Investigating Genetic and Epigenetic Regulators of Human Naive Pluripotent Stem Cells

by

Nilya KARASÜRMELİ

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

Master of Science

in

Molecular Biology and Genetics

September 20, 2023

Investigating Genetic and Epigenetic Regulators of Human Naive Pluripotent Stem Cells

Koç University

Graduate School of Sciences and Engineering

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

Nilya KARASÜRMELİ

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)

Prof. Nesrin ÖZÖREN

Prof. Tuğba BAĞCI ÖNDER

Date: 20.09.2023

ABSTRACT

Investigating Genetic and Epigenetic Regulators of Human Naive Pluripotent Stem Cells

Nilya KARASÜRMELİ

Master of Science in Molecular Biology and Genetics

20.09.2023

Pluripotent stem cells can exist in two primary cellular states referred to as naive or primed. Naive pluripotent stem cells possess enhanced developmental potency compared to conventional primed states. However, generating and maintaining human naive pluripotent stem cells remains challenging. Developing tools to isolate naive cells and elucidate associated genetic and epigenetic regulators is therefore essential. We aimed to create a reporter for isolating naive cells from heterogeneous cultures, optimize naive induction protocols, and uncover regulators of naive status acquisition. Based on expression data, we selected KLF17 as a gene specifically expressed in naive pluripotent stem cells. Therefore, we engineered a KLF17-T2A-EGFP fluorescent reporter in human induced pluripotent stem cells (iPSCs) using CRISPR/Cas9-based genome editing. We tested this GFP-based reporter system during naive conversion steps. Quantitative PCR analysis of EGFP expression showed naive-specific activation, indicating successful construction of the reporter.

Optimizing the conversion of naive stem cells is an ongoing challenge. We tested the effects of overexpressing KLF17 and NANOG/LIN28A on naive induction. qPCR results demonstrated these transcription factors increased activation of naive markers compared to controls, suggesting they may accelerate resetting. We also aimed to assess the effects of KLF17 and NANOG overexpression on naive and primed reprogramming efficiency from somatic cells. The results showed that overexpression of KLF17 did not affect reprogramming efficiency. However, NANOG overexpression increased the efficiency of both primed and naive reprogramming.

In addition, we investigated the impacts of knocking out transcription factors KLF5, KLF17, and LIN28A on naive and primed reprogramming. KLF5 knockout did not affect either reprogramming system. But KLF17 knockout specifically decreased naive reprogramming efficiency while not affecting primed reprogramming. In contrast, LIN28A knockout reduced the efficiency of both primed and naive reprogramming. Lastly, we tested small molecule inhibitors of epigenetic regulators on primed and naive reprogramming. The inhibitors iCBP112, iBRD9, RN1, A83-01, and EPZ5676 all significantly increased reprogramming efficiency under both primed and naive conditions. Taken together, these results demonstrated that both transcription factors and epigenetic regulators affect the acquisition of primed vs. naive identity from somatic cells.

ÖZETÇE

İnsan Naif Pluripotent Kök Hücrelerinin Genetik ve Epigenetik Düzenleyicilerinin Araştırılması

Nilya KARASÜRMELİ

Moleküler Biyoloji ve Genetik, Yüksek Lisans

20 Eylül 2023

Pluripotent kök hücreler, naif veya primed (meyilli) olarak adlandırılan iki temel hücresel durumda bulunabilirler. Naif pluripotent kök hücreler, geleneksel hazır durumlarla karşılaştırıldığında gelişimsel yetenek açısından üstün bir konuma sahiptirler. Ancak insan naif pluripotent kök hücrelerinin üretilmesi ve korunması zorlu bir süreçtir. Bu nedenle, naif hücreleri izole etmek ve ilişkilendirilen genetik ve epigenetik düzenleyicileri açığa çıkarmak için araçlar geliştirmek önemlidir. Amacımız, heterojen kültürlerden naif hücreleri izole etmek için bir haberci oluşturmak, naif indüksiyon protokollerini optimize etmek ve naif durum kazanımının düzenleyicilerini ortaya çıkarmaktı. İfade verilerine dayanarak, naif pluripotent kök hücrelerde özellikle ifade edilen bir gen olan KLF17'yi seçtik. Bu nedenle, insan uyarılmış pluripotent kök hücrelerinde (iPSCs) KLF17-T2A-EGFP floresan bir haberci oluşturmak için CRISPR/Cas9 tabanlı genom düzenleme kullanarak bir KLF17-T2A-EGFP floresan habercisini oluşturduk. Bu GFP tabanlı haberci sistemini naif dönüşüm adımları sırasında test ettik. EGFP ifadesinin kantitatif PCR analizi, naif özgü aktivasyonu gösterdi ve bu da raporun başarılı bir şekilde oluşturulduğunu gösterdi.

Naif kök hücrelerin dönüşümünü optimize etmek devam eden bir zorluktur. KLF17 ve NANOG/LIN28A'nın naif indüksiyon üzerindeki etkilerini test ettik. Kantitatif PCR sonuçları, bu transkripsiyon faktörlerinin kontrollere göre naif belirteçlerin aktivasyonunu artırdığını gösterdi, bu da onların sıfırlamayı hızlandırabileceğini düşündürdü. Ayrıca, KLF17 ve NANOG aşırı ifadesinin somatik hücrelerden hem primed hem de naif yeniden programlamada etkilerini değerlendirmeyi amaçladık. Sonuçlar, KLF17 aşırı ifadesinin yeniden programlama verimini etkilemediğini gösterdi. Ancak NANOG aşırı ifadesi hem primed hem de naif yeniden programlamanın verimini artırdı.

Ayrıca, naif ve primed yeniden programlama üzerinde KLF5, KLF17 ve LIN28A transkripsiyon faktörlerinin nakavt etkilerini araştırdık. KLF5 nakavtı her iki programlamayı da etkilemedi. Ancak KLF17 nakavtı özellikle naif yeniden programlamanın verimini azalttı, primed yeniden programlamayı etkilemedi. Bunun aksine, LIN28A nakavtı hem primed hem de naif yeniden programlamanın verimini azalttı. Son olarak, epigenetik düzenleyicilerin küçük molekül inhibitörlerini primed ve naif yeniden programlama üzerinde test ettik. İCBP112, İBRD9, RN1, A83-01 ve EPZ5676 inhibitörleri hem primed hem de naif koşullar altında yeniden programlama verimini önemli ölçüde artırdı. Tüm bu sonuçlar, somatik hücrelerden primed ve naif kimlik kazanımını etkileyen hem transkripsiyon faktörlerinin hem de epigenetik düzenleyicilerin rolünü göstermektedir.

ACKNOWLEDGEMENTS

I sincerely thank my supervisor, Prof. Tamer Önder, for the opportunity to perform my thesis research under his guidance. He taught me lab techniques and scientific thinking from the basics onward. His mentoring throughout my master's enabled me to successfully complete this thesis. I could not have done it without his guidance and support.

I thank my thesis committee members, Prof. Nesrin Özören and Prof. Tuğba Bağcı Önder, for their participation and constructive feedback, which strengthened my work.

I would like to thank my colleagues at the Önder Lab, whom I've been working with over the past two years. It has been a valuable experience to collaborate with them during my time in the lab. I am happy because of knowing Sarra. You have a kind heart and a thoughtful approach. I enjoyed our daily lunch breaks, where we used to carry out all sorts of conversations that I will always remember. Your friendship has meant a lot to me, and you will always hold a special place in my life. Elifsu, getting to know you has been so valuable. You have always been willing to help me. Our conversations have meant a lot to me, and I hope we will stay in touch in the future. I am glad that I got to know you. Arda, as I've learned so much from you. Thank you for all of your assistance and instruction during my time in the lab. Mükrime, you were always so helpful and friendly to everyone. It was a pleasure working with a cheerful soul like you. Ömer, your friendly nature always makes the work environment much more enjoyable. It was nice to work with you. Morteza, the conversations we had were meaningful to me. Thank you for taking the time to share your perspectives and thoughts with me.

I had the pleasure of working with Erdem, Elif Naz, and Batuhan. They were dedicated and intelligent people. I am grateful for all the help they provided me. Elif Naz, your kind

personality made collaborating enjoyable. I appreciate the opportunity we had to work together. I hope wonderful things are ahead for you.

I was fortunate to share a room with such wonderful people like Maryam, Nilüfer, and Beste. It was a pleasure getting to know you all and sharing with you a place I called home. Maryam, your friendship made a significant impact on my life. I am grateful for the constant support you provided. Thank you for everything you've done for me. I value the time we spent together and the memories we made.

My family has always supported me through everything. I'm truly thankful for their unconditional love, which has always empowered me to be the person I am now. I feel so fortunate to have them in my life. Their care and guidance are invaluable to me and knowing that I can rely on them is a blessing. I feel lucky to be surrounded by their love. My life is better with them in it. I'm so grateful they are my family. I love you all.

TABLE OF CONTENTS

LIST OF TABLES x

LIST OF FIGURES xi

ABBREVIATIONS xv

Chapter 1: INTRODUCTION 1

1.1 Key Features and Functional Properties of Stem Cells 1

1.2 The Potency and Differentiation Capacity of Stem Cells 2

1.3 Pluripotent Stem Cells and Their Properties 3

1.3.1 The History of Pluripotent Stem Cells 4

1.3.2 Primed Pluripotent Stem Cells 8

1.3.3 Naive Pluripotent Stem Cells 12

1.4 Krüppel Like Factor 17 22

1.5 Somatic Cell Reprogramming 24

1.6 Epigenetic and Chromatin Factors 26

Chapter 2: METHODS 29

2.1 Cloning of Single Guide RNA Oligonucleotides into the LentiCRISPR v2 Vector 29

2.2 Virus Production and Concentration 32

2.3 Culturing dH1f Cells 33

2.4 Culturing iPSCs 33

2.5 Culturing Naive PSCs 34

2.6 Viral Infection of Fibroblasts and iPSCs 34

2.7 Reprogramming dH1f Cells to Primed iPSCs with E7 Medium 35

2.8 Reprogramming dH1f Cells to Naive iPSCs with PGXL and HENSM Media 36

2.9 Conversion of primed iPSCs to Naive iPSCs 37

2.10 Transfection Techniques for iPSCs 39

2.11 cDNA Synthesis and Real-Time Quantitative PCR (RT-qPCR) 40
2.12 T7 Endonuclease I Assay 42
2.13 Q5 Site-Directed Mutagenesis 42
2.14 The Polymerase Chain Reaction (PCR) 43
2.15 Immunostaining of Tissue Culture Plates 46
Chapter 3: RESULTS 48
3.1 KLF17 Is Specifically Expressed in Naive but Not Primed Pluripotent Cells 48
3.2 Preparation of The KLF17 Knock-in Targeting Vector Expressing EGFP and Cloning of CRISPR Guide RNAs 49
3.3 PCR Validation of KLF17 Knock-in Clones Generated by Genome Editing 56
3.4 Introducing PAM Mutations into The KLF17 Targeting Cassette 65
3.5 PCR Screening of PAM-Mutated KLF17 Clones 67
3.6 PCR Validation of PGK-PuroR Excision by Cre Recombination in Clone 2E1 70
3.7 Assessing GFP Reporter Expression in the 2E1 c10 Knock-in Cell Line 74
3.8 KLF17, NANOG, and LIN28A Overexpression Effect on Naive State Induction 78
3.9 Investigating NANOG and KLF17 Overexpression Effect on Reprogramming Efficiency 81
3.10 Assessing the Effect of KLF5, LIN28A, and KLF17 Knockouts on Naive Reprogramming 84
3.11 Enhancing Naive Reprogramming Efficiency with Inhibitors of Epigenetic Regulators 95
Chapter 4: DISCUSSION 103
4.1 Preventing Excess Cas9 Activity Following Desired On-Target Genome Editing 103

4.2 Strategies for GFP Signal Detection in CRISPR-Engineered Cells.....104
4.3 Enhancing Naive Conversion through NANOG and KLF17 Overexpression.....106
4.4 Different Effects of NANOG and KLF17 Overexpression on Reprogramming.....107
4.5 KLF5 and KLF17 Knockout Effect on Reprograming 108
4.6 LIN28A Knockout Reduces Both Naive and Primed Cellular Reprogramming.....109
4.7 Improving Naive Reprogramming Efficiency with Small Epigenetic Inhibitors 110
BIBLIOGRAPHY 112
APPENDIX A: LIST OF PLASMIDS 135

LIST OF TABLES

Table 2. 1 Single guide RNA sequences. 31

Table 2. 2 Composition of defined stem cell culture media E8, HENSM, and PGXL. 37

Table 2.3 Forward and reverse primers list used for RT-qPCR in 5' to 3' direction. 41

Table 2. 4 Forward and reverse primer sequences listed in the 5' to 3' direction. 44

Table 2. 5 Primer combinations with annealing temperatures. 45

Table 2. 6 List of primary and secondary antibodies. 47

Table 3. 1 PCR screening of clones. 57

Table 3. 2 Expected PCR products for OUT and IN PCR. 71

Table 3. 3 List of inhibitors used during reprogramming with their functions. 96

Appendix A - Table 1. 1 List of plasmids constructed and utilized in experimental procedures. 135

LIST OF FIGURES

Figure 1. 1 Waddington's epigenetic landscape.. 3

Figure 1. 2 Comparison of naive and primed pluripotent stem cell states.....21

Figure 1. 3 Distinguishing functional characteristics of naive versus primed pluripotent stem cells.....21

Figure 1. 4 Domain structures of the Krüppel-like factor (KLF) family of transcription factors.. 24

Figure 1. 5 Reprogramming and conversion of naive and primed pluripotent states.....26

Figure 3. 1 Transcriptional datasets confirming naive-specific expression of KLF17. 49

Figure 3. 2 Overview of experimental strategy for generating KLF17 reporter by CRISPR/Cas9n. 50

Figure 3. 3 Schematic representation of human KLF17 (Krüppel Like Factor 17) gene structure, and the positioning of T7 primers and guide RNA target site for knock-in strategy.....51

Figure 3. 4 The Sanger sequencing results for cloning the sgRNAs into LentiCRISPR v2 (Plasmid #52961), respectively hKLF17 Guide 1, hKLF17 Guide 2, and hKLF17 Guide 3. 52

Figure 3. 5 T7 assay results for PCR products targeting the KLF17 gene using guide 1, guide 2, and guide 3. 53

Figure 3. 6 The Sanger sequencing results for cloning the sgRNAs into pSpCas9n(BB)-2A-Puro (PX462) V2.0 (Plasmid #62987), respectively hKLF17 Guide 1, and hKLF17 Guide 3.....53

Figure 3. 7 Schematic representation of T-2A-EGFP-PGK-Puro (Plasmid #83344) showing the restriction sites for homology arm cloning. 54

Figure 3. 8 Targeted knock-in of KLF17 using CRISPR/Cas9 nickase-mediated homology-directed repair.....55

Figure 3. 9 Schematic showing the primer design for random integration PCR. 57

Figure 3. 10 Schematic of the non-integrated KLF17 locus (top) and the targeted integrated locus (bottom). 59

Figure 3. 11 Schematic of the non-integrated KLF17 locus (top) and the targeted locus with 5′ integration (bottom). 60

Figure 3. 12 Schematic of the non-integrated KLF17 locus (top) and the targeted locus with 3′ integration (bottom). 61

Figure 3. 13 PCR screening of clones. 62

Figure 3. 14 Sanger sequencing results of clone 6 PCR products.....64

Figure 3. 15 Sanger sequencing of PAM mutation incorporation into KLF17 targeting vector. 67

Figure 3. 16 Screening of 93 clones after introducing PAM mutations by EGFP PCR. 68

Figure 3. 17 PCR validations of EGFP integrated clones.....69

Figure 3. 18 Sanger sequencing results of clone 2E1 PCR products. 70

Figure 3. 19 *KLF17* locus after Cre recombinase treatment. 71

Figure 3. 20 Screening of 24 clones after CRE recombination by OUT PCR.....72

Figure 3. 21 IN PCR primer locations (arrows) inside the 5' and 3' homology arms. ...73

Figure 3. 22 PCR Validation of CRE recombination. 73

Figure 3. 23 Sanger sequencing results of 2E1 C10 PCR product.....74

Figure 3. 24 Overview of the stage-wise protocol for resetting primed iPSCs to naive state.....75

Figure 3. 25 Phase contrast microscopy of iPSCs during transition from primed to naive pluripotency. 76

Figure 3. 26 Quantitative PCR data displayed as relative mRNA expression levels for KLF17, SUSD2 and DNMT3L. 77

Figure 3. 27 Schematic showing the EGFP and Knock-in qPCR primer design. 78

Figure 3. 28 qPCR detection of EGFP and knock-in alleles. 78

Figure 3. 29 qPCR analysis of naive marker expression during naive conversion after KLF17 overexpression. 79

Figure 3. 30 qPCR analysis of NANOG and LIN28A expression during naive conversion after NANOG/LIN28A overexpression. 80

Figure 3. 31 qPCR analysis of naive marker expression during naive conversion after NANOG/LIN28A overexpression. 81

Figure 3. 32 Relative mRNA expression levels in dH1f cells after lentiviral infection. 82

Figure 3. 33 Experimental workflow for reprogramming dH1f cells with OSKM and NANOG/KLF17 overexpression. 83

Figure 3. 34 Number of Tra-1-60 positive colonies of reprogrammed dH1f cells with NANOG and KLF17 overexpression. 84

Figure 3. 35 Experimental workflow for generating KLF5, KLF17, and LIN28A knockout dH1f cells and reprogramming knockout dH1f cells with OSKM. 86

Figure 3. 36 T7 assay results for Knockout dH1f cells. 86

Figure 3. 37 Number of Tra-1-60 positive colonies of KLF5 knockout dH1f cells after reprogramming. 89

Figure 3. 38 Number of Tra-1-60 positive colonies of KLF17 knockout dH1f cells after reprogramming. 91

Figure 3. 39 Quantification of TRA-1-60 and KLF17 positive colonies after reprogramming of KLF17 knockout dH1f cells under naive conditions.....92

Figure 3. 40 Representative immunofluorescence images. 92

Figure 3. 41 Number of Tra-1-60 positive colonies of LIN28A knockout dH1f cells after reprogramming.. 95

Figure 3. 42 Experimental workflow for reprogramming dH1f cells with OSKM and epigenetic inhibitors. 96

Figure 3. 43 Number of Tra-1-60 positive colonies dH1f cells after reprogramming with inhibitors. 99

Figure 3. 44 Tra-1-60 staining of dH1f cells after reprogramming with combination of inhibitors. 102

ABBREVIATIONS

iPSC	Induced Pluripotent Stem Cell
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
RT-qPCR	Real Time Quantitative Polymerase Chain Reaction
ICM	Inner Cell Mass
ESC	Embryonic Stem Cells
mESC	Mouse Embryonic Stem Cell
hESC	Human Embryonic Stem Cell
iMEF	Irradiated Mouse Embryonic Fibroblast
TF	Transcription Factor
OSKM	OCT4, SOX2, KLF4 and c-MYC
HDAC	Histone Deacetylase
P300/CBP	EP300/CREB-Binding Protein
RNA	Ribonucleic acid
DNA	Deoxyribonucleic acid
PCR	Polymerase chain reaction
sgRNA	Single Guide RNA
DMSO	Dimethyl Sulfoxide
DMEM	Dulbecco's Modified Eagle Medium
PBS	Phosphate-Buffered Saline
FBS	Fetal Bovine Serum
EDTA	Ethylenediaminetetraacetic acid
PEG	Polyethylene Glycol
Ct	Threshold Cycle
cDNA	Complementary DNA

HRP	Horseradish Peroxidase
RPKM	Reads Per Kilobase per Million mapped reads
NT1	Non-Targeting 1
LIF	Leukemia inhibitory factor
FGF2	Fibroblast growth factors 2
TGFβ	Transforming growth factor-β
BMP	Bone Morphogenetic Protein
PFA	Paraformaldehyde
HRP	Horseradish peroxidase
HENSM	Human Enhanced Naive Stem cell Medium
CRM1/2	Chemical Resetting Medium 1/2
EGFP	Enhanced Green Fluorescent Protein
SCID	Severe combined immunodeficiency
BSA	Bovine Serum Albumin
PAM	Protospacer Adjacent Motif
PTM	Posttranslational Modifications

Chapter 1:INTRODUCTION

1.1 Key Features and Functional Properties of Stem Cells

Stem cells are undifferentiated cells that have the potential to divide and differentiate into a variety of specialized cell types. Stem cells are unique cells with crucial characteristics, including self-renewal, potency, differentiation potential, regenerative capacities, and plasticity. The functional capabilities of stem cells highlight their integral roles in development, tissue formation, and maintaining homeostasis.

Self-renewal refers to the capability of stem cells to divide to generate new stem cells, and the self-renewal process is tightly regulated, which enables persistence over extended durations. Both intrinsic transcriptional networks and extrinsic signals from the stem cell niche and intrinsic transcriptional networks have roles in regulating self-renewal characteristics. Wnt, Notch, TGF-beta signals, and the neighbor cells promote self-renewal via signaling activating pathways (Reya et al., 2003; Duncan et al., 2005; Watabe & Miyazono, 2008). Additionally, SOX2, OCT4, and NANOG are key transcription factors that maintain the undifferentiated state and induce proliferation (Niwa et al., 2000; Masui et al., 2007; Loh et al., 2006).

The capacity of stem cells to differentiate into various cell types is called potency. Potency is regulated by epigenetic modifications and transcription factors. Thus, parts of the genome become silent through epigenetic changes as differentiation progresses, limiting cells to specific destinies, and epigenetic modifications also activate linage-specific genes (Ladewig et al., 2013). The type of stem cell determines the level of potency (Jaenisch & Young, 2008).

1.2 The Potency and Differentiation Capacity of Stem Cells

According to their developmental stage and capacity for differentiation, stem cells are divided into numerous types, which were also explained in the epigenetic Landscape metaphor, which Waddington proposed, and it shows cell fate determination during development and differentiation. Totipotent cells exist at the top of the epigenetic landscape, which means that, at this stage, the totipotent cells can differentiate into all cell types, including embryonic and extraembryonic tissues that are present in the early phases of embryonic development. Totipotent cells have unrestricted cell fate and the greatest differentiation potential (Waddington, 1957; Surani et al., 1990).

The restriction from totipotency to pluripotency is represented by a ridge that the ball eventually comes across as it keeps rolling downward. Pluripotent cells have the potential to differentiate into cells of all three primary germ layers: ectoderm, mesoderm, and endoderm, and cell types of the adult organism, but extraembryonic tissues cannot be produced by pluripotent cells (Figure 1.1) (Murry & Keller, 2008).

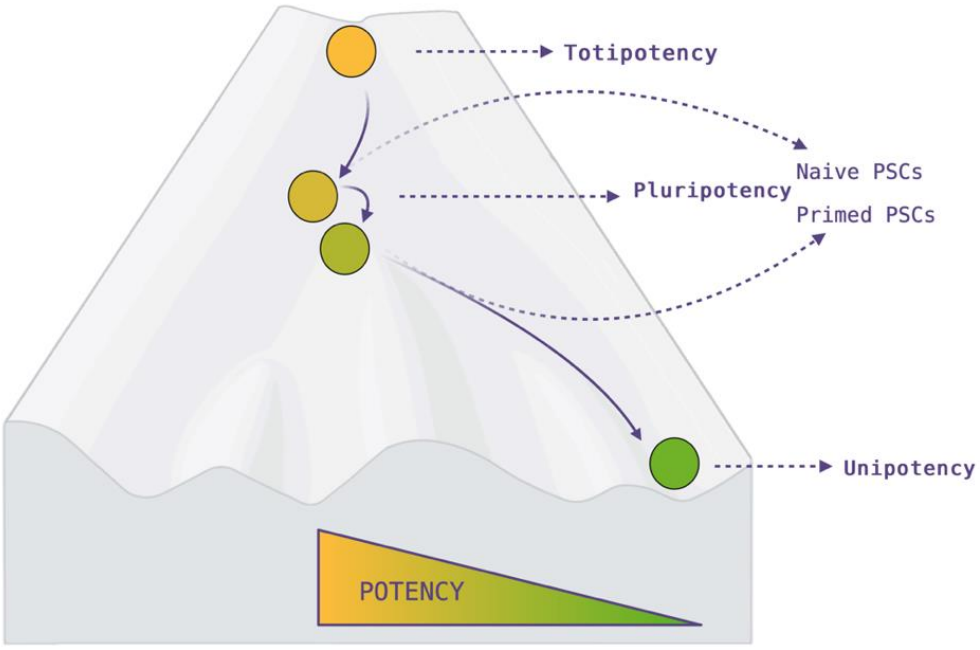


Figure 1. 1 Waddington's epigenetic landscape. This metaphorical model represents cell fate determination and differentiation as marbles rolling down ridges and valleys of an epigenetic landscape. The starting point at the top represents a totipotent stem cell. As the marble rolls down, options become restricted into distinct valleys representing commitment to specific cell fates and loss of developmental potency.

1.3 Pluripotent Stem Cells and Their Properties

The inner cell mass of the blastocyst during embryonic development is pluripotent. From the inner cell mass, pluripotent embryonic stem cells can be derived (Evans & Kaufman, 1981; Thomson et al., 1998). ESCs are considered pluripotent, which means that they are capable of differentiating into all embryonic cell types, including extraembryonic tissues. ESCs show unlimited self-renewal capacity in vitro without losing developmental potency (Murry & Keller, 2008; Nichols & Smith, 2009). SSEA4, TRA-1-60, and TRA-1-81 are the most known cell surface markers that can be used to identify undifferentiated human ESCs (Adewumi et al., 2007). Additionally, OCT4, SOX2, and NANOG are core transcription factors that help to maintain the pluripotency state of ESCs (Boyer et al., 2005).

Induced pluripotent stem cells (iPSCs) are adult cells that have been reprogrammed to return to a pluripotent state, enabling them to develop into any cell type. Induction of pluripotency in somatic cells can be achieved by using defined transcription factors OCT3/4, SOX2, KLF4, and MYC. Thus, it is shown that somatic cell fates could be reversed to a pluripotent state by introducing defined transcription factors (Takahashi & Yamanaka, 2006). Human iPSCs demonstrate self-renewal, similar to ESCs, and can generate all three primary germ layers: ectoderm, endoderm, and mesoderm (Takahashi et al., 2007).

1.3.1 The History of Pluripotent Stem Cells

The First Derivation of Mouse Embryonic Stem Cell

The first embryonic stem cell (ESC) lines were derived from mouse embryos in 1981. Cells of the Inner cell mass (ICM) of mouse blastocyst were plated on mitotically inactivated mouse embryonic fibroblast (MEF) feeder layers. These ICM-derived ESCs formed undifferentiated colonies and were passaged to establish a stable pluripotent cell line. The pluripotency feature of these cells was detected by alkaline phosphatase activity, which is used as a marker for undifferentiated pluripotent stem cells. Additionally, the cultured ESCs had large nuclei compared to their cytoplasm and distinguished nucleoli, which were considered as features for undifferentiated conditions. Karyotype analysis was done for these cultured ESCs. In contrast to previous attempts, the chromosomal composition of these cells was normal after extended culture. When ECSs were injected into mice, teratomas with various differentiated cell types from three germ layers formed, indicating their developmental pluripotency. ES cells, when injected into mouse blastocysts, contributed to normal embryonic development (Evans et al., 1981).

The First Derivation of Human Embryonic Stem Cell

The first human ESCs were derived in 1998. The inner cell mass of the human blastocyst produced from in vitro fertilization was isolated and plated onto MEF feeder layers. After continuous culture of human ESCs, five cell lines were established, which were named H1, H7, H9, H13, and H14. These cells had rounded, multicellular colonies with distinct borders and a high nucleus-to-cytoplasm ratio, which is considered the hallmark morphology of undifferentiated ESCs. The expression of *OCT4*, *TRA-1-60*, *TRA-1-81*, *SSEA-4*, and alkaline phosphatase activity was used to confirm the pluripotency of human ESCs. *OCT4*, *TDGF1*, *NANOG*, and *SOX2* expression levels were high, which was consistent with the expectation. Moreover, isolated, and cultured human

ESCs were injected into severe combined immunodeficient (SCID) mice to show teratoma formation containing tissues from the three germ layers, which are ectoderm, mesoderm, and endoderm. Furthermore, despite continuous passaging, because of high levels of telomerase activity, the cells had normal karyotypes (Thomson et al., 1998).

The Importance of LIF and Downstream Signaling Pathways for ESC Cultures

It was found that mouse ESCs cultured in vitro could be maintained in an undifferentiated state by using some particular components in serum. The addition of leukemia inhibitory factor (LIF) into mouse ESC cultures prevented differentiation of these cells while maintaining pluripotency markers and alkaline phosphatase activity. Likewise, the depletion of LIF from the medium gave rise to differentiating ESCs. It was also demonstrated that using LIF made it possible for mouse ESCs to be cultured without a feeder layer while inhibiting differentiation (Williams et al., 1988; Smith et al., 1988). It was established that leukemia inhibitory factor (LIF), working through the gp130 receptor subunit and related JAK kinases, is crucial for mouse ESC self-renewal. LIF activates the transcription factor STAT3, and LIF-independent self-renewal could be achieved by constitutively active STAT3, which was confirmed based on the morphology of cells, staining of alkaline phosphatase, and expression of pluripotency markers OCT3/4 and REX1 (Matsuda et al., 1999; Niwa et al., 1998).

Functional Role of NANOG in ESC Fate Control

Functional expression screening revealed NANOG to be a unique intrinsic factor, and independent of LIF/STAT3, NANOG maintains ESC self-renewal. In the absence of LIF, pluripotency and self-renewal of mouse ESCs could be supported by *Nanog* overexpression, which was confirmed by the expression of markers and morphology of colonies. When injected into mouse blastocysts, Nanog-overexpressing clones maintained functional pluripotency as indicated by teratoma formation and chimera

contribution. Under standard ESC culture conditions, Loss of *Nanog* expression caused spontaneous differentiation and a decline in self-renewal. It was concluded that Nanog is a crucial element of the regulatory circuitry maintaining the undifferentiated ESC state and pluripotency (Chambers et al., 2003).

Nanog knockout mouse embryos were not able to form epiblast, and *Nanog* knockout caused the generation of parietal endoderm-like cells. Thus, it was found that for the establishment of pluripotent epiblast, NANOG is a necessary element. Under standard culture conditions, *Nanog* knockout ES cells spontaneously differentiated into extraembryonic endoderm lineage, which was detected by analyzing the expression of pluripotency markers *Sox2*, *Oct4*, and *FoxD3*. This was also confirmed by rescue experiments done, and self-renewal and pluripotency marker expression was restored by re-expression of *Nanog* in knockout ES cells. It was elucidated that NANOG is an indispensable and non-redundant element for maintaining pluripotency in vivo in the developing epiblast and in vitro in ES cells (Mitsui et al., 2003).

Derivation of Novel Epiblast Stem Cells

Stem cell lines can also be derived from post-implantation stage mouse epiblast tissue, and these are called post-implantation epiblast-derived stem cells (EpiSCs). EpiSCs require FGF2 and ACTIVIN A signaling for self-renewal and pluripotency marker expression, and unlike mouse ESCs, they do not depend on LIF/STAT3 signaling pathways. In contrast to mouse ESCs, several transcriptional factors, such as *Klf2*, *Klf4*, and *Tbx3*, were repressed in EpiSCs. In addition, EpiSCs expressed *Otx2*, *Brachyury*, and *Sox17*, which are silent in mouse ESCs. Even though human ESCs and mouse EpiSCs had similar hypomethylated DNA methylation patterns, this pattern was different from mouse ESCs. Unlike the compact colonies, EpiSCs had flattened epithelial morphology. Thus, it was stated that based on differences in signaling, gene expression, and

morphology, EpiSC lines reflect a different pluripotent state in comparison to mouse ESCs, and there were some similarities between EpiSCs and human ESCs (Tesar et al., 2007; Brons et al., 2007).

The Concept of Two Distinct Pluripotent States "Naive" and "Primed"

In 2009, the naive and primed states of pluripotency exhibited by embryonic stem cells (ESCs) were distinguished. Mouse ESCs were called "naive" pluripotent stem cells, representing the ground state in developmental stages. In contrast, EpiSCs were called "primed" pluripotent stem cells, corresponding to the primed state. Naive ESCs were derived from pre-implantation embryos, such as mouse ESCs, whereas Primed ESCs were derived from post-implantation epiblast, such as mouse epiblast stem cells and human ESCs. It was stated that the naive state depends on LIF/STAT3 signaling and exhibits two active X chromosomes in female cells, and naive stem cells efficiently contribute to chimeras. However, the primed state relies on FGF/ACTIVIN signaling, undergoes X-inactivation, and developmental potency is limited. In addition, global DNA hypomethylation happens in the naive state, but the primed state has methylated genomes (Nichols & Smith, 2009).

Mouse ESCs can be cultured in standard serum with the addition of LIF or under serum-free conditions containing LIF, MEK inhibitor PD0325901, and GSK3 inhibitor CHIR99021 (2i/LIF). ESCs cultured in 2i/LIF showed upregulation of inner cell mass-specific genes and hypomethylation compared serum+LIF, and *Oct4*, *Nanog* highly is expressed in 2i/LIF. In contrast, ESCs in serum+LIF condition downregulated these self-renewal markers following prolonged culture. It was concluded that MEK/ERK signaling pathway imbibition led to transcriptional and epigenetic changes towards the ground state similar to pre-implantation epiblast, whereas ESCs cultured in serum+LIF resembled a developmentally primed pluripotent state. (Ying et al., 2008; Nichols et al., 2009).

1.3.2 Primed Pluripotent Stem Cells

Primed pluripotent cells can differentiate into all somatic cell types in vitro but have limited potential in vivo. Primed PSCs correspond to the developmental state of late epiblast cells in post-implantation embryos in vivo, and they cannot contribute to chimeras when injected into pre-implantation embryos (Nichols & Smith, 2009). This indicates that primed cells have limitations in recapitulating early embryonic development (Weinberger et al., 2016; Rossant et al., 2017). Primed cells can be obtained from embryos, or standard reprogramming of human somatic cells also generates induced pluripotent stem cells (iPSCs) that are in the primed state (Brons et al., 2007; Wernig et al., 2007; Yu et al., 2007). Primed PSCs display distinct morphological, metabolic, molecular, and functional characteristics that prime them for lineage specification (Weinberger et al., 2016). Understanding pluripotency regulation in early human development requires an in-depth comprehension of the transcriptional circuitry controlling the primed state.

The autoregulatory loop made up of the transcription factors OCT4, SOX2, and NANOG sustains the gene expression profile underlying the primed pluripotent state. OCT4, SOX2, and NANOG co-occupy and regulate several downstream target genes, and OCT4 and SOX2 bind to regulatory regions of genes such as promoters and enhancers, which are essential for self-renewal and pluripotency in human ESCs and mouse ESCs. In particular, OCT4 is a crucial transcription factor necessary for ESC identity by enabling SOX2/NANOG binding, and OCT4, SOX2, and NANOG circuitry is conserved between mouse and human ESCs (Kim et al., 2008; Boyer et al., 2005). Additional stage-specific transcription factors control lineage priming and differentiation in primed PSCs by acting downstream of the OCT4-SOX2-NANOG circuitry (Buecker et al., 2014). These include the transcription factors OTX2, GATA6, SOX17, FOXA2,

and PAX6, which prepare cells for mesendoderm, endoderm, mesoderm, and ectoderm fates, respectively. (Acampora et al., 2013; Seguin et al., 2008; Fujikura et al., 2002; Teo et al., 2011; Zhang et al., 2010). Lineage markers in primed human PSC cells differ in intensity due to the heterogeneous expression of these lineage-priming transcription factors.

Signaling Pathways Regulating Primed Pluripotency

For self-renewal and pluripotency maintenance, primed PSCs depend on Activin/Nodal signaling through the TGF pathway. Nodal and Activin A activate SMAD2/3 factors that control gene expression to maintain primed pluripotency. (Vallier et al., 2005; Xu et al., 2008). Pluripotency genes like NANOG are activated by SMAD2/3 (James et al., 2005; Vallier et al., 2009). The genomic binding and transcriptional activity of SMAD2/3 in response to TGFβ signaling is regulated by lineage-determining transcription factors in a cell-type-dependent manner, so SMAD2/3 also represses differentiation genes (Mullen et al., 2011). Therefore, the TGFβ pathway supports primed pluripotency through SMAD2/3 transcriptional regulation.

Fibroblast growth factor (FGF) ligands activate FGF receptors (FGFRs), resulting in downstream signaling via the PI3K/AKT and MAPK/ERK pathways (Kunath et al., 2007; Greber et al., 2010). PI3K/AKT signaling pathway works with Activin/Nodal to sustain pluripotency, but ERK signaling gives rise to differentiation, so inhibition of ERK pathway stabilizes primed human pluripotent stem cells in an undifferentiated state (Vallier et al., 2005; Greber et al., 2010). FGF signals must be tightly regulated to balance self-renewal and differentiation in primed pluripotent stem cells.

Bone morphogenetic protein (BMP) signaling works in concert with FGF and Activin/Nodal signaling to sustain pluripotency and self-renewal in primed human pluripotent stem cells (Vallier et al., 2005). BMP ligands cause BMP receptors to become

active, which phosphorylates SMAD proteins. In order to balance self-renewal vs. differentiation in primed human pluripotent stem cells, this BMP-SMAD signaling works in combination with other pathways to control gene expression. BMP signaling must be precisely modulated for primed human pluripotent stem cells to be properly maintained (Qi et al., 2004; Xu et al., 2008).

Epigenetic Features Primed Pluripotency

The promoters and enhancers of developmental regulators are particularly hypermethylated globally in primed PSCs compared to the naive state, and the lineage potential is constrained by increased methylation, which is linked to the silencing of early developmental genes. Nevertheless, essential pluripotency genes, such as *OCT4*, *SOX2*, and *NANOG,* remain in primed cells with low levels of methylation, which may allow for the ability to self-renew. Primed PSCs exhibit hypomethylation compared to differentiated cells, so differentiated cells have strong methylation patterns at promoters of key developmental regulators that are unmethylated and active in primed PSCs (Lister et al., 2009). In primed cells, de novo DNA methyltransferases, DNMT3A and DNMT3B, are strongly expressed to maintain DNA methylation patterns (Nazor et al., 2012).

Histone modifications are essential for controlling the gene expression patterns that underlie the ability of primed pluripotent stem cells to self-renew and differentiate. In the primed state compared to the naive state, the polycomb repressive complex 2 (PRC2) deposits the repressive mark H3K27me3 more widely across the genome. Increased H3K27me3 methylation by PRC2 prevents the activation of developmental genes while preparing them for induction during differentiation (Weinberger et al., 2016). Additionally, the "bivalent" chromatin regions of numerous important developmental regulatory genes in PSCs contain both activating (H3K4me3) and repressive (H3K27me3) histone modifications. When PSCs are not differentiated, bivalency keeps

developmental genes inactive but prepares them for activation when differentiation occurs. Bivalent domains could serve as a strategy to preserve pluripotency while keeping developmental genes ready for lineage-specific activation (Bernstein et al., 2006). Therefore, in comparison to earlier developmental stages, primed cells have more compacted, lineage-restricted chromatin globally (Hawkins et al., 2010).

Metabolic Characteristics of Primed PSCs

In contrast to naive pluripotent stem cells (PSCs), primed PSCs, such as human ESCs and epiblast stem cells (EpiSCs), mainly rely on glycolysis rather than oxidative phosphorylation for energy production (Zhou et al., 2012). The Warburg-effect metabolic state supports the biosynthetic requirements of rapid proliferation and self-renewal linked to the pluripotent state (Mathieu & Ruohola-Baker, 2017). Despite having more mitochondria than naive PSCs, primed PSCs have lower mitochondrial respiration and immature mitochondrial structure. The uncoupling protein UCP2 expression also limits oxidative phosphorylation in primed PSCs (Zhang et al., 2011). The transcription factor HIF1α has a role in controlling metabolic shift by activating glycolysis while suppressing oxidative phosphorylation (Zhou et al., 2012). Overall, the distinct metabolic profile of primed PSCs, which is characterized by aerobic glycolysis, anabolism, and decreased mitochondrial respiration, enables the rapid buildup of biomass and high rates of energy-demanding proliferation and self-renewal.

Markers to Distinguish of Primed PSCs

Primed PSCs express unique surface proteins, which distinguish them from naive PSCs. These include the glycolipids TRA-1-60, TRA-1-81, and the keratan sulfate antigens SSEA3 and SSEA4. Primed PSCs also express markers like CD24, CD57, and CD90 (Collier & Rugg-Gunn, 2018; Chan et al., 2013; Characterization of human embryonic stem cell lines by the International Stem Cell Initiative, 2007). At the

transcriptional level, primed PSCs uniquely express the homeobox protein OTX2 (Tesar et al., 2007). *SOX2*, *OCT4*, and *NANOG* are expressed in both primed and naive PSCs (Boyer et al., 2005). Altogether, primed PSCs can be distinguished from naive cells by the specific surface protein and transcription factor profile.

1.3.3 Naive Pluripotent Stem Cells

In contrast to primed cells, naive pluripotent cells reflect an earlier developmental state resembling pre-implantation epiblast (Nichols & Smith, 2009). Naive PSCs closely resemble the inner cell mass in their molecular profile, transcriptional state, and developmental potential. As such, they can extensively contribute to chimeric embryos when introduced into blastocysts (Theunissen et al., 2016; Bredenkamp et al., 2019). Single-cell RNA sequencing analysis of existing naive human iPSC lines and human embryos has identified markers that distinguish naive cells from conventional primed cells (Takashima et al., 2014; Blakeley et al., 2015; Messmer et al., 2019; Goodwin et al., 2020). Although primed PSCs can differentiate into any somatic lineage, they only resemble post-implantation epiblast cells. They cannot be used to mimic earlier stages of embryonic development. (Brons et al., 2007). Understanding the molecular differences between primed and naive states is important for optimizing stem cell culture conditions and applications. Specifically, capturing the naive state may enable stem cells to enhance functional capabilities. Generating stable naive state human PSCs would enable to model major events like gastrulation, comprehend details of human embryogenesis, and generate chimeric models for human tissues and organs.

Transcriptional Regulation of Naive Pluripotency

Although core regulators of pluripotency like OCT4, SOX2, and NANOG are expressed in both naive and primed cells, other regulators are unique to the naive state.

KLF2, KLF4, KLF5, TBX3, TFCP2L1, ESRRB, and TET1/2 are the additional regulators that are specifically necessary for naive but not primed pluripotency. The most crucial naive-specific factors are KLF2, KLF4, TFCP2L1, and ESRRB. TFCP2L1 and NANOG co-occupy promoters of naive genes, including *Klf2/4* (Dunn et al., 2014). KLF2, KLF4, and KLF5 maintain self-renewal and are highly expressed in mouse ESCs. Since KLF2, KLF4, and KLF5 are integral parts of the core pluripotency circuitry in mouse ESCs, they interact with other pluripotency transcription factors, including OCT4, SOX2, and NANOG (Jiang et al., 2008; Hall et al., 2009). TFCP2L1 is another important pluripotency factor mediating LIF-STAT3 signaling in mouse ESCs. TFCP2L1 and NANOG are also located in regulatory regions of naive state such as *Klf2/4* (Martello et al., 2013). ESRRB is the pluripotency transcription factor, and ESRRB is a direct target of NANOG in mouse ESCs, and NANOG binding to regulatory regions of *Esrrb* locus regulates *Esrrb* expression (Festuccia et al., 2012). In addition, the upregulation of pre-implantation-associated genes like *KLF17*, *DPPA3*, *DPPA5*, and *TFCP2L1* is another hallmark of naive stem cells (Guo et al., 2017).

Signaling Pathways Regulating Naive Pluripotency

LIF/JAK/STAT3 Pathway is one of the main routes to regulate naive pluripotency. JAK kinases are activated by LIF, which allows them to phosphorylate and activate the STAT3 transcription factor (Niwa et al., 2009). *Klf4*, *Tfcp2l1*, and *Sox2* are naive pluripotency genes that are upregulated by phosphorylated STAT3 (Martello & Smith, 2014). Therefore, the naive gene regulatory network is maintained by LIF-mediated STAT3 activation. Additionally, LIF prevents differentiation by inhibiting MAPK/ERK signaling via a negative feedback loop (Kim et al., 2008). Overall, in order to maintain the gene regulatory network that supports the naive pluripotent state, the

LIF/JAK/STAT3 pathway is induced, which simultaneously upregulates pluripotency factors and blocks differentiation signals.

Regulation of the FGF/MEK/ERK signaling pathway is crucial for maintaining a naive state. Key naive Transcription factors such as KLF2 and KLF4 are phosphorylated and inhibited when FGF initiates the MEK/ERK kinase cascade. Thus, in order to maintain a naive state, the FGF/ERK signaling pathway must be suppressed as it promotes transition toward primed pluripotency (Hamilton et al., 2019; Kunath et al., 2007). In order to induce primed pluripotency in human ESCs, FGF/ERK works together with Activin/NODAL signaling (Vallier et al., 2017). Critical pluripotency factors, including NANOG, are suppressed by ERK-mediated KLF2 and KLF4 inhibition. ERK activation also suppresses *SOX2* expression. Due to these pathways, FGF/ERK inhibition is necessary for the naive induction or maintenance of both mouse and human ESCs (Hamilton et al., 2019; Theunissen et al., 2014).

The Wnt/GSK3/β-Catenin signaling pathway is another important mechanism that regulates the naive state. For naive ESCs, regulating GSK3 and Wnt signaling balances self-renewal and differentiation. GSK3 is a kinase that targets β-catenin for degradation, so Wnt signaling represses GSK3 function, enabling β-catenin to accumulate and activate gene transcription (Ten Berge et al., 2011). GSK3 inhibition stabilizes TCP2L1, which helps maintain the naive state. Wnt/GSK3 regulation is not necessarily required for mESC self-renewal if the LIF/STAT3 signaling is strong enough. In human ESCs, GSK3 inhibition helps induction but not maintenance of the naive state (Theunissen et al., 2014; Sokol, 2011).

Another important mechanism that controls the naive state is the TGFβ signaling pathway. For the maintenance of naive pluripotent stem cells, the TGF pathway must be blocked. Through the TGFβ receptor complex, TGF ligands induce Smad2/3 transcription

factors to become active, and activated Smad2/3 promotes lineage-specifying transcription factors to induce differentiation (Oshimori & Fuchs, 2012). In contrast, TGFβ pathway inhibition by targeting ALK4/5/7 induces the transition from the primed to the naive state (Guo et al., 2017). In addition, TGFβ signaling activation suppresses *Nanog* expression, which is a core transcription factor that sustains the pluripotency state (Xu et al., 2008). Therefore, to maintain naive status, blocking TGFβ signaling is critical for both mouse and human ESCs (Gafni et al., 2013). However, TGFβ signaling inhibition alone is insufficient to obtain a naive state in human ESCs, demonstrating that TGFβ works with other pathways like FGF/ERK (Theunissen et al., 2014).

The Hippo signaling pathway has an essential role in promoting naive pluripotency. YAP and TAZ are the members of the Hippo pathway and regulate the activity and localization of transcriptional-enhanced associate domain (TEAD) transcription factors. When YAP/TAZ are located in the nucleus, they inhibit TEAD factors to induce lineage-specific gene expression, which maintains the naive state in mouse ESCs. When YAP/TAZ are translocated to the nucleus, they bind TEAD factors and activate the transcription of lineage-specific genes (Lian et al., 2010). In addition, YAP induces human naive pluripotency by occupying and activating naive enhancers. Nuclear YAP interacts with naive transcription factors like TFAP2C, NANOG, and KLF17 to upregulate naive genes (Qin et al., 2016). Overall, regulation of YAP/TAZ expression and localization helps to balance between naive state and differentiation.

Epigenetic Features of Naive Pluripotency

Comparable to primed ESCs and differentiated somatic cells, naive embryonic stem cells (ESCs) have a globally hypomethylated DNA state. Although the whole genome is hypomethylated, enhancers, gene bodies, and promoters show the most significant hypomethylation (Leitch et al., 2013; Takashima et al., 2014). Similarly, in

naive ESCs, pluripotency genes like *NANOG* and *OCT4* promoters are generally unmethylated. In contrast, these locations are highly methylated in somatic and primed pluripotent cells (Veillard et al., 2014). This hypomethylated DNA allows an open chromatin state and indiscriminate gene expression program that prepares the naive cells for differentiation potential. In summary, naive ESCs can be distinguished from primed pluripotent and differentiated cell types by global DNA hypomethylation at regulatory sites. This crucial epigenetic support allows the naive pluripotent state's plasticity and potency.

Repressive histone marks, associated with heterochromatin formation and gene silencing like H3K9me3 and H3K27me3, are absent in naive ESCs. In primed ESCs, H3K27me3 silences the developmental regulators, while in naive ESCs, H3K27me3 is depleted (Azuara et al., 2006; Bernstein et al., 2006). Naive ESCs, on the other hand, contain active histone modifications that support an open chromatin state and transcriptional activation, so the promoters and enhancers of developmental genes are enriched with H3K4me3, H3K36me3, H3K27ac, and H3K9ac in the naive state (Marks et al., 2012). When compared to primed PSCs, naive cells have more active H3K4me3 deposited throughout the genome at enhancers and promoters. The increased H3K4me3 is associated with the activation of naive transcription factors like KLF2/4/5 and TFCP2L1 to enable the naive state (Hanna et al., 2010). Additionally, genome-wide H3K27ac levels are high in naive PSCs that activate enhancers to obtain an open chromatin state (Rada-Iglesias et al., 2011; Mikkelsen et al., 2007). The distinct histone modification landscape between naive and primed pluripotent stages regulates chromatin accessibility and transcriptional competence (Atlasi & Stunnenberg, 2017; Hackett & Surani, 2014).

Naive embryonic stem cells (ESCs) have two active X chromosomes, which resemble the pre-implantation epiblast in vivo (Mak et al., 2004; Silva et al., 2009). NANOG, PRDM14, and REX1 are transcription factors that suppress Xist in the naive stage, keeping both X chromosomes active (Sahakyan et al., 2017; Navarro et al., 2008; Leitch et al., 2013). In contrast, primed ESCs have X chromosome inactivation with one inactive X (Silva et al., 2009). During the naive-to-primed transition, repression of naive transcription factors and overexpression of Xist results in Xist coating and inactivation of one X chromosome (Sahakyan et al., 2017; Gafni et al., 2013). One important epigenetic distinction between the pluripotent phases is the different X chromosome status. Naive ESCs have greater developmental potential due to two active X chromosomes, allowing for enhanced X-linked pluripotency gene expression.

Metabolic Features of Naive Pluripotency

Human naive cells require both oxidative phosphorylation and glycolysis, whereas mouse naive PSCs only rely on oxidative phosphorylation. In contrast to conventionally primed human PSCs, oxidative phosphorylation demand is higher in a naive state (Sperber et al., 2015). Human naive PSCs also have higher levels of mitochondrial DNA and biomass than mice (Takashima et al., 2014). In line with the balanced metabolic program, oxidative phosphorylation and glycolytic genes are increased at the gene expression level in human naive PSCs. Metabolomic analysis reveals species-specific variations in the supplying of the TCA cycle, with a reduced contribution of glutamine in human cells (Sperber et al., 2015). Compared to Mouse ESCs or conventional human primed PSCs, human naive PSCs maintain a more balanced mitochondrial/glycolytic condition.

Generation and Maintenance of Naive PSCs

In mice, primed PSCs can be converted by culturing in media containing LIF, GSK3β, and MEK inhibitors. This 2i media containing MEK and GSK3 inhibitors induce and maintain mouse ESCs in a naive ground state (Ying et al., 2008). However, converting human PSCs to a naive state is much more challenging, and the 2iLIF condition is also not sufficient to derive and maintain human naive PSCs.

The ideal approach to obtain human naive PSCs is to directly derive them from pre-implantation blastocyst when the epiblast is in a naive ground state. However, conventional culture conditions produce human ESCs with a primed state that have a post-implantation identity (Thomson et al., 1998; Reubinoff et al., 2000). By combining MEK and GSK3 suppression with BMP and TGFβ blocking, naive epiblast cells from dissociated human embryos can also be harvested in vitro (Guo et al., 2016; Guo et al., 2017). Further culture condition optimizations through the inhibition of PKC signaling resulted in establishing a t2iL+Gö medium that can continuously maintain isolated early embryo epiblast cells without any genetic manipulation (Guo et al., 2017).

Another approach to obtain naive PSCs is converting human primed PSCs to naive state. Overexpression of key naive transcription factors NANOG and KLF2 in primed PSCs and culture conditions that contain MEK and GSK3 inhibitors can convert primed state cells to naive state (Hanna et al., 2010; Takashima et al., 2014).

In a further refinement, conventional primed human PSCs were converted into a naive state using a 3iL medium with MEK, GSK3β, and p38 inhibitors, along with LIF, FGF2, and TGFβ modulation. Naive human PSCs obtained by this protocol showed a dome-shaped morphology and characteristics similar to mouse ground-state ESCs, including dependency on LIF/STAT3 signaling. Additionally, upregulation of naive transcription factors, global DNA hypomethylation, X activation in female lines, and

contribution to interspecies chimeras when injected into mouse morula were observed for converted naive cells. Therefore, it was stated that naive human pluripotent stem cells, which resemble human pre-implantation epiblast in terms of transcriptional, epigenetic, and functional features, can derived and maintained by modulating signaling pathways (MEK, GSK3β, p38) and using LIF, FGF2 without a need for transgenes (Gafni et al., 2013).

In addition, transgene-free resetting of human primed PSCs to a naive state can be achieved by stepwise blocking of MEK/ERK and GSK3 followed by subsequent inhibition of TGFβ signaling pathways by using inhibitors enables transition of primed PSCs to a naive state (Theunissen et al., 2014). The Hippo, WNT, and JAK-STAT pathways regulation with stepwise MEK/ERK and TGFβ inhibition can further enhance transgene-free primed-to-naive conversion (Pastor et al., 2016; Kilens et al., 2018). Adding IWR1, XAV939, or IWP2 helps stabilize the reacquisition of naive status (Qin et al., 2016).

The third approach is the direct reprogramming of somatic cells into naive PSCs using Yamanaka factors OCT4, SOX2, KLF4, and MYC (OSKM). In the literature, two techniques have been employed for this. The first one is OSKM factors transfection to fibroblast with the Sendai virus transfers to fibroblasts, and cells are cultured on a medium containing signaling pathway inhibitors, such as rNHSM, rRSeT, r5iLAF, and t2iLGöY (Kilens et al., 2018; Liu et al., 2017). The second, synthetic mRNAs are used to provide reprogramming factors for fibroblasts, and cells are then cultured on PGXL media that contains PKC and Wnt inhibitors (Bredenkamp et al., 2019; Bayerl et al., 2021). In addition, modifying OSKM induction by co-expressing NANOG, KLF2, or other naive transcription factors directly converts fibroblasts into expandable human naive iPSCs without transitioning through a primed state (Liu et al., 2017).

Building on these foundations, optimized media for naive derivation include:

5i/L/A - GSK3i, MEKi, BRAFi, ROCKi, SRCi plus LIF and ALKi (Theunissen et al., 2014)

t2iL+Gö - GSK3i, MEKi, ROCKi plus LIF and PKCi Gö6983 (Takashima et al., 2014)

NHSM - GSK3i, MEKi, ROCKi, SRCi, LIF, ALKi, SPRYi (Pastor et al., 2016)

3iL - GSK3i, MEKi, ROCKi, LIF (Guo et al., 2017)

These combine modulation of WNT, MAPK, TGFβ, FGF, and PKC pathways using small molecules and cytokines.

Markers to Distinguish of Naive PSCs

For mouse Naive ESCs, Klf4, Klf5, Tbx3, Tfcp2l1, Esrrb, Dppa3, Klf2, Prdm14, Dnmt3l, Zfp42 (Rex1), Pecam1 are stated as naive-specific genes (Marks et al., 2012; Boroviak et al., 2015). In contrast, lineage markers like Fgf5, Otx2, Pdgfra, Hoxa1, and Hoxb1 are downregulated (Boroviak et al., 2015).

For human Naive ESCs, NANOG, KLF2, KLF4, KLF5, KLF17, DPPA3, DPPA5, TBX3, ESRRB, ZFP42, TFCP2L1 are considered as naive cell markers (Chan et al., 2013; Theunissen et al., 2014). On the other hand, primed markers like OTX2, FOXA2, SOX11, SOX21, ZIC2, ZIC5, and HOXD1 are expressed at low levels (Sahakyan et al., 2017). Additionally, naive cells uniquely express surface markers such as CD7, CD75 and CD130 (Collier & Rugg-Gunn, 2018; Collier et al., 2017).

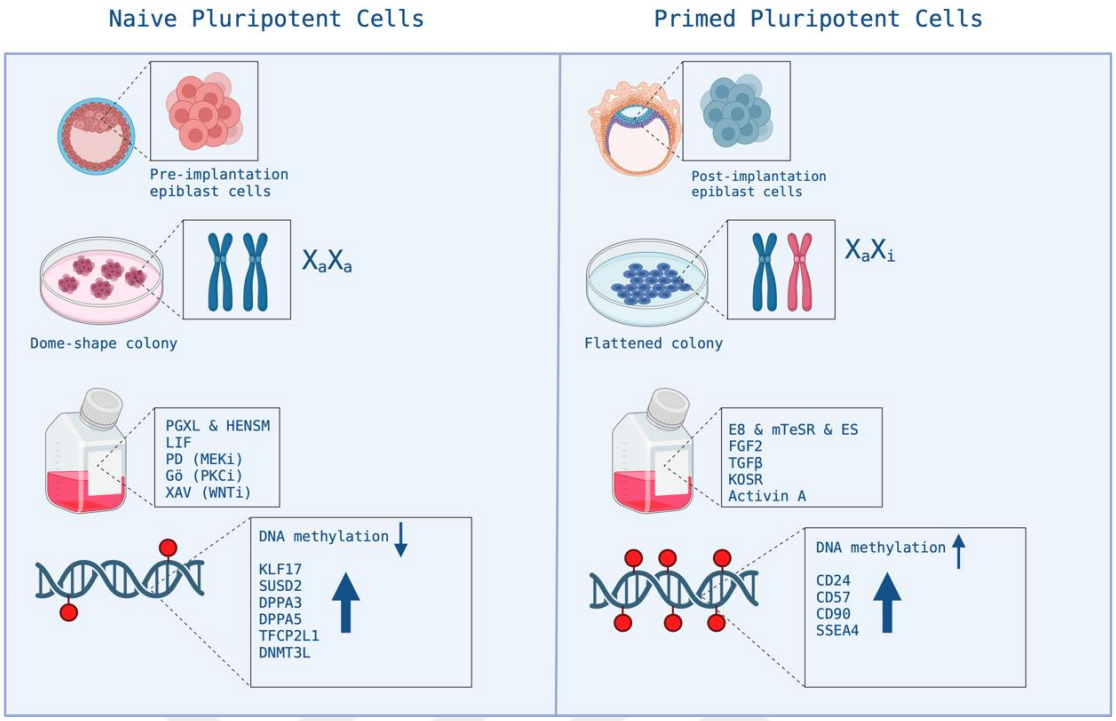


Figure 1. 2 Comparison of naive and primed pluripotent stem cell states. This diagram highlights key differences between naive and primed pluripotent stem cells across various properties including morphology, signaling, metabolism, and developmental potency.

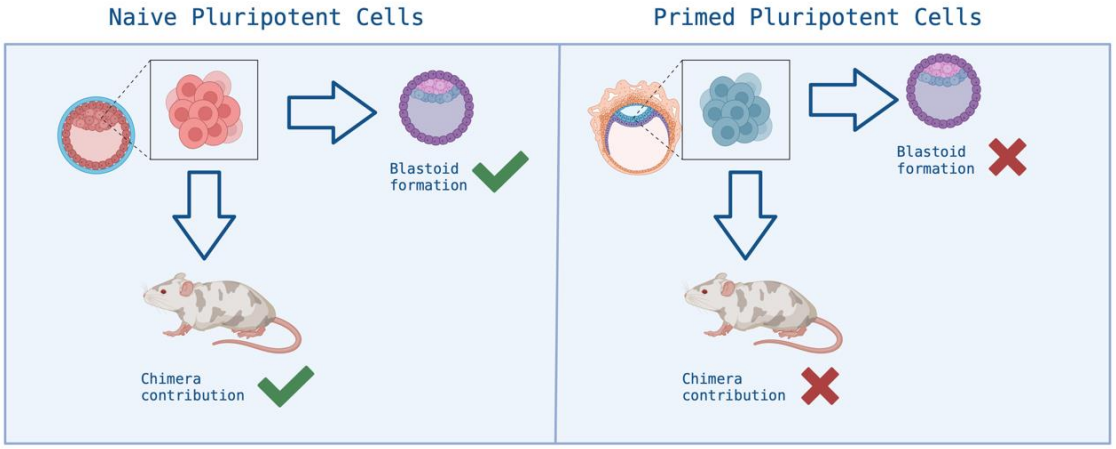


Figure 1. 3 Distinguishing functional characteristics of naive versus primed pluripotent stem cells.

The Significance of Investigating Naive Pluripotency

The early developmental events and cellular differentiation in human embryos cannot be examined in their natural setting (in vivo). While model organisms like mice have offered valuable insights, they have limitations in fully replicating human embryonic development. For example, key molecular processes in early human embryogenesis differ from those in the mouse. Therefore, research using mouse models alone is insufficient to fully understand early human development. Advancing cellular models that more closely mimic human embryo development would significantly benefit three critical areas: improving in vitro fertilization (IVF) embryo implantation success rates, elucidating the mechanisms of organ formation, and revealing the pathophysiology underlying developmental defects in the fetal stage.

A key challenge is generating cell lines with broad developmental potential to generate embryo-like structures in vitro. In summary, advanced in vitro human cellular models that reflect early embryonic development will allow research into critical aspects of embryology, implantation, organogenesis, and developmental disorders that cannot be fully addressed with current model organisms.

1.4 Krüppel Like Factor 17

The KLF family (Krüppel-like factors) comprises DNA-binding proteins that regulate gene expression by binding to GC-rich regions in promoters. KLFs contain three C-terminal C2H2 zinc fingers that facilitate DNA binding, with conserved sequences between fingers determining binding specificity. Over 20 KLF family members have been identified in mammals (Figure 1.4). They control diverse biological processes, including cell growth, proliferation, differentiation, development, and responses to external stressors (McConnell & Yang, 2010).

KLF17 is one transcription factor belonging to the Krüppel-like factor (KLF) family. KLF17 is a transcriptional regulator essential for maintaining and promoting naive pluripotency in human ESCs and has zinc finger DNA binding domains that bind to GC-rich regions in promoters to control gene expression. Of the genes specifically expressed in naive pluripotent stem cells, the transcription factor KLF17 is among the most well-studied. KLF17 binds to promoter regions of many active genes in naive cells (Bayerl et al., 2021). KLF17 is specifically enriched in the naive state of human ESCs and is expressed in the NANOG/SOX2 positive epiblast cells of the human blastocyst, suggesting a role in regulating naive pluripotency (Lea et al., 2021). Knockout of KLF17 in human ESCs led to the collapse of naive pluripotency and differentiation when cells are cultured in naive conditions (Bayerl et al., 2021). However, Lae et al. stated that the knockout of KLF17 does not block the derivation of naive hESCs, demonstrating that KLF17 is not strictly essential for establishing the naive state. In addition, Overexpression of KLF17 in human primed ESCs is sufficient to induce the upregulation of naive markers, and KLF17 overexpression allows primed cells to reset to the naive state (Lea et al., 2021). Based on these findings, KLF17 expression appears to be a defining signature of naive pluripotent stem cells. However, a reporter system to isolate naive cells based on KLF17 expression has not yet been developed. In summary, the transcription factor KLF17 may have utility as a marker to monitor the naive state and could potentially be engineered into a reporter to enrich naive stem cell derivatives.

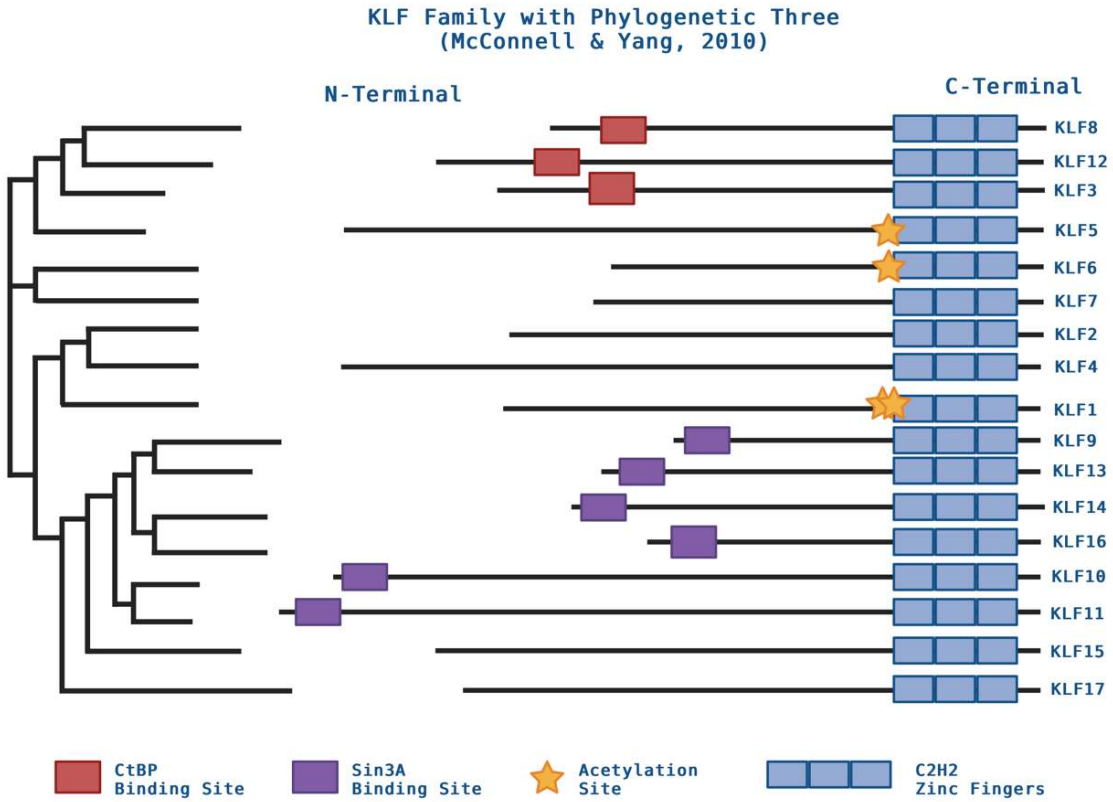


Figure 1. 4 Domain structures of the Krüppel-like factor (KLF) family of transcription factors. Schematic shows the conserved C-terminal zinc finger DNA binding domain and variable N-terminal regulatory domains among KLF family members. Adapted from McConnell and Yang, 2010.

1.5 Somatic Cell Reprogramming

Somatic cell reprogramming is the process of returning differentiated cells to a pluripotent state. In 1958, Gurdon et al. first demonstrated this concept by transferring a nucleus from an intestinal epithelial cell into an enucleated frog egg, producing tadpoles that developed into normal adults (Gurdon et al., 1958). This nuclear transfer technique showed that the nuclei of mammalian differentiated cells retain all of the genetic material needed for embryonic development. Wilmut et al. produced live lambs in mammals in 1996 by inserting the nuclei of mammary gland cells into enucleated sheep eggs (Wilmut

et al., 1997). In 1998, Wakayama et al. expanded this method to produce cloned animals from adult nuclei (Wakayama et al., 1998). However, the mechanisms underlying nuclear reprogramming remain largely unknown.

Takahashi and Yamanaka developed an alternative strategy utilizing transcription factors in 2006. Twenty-four candidate genes used in embryonic stem cell culture were screened, and introducing four factors, OCT4, SOX2, KLF4, and MYC, reprogrammed mouse fibroblasts to induced pluripotent stem cells (iPSCs) (Figure 1.4) (Takahashi & Yamanaka, 2006). In terms of shape, ability to self-renew, and pluripotency, these iPSCs were similar to embryonic stem cells. iPSCs can differentiate into all three germ layers - endoderm, mesoderm, and ectoderm. However, reprogramming efficiency with Yamanaka factors remains low at around 0.1-1% of starting cells (Takahashi et al., 2007; Takahashi & Yamanaka, 2006). In order to improve the process, the fundamental mechanisms governing reversion to pluripotency have been studied. Reprogramming has distinct phases, which are characterized by changes in metabolism, cell signaling, chromatin, and gene expression (Apostolou & Hochedlinger, 2013). To increase reprogramming efficiency, various factor combinations, small compounds and inhibitors, microRNAs, modified mRNAs, and non-integrating vectors have been explored (Malik & Rao, 2013; Hou et al., 2013).

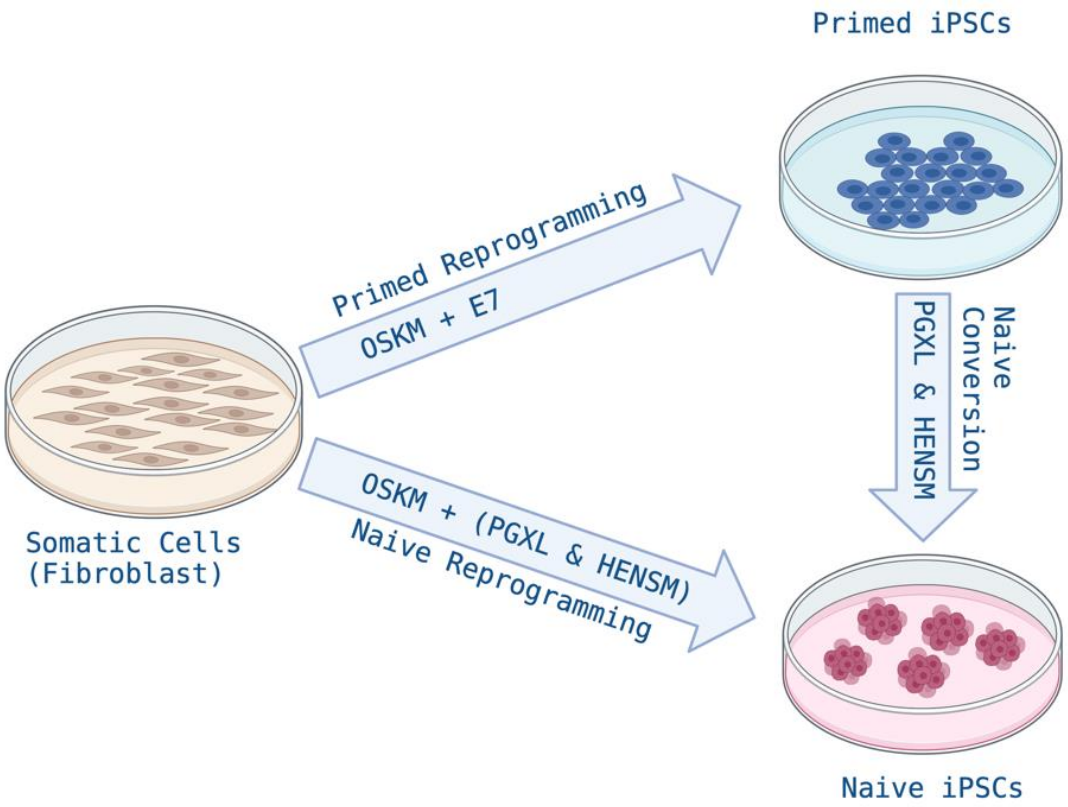


Figure 1. 5 Reprogramming and conversion of naive and primed pluripotent states. This diagram depicts reprogramming from somatic cells to naive or primed pluripotency using OSKM in specific culture media. It also shows conversion the primed cells to naive states via culture condition changes.

1.6 Epigenetic and Chromatin Factors

During reprogramming, extensive chromatin remodeling and epigenetic modifications occur to transition between differentiated and pluripotent cell states. Chromatin-modifying proteins are crucial epigenetic regulators that control chromatin architecture and gene expression. Chromatin factors can be divided into different categories based on their function as readers, writers, or erasers of histone modifications. Reader proteins recognize and bind to particular histone post-translational modifications (PTMs), bringing other regulatory elements to modify chromatin structure and

transcriptional activity. Chromatin reader proteins such as HP1 and BRD4 recognize epigenetic modifications and recruit writer modifiers to control gene expression (Sridharan et al., 2013). During reprogramming, dynamic expression of readers that recognize chromatin state changes is crucial. Writers are the enzymes that add chromatin modifications like DNA methylation, histone methylation, or acetylation (Simonsson & Gurdon, 2004; Zhong & Jin, 2009). Resetting epigenetic memory requires the controlled activity of writer proteins that alter the epigenetic landscape of the genome. In contrast to writers, erasers remove chromatin modifications. For example, eraser proteins like TET and histone deacetylases remove modifications and reverse repression (Gao et al., 2013; Ryall et al., 2015). During reprogramming, erasers must remove somatic epigenetic markers to obtain iPSCs. Overall, by changing the balance between chromatin readers, writers, and erasers, the chromatin state and gene expression can be globally altered, ultimately enabling differentiated cells to be returned to their undifferentiated state.

The objectives of this study were three-fold:

1. To develop a reporter system that enables specific labeling and isolation of naive pluripotent stem cells from heterogeneous populations. This was achieved by engineering a fluorescent reporter under the control of the KLF17 promoter using CRISPR-Cas9 genome editing.

2. To optimize and standardize robust protocols for generating naive human pluripotent stem cells. Systematic approaches were taken to identify optimal combinations of culture conditions, media, and small molecules that support the capture and maintenance of the naive state from somatic cells via reprogramming and from primed cells via conversion.

3. To elucidate the roles of genetic and epigenetic regulators during naive state induction. Small molecule inhibitors were used to determine if modulating the activity of these

factors enhances reprogramming and conversion efficiencies when generating naive pluripotent stem cells.

In summary, the central aims were to develop tools to identify naive stem cells, improve methods to produce naive stem cells, and gain mechanistic insights into epigenetic barriers that hinder naive fate acquisition.

Chapter 2:METHODS

2.1 Cloning of Single Guide RNA Oligonucleotides into the LentiCRISPR v2 Vector

To knock out the target genes, single guide RNAs (sgRNAs) were designed against each gene (Table 2.1). The sgRNA sequences were modified to add overhangs compatible with cloning into the BsmBI (NEB R0739S) sites of LentiCRISPR v2 (Addgene #52961). Specifically, a 5' CACCG overhang was added to the top strand oligos, and 5' AAAC overhangs were added to the bottom strands. A 3' C overhang was also added to the bottom oligos. These modifications allow directional cloning of the annealed oligos into the lentiviral vector. The full sgRNA oligonucleotide sequences with added overhangs were commercially synthesized by Macrogen.

The pLentiCRISPR_v2 vector digested with BsmBI restriction endonuclease was dephosphorylated by adding 2 µL of Antarctic phosphatase and incubating at 37°C for 30 minutes. This dephosphorylation step prevents the self-ligation of the vector backbone. The dephosphorylated digested vector was then run on a 1% agarose gel for visualization. Since BsmBI has two cutting sites in pLentiCRISPR_v2, digestion releases a ~2 kb fragment from the vector. The gel confirmed the removal of this ~2 kb band, indicating complete digestion by BsmBI. The remaining ~13 kb linearized pLentiCRISPR_v2 backbone was excised and purified from the gel using the NucleoSpin Gel and PCR clean-up kit (MN 740609.50). This removes the excised 2 kb fragment, isolating just the lentiviral vector backbone for sgRNA insertion. The concentration of the purified backbone was quantified by measuring absorbance at 260nm using a NanoDrop spectrophotometer.

The sgRNA oligos were prepared for ligation by phosphorylation and annealing. The oligos were first phosphorylated at their 5' ends using T4 polynucleotide kinase (T4

PNK) (NEB M0201S) to enable ligation into the vector backbone. The top and bottom oligo strands were then annealed to generate double-stranded DNA fragments containing the guide sequences. Annealing was performed by preparing an oligo mix and incubating first at 37°C for 30 minutes to facilitate annealing, then at 95°C for 5 minutes to denature any remaining single-stranded regions, and finally ramping down to 25°C slowly at 0.1°C/sec to enable proper hybridization. The annealed oligos were diluted 1:100 in water to yield suitable ligation concentrations.

For the ligation reaction, 50 ng of the gel-purified digested pLentiCRISPR_v2 backbone was combined with 1 µL of diluted annealed oligos and ligated using T4 DNA ligase (NEB #M0202) at room temperature for 4 hours. Then 5 µL of the ligation mix was transformed into 50 µL of Stbl3 competent E. coli cells. Plasmids containing the sgRNA inserts were purified from single colonies using a miniprep kit. Successful oligo cloning was verified by Sanger sequencing using the LKO.1 5' (5'-GACTATCATATGCTTACCGT-3') sequencing primer that binds within the vector backbone region.

To clone hKLF17 Guide 1 and hKLF17 Guide 3 into pSpCas9n(BB)-2A-Puro (PX462) V2.0 (Addgene #62987), BbsI (NEB R0539S) cloning was performed instead of BsmBI. The pX462 plasmid was digested with BbsI, which cuts this vector at two sites and removes a small 76 bp fragment. Since the excised fragment is so small, the digested plasmid was directly purified using a PCR clean-up kit rather than being run on a gel. The cloning process followed the same steps of backbone digestion, oligo preparation, ligation, transformation, and sequencing.

Table 2. 1 Single guide RNA sequences. The gRNAs were cloned into LentiCRISPR_v2 listed in the 5' to 3' direction.

KLF17 Guide 1 Top	**CACCG**CAGGACAGTCCTCCTGCTGC
KLF17 Guide 1 Bottom	**AAAC**GCAGCAGGAGGACTGTCCTG**C**
KLF17 Guide 2 Top	**CACCG**GTCCCGGCCGATGAGTCTTC
KLF17 Guide 2 Bottom	**AAAC**GAAGACTCATCGGCCGGGAC**C**
KLF17 Guide 3 Top	**CACCG**TCTTCTGGTGTTGCTTGAGA
KLF17 Guide 3 Bottom	**AAAC**TCTCAAGCAACACCAGAAGA**C**
KLF5 (2_1) KO gRNA Top	**CACCG**CAGTCGTAGACCAGTTCTTC
KLF5 (2_1) KO gRNA Bottom	**AAAC**GAAGAACTGGTCTACGACTG**C**
KLF5 (2_2) KO gRNA Top	**CACCG**ACTGCAGTGAAACAATTCCA
KLF5 (2_2) KO gRNA Bottom	**AAAC**TGGAATTGTTTCACTGCAGT**C**
KLF5 (2_3) KO gRNA Top	**CACCG**TCCCTGAGTTCACCAGTATA
KLF5 (2_3) KO gRNA Bottom	**AAAC**TATACTGGTGAACTCAGGGA**C**
KLF17 (3_1) KO gRNA Top	**CACCG**TGAGCTTAGACGACATATGC
KLF17 (3_1) KO gRNA Bottom	**AAAC**GCATATGTCGTCTAAGCTCA**C**
KLF17 (3_2) KO gRNA Top	**CACCG**ATGAGCTTAGACGACATATG
KLF17 (3_2) KO gRNA Bottom	**AAAC**CATATGTCGTCTAAGCTCAT**C**
KLF17 (3_3) KO gRNA Top	**CACCG**GAGGCCATATTCTTGCAACT
KLF17 (3_3) KO gRNA Bottom	**AAAC**AGTTGCAAGAATATGGCCTC**C**
LIN28A (2_1) KO gRNA Top	**CACCG**AGTGGTTCAACGTGCGCATG
LIN28A (2_1) KO gRNA Bottom	**AAAC**CATGCGCACGTTGAACCACT**C**
KLF5 (2_3) KO gRNA Top	**CACCG**TCCCTGAGTTCACCAGTATA

KLF5 (2_3) KO gRNA Bottom	**AAAC**TATACTGGTGAACTCAGGGA**C**
LIN28A (2_2) KO gRNA Top	**CACCG**CAGTGGATGTCTTTGTGCAC
LIN28A (2_2) KO gRNA Bottom	**AAAC**GTGCACAAAGACATCCACTG**C**
LIN28A (2_3) KO gRNA Top	**CACCG**CTGTCCATGACCGCCCGCGC
LIN28A (2_3) KO gRNA Bottom	**AAAC**GCGCGGGCGGTCATGGACAG**C**

2.2 Virus Production and Concentration

Lentiviral plasmids were packaged into viral particles using HEK 293T cells cultured in D10 media (DMEM (Sigma D6429) with 10% FBS (Biowest S181H) and Penicillin-Streptomycin (Biowest L0022)). HEK 293T cells were seeded at 2.5 million cells per 10cm plate a day before transfection. For transfection, plasmid mixes were prepared containing the lentiviral backbone plasmid, packaging plasmids (2250 ng psPAX2 (Addgene #12260), 250 ng pCMV-VSV-G (Addgene #8454) and 2500 ng Vector), and in addition to this plasmid combination, 20 µL of FuGENE (Promega E2312) was mixed with 185 µL of DMEM. The plasmid-FuGENE mix was incubated at room temperature for 30 minutes to allow complex formation, then added dropwise to the HEK 293T cells. After 16 hours, the media was changed to fresh D10. Virus-containing media was harvested at 48- and 72 hours post-transfection. The media was centrifuged at 1500rpm for 5 minutes and filtered through a 0.45µm filter to remove cells. The clarified filtered media was aliquoted and stored at -80°C or concentrated using polyethylene glycol (PEG) precipitation. For concentration, 50% PEG 8000 (polyethylene glycol, Sigma P5413) was added at a 1:4 ratio to the viral media, incubated at 4°C for 2 days, then centrifuged at 2500 rpm for 20 minutes. The viral pellet was resuspended in 1x PBS (Biowest L0615) at 1:100 of the initial volume and aliquoted for storage at -80°C.

2.3 Culturing dH1f Cells

dH1f cells were cryopreserved in freezing media composed of 90% FBS and 10% DMSO. To thaw, frozen vials were incubated at 37°C until melting, and the freezing media was removed by centrifuging at 1300rpm for 4 minutes. The cell pellet was resuspended in D10 medium and seeded into a 10 cm plate. The thawed cells were cultured at 37°C with 5% CO2. Media was changed to fresh D10 every 2 days, and cells were passaged at ~80% confluency. For passaging, cells were detached by incubating with trypsin (Multicell 325-542-EL) for 3 minutes at 37°C. Trypsin was inactivated by the addition of D10 medium. The cell suspension was counted using a hemocytometer and re-seeded at desired densities based on downstream applications.

2.4 Culturing iPSCs

Induced pluripotent stem cells (iPSCs) were cultured on Geltrex-coated (Gibco A1413302) plates in homemade mTeSR medium. The mTeSR medium was prepared in-house according to published recipes. iPSCs were fed daily with fresh mTeSR medium and passaged every 4-5 days before colonies reached > 70% confluency. For passaging, iPSC colonies were rinsed gently with PBS, and 0.5 mM EDTA in PBS was then added to the cells and incubated at 37°C for 4 minutes to dissociate cell-cell connections within colonies. After removing the EDTA solution, colonies were further detached by gentle pipetting. The desired number of detached colony fragments were transferred into new Geltrex-coated plates with mTeSR medium supplemented with 10μM Y-27632 ROCK inhibitor to support cell survival after passaging. The next day, the medium was changed to mTeSR without Y-27632. Cultures were incubated at standard culture conditions of 37°C, 5% CO2, and 5% O2.

2.5 Culturing Naive PSCs

Naive human PSCs were maintained in a naive state using PGXL culture conditions. Naive PSCs were cultured on irradiated mouse embryonic fibroblast (MEF) feeder layers in PGXL medium, which is composed of a 1:1 mix of Neurobasal medium (Gibco 21103049) and DMEM/F-12 (Gibco 21331020), 1% N2 (100X) and 2% B27 (50X) supplements (homemade prepared according to published recipes), 20 ng/mL human LIF, 1% Glutamax (Gibco 35050061), 1% Penicillin-Streptomycin, 100 μM β-mercaptoethanol (Gibco 21985023), and the following small molecule inhibitors: 1 μM PD0325901, 2 μM Gö6983 and 2 μM XAV939 was added (Table 2.2). To passage naive PSCs, colonies were washed with PBS and dissociated into single cells using 0.5 mM EDTA. The cells were seeded onto fresh MEF feeder layers in PGXL media containing ROCK inhibitor Y-27632 to support survival. The next day, the PGXL medium was changed to remove Y-27632. Naive cultures were passaged every 3-4 days upon reaching 70-80% confluency. They were maintained at 5% O2 and 5% CO2 at 37°C.

2.6 Viral Infection of Fibroblasts and iPSCs

For fibroblast infections, cells were seeded at 100,000 cells per well in a 6-well plate or 50,000 cells per well in a 12-well plate. Unconcentrated viral supernatant was added to the wells along with D10 medium and 1:1000 protamine sulfate (8 mg/mL, Sigma P4505) to improve transduction efficiency. After 16 hours, the fibroblast medium was changed to fresh D10. The protocol was adapted for infecting iPSCs. iPSCs were seeded at 70% confluency on the day prior to infection. The concentrated viral supernatant was used to achieve a high infection rate. The viruses were diluted in the E8 medium optimized for iPSC culture. The E8 medium used for culturing induced pluripotent stem cells contains 20 μg/ml insulin (ProSpec CYT-270), 10 μg/ml optiferrin

(InVitra 777TRF029), 13.6 ng/ml selenite, 64 µg/ml L-ascorbic acid (Sigma A8960), 1% Penicillin-Streptomycin, 1 ng/ml TGF-β1 (PeproTech 100-21), and 100 ng/ml FGF2 (homemade) in DMEM/F-12 medium. After 16 hours, the E8 medium was refreshed. Selection with 1µg/mL puromycin was then carried out for 3 days to select stably transduced cells.

2.7 Reprogramming dH1f Cells to Primed iPSCs with E7 Medium

Reprogramming experiments with dH1f cells were initiated by seeding the cells at 50,000 per well onto 12-well plates. To deliver the Yamanaka reