	RNF138
0,533306964	6,44861E-11	TBC1D2
0,53130314	2,52386E-05	ATF5
0,530576875	6,30192E-05	ARL4A
0,529963632	7,10059E-21	VEGFA
0,529188379	0,006826655	B3GNT5
0,529182588	0,000479736	NEK2
0,526795767	0,003131906	BRI3BP
0,523643437	4,04822E-07	PHTF2

0,523088794	4,09357E-22	H2AFZ
0,52266932	4,66212E-09	SPP1
0,5224167	2,5141E-05	LURAP1L
0,520111354	0,000124757	SMOX
0,520052155	8,32311E-12	TGFBR1
0,519899104	1,34506E-09	IER3
0,517902722	6,20617E-13	REEP3
0,516752656	0,000626603	S1PR1
0,515429103	3,81129E-17	PCK2
0,515050957	6,64357E-36	SLC7A5
0,513919165	1,29877E-10	LBR
0,511522478	0,000134394	AP1AR
0,510529697	4,72726E-17	TEK
0,508369692	6,00853E-17	OLFML3
0,505233585	0,005045517	UNC5C
0,50479654	0,005166216	DOCK9
0,504408821	3,14837E-07	SERTAD4
0,503526866	1,75504E-11	FHL3
0,503255533	0,002324458	TMC6
0,503226601	0,000605461	CTH
0,501778887	0,001601178	GLRB
0,500387093	0,000776513	IL20RB

Table 37. post OSKM upregulated genes in shSETD2-3		
postOSKM shSETD2-3		
log2FoldChange	padj	gene name
3,024027629	2,1956E-92	LEFTY2
2,805833579	1,12031E-78	RBP1
2,632328082	2,40429E-79	DHRS3
2,254267278	1,68683E-66	GUCY1A3
1,952399767	4,73896E-19	DSG3
1,844814556	1,30616E-15	TMEM155
1,83989472	3,35353E-26	NEBL
1,792868248	5,68096E-62	PTPRB
1,739308789	4,45983E-21	SCEL
1,7036663	4,50344E-27	ARL6IP1
1,703287296	2,0982E-275	KRT16
1,701177557	1,34387E-32	CLSTN2
1,682375148	8,13128E-39	L1TD1
1,636977199	1,26873E-19	MFAP3L
1,610362829	1,20631E-25	AC105243.1
1,592114842	3,53106E-66	KRT17
1,588133926	4,85863E-23	GALNT12
1,563572084	7,53553E-27	STMN2
1,531442013	1,02454E-11	LEFTY1
1,517206055	9,99579E-11	PDGFB
1,513412551	2,39696E-14	PRR15
1,509516378	1,25008E-18	SYT2
1,499893105	1,83937E-86	SPP1
1,474091725	2,76369E-15	FNDC5
1,459307611	5,9014E-34	INHBE
1,44614734	5,66338E-10	PCDH15
1,443724303	9,66241E-14	NANOG
1,44084086	1,87044E-15	SCGB3A2
1,423610186	1,46419E-13	HPGD
1,419849196	1,11453E-46	SERPINB9
1,410771471	1,42958E-12	SYT6
1,400617145	2,08952E-12	AC105460.1
1,376574622	3,73083E-27	SLC16A9
1,373832983	2,32937E-11	APLNR
1,361098401	1,54476E-17	PLA2G16
1,358672606	6,23819E-31	ISLR
1,356589201	1,14627E-51	CCL2
1,354079738	4,05474E-09	TMEM178A
1,348851447	4,43353E-23	FOXS1

1,344164668	4,31734E-20	EPHA3
1,342414582	2,60637E-21	BMP6
1,341574753	1,33139E-11	KRT5
1,33926867	3,83083E-08	KCNE5
1,334042921	4,69575E-08	ISLR2
1,33287146	9,1481E-145	CDKN2B
1,331027832	7,37891E-10	PRODH
1,323963053	2,10047E-08	FGFBP2
1,322367445	9,51211E-29	PMEPA1
1,314538905	9,35427E-41	ST6GAL1
1,314042129	2,727E-28	ESRG
1,310729257	5,677E-08	CD200
1,301127827	7,40616E-09	IQCA1
1,299934726	7,23639E-14	ALDH1A2
1,293214338	4,66591E-08	CDH1
1,287123642	5,37959E-22	SEMA6A
1,283042878	6,02E-11	F11R
1,282679141	9,73001E-23	MATN3
1,281207367	7,27444E-09	LRFN5
1,275251313	2,56479E-17	TPTEP1
1,257055991	2,33858E-19	MDFI
1,255817399	7,26223E-20	CTSO
1,25035962	7,27743E-24	CHD7
1,229802934	6,95533E-32	LXN
1,227816448	4,1102E-17	MECOM
1,225638465	9,7639E-17	LGR6
1,223027864	4,78676E-09	LAMP3
1,219486136	1,8451E-131	KRT6A
1,212536551	3,90676E-09	RAB20
1,209469848	5,00548E-10	FA2H
1,205484525	3,18863E-07	CD177
1,204543171	1,09312E-08	PLEKHG1
1,194805377	1,78032E-18	DENND2A
1,193970457	1,2314E-63	PCDH1
1,193825354	4,61974E-07	RASSF6
1,186523483	1,11055E-10	ADGRV1
1,186510805	6,77591E-07	MIR30A
1,181281325	9,7319E-20	PIK3AP1
1,173819743	2,15116E-06	AP000812.2
1,160694712	1,16312E-61	SPSB1
1,159307012	1,72315E-17	PDE4D
1,156815222	6,58129E-31	KIAA1755

1,153481283	3,45342E-08	KDR
1,145182487	1,32067E-58	DSG2
1,14513229	4,14577E-78	MIAT
1,136857214	5,18199E-39	F2RL1
1,13474763	1,12044E-07	CNTN1
1,126786199	3,14657E-13	IGSF3
1,121680111	7,97991E-75	FAP
1,121057511	1,92069E-23	ANXA3
1,120217874	1,02733E-06	CFI
1,11925823	5,44691E-06	AL450472.1
1,115933107	8,2004E-12	LIPG
1,110691077	1,53147E-06	ITPKA
1,106453659	3,13127E-11	NRARP
1,104289651	3,18243E-12	BCHE
1,098469628	8,40674E-76	ACKR3
1,098209893	4,05633E-06	C13orf42
1,097789321	7,58491E-06	COL2A1
1,096117348	1,52937E-51	RARB
1,091585607	1,60981E-06	TRIML2
1,089982977	1,69068E-39	STRA6
1,087795718	5,15458E-19	HHIP
1,084816951	4,23402E-25	CALB1
1,081750542	2,31006E-38	FIBIN
1,07989589	9,76186E-08	PCLO
1,078811109	9,54532E-06	MIR17HG
1,072908732	1,28459E-08	EMB
1,070329358	2,14224E-05	PCSK1
1,064183918	3,5214E-67	JUP
1,05973512	1,08496E-17	RGS2
1,044263776	1,98305E-08	IGFL2
1,042267269	2,50949E-16	SEC14L2
1,037530672	9,81467E-06	NPR1
1,036123426	3,28547E-05	CHRNA9
1,035973357	1,4928E-08	CDHR1
1,0351825	8,3217E-07	SSC4D
1,031989694	3,7591E-05	AC008517.1
1,028336267	8,31555E-30	PTHLH
1,027747958	1,72944E-05	TTYH1
1,025509694	4,92672E-05	WDFY4
1,009147786	7,01182E-05	CDS1
1,00800368	5,45482E-06	FILIP1
1,007040974	2,27526E-05	NLRP12

1,005960801	1,46254E-06	HTR1B
0,99985115	8,38182E-05	SFRP4
0,99770976	3,33495E-06	CXADR
0,997056825	6,00789E-05	MAP1LC3C
0,990877037	3,72006E-11	LOXL4
0,983069146	2,60353E-23	PCDH10
0,980943862	1,13616E-11	MBP
0,979473254	2,65666E-21	AK4
0,978222657	7,44935E-07	EMCN
0,976480067	5,81653E-12	CACHD1
0,973366857	0,00014334	RAPGEF5
0,972105774	8,55232E-05	CLIC5
0,971333625	2,0732E-12	CAVIN2
0,971038654	5,69782E-05	SERPINA5
0,969977437	4,16272E-08	MCTP1
0,968954781	3,89582E-05	DIRC3
0,968804146	0,000147861	HTR1D
0,967753691	1,5466E-18	NCF2
0,967004968	5,93265E-38	MAN2A1
0,966102638	9,61399E-06	ASXL3
0,965899553	7,46132E-07	ACSS3
0,965834717	0,000131912	RARRES1
0,956494844	2,27298E-05	HHLA1
0,954705713	6,9604E-113	KRT8
0,953759859	0,000149525	SMIM10L2A
0,95223446	1,97796E-09	LIN28B
0,949669688	4,02456E-05	RRAGD
0,94922689	4,62037E-07	MT1X
0,948435279	0,00022444	APELA
0,947792024	0,000226792	RHBDL3
0,946193677	0,000221946	IGDCC3
0,946143466	9,02193E-06	WISP1
0,944792328	3,42522E-24	POU5F1
0,944211311	2,52608E-98	RAB1A
0,94285735	4,89667E-35	LAMC2
0,942643482	0,000225494	RGCC
0,939861849	2,36366E-07	PTPRE
0,931384607	6,90292E-57	FERMT1
0,93125401	1,46419E-13	KIAA1211
0,930409591	3,68726E-05	COL4A3
0,930190327	5,00917E-05	COL10A1
0,928566207	0,000272003	AL117378.1

