Identification of MRG15 (MORFL1) as a barrier to somatic cell reprogramming

by

Elifsu Kartal

A Dissertation Submitted to the
Graduate School of Health Sciences
in Partial Fulfillment of the Requirements for
the Degree of

Master of Science

in

Cellular and Molecular Medicine

KOÇ
ÜNİVERSİTESİ

Identification of MRG15 (MORFL1) as a barrier to somatic cell reprogramming

Koç University

Graduate School of Health Sciences

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

Elifsu Kartal

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:

_____

Prof. Tamer Önder (Advisor)

_____

Assoc. Prof Gözde Korkmaz

_____

Assoc. Prof Perinur Bozaykut Eker

Date: _____

To my grandfather in heaven Mehmet Kartal, whose steps I followed in laboratory.

ABSTRACT

Identification of MRG15 (MORFL1) as a barrier to somatic cell reprogramming

Elifsu Kartal

Master of Science in Cellular Molecular Medicine

01.22.2024

Reprogramming somatic cells into induced pluripotent stem cells (iPSCs) is specified by the expression of OCT4, SOX2, KLF4, and c-MYC transcription factors. However, reprogramming has a limited success rate, suggesting the presence of barriers to this conversion. Epigenetic mechanisms such as DNA methylation and post-translational modifications on histone tails can hinder reprogramming. Preliminary studies demonstrated that SETD2, H3 lysine 36 tri-methyltransferase hinders cellular reprogramming. A CRISPR/Cas9 screen on H3K36me3-reader proteins yielded promising results in reprogramming efficiency, specifically with MRG15. In this thesis, I investigated the molecular processes behind MRG15's role as a barrier to iPSC production.

First, I examined the effect of a SETD2 inhibitor in cellular reprogramming. This inhibitor selectively inhibited SETD2 activity, decreased H3K36me3 levels, and, similar to SETD2 knockdown, increased reprogramming efficiency. Next, I tested MRG15 knockout fibroblasts for reprogramming efficiency, and it increased iPSC generation four-fold. Overexpression of wildtype MRG15 rescued the knockout phenotype, confirming that the increase in reprogramming efficiency is not caused by Cas9 off-target events. I hypothesized that MRG15 acts downstream of SETD2 to block cellular reprogramming. However, double knockout of MRG15 and SETD2 yielded an additive increase in iPSC generation, which suggests MRG15 is not a direct SETD2 downstream effector.

Next, I created MRG15 knockout iPSCs and examined OCT4, SOX2, NANOG, and SSEA4 pluripotency markers. Results indicated that iPSC maintenance does not require MRG15. RNA sequencing was performed with MRG15 knockout fibroblast samples, and gene set enrichment analysis demonstrated that MRG15 knockout samples are significantly enriched in pluripotent gene sets. Additionally, MRG15 knock-out upregulated cell cycle-related gene set expression, which may indicate that MRG15 knockout enhances iPSC production through higher proliferation.

MRG15 can also be found in the SIN3B/HDAC2 repressive histone deacetylation complex, therefore, in a final set of experiments, I observed that MRG15 knockout led to a four-fold increase in H3K14 acetylation.

Taken together, these results demonstrate that MRG15 and SETD2 operate as obstacles to cellular reprogramming through separate mechanisms; inhibiting these factors enhances the currently poor efficiency of generating iPSCs.

ÖZETÇE

Somatik Hücre Yeniden Programlamaya Engel Olarak MRG15 (MORFL1) Tanımlanması

Elifsu Kartal

Hücresel ve Moleküler Tıp, Yüksek Lisans

01.22.2024

Hücresel programlamayla somatik hücreleri uyarılmış pluripotent kök hücrelere (UPKH) dönüştürmek, OCT4, SOX2, KLF4 ve c-MYC transkripsiyon faktörlerinin ifadesi tarafından belirlenir. Ancak, programlamanın sınırlı bir başarı oranı vardır, bu da bu dönüşüme engellerin varlığını gösterir. DNA metilasyonu ve histon kuyruklarındaki translasyonel sonrası modifikasyonlar gibi epigenetik mekanizmalar, programlamayı engelleyebilir. Ön çalışmalar, SETD2, H3 lizin 36 tri-metiltransferazının hücresel programlamayı engellediğini göstermiştir. H3K36me3-okuyucu proteinler üzerinde yapılan bir CRISPR/Cas9 taraması, özellikle MRG15 ile birlikte programlama verimliliğinde umut verici sonuçlar vermiştir. Bu tezde, MRG15'in iPSC üretimindeki rolünü araştırdım.

İlk olarak, SETD2 inhibitörünün hücresel programlama üzerindeki etkisini inceledim. Bu inhibitör, selektif olarak SETD2 aktivitesini engelledi, H3K36me3 seviyelerini azalttı ve SETD2 susturulmuş hücrelerinde benzer şekilde programlama verimliliğini artırdı. Sonra, MRG15 silinmiş fibroblastlar üzerinde programlama verimliliğini test ettim ve UPKH üretimini dört kat artırdığını gözlemledim. Doğal fenotip MRG15'in aşırı ifadesi, UPKH yapılmış fenotipi kurtardı ve bu da yeniden proglamanın verimliliğinin Cas9 hedef dışı olaylarından kaynaklanmadığını doğruladı. MRG15'in SETD2'nin alt akışında hücresel programlamayı engelleyen bir faktör olduğunu hipotez ettim. Ancak, MRG15 ve SETD2 susturulmuş hücrelerde, UPKH üretiminde a ek bir artış gözlemledim. Bu da MRG15'in doğrudan bir SETD2 alt akış etkileyeni olmadığını gösteriyor.

Daha sonra, MRG15 silinmiş UPKH'leri oluşturdum ve OCT4, SOX2, NANOG, ve SSEA4 pluripotensi işaretleyicilerini inceledim. Sonuçlar, UPKH bakımının MRG15'e ihtiyaç duymadığını gösterdi. MRG15 silinmiş fibroblast örnekleri ile RNA sekanslama yapıldı ve gen seti zenginleştirme analizi, MRG15 silinmiş örneklerinin pluripotent gen setlerinde önemli ölçüde zenginleştiğini gösterdi. Ayrıca, MRG15 silinmiş, hücre döngüsü ile ilgili gen seti ifadesini yükseltti, bu da MRG15 silinmiş UPKH üretimini daha yüksek çoğalma aracılığıyla artırdığını gösterebilir.

MRG15 aynı zamanda SIN3B/HDAC2 baskılayıcı histon deasetilasyon kompleksinde bulunabilir, bu nedenle son bir dizi deneyde MRG15 silinmiş hücrelerde H3K14 asetilasyonunda dört kat artışa neden olduğunu gözlemledim.

Bu sonuçlar, MRG15 ve SETD2'nin, şu anda zayıf olan UPKH üretim verimliliğini artırmak için ayrı mekanizmalarla hücresel programlamaya engel olduğunu göstermektedir.

ACKNOWLEDGEMENTS

I would like to start by thanking my supervisor, Prof.Dr.Tamer Önder. I was honored when Dr.Önder accepted to work with me. I have tried to be a great student; however, sometimes, I lose my motivation. You always see me as a promising student who has yet to achieve her full potential. Thank you for allowing me to study this project.

I also thank Assoc. Prof. Perinur Bozaykut and Prof. Gözde Korkmaz, members of my thesis committee, for agreeing to serve as my jury members and for their helpful and encouraging feedback on my findings and potential directions for the study.

I gratefully acknowledge the use of the services and facilities of the Koç University Research Center for Translational Medicine (KUTTAM), funded by the Presidency of Turkey, Head of Strategy and Budget.

I thank Önder Lab members and alums for their support and help throughout my Master's. I would like to thank Nilya Karasürmeli for presenting her friendship at all times. I miss singing with you when we work in the cell culture. How often have I come to you for endless technical problems, and you helped me right then and there? Next, as promised, I would like to thank one of the building blocks of Önder lab, Arda Odabaş. He is the ResearchGate of the lab and KUTTAM and knows everything related to molecular biology. Like Nilya, countless times you helped me without asking for anything in return, and I will be forever grateful for that. Outside the lab, I had many car rides with you and Canan, greatly improving my days. Also, Canan, thank you for the coffee breaks, which I enjoyed and greatly appreciated during the last months of my Masters. In addition, I would like to thank Mükrime Altun for our office chats across our desks. We shared our snacks and problems. I last member of the Önder lab I want to thank is, Zeynep Umay Gürbüz. You were a great teammate, and I enjoyed working with you on this project and being your supervisor. You are the only person with the same desperation and sarcasm as me in the lab, so I believe we shared good laughs.

I would like to Sevilay Münire Girgin aka SMG lab. It is hard to describe your place in my heart and life and might require another thesis. I came here because you prepared a tour of Koç University; I could start my studies because you opened your house to me.

But this was just the beginning; not only did you open your home, but you opened your life to me. You are one of the main reasons I could finish this degree since you never let me give up. I have always admired you as a scientist and hope to reach your level someday. Countless times, we sit on the engineering court and look up to the sky. At that time, we were gloomy and stressed, but I always knew the time we shared was special, and even being gloomy with you was something I enjoyed. We have regrets about some of our decisions, but I never regret coming here and building a life with you. Our living room in Kiptaş holds many laughs, deep conversations, and cries; I will cherish them forever.

Finally, I would like to thank my family, my mother, Sevilay Karlı Kartal, my father, Zafer Kartal, and my fiancée, Onur Ciddi, for their endless support and love. I am so grateful to have you as my family.

TABLE OF CONTENTS

Chapter 1: Introduction 1

1.1 Mammalian Embryogenesis 1

1.2 Stem Cells 2

1.3 Somatic Cell Reprogramming 2

1.4 Roadblocks Against iPSC Generation 4

1.5 Epigenetic changes upon Reprogramming 4

1.6 Chromatin Associated Factors & Histone Modifications 5

1.7 H3K36me3 mark and its only catalyzer SETD2 7

1.8 MRG15 structure 8

1.9 Discovery of MRG15 and MORF family 10

1.10 MRG15 homologs among species 11

1.11 DNA repair mechanism by MRG15 12

1.12 Alternative Splicing Regulation by MRG15 13

1.13 MRG15 is part of HAT and HDAC complexes 14

1.14 The main question addressed 15

Chapter 2: METHODS 17

2.1 Cloning of sgRNA to lentiCRISPR_v2_blast 17

2.2 Cloning of Pam Mutant MRG15 into PWZL-BLAST 20

2.3 Viral Packaging 22

2.4 Culture of dH1f 23

2.5 Viral Infection 24

2.6 Epigenetic Reprogramming of Fibroblasts 24

2.7 RNA Isolation & cDNA Synthesis 26

2.8 Quantitive PCR (qPCR) 27

2.9 Whole Cell Protein Isolation 30
2.10 Histone Isolation 30
2.11 Western Blot 31
2.12 Flow Cytometry 33
2.13 Proliferation Assay 34
2.14 Immunofluorescence Staining 34
2.15 RNA- Sequencing and GSEA 35
2.16 Data Analysis 36
Chapter 3: RESULTS 37
3.1 SETD2 inhibitor noticeably diminish H3K36me3 expression whereas increase H3K36me2 levels 37
3.2 SETD2 inhibition increases reprogramming efficiency. 40
3.3 SETD2 inhibition does not enhance proliferation. 42
3.4 Combination of iSETD2 with other inhibitors has increased reprogramming efficiency 43
3.5 MRG15 KO increases reprogramming efficiency 46
3.6 MRG15 KO phenotype is independent of CRISPR/Cas9 off-target effect 47
3.7 Coupling MRG15KO with SETD2 KD has combinatorial effect in terms of reprogramming efficiency 49
3.8 MRG15KO has enriched reprogramming efficiency without Myc expression 50
3.9 MRG15KO yielded an additive effect with a combination of iBRD9 and CBP30 51
3.10 MRG15KO iPSC can be established with transient expression of OSKM factors 52
3.11 MRG15KO iPSC clones express all four pluripotency markers 53
3.12 MRG15 KO changes the whole transcriptome 54
3.13 MRG15KO transcriptome enriched in pluripotency-related gene sets 57

3.14 MRG15KO transcriptome is enriched in cell cycle pathways 59
3.15 MRG15KO does not change the global mRNA splicing profile 61
3.16 MRG15KO cells have increased histone acetylation 63
Chapter 4: DISCUSSION 65
4.1 SETD2 inhibitor can replace SETD2 knockdown 65
4.2 MRG15 and SETD2 act as an epigenetic barrier through different pathways 66
4.3 MRG15 is dispensable for iPSC survival and maintenance 67
4.4 MRG15 regulates alternative splicing towards late reprogramming 68
4.5 MRG15 is responsible for cell cycle regulation 69
4.6 MRG15KO leads hyperacetylated H3K14 70
Bibliography 72

Chapter 1:
INTRODUCTION

1.1 Mammalian Embryogenesis

Despite the vast evolutionary timespan separating mammals, their pre-implantation development processes share striking similarities (Gerri et al., 2020). This journey begins with the fertilization of an egg by a sperm, forming a zygote that attaches to the uterine wall(Gerri et al., 2020). The zygote then divides multiple times without its overall volume increasing. In humans, cell polarization begins at the eight-cell stage, signaling the start of differentiation into distinct cell types(Zhu & Zernicka-Goetz, 2020).

Around five days after fertilization, the embryo forms an early blastocyst, consisting of three essential structures: the trophectoderm, which includes the outer layer and contributes to the placenta; the inner cell mass, which gives rise to the embryo itself; and the blastocoel, a fluid-filled cavity within the blastocyst(Zhu & Zernicka-Goetz, 2020).

The late blastocyst is further characterized by the trophectoderm surrounding the embryo, the primitive endoderm lining the blastocoel, and the epiblast, which sits between the trophectoderm and primitive endoderm and forms the embryo(Zhu & Zernicka-Goetz, 2020).

While the exact mechanisms controlling cell fate commitment within the blastocyst are still being investigated, specific markers have been identified: YAP1 expression is restricted to the trophectoderm. In contrast, NANOG expression is particular to the epiblast(Zhu & Zernicka-Goetz, 2020).

Research on human blastocyst development faces ethical and practical challenges, leading to reliance on animal models, primarily mice. However, it's crucial to note that mice and humans differ in implantation timing and blastocyst morphology, highlighting the need for alternative models(Gerri et al., 2020).

To overcome these challenges, scientists are exploring two promising in vitro approaches: cultivating stem cells that mimic post-implantation stages to study cell fate decisions and morphogenesis and engineering cell culture systems that simulate the uterine environment to directly observe intact embryos during post-implantation development(Gerri et al., 2020).

These advancements offer great potential for unraveling the complexities of early human development, leading to a deeper understanding of developmental disorders and infertility and potentially paving the way for improved reproductive technologies.

1.2 Stem Cells

Stem cells are remarkable, with unique abilities that make them invaluable for research and potential therapies. Their two defining characteristics are self-renewal and differentiation potential(Zakrzewski et al., 2019).Stem cells can replicate themselves indefinitely, maintaining a constant supply of these versatile cells(Lu et al., n.d.). This ability is essential for their role in tissue maintenance and their potential use in regenerative medicine. Stem cells can transform into specialized cell types, such as muscle cells, nerve cells, or blood cells. This remarkable capacity promises to repair damaged tissues and organs, replace lost cells, and potentially cure diseases(Zakrzewski et al., 2019).

However, stem cells are all different in their differentiation potential. They are classified based on the range of cell types they can generate, including totipotent stem cells and pluripotent stem cells(Zakrzewski et al., 2019). These cells possess the most extensive differentiation potential, capable of forming any cell type in the human body, including both embryonic and extraembryonic tissues. Totipotent stem cells are only found in the fertilized egg (zygote) and have not yet been successfully created in the laboratory(Xi et al., 2022).

While not as versatile as totipotent cells, pluripotent stem cells can still differentiate into various cell types. They can form cells from all three germ layers of the developing embryo, which give rise to the multiple tissues and organs of the body(Lu et al., n.d.). Pluripotent stem cells can be found in the inner cell mass of pre-implantation embryos, known as embryonic stem cells (ESCs), or they can be generated in the laboratory through the reprogramming of adult cells, resulting in induced pluripotent stem cells (iPSCs)(Zakrzewski et al., 2019).

1.3 Somatic Cell Reprogramming

For nearly 70 years, scientists have actively investigated the possibility of converting typical somatic cells, which comprise most of our body's tissues, into pluripotent stem

cells, which can transform into various specialized cell types(Srivastava & DeWitt, 2016). This conversion offers a powerful tool for research and potential therapies, and three main methods have emerged to achieve it: somatic cell nuclear transfer (SCNT), cell fusion, and ectopic expression of epigenetic factors(Hanna et al., 2010).

SCNT was first demonstrated by Gurdon et al. in 1958, and it involves transferring the nucleus of a somatic cell into an egg cell with its nucleus removed(Lensch & Mummery, 2013). This process rewinds the developmental clock of the somatic cell, prompting it to revert to an embryonic-like state with pluripotent capabilities(Lensch & Mummery, 2013).

While less frequently used, cell fusion provides a faster route to reprogramming. In this method, scientists merge a stem cell with a somatic cell, creating a hybrid cell known as a heterokaryon, which temporarily has two separate nuclei(Srivastava & DeWitt, 2016). Over a few days, these nuclei fuse, resulting in a cell that has acquired pluripotency. Techniques such as using inactivated Sendai virus, polyethylene glycol, or electrical pulses can facilitate this fusion process(Hanna et al., 2010; Srivastava & DeWitt, 2016).

In 2006, Takahashi and Yamanaka made a groundbreaking discovery by identifying four specific genes, collectively called the "Yamanaka factors" or "Yamanaka cocktail "(Takahashi & Yamanaka, 2006). These genes (OCT4, SOX2, KLF4, and c-MYC) can directly reprogram somatic cells into pluripotent stem cells when introduced into them. This achievement earned Takahashi and Yamanaka the Nobel Prize in Physiology in 2012(Lensch & Mummery, 2013).

Developing these methods for generating iPSCs has opened up exciting possibilities for researchers. It allows them to study cell plasticity (the ability of cells to change their fate), understand the mechanisms that control cell fate decisions, and explore the potential of iPSCs for regenerative medicine, hoping to treat various diseases and repair damaged tissues(Srivastava & DeWitt, 2016). This technology holds immense promise for advancing our knowledge of human development and unlocking new therapeutic approaches.

1.4 Roadblocks Against iPSC Generation

Although using epigenetic factors to achieve iPSCs is a significant achievement, it has limitations, such as the extended time to create pluripotent cells and the unpredictable aspects of certain factors(Vierbuchen & Wernig, 2012). Somatic cell identity and its epigenetic landscape are the leading causes. However, using small molecules such as epigenetic identity disruptors (HDAC, Polycomb, etc.) increases reprogramming efficiency and overall iPSC yield(Hanna et al., 2010). Transcription factor binding sites must be available and open to achieve effective reprogramming (Vierbuchen & Wernig, 2012). However heterochromatic regions are tightly closed by histones impossible for protein binding. Chromatin structure of DNA influences cis- and trans-interactions, therefore regulating cellular identity(Arabacı et al., 2021).

1.5 Epigenetic changes upon Reprogramming

Epigenome is inheritable alterations in chromatin, including histone post-translational modifications and DNA methylation which influence gene expression (Buganim et al., 2013). Epigenome organizations and inheritance during embryogenesis are still being investigated (Buganim et al., 2013). For example, histone & DNA mark inheritance during DNA replication is unclear. Modifications in chromatin structure, nucleosome remodeling, and use of histone variants make the unique epigenetic mark for each cell type. To achieve pluripotency, one must disrupt this epigenetic landmark and shape it into stem cell-like (Buganim et al., 2013).

RNAi and gene expression analysis identified three reprogramming stages: initiation, maturation, and stabilization. In the initiation, stage cells have an increased signal, which induces MET (mesenchymal to epithelial transition)(Buganim et al., 2013). Surprisingly, during maturation and stabilization of reprogramming, OSKM factors are repressed. Also, during later stages, genes related to gonads, gametes, cytoskeletal dynamics, and signaling pathways are upregulated(Buganim et al., 2013).

1.6 Chromatin Associated Factors & Histone Modifications

When a somatic cell transforms into a stem-cell-like state, it is essential to eliminate the epigenetic characteristics of the original somatic cell. This process involves several vital alterations, such as the restructuring of chromatin, the removal of methyl groups from DNA in the promoter regions of crucial pluripotency genes like Nanog, Sox2, and Oct4, the reactivation of the X chromosome, which is usually inactive in somatic cells, and comprehensive restoration of modifications to histone proteins after translation. More than 100 distinct histone post-translational modifications have been identified. Out of these 100, the alterations related to lysine methylation and acetylation have been intensively researched and are considered essential in reprogramming and acquiring a stem-cell-like epigenome. As indicated by previous studies, DNA demethylation and reactivation of the X-chromosome occur towards the end of the reprogramming process.

Conversely, alterations in histone modifications can be observed immediately after the introduction of reprogramming factors. Following the induction of transcription factors, there is a rapid and significant increase in the deposition of newly synthesized histone H3 dimethylated at lysine 4 (H3K4me2) at both promoter and enhancer sites of pluripotency-related genes such as Sal4 and FGF4. A progressive decrease in H3K27me3 levels and promoter hypomethylation in crucial areas were observed in the early stages of reprogramming.

During the initial stages of the reprogramming process, modifications primarily occur in CpG islands. These areas are susceptible to transcription factor activity, allowing changes to happen more quickly. Simultaneously, the promoters of somatic genes experience a decrease in H3K4me2, indicating the early downregulation of MEF (Mouse Embryonic Fibroblasts) markers, such as thymus cell antigen one theta (Thy1) and periostin, osteoblast-specific factor (Postn). Many somatic gene enhancers experience a loss of H3K4me2, resulting in hypermethylation and subsequent silencing throughout later stages. Hence, alterations in the expression of crucial MEF identity factors and early pluripotency genes due to epigenetic modifications may serve as an initial stage in transforming a somatic cell into a pluripotent state.

While histone marks undergo significant changes during reprogramming, it is unclear which chromatin modifiers play a role in changing the epigenomic landscape of somatic

cells and how they target genes with altered expression, which is necessary for conversion(Srivastava & DeWitt, 2016). Epigenetically changed areas are likely marked by OSKM binding sites across the genome. Research suggests that OCT4 interacts with WD repeat protein 5 (WDR5), a significant member of the human Trithorax complex, on pluripotency gene promoters, ensuring global and localized H3K4me3 distribution (Arabacı et al., 2021). The H3K27 demethylase enzyme UTX interacts with OSK (OCT4, SOX2, and KLF4) to remove the restrictive mark H3K27me3 from pluripotency genes such as Fgf4, Sall4, Sall1, and Utf1. Loss of UTX leads to abnormal H3K27me3 distribution and inhibits reprogramming. Nanog can recruit TET1 and TET2, two methylcytosine hydroxylase family members, to boost the expression of specific reprogramming target genes, including Nanog, Esrrb, and Oct4(Arabacı et al., 2021). Research suggests that TET1 and TET2 have a role in demethylating and reactivating genes and regulatory areas crucial for pluripotency. PARP1 is a parallel in establishing early epigenetic marks during somatic cell reprogramming by regulating 5mC modification. The BAF chromatin-remodeling complex, consisting of BRG1 (SMARCA4) and BAF155 (SMARCC1), increases reprogramming by creating a euchromatic chromatin state and boosting factor binding to important gene promoters(Arabacı et al., 2021).

Overexpression of BRG1 and BAF155 leads to OSKM-mediated demethylation of pluripotency genes like Oct4, Nanog, and Rex1 (Zfp42), enhancing iPSC conversion.

Numerous chromatin modifiers play a role in resetting the epigenome of programmable cells. To convert human fibroblasts to iPSCs, the H3K9 methyltransferases EHMT1 and SETDB1 and five components of the Polycomb repressive complexes (PRCs) (BMI1 and RING1 from PRC1, EZH2, EED, and SUZ12 from PRC2) are needed to reset the epigenome of somatic cells. Loss of these genes significantly decreases iPSC formation. Similarly, the H3K79me2 methyl-transferase DOT1L hinders early to middle-phase reprogramming(Onder et al., 2012). It counteracts gene repression.H3K79me3 is linked to transcriptional elongation and is higher in gene bodies of housekeeping and cell type-specific genes in somatic cells(Onder et al., 2012). The barrier function of DOT1L was discovered using a shRNA study of key chromatin regulators in human reprogramming. Inhibiting it genetically or pharmacologically enhances reprogramming efficiency and can replace KLF4 and c-MYC(Onder et al., 2012).

A recent chemical screen on chromatin found new compounds that function with DOT1L inhibitors. Bromodomain inhibitors of CBP and EP300 have been extensively studied. CBP and EP300 are significant proteins with HAT domains and bromodomains used for chromatin binding and histone acetylation recognition(Ebrahimi et al., 2019). These coactivators maintain the activity of cell-type-specific enhancers. Early EP300/CBP bromodomain inhibition enhances reprogramming efficiency and facilitates the efficient generation of human iPSCs with OCT4 and SOX2 when combined with DOT1L inhibitors(Ebrahimi et al., 2019). EP300/CBP bromodomain inhibition reduces histone acetylation and chromatin accessibility at enhancers of somatic-specific factors such as PRRX1. These findings support the idea that the loss of H3K27 acetylation from somatic-specific enhancers is an initial step in reprogramming(Ebrahimi et al., 2019). Complete catalytic inhibition of these two HATs hindered reprogramming and prevented pluripotency gene expression, suggesting a unique role for their bromodomains in maintaining somatic cell identity(Arabacı et al., 2021).

1.7 H3K36me3 mark and its only catalyzer SETD2

H3K36 methylation, a histone alteration that remains constant, governs the arrangement of chromatin and numerous biological processes. H3K36 residue can be mono-, di-, or tri-methylated(Yu et al., 2023). The locations and molecular roles of di- and tri-methylation of H3K36 are varied, but mono-methylation serves as an intermediary level for higher methylation status(Arabacı et al., 2021). H3K36 di-methylation (H3K36me2) is present at intergenic areas of the genome, while H3K36 tri-methylation (H3K36me3) is found in gene bodies actively transcribing(Arabacı et al., 2021). Genome-wide analyses have demonstrated that H3K36me1 and -me2 are predominantly located near the 5'end regions, while H3K36me2 and -me3 exhibit more significant levels towards the 3'ends. Multiple studies have demonstrated that the process of H3K36 methylation is a protective mechanism against the transformation of somatic cells(Arabacı et al., 2021). Ascorbic acid, often Vitamin C, was used as a cofactor for H3K36 demethylases KDM2A/B to enhance the formation of induced pluripotent stem cells (iPSC)(Arabacı et al., 2021). Ascorbic acid also enhanced the expression of cell cycle regulators by demethylating H3K36me2/3. Increased expression of KDM2B, a

demethylase specific to H3K36me2, can improve the effectiveness of reprogramming by early reactivation of genes relevant to pluripotency(Arabacı et al., 2021).

The catalytic domain of these methyltransferases, known as SET (Suppressor of Variegation 3-9, Enhancer of Zeste, and Trithorax), transfers a methyl group to H3K36 by utilizing S-adenosylmethionine(Yu et al., 2023). SETD2 and H3K36me3 primarily facilitate the transcription of actively transcribed euchromatin regions. The study of Set2's SRI domain interaction with phosphorylated CTD of RNAPII in Saccharomyces cerevisiae elucidated the correlation between H3K36 methylation and transcription(Keogh et al., 2005). The interaction between these factors involves the coordination of transcriptional elongation, RNA splicing, and preventing erroneous intragenic transcripts by H3K36 methylation(Carrozza et al., 2005).

SETD2 knockdown abolishes H3K36me3, indicating that it is an essential enzyme responsible for the trimethylation of H3K36, a persistent alteration of histones(Li et al., 2016). Multiple effector proteins identify H3K36me3, a modification associated with alternative splicing, DNA replication and repair, transcriptional repression, DNA methylation, and dosage compensation (Z. Chen et al., 2018). Therefore, these reader proteins facilitate additional SETD2 functions.

While multiple investigations have indicated that H3K36 methylation may impede reprogramming, the specific function of H3K36me3 in this process remains unclear. Given the wide range of interactions and roles that H3K36me3 has in the cell, it is probable that it plays a crucial role in regulating somatic cell reprogramming. Ozcimen (2018) showed that when the gene SETD2 is suppressed, epigenetic reprogramming of human fibroblasts is enhanced. When SETD2 is suppressed, there is a considerable drop in the overall amount of H3K36me3, although other essential chromatin marks remain unchanged (Ozcimen, 2018). These findings indicate that the process of iPSC production is hindered by SETD2-mediated H3K36 tri-methylation.

1.8 MRG15 structure

MRG15, also known as MORF4L1, is one of the reader proteins of H3K36me3. MRG15 has two variant isoform 1 and isoform 2(Bertram et al., 1999). Isoform 2 has an alternate exon (exon 4) in the coding region in-frame and has a longer chromodomain

than isoform 1(Devoucoux et al., 2022) (Figure 1.1). The human MRG15 protein comprises a chromodomain at the N-terminal and a conserved MRG domain at the C-terminal, connected by a flexible region(Xie et al., 2015) (Figure 1.1). The MRG domain is always preserved in all MRG proteins, and the crystal structure of the MRG domain in humans was identified in 2006(Zhang, Du, et al., 2006).

The protein assumes a mostly α-helical shape and is an adapter module facilitating interactions with other proteins in nuclear protein complexes(Xie et al., 2015). Site-directed mutagenesis experiments demonstrate that several hydrophobic amino acids provide a shallow hydrophobic cavity to engage with the N-terminal region of PAM14(Zhang, Zhao, et al., 2006). The precise role of the chromodomain of MRG15 has not been fully comprehended. Nevertheless, the fact that the chromodomain is conserved in numerous MRG15 homologs highlights its significant role in functionality(Devoucoux et al., 2022). Prior biochemical and structural investigations have demonstrated that chromo and chromo-like domains, such as Tudor and PWWP domains, are responsible for recognizing and interacting with histones or other proteins that contain modified residues, such as methylated lysines or arginines, in nucleoprotein complexes like HAT and HDAC complexes(Zhang, Du, et al., 2006). These domains play crucial roles in modifying chromatin structure, which ultimately results in the activation or repression of numerous genes(Zhang, Du, et al., 2006).

The structure of MRG15's chromodomain is found to be more comparable to the chromo barrel domain of Drosophila MOF (dMOF) than the normal HP1/Pc chromo domain(Zhang, Du, et al., 2006). It has been reported that the chromodomain of MRG15 can bind to methylated H3K36. The combined structural and biochemical evidence indicates that the MRG15 chromo domain likely acts as an adapter module, facilitating interaction with a modified histone in a manner distinct from the HP1/Pc chromo domains(Zhang, Du, et al., 2006).

The MRG domain is exclusively present in a few proteins, including all three members of the MORF family and the MSL3 protein(Mccarthy et al., 2022). The MSL3 protein acts as a subunit of the dosage compensation complex, which controls gene expression on the X chromosome(Ae & Pereira-Smith, 2008). MRG domain can interact with many proteins, such as MRGBP, MRFAP1, Pf1, and PALB2, interact. These interactions occur in specific complexes(M. Chen et al., 2009). Crystallographic and nuclear magnetic resonance (NMR) analyses have revealed that the MRG domain

consists primarily of helical structures(Xie et al., 2015). The core of the domain is composed of two helical hairpins oriented at right angles to each other. Previous studies have uncovered a two-part structural pattern in Pf1, consisting of a helix and a segment that contains an FxLP sequence(Kumar et al., 2012). This pattern is in an elongated shape and interacts explicitly with neighboring surfaces of the MRG domain(Kumar et al., 2012). Both structural motifs are crucial for the strong interaction with a binding affinity in the nanomolar range, as each segment contributes equally to the overall binding energy. While all the MRG interactors mentioned above have the FxLP motif in common, a Pf1-like helical motif is not easily identifiable by sequence analysis(Kumar et al., 2012). The article suggests that there may be many processes by which MRG recognizes its targets, indicating the versatility of the MRG domain in engaging with different targets(Kumar et al., 2012).

1.9 Discovery of MRG15 and MORF family

An article published in 1999 by Pereira et al. presents the process of cloning and identifying a gene called mortality factor 4 (MORF 4) located on chromosome 4(Bertram et al., 1999). Introducing MORF4 into two complementation group B cell lines induces a phenotype similar to senescence(Bertram et al., 1999). They have reported that MORF4 belongs to a unique family of genes with patterns identical to transcription factor genes. MORF 4 is a gene that lacks introns and is a pseudogene(Bertram et al., 1999). While the majority of processed pseudogenes are not transcribed and so do not produce functional proteins due to various genetic alterations such as frameshifts, point mutations, insertions, or deletions, there are instances where these genes do exhibit functionality(Bertram et al., 1999). Bertram et al. also identified other members of the now-known MORF family: MRG15 and MRGX. The MRG family of proteins possesses evolutionarily conserved nuclear localization signals (NLS), ATP/GTP binding, helix-loop-helix, Leucine zipper, and C-terminal MRG motifs(Bertram et al., 1999) (Figure 1.1). These motifs facilitate physical interactions with nucleoproteins, including the tumor suppressor Rb. MRG15 stands out from other family members due to a chromatin-binding domain (chromodomain) at its N-terminus (Bertram et al., 1999)(Figure 1.1).

On the other hand, MRGX is a protein with a molecular weight of 31 kilodaltons, found only in mammalian cells(Bertram et al., 1999). It does not have the chromodomain but is otherwise almost identical to MRG15, with a similarity of 75%. Both MRG15 and MRGX have been demonstrated to operate as transcription factors that regulate the activation or suppression of the B-myb promoter in a manner that depends on the cell type(Bertram et al., 1999). Additionally, when a Gal4-Mrg15 fusion protein is attached to a promoter, it suppresses a neighboring reporter gene. Deletion investigation revealed that the chromodomain of MRG15 and the helix-loop-helix and Leucine zipper motifs of both MRG15 and MRGX are essential for their involvement in transcriptional regulation(Bertram et al., 1999). In addition, MRG15 and MRGX are found in the cell nucleus and have been observed to co-localize with hyperphosphorylated RNA polymerase II at specific regions involved in active transcription elongation(Bertram et al., 1999).

1.10 MRG15 homologs among species

MRG15 is a highly conserved protein with homologous from yeast to C.elegans (Devoucoux et al., 2022). MRG15 homolog in S.cerevisae Eaf3 (Esa1-associated factor 3) is a component of the NuA4 histone acetylase and Rpd3 histone deacetylase complexes(Joshi & Struhl, 2005). Regulating the histone acetylation pattern that differentiates promoters from coding areas is crucial(Joshi & Struhl, 2005). Like MRG15, Eaf3 possesses a chromodomain and can bind methylated lysine residues(Joshi & Struhl, 2005)(Figure 1.1). C.elegans homolog of MRG15 is MRG-1, contains conserved chromo-domain and acts as a barrier for germ cell conversion induced by transcription factors(Hajduskova et al., 2019). In the nematode C. elegans, MRG-1 is involved in several essential processes, including chromosomal pairing, preserving the integrity of the genome, suppressing genes located on the X chromosome, and controlling cell division in the reproductive cells(Hajduskova et al., 2019). MRG-1 engages with SET-26, which facilitates the process of restrictive histone H3K9 methylation(Hajduskova et al., 2019). Hadjuskova et al. reported that MRG-1 primarily binds to genomic regions with active histone marks, such as H3K36me3 and H3K4me3(Hajduskova et al., 2019).

In male Drosophila, the chromodomain (CD) of Msl3 within the MSL complex recognizes H3K36me3 marks, while the histone acetyltransferase (HAT)(Mccarthy et al., 2022). Males absent on the first (MOF) adds acetylation to histone H4 lysine 16 (H4K16ac)(Mccarthy et al., 2022). Set2, Msl3, and the Ada2a-containing HAT complex (ATAC) are identified as promoters of germ stem cell (GSC) development(Mccarthy et al., 2022). During the early stages of oogenesis, it has been observed that the Msl3 gene is active. Msl3 promotes the transcription of several members of the stem cell population, known as the SC, through a process called HAT-mediated transcription(Mccarthy et al., 2022). Additionally, Msl3 also promotes the transcription of a specific variant of the eukaryotic ribosomal protein S19, known as eRpS19/RpS19, which is specific to the germline. This translation process ultimately leads to the differentiation of germ stem cells (GSC) into oocytes(Mccarthy et al., 2022).

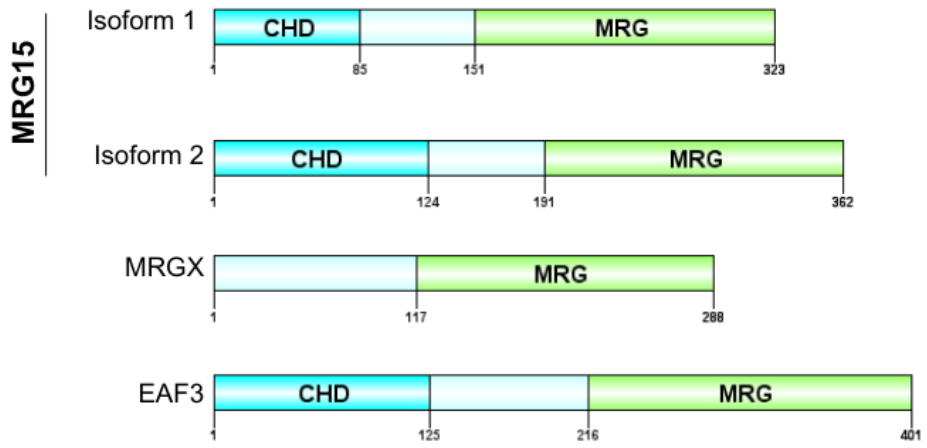


Figure 1.1: A schematic diagram illustrating the spliced variations of MRG15, MRGX, and the functional domains of the yeast Eaf3. CHD: chromodomain, MRG: MRG domain. Adapted from “MRG Proteins Are Shared by Multiple Protein Complexes With Distinct Functions” by Devoucoux M. et al., 2022.

1.11 DNA repair mechanism by MRG15

Bleuyard et al. have demonstrated that PALB2 connects with active genes by interacting with its primary binding partner, MRG15(Bleuyard et al., 2017). Cells with a PALB2 variant with missense mutations that interfere with MRG15 binding are more sensitive to the topoisomerase I (TOP1) inhibitor camptothecin (CPT)(Bleuyard et al., 2017). These cells also have higher levels of abnormal metaphase chromosomes and DNA stress in gene bodies. However, when DNA replication is prevented, these effects are reduced. Their 2018 article argued that the continuous binding of PALB2 to chromatin helps safeguard active genes from stress caused by DNA replication(Bleuyard et al., 2017). H3K36me3 binds the MRG15-PALB2 complex to intact chromatin, allowing PALB2 to be readily accessible at active genes. This mechanism safeguards particularly susceptible areas of the genome from DNA damage(Bleuyard et al., 2017).

1.12 Alternative Splicing Regulation by MRG15

After identifying MORF4 and the MORF family, the Pereira group has worked on this small nuclear protein for almost a decade. In 2010, they reported that H3K36me3 plays a significant role in the management of AS(Luco et al., 2010). They have suggested the presence of adapter systems comprising histone modifications, a chromatin-binding protein that interprets the histone marks, and an interacting splicing regulator(Luco et al., 2010). These complexes serve as a mechanism for transmitting epigenetic information to the pre-mRNA processing machinery. They likely facilitate the recruitment of certain splicing regulators to the pre-mRNA, determining the splicing outcome(Luco et al., 2010). Their findings demonstrate that a specific group of genes dependent on PTB are regulated by a complex involving H3-K36me3, its associated protein MRG15, and the splicing regulator PTB(Luco et al., 2010). Findings in the 2010 paper support the notion of a direct connection between histone modifications and the splicing machinery. However, histone marks can also indirectly influence the selection of splice sites(Luco et al., 2010). There is a lot of evidence that shows the importance of RNA polymerase II elongation rate and higher-order chromatin structure in determining splicing outcomes. Histone modifications work together with these mechanisms(Luco et al., 2010).

Six years after the first article, Pereira's then-student Tominaga group published "MRG15 is required for pre-mRNA splicing and spermatogenesis," where they generated male mice with a conditional knockout (cKO) of MRG15 in their germline were unable to reproduce due to a halt in sperm development at the round spermatid stage(Iwamori et al., 2016). They observed no notable changes in the meiotic division process and histone acetylation modification. 66 germ cell-expressed genes experienced the loss of particular mRNA sequences when MRG15 was absent, but four genes in the MRG15 cKO testes retained unique intronic regions in their mRNAs(Iwamori et al., 2016). Specifically, introns were preserved in the messenger RNAs (mRNAs) that encode the transition proteins responsible for replacing histones during the condensation of sperm chromatin(Iwamori et al., 2016). MRG15 is shown to coexist with splicing factors PTBP1 and PTBP2 at H3K36me3 sites, located explicitly between the exons and sole intron of transition nuclear protein 2 (Tnp2) in round spermatids(Iwamori et al., 2016). Iwamori et al. demonstrate that MRG15 plays a crucial role in pre-mRNA splicing during spermatogenesis.

1.13 MRG15 is part of HAT and HDAC complexes

MRG15 found in histone modifying complexes such as histone acetyltransferase (HAT) and histone deacetylase (HDAC) complexes. (Devoucoux et al., 2022). However, the precise mechanisms by which it operates are yet unknown. Although there has been substantial research on yeast homologs EAF3 and the NuA4/RPD3S complexes, our understanding of MRG15's interactions within mammalian HAT (Tip60) and HDAC (Sin3B/HDAC2) complexes is limited (Carrozza et al., 2005; Doyon & Côté, 2004). A recent study provided insight into its duality. The 2022 investigation conducted by Guan et al. confirmed MRG15's direct contact with the Sin3B/HDAC2 complex in living organisms, indicating its active involvement in deacetylation. This is consistent with the findings of Guan et al., who put out the notion of a "seeding mark" in yeast. In this paradigm, the Rpd3S-NuA3/NuA4 complex ensures an equilibrium between acetylation and deacetylation. Rpd3S eliminates the majority of N-terminal acetylation modifications, except for H3K9ac(Guan et al., 2023). H3K9ac, in conjunction with

H3K36me3, serves as a "seeding mark" that facilitates the restoration of hyperacetylation by NuA3 and NuA4(Guan et al., 2023).Remarkably, the results obtained by Guan et al. are consistent with this observation, as they demonstrate that Sin3B/HDAC2 removes acetyl groups from H3K14 and H3K27 while leaving H3K9 unaffected. This implies a shared system among different species, in which some acetylation marks act as attachment points for opposing modifying enzymes(Guan et al., 2023). This mechanism guarantees a continuous interaction between acetylation and deacetylation processes. In addition to deacetylation, Sin3B/HDAC2, typically accompanied by HDAC1, is vital in regulating the cell cycle(Adams et al., 2018). Their coexistence with E2F4, p107, and p130 at genes such as cyclin A and E2F1 inhibits their production, thus impeding cells from reinitiating the cell cycle. The presence of abnormalities in Sin3B can result in severe outcomes, as demonstrated by embryonic death and developmental abnormalities in Sin3B-knockout mice(Adams et al., 2018). Ultimately, MRG15 seems crucial in preserving the intricate equilibrium between chromatin acetylation and deacetylation. It is involved in both HAT and HDAC complexes, like its yeast equivalent EAF3, indicating a shared mechanism for dynamically modifying chromatin(Adams et al., 2018). Additional investigation into the precise connections of MRG15 and its involvement in cell cycle regulation has significant potential for enhancing our comprehension of intricate biological mechanisms and potentially revealing treatments.

1.14 The main question addressed

Although MRG15 has been associated with maintaining cellular identity, its specific function as a hindrance to epigenetic reprogramming is still poorly understood. The protein's wide range of interactions with chromatin modifiers and its complex functions indicate that it may play a crucial role in regulating the process of somatic cell reprogramming. The study conducted by Bayırbaşı in 2021 showed that the absence of MRG15 produces similar effects to the reduction of SETD2 in improving reprogramming effectiveness. This indicates that MRG15 is a new protein obstacle in the SETD2 pathway. This thesis seeks to elucidate the molecular pathways via which the loss of MRG15 promotes reprogramming, providing insight into its function in preserving cellular identity.

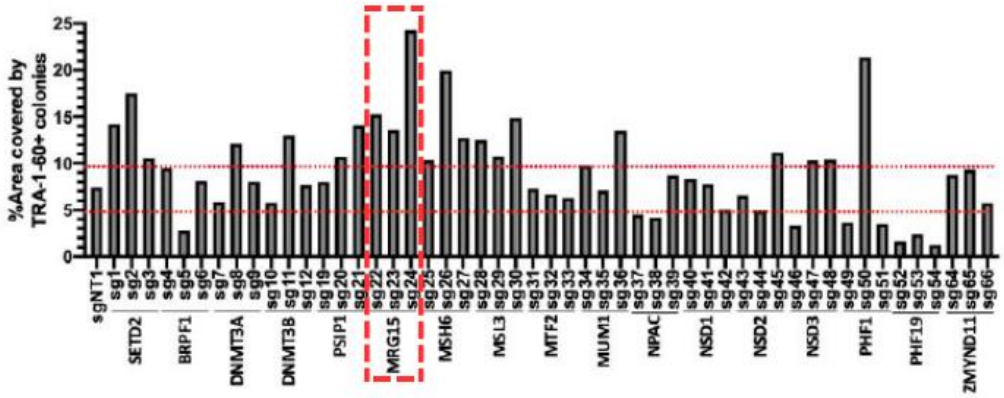


Figure 1.2: Preliminary data showing CRISPR/Cas9-knockout screening of all mammalian H3K36me3 readers in OSKM-reprogramming. MRG15 targeting gRNAs TRA-1-60 colony counts are highlighted with red dots and retrieved from "Understanding How Loss of Histone H3 Lysine 36 Trimethylation Enhances Somatic Cell Reprogramming" by B. Bayırbaşı, 2021.

Chapter 2:
METHODS

2.1 Cloning of sgRNA to lentiCRISPR_v2_blast

sgRNA was designed by Bayırbaşı B. (2021) using GeCKOv2((Addgene#1000000048) and Brunello (Addgene#73178) CRISPR libraries (Sanjana et al., 2014)(Table 2.1). The BsmBI cut sites were modified by adding compatible ends: "CACCG" for the 5' ends of top oligos, "AAAC" for the 5' ends of bottom oligos, and "C" for the 3' ends of bottom oligos. These modified sequences were then synthesized by Macrogen (Bayırbaşı, 2021).

Table.2-1 Sequences of sgRNAs cloned into pLentiCRISPR_v2 vector_Blast

Oligo			**Sequence (5'-3')**
sgNT1	Control	Top	CACCGACGGAGGCTAAGCGTCGCAA
		Bottom	AAACTTGCGACGCTTAGCCTCCGTC
sg22	sgMRG15-1	Top	CACCGAAATACTTCATACATTACAG
		Bottom	AAACCTGTAATGTATGAAGTATTTC

The cloning of sgRNAs into the pLentiCRISPR_version2 (pLentiCRISPR_v2) backbone (Addgene #83480) was performed using the CRISPR cloning methodology developed by the Zhang Lab, as described in the publications by Shalem et al. (2014) and

Sanjana et al. (2014). To begin, the pLentiCRISPR_v2 vector was prepared for cloning. This involved digesting 8 µg of the pLentiCRISPR_v2 backbone with the BsmBI enzyme at a temperature of 55 ℃ for a duration of 3 hours. This process was carried out in 4 independent processes (Table 2.2).

Table 2-2 BsmBI digestion of pLentiCRISPR_v2 vector

Reagents	**1x Reaction**
pLentiCRISPR_v2_BLAST	5000 ng
BsmbI	2 uL
NEB 3.1	5 uL
ddH2O	Up to 50 uL

The vector that had undergone digestion was subsequently dephosphorylated by being exposed to 2 µL of Antarctic phosphatase (NEB M0289S) at a temperature of 37 ℃ for a duration of 30 minutes. This step was carried out to inhibit self-ligation. The vector was subjected to electrophoresis on an agarose gel with a concentration of 0.8%. The BsmBI restriction endonuclease cleaves the pLentiCRISPR_v2 vector at two specific locations, resulting in the release of a 2 kb fragment. Following identifying a 2-kbp fragment on the gel, the remaining plasmid (about 13 kbp) was isolated from the gel using the NucleoSpin Gel and PCR Clean-up kit (MN 740609.50). The concentration of the purified backbone was measured using a NanoDrop spectrophotometer.

sgRNA oligos were produced for cloning using the processes of phosphorylation and annealing. Initially, they underwent phosphorylation using T4 PNK (NEB M0201S). Subsequently, the top and bottom oligonucleotides were combined with an enzyme reaction (Table 2.3) and subjected to annealing in a thermal cycler. The thermal cycler program was as follows: 37 ℃ for a duration of 30 minutes, 95 ℃ for 5 minutes, followed by a gradual decrease in temperature to 25 ℃ at a rate of 0.1 ℃ per second. Ultimately, the annealed oligos were diluted by a factor 1/1000 using dH2O.

Table 2-3 sgRNA Annealing Reaction Mixture

Reagents	**1x Reaction**
Top Oligo (100 uM)	1 uL
Bottom Oligo (100 uM)	1 uL
NEB T4 PNK	0.5 uL
NEB 10X T4 Ligase Buffer	1 uL
dH2O	6.5 uL

The digested backbone, amounting to 50 nanograms, was ligated with one microliter of diluted oligos using T4 DNA ligase (NEB #M0202). The ligation mixture, as specified in Table 2.4, was allowed to incubate at room temperature for a duration of 1 hour. Subsequently, a volume of 5 µL of the ligation product was introduced into a solution containing 50 µL of Stbl3 bacteria. The plasmid was isolated using a small kit (MN 740588.50) from a selected individual colony. Ultimately, the confirmation of cloning was accomplished using Sanger sequencing using the hU6.F sequencing primer (5'-GAGGGCCTATTTCCCATGATT-3').

Table 2-4 sgRNA Cloning Ligation Reaction Mix

Reagents	**1x Reaction**
BmbI cut vector	50 ng
1:1000 oligo mix	1 uL
NEB T4 Ligase	1 uL
NEB 10x T4 Buffer	2 uL
dH2O	Up to 20 uL

2.2 Cloning of Pam Mutant MRG15 into PWZL-BLAST

The cDNA sequence of MRG15 was isolated from dH1f cell line by Ömer Ergül. He designed a primer set targeting MRG15 cDNA with restriction sites added at the end of the primer sequences (Table 2.5) Later, the amplified cDNA was inserted into the pUC19 vector (Addgene #50005). The pam mutation was acquired using the Q5 site-directed mutagenesis kit (NEB E0554). The appropriate primers (refer to Table 2.5) employed in Q5 site-directed mutagenesis were constructed via the online tool NEBaseChanger.

Table 2-5 Primers used in cDNA amplification and Q5 mutagenesis.

Primers	**Sequences (5'-3')**
BamHI-MRG15-Fwd	gaaaggatccaccATGGCGCCGAAGCAG
XhoI-MRG15-Rev	gaaactcgagTCACACAGCTTTCCGATGGTAC
g22 mutation top	AATACTTCATtcacTACAGTGGTTG
g22 mutation bottom	TCACTTGTTTGTCCTTTATG

The reaction mixture is indicated in Table 2.6, and thermocycling conditions are listed in Table 2.7.

Table 2-6 Reaction Mixture for Q5 Site-Directed Mutagenesis

Reagents	**1x Reaction**
NEB 2x Q5 Master Mix	12.5 uL
Forward Primer (10 uM)	1.25 uL
Reverse Primer (10 uM)	1.25 uL
pUC19-MRG15	5 ng

dH2O	Up to 25 uL

Table 2-7 Thermocycling conditions for PCR

Repeat	**STEP**	**Temperature**	**Time**
1	Initial Denaturation	98 ℃	30 sec
25	Denaturation	98 ℃	10 sec
25	Annealing	60 ℃	30 sec
25	Elongation	72 ℃	2 minutes
1	Final Extension	72 ℃	8 minutes
1	Hold	4℃	-

Following the generation of substitutions using PCR, a volume of 1 μL of PCR product was subjected to treatment with KLD (kinase, ligase & DpnI) mix at room temperature for a duration of 5 minutes. This treatment aimed to eliminate unmodified vectors, as indicated in Table 2.8.

Table 2-8 Reaction Mixture for KLD Treatment

Reagents	**1x Reaction**
PCR product	**1 uL**
10x KLD Mix	**5 uL**
2x KLD Buffer	**1 uL**
dH2O	**3 uL**

1 µL of KLD-treated mix was transformed into 50 µL of Stbl3 bacteria following the manufacturer's instructions (NEB E0554). Single colonies were picked from LB-agar plates and inoculated for plasmid purification. Then, the isolated plasmid was sequenced with primer M13 Reverse(5'- CAGGAAACAGCTATGAC-3') to verify mutagenesis.

2.3 Viral Packaging

293T cells were grown in DMEM media supplemented with 10% FBS (referred to as D10), which was made using DMEM 1X, 10% heat-inactivated FBS, and 1% Penicillin/Streptomycin. The cells were maintained at 37 ℃ with 5% CO2 for viral packaging of lenti- and retro-plasmids. Prior to transfection, 293T cells were cultured at a density of 2x106 cells per 10 cm plate. The transfection process involved assembling the necessary plasmids in specific quantities, as indicated in Table 2.9.

Table 2-9 Plasmid constituents for viral packaging

Plasmid	**Lentiviral**	**Retroviral**
pCMV-VSV-G (Addgene #8454)	250 ng	250 ng
psPAX2 (Addgene #12260)	2250 ng	-
pUMVC (Addgene #8449)	-	2250 ng
Vector	2500 ng	2500 ng
DMEM	up to 200 µL	up to 200 µL

In addition to this mixture of plasmids, 20 µL of FuGENE (Promega E2312) was diluted with 180 µL of DMEM. Subsequently, the plasmid mixture was applied onto FuGENE and the entire mixture was subjected to incubation at room temperature for a duration of 30 minutes. Following incubation, the mixture was gradually put onto 293T cells. After a duration of 16 hours, the medium of 293T cells was replaced with new D10. The medium of 293T cells was taken at 48 and 72 hours post-transfection. The medium

was filtered using a 0.45 μm filter to remove the 293T cells. Ultimately, the medium containing viruses was divided into smaller portions and preserved at a temperature of -80 °C.

Preparing a concentrated virus requires the utilization of PEG 8000 (polyethylene glycol, Sigma P5413). A mixture of 50% PEG 8000 was created by combining it with 1X PBS. This mixture was then added to the collected and filtered virus-containing medium at a volume equal to one-fourth of the medium's volume. Following a two-day incubation period with polyethylene glycol (PEG) at a temperature of 4 °C, the mixture was subjected to centrifugation at a speed of 2000 rpm for a total of 20 minutes, after which the liquid portion above the sediment was removed. The pellet was reconstituted with 1X PBS at a dilution of 1/100 of the original volume. Subsequently, it was divided into smaller portions (50-100 μL) and preserved at a temperature of -80 °C.

2.4 Culture of dH1f

dH1f cells were cryopreserved in a freezing solution consisting of 50% fetal bovine serum (FBS), 40% Dulbecco's Modified Eagle Medium (D10), and 10% dimethyl sulfoxide (DMSO). The frozen dH1f cells were thawed at a temperature of 37 °C for a duration of 2 minutes, and the freezing medium was subsequently removed by centrifugation at a speed of 1400 rpm for a duration of 4 minutes. The cell pellet was reconstituted with D10 media and distributed onto plates at a density of 1 million cells per 10 cm plate. Subsequently, the cells were incubated at a temperature of 37 °C in an environment with 5% CO2 concentration. The cell culture media was replaced with fresh D10 every 2 days, and the cells were subcultured when they achieved 80% confluency. The dH1f cells were detached using trypsinization at a temperature of 37 °C for a duration of 3 minutes. Subsequently, trypsin was rendered inactive by the addition of a D10 medium. The cells were counted and distributed at a specific concentration based on the experiment using a hemocytometer.

2.5 Viral Infection

Viral infections were conducted either in a 6-well plate or a 12-well plate, depending on the experimental requirements. A total of 100,000 dH1f cells were put into each well of a 6-well plate, while 50,000 cells were placed into each well of a 12-well plate. The dH1f cells were placed in a 6-well plate and exposed to 1 mL of undiluted virus, 500 µL of D10 media, and 1.5 µL of protamine sulfate (8 mg/mL, Sigma P4505). The cells that were placed in a 12-well plate were exposed to 500 µL of virus that was not concentrated, along with 200 µL of D10 medium and 0.7 µL of protamine sulfate. After a 16-hour period following viral infection, the cell media was replaced with new D10. then, a viral infection was administered once more after 24 hours, and the culture medium was then replaced following an overnight incubation period. The choice of antibiotics for infected cells was based on the presence of the resistance gene in the transduced plasmid. Puromycin was administered at a concentration of 1 µg/mL, whilst blasticidin was employed at a concentration of 10 µg/mL. The process of selecting antibiotics for dH1f cells typically took two days.

2.6 Epigenetic Reprogramming of Fibroblasts

The reprogramming of dH1f cells was conducted in 12-well plates, with each well containing 5x104 dH1f cells. The transfer of Yamanaka factors involved the use of two plasmids: pSIN4-EF2-O2S (Addgene #21162), which expresses OCT4 and SOX2, and pSIN4-CMV-K2M (Addgene #21164), which expresses KLF4 and c-MYC. The cells were inoculated with unenriched viruses from pSIN4-EF2-O2S and pSIN4-CMV-K2M plasmids the day after being placed onto 12-well plates. The cells were incubated for a duration of 16 hours. Subsequently, the viral transduction medium was replaced with a new batch of D10 media. This modification was approved as the initial alteration in the 21-day reprogramming process. Additionally, a fresh D10 medium change was performed on the third and fifth days. On the sixth day, 1/8 of the cells were moved onto the vitronectin-coated plate by trypsinization. The following day, after transferring onto vtn-coated plates, the cells' media was replaced with E7 medium (Table 2.10). During the first 14 days of reprogramming, the cell media was replaced every two days with a fresh

medium. Subsequently, the medium was refreshed on a daily basis with the addition of TGFB.

After the 18-day reprogramming process, the cells were treated with 4% paraformaldehyde at room temperature for 15 minutes to stop further changes. Subsequently, the plates underwent three rounds of washing with PBS. To quantify the iPSC colonies that were produced, we used labeling with the stem cell marker TRA-1-60. The fixed cells were cultured with the biotin-conjugated TRA-1-60 primary antibody (BioLegend 330604) for an extended period at a temperature of 4 ℃. The antibody was diluted at a ratio of 1:250 using a staining buffer that consisted of 3% Trition-X and 3% FBS in PBS. Following five washes with phosphate-buffered saline (PBS), the cells were subjected to a two-hour treatment at room temperature with streptavidin-conjugated horseradish peroxidase (HRP) from BioLegend (catalog number 405210). The HRP enzyme was diluted at a ratio of 1:500 using the staining buffer. After undergoing five rounds of washing with PBS, the cells were subjected to DAB solution (Table 2.11), which serves as the substrate for the HRP enzyme, for a duration of 15-20 minutes at room temperature. Induced pluripotent stem cells (iPSC) that are bound to horseradish peroxidase (HRP) produce brown precipitates when treated with 3,3'-diaminobenzidine (DAB). Consequently, the plates underwent scanning, and the numbers of TRA-1-60+ iPSC colonies were determined by ImageJ analysis.

Table 2-10 E8 Medium Ingredients

Reagent	**Final Concentration**	**Volume**
DMEM/F12		495 mL
Penicillin/Streptomycin		5 mL
Insulin	20 ug/ml	
holo-Transferrin	10 ug/ml	
Selenite	13.6 ng /mL	
L-ascorbic acid	64 ug/ml	
FGF2	100 ng/ml	

TGF-Beta1	1 ng/ml	

Table 2-11 DAB Solution

Reagent	**Volume for 1x**
DAB	250 uL
Hydrogen Peroxide (%0.37)	250 uL
Nickel	250 uL
PBS	5 mL
NaOH	Adjust to pH 7.4

2.7 RNA Isolation & cDNA Synthesis

The NucleoSpin RNA kit (MN 740955.50) was used for RNA isolation. If the intention was to sequence the extracted RNA, RNA isolation was carried out using the direct-zol RNA miniprep kit (Zymogen R2050). Subsequently, a quantity of 1000 nanograms of RNA was employed to produce complementary DNA (cDNA). The initial stage of cDNA synthesis involved the assembly of RNA, dNTP, and random hexamers in PCR tubes using indicated amounts (Table 2.12).

Table 2-12 First cDNA reaction

Reagents	**Catalog No**	**1x Reaction**
dNTP (2 mM)	Thermo Scientific R0192	2.5 µL
Hexanucleotide Mix (2.5 mg/mL)	Roche 11277081001	2 uL
Template RNA		500 ng

dH2O		Up to 16/5 uL

The combination was subjected to incubation at a temperature of 65 ℃ for a duration of 5 minutes, followed by immediate placement on ice. Subsequently, a reverse transcriptase mixture was made according to Table 2.13 and added to the initial mixture. The entire mixture was left to develop at room temperature for a duration of 10 minutes.

Table 2-13 Reverse Transcription Assembly

Reagents	**Catalog No**	**1x Reaction**
5x First Strand Buffer	Invitrogen P/N Y02321	5 uL
DTT (0.1 uM)	Invitrogen P/N Y00147	2 uL
RNAsin	Promega N251B	0.5 uL

Following a 10-minute incubation period, 1 µL of MMLV RT enzyme (Invitrogen 28025-013) was added to each sample. The reaction was then carried out at a temperature of 37 ℃ for a duration of 1 hour and subsequently deactivated at 70 ℃ for 15 minutes. Afterward, the cDNA products were diluted by a factor of ¼ using 75 µL of water free from DNase and RNase enzymes. The diluted cDNA products were then stored at a temperature of -20 ℃.

2.8 Quantitive PCR (qPCR)

qPCR was conducted to assess the relative mRNA expression levels derived from cDNA samples. The SYBR mixture was produced and used with cDNA in 96-well PCR plates for qPCR. The given information is presented in Table 2.14.

Table 2-14 Reaction Mixture

Reagents	**1x Reaction**
2x SYBR Green (Roche 12054420)	10 uL
Primer mix (2.5 μM)	2 uL
cDNA	6 uL
dH2O	2 uL

The experiment was conducted using the Roche Light Cycler 480II. The ΔΔCT method was applied to the Ct data collected from LightCycler, with Beta-Actin expression serving as the internal control. The primers utilized for reverse transcription quantitative polymerase chain reaction (RT-qPCR) are documented in Table 2.15.

Table 2-15 Primer List for RT-qPCR

Primers	**Direction**	**Sequences (5'-3')**
Beta-Actin	**Forward**	CCACGAAACTACCTTCAACTCC
	Reverse	GTGATCTCCTTCTGCATCCTGT
Setd2	**Forward**	**GGAAAACACAACAACTGAACGAG**
	Reverse	**TGAGTTTGCTGGTCTGGGTCT**
Mrg15	**Forward**	**AGCAATGTTGGCTTATACACCTC**

	Reverse	**AGCTTTCCGATGGTACTCAGG**
Gapdh	**Forward**	CTCCTGCACCACCAACTGCT
	Reverse	GGGCCATCCACAGTCTTCTG
Oct4	**Forward**	**AGCGAACCAGTATCGAGAAC**
	Reverse	**TTACAGAACCACACTCGGAC**
Sox2	**Forward**	**AGCTACAGCATGATGCAGGA**
	Reverse	**GGTCATGGAGTTGTACTGCA**
Klf-4	**Forward**	**TCTCAAGGCACACCTGCGAA**
	Reverse	**TAGTGCCTGGTCAGTTCATC**
Nanog	**Forward**	**TGATTTGTGGGCCTGAAGAAA**
	Reverse	**TGGTGGTAGGAAGAGTAAAG**

2.9 Whole Cell Protein Isolation

The cell pellets were subjected to lysis on ice for a duration of 60 minutes using the lysis buffer containing the reagents specified in Table 2.16. During the lysis process, the samples were vigorously mixed using a vortex at 15-minute intervals. Subsequently, the sample was subjected to centrifugation at a speed of 13000 revolutions per minute for a duration of 10 minutes, resulting in the formation of a pellet containing cell debris. The supernatant was obtained and preserved at a temperature of -80 ℃ for subsequent investigations.

Table 2-16 Whole Cell Lysis Buffer Mix

Reagants	Stock Concentration	Final Concentration
RIPA Lysis Buffer (EcoTech RIPA-100)	1x	1x
PMSF	100 mM	1 mM
CompleteProtease Inhibitor Coctail (Roche 11697498001)	10x	1x

2.10 Histone Isolation

The histone extraction procedure was conducted using the protocol provided on the Abcam website (Abcam, n.d.). The fibroblast pellets were disrupted using a Triton Extraction Buffer (Table 2.17) at a low temperature for 10 minutes with moderate shaking. Subsequently, the cells were subjected to centrifugation at a speed of 2000 revolutions per minute for a duration of 10 minutes, and the supernatant was discarded. The cell pellet was rinsed with half the volume of TEB and subjected to centrifugation once more. The pellet was reconstituted with 0.2 N hydrochloric acid and incubated at a temperature of 4 ℃ for the duration of the night. During this incubation period, acid causes the extraction of histones into the supernatant. The following day, the sample underwent centrifugation at a speed of 2000 revolutions per minute for a duration of 10

minutes. The supernatant, which contained histones, was then collected. To neutralize the samples, a solution of 0.1 M NaOH was added to the supernatant in a volume that is one-fifth of the amount of HCl utilized. The isolated histones were preserved at -80 ℃ for extended periods or at -20 ℃ for shorter durations prior to western blotting.

Table 2-17 Triton Extraction Buffer Mix

Reagent	**Final Concentration**
Triton X100	%0.5
PMSF	2mM
NaN3	%0.02

2.11 Western Blot

The concentrations of the isolated proteins were measured using the BCA protein assay kit (Pierce 23227). The color development resulting from the BCA reaction was measured using a plate reader at a wavelength of 562 nm. Subsequently, concentrations were determined based on the absorbance values of BSA standards. The BSA protein, provided with Pierce 23227, was utilized in seven distinct concentrations (2, 1.5, 1, 0.75, 0.5, 0.37, and 0.25 µM) as reference standards in the BCA analysis.

After calculating protein concentrations, the proteins were ready for sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). To begin, a protein loading solution was produced by combining 900 µL of 4X Laemmli sample buffer (Biorad 1610747) with 100 µL of 5M 2-mercaptoethanol. Subsequently, protein concentrations were modified to the appropriate volume using 1X loading buffer and dH2O, and the proteins were subjected to boiling at 95 ℃ for a duration of 15 minutes. The boiled samples were promptly placed and stored on ice. If they were utilized in subsequent research, they were preserved at a temperature of -20 ℃.

The prepared proteins were placed into a polyacrylamide gel (Mini-Protean 4568084) and subjected to electrophoresis using 1X TGS running buffer at a current of 35 mA per gel for a duration of 45-60 minutes. The 1X TGS (Tris/Glycine/SDS) buffer was created by diluting the 10X TGS buffer with distilled water (BioRad 161-0772). The proteins were transferred from the gel onto a methanol-activated PVDF membrane once the running process was finished. The transfer was conducted utilizing the Trans-Blot Turbo transfer device (BioRad 1704150) in a semi-dry manner. To prepare the transfer solution, a 1X concentration of transfer buffer was made by diluting 10X transfer buffer TG (Tris/Glycine) with 70% water and 20% methanol. The stacks employed in the transfer mechanism were saturated with 1X transfer buffer. Following the transfer, the PVDF membrane was treated with a solution of 5% non-fat dried milk in TBS-T (composed of 24 g Trizma Base, 88 g NaCl in 1L dH2O with pH 7.6 and supplemented with 1 mL Tween-20) for 1 hour at room temperature with agitation. Subsequently, the membrane was incubated overnight at 4 °C with the first antibody by shaking. The initial antibodies were employed at a dilution of 1:1000, using a primary antibody buffer consisting of 2% BSA in TBS-T. Following a 10-minute incubation with antibodies, the membranes were thoroughly cleaned with TBS-T by shaking multiple times. Next, the samples were exposed to a secondary antibody that was diluted 1:5000 with HRP conjugation. The incubation took place at room temperature for 1 hour with continuous shaking. The membrane underwent many washes with TBS-T. Subsequently, the membranes were seen using the Li-cor Odyssey Fc imaging system, utilizing enhanced chemiluminescence (ECL, Thermo Scientific 32106) as the substrate for HRP. The antibodies employed in western blot tests were enumerated in Table 2.18.

Table 2-18 Antibody List for Western Blot

Antibody	**Catalog No**	**Source**
Tri-Methyl-Histone H3 (Lys36) (D5A7)	Cell Signaling 4909S	Rabbit
Di-Methyl-Histone H3 (Lys36) (C75H12)	Cell Signaling 2901S	Rabbit

Acetyl-Histone H3 (Lys9) (C5B11)	Cell Signaling 9649	Rabbit
Acetyl-Histone H3 (Lys14) (D4B9)	Cell Signaling 7627	Rabbit
Acetyl-Histone H3K27 (D5E4)	Cell Signaling 8173	Rabbit
Histone H3 (D1H2) Total	Cell Signalling 4499	Rabbit
MORF4L1/MRG15 (D2Y4J)	Cell Signaling 14098	Rabbit
Beta-tubulin	Abcam ab6046	Rabbit
Beta Actin antibody	Abcam ab8227	Rabbit
Goat Anti-Rabbit IgG H&L (HRP)	Abcam ab95071	Goat

2.12 Flow Cytometry

The identification of TRA-1-60 + fibroblasts was performed utilizing the BD Accuri C6 instrument, with the utilization of negative and positive controls to establish gating parameters. TRA-1-60 expression on the surface of cells was examined on the sixth day of the reprogramming process. The cells were dissociated using 0.3 mL of a 0.05% trypsin-EDTA solution from each well of a 12-well plate and then placed in an incubator at a temperature of 37°C for a duration of 5 minutes. 0.3 milliliters of a 10% solution of FBS in PBS was added to each well and then subjected to centrifugation at a speed of 1500 revolutions per minute for a duration of 5 minutes. The pellets were suspended again in a solution of 1% FBS in PBS and subjected to centrifugation as previously described. The pellets were suspended in 0.25 µL of PE-conjugated anti-human TRA-1-60-R antibody (Biolegend, catalog no. 330610) in a solution of 1% FBS and PBS. The mixture was then incubated on ice in the dark for 1 hour. The cells were rinsed twice with PBS

containing 1% FBS using the same centrifugation settings as previously. The pellets were suspended in a 100 µL solution of 1% FBS in PBS. Reprogramming efficiency was evaluated using an Accuri C6 flow cytometer (BD) by placing gates on the GFP and PE channels based on the negative control.

2.13 Proliferation Assay

A total of 1000 fibroblasts were distributed evenly into each well of 96-well plates, with six replicates for each condition. Cell proliferation was assessed on the following day after seeding, designated as day 0, and on the sixth day using the CellTiter-Glo Luminescent Cell Viability Assay (Promega). The cells' media was changed every two days, and each group received inhibitors three times. The assay followed the manufacturer's instructions, utilizing luminometric readings with a plate reader (Synergy H1 Reader, BioTek).

Another method was counting cells with a hematocytometer. 10000 cells were seeded as in replicates, and on day one, day three, and day five, triplicates were trypsinized and counted.

2.14 Immunofluorescence Staining

To conduct immunofluorescence staining, the cultivated cells in tissue culture plates were initially treated with 4% paraformaldehyde (PFA) for a duration of 15-30 minutes at room temperature. Subsequently, the cells were rinsed with PBS. Plates can be stored at a temperature of 4°C following fixing. Subsequently, the cells were treated with a 0.1% solution of Triton X-100 in PBS for a duration of 10 minutes to induce permeabilization. The cells were subsequently treated with a blocking solution consisting of 3% BSA and 5% goat serum diluted in PBS and incubated for 30 minutes. Following the blocking step, the primary antibodies were diluted with a blocking buffer. The diluted antibodies were then introduced to the cells and incubated overnight on a rotator at a temperature of 4 degrees Celsius. The primary antibodies are listed in Table 2.19. After the primary antibody incubation, the cells were rinsed thrice with PBS for 5 minutes each. The secondary antibodies were diluted in a blocking buffer. The diluted antibodies were then applied to the cells and left for 1 hour at room temperature while being shielded from light. The secondary antibodies are listed in Table 2.19. Following the incubation with

the secondary antibodies, the cells were rinsed three times with PBS for a duration of 5 minutes per rinse. Subsequently, the nuclei were stained with DAPI at a dilution of 1:5000 in PBS for a duration of 10 minutes. This was then proceeded by a final wash with PBS. The Cytation 5 automated fluorescence microscope was used to capture images of stained cell samples. Each well of a 48-well plate was filled with 150 µL of solution for all antibody incubation phases.

Table 2-19 Antibody List for Immunofluorescence

Antibody	**Catalog No**	**Source**
NANOG	ab21624	Rabbit
OCT4	ab19857	Rabbit
SOX2	PCRP-SOX2-1B3	Mouse
SSEA4	MC-813-70	Mouse
anti-rabbit Alexa 488		Goat

2.15 RNA- Sequencing and GSEA

For RNA Sequencing RNA isolation from respective samples collected as mentioned above. Transcriptome_Denovo_Seq performed by Macrogen Europe B.V. with 50M throughput and read length. The Deseq analysis performed by Hamzah Syed and rMAT analysis performed by Marco Preussner. GO gene set analysis and GSEA performed by me. For GSEA, protocols in their website followed (GSEA,n.d.) First gene count file converted to gene cluster text file format (*.gct). The version string, which is constant for this file format, appears in the first line.The size of the data table that takes up the rest of the file is indicated by the numbers in the second line. The amount of data columns does not include the name and description fields. A list of identifiers for the samples linked to each of the remaining columns in the file is contained in the third line. This converted excel file is saved as “tab delimited text” and at the end of the file name *.gct

should be included. Then this data is loaded to GSEA program. In the run GSEA page expression data set is selected as *.gct file. Gene sets database can be choosed as preferred, for my experiments I have run all gene sets. Phenotype label is constituted by create-on-the-fly-phenotype option. I have triplicates so I entered each triplicate group name as NT-1, NT-2, NT3 etc. as it is labeled in the expression data set. No collapse is chosen and the analysis run.

2.16 Data Analysis

For data analysis Graphpad Prism version 8.4.2 is used. T-tests and One way ANOVA is utilized and satiated under the related figure.

Chapter 3:
RESULTS

3.1 SETD2 inhibitor noticeably diminish H3K36me3 expression whereas increase H3K36me2 levels

Previous studies by Özçimen B. demonstrated that upon SETD2 silencing, H3K36me3 levels decrease, whereas H3K36me2 levels increase. This results from a lack of trimethylation activity and increased epigenetic reprogramming efficiency. To test the SETD2 inhibitor's catalytic activity, I treated dh1f cells with various concentrations of SETD2 inhibitor (EPZ-719). I have received two isotopes of the same inhibitor, isotope 1 SETD2 enzymatic activity. I treated dh1f cells with various concentrations of SETD2 inhibitor (EPZ-719). I have received two isotopes of the same inhibitor, isotope 1 trans stereoisomer and isotope two cis stereoisomer (Figure 3.1). A recently developed SETD2 inhibitor) and its inactive stereoisomer was tested (Lampe et al., 2021). I treated dh1f cells with various concentrations of SETD2 inhibitor, EPZ-719-2 and its stereoisomer EPZ719-1. (Figure 3.1). I have tested each isomer under the same conditions (Figure 3.2.a). Protein samples from EPZ-719-2 treated cells demonstrated diminished H3K36me3 levels in every tested inhibitor concentration (Figure 3.2.b). However, isomer 1 was less active than the second since it could only reduce tri-methylation at 10 uM concentration (Figure 3.2.a). This suggests that SETD2 inhibitors successfully repress SETD2 catalytic activity. The increase in H3K36me2 levels upon inhibition of SETD2 was in line with the literature.(Yu et al., 2023).

Cells that have diminished the tri-methyl mark have increased di-methyl levels (Figure 3.2.b) Similar to literature diminished H3K36me3 levels led to increase in h3K36me2 levels.(Keogh et al., 2005) (Figure 3.2.b). Cells that were treated with 3 uM EPZ-719-1 inhibitor have increased H3K36me2 levels with little to no decrease in the H3K36me3 mark (Figure 3.2.b). This may suggest that this stereoisomer may have an off-target effect on di-methylase catalytic activity. Since EPZ-719-2 is more potent than EPZ-719-1 in catalytic inhibition, I have decided to use EPZ-719-2.

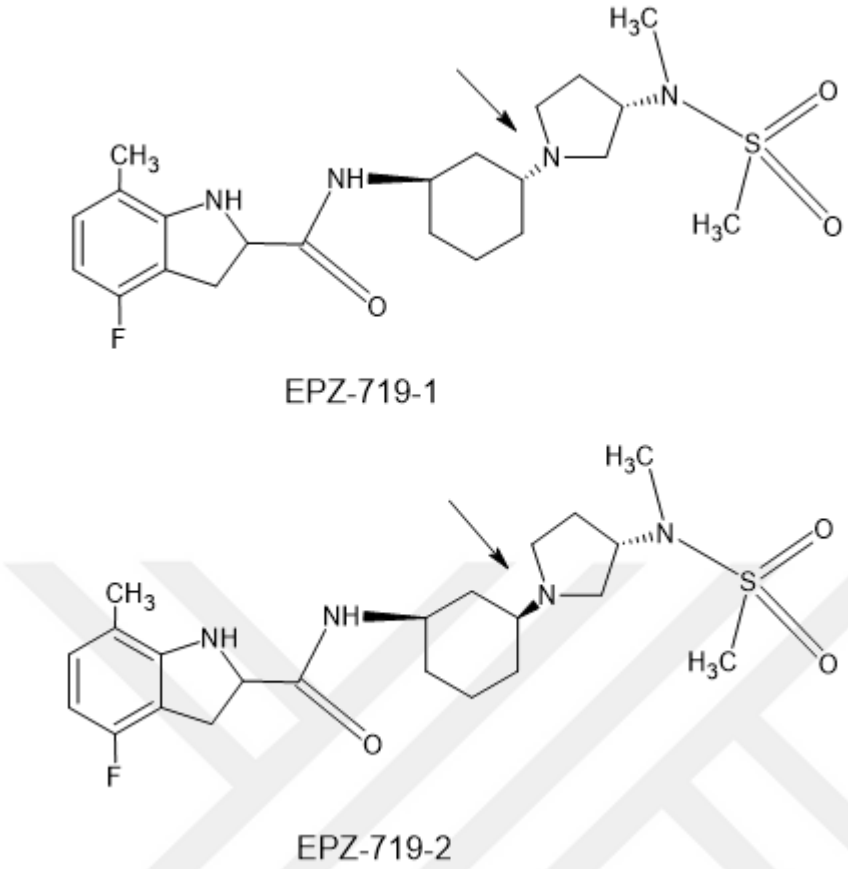


Figure 3.1 Chemical structures of SETD2 inhibitor EPZ-719. Trans isomer P1 (top) and cis isomer P2(bottom). Arrows indicate trans/cis position. (Draw in ACD Labs ChemSketch)

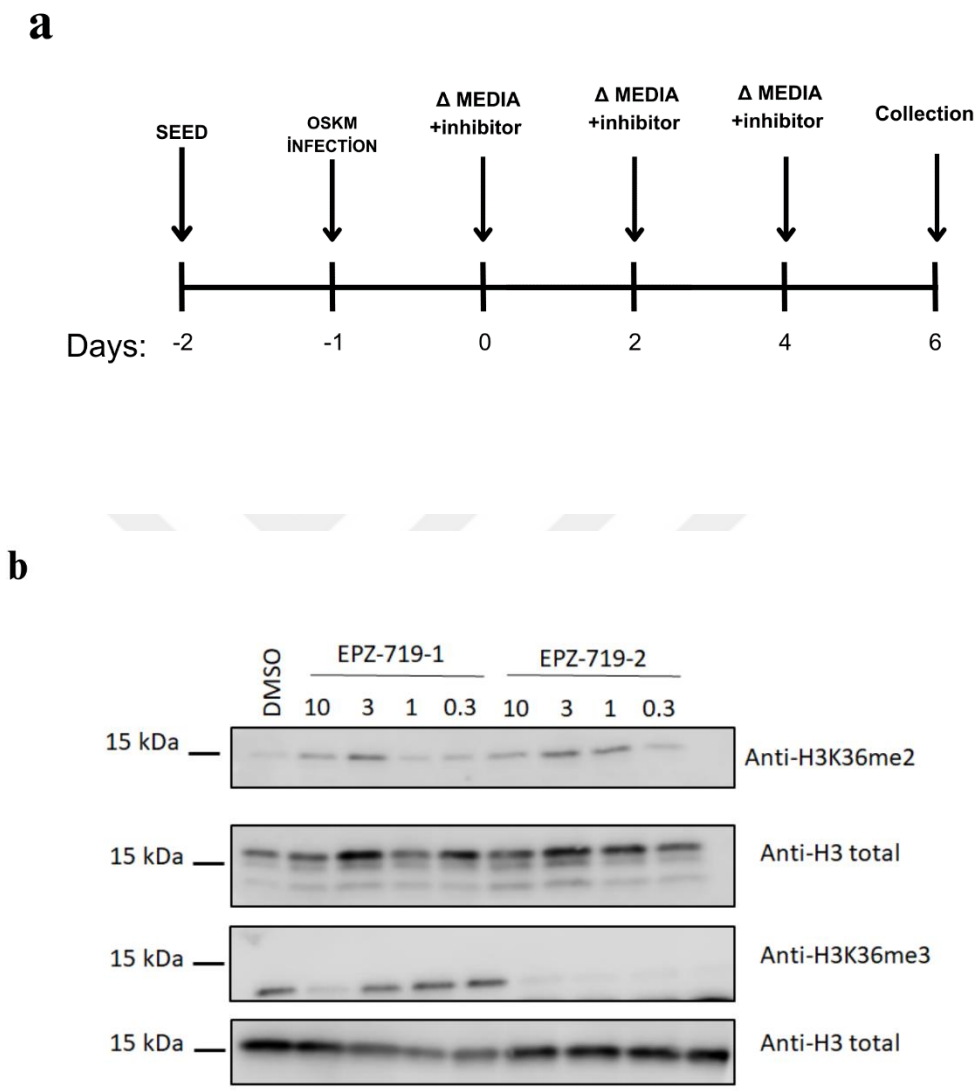


Figure 3.2 EPZ-9-719-2 treated cells have diminished H3K36me3 levels. levels (a) The schematic depicts the time intervals for treatment of EPZ-719-1 and EPZ-719-2. (b) Western blots for histone H3 lysine 36 di-methylation (H3K36me2) and H3 lysine 36 tri-methylation (H3K36me3) in cells treated for EPZ-719-1 and EPZ-719-2 with the indicated micro molar concentrations. DMSO was used as a negative control. Total histone H3 was used as a loading control.

3.2 SETD2 inhibition increases reprogramming efficiency.

To gain insight into the EPZ-719-2's (also referred to as iSETD2) effect on epigenetic reprogramming, I have conducted two separate experiments to determine the optimal time course of treatment and the dose at which maximal increase in reprogramming is achieved. I have tested inhibitor concentrations ranging from 1uM to 0.1 uM. EPZ-719-1 did not yield significantly higher iPSC colonies than DMSO except for 0.1 uM treatment (Figure 3.3.b). I EPZ-719-1 0.1 uM increased iPSC colonies compared to NT (Figure 3.3b). However, H3K36me3 levels did not decreased EPZ-719-1 treated cells (Figure 3.2.a) This may suggest that EPZ-719-1 0.1 uM increased iPSC colonies independent from H3K36me3 levels decrease. EPZ- (Figure 3.2.a).

EPZ-719-2 treatment resulted in significantly higher numbers of iPSC colonies than the DMSO group except at 0.1 uM (Figure 3.3.b). The P-value of 1 uM treatment was lower than 0.3 uM treatment; therefore, I have decided to continue using EPZ-719-2 at 1 uM for all subsequent experiments (Figure 3.3.b).

In the second set of experiments, In the time course experiment, I tested three different time windows: from day 0 to day 6 post-OSKM expression at which time cells were passaged onto MEFs, after passaging until the end of reprogramming and during the entire period of reprogramming (Figure 3.3.a). I have observed that cells treated with iSETD2 for the initial 6 days produced significantly higher iPSC colonies than the DMSO-treated group (Figure 3.3c). This suggests that iSETD2 treatment at an earlier stage of reprogramming alters the epigenetic landscape to facilitate acquisition of pluripotent identity.

Taken together, these results suggests that catalytically inhibition of SETD2 in the first six days of reprogramming enhances reprogramming efficiency.

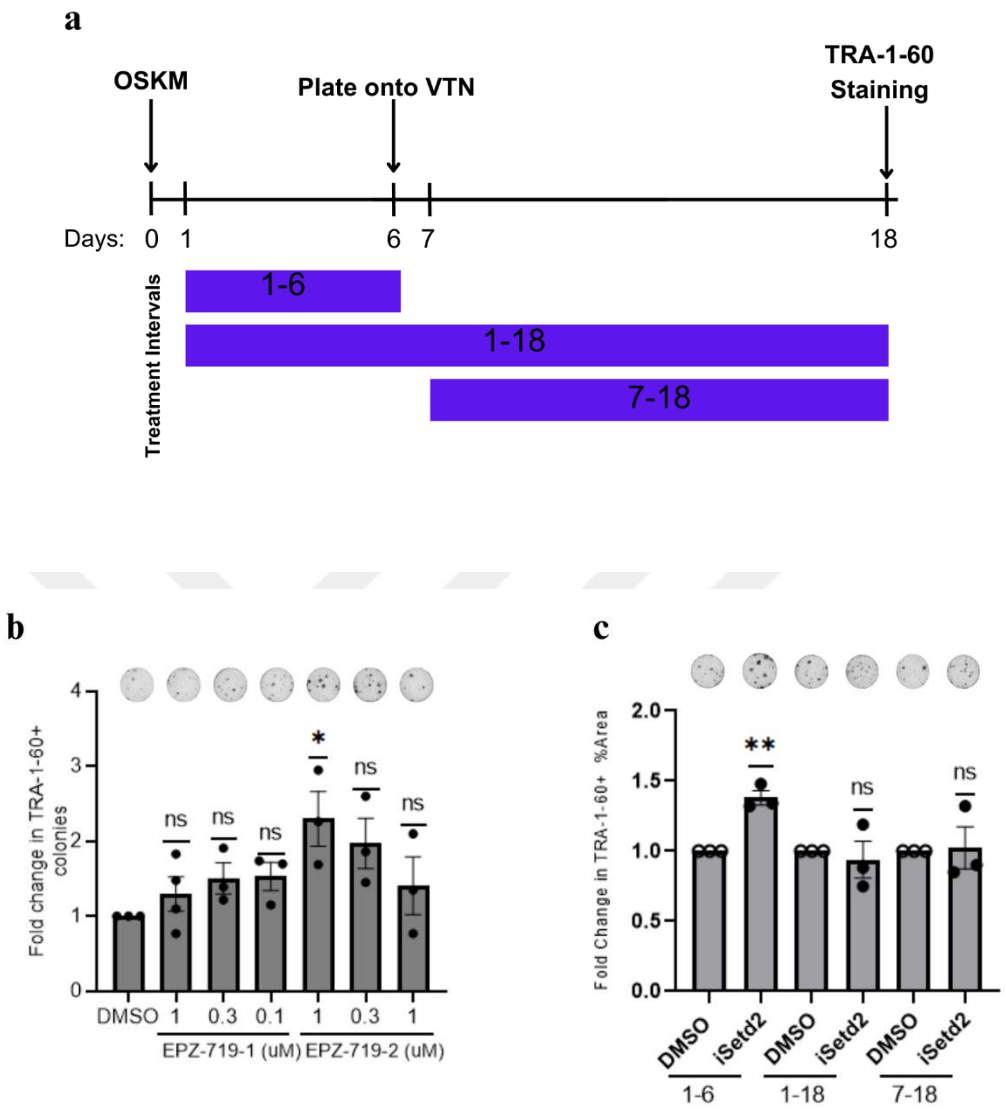


Figure 3.3 EPZ-719-2 is the most potent when using the first weeks of reprogramming in terms of reprogramming efficiency(a) The schematic depicts the time intervals for treatment of EPZ-719-2 during reprogramming. (b) Fold change in the number of TRA-1-60-positive colonies with the indicated compound treatments compared to DMSO control. Representative well images are shown above. Bar graphs show the mean, and error bars represent the standard error of the mean. n=3, biological replicates for each treatment with three technical replicates. The Unpaired t-test calculates P-values. The exact p-values from left to right are 0.3276,0.0711,0.0496,0.0234,0.0449,0.3504.(c) Fold change in percent area covered by

TRA-1-60-positive colonies with the indicated period treatments compared to DMSO control. Representative well images are shown above. Bar graphs show the mean, and error bars represent the standard error of the mean. n=3, biological replicates for each treatment with three technical replicates. P-values are calculated by Unpaired t-test. * denotes $p<0.05$, ** denotes $p<0.005$, p-values from left to right are 0.0016,0.6696,0.8832

3.3 SETD2 inhibition does not enhance proliferation.

After establishing the effect of SETD2 catalytic inhibitor on overall reprogramming efficiency, I next investigated whether SETD2 inhibition accelerates the emergence of reprogrammed cells. To this end, I determined the percentage TRA-1-60 positive cells on day six after OSKM expression which is an early time point during reprogramming. EPZ-719-2 treated cells do not have a significantly higher TRA-1-60 + population than DMSO-treated cells (Figure 3.4.a). This data suggests that iSETD2 treatment does not accelerate reprogramming.

A previous study stated that SETD2 is related to the proliferation of gastric cancer types and that overexpression of SETD2 significantly decreased the proliferation of different cancer cell lines(Chen et al., 2018). Proliferation is also an important characteristic of pluripotent cell types(Takahashi & Yamanaka, 2006). Thus, I have conducted a proliferation assay with iSETD2. However, similar to genetic loss of SetD2 (Ozcimen,2018), diminished SETD2 catalytic activity did not result in increased proliferation of fibroblasts compared to DMSO-treated fibroblasts (Figure 3.4.b). This may suggest that a decrease in SETD2 levels may cause an increase in proliferation in a cancerous epigenetic landscape; however, the same phenotype is not observed during reprogramming.

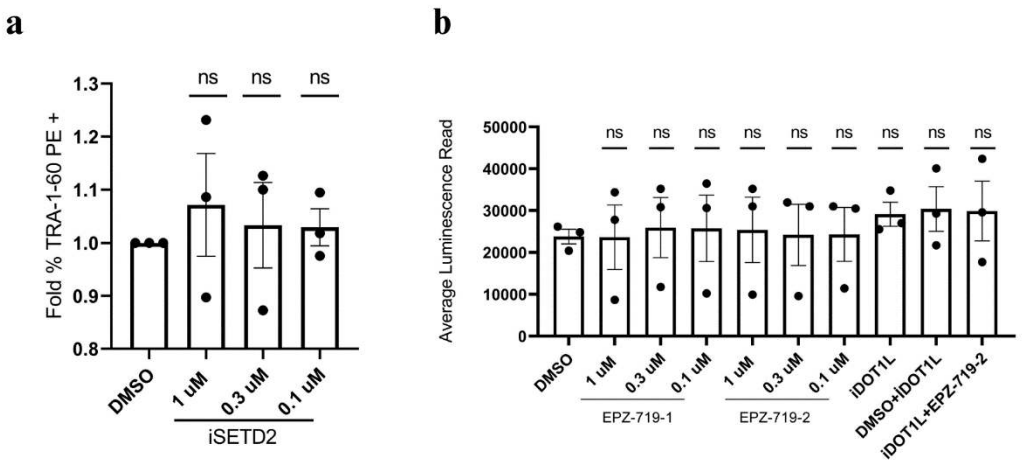


Figure 3.4 The Effect of SETD2 inhibition could not observed on the sixth day. (a) Fold change in the number of TRA-1-60-positive cells with the indicated concentration treatments of EPZ-719-2 compared to DMSO control. Bar graphs show the mean, and error bars represent the standard error of the mean. n=10000, biological replicates for each treatment with three technical replicates. The Unpaired t-test calculates P-values. P-values from left to right are 0.5006,0.7023,0.4472 (b) Average Luminescence Read for cells treated with indicated EPZ-719-1 and EPZ-719-2 various concentrations, iDOT1L and combinations compared to DMSO. Bar graphs show the mean, and error bars represent the standard error of the mean. n=6, biological replicates for each treatment with three technical replicates. The Unpaired t-test calculates P-values. p-values from left to right are: 0.9837,0.7874,0.8201,0.8254, 0.9603, 0.9426, 0.1885,0.3061, 0.4525.

3.4 Combination of iSETD2 with other inhibitors has increased reprogramming efficiency.

I next investigated the combined effect of iSETD2 with inhibitors proven to increase reprogramming efficiency. A83 is an inhibitor of ALK-5, which is a type I TGFB receptor, and it was previously reported to help maintain the pluripotent state of iPSC culture(Figure 3.5.a) (Tojo et al., 2005). The combination of iSETD2 with A83 did not yield significantly higher reprogramming efficiency. RN1 is another small molecule reported to increase iPSC generation (Rivers et al., 2015). It inhibits the catalytic activity of LSD1 lysine-specific histone demethylase(Rivers et al., 2015). However, similar to A83, RN1 did not show a significant increase in iPSC generation when combined with

iSETD2(Figure 3.5.a). Sevinç et al. have reported that using iBRD9, an inhibitor that targets a subunit of noncanonical complex BAF, has yielded increased iPSC colonies in reprogramming(Sevinç et al., 2022). ncBAF has been shown to locate cell type-specific enhancers, and its depletion leads to a decrease in proliferation and G1 arrest(Sevinç et al., 2022). Similar to A83 and RN1, iBRD9 did not yield a significant increase in reprogramming when combined with iSETD2 (Figure 3.5.a). CBP and EP300, also known as KAT3A and KAT3B, are histone acetyltransferases that act as transcriptional coactivators(Ebrahimi et al., 2019). Ebrahimi et al. have reported an inhibitor targeting these coactivators, namely CBP30, yielding a nearly 4-fold increase in reprogramming(Ebrahimi et al., 2019). I have observed that combining iSETD2 with CBP30 has a significantly additive effect(Figure 3.5.a). Like CBP30, a combination of iDOT1L and iSETD2, also yielded higher iPSC colonies than iDOT1L alone(Figure 3.5.a). DOT1L is an H3K79 methyl transferase able to catalyze all three methylation forms, and this specific histone mark is associated with active genes. Onder et al. have reported that using iDOT1L during reprogramming has yielded 4-fold higher iPSC colonies(Onder et al., 2012). These experiments suggest that iSETD2 has an additive effect when combined with potent inhibitors such as CBP30 and iDOT1L (Figure 3.5.a).

Chemical reprogramming using chemical inhibitors has been reported for the past decade(Guan et al., 2022). I wanted to investigate the iSETD2 effect on small inhibitor cocktails. Both Sevinç et al. and Ebrahimi et al. have reported that using iDOT1L has yielded significantly higher iPSCs than using only one inhibitor (Ebrahimi et al., 2019; Sevinç et al., 2022). I tested inhibitor cocktails with iDOT1L, iBRD9, and CBP30(Figure 3.5.b). The combination of iDOT1L & iBRD9 showed significantly higher reprogramming efficiency than the minus iSETD2 group(Figure 3.5.b). Oppositely, the combination of iDOT1L & iCBP30 did not show higher reprogramming efficiency than the minus iSETD2 group(Figure 3.5.b). Interestingly, the combination of BRD9 with iSETD2 yields higher efficiency in the presence of iDOT1L, and the combination of CBP30 yields higher efficiency unlike in the presence of iDOT1L(Figure 3.5.b). This data shows that SETD2, BRD9, CBP30, and iDOT1L have complex interactions, and their combinations should be adjusted to the highest efficiency(Figure 3.5.b).

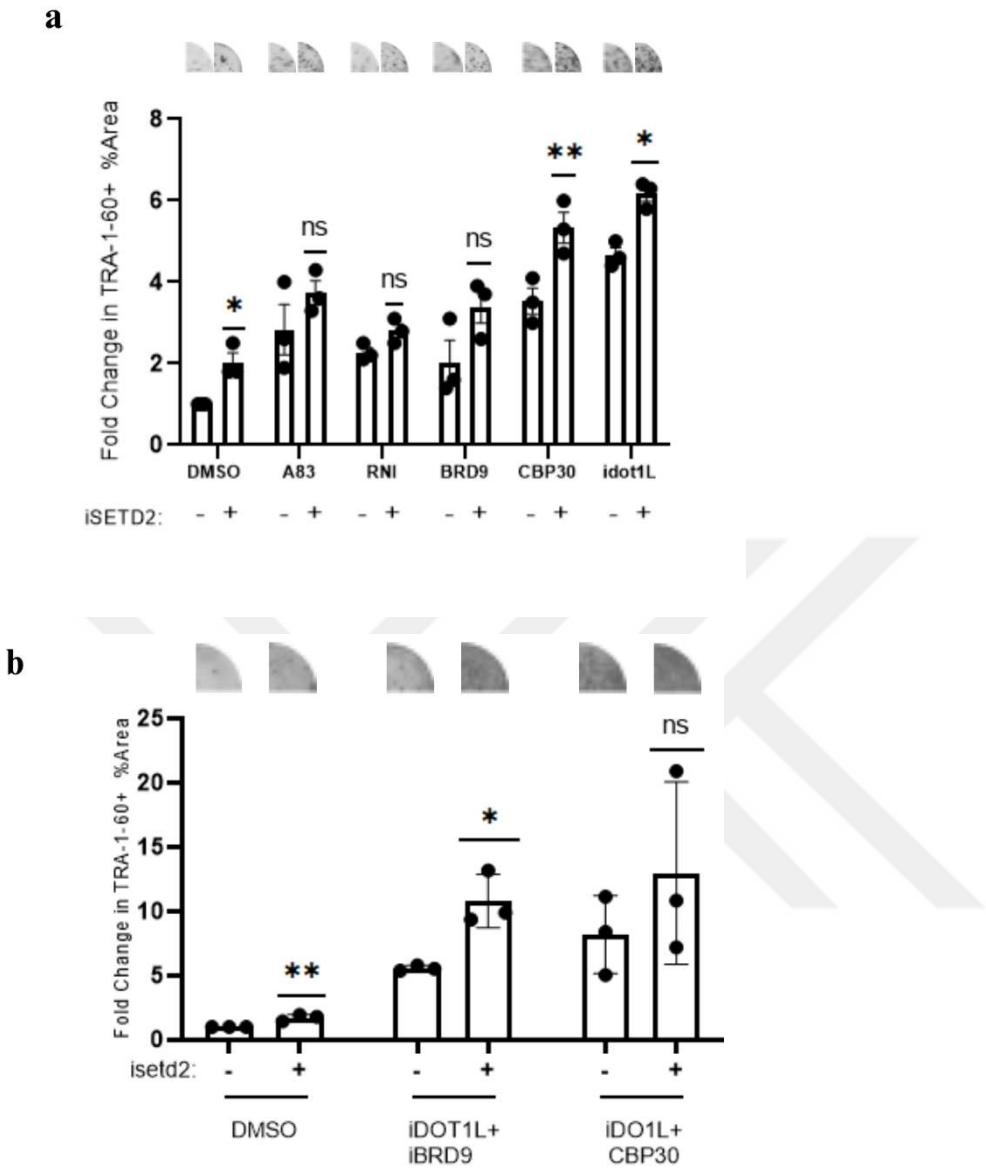


Figure 3.5 iSETD2 displays an additive increase in reprogramming with CBP30 and iDOT1L. (a)Fold change in percent area covered by TRA-1-60-positive colonies with the indicated inhibitor treatment compared to minus iSETD2 group. Representative well images are shown above. Bar graphs show the mean, and error bars represent the standard error of the mean. n=3, biological replicates for each treatment with three technical replicates. P-values are calculated by Unpaired t-test. * denotes $p<0.05$, ** denotes $p<0.005$, p-values from left to right are 0.0114, 0.2590, 0.0647, 0.116, 0.0216, 0.0042. (B) Fold change in percent area covered by TRA-1-60-positive colonies indicated inhibitor combination treatment compared to the minus iSETD2 group. Representative well images are shown above. Bar graphs show the mean, and error bars represent the standard error of the mean. n=3, biological replicates for each treatment with three

technical replicates. An Unpaired t-test calculates P-values. P-values from left to right are 0.0088, 0.0119, and 0.3436.

3.5 MRG15 KO increases reprogramming efficiency

In a previous study, a CRISPR/Cas9 knock-out strategy targeting SETD2 reader proteins was utilized and MRG15 was identified as the most prominent candidate to phenocopy SETD2 knockdown in terms of reprogramming efficiency (Bayirbasi,2021). In confirmatory experiments, I observed that MRG15 knockout results in a 5-fold increase in iPSC colonies compared to NT-infected fibroblasts (Figure 3.6.a). Western blotting confirmed greatly diminished MRG15 protein levels in MRG15-targeted gRNA expressing cells(g22) compared to control gRNA (NT1) expressing cells (Figure 3.6.b). This suggests that MRG15 is a prominent novel barrier to epigenetic reprogramming.

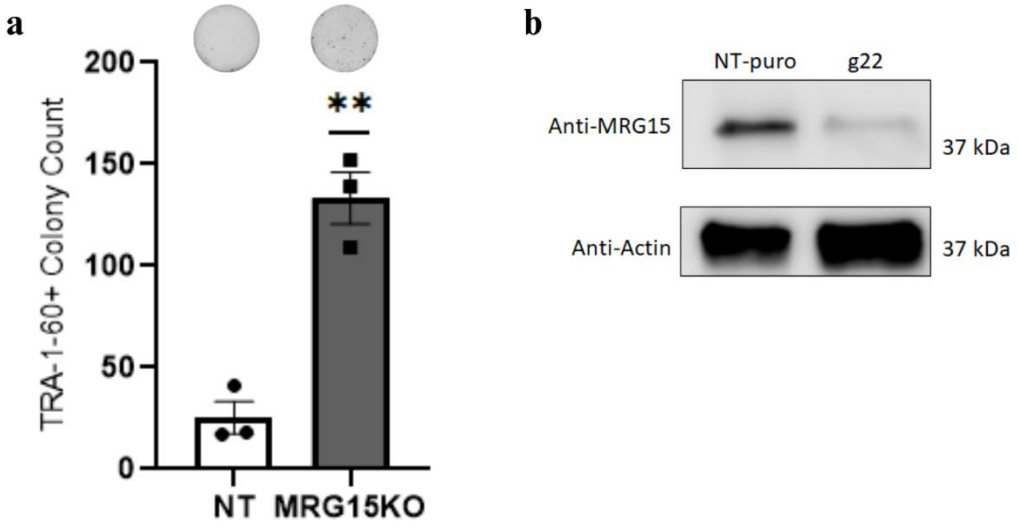


Figure 3.6 MRG15 acts as a barrier to epigenetic reprogramming. (a) The number of TRA-1-60 positive colonies with the MRG15KO compared to the NT group. Representative well images are shown above. Bar graphs show the mean, and error bars represent the standard error of the mean. n=3, biological replicates. The Unpaired t-test calculates P-values. * denotes $p<0.05$. The exact p-value is 0.0019. (b)Western blots for MRG15 in cells treated with guide RNA targeting MRG15 (g22). NT was used as control. Total actin was used as a loading control.

3.6 MRG15 KO phenotype is independent of CRISPR/Cas9 off-target effect

Next, I wanted to establish that the reprogramming phenotype was not due to potential off-target effects of CRISPR.Cas9 by performing a genetic rescue experiment. To this end, I cloned MRG15 cDNA to a retroviral expression vector (PWZL-BLAST), which was rendered immune to gRNA with two silent mutations in the PAM region of MRG15-targeting guide 22. I verified the introduction of silent mutations via site-directed mutagenesis by Sanger Sequencing the region of interest onMRG15 cDNA (Figure 3.7.a).

I confirmed that I have successfully expressed wild-type MRG15 with Western Blot analysis (Figure 3.7.b). Next, I measure epigenetic reprogramming efficiency upon wild-type expression of MRG15. Overexpression of wild type PAM mutant MRG15 cDNA in MRG15KO cells rescued the phenotype and showed similar reprogramming efficiency as NT cells (Figure 3.7.c). We can therefore conclude that the increase in iPSC generation efficiency was due to the absence of MRG15 and not CRISPR/Cas9 off-target events.

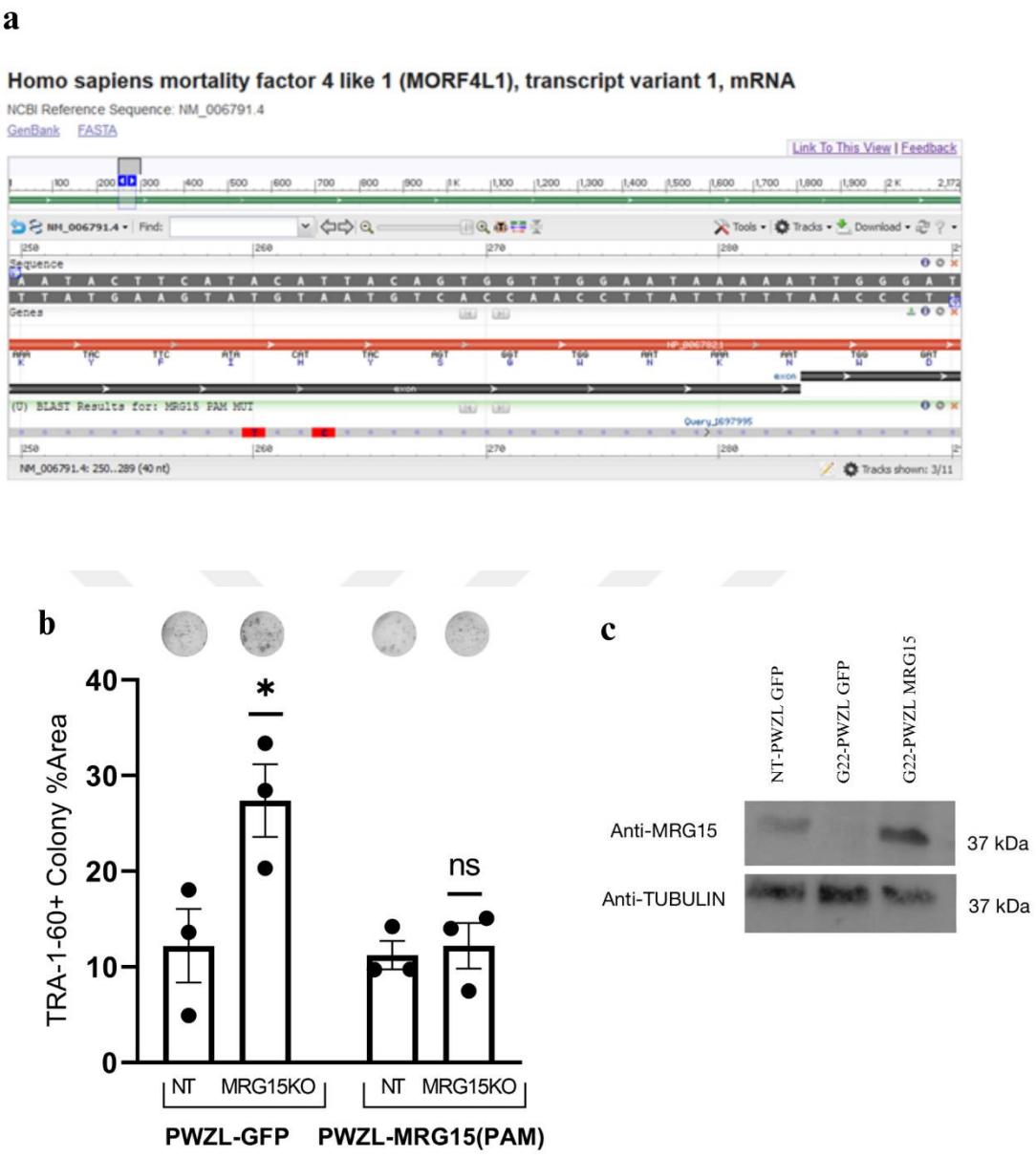


Figure 3.7 Overexpression of MRG15 reverts phenotype. (a) Sanger results compared to wildtype MRG15 transcript variant 1. Sanger result is at the bottom, and the reference sequence is at the top. Mismatches are highlighted in red. (b)Western blots for MRG15 in cells treated with guide RNA targeting MRG15 (g22) and MRG15 overexpression (PWZL- MRG15). PWZL-GFP and NT were used as controls. Total actin was used as a loading control.(c) Percent area covered by TRA-1-60-positive colonies with the MRG15KO compared to the NT group. Representative well images are shown above. Bar graphs show the mean, and error bars represent the standard error of the mean.

n=3, biological replicates. The Unpaired t-test calculates P-values. * denotes p<0.05. The exact p-values are 0.0486, 0.7441.

3.7 Coupling MRG15KO with SETD2 KD has combinatorial effect in terms of reprogramming efficiency

Next, I investigated whether MRG15 is works as downstream effector of H3K36me3 in context of cellular reprogramming. To test this hypothesis, I generated fibroblasts where both MRG15 and SETD2 was repressed by genetic means (MRG15 knockout via CRISPR-Cas9 and SETD2 knockdown via shRNA). Expression levels of both genes were verified using qPCR (Figure 3.8.a). I expected to observe no significant additive effect when both SETD2 and MRG15 levels are diminished if MRG15 is indeed working as a downstream of H3K36 tri-methylation However, MRG15KO cells that expressed SETD2 shRNA generated almost 2-fold more iPSC colonies than MRG15KO alone or SETD2 shRNA alone expression (Figure 3.8.b). I tested this hypothesis again using iSETD2, which yielded a similar result (Figure 3.8.c). This set of data suggest that MRG15's barrier function to reprograming is not dependent on SETD2.

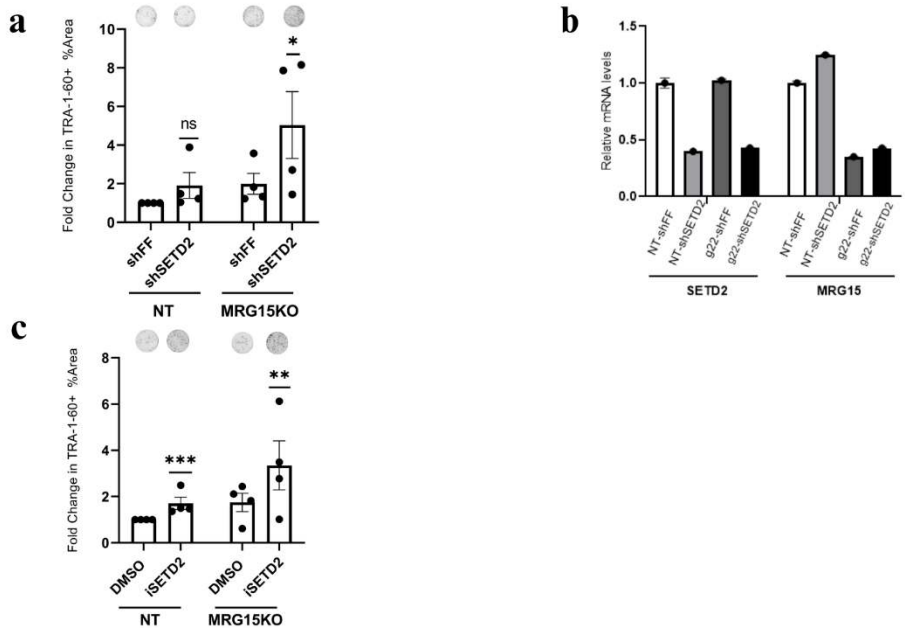


Figure 3.8 MRG15 knockout with SETD2 knockdown has an additive increase in reprogramming efficiency. (a)Fold change in expression levels of SETD2 and MRG15 upon SETD2 knockdown and MRG15 knockout compared to NT and shFF. (b) Fold

change in percent area covered by TRA-1-60-positive colonies with the MRG15KO and SETD2 knockdown normalized to the NT-shFF group. NT-FF compared to MRG15KO-FF and NT-SETD2 compared to MRG15KO-SETD2. Representative well images are shown above. Bar graphs show the mean, and error bars represent the standard error of the mean. n=3, biological replicates for each treatment with four technical replicates. The Two Way ANOVA calculates P-values. * denotes $p<0.05$. The exact p-values are 0.2297, 0.0178. (c) Fold change in percent area covered by TRA-1-60-positive colonies with the MRG15KO treated with either DMSO or iSETD2 normalized to the NT cells treated with the DMSO group. DMSO-treated NT cells compared to MRG15KO cells, and iSETD2-treated NT cells compared to MRG15KO cells. Representative well images are shown above. Bar graphs show the mean, and error bars represent the standard error of the mean. n=3, biological replicates for each treatment with four technical replicates. The Unpaired t-test calculates P-values. * denotes $p<0.05$, ** denotes $p<0.005$, *** denotes < 0.0005.The exact p-values are 0.0005, 0.0034.

3.8 MRG15KO has enriched reprogramming efficiency without Myc expression

After establishing that MRG15 is a novel epigenetic barrier working independently of SETD2, I investigated whether its loss can lead to the reprogramming in the absence of one or more of original Yamanaka factors. Previous studies with similar epigenetic barriers have yielded iPSC colonies in the absence of KLF-4 and MYC (Ebrahimi et al., 2019). I have tested only OCT4 and SOX2 expression; however, this condition did not yield any iPSC colonies, even in the presence of iDOT1L (Figure 3.9). I next tested omitting MYC and introduced OSK to control and MRG15 KO cells. MRG15KO generated significantly higher iPSC colonies than control cells in the presence of iDOT1L (Figure 3.9). This data implies that even though MRG15KO increases reprogramming efficiency; it cannot replace KLF4 and MYC on its own.

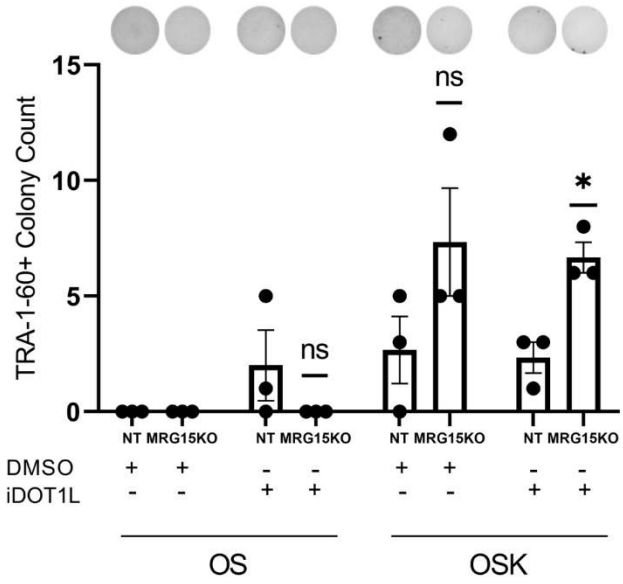


Figure 3.9 MYC is unessential for MRG15KO epigenetic reprogramming. The number of TRA-1-60 positive colonies with the MRG15KO compared to the NT group treated with OS, OSK DMSO, and iDOT1L. NT cells compared to MRG15KO cells in each group. The first group yielded no colony; thus, a t-test could not be performed. Representative well images are shown above. Bar graphs show the mean, and error bars represent the standard error of the mean. n=3, biological replicates. The Unpaired t-test calculates P-values. * denotes $p<0.05$. The exact p-values are 0.2606, 0.1648, 0.0101.

3.9 MRG15KO yielded an additive effect with a combination of iBRD9 and CBP30

Next, I wanted to study combinatorial effect of MRG15KO cells with previously established inhibitors such as iBRD9, İDOT1L and CBP30. Except for iDOT1L, the other two inhibitors showed increased reprogramming when combined with MRG15KO (Figure 3.10). This may suggest that MRG15KO can be coupled with various small molecules to improve reprogramming efficiency.

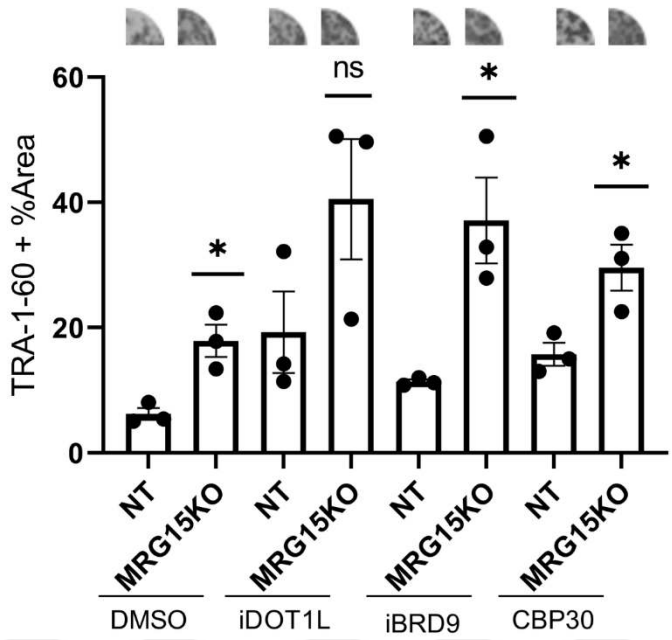


Figure 3.10 MRG15KO increases the iPSC yield of iBRD9 and CBP30.

Fold change in percent area covered by TRA-1-60-positive colonies with the indicated inhibitor treatment. MRG15KO was compared to NT cells in each treatment. Representative well images are shown above. Bar graphs show the mean, and error bars represent the standard error of the mean. n=3, biological replicates. P-values are calculated by Unpaired t-test. * denotes $p<0.05$. p-values from left to right are 0.0132, 0.1401, 0.0200, and 0.0280.

3.10 MRG15KO iPSC can be established with transient expression of OSKM factors

MRG15KO acts as a barrier to the epigenetic barrier of somatic to pluripotent state transition. However, it was unclear that MRG15KO iPSC is maintainable. I have established MRG15KO iPSCs from dh1f cells by using retroviruses that has Yamanaka factors with GFP reporters. OSKM silenced colonies are desired thus I isolated iPSC colonies that were GFP negative (Figure 3.11.a,b). I have isolated various clones and expanded three of them: clone 8, clone10, and clone12. I have verified that iPSC clones were MRG15KO by Western Blot analysis. There were low MRG15 signals in all three clones; however, compared to NO iPSC, they can be ignored (Figure 3.11.c). This data suggests that MRG15 does not affect iPSC viability.

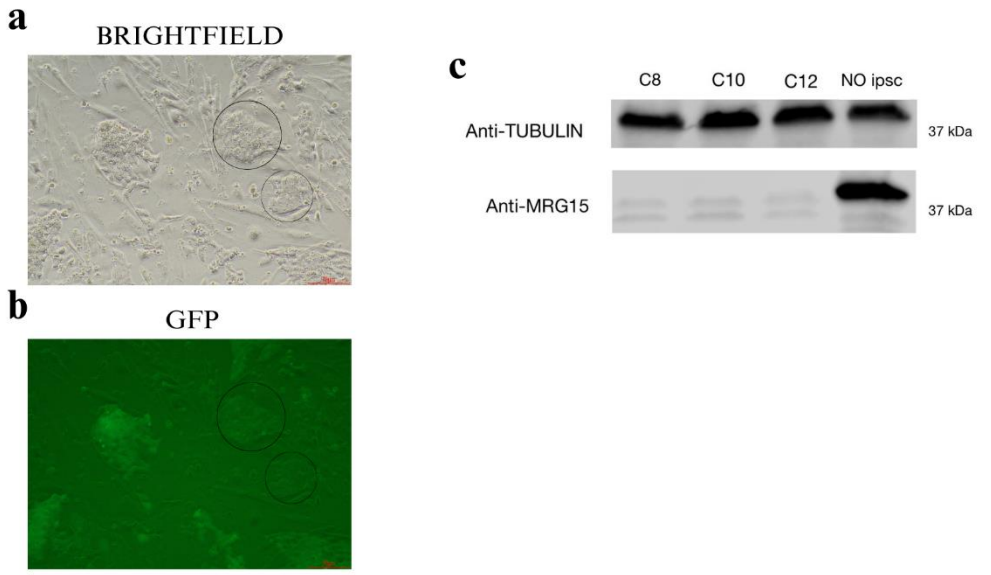


Figure 3.11 MRG15 is dispensable for iPSC viability. (a) Brighfield and GFP fluorescence images of colonies derived by MRG15KO show typical iPSC morphology and silencing of retroviral GFP transgene. (b) Western blot for MRG15 from iPSC clones generated from NO iPSC or MRG15KO fibroblasts. Tubulin serves as a loading control.

3.11 MRG15KO iPSC clones express all four pluripotency markers

To verify the selected clones were indeed iPSCs, I performed immunofluorescence staining for established pluripotency markers such as OCT4, SOX2, NANOG, and SSEA4. I have used NO iPSCs as a positive control (Alici-Garipcan et al., 2020). All clones were positive for the expression of pluripotency markers (Figure 3.12.a,b,c,d).

Additionally, I determined the mRNA levels of OCT4, SOX2, KLF4, and cMYC in selected MRG15KO iPSC clones using qPCR. Similar to immunostaining findings, qPCR results showed that the clones are positive for pluripotency markers (Figure 3.12.e,f). These sets of data further strengthen the conclusion that MRG15 is not necessary for iPSC generation and maintenance

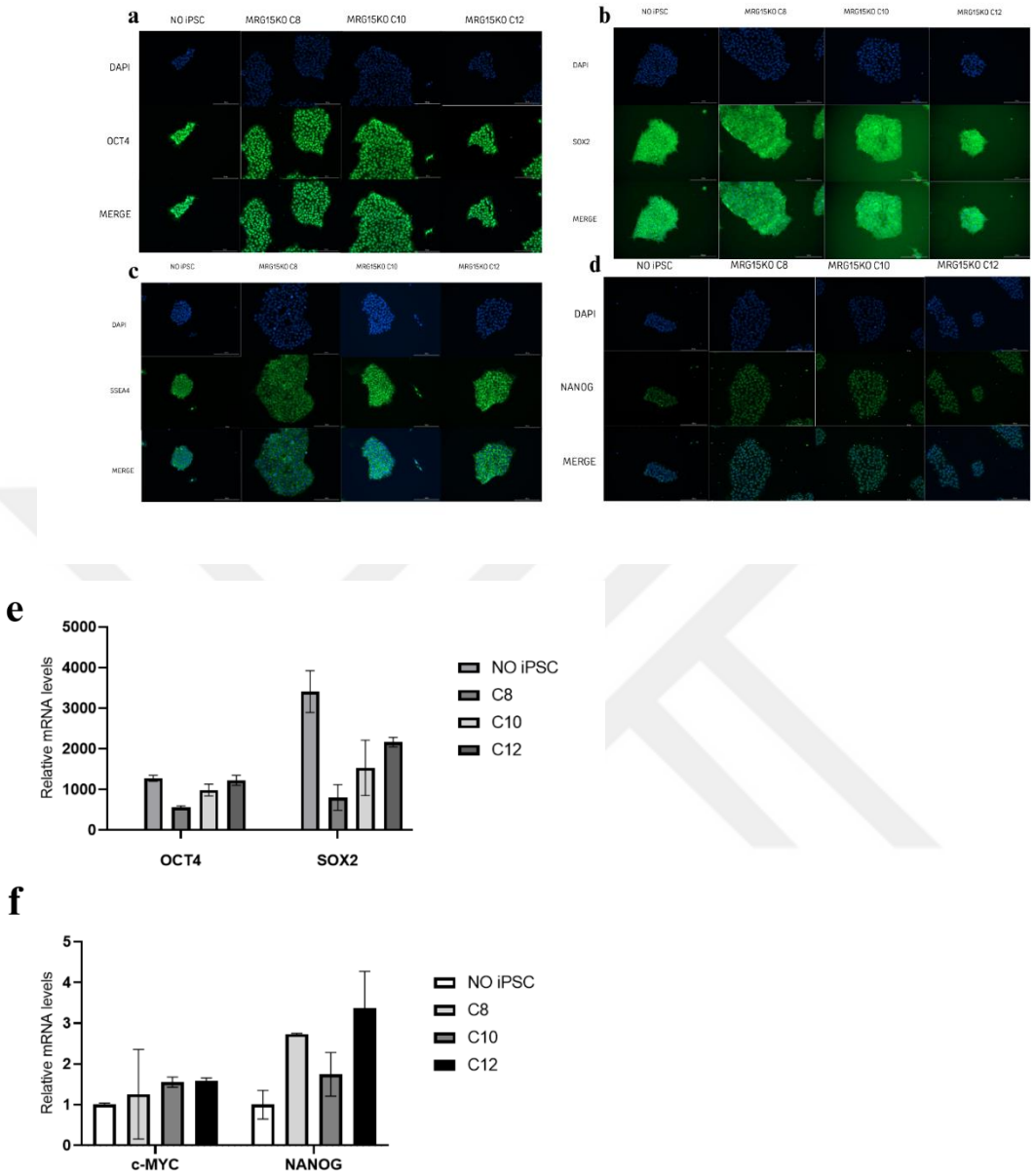
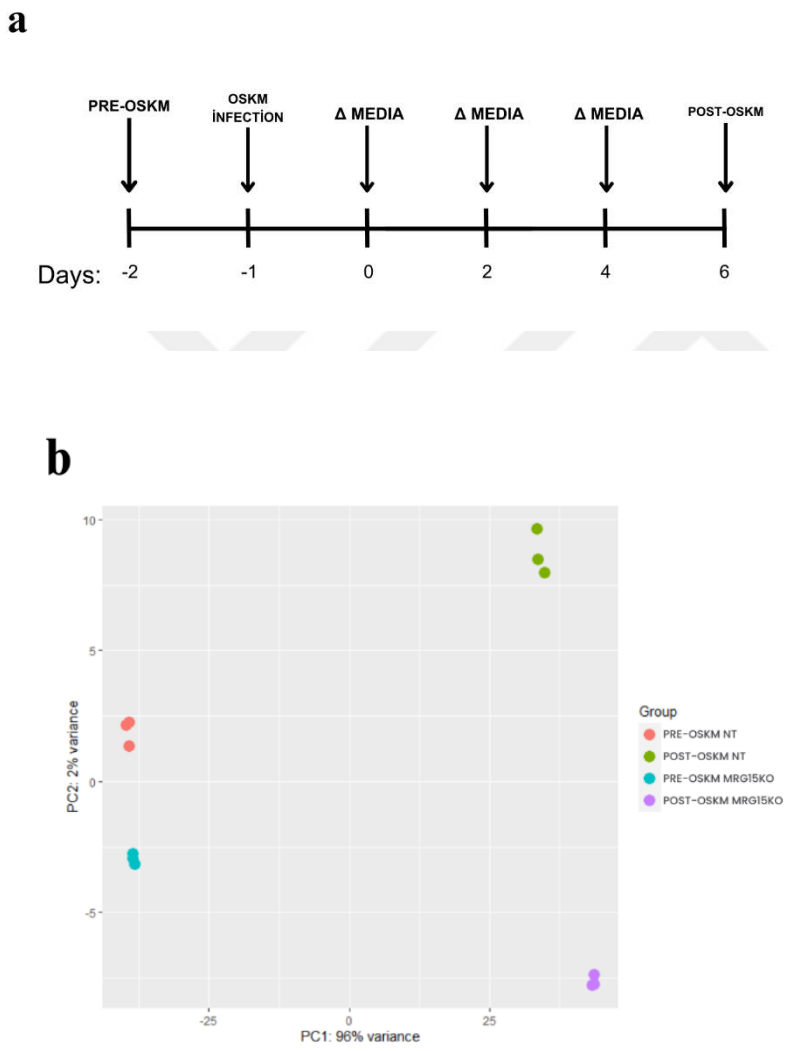


Figure 3.12 MRG15KO can produce viable iPSCs. Immunofluorescence for OCT4 (a), SOX2 (b), SSEA4 (c), and NANOG (d) in iPSCs generated from control NO iPSC and MRG15KO clones(8-10-12) cells. DAPI was used to stain the nuclei. Scale bars denote 200 μm. qPCR results for mRNA levels of MRG15KO clones for Oct4, Sox2(e) and Myc, Nanog (f). Expressions normalized to GAPDH mRNA levels.

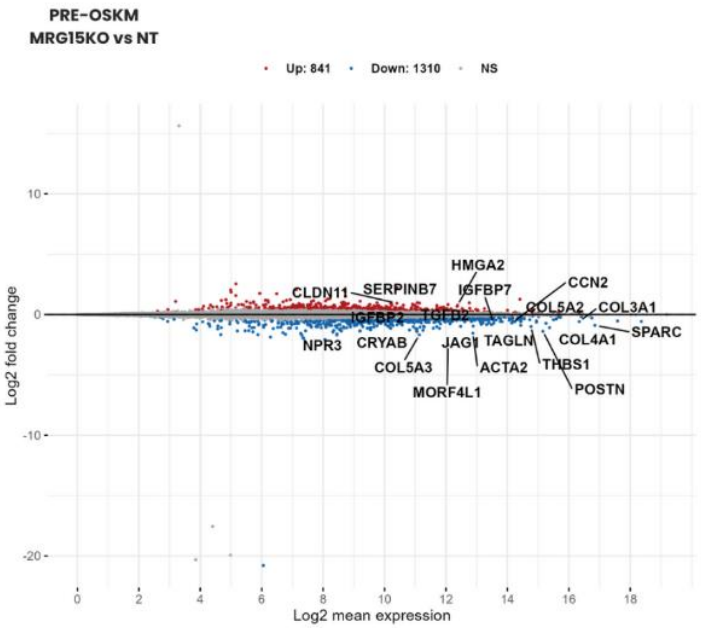
3.12 MRG15 KO changes the whole transcriptome

After establishing MRG15 as a barrier to epigenetic reprogramming and not a requisite factor for iPSC maintenance, I investigated the mechanism by which MRG15

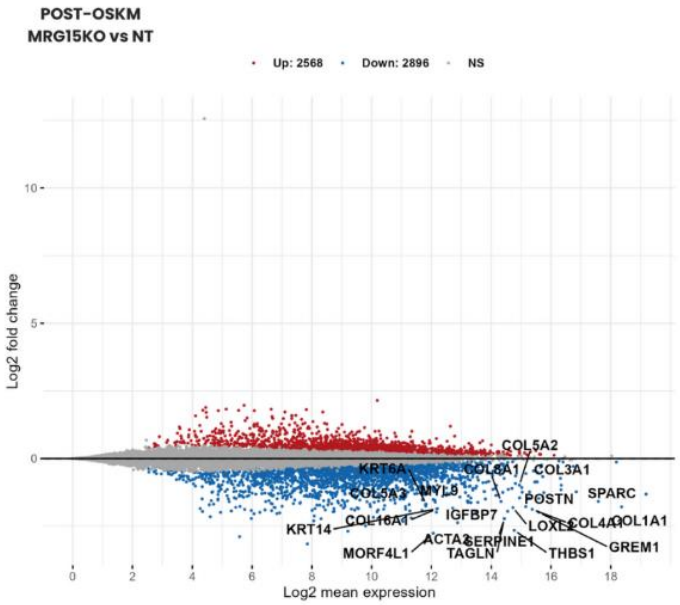
prevents reprogramming by performing RNA-sequencing on MRG15KO fibroblasts prior to (pre-OSKM) and 6 days after initiation of reprogramming (post-OSKM) (Figure 3.13.a). Replicates were clustered in the PCA plot, indicating technical reproducibility (Figure 3.13.b). MRG15KO changed global expression levels in both time points; however, this difference is more prominent in post-OSKM (Figure 3.13.c,d). This can imply that MRG15 has a role in protecting cellular identity.



c


PRE-OSKM
MRG15KO vs NT
Up: 841
Down: 1310
NS
Log2 fold change
Log2 mean expression
HMGA2
CCN2
CLDN11
SERPINB7
IGFBP7
COL5A2
COL3A1
IGFBP2
TGFB2
SPARC
NPR3
CRYAB
JAG1
TAGLN
COL4A1
COL5A3
ACTA2
THBS1
MORF4L1
POSTN

d


POST-OSKM
MRG15KO vs NT
Up: 2568
Down: 2896
NS
Log2 fold change
Log2 mean expression
COL5A2
KRT6A
COL6A1
COL3A1
COL5A3
MYL9
POSTN
SPARC
COL16A1
IGFBP7
LOXL2
COL4A1
COL1A1
KRT14
ACTA2
SERPINE1
GREM1
MORF4L1
TAGLN
THBS1

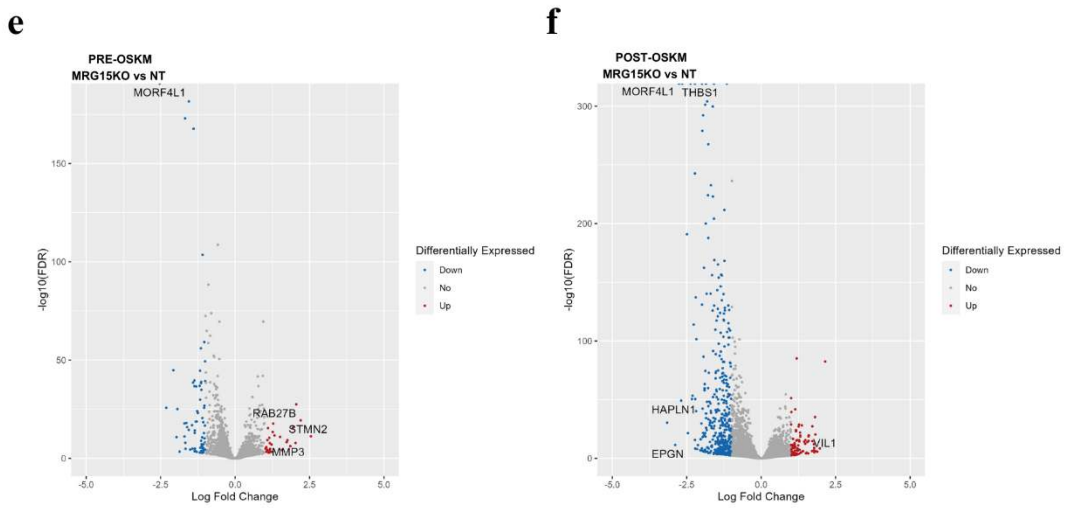


Figure 3.13 MRG15KO changes the expression profile of the cell both pre and post-OSKM. (a)The schematic depicts the time intervals for RNA-Seq sample collection. (b) Principal component analysis of transcriptome data with indicated cells. MA Plot for NT vs MRG15 KO pre-OSKM(c) and post-OSKM(d) samples. Volcano Plot for NT vs MRG15 KO pre-OSKM(e) and post-OSKM(f) samples

3.13 MRG15KO transcriptome enriched in pluripotency-related gene sets

Previous experiments demonstrated that MRG15 knock out cells enhance reprogramming efficiency (Figure 3.6, 3.9). I hypothesized that MRG15 knockout fibroblasts enriched in pluripotency related genes. In order test this hypothesis I utilized GSEA.

Gene Set Enrichment Analysis (GSEA) enhances traditional gene expression analysis by focusing on gene sets rather than individual genes. This method, refined for broader applicability, facilitates the interpretation of large-scale experiments, offering advantages over single-gene approaches. GSEA identifies pathways and processes, improving reproducibility and interpretability. It boosts signal-to-noise ratio by considering cross-correlations among gene set members and aids in defining gene subsets through leading-edge analysis. Comparisons with other tools highlight GSEA's unique features, such as considering all genes in an experiment and assessing significance

through class label permutations. The flexibility of GSEA is demonstrated by its molecular signature database, encompassing 34,450 gene sets based on various criteria, with the potential for further expansion through genetic and chemical perturbations.

In GSEA plots, pluripotency-related gene sets were also enriched in pre- and post-OSKM samples (Figure 3.14.a, c). In pre-OSKM samples pluripotency related genes such as RAB25, CD24, BMP2 and FOXO1 upregulated and in post-OSKM samples; LIN28A, DNMT3B and EPCAM. Likewise, fibroblast-specific gene sets were negatively enriched in MRG15KO samples (Figure 3.14.b, d). In pre-OSKM samples fibroblast related genes such as, COL5A3, POSTN, MYOCD downregulated and in post-OSKM samples; ACTA2, ADAM12, COL12A1. Even before transitioning to the pluripotent state, MRG15KO fibroblasts are closer to the pluripotent epigenetic landscape than the control group. However, as expected, the enrichment score is more striking in post-OSKM samples since they have received Yamanaka factors (Figure 3.14. c, d,e). This result supports the previous claim that MRG15KO fibroblasts are more prone to transition to iPSCs, and therefore, MRG15 is a roadblock against reprogramming.

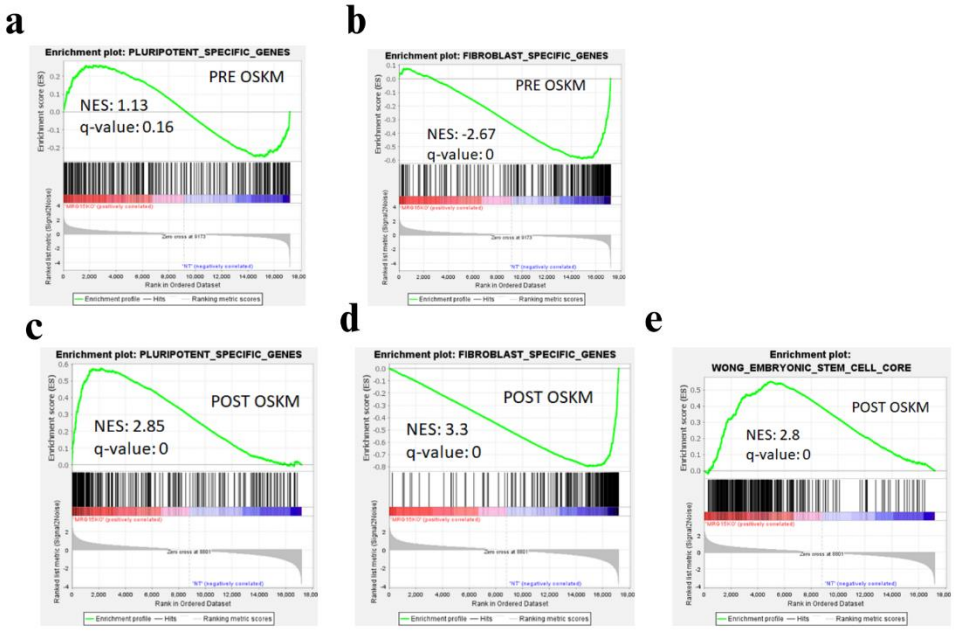


Figure 3.14 GSEA analysis indicates that MRG15KO transcriptome is more towards pluripotency than NT. (a) Gene set enrichment analysis (GSEA) of transcriptome data comparing MRG15KO fibroblasts to NT fibroblasts for pluripotent specific gene set in PRE-OSKM samples. NES: normalized enrichment score, q val: False discovery rate (FDR) q-value. (b) GSEA of transcriptome data comparing MRG15KO fibroblasts to NT

fibroblasts for fibroblasts specific gene set in PRE-OSKM samples. (c) GSEA of transcriptome data comparing MRG15KO fibroblasts to NT fibroblasts for pluripotent specific gene set in POST-OSKM samples. (d) GSEA of transcriptome data comparing MRG15KO fibroblasts to NT fibroblasts for fibroblast-specific gene set in POST-OSKM samples. (e) GSEA of transcriptome data comparing MRG15KO fibroblasts to NT fibroblasts for Wong Embryonic Stem Cell Core gene set in POST-OSKM samples.

3.14 MRG15KO transcriptome is enriched in cell cycle pathways

The GO enrichment graph increased cell cycle-related gene sets in pre-and post-OSKM samples (Figure 3.15.a & 3.16.a). However, pre-OSKM is more enriched in cell cycle-related gene sets than pos-OSKM samples (Figure 3.15.c,d,e).

After this analysis, I wanted to investigate this effect and performed a simple proliferation assay where I counted cells at three separate time points (Figure 3.15.f). MRG15KO cells showed similar linear regression in terms of proliferation compared to NT.

Another striking finding was that DREAM targets gene set enrichment in pre-and post-OSKM samples with high NES values (Figure 3.15.b & 3.16.b). DREAM complex homologs were first identified in C.elegnas and D.melanogaster(Sadasivam & Decaprio, 2013). This complex mainly consists of DP, RB-like protein, E2F, and MUVB, with many subunits from the LIN family(Sadasivam & Decaprio, 2013). Litovchik et al. first reported the DREAM complex in 2007 as an evolutionarily highly conserved cell cycle regulator complex(Litovchick et al., 2007). However, it was Fischer et al. who identified all DREAM complex gene targets with a meta-analysis approach in 2016, and I have used the same gene sets in my analysis(Fischer et al., 2016). DREAM complex occupies its target gene loci during G0, in other words, the quiescence state, and it is argued that the DREAM complex works as a guardian to transition from a quiescence state to a proliferative state(Sadasivam & Decaprio, 2013). Therefore, enrichment in the DREAM target gene set in MRG15KO implies that cells are not quiescent and lean towards a proliferative state.

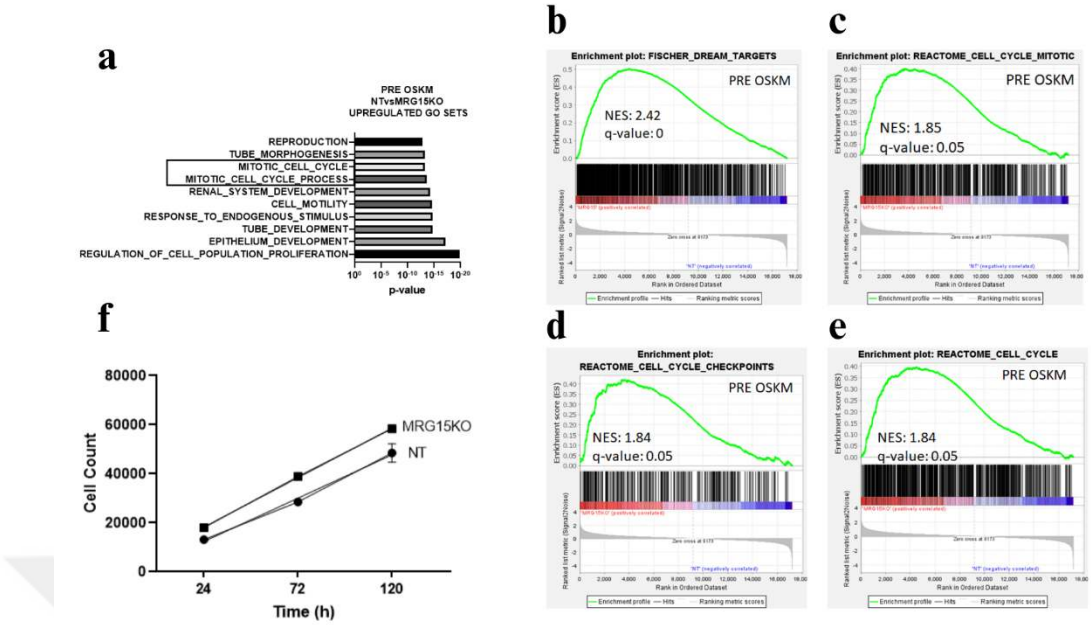


Figure 3.15 PRE-OSKM samples are enriched in cell cycle-related gene sets. (a) Top ten gene ontology (GO) gene sets enriched in upregulated genes upon MRG15KO in PRE-OSKM samples. P-values were calculated by hypergeometric distribution. (b) Gene set enrichment analysis (GSEA) of transcriptome data comparing MRG15KO fibroblasts to NT fibroblasts for Fischer DREAM targets gene set in PRE-OSKM samples. NES: normalized enrichment score, q val: False discovery rate (FDR) q-value. (c) Gene set enrichment analysis (GSEA) of transcriptome data comparing MRG15KO fibroblasts to NT fibroblasts for Reactome cell cycle mitotic gene set in PRE-OSKM samples. (d) Gene set enrichment analysis (GSEA) of transcriptome data comparing MRG15KO fibroblasts to NT fibroblasts for Reactome cell cycle checkpoints gene set in PRE-OSKM samples. (e) Gene set enrichment analysis (GSEA) of transcriptome data comparing MRG15KO fibroblasts to NT fibroblasts for Reactome Cell Cycle gene set in PRE-OSKM samples. (f) Proliferation assay, MRG15 compared to NT fibroblasts. n=3 biological replicates.

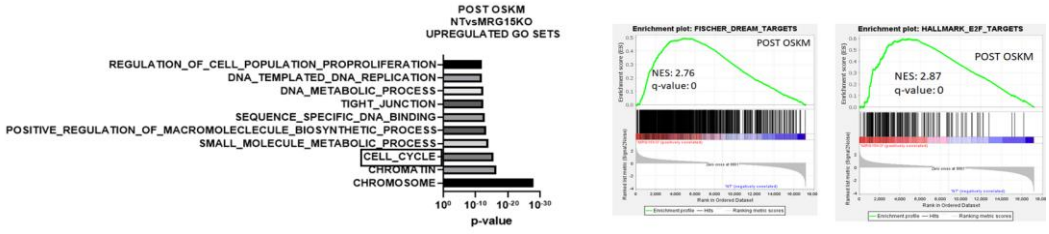


Figure 3.16 PSOT-OSKM samples have enriched DREAM target gene expression (a)Top ten gene ontology (GO) gene sets enriched in upregulated genes upon MRG15KO

in POST-OSKM samples. P-values were calculated by hypergeometric distribution. (b) Gene set enrichment analysis (GSEA) of transcriptome data comparing MRG15KO fibroblasts to NT fibroblasts for Fischer DREAM targets gene set in POST-OSKM samples. NES: normalized enrichment score, q val: False discovery rate (FDR) q-value. (c) Gene set enrichment analysis (GSEA) of transcriptome data comparing MRG15KO fibroblasts to NT fibroblasts for Hallmark E2F targets gene set in POST-OSKM samples.

3.15 MRG15KO does not change the global mRNA splicing profile

It has previously been reported that MRG15 regulates alternative splicing with PTB (Luco et al., 2010). Bayırbaşı B. has observed that upon MRG15 depletion, certain protein alternative splicing direction changes. She has focused on proteins with somatic-specific and pluripotent-specific isoforms. MBD2 was one of the prominent candidates since MRG15KO has significantly higher pluripotent isoform expression than NT. She hypothesized that the MRG15 working mechanism could be achieved through regulating alternative splicing of pluripotent-specific isoforms. However, rMAT analysis on MRG15KO RNA-Seq results, there were no significant differences between NT and MRG15KO cells in pre-and post-OSKM samples (Figure 3.17 b& c). This implied that MRG15KO did not alter alternative splicing globally; however, it can regulate certain transcripts.

Interestingly, the GSEA showed that cellular pathways such as mRNA processing or spliceosome are enriched in MRG15KO cells (Figure 3.18 a,b,c,d,e). Notably, these pathways are enriched only in post-OSKM samples (Figure 3.18 a,b,c,d,e). This may imply that RNA-Seq analysis later points of reprogramming could reveal global changes in alternative splicing profiles since, on day six, spliceosome genes are newly upregulated.

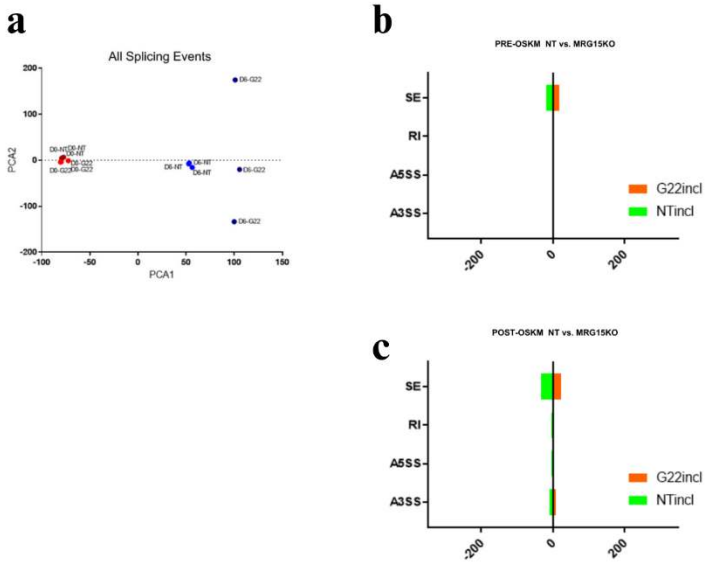


Figure 3.17 MRG15 KO does not lead to a significant change in Alternative Splicing modes. (a) Principle component analysis of all alternative splicing events with indicated cells.(b) Replicate Multivariate analysis of Transcript Splicing (rMAT) graph for PRE-OSKM NT and MRG15KO sample. n=3 biological replicates. (c)Replicate Multivariate analysis of Transcript Splicing (rMAT) graph for POST-OSKM NT and MRG15KO sample. n=3 biological replicates

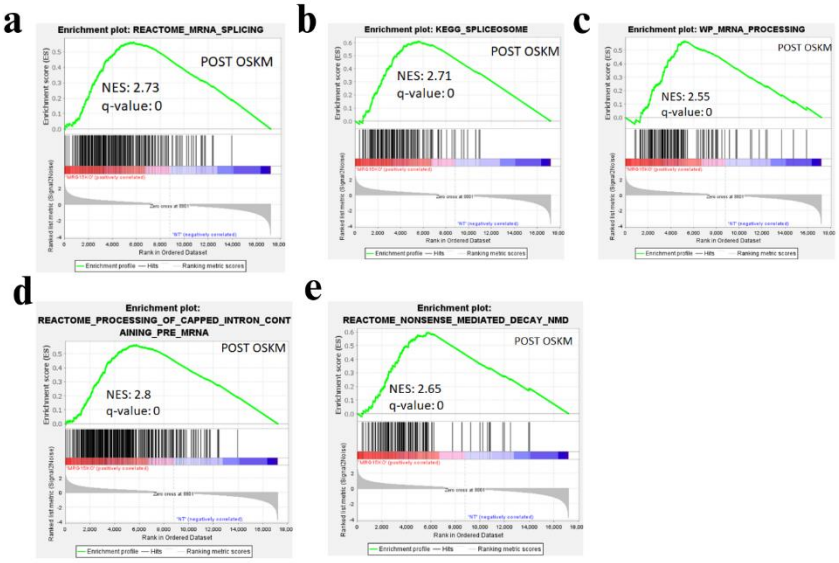


Figure 3.18 MRG15 KO cells have increased expression of alternative splicing-related gene sets. (a) Gene set enrichment analysis (GSEA) of transcriptome data

comparing MRG15KO fibroblasts to NT fibroblasts for Reactome mRNA splicing gene set in POST-OSKM samples. NES: normalized enrichment score, q val: False discovery rate (FDR) q-value. (b) GSEA of transcriptome data comparing MRG15KO fibroblasts to NT fibroblasts for KEGG Splicesome gene set in POST-OSKM samples. (c) GSEA of transcriptome data comparing MRG15KO fibroblasts to NT fibroblasts for WP mRNA processing gene set in POST-OSKM samples. (d) GSEA of transcriptome data comparing MRG15KO fibroblasts to NT fibroblasts for Reactome processing of capped intron-containing pre-mRNA targets gene set in POST-OSKM samples. (e) GSEA of transcriptome data comparing MRG15KO fibroblasts to NT fibroblasts for Reactome nonsense-mediated decay gene set in POST-OSKM samples.

3.16 MRG15KO cells have increased histone acetylation

In the last part of my thesis project, I investigated MRG15's relationship with other chromatin-associated factors. As mentioned previously, MRG15 has been reported to be found in HDAC2/SIN3B complex in vitro CRYO-EM analysis, and also in vivo IP analysis shows MRG15's association with HDAC2/SIN3B complex (Feoli et al., 2021). It was reported that HDAC2 is responsible for the deacetylation of H3K9, H3K14, and H3K27(Adams et al., n.d.). I have observed the most striking difference in H3K14ac levels as upon MRG15 depletion acetylation intensity increased 4-fold (Figure 3.19 b). This can imply that MRG15KO leads to lower HDAC activity and resulting change in various histone marks that lead to a switch in tissue-specific gene expression.

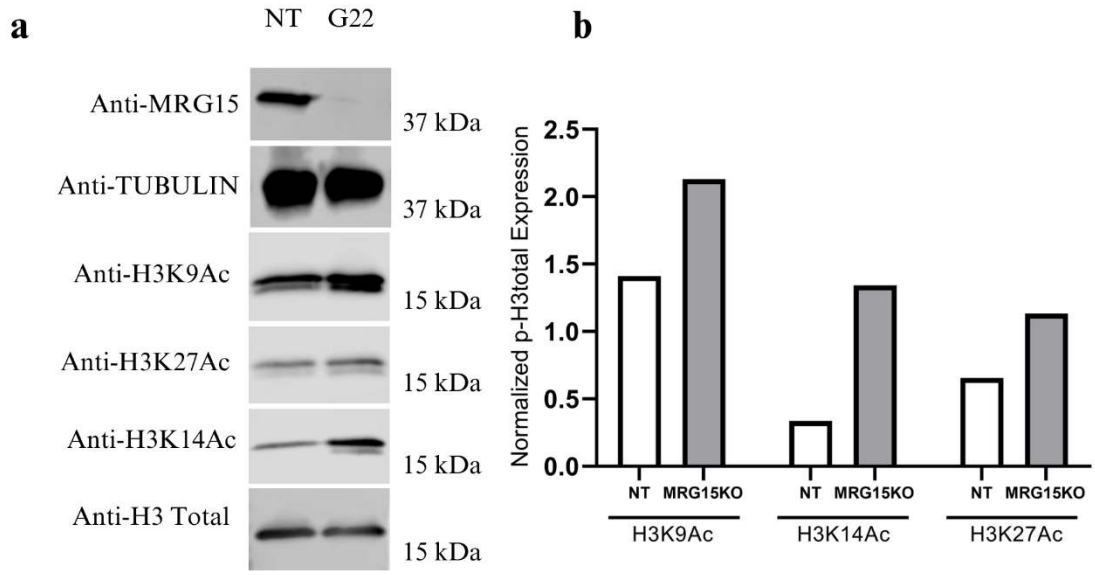


Figure 3.19 MRG15KO leads to an increase in acetylation in specific histone marks. (a) Western blots for histone H3 lysine 9 acetylation (H3K9ac), histone H3 lysine 27 acetylation (H3K27ac), and H3 lysine 14acetylation (H3K14ac) in NT (left) or MRG15KO (right) for the. Total histone H3 was used as a loading control. (b) Qualitative representation of Western Blot signals derived from signal intensity area in ImageJ. Signals are normalized to each cell's H3 signal.

Chapter 4:
DISCUSSION

My thesis examined how MRG15 deletion promotes somatic cell reprogramming into pluripotent cells. The primary inquiry of this thesis arose from a prior investigation in which the inhibition of SETD2 was discovered to increase the efficiency of reprogramming, and the knockout of MRG15 replicated this characteristic (Ozcimen, 2018; Bayırbaşı, 2021). MRG15 is among the many receptors of H3K36me3. Thus, this study proposes that MRG15 may hinder somatic cell reprogramming due to its capabilities as a reader of H3K36me3. The discovery was significant as the precise function of MRG15 in reprogramming has not yet been fully understood.

The significant lack of efficiency in somatic cell reprogramming indicates the existence of obstacles that impede the reprogramming process in somatic cells. Epigenetic alterations and the entities that control them are significant contenders for these obstacles. MRG15 is a small, stable nuclear protein in the SIN3B-HDAC2 and TIP60-HAT complexes(Zhang, Zhao, et al., 2006). These two complexes are responsible for the acetylation and deacetylation of many histone marks, activating or deactivating gene expression(Chen et al., 2009). This process is crucial for regulating cell proliferation, apoptosis, and lineage commitment(Bainor et al., 2018; Garcia-Sanz et al., 2014; Park et al., 2010). Studying the role of MRG15 in reprogramming is vital due to its involvement in maintaining a complicated equilibrium of histone acetylation.

4.1 SETD2 inhibitor can replace SETD2 knockdown

The previous studies conducted by Özçimen B. revealed that upon the silencing of the SETD2 gene, there is a reduction in H3K36me3 levels while an increase in H3K36me2 levels occurs. This shift is attributed to the diminished trimethylation activity and an enhancement in epigenetic reprogramming efficiency. To assess the catalytic activity of a SETD2 inhibitor (EPZ-719), I treated dh1f cells with varying concentrations of the inhibitor, which exists in two isotopic forms (isotope 1 trans stereoisomer and isotope 2 cis stereoisomer). The results demonstrated that EPZ-719-2 exhibited greater catalytic repression, leading to a decrease in H3K36me3 levels across all tested concentrations compared to isotope 1.

Furthermore, the inability of isotope 1 to efficiently reduce tri-methylation at concentrations below 10 uM hinted at its lower catalytic activity. This data reaffirms that SETD2 inhibitors effectively suppress SETD2 catalytic activity, considering SETD2 as the sole known tri-methylase responsible for catalyzing the tri-methyl mark. Consequently, the cellular response to SETD2 inhibition manifested in increased H3K36me2 levels when tri-methylation was impaired, resembling observations in the existing literature.

Next, I focused on exploring the impact of SETD2 inhibition on reprogramming efficiency. Through time course and dose treatment experiments, I observed that applying iSETD2 during the early stages of reprogramming significantly increased induced pluripotent stem cell (iPSC) colony formation. Notably, EPZ-719-2 exhibited the highest potency in enhancing reprogramming efficiency during the initial weeks of the process. This data suggests that rather than irrevocably decreasing SETD2 protein levels, using SETD2 inhibitor three times can achieve the same effect in a much safer manner. Thus, using iSETD2 in a limited time prevents us from disrupting the non-amendable epigenetic landscape.

4.2 MRG15 and SETD2 act as an epigenetic barrier through different pathways

Bayırbaşı B. has hypothesized that MRG15 is a downstream effector of SETD2 since MRG15 is the most powerful candidate among H3K36me3 readers regarding iPSC generation. MRG15's role in cellular reprogramming has not been reported yet; however, Pereira et al. indicate this protein's role in senescence regulation and alternative splicing(Luco et al., 2010; Tominaga et al., 2010). Both biological pathways can impede cellular reprogramming.

I have performed two experiments to test if MRG15 is working downstream of SETD2. First, I used sgRNA and shRNA to decrease expression levels of both SETD2 and MRG15 and perform cellular reprogramming. MRG15KO-SETD2KD fibroblasts produce more iPSCs than MRG15KO or SETD2KD alone. In the second experiment setup, I used SETD2 inhibitor instead of shRNAs. I have verified that iSETD2 inhibits catalytic activity and increases reprogramming efficiency. MRG15KO treated with iSETD2 had more TRA-1-60+ cells than MRG15KO cells treated with DMSO control.

These data suggest that MRG15 has different routes of regulation other than the H3K36me3 reader.

In a 2006 paper, Zhang et al. performed in vivo binding experiments with MRG15 and observed MRG15 chromodomain binding to H3K36me3 and H3K36me2(Zhang, Du, et al., 2006). Later, in 2015, Ruan et al. observed that Eaf3 MRG15 yeast homolog binding to H3K36me2 is almost as strong as H3K36me3 in vivo(Ruan et al., 2015). These findings illustrate that MRG15 is not only an H3K36me3 reader but also an H3K36me2 reader. In a 2020 paper, Difiore et al. investigated the roles of H3K36 methylation states in yeast and stated that for Eaf3 (MRG15 homolog), binding to H3K36me3 or me2 is interchangeable(DiFiore et al., 2020). Thus, inhibition of SETD2 does not alter MRG15 reader role, and MRG15 binds to H3K36me2 instead. However, the absence of MRG15 may alter the binding of HDAC and HAT to the chromatin completely and have a different downstream effect than SETD2 knockdown.

4.3 MRG15 is dispensable for iPSC survival and maintenance

MRG15KO serves as a hindrance to the epigenetic barrier during the transition from somatic to pluripotent states. To assess the viability of MRG15KO-induced pluripotent stem cells (iPSCs), I generated MRG15KO iPSCs from dh1f cells through transient expression of Yamanaka factors with GFP reporters. The transient expression was lost after two weeks of infection, and GFP-negative iPSC colonies were isolated. Various clones were obtained, and three of them, namely clone 8, clone 10, and clone 12, were expanded. Western Blot analysis confirmed MRG15KO in these clones, with low MRG15 signals compared to NO iPSC, suggesting that MRG15 does not significantly impact iPSC viability.

MRG15KO iPSC clones were subjected to immunofluorescence (IF) staining for pluripotency markers OCT4, SOX2, NANOG, and SSEA4 to validate iPSC status further. All clones exhibited similar signaling patterns to NO iPSCs across these markers. Additionally, RT-PCR analysis, replacing SSEA4 with c-MYC, reiterated the positive expression of pluripotency markers in MRG15KO clones. These results affirm that MRG15 is dispensable for generating and maintaining viable iPSCs.

4.4 MRG15 regulates alternative splicing towards late reprogramming

Previous reports have highlighted the role of MRG15 in regulating alternative splicing, particularly in association with PTB (polypyrimidine tract-binding protein)(Luco et al., 2010). Building upon this, observations by Bayırbaşı B. indicated that MRG15 depletion led to changes in the alternative splicing direction of specific proteins, focusing on those exhibiting somatic-specific and pluripotent-specific isoforms. MBD2 emerged as a notable candidate, given that MRG15KO displayed significantly higher expression of pluripotent isoforms than non-targeted (NT) cells. The working hypothesis suggested that MRG15 might achieve its effects by modulating the alternative splicing of pluripotent-specific isoforms (Bayırbaşı,2021).

However, the results of rMAT analysis on MRG15KO RNA-Seq data revealed no significant differences between NT and MRG15KO cells in pre- and post-OSKM samples (Figure 3.17 b&c). This suggested that MRG15KO did not induce global alterations in alternative splicing; rather, it appeared to regulate specific transcripts selectively.

Interestingly, the Gene Set Enrichment Analysis (GSEA) plot indicated that MRG15KO cells exhibited enriched cellular pathways related to mRNA processing and the spliceosome (Figure 3.18 a,b,c,d,e). Notably, these pathway enrichments were observed only in post-OSKM samples. This observation implies that RNA-Seq analysis at later stages of reprogramming may unveil global changes in alternative splicing profiles. Specifically, on day six, splicesome genes showed a newly upregulated pattern, suggesting a dynamic shift in alternative splicing profiles during the later stages of the reprogramming process.

In their study, Ohta et al. (2013) conducted a comprehensive examination of splicing pattern changes during the reprogramming of mouse embryonic fibroblasts (MEFs) into induced pluripotent stem cells (iPSCs)(Ohta et al., 2013). They employed high-throughput RNA sequencing to analyze transcriptome dynamics at various reprogramming stages(Ohta et al., 2013). This approach allowed them to identify distinct splicing transitions correlating with the reprogramming process(Ohta et al., 2013). The study revealed that these transitions were temporally coordinated and could be classified into different subgroups, suggesting a complex yet organized regulation of gene expression during cell reprogramming(Ohta et al., 2013). These subgroups were MEF (

pre-OSKM), Day 4 (SSEA1-), Day 8 (SSEA4+), Day 14 (Nanog +), and iPSC. They have analyzed many genes for splicing events and divided their exclusion and inclusion ratio analysis(Ohta et al., 2013). Although splicing switches have been observed in each stage of cellular reprogramming, reprogrammed cells tend to have this switch late reprogramming around day 8-14. MRG15KO has increased alternative splicing regulator expression on day 6; this could translate into a significant global splicing profile difference after day 8.

4.5 MRG15 is responsible for cell cycle regulation

To understand MRG15's role in the epigenetic landscape, I collected samples of MRG15KO cells for RNA Sequencing. The GO enrichment analysis revealed increased cell cycle-related gene sets in pre-and post-OSKM samples, as depicted in Figure 3.15. a and 3.16.a. Notably, the pre-OSKM samples exhibited a higher enrichment in cell cycle-related gene sets than post-OSKM samples, as shown in Figure 3.15.c, d, and e. To investigate this observation further, I conducted a proliferation assay, counting cells at three distinct time points (Figure 3.15.f). Interestingly, MRG15KO cells displayed a similar linear regression in proliferation compared to the non-targeted (NT) cells. Another noteworthy discovery was the enrichment of the DREAM targets gene set in pre-and post-OSKM samples, characterized by high NES values, as illustrated in Figures 3.15.b and 3.16.b.

A 2018 paper by Baynor et al. provided evidence that the scaffold protein Sin3B forms an association with the DREAM complex, a master regulator controlling the transcriptional activity of cell-cycle genes during quiescence(Bainor et al., 2018). Their findings suggest that Sin3B carries repressive activity linked to the DREAM complex, as its inactivation leads to hyper-acetylation of DREAM target promoter regions and the transcriptional derepression of corresponding genes(Bainor et al., 2018). Despite the inactivation of Sin3B not being sufficient for quiescent cells to resume proliferation, it is suggested that under specific conditions, Sin3B inactivation may allow quiescent cells to re-enter the cell cycle(Bainor et al., 2018). Interestingly, Sin3B inactivation prevents oncogene-induced senescence, resembling the effect of DREAM complex inactivation. In my observations, MRG15KO increased DREAM target gene expression similar to the

Sin3B knock-out phenotype Baynor et al. have observed(Bainor et al., 2018). Although the authors did not mention MRG15 as part of the Sin3b HDAC complex, MRG15 may be important for Sin3B to locate and interact with the DREAM complex, resulting in a weaker effect on DREAM repression. This may suggest that upon OSKM infection, MRG15KO DREAM targets additionally force cells to proliferate and thus yield higher iPSC generation.

4.6 MRG15KO leads hyperacetylated H3K14

In the concluding segment of my thesis project, I delved into the relationship between MRG15 and other SIN3B/HDAC2. Previous studies had indicated that MRG15 is present in the HDAC2/SIN3B complex, as revealed by in vitro CRYO-EM analysis, and in vivo IP analysis further affirmed MRG15's association with the HDAC2/SIN3B complex(Jelinic et al., 2011; Wan et al., 2023). HDAC2 is known for its role in the deacetylation of histone marks, specifically H3K9, H3K14, and H3K27 (Wan et al., 2023). Upon depleting MRG15, I observed elevated acetylation levels in all three histone marks (Figure 3.19 a,b). The most notable difference was observed in H3K14ac levels (Figure 3.19 b).

In their 2012 study, Karmodiya et al. introduced a new H3K14ac antibody, revealing the co-occurrence of H3K9ac and H3K14ac with other active histone modifications in mES cells (Karmodiya et al., 2012). Their findings suggest a mechanistic cross-talk between H3K4me3 and H3 acetylation, influencing Pol II transcription initiation. The study highlights the correlation between H3K9ac and H3K14ac at promoters with CpG content, emphasizing their role in regulating gene activity (Karmodiya et al., 2012). Additionally, H3K9ac and H3K14ac are identified as markers for active enhancers, distinguishing them from inactive/poised enhancers. The study also demonstrates the presence of H3K9ac and H3K14ac at bivalent promoters and reveals that H3K14ac, unlike H3K9ac, is selectively associated with repressive marks on a subset of inactive promoters (Karmodiya et al., 2012). Overall, the research underscores the diverse roles of H3K9ac and H3K14ac in gene regulation, offering insights into their dynamic involvement in chromatin modifications (Karmodiya et al., 2012). Thus, due to the

complex nature of this histone mark effect of H3K14 hyperacetylation, it should be studied further to understand its possible effect on iPSC generation.

In summary, this thesis examined how loss of MRG15 improves the reprogramming of human somatic cells. RNA-Seq analysis demonstrated that MRG15 controls the expression of cell cycle genes, and when MRG15 is deleted, cells transition towards a transcriptome profile that promotes cell proliferation. While MRG15 does not alter the overall alternative splicing pattern on a global scale, it can influence later reprogramming phases. Therefore, the data suggest that the deletion of MRG15 impacts reprogramming efficiency by controlling the expression of DREAM target genes.

BIBLIOGRAPHY

Adams, G. E., Chandru, A., & Cowley, S. M. (2018). Co-repressor, co-activator and general transcription factor: The many faces of the Sin3 histone deacetylase (HDAC) complex. *Biochemical Journal*, *475*(24), 3921–3932. https://doi.org/10.1042/BCJ20170314

Adams, M. K., Banks, C. A. S., Thornton, J. L., Kempf, C. G., Zhang, Y., Miah, S., Hao, Y., Sardiu, M. E., Killer, M., Hattem, G. L., Murray, A., Katt, M. L., Florens, L., & Washburn, M. P. (2020). Differential Complex Formation via Paralogs in the Human Sin3 Protein Interaction Network. *Molecular and Cellular Proteomics*, *19*(9), 1468–1484. https://doi.org/10.1074/mcp.RA120.002078

Ae, S. N. G., & Pereira-Smith, O. (2008). MRGing Chromatin Dynamics and Cellular Senescence. *Cell Biochem Biophys* , *50*, 133–141. https://doi.org/10.1007/s12013-008-9006-7

Alici-Garipcan, A., Özçimen, B., Süder, I., Ülker, V., Önder, T. T., & Özören, N. (2020). NLRP7 plays a functional role in regulating BMP4 signaling during differentiation of patient-derived trophoblasts. *Cell Death & Disease*. https://doi.org/10.1038/s41419-020-02884-1

Arabacı, D. H., Terzioğlu, G., Bayırbaşı, B., & Önder, T. T. (2021). Going up the hill: chromatin-based barriers to epigenetic reprogramming. *FEBS Journal*, *288*(16), 4798–4811. https://doi.org/10.1111/febs.15628

Bainor, A. J., Saini, S., Calderon, A., Casado-Polanco, R., Giner-Ramirez, B., Moncada, C., Cantor, D. J., Ernlund, A., Litovchick, L., & David, G. (2018). The HDAC-Associated Sin3B Protein Represses DREAM Complex Targets and Cooperates with APC/C to Promote Quiescence. *Cell Reports*, *25*(10), 2797-2807.e8. https://doi.org/10.1016/j.celrep.2018.11.024

Bayırbaşı, Büşra. *Understanding How Loss of Histone H3 Lysine 36 Trimethylation Enhances Somatic Cell Reprogramming*. Koç Üniversitesi, Graduate School of Science and Engineering, Master of Science Thesis, 2021.

Bertram, M. J., Bé Rubé, N. G., Hang-swanson, X., Ran, Q., Leung, J. K., Bryce, S., Spurgers, K., Bick, R. J., Baldini, A., Ning, Y., Clark, L. J., Parkinson, E. K., Barrett, J. C., Smith, J. R., Pereira-smith, O. M., & Gough, P. (1999). Identification of a Gene That Reverses the Immortal Phenotype of a Subset of Cells and Is a

Member of a Novel Family of Transcription Factor-Like Genes †. In *MOLECULAR AND CELLULAR BIOLOGY* (Vol. 19, Issue 2).

Bleuyard, J. Y., Fournier, M., Nakato, R., Couturier, A. M., Katou, Y., Ralf, C., Hester, S. S., Dominguez, D., Rhodes, D., Humphrey, T. C., Shirahige, K., & Esashi, F. (2017). MRG15-mediated tethering of PALB2 to unperturbed chromatin protects active genes from genotoxic stress. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(29), 7671–7676. https://doi.org/10.1073/pnas.1620208114

Buganim, Y., Faddah, D. A., & Jaenisch, R. (2013). Mechanisms and models of somatic cell reprogramming. *Nature Publishing Group*. https://doi.org/10.1038/nrg3473

Carrozza, M. J., Li, B., Florens, L., Suganuma, T., Swanson, S. K., Lee, K. K., Shia, W. J., Anderson, S., Yates, J., Washburn, M. P., & Workman, J. L. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell*, *123*(4), 581–592. https://doi.org/10.1016/j.cell.2005.10.023

Chen, M., Takano-Maruyama, M., Pereira-Smith, O. M., Gaufo, G. O., & Tominaga, K. (2009). MRG15, a component of HAT and HDAC complexes, is essential for proliferation and differentiation of neural precursor cells. *Journal of Neuroscience Research*, *87*(7), 1522–1531. https://doi.org/10.1002/jnr.21976

Chen, Z., Raghoonundun, C., Chen, W., Zhang, Y., Tang, W., Fan, X., & Shi, X. (2018). SETD2 indicates favourable prognosis in gastric cancer and suppresses cancer cell proliferation, migration, and invasion. *Biochemical and Biophysical Research Communications*, *498*(3), 579–585. https://doi.org/10.1016/j.bbrc.2018.03.022

Devoucoux, M., Roques, C., Lachance, C., Lashgari, A., Joly-Beauparlant, C., Jacquet, K., Alerasool, N., Prudente, A., Taipale, M., Droit, A., Lambert, J. P., Hussein, S. M. I., & Côté, J. (2022). MRG Proteins Are Shared by Multiple Protein Complexes With Distinct Functions. *Molecular and Cellular Proteomics*, *21*(7). https://doi.org/10.1016/j.mcpro.2022.100253

DiFiore, J. V., Ptacek, T. S., Wang, Y., Li, B., Simon, J. M., & Strahl, B. D. (2020). Unique and Shared Roles for Histone H3K36 Methylation States in Transcription Regulation Functions. *Cell Reports*, *31*(10). https://doi.org/10.1016/j.celrep.2020.107751

Doyon, Y., & Côté, J. (2004). The highly conserved and multifunctional NuA4 HAT complex. In *Current Opinion in Genetics and Development* (Vol. 14, Issue 2, pp. 147–154). https://doi.org/10.1016/j.gde.2004.02.009

Ebrahimi, A., Sevinç, K., Sevinç, G. G., Cribbs, A. P., Philpott, M., Uyulur, F., Morova, T., Dunford, J. E., Göklemez, S., Arı, Ş., Oppermann, U., & Önder, T. T. (2019). Bromodomain inhibition of the coactivators CBP/EP300 facilitate cellular reprogramming. *Nature Chemical Biology*, *15*(May), 519–528. https://doi.org/10.1038/s41589-019-0264-z

Feoli, A., Pisapia, V., Viviano, M., Castellano, S., Bartoschik, T., & Sbardella, G. (2021). Development of a Microscale Thermophoresis-Based Method for Screening and Characterizing Inhibitors of the Methyl-Lysine Reader Protein MRG15. *SLAS Discovery*, *26*(1), 77–87. https://doi.org/10.1177/2472555220949166

Fischer, M., Grossmann, P., Padi, M., & DeCaprio, J. A. (2016). Integration of TP53, DREAM, MMB-FOXM1 and RB-E2F target gene analyses identifies cell cycle gene regulatory networks. *Nucleic Acids Research*, *44*(13), 6070–6086. https://doi.org/10.1093/nar/gkw523

Garcia-Sanz, P., Quintanilla, A., Lafita, M. C., Moreno-Bueno, G., García-Gutierrez, L., Tabor, V., Varela, I., Shiio, Y., Larsson, L. G., Portillo, F., & Leon, J. (2014). Sin3b Interacts with Myc and decreases Myc levels. *Journal of Biological Chemistry*, *289*(32), 22221–22236. https://doi.org/10.1074/jbc.M113.538744

Gerri, C., Menchero, S., Mahadevaiah, S. K., Turner, J. M. A., & Niakan, K. K. (2020). Human Embryogenesis: A Comparative Perspective. *Annu. Rev. Cell Dev. Biol. 2020*, *36*, 411–440. https://doi.org/10.1146/annurev-cellbio-022020

Guan, H., Wang, P., Zhang, P., Ruan, C., Ou, Y., Peng, B., Zheng, X., Lei, J., Li, B., Yan, C., & Li, H. (2023). Diverse modes of H3K36me3-guided nucleosomal deacetylation by Rpd3S. *Nature*, *620*(7974), 669–675. https://doi.org/10.1038/s41586-023-06349-1

Guan, J., Wang, G., Wang, J., Zhang, Z., Fu, Y., Cheng, L., Meng, G., Lyu, Y., Zhu, J., Li, Y., Wang, Y., Liuyang, S., Liu, B., Yang, Z., He, H., Zhong, X., Chen, Q., Zhang, X., Sun, S., … Deng, H. (2022). Chemical reprogramming of human somatic cells to pluripotent stem cells. *Nature* |, *605*. https://doi.org/10.1038/s41586-022-04593-5

Hajduskova, M., Baytek, G., Kolundzic, E., Gosdschan, A., Kazmierczak, M., Ofenbauer, A., Del Rosal, M. L. B., Herzog, S., Ul Fatima, N., Mertins, P., Seelk-Müthel, S., & Tursun, B. (2019). MRG-1/MRG15 is a barrier for germ cell to neuron reprogramming in Caenorhabditis elegans. *Genetics*, *211*(1), 121–139. https://doi.org/10.1534/genetics.118.301674

Hanna, J. H., Saha, K., & Jaenisch, R. (2010). Pluripotency and cellular reprogramming: Facts, hypotheses, unresolved issues. In *Cell* (Vol. 143, Issue 4, pp. 508–525). Elsevier B.V. https://doi.org/10.1016/j.cell.2010.10.008

Iwamori, N., Tominaga, K., Sato, T., Riehle, K., Iwamori, T., Ohkawa, Y., Coarfa, C., Ono, E., Matzuk, M. M., Bortvin, A., & Braun, R. E. (2016). MRG15 is required for pre-mRNA splicing and spermatogenesis. *PNAS*, 5408–5415. https://doi.org/10.1073/pnas.1611995113

Jelinic, P., Pellegrino, J., & David, G. (2011). A Novel Mammalian Complex Containing Sin3B Mitigates Histone Acetylation and RNA Polymerase II Progression within Transcribed Loci. *Molecular and Cellular Biology*, *31*(1), 54–62. https://doi.org/10.1128/mcb.00840-10

Joshi, A. A., & Struhl, K. (2005). Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to pol II elongation. *Molecular Cell*, *20*(6), 971–978. https://doi.org/10.1016/j.molcel.2005.11.021

Karmodiya, K., Krebs, A. R., Oulad-Abdelghani, M., Kimura, H., & Tora, L. (2012). *H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells*. https://doi.org/10.1186/1471-2164-13-424

Keogh, M. C., Kurdistani, S. K., Morris, S. A., Ahn, S. H., Podolny, V., Collins, S. R., Schuldiner, M., Chin, K., Punna, T., Thompson, N. J., Boone, C., Emili, A., Weissman, J. S., Hughes, T. R., Strahl, B. D., Grunstein, M., Greenblatt, J. F., Buratowski, S., & Krogan, N. J. (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell*, *123*(4), 593–605. https://doi.org/10.1016/j.cell.2005.10.025

Kumar, G. S., Chang, W., Xie, T., Patel, A., Zhang, Y., Wang, G. G., David, G., & Radhakrishnan, I. (2012). Sequence requirements for combinatorial recognition of histone H3 by the MRG15 and Pf1 subunits of the Rpd3S/Sin3S corepressor

complex. *Journal of Molecular Biology*, *422*(4), 519–531. https://doi.org/10.1016/j.jmb.2012.06.013

Lampe, J. W., Alford, J. S., Ann Boriak-Sjodin, P., Brach, D., Cosmopoulos, K., Duncan, K. W., Eckley, S. T., Foley, M. A., Harvey, D. M., Motwani, V., Munchhof, M. J., Raimondi, A., Riera, T. V, Tang, C., Thomenius, M. J., Totman, J., & Farrow, N. A. (2021). Discovery of a First-in-Class Inhibitor of the Histone Methyltransferase SETD2 Suitable for Preclinical Studies. *Cite This: ACS Med. Chem. Lett. 2021*, *12*, 1539–1545. https://doi.org/10.1021/acsmedchemlett.1c00272

Lensch, M. W., & Mummery, C. L. (2013). From stealing fire to cellular reprogramming: A scientific history leading to the 2012 Nobel Prize. In *Stem Cell Reports* (Vol. 1, Issue 1, pp. 5–17). https://doi.org/10.1016/j.stemcr.2013.05.001

Li, J., Kluiver, J., Osinga, J., Westers, H., Van Werkhoven, M. B., Seelen, M. A., Sijmons, R. H., Van Den Berg, A., & Kok, K. (2016). Functional Studies on Primary Tubular Epithelial Cells Indicate a Tumor Suppressor Role of SETD2 in Clear Cell Renal Cell Carcinoma 1. *NEO*, *18*, 339–346. https://doi.org/10.1016/j.neo.2016.04.005

Litovchick, L., Sadasivam, S., Florens, L., Zhu, X., Swanson, S. K., Velmurugan, S., Chen, R., Washburn, M. P., Liu, X. S., & DeCaprio, J. A. (2007). Evolutionarily Conserved Multisubunit RBL2/p130 and E2F4 Protein Complex Represses Human Cell Cycle-Dependent Genes in Quiescence. *Molecular Cell*, *26*(4), 539–551. https://doi.org/10.1016/j.molcel.2007.04.015

Lu, Y., Loh, Y.-H., Li, H., Cesana, M., Ficarro, S. B., Parikh, J. R., Salomonis, N., Delon Toh, C.-X., Andreadis, S. T., John Luckey, C., Collins, J. J., Daley, G. Q., & Marto, J. A. (n.d.). *Cell Stem Cell, Volume 15 Alternative Splicing of MBD2 Supports Self-Renewal in Human Pluripotent Stem Cells*.

Luco, R. F., Pan, Q., Tominaga, K., Blencowe, B. J., Pereira-Smith, O. M., & Misteli, T. (2010). Regulation of alternative splicing by histone modifications. *Science*, *327*(5968), 996–1000. https://doi.org/10.1126/science.1184208

Mccarthy, A., Sarkar, K., Martin, E. T., Upadhyay, M., Jang, S., Williams, N. D., Forni, P. E., Buszczak, M., & Rangan, P. (2022). *Msl3 promotes germline stem cell differentiation in female Drosophila*. https://doi.org/10.1242/dev.199625

Ohta, S., Nishida, E., Yamanaka, S., & Yamamoto, T. (2013). Global Splicing Pattern Reversion during Somatic Cell Reprogramming. *Cell Reports*, *5*(2), 357–366. https://doi.org/10.1016/j.celrep.2013.09.016

Onder, T. T., Kara, N., Cherry, A., Sinha, A. U., Zhu, N., Bernt, K. M., Cahan, P., Ogan Mancarci, B., Unternaehrer, J., Gupta, P. B., Lander, E. S., Armstrong, S. A., & Daley, G. Q. (2012). *Chromatin-modifying enzymes as modulators of reprogramming*. https://doi.org/10.1038/nature10953

Özçimen, Büşra. *The Role of Histone H3 Lysine 36 Methylation in Reprogramming of Fibroblasts and on Induced Pluripotent Stem Cell Generation*. Koç Üniversitesi, Graduate School of Science and Engineering, Doctor of Philosophy Thesis, 2018.

Park, J. H., Sun, X.-J., & Roeder, R. G. (2010). The SANT Domain of p400 ATPase Represses Acetyltransferase Activity and Coactivator Function of TIP60 in Basal p21 Gene Expression . *Molecular and Cellular Biology*, *30*(11), 2750–2761. https://doi.org/10.1128/mcb.00804-09

Rivers, A., Vaitkus, K., Ruiz, M. A., Ibanez, V., Jagadeeswaran, R., Kouznetsova, T., DeSimone, J., & Lavelle, D. (2015). RN-1, a potent and selective lysine-specific demethylase 1 inhibitor, increases γ-globin expression, F reticulocytes, and F cells in a sickle cell disease mouse model. *Experimental Hematology*, *43*(7), 546-553.e3. https://doi.org/10.1016/j.exphem.2015.04.005

Ruan, C., Lee, C. H., Cui, H., Li, S., & Li, B. (2015). Nucleosome contact triggers conformational changes of Rpd3S driving high-affinity H3K36me nucleosome engagement. *Cell Reports*, *10*(2), 204–215. https://doi.org/10.1016/j.celrep.2014.12.027

Sadasivam, S., & Decaprio, J. A. (2013). The DREAM complex: master coordinator of cell cycle-dependent gene expression. In *NATURE REVIEWS | CANCER* (Vol. 13). www.nature.com/reviews/cancer

Sevinç, K., Sevinç, G. G., Cavga, A. D., Philpott, M., Kelekçi, S., Can, H., Cribbs, A. P., Yıldız, A. B., Yılmaz, A., Ayar, E. S., Arabacı, D. H., Dunford, J. E., Ata, D., Sigua, L. H., Qi, J., Oppermann, U., & Onder, T. T. (2022). BRD9-containing non-

canonical BAF complex maintains somatic cell transcriptome and acts as a barrier to human reprogramming. *Stem Cell Reports*, *17*(12), 2629–2642. https://doi.org/10.1016/j.stemcr.2022.10.005

Srivastava, D., & DeWitt, N. (2016). In Vivo Cellular Reprogramming: The Next Generation. In *Cell* (Vol. 166, Issue 6, pp. 1386–1396). Cell Press. https://doi.org/10.1016/j.cell.2016.08.055

Takahashi, K., & Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, *126*(4), 663–676. https://doi.org/10.1016/j.cell.2006.07.024

Tojo, M., Hamashima, Y., Hanyu, A., Kajimoto, T., Saitoh, M., Miyazono, K., Node, M., & Imamura, T. (2005). The ALK-5 inhibitor A-83-01 inhibits Smad signaling and epithelial-to-mesenchymal transition by transforming growth factor-β β β β. *Japanese Cancer Association*, *November*. https://doi.org/10.1111/j.1349-7006.2005.00103.x

Tominaga, K., Tominaga, E., Ausserlechner, M. J., & Pereira-Smith, O. M. (2010). The cell senescence inducing gene product MORF4 is regulated by degradation via the ubiquitin/proteasome pathway. *Experimental Cell Research*, *316*(1), 92–102. https://doi.org/10.1016/j.yexcr.2009.09.015

Vierbuchen, T., & Wernig, M. (2012). Molecular Roadblocks for Cellular Reprogramming. In *Molecular Cell* (Vol. 47, Issue 6, pp. 827–838). https://doi.org/10.1016/j.molcel.2012.09.008

Wan, M. S. M., Muhammad, R., Koliopoulos, M. G., Roumeliotis, T. I., Choudhary, J. S., & Alfieri, C. (2023). Mechanism of assembly, activation and lysine selection by the SIN3B histone deacetylase complex. *Nature Communications*, *14*(1). https://doi.org/10.1038/s41467-023-38276-0

Xi, C., Sun, J., Xu, X., Wu, Y., Kou, X., Zhao, Y., Shen, J., Dong, Y., Chen, K., Su, Z., Liu, D., Ye, W., Liu, Y., Zhang, R., Xu, Y., Wang, H., Hao, L., Wu, L., & Gao, S. (2022). Mettl14-driven senescence-associated secretory phenotype facilitates somatic cell reprogramming. *Stem Cell Reports*, *17*(8), 1799–1809. https://doi.org/10.1016/j.stemcr.2022.06.012

Xie, T., Zmyslowski, A. M., Zhang, Y., & Radhakrishnan, I. (2015). Structural Basis for Multi-specificity of MRG Domains. *Structure*, *23*(6), 1049–1057. https://doi.org/10.1016/j.str.2015.03.020

Yu, M., Qian, K., Wang, G., Xiao, Y., Zhu, Y., & Ju, L. (2023). Histone methyltransferase SETD2: An epigenetic driver in clear cell renal cell carcinoma. In *Frontiers in Oncology* (Vol. 13). Frontiers Media S.A. https://doi.org/10.3389/fonc.2023.1114461

Zakrzewski, W., Dobrzyński, M., Szymonowicz, M., & Rybak, Z. (2019). Stem cells: Past, present, and future. In *Stem Cell Research and Therapy* (Vol. 10, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s13287-019-1165-5

Zhang, P., Du, J., Sun, B., Dong, X., Xu, G., Zhou, J., Huang, Q., Liu, Q., Hao, Q., & Ding, J. (2006). Structure of human MRG15 chromo domain and its binding to Lys36-methylated histone H3. *Nucleic Acids Research*, *34*(22), 6621–6628. https://doi.org/10.1093/nar/gkl989

Zhang, P., Zhao, J., Wang, B., Du, J., Lu, Y., Chen, J., & Ding, J. (2006). The MRG domain of human MRG15 uses a shallow hydrophobic pocket to interact with the N-terminal region of PAM14. *Protein Science*, *15*(10), 2423–2434. https://doi.org/10.1110/ps.062397806

Zhu, M., & Zernicka-Goetz, M. (2020). Principles of Self-Organization of the Mammalian Embryo. In *Cell* (Vol. 183, Issue 6, pp. 1467–1478). Cell Press. https://doi.org/10.1016/j.cell.2020.11.003