0,927674076	7,45142E-05	SLC44A3-AS1
0,926784734	9,69231E-05	FAM20A
0,926335887	1,14193E-05	CYP2S1
0,926249024	0,000236546	LINC00678
0,9256872	1,35203E-15	SOX2
0,921980524	4,74338E-81	NRP2
0,920553067	0,000256748	GPM6A
0,919417387	1,35231E-06	CASC9
0,914478915	7,20426E-06	KLRG2
0,913644621	0,00026259	LINC01239
0,910411193	1,38425E-20	DPPA4
0,909982335	8,83624E-17	CRYBG1
0,904686897	4,14497E-28	PRR5L
0,895230357	0,000102296	ZNF385B
0,894084936	2,69943E-69	TEK
0,893625094	2,40667E-07	TSPAN12
0,891909968	8,17585E-10	DOCK9
0,890764249	6,25816E-05	SALL4
0,889615004	5,78239E-06	TMC7
0,888725702	9,45956E-05	HOOK1
0,887446833	0,000259776	Z93241.1
0,885321289	0,000562547	ERVH-1
0,881690476	9,79032E-05	TIE1
0,88038828	1,83418E-09	SLC22A23
0,87758844	0,000756232	MACROD2
0,872376546	3,59949E-06	LIMCH1
0,871511217	3,75459E-16	GUCY1B3
0,867403589	1,42055E-06	HEYL
0,867308762	0,000333797	AL590004.4
0,866835764	7,30606E-11	NEGR1
0,866612514	3,18292E-15	ITGB4
0,86393191	4,19917E-13	SLC39A8
0,862806902	1,33634E-60	NNMT
0,862259909	1,44972E-21	SNHG15
0,859026819	0,000999134	TEX41
0,858862486	0,000961851	LINC00973
0,858058408	1,4251E-36	ENC1
0,855195734	0,000443593	CXXC4
0,854560765	1,03E-05	LRRC15
0,853609717	0,000415302	MAPK10
0,853128146	6,38602E-05	C6orf15
0,851156691	0,000407388	KYNU

0,850900993	1,71575E-15	SIK1
0,848882156	0,001139698	ATP12A
0,84560464	0,001182249	HS3ST5
0,844664852	0,000383094	TSPAN1
0,842330678	1,44379E-28	CHST2
0,84132136	1,69308E-41	F3
0,840398937	0,000701875	VWDE
0,838695544	0,001176935	ERVW-1
0,837794329	0,001236446	KLF3-AS1
0,837295517	5,2879E-117	ADAM19
0,835955665	0,000903712	NPSR1
0,834080468	4,92475E-05	FZD3
0,833212171	1,6461E-123	PLOD2
0,832301812	2,88541E-05	RASEF
0,829006066	0,001642461	MYLK2
0,826472848	0,001022357	CPVL
0,82120429	0,000290803	FBP1
0,820312547	0,000236546	TDGF1
0,819688063	1,6964E-151	KRT18
0,815575343	2,07507E-07	RNF125
0,81550267	0,000226952	SLCO3A1
0,813913693	0,001115853	RNY3
0,810543082	7,50271E-07	SIX2
0,810153962	7,39281E-09	PTPRU
0,807945278	0,002144107	LINC01559
0,805542598	2,08247E-18	KRT14
0,805174208	0,002194381	SLC26A7
0,805018616	0,000489522	LINC01583
0,804043684	1,19312E-33	PYCR1
0,804017715	5,75856E-05	C3orf52
0,803369137	0,001577076	PLLP
0,802122196	0,000685813	SPINK1
0,800951838	6,87094E-64	BHLHE40
0,800201933	0,002439691	ZFHX4-AS1
0,798201153	0,002507368	B3GALT4
0,797926836	0,00156737	ENPEP
0,797847817	1,36762E-25	SERTAD4
0,797326982	0,001949005	LINC01592
0,793740764	0,000949884	AL590617.2
0,792663716	0,000276045	NOX4
0,791978402	1,147E-40	KIAA1522
0,7912741	0,001289585	BIK

0,790880565	2,7131E-42	CHN1
0,789270522	1,50896E-05	TSPAN11
0,789013303	2,10134E-07	TMEM154
0,788741808	0,002419623	ERG
0,78818237	0,002105352	SLC44A5
0,787490195	2,82404E-27	ST6GAL2
0,78565144	0,002513191	TMEM88
0,784835843	0,001353245	SLC13A5
0,784717738	0,000244956	TRIM6
0,783999959	0,00240308	HPDL
0,779862433	7,21529E-16	NGEF
0,778686394	5,62272E-09	RNF138
0,777481441	4,04934E-10	PHKA1
0,774960982	1,43032E-10	CD36
0,774566416	0,001925609	WIPF3
0,774428686	5,2876E-37	ANK2
0,773998617	0,000120315	NRK
0,77380747	2,15073E-07	G0S2
0,773778522	0,00365621	EEF1B2P6
0,773203081	3,43704E-09	CD83
0,772654786	0,000665447	OCLN
0,771672185	2,5137E-21	GPC4
0,770917716	3,44982E-12	HIC1
0,769157903	1,07513E-05	COL4A4
0,769008045	0,000111782	UCP2
0,76802205	0,004056299	RASL11B
0,767628711	0,000348243	HTR7
0,766565974	0,002903634	LINC00571
0,765420377	9,93215E-37	ASS1
0,764882174	2,81583E-12	ZNF385A
0,763821057	0,000865225	PLXNC1
0,761671038	0,001385493	MYO1G
0,760789939	0,002763678	FMNL1
0,757836667	0,004239183	MYCL
0,75673698	0,003674197	FHDC1
0,756178581	0,004141071	ADD2
0,755322435	0,000442978	BCL11A
0,754251979	0,004323225	RNASE1
0,753271101	5,95466E-07	ZNF521
0,752390502	0,002914806	HHLA2
0,751508039	5,58767E-16	OSBPL3
0,751280168	0,00182331	RTP1

0,750868049	7,59842E-07	DTNA
0,7507013	0,005027228	MSI1
0,749638506	0,003503394	TNFSF15
0,749530821	3,1142E-143	VCAN
0,749384573	0,000164171	MYO5C
0,749275021	5,2876E-37	SKIL
0,748544506	0,000170066	ZNF365
0,748313946	0,003973078	AC021231.2
0,7480721	4,40645E-31	CLCN3
0,747112587	1,64333E-08	ARL15
0,746198944	0,000838428	RPS23P8
0,745761997	0,004972721	RNF128
0,745729161	0,00525858	CACNA2D2
0,743834207	9,72737E-06	AZIN2
0,743423364	0,005797228	TRIM9
0,743140829	8,47095E-06	FAM213A
0,742277116	1,35359E-33	GPC6
0,741827688	0,000362892	IQGAP2
0,741570204	5,22513E-09	NRCAM
0,739332786	0,006122986	KRT81
0,735264464	5,42721E-22	TUBB2B
0,731899607	0,000149194	TRIM71
0,731585296	0,002196416	SOX21-AS1
0,72954772	0,000289879	SH3D21
0,729421967	1,34111E-24	CD55
0,729000572	2,65559E-48	OLFML3
0,728914663	3,4967E-06	GPAT3
0,727767595	1,9111E-126	DSP
0,727764194	0,000248466	WNT2
0,727374923	0,006750292	ASRGL1
0,727119796	0,002450052	SNORA53
0,726115402	0,000144116	PRTG
0,72466905	0,003662291	NRN1
0,723212495	0,006333797	CACNA1I
0,723100353	7,02047E-07	RHOU
0,723098318	2,49797E-33	DDIT4
0,723031516	0,000421952	LYPD1
0,722875386	0,005868163	RIMS2
0,722706763	0,006431577	LEF1
0,722637979	0,007085041	CFAP45
0,720614205	0,00236757	DNAJC12
0,71926385	2,31852E-81	SLC2A1

0,718809688	0,005359211	MUC4
0,717924048	3,2749E-23	SORBS2
0,717421509	0,001604613	MAP2K6
0,71612894	0,007018192	AC078850.1
0,714807277	0,007080775	KRT13
0,714034677	0,002537944	LINC00520
0,713798968	0,00010965	LINC00702
0,713106449	4,75662E-06	ARRDC4
0,712035582	0,007922048	ARHGAP30
0,711202603	0,008885802	PLBD1
0,710256987	0,001105274	IRX3
0,709463039	7,03606E-07	MSC
0,707028529	0,004189591	BRINP3
0,705211856	0,003781565	ATAD3C
0,704984533	0,007677788	LCP1
0,704103336	0,000444797	PARP8
0,703532041	0,002129065	CHRNA1
0,703491466	0,00013294	LINC00511
0,702886121	5,37759E-09	CTH
0,702869072	0,007457187	ONECUT2
0,701066218	0,009381722	C1orf115
0,699823582	9,01178E-10	AP1AR
0,699779606	0,009523129	RAPGEF4
0,699577058	5,76326E-82	YWHAG
0,698580929	1,8687E-18	RFLNB
0,698443535	0,000327991	NECTIN4
0,695739267	1,97662E-27	NUS1
0,694013935	0,007515294	TMCC3
0,693953786	0,006677965	AL591845.1
0,693873981	0,005974899	PCDH19
0,69196925	0,007485153	BEX2
0,691187127	0,000108031	MAL2
0,688306036	6,12522E-06	NHSL1
0,686976252	2,07478E-11	CLCN5
0,686664495	0,000115806	SEMA3G
0,684287871	0,000221946	FRMD5
0,684273565	0,000122203	KIF21A
0,681648261	1,24255E-24	MRAS
0,681492006	0,006791938	SHPK
0,681229193	0,007778952	KRT6C
0,679328866	0,004239183	FRMD4B
0,678917313	0,003400443	AC104083.1

0,678598545	1,05125E-19	PMAIP1
0,676710916	9,74469E-06	GJB2
0,675400765	3,34608E-26	UNC5B
0,674416085	3,21888E-19	FAM60A
0,674087027	6,77728E-08	HAS2
0,671991338	8,7156E-08	CCND2
0,671769179	0,009371159	ELMOD1
0,671019424	1,15606E-11	IL17RD
0,670102861	7,22079E-05	MBOAT1
0,665677098	0,006377525	RNF144A-AS1
0,664000939	0,001641024	ATP2A3
0,663537938	0,009269848	DNM3
0,663404602	1,2181E-15	MID1IP1
0,66303916	8,91849E-07	MXD1
0,662972463	1,3294E-07	SMOX
0,662396532	0,005323163	LINC00437
0,662370998	1,85584E-44	RCC2
0,662365443	1,41582E-21	TBC1D2
0,662106714	0,005343698	PLCL2
0,659720255	0,000193723	VSNL1
0,659363259	2,92932E-42	CBX5
0,659099842	0,005906996	PRKAA2
0,658867511	0,001744049	NHS
0,658747299	4,50285E-12	SLC44A1
0,658408695	4,79064E-15	GLIPR2
0,657736842	4,9282E-14	CNOT6
0,65695552	3,71866E-06	CDCA7
0,65350906	1,17801E-45	TGOLN2
0,650320723	1,26995E-07	OSBPL6
0,650019646	3,95329E-07	KCNK1
0,649366143	0,000340763	PKIA
0,64930164	0,001572703	STARD8
0,648848656	0,000366157	PTP4A3
0,648253613	2,40963E-10	UBASH3B
0,645090528	0,007797228	C2orf27A
0,642921465	0,002844499	RAB3IP
0,642911186	0,000908974	MYO7B
0,64220651	1,09312E-08	MINDY3
0,642113734	4,54419E-05	LFNG
0,641619596	0,001995804	IFITM10
0,641489687	4,08224E-16	TSC22D3
0,640118256	8,20765E-08	CELSR1

0,639774566	1,28638E-16	IER3
0,6390179	0,000352548	BCL2L11
0,638924256	0,006565902	CDCP1
0,6385869	7,71506E-05	BMF
0,636910168	7,50977E-09	GPR1
0,63640597	2,28919E-12	AVL9
0,636114891	0,003681223	HSPA12B
0,63598575	0,006294721	CYYR1
0,634806788	0,002139587	NTNG1
0,634570014	2,48534E-05	SERPINB7
0,634338578	0,008208932	RHOF
0,634126574	0,002830613	FBXO32
0,633330692	0,000499007	HGNC:18790
0,633262062	2,08538E-81	PYGB
0,632872384	1,62612E-41	SYNPO
0,632576308	0,00011968	GRASP
0,632024083	0,001727012	CA2
0,631724374	0,000492233	ST8SIA4
0,631648587	1,39616E-06	ITGB3
0,631486226	3,67204E-30	HACD3
0,630486705	8,64367E-09	PELI1
0,630371166	0,000744436	ANO4
0,629455362	8,45243E-05	BRI3BP
0,626047312	0,002501062	AC073621.1
0,623698155	0,002269326	KIF21B
0,6233772	2,43648E-54	P4HA1
0,623307618	8,39113E-06	CSRNP1
0,622838881	1,29014E-47	SERPINE2
0,622202307	1,14206E-06	HEY1
0,620826004	1,2055E-09	POLR2K
0,620662617	0,006042276	CYP26A1
0,619590017	0,00177598	DAPK1
0,619245596	6,47563E-05	TMC6
0,6170235	0,004207111	AL356488.2
0,613788104	7,71968E-06	SLC39A11
0,612544866	0,00494868	ZNF674
0,611440844	1,57074E-20	ATXN7L3B
0,611138942	7,10385E-12	TRIM24
0,610932374	8,31483E-11	KIF26A
0,609455662	1,83368E-34	MAP7D1
0,608719181	0,006289618	SFTA1P
0,608506316	6,14934E-20	PTGFRN

0,608356485	1,11953E-11	SDC1
0,605636205	2,83975E-13	IL1R1
0,605007554	1,82232E-07	KLHL4
0,604108501	0,004351256	PTGS2
0,603644589	2,6107E-40	ROBO1
0,602301196	1,15305E-27	ABLIM1
0,60123556	0,000149828	MGP
0,600683868	3,75575E-13	KCNG1
0,60019198	0,00097824	ZC3H12A
0,598815541	2,20888E-11	PKDCC
0,598431637	4,96863E-13	AQP3
0,596290715	1,79727E-13	FABP5
0,595939454	2,23429E-15	TCF4
0,595938469	0,005113358	GALNT3
0,595747941	0,008392477	SPAG17
0,595712524	0,004445365	COL25A1
0,595440126	0,000250355	METTL7A
0,594845547	0,000544645	PPP1R13L
0,593502492	3,79915E-20	MBOAT2
0,592250558	5,37277E-35	IDH1
0,591887993	0,004341057	MARS2
0,591105639	8,35216E-08	GCNT1
0,590410433	0,008879618	ZBED2
0,589974854	1,42609E-14	LAMB3
0,589804035	0,007922048	SELL
0,589720877	0,000872169	TMEM170A
0,586981041	0,000599159	IL32
0,586822707	0,003298949	ICOSLG
0,586185363	3,47463E-61	CKB
0,585840158	3,16885E-12	BNIP3
0,585443149	9,5593E-08	SLC38A5
0,584338738	4,4892E-17	UHMK1
0,583592033	7,48406E-23	ITGA6
0,583278231	1,60352E-22	PDGFC
0,582231477	2,86786E-09	ADA
0,581706088	0,00338128	CD3EAP
0,580320592	1,63645E-22	GRB10
0,579094299	1,76458E-11	ZNF827
0,578187854	6,97125E-06	PAPPA2
0,577829475	0,000903291	MAP2
0,577224464	0,000693191	WFDC21P
0,577196192	0,008902444	PDZD2

0,576773893	1,40412E-13	VLDLR
0,576716233	2,83229E-10	SUDS3
0,576338839	0,000204312	VEPH1
0,576136853	1,05457E-07	HIC2
0,575343799	1,01313E-13	NARF
0,574725568	0,000324842	ARHGAP26
0,574464732	0,002522617	HIST1H4H
0,574259996	0,002190224	EPB41L4B
0,568990838	5,92875E-05	SCRG1
0,568731095	0,000639591	NDUFA4L2
0,56371258	3,21584E-07	LARS2
0,563407399	1,09533E-12	A4GALT
0,563335407	0,004849042	RASGRF2
0,563108307	0,000890737	KIF26B
0,562830879	4,24341E-12	ITGA4
0,562181901	1,00637E-10	RAB12
0,561671247	0,00095673	TUBB4A
0,561493268	2,19489E-05	DUSP16
0,56142142	0,008673609	CCDC33
0,560228359	1,89067E-84	GSN
0,558681335	0,000132516	ZNF300
0,558226738	1,24952E-05	RUNX1T1
0,558171536	0,008713411	FOXO1
0,558097789	2,01303E-73	BGN
0,557988138	0,005387598	RPLP0P2
0,556809059	6,42753E-09	VASH1
0,556585449	1,8314E-07	MLX
0,554925979	0,000118649	IL20RB
0,553802405	0,000222689	AC098864.1
0,553507835	3,07601E-12	RFK
0,551832489	4,17777E-13	PYGL
0,551317866	1,51975E-05	UBE2D1
0,550797332	2,03401E-31	EZR
0,550506729	0,009744039	NOG
0,550164397	7,61126E-05	RNF144A
0,548408289	1,32823E-14	ADGRG6
0,548050412	6,20348E-05	PAPLN
0,547249499	2,83573E-95	TGFBI
0,546265521	1,28871E-06	TMEM41B
0,546115062	6,69674E-05	RUNX2
0,545424782	8,45112E-07	AMMECR1
0,542537271	0,008922954	SLC22A3

0,541771632	7,87125E-18	CDC42EP1
0,541620906	4,1388E-06	ATP10A
0,540338506	4,58028E-09	SAT1
0,539710412	5,00187E-08	NFATC2
0,539079691	5,92885E-21	EIF1AX
0,538086666	7,11885E-05	FGFR3
0,53721824	5,55998E-13	PHLDA3
0,534360571	5,48962E-20	SLC20A1
0,533290589	6,01073E-11	JUNB
0,532231514	5,1115E-11	CCPG1
0,530441357	0,004971425	TLR3
0,530258073	5,9775E-49	CRABP2
0,528435968	8,62817E-18	REEP3
0,528324389	1,58037E-11	ZNF24
0,527848313	3,89515E-06	PDLIM3
0,525815295	2,65565E-16	KLF10
0,523568885	0,000451082	KDM7A
0,522029289	0,009646758	KLF5
0,521543338	0,000981366	FAR2
0,521440088	3,59676E-05	PCDH9
0,521087986	0,007690795	PSD4
0,520732856	7,70245E-11	TPST2
0,520089995	3,59482E-06	AGPAT4
0,519951866	6,33762E-06	PGM2L1
0,51810861	0,008731767	COL26A1
0,517352089	0,000493259	INHBA
0,516481712	0,000450887	USP54
0,516403705	6,34347E-06	ZNF608
0,515353609	0,002191786	GNG2
0,515042822	1,2745E-05	MIGA1
0,51229749	3,54664E-15	AGO1
0,511324581	5,1649E-06	PHTF2
0,511313033	1,81195E-05	KCTD12
0,511000148	1,0487E-07	KLF9
0,51095052	2,29999E-10	GFPT2
0,510267656	0,004634373	INAFM1
0,510099247	5,93137E-11	ITPR2
0,509629544	5,20981E-16	TMEM181
0,507742667	1,7782E-06	SOX9
0,50747787	9,09108E-05	OCIAD2
0,504350236	0,007219608	B3GALT5
0,50111954	1,46561E-14	CAMK2D

0,500453839	0,00087458	MYCBP

Table 38. Fibroblast Gene Set						
LINC01279	TSHZ1	EPS8	CDKN2C	NR3C1	ANO3	TCEAL1
DCN	SNAI2	SYNPO2	PTGS1	C1S	MMP3	NEGR1
TNFRSF11B	SLC16A4	DAB2	CDC42EP5	MXRA5	RCAN2	SLFN5
NNMT	MFAP5	ARSJ	AHR	VASN	COMMD3	CHSY3
SRGN	RGS4	ATP10D	COL16A1	CFL2	IL1R1	MIR31HG
GREM1	DKK1	SAMD9	FKBP9	FIBIN	CBR3	GBE1
GPNMB	VEGFC	PRR16	FAS	OSR1	LHFP	KDELR3
LUM	TGFB1I1	COL15A1	CDH11	CORIN	FOSL2	MRVI1
ITGBL1	INHBA	PDGFRB	BST1	MKX	MFSD1	SLC38A11
NUPR1	THBS2	CCDC92	CLIP4	TSPO	ZHX1	NDRG1
GBP3	VGLL3	SSPN	TGFBI	EVA1A	ZBTB20	PALMD
GBP1	DDR2	EFEMP2	AHNAK2	IFI6	HCFC2	COMP
FBN1	FOXC1	SPOCK1	SEMA3C	PCOLCE	WNT5A	CRIM1
LPPR4	PLSCR4	FHL1	PLAT	EMILIN1	SCN9A	RRAS
CCDC80	COL5A1	AMIGO2	PRSS23	LINC01133	F2RL2	MARCH4
NTN4	MFAP4	DNM3OS	COL12A1	WFDC21P	NFIC	SAMD9L
SYNC	ASPN	PTRF	COL5A2	MYOCD	LOC101927263	GEM
FBLN5	EMP3	ADAMTS1	MBNL1-AS1	CD44	GLT8D2	ATP2B4
FOXF2	NPR3	CAV1	HMCN1	ADAMTS2	TFPI2	GNPDA2
MIR100HG	SLC22A4	CPQ	KRTAP1-5	NEK7	SRPX	ULBP2
PRRX1	NID2	LPAR1	MAGI2-AS3	CASP4	IL10RB	MGARP
GLIPR1	RNASE4	TIMP2	CYP1B1	RARRES1	CDH13	FLJ32255
SERPINB2	NEXN	IGFBP5	CRIP1	FAM114A1	LIMS2	CD59
OXTR	S100A4	SLFN11	NABP1	SELM	IFIT1	IRF9
LOX	C1R	ITGA4	CRLF1	ANGPT1	MEIS2	ELN
SULF1	DAAM2	RHOJ	COL1A2	IL15	NR2F2	C1orf198
BGN	LY96	ALPK2	IL6ST	RGL1	FBLN2	COL5A3
CRYAB	CFH	PPP1R3C	MVP	SPOCD1	COPZ2	TMEM117
AHRR	IGFBP7	MAMLD1	ENPP2	ECM2	LOXL1	KAT2B
ADAMTS5	SLC1A4	CAV2	CALHM2	MXRA7	CCDC176	ANXA1
MSRB3	NT5E	EFEMP1	ACKR4	CSPG4	SPHK1	PRRX2
MGLL	MYOF	SERPINE1	C15orf52	B3GALT2	ECM1	DRAM1
SQRDL	IGFBP6	TWIST1	C12orf75	PLAGL1	CCPG1	FOXC2
PEAR1	PRSS12	SFXN3	IGFBP4	OMD	TBC1D2	COL6A2
FAP	CEMIP	CSRP1	MAP3K7CL	IFIT5	ADGRL4	MEGF6
NOV	COX7A1	ADGRA2	THBS1	MORC4	GLIS3	ALCAM
COL8A1	LURAP1L	SH3RF3	PARVA	MSC-AS1	VIT	SLC16A7
CYBRD1	FILIP1L	ADAM12	MIR143HG	PDE4DIP	LXN	
COL6A3	TMEM119	PLP2	MGP	DKK3	PCGF5	
OSR2	SIX1	BICC1	SGIP1	ACTA2	IL13RA1	
IGFBP3	TMEM173	WFDC1	ZNF25	TLCD2	FRMD6	
TIMP3	EHD2	CD248	ZEB1	LTBP2	TRAM2	
PTX3	IFIT3	IFI16	HEG1	TIMP1	MMP1	
POSTN	EBF1	GNG11	HTRA1	SEPT11	C11orf87	
COL3A1	CNRIP1	MRGPRF	KCNE4	LGALS1	LINC00968	

Table 39. Pluripotent Gene Set						
LIN28A	SEMA6A	COCH	DEPDC1B	ADAMTS19	EOMES	GUCA1A
TRIM71	SLC7A3	HS6ST2	NPTX2	FABP5	LMNB1	ACTN3
APELA	IDO1	CTSV	IGFBPL1	ZNF850	TCEAL2	ERVH48-1
LIN28B	SCG3	ERVH-1	SALL3	DSG2	LRRC8B	GNAS-AS1
LRRN1	BEX2	SOX2	ADGRG2	PDPN	ZNF738	KCNN2
EPCAM	LOC101927746	KIAA1804	PRDM14	ZSCAN10	LINC00992	FXYD6
CD24	FAM60A	FOXD3-AS1	CYYR1	HELLS	MARVELD3	KCNK12
PTPRZ1	GPR160	CNTNAP2	DPPA2	TERF1	CDC25A	PMAIP1
L1TD1	VANGL2	SNX10	BCL11A	PRKCZ	PRIMA1	PLEKHH1
ZIC2	AASS	VAT1L	ELOVL7	GALNT12	GPR27	PPAT
PROM1	ACTA1	NMRK2	FAM213A	SERPINB9	SLC2A3	NRARP
LEFTY1	HOOK1	GPR19	TET1	PALD1	LINC01224	CDC7
USP44	CD200	KDR	GCNT2	CGN	ATP1A2	GCA
POU5F1P3	NLGN4X	DDX25	FRAT2	CHST9	GNAL	WDHD1
MAL2	NAP1L2	DMKN	CER1	RPS6KA1	FOXO1	TMEM125
DNMT3B	UNC5D	CPVL	CADPS2	RPRM	RAB39B	FAM150B
LINC01405	ZFP42	CA2	PLBD1	ARID3B	PLEKHA7	TAF4B
NTS	SLAIN1	C9orf135	MAP7	SYT6	SORL1	IGF2BP3
TSTD1	TDRP	FLVCR1	DAZL	PIM2	TMEM133	SPINT2
OTX2	NELL2	CST1	CRISPLD1	SCG5	SFRP2	CLDN6
TOX3	SALL2	ALMS1-IT1	PIPOX	ADD2	HOMER1	PPAP2C
ESRP1	TPD52	B4GALT6	TNNI3	KRTCAP3	RRAGD	GRHL2
POU5F1P4	VAMP8	C1orf106	PMEL	GPR143	JARID2	DNMT3A
VASH2	LECT1	SYNGR3	PAIP2B	RASGEF1A	PAX6	KCNK5
MIR302B	GPM6B	COBL	STRBP	CMTM8	TMEM27	PKIB
NMU	TUNAR	KIF21A	FOXH1	CHMP4C	FAM46B	LAPTM4B
FREM2	WNK3	ZNF165	PHC1	JAKMIP2-AS1	SOX11	CXCL5
HLA-DPB2	PRSS8	RNF125	TNFRSF21	IGDCC3	CHST6	STK26
PODXL	LINC00458	PRTG	FAM110C	PRSS16	STOX2	IQGAP2
FGF13	ZNF204P	SCGB3A2	FAM221A	RLN2	KIAA1551	APOC1
DPPA4	CDH3	CLDN10	LINGO1	MIAT	HPGD	GDF3
PLP1	SLC16A9	PLS1	MYO5C	SLC1A6	PITX2	LPAR4
GPC4	FZD3	FOS	EGLN3	ITM2A	RARRES2	CLGN
ZIC3	APOA2	GAL	LCK	ZNF711	CYP4X1	ORC1
GLDC	HERC5	TUBB2B	LINC00698	GRTP1	ELAVL2	PRKCQ
SALL4	BMP2	LINC00545	ZNF114	ZNF462	BEND3	MND1
FAM169A	RAB25	CA14	PRKX	SLC35F2	PPM1E	MRAP2
NANOG	CDH1	VWDE	PWAR5	SALL1	FAM117A	NEFL
CXADR	UGT8	LINC01356	COL9A3	LSR	PHF21B	
LOC101927668	F11R	NAP1L3	KCNS3	EPHX3	MTF2	
LOC101060391	CYP2S1	SBK1	LRAT	GYG2	FOXA3	
GABRB3	VSNL1	BST2	NPPB	PELI1	LINC01021	
RBPMS2	CALB1	KIF5C	ENPP5	KCND2	SLC16A1	
POU5F1B	NFE2L3	SMIM10L2A	CENPV	FZD5	SYNE4	
CYP26A1	EDNRB	ERVMER34-1	CACNA2D2	LOC101926975	ZDHHC23	

REFERENCES

1. Stadtfeld, M. and K. Hochedlinger, *Induced pluripotency: history, mechanisms, and applications.* Genes Dev, 2010. **24**(20): p. 2239-63.
2. Schlesinger, S. and S.P. Goff, *Retroviral Transcriptional Regulation and Embryonic Stem Cells: War and Peace.* Molecular and Cellular Biology, 2015. **35**(5): p. 770.
3. Ying, Q.L., et al., *BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3.* Cell, 2003. **115**(3): p. 281-92.
4. Irina Klimanskaya, E.A.K., Robert Lanza, *Embryonic Stem Cells*, in *Principles of Tissue Engineering*, R.L. Robert Lanza, Joseph Vacanti, Editor. 2014. p. 565-579.
5. Itskovitz-Eldor, J., et al., *Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers.* Mol Med, 2000. **6**(2): p. 88-95.
6. Stevens, L.C. and C.C. Little, *Spontaneous Testicular Teratomas in an Inbred Strain of Mice.* Proc Natl Acad Sci U S A, 1954. **40**(11): p. 1080-7.
7. Stevens, L.C., *Origin of testicular teratomas from primordial germ cells in mice.* J Natl Cancer Inst, 1967. **38**(4): p. 549-52.
8. Evans, M.J., *The isolation and properties of a clonal tissue culture strain of pluripotent mouse teratoma cells.* J Embryol Exp Morphol, 1972. **28**(1): p. 163-76.
9. Stevens, L.C., *The development of transplantable teratocarcinomas from intratesticular grafts of pre- and postimplantation mouse embryos.* Dev Biol, 1970. **21**(3): p. 364-82.
10. Evans, M.J. and M.H. Kaufman, *Establishment in culture of pluripotential cells from mouse embryos.* Nature, 1981. **292**(5819): p. 154-6.
11. Martin, G.R., *Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells.* Proc Natl Acad Sci U S A, 1981. **78**(12): p. 7634-8.
12. Bradley, A., et al., *Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines.* Nature, 1984. **309**(5965): p. 255-6.
13. Thomson, J.A., et al., *Embryonic stem cell lines derived from human blastocysts.* Science, 1998. **282**(5391): p. 1145-7.

14. Schwartz, S.D., et al., *Human embryonic stem cell-derived retinal pigment epithelium in patients with age-related macular degeneration and Stargardt's macular dystrophy: follow-up of two open-label phase 1/2 studies.* Lancet, 2015. **385**(9967): p. 509-16.
15. Spemann, H., *Embryonic Development and Induction*. 1988: Garland Pub.
16. 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.
17. 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.
18. Campbell, K.H., et al., *Sheep cloned by nuclear transfer from a cultured cell line.* Nature, 1996. **380**(6569): p. 64-6.
19. Williams, N., *Death of Dolly marks cloning milestone.* Curr Biol, 2003. **13**(6): p. R209-10.
20. Tada, M., et al., *Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells.* Curr Biol, 2001. **11**(19): p. 1553-8.
21. 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.
22. Takahashi, K., et al., *Induction of pluripotent stem cells from adult human fibroblasts by defined factors.* Cell, 2007. **131**(5): p. 861-72.
23. Blelloch, R., et al., *Generation of induced pluripotent stem cells in the absence of drug selection.* Cell Stem Cell, 2007. **1**(3): p. 245-7.
24. Nakagawa, M., et al., *Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts.* Nat Biotechnol, 2008. **26**(1): p. 101-6.
25. Yu, J., et al., *Induced pluripotent stem cell lines derived from human somatic cells.* Science, 2007. **318**(5858): p. 1917-20.
26. Kim, J.B., et al., *Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors.* Nature, 2008. **454**(7204): p. 646-50.
27. Maherali, N. and K. Hochedlinger, *Guidelines and techniques for the generation of induced pluripotent stem cells.* Cell Stem Cell, 2008. **3**(6): p. 595-605.
28. Stadtfeld, M., et al., *Induced pluripotent stem cells generated without viral integration.* Science, 2008. **322**(5903): p. 945-9.
29. Zhou, W. and C.R. Freed, *Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells.* Stem Cells, 2009. **27**(11): p. 2667-74.

30. Fusaki, N., et al., *Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome.* Proc Jpn Acad Ser B Phys Biol Sci, 2009. **85**(8): p. 348-62.
31. Okita, K., et al., *Generation of mouse induced pluripotent stem cells without viral vectors.* Science, 2008. **322**(5903): p. 949-53.
32. Yu, J., et al., *Human induced pluripotent stem cells free of vector and transgene sequences.* Science, 2009. **324**(5928): p. 797-801.
33. Jia, F., et al., *A nonviral minicircle vector for deriving human iPS cells.* Nat Methods, 2010. **7**(3): p. 197-9.
34. Zhou, H., et al., *Generation of induced pluripotent stem cells using recombinant proteins.* Cell Stem Cell, 2009. **4**(5): p. 381-4.
35. Warren, L., et al., *Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA.* Cell Stem Cell, 2010. **7**(5): p. 618-30.
36. Hou, P., et al., *Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds.* Science, 2013. **341**(6146): p. 651-4.
37. Guenther, M.G., et al., *Chromatin Structure and Gene Expression Programs of Human Embryonic and Induced Pluripotent Stem Cells.* Cell stem cell, 2010. **7**(2): p. 249-257.
38. Hochedlinger, K. and K. Plath, *Epigenetic reprogramming and induced pluripotency.* Development (Cambridge, England), 2009. **136**(4): p. 509-523.
39. Holliday, R., *Epigenetics: A Historical Overview.* Epigenetics, 2006. **1**(2): p. 76-80.
40. Van Speybroeck, L., *From epigenesis to epigenetics: the case of C. H. Waddington.* Ann N Y Acad Sci, 2002. **981**: p. 61-81.
41. Feinberg, A.P. and B. Tycko, *The history of cancer epigenetics.* Nat Rev Cancer, 2004. **4**(2): p. 143-53.
42. Lyko, F., B.H. Ramsahoye, and R. Jaenisch, *DNA methylation in Drosophila melanogaster.* Nature, 2000. **408**(6812): p. 538-40.
43. Ehrlich, M., et al., *Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells.* Nucleic Acids Res, 1982. **10**(8): p. 2709-21.
44. Jaenisch, R., et al., *DNA methylation, retroviruses, and embryogenesis.* J Cell Biochem, 1982. **20**(4): p. 331-6.

45. Yin, Y., et al., *Impact of cytosine methylation on DNA binding specificities of human transcription factors.* Science, 2017. **356**(6337).
46. Issa, J.P., *CpG-island methylation in aging and cancer.* Curr Top Microbiol Immunol, 2000. **249**: p. 101-18.
47. Jin, B., Y. Li, and K.D. Robertson, *DNA methylation: superior or subordinate in the epigenetic hierarchy?* Genes Cancer, 2011. **2**(6): p. 607-17.
48. Deaton, A.M. and A. Bird, *CpG islands and the regulation of transcription.* Genes Dev, 2011. **25**(10): p. 1010-22.
49. Lewis, J.D., et al., *Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA.* Cell, 1992. **69**(6): p. 905-14.
50. Meehan, R.R., J.D. Lewis, and A.P. Bird, *Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA.* Nucleic Acids Res, 1992. **20**(19): p. 5085-92.
51. Mikkelsen, T.S., et al., *Dissecting direct reprogramming through integrative genomic analysis.* Nature, 2008. **454**(7200): p. 49-55.
52. Bhutani, N., et al., *A critical role for AID in the initiation of reprogramming to induced pluripotent stem cells.* The FASEB Journal, 2013. **27**(3): p. 1107-1113.
53. Gao, Y., et al., *Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming.* Cell Stem Cell, 2013. **12**(4): p. 453-69.
54. Luger, K. and T.J. Richmond, *The histone tails of the nucleosome.* Curr Opin Genet Dev, 1998. **8**(2): p. 140-6.
55. Kornberg, R.D. and Y. Lorch, *Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome.* Cell, 1999. **98**(3): p. 285-94.
56. van Holde, K. and J. Zlatanova, *Chromatin architectural proteins and transcription factors: a structural connection.* Bioessays, 1996. **18**(9): p. 697-700.
57. Wolffe, A.P. and J.J. Hayes, *Chromatin disruption and modification.* Nucleic Acids Res, 1999. **27**(3): p. 711-20.
58. Hecht, A., et al., *Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins: a molecular model for the formation of heterochromatin in yeast.* Cell, 1995. **80**(4): p. 583-92.
59. Edmondson, D.G., M.M. Smith, and S.Y. Roth, *Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4.* Genes Dev, 1996. **10**(10): p. 1247-59.

60. Gross, D.S. and W.T. Garrard, *Nuclease hypersensitive sites in chromatin.* Annu Rev Biochem, 1988. **57**: p. 159-97.

61. McGhee, J.D., et al., *Orientation of the nucleosome within the higher order structure of chromatin.* Cell, 1980. **22**(1 Pt 1): p. 87-96.

62. Norton, V.G., et al., *Nucleosome linking number change controlled by acetylation of histones H3 and H4.* J Biol Chem, 1990. **265**(32): p. 19848-52.

63. Lee, D.Y., et al., *A positive role for histone acetylation in transcription factor access to nucleosomal DNA.* Cell, 1993. **72**(1): p. 73-84.

64. Brownell, J.E. and C.D. Allis, *Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation.* Curr Opin Genet Dev, 1996. **6**(2): p. 176-84.

65. Struhl, K., *Histone acetylation and transcriptional regulatory mechanisms.* Genes Dev, 1998. **12**(5): p. 599-606.

66. Grunstein, M., *Histone acetylation in chromatin structure and transcription.* Nature, 1997. **389**(6649): p. 349-52.

67. Thorne, A.W., et al., *Patterns of histone acetylation.* Eur J Biochem, 1990. **193**(3): p. 701-13.

68. Wade, P.A., *Transcriptional control at regulatory checkpoints by histone deacetylases: molecular connections between cancer and chromatin.* Hum Mol Genet, 2001. **10**(7): p. 693-8.

69. Ito, K., P.J. Barnes, and I.M. Adcock, *Glucocorticoid receptor recruitment of histone deacetylase 2 inhibits interleukin-1beta-induced histone H4 acetylation on lysines 8 and 12.* Mol Cell Biol, 2000. **20**(18): p. 6891-903.

70. Bjerling, P., et al., *Functional divergence between histone deacetylases in fission yeast by distinct cellular localization and in vivo specificity.* Mol Cell Biol, 2002. **22**(7): p. 2170-81.

71. Mali, P., et al., *Butyrate Greatly Enhances Derivation of Human Induced Pluripotent Stem Cells by Promoting Epigenetic Remodeling and the Expression of Pluripotency-Associated Genes.* Stem cells (Dayton, Ohio), 2010. **28**(4): p. 713-720.

72. McCool, K.W., et al., *The role of histone acetylation in regulating early gene expression patterns during early embryonic stem cell differentiation.* J Biol Chem, 2007. **282**(9): p. 6696-706.

73. Karantzali, E., et al., *Histone deacetylase inhibition accelerates the early events of stem cell differentiation: transcriptomic and epigenetic analysis.* Genome Biology, 2008. **9**(4): p. R65-R65.

74. Rea, S., et al., *Regulation of chromatin structure by site-specific histone H3 methyltransferases.* Nature, 2000. **406**(6796): p. 593-9.

75. Strahl, B.D., et al., *Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena.* Proc Natl Acad Sci U S A, 1999. **96**(26): p. 14967-72.

76. Kouzarides, T., *Histone methylation in transcriptional control.* Curr Opin Genet Dev, 2002. **12**(2): p. 198-209.

77. Lachner, M. and T. Jenuwein, *The many faces of histone lysine methylation.* Curr Opin Cell Biol, 2002. **14**(3): p. 286-98.

78. Margueron, R., P. Trojer, and D. Reinberg, *The key to development: interpreting the histone code?* Curr Opin Genet Dev, 2005. **15**(2): p. 163-76.

79. Martin, C. and Y. Zhang, *The diverse functions of histone lysine methylation.* Nat Rev Mol Cell Biol, 2005. **6**(11): p. 838-49.

80. Fraga, M.F., et al., *Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer.* Nat Genet, 2005. **37**(4): p. 391-400.

81. Hake, S.B., A. Xiao, and C.D. Allis, *Linking the epigenetic 'language' of covalent histone modifications to cancer.* Br J Cancer, 2004. **90**(4): p. 761-9.

82. Schneider, R., A.J. Bannister, and T. Kouzarides, *Unsafe SETs: histone lysine methyltransferases and cancer.* Trends Biochem Sci, 2002. **27**(8): p. 396-402.

83. Varambally, S., et al., *The polycomb group protein EZH2 is involved in progression of prostate cancer.* Nature, 2002. **419**(6907): p. 624-9.

84. Hess, J.L., *Mechanisms of transformation by MLL.* Crit Rev Eukaryot Gene Expr, 2004. **14**(4): p. 235-54.

85. Zhang, Y. and D. Reinberg, *Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails.* Genes Dev, 2001. **15**(18): p. 2343-60.

86. Dong, X. and Z. Weng, *The correlation between histone modifications and gene expression.* Epigenomics, 2013. **5**(2): p. 113-6.

87. Sanders, S.L., et al., *Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage.* Cell, 2004. **119**(5): p. 603-14.

88. Byvoet, P., et al., *The distribution and turnover of labeled methyl groups in histone fractions of cultured mammalian cells.* Arch Biochem Biophys, 1972. **148**(2): p. 558-67.
89. Thomas, G., H.W. Lange, and K. Hempel, *[Relative stability of lysine-bound methyl groups in arginie-rich histones and their subfrations in Ehrlich ascites tumor cells in vitro].* Hoppe Seylers Z Physiol Chem, 1972. **353**(9): p. 1423-8.
90. Shi, Y., et al., *Histone demethylation mediated by the nuclear amine oxidase homolog LSD1.* Cell, 2004. **119**(7): p. 941-53.
91. Clissold, P.M. and C.P. Ponting, *JmjC: cupin metalloenzyme-like domains in jumonji, hairless and phospholipase A2beta.* Trends Biochem Sci, 2001. **26**(1): p. 7-9.
92. Katoh, M. and M. Katoh, *Identification and characterization of JMJD2 family genes in silico.* Int J Oncol, 2004. **24**(6): p. 1623-8.
93. Fischle, W., Y. Wang, and C.D. Allis, *Binary switches and modification cassettes in histone biology and beyond.* Nature, 2003. **425**(6957): p. 475-9.
94. Onder, T.T., et al., *Chromatin-modifying enzymes as modulators of reprogramming.* Nature, 2012. **483**(7391): p. 598-602.
95. Sridharan, R., et al., *Proteomic and genomic approaches reveal critical functions of H3K9 methylation and Heterochromatin Protein-1γ in reprogramming to pluripotency.* Nature cell biology, 2013. **15**(7): p. 872-882.
96. Liang, G., J. He, and Y. Zhang, *Kdm2b promotes induced pluripotent stem cell generation by facilitating gene activation early in reprogramming.* Nature cell biology, 2012. **14**(5): p. 457-466.
97. Eram, M.S., et al., *Kinetic characterization of human histone H3 lysine 36 methyltransferases, ASH1L and SETD2.* Biochim Biophys Acta, 2015. **1850**(9): p. 1842-8.
98. Barrand, S., I.S. Andersen, and P. Collas, *Promoter-exon relationship of H3 lysine 9, 27, 36 and 79 methylation on pluripotency-associated genes.* Biochem Biophys Res Commun, 2010. **401**(4): p. 611-7.
99. Liang, G., J. He, and Y. Zhang, *Kdm2b promotes induced pluripotent stem cell generation by facilitating gene activation early in reprogramming.* Nat Cell Biol, 2012. **14**(5): p. 457-66.
100. Faber, P.W., et al., *Huntingtin interacts with a family of WW domain proteins.* Hum Mol Genet, 1998. **7**(9): p. 1463-74.

101. Mao, M., et al., *Identification of genes expressed in human CD34(+) hematopoietic stem/progenitor cells by expressed sequence tags and efficient full-length cDNA cloning.* Proc Natl Acad Sci U S A, 1998. **95**(14): p. 8175-80.
102. Zhang, Q.H., et al., *Cloning and functional analysis of cDNAs with open reading frames for 300 previously undefined genes expressed in CD34+ hematopoietic stem/progenitor cells.* Genome Res, 2000. **10**(10): p. 1546-60.
103. Nagase, T., et al., *Prediction of the coding sequences of unidentified human genes. XIX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.* DNA Res, 2000. **7**(6): p. 347-55.
104. Rega, S., et al., *Identification of the full-length huntingtin- interacting protein p231HBP/HYPB as a DNA-binding factor.* Mol Cell Neurosci, 2001. **18**(1): p. 68-79.
105. Sun, X.J., et al., *Identification and characterization of a novel human histone H3 lysine 36-specific methyltransferase.* J Biol Chem, 2005. **280**(42): p. 35261-71.
106. Edmunds, J.W., L.C. Mahadevan, and A.L. Clayton, *Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation.* EMBO J, 2008. **27**(2): p. 406-20.
107. Ibanez, G., et al., *A high throughput scintillation proximity imaging assay for protein methyltransferases.* Comb Chem High Throughput Screen, 2012. **15**(5): p. 359-71.
108. Sudol, M., et al., *Characterization of a novel protein-binding module--the WW domain.* FEBS Lett, 1995. **369**(1): p. 67-71.
109. Macias, M.J., et al., *Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide.* Nature, 1996. **382**(6592): p. 646-9.
110. Rebehmed, J., et al., *Expanding the SRI domain family: a common scaffold for binding the phosphorylated C-terminal domain of RNA polymerase II.* FEBS Lett, 2014. **588**(23): p. 4431-7.
111. Youdell, M.L., et al., *Roles for Ctk1 and Spt6 in regulating the different methylation states of histone H3 lysine 36.* Mol Cell Biol, 2008. **28**(16): p. 4915-26.
112. Yuan, W., et al., *Heterogeneous nuclear ribonucleoprotein L Is a subunit of human KMT3a/Set2 complex required for H3 Lys-36 trimethylation activity in vivo.* J Biol Chem, 2009. **284**(23): p. 15701-7.
113. Gao, Y.G., et al., *Autoinhibitory structure of the WW domain of HYPB/SETD2 regulates its interaction with the proline-rich region of huntingtin.* Structure, 2014. **22**(3): p. 378-86.

114. Gibson, M.D., et al., *PHF1 Tudor and N-terminal domains synergistically target partially unwrapped nucleosomes to increase DNA accessibility.* Nucleic Acids Res, 2017. **45**(7): p. 3767-3776.
115. Li, H., et al., *Polycomb-like proteins link the PRC2 complex to CpG islands.* Nature, 2017. **549**(7671): p. 287-291.
116. Yoh, S.M., J.S. Lucas, and K.A. Jones, *The Iws1:Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation.* Genes Dev, 2008. **22**(24): p. 3422-34.
117. Zhu, K., et al., *SPOP-containing complex regulates SETD2 stability and H3K36me3-coupled alternative splicing.* Nucleic Acids Res, 2017. **45**(1): p. 92-105.
118. Ebmeier, C.C., et al., *Human TFIIH Kinase CDK7 Regulates Transcription-Associated Chromatin Modifications.* Cell Rep, 2017. **20**(5): p. 1173-1186.
119. Zhang, L.Q. and Q.Z. Li, *Estimating the effects of transcription factors binding and histone modifications on gene expression levels in human cells.* Oncotarget, 2017. **8**(25): p. 40090-40103.
120. Ferrari, K.J., et al., *Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity.* Mol Cell, 2014. **53**(1): p. 49-62.
121. Carrozza, M.J., et al., *Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription.* Cell, 2005. **123**(4): p. 581-92.
122. Carvalho, S., et al., *Histone methyltransferase SETD2 coordinates FACT recruitment with nucleosome dynamics during transcription.* Nucleic Acids Res, 2013. **41**(5): p. 2881-93.
123. Baubec, T., et al., *Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation.* Nature, 2015. **520**(7546): p. 243-7.
124. Neri, F., et al., *Intragenic DNA methylation prevents spurious transcription initiation.* Nature, 2017. **543**(7643): p. 72-77.
125. de Almeida, S.F., et al., *Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36.* Nat Struct Mol Biol, 2011. **18**(9): p. 977-83.
126. Sen, A., et al., *Smooth, an hnRNP-L Homolog, Might Decrease Mitochondrial Metabolism by Post-Transcriptional Regulation of Isocitrate Dehydrogenase (Idh) and Other Metabolic Genes in the Sub-Acute Phase of Traumatic Brain Injury.* Front Genet, 2017. **8**: p. 175.

127. Chang, W.H., et al., *Modulation the alternative splicing of GLA (IVS4+919G>A) in Fabry disease.* PLoS One, 2017. **12**(4): p. e0175929.
128. Cheng, T.L., et al., *Regulation of mRNA splicing by MeCP2 via epigenetic modifications in the brain.* Sci Rep, 2017. **7**: p. 42790.
129. Gatto, S., et al., *ICF-specific DNMT3B dysfunction interferes with intragenic regulation of mRNA transcription and alternative splicing.* Nucleic Acids Res, 2017. **45**(10): p. 5739-5756.
130. Chen, K., et al., *Methyltransferase SETD2-Mediated Methylation of STAT1 Is Critical for Interferon Antiviral Activity.* Cell, 2017. **170**(3): p. 492-506 e14.
131. Park, I.Y., et al., *Dual Chromatin and Cytoskeletal Remodeling by SETD2.* Cell, 2016. **166**(4): p. 950-962.
132. Diao, Y.F., et al., *Dynamic changes of SETD2, a histone H3K36 methyltransferase, in porcine oocytes, IVF and SCNT embryos.* PLoS One, 2018. **13**(2): p. e0191816.
133. Yu, X.X., et al., *Ascorbic acid induces global epigenetic reprogramming to promote meiotic maturation and developmental competence of porcine oocytes.* Sci Rep, 2018. **8**(1): p. 6132.
134. Zuo, X., et al., *The histone methyltransferase Setd2 is required for expression of acrosin-binding protein 1and protamines and essential for spermiogenesis in mice.* J Biol Chem, 2018.
135. Hu, M., et al., *Histone H3 lysine 36 methyltransferase Hypb/Setd2 is required for embryonic vascular remodeling.* Proc Natl Acad Sci U S A, 2010. **107**(7): p. 2956-61.
136. Zhang, Y., et al., *H3K36 histone methyltransferase Setd2 is required for murine embryonic stem cell differentiation toward endoderm.* Cell Rep, 2014. **8**(6): p. 1989-2002.
137. Zhang, Y.L., et al., *Setd2 deficiency impairs hematopoietic stem cell self-renewal and causes malignant transformation.* Cell Res, 2018. **28**(4): p. 476-490.
138. Zhou, Y., et al., *Setd2 regulates quiescence and differentiation of adult hematopoietic stem cells by restricting RNA polymerase II elongation.* Haematologica, 2018.
139. Ohhata, T., et al., *Histone H3 Lysine 36 Trimethylation Is Established over the Xist Promoter by Antisense Tsix Transcription and Contributes to Repressing Xist Expression.* Mol Cell Biol, 2015. **35**(22): p. 3909-20.

140. Lu, C., et al., *Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape.* Science, 2016. **352**(6287): p. 844-9.
141. Nohr, E., et al., *Diagnostic value of histone 3 mutations in osteoclast-rich bone tumors.* Hum Pathol, 2017. **68**: p. 119-127.
142. Bjerke, L., et al., *Histone H3.3. mutations drive pediatric glioblastoma through upregulation of MYCN.* Cancer Discov, 2013. **3**(5): p. 512-9.
143. Zhuang, L., et al., *Depletion of Nsd2-mediated histone H3K36 methylation impairs adipose tissue development and function.* Nat Commun, 2018. **9**(1): p. 1796.
144. Kernohan, K.D., et al., *H3.1 K36M mutation in a congenital-onset soft tissue neoplasm.* Pediatr Blood Cancer, 2017. **64**(12).
145. Clouaire, T. and G. Legube, *DNA double strand break repair pathway choice: a chromatin based decision?* Nucleus, 2015. **6**(2): p. 107-13.
146. Aymard, F., et al., *Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks.* Nat Struct Mol Biol, 2014. **21**(4): p. 366-74.
147. Bhattacharjee, P., et al., *Epigenetic alteration of mismatch repair genes in the population chronically exposed to arsenic in West Bengal, India.* Environ Res, 2018. **163**: p. 289-296.
148. Carvalho, S., et al., *SETD2 is required for DNA double-strand break repair and activation of the p53-mediated checkpoint.* Elife, 2014. **3**: p. e02482.
149. Pfister, S.X., et al., *SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability.* Cell Rep, 2014. **7**(6): p. 2006-18.
150. Li, L. and Y. Wang, *Cross-talk between the H3K36me3 and H4K16ac histone epigenetic marks in DNA double-strand break repair.* J Biol Chem, 2017. **292**(28): p. 11951-11959.
151. Mar, B.G., et al., *SETD2 alterations impair DNA damage recognition and lead to resistance to chemotherapy in leukemia.* Blood, 2017. **130**(24): p. 2631-2641.
152. Zhang, P., et al., *Structure of human MRG15 chromo domain and its binding to Lys36-methylated histone H3.* Nucleic Acids Res, 2006. **34**(22): p. 6621-8.
153. Bleuyard, J.Y., et al., *MRG15-mediated tethering of PALB2 to unperturbed chromatin protects active genes from genotoxic stress.* Proc Natl Acad Sci U S A, 2017. **114**(29): p. 7671-7676.

154. Li, F., et al., *The histone mark H3K36me3 regulates human DNA mismatch repair through its interaction with MutSalpha.* Cell, 2013. **153**(3): p. 590-600.
155. Yamamoto, H. and K. Imai, *Microsatellite instability: an update.* Arch Toxicol, 2015. **89**(6): p. 899-921.
156. Wong, M.C.S., et al., *Incidence and mortality of kidney cancer: temporal patterns and global trends in 39 countries.* Sci Rep, 2017. **7**(1): p. 15698.
157. Liao, L., J.R. Testa, and H. Yang, *The roles of chromatin-remodelers and epigenetic modifiers in kidney cancer.* Cancer Genet, 2015. **208**(5): p. 206-14.
158. Bi, M., et al., *Genomic characterization of sarcomatoid transformation in clear cell renal cell carcinoma.* Proc Natl Acad Sci U S A, 2016. **113**(8): p. 2170-5.
159. Ho, T.H., et al., *Loss of histone H3 lysine 36 trimethylation is associated with an increased risk of renal cell carcinoma-specific death.* Mod Pathol, 2016. **29**(1): p. 34-42.
160. Cancer Genome Atlas Research, N., *Comprehensive molecular characterization of clear cell renal cell carcinoma.* Nature, 2013. **499**(7456): p. 43-9.
161. Kanu, N., et al., *SETD2 loss-of-function promotes renal cancer branched evolution through replication stress and impaired DNA repair.* Oncogene, 2015. **34**(46): p. 5699-708.
162. Liu, W., et al., *Decreased Expression of SETD2 Predicts Unfavorable Prognosis in Patients With Nonmetastatic Clear-Cell Renal Cell Carcinoma.* Medicine (Baltimore), 2015. **94**(45): p. e2004.
163. Piva, F., et al., *Computational analysis of the mutations in BAP1, PBRM1 and SETD2 genes reveals the impaired molecular processes in renal cell carcinoma.* Oncotarget, 2015. **6**(31): p. 32161-8.
164. Newbold, R.F. and K. Mokbel, *Evidence for a tumour suppressor function of SETD2 in human breast cancer: a new hypothesis.* Anticancer Res, 2010. **30**(9): p. 3309-11.
165. Patani, N., et al., *Histone-modifier gene expression profiles are associated with pathological and clinical outcomes in human breast cancer.* Anticancer Res, 2011. **31**(12): p. 4115-25.
166. Kim, J.Y., et al., *Genetic and Clinical Characteristics of Phyllodes Tumors of the Breast.* Transl Oncol, 2017. **11**(1): p. 18-23.

167. Piscuoglio, S., et al., *Massively parallel sequencing of phyllodes tumours of the breast reveals actionable mutations, and TERT promoter hotspot mutations and TERT gene amplification as likely drivers of progression.* J Pathol, 2016. **238**(4): p. 508-18.
168. Fontebasso, A.M., et al., *Mutations in SETD2 and genes affecting histone H3K36 methylation target hemispheric high-grade gliomas.* Acta Neuropathol, 2013. **125**(5): p. 659-69.
169. Nomura, M., et al., *Distinct molecular profile of diffuse cerebellar gliomas.* Acta Neuropathol, 2017. **134**(6): p. 941-956.
170. Vanan, M.I. and D.D. Eisenstat, *Management of high-grade gliomas in the pediatric patient: Past, present, and future.* Neurooncol Pract, 2014. **1**(4): p. 145-157.
171. Stadtfeld, M., et al., *Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse.* Cell Stem Cell, 2008. **2**(3): p. 230-40.
172. Li, R., et al., *A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts.* Cell Stem Cell, 2010. **7**(1): p. 51-63.
173. Samavarchi-Tehrani, P., et al., *Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming.* Cell Stem Cell, 2010. **7**(1): p. 64-77.
174. Brambrink, T., et al., *ES cells derived from cloned and fertilized blastocysts are transcriptionally and functionally indistinguishable.* Proc Natl Acad Sci U S A, 2006. **103**(4): p. 933-8.
175. Marion, R.M., et al., *Telomeres acquire embryonic stem cell characteristics in induced pluripotent stem cells.* Cell Stem Cell, 2009. **4**(2): p. 141-54.
176. Bernstein, B.E., et al., *A bivalent chromatin structure marks key developmental genes in embryonic stem cells.* Cell, 2006. **125**(2): p. 315-26.
177. Meshorer, E., et al., *Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells.* Dev Cell, 2006. **10**(1): p. 105-16.
178. Golipour, A., et al., *A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network.* Cell Stem Cell, 2012. **11**(6): p. 769-82.
179. Polo, J.M., et al., *A molecular roadmap of reprogramming somatic cells into iPS cells.* Cell, 2012. **151**(7): p. 1617-32.
180. Mattout, A., A. Biran, and E. Meshorer, *Global epigenetic changes during somatic cell reprogramming to iPS cells.* J Mol Cell Biol, 2011. **3**(6): p. 341-50.

181. Koche, R.P., et al., *Reprogramming factor expression initiates widespread targeted chromatin remodeling.* Cell Stem Cell, 2011. **8**(1): p. 96-105.
182. Ang, Y.S., et al., *Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network.* Cell, 2011. **145**(2): p. 183-97.
183. Mansour, A.A., et al., *The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming.* Nature, 2012. **488**(7411): p. 409-13.
184. Costa, Y., et al., *NANOG-dependent function of TET1 and TET2 in establishment of pluripotency.* Nature, 2013. **495**(7441): p. 370-4.
185. Wang, T., et al., *The histone demethylases Jhdm1a/1b enhance somatic cell reprogramming in a vitamin-C-dependent manner.* Cell Stem Cell, 2011. **9**(6): p. 575-87.
186. Buckley, S.M., et al., *Regulation of pluripotency and cellular reprogramming by the ubiquitin-proteasome system.* Cell Stem Cell, 2012. **11**(6): p. 783-98.
187. Chen, J., et al., *H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs.* Nat Genet, 2013. **45**(1): p. 34-42.
188. Liu, Y., et al., *Reprogramming of MLL-AF9 leukemia cells into pluripotent stem cells.* Leukemia, 2014. **28**(5): p. 1071-1080.
189. Baek, S., et al., *Electromagnetic Fields Mediate Efficient Cell Reprogramming into a Pluripotent State.* ACS Nano, 2014. **8**(10): p. 10125-10138.
190. Goyal, A., S.L. Chavez, and R.A. Reijo Pera, *Generation of human induced pluripotent stem cells using epigenetic regulators reveals a germ cell-like identity in partially reprogrammed colonies.* PLoS One, 2013. **8**(12): p. e82838.
191. Nagamatsu, G., et al., *Optimal Ratio of Transcription Factors for Somatic Cell Reprogramming.* The Journal of Biological Chemistry, 2012. **287**(43): p. 36273-36282.
192. Rao, R.A., et al., *Ezh2 mediated H3K27me3 activity facilitates somatic transition during human pluripotent reprogramming.* Scientific Reports, 2015. **5**: p. 8229.
193. Rayasam, G.V., et al., *NSD1 is essential for early post-implantation development and has a catalytically active SET domain.* The EMBO Journal, 2003. **22**(12): p. 3153-3163.
194. Sun, H., et al., *Lysine-specific histone demethylase 1 inhibition promotes reprogramming by facilitating the expression of exogenous transcriptional factors and metabolic switch.* Scientific Reports, 2016. **6**: p. 30903.

195. Wang, W.-P., et al., *The EP300, KDM5A, KDM6A and KDM6B Chromatin Regulators Cooperate with KLF4 in the Transcriptional Activation of POU5F1.* PLOS ONE, 2012. **7**(12): p. e52556.
196. Kidder, B.L., et al., *Extended Self-Renewal and Accelerated Reprogramming in the Absence of Kdm5b.* Molecular and Cellular Biology, 2013. **33**(24): p. 4793-4810.
197. Chen, J., et al., *H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs.* Nature Genetics, 2012. **45**: p. 34.
198. Wei, J., et al., *KDM4B-mediated reduction of H3K9me3 and H3K36me3 levels improves somatic cell reprogramming into pluripotency.* Scientific Reports, 2017. **7**(1): p. 7514.
199. Ciccone, D.N., et al., *KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints.* Nature, 2009. **461**: p. 415.
200. Mansour, A.A., et al., *The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming.* Nature, 2012. **488**: p. 409.
201. Hanahan, D., J. Jessee, and F.R. Bloom, *Plasmid transformation of Escherichia coli and other bacteria.* Methods Enzymol, 1991. **204**: p. 63-113.
202. Jozefczuk, J., K. Drews, and J. Adjaye, *Preparation of mouse embryonic fibroblast cells suitable for culturing human embryonic and induced pluripotent stem cells.* J Vis Exp, 2012(64).
203. Onder, T.T., et al., *Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways.* Cancer Res, 2008. **68**(10): p. 3645-54.
204. Park, I.H., et al., *Reprogramming of human somatic cells to pluripotency with defined factors.* Nature, 2008. **451**(7175): p. 141-6.
205. Schmittgen, T.D. and K.J. Livak, *Analyzing real-time PCR data by the comparative C(T) method.* Nat Protoc, 2008. **3**(6): p. 1101-8.
206. Mahmood, T. and P.C. Yang, *Western blot: technique, theory, and trouble shooting.* N Am J Med Sci, 2012. **4**(9): p. 429-34.
207. Cribbs, A., et al., *Inhibition of histone H3K27 demethylases selectively modulates inflammatory phenotypes of natural killer cells.* J Biol Chem, 2018. **293**(7): p. 2422-2437.
208. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15545-50.

209. Yuan, H., et al., *Histone methyltransferase SETD2 modulates alternative splicing to inhibit intestinal tumorigenesis.* J Clin Invest, 2017. **127**(9): p. 3375-3391.
210. Chan, E.M., et al., *Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells.* Nat Biotechnol, 2009. **27**(11): p. 1033-7.
211. Li, J., et al., *Functional Studies on Primary Tubular Epithelial Cells Indicate a Tumor Suppressor Role of SETD2 in Clear Cell Renal Cell Carcinoma.* Neoplasia, 2016. **18**(6): p. 339-46.
212. Chen, Z., et al., *SETD2 indicates favourable prognosis in gastric cancer and suppresses cancer cell proliferation, migration, and invasion.* Biochem Biophys Res Commun, 2018. **498**(3): p. 579-585.
213. Fellmann, C., et al., *An optimized microRNA backbone for effective single-copy RNAi.* Cell Rep, 2013. **5**(6): p. 1704-13.
214. Hotta, A. and J. Ellis, *Retroviral vector silencing during iPS cell induction: an epigenetic beacon that signals distinct pluripotent states.* J Cell Biochem, 2008. **105**(4): p. 940-8.
215. Lewis, P.W., et al., *Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma.* Science, 2013. **340**(6134): p. 857-61.
216. Fang, D., et al., *The histone H3.3K36M mutation reprograms the epigenome of chondroblastomas.* Science, 2016. **352**(6291): p. 1344-8.
217. Lee, K.C., W.K. Wong, and B. Feng, *Decoding the Pluripotency Network: The Emergence of New Transcription Factors.* Biomedicines, 2013. **1**(1): p. 49-78.
218. van den Hurk, M., et al., *Transcriptional and epigenetic mechanisms of cellular reprogramming to induced pluripotency.* Epigenomics, 2016. **8**(8): p. 1131-49.
219. Qin, H., et al., *Epigenetic Control of Reprogramming and Transdifferentiation by Histone Modifications.* Stem Cell Rev, 2016. **12**(6): p. 708-720.
220. Johnson, D.W., et al., *Assignment of human transforming growth factor-beta type I and type III receptor genes (TGFBR1 and TGFBR3) to 9q33-q34 and 1p32-p33, respectively.* Genomics, 1995. **28**(2): p. 356-7.
221. Maherali, N. and K. Hochedlinger, *Tgfbeta signal inhibition cooperates in the induction of iPSCs and replaces Sox2 and cMyc.* Curr Biol, 2009. **19**(20): p. 1718-23.
222. Tan, F., et al., *Inhibition of transforming growth factor beta (TGF-beta) signaling can substitute for Oct4 protein in reprogramming and maintain pluripotency.* J Biol Chem, 2015. **290**(7): p. 4500-11.

223. Kuno, A., K. Nishimura, and S. Takahashi, *Time-course transcriptome analysis of human cellular reprogramming from multiple cell types reveals the drastic change occurs between the mid phase and the late phase.* BMC Genomics, 2018. **19**(1): p. 9.
224. Kurpinski, K., et al., *Transforming growth factor-beta and notch signaling mediate stem cell differentiation into smooth muscle cells.* Stem Cells, 2010. **28**(4): p. 734-42.
225. Ruetz, T., et al., *Constitutively Active SMAD2/3 Are Broad-Scope Potentiators of Transcription-Factor-Mediated Cellular Reprogramming.* Cell Stem Cell, 2017. **21**(6): p. 791-805 e9.
226. Dahle, O. and M.R. Kuehn, *Polycomb determines responses to smad2/3 signaling in embryonic stem cell differentiation and in reprogramming.* Stem Cells, 2013. **31**(8): p. 1488-97.
227. Hoofnagle, M.H., et al., *Myocardin is differentially required for the development of smooth muscle cells and cardiomyocytes.* Am J Physiol Heart Circ Physiol, 2011. **300**(5): p. H1707-21.
228. Marchand, M., et al., *Transcriptomic signature of trophoblast differentiation in a human embryonic stem cell model.* Biol Reprod, 2011. **84**(6): p. 1258-71.
229. King, M., et al., *T-genes and limb bud development.* Am J Med Genet A, 2006. **140**(13): p. 1407-13.
230. Pfannkuche, K., et al., *Cardiac myocytes derived from murine reprogrammed fibroblasts: intact hormonal regulation, cardiac ion channel expression and development of contractility.* Cell Physiol Biochem, 2009. **24**(1-2): p. 73-86.
231. Festuccia, N., et al., *Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells.* Cell Stem Cell, 2012. **11**(4): p. 477-90.
232. Qin, S. and J. Min, *Structure and function of the nucleosome-binding PWWP domain.* Trends Biochem Sci, 2014. **39**(11): p. 536-47.
233. Wen, H., et al., *ZMYND11 links histone H3.3K36me3 to transcription elongation and tumour suppression.* Nature, 2014. **508**(7495): p. 263-8.
234. Pawlak, M. and R. Jaenisch, *De novo DNA methylation by Dnmt3a and Dnmt3b is dispensable for nuclear reprogramming of somatic cells to a pluripotent state.* Genes Dev, 2011. **25**(10): p. 1035-40.

Vita

Burcu Özçimen was born in Konya on January 14, 1984. She received her BSc degree in Molecular Biology and Genetics from Izmir Institute of Technology in 2010. She attended to the Master of Science program in Medical Biology and Genetics at Dokuz Eylül University between 2010 and 2013. She was accepted to the PhD program in Molecular Biology and Genetics in 2013 at Koç University. Her research interests included stem cell biology and epigenetics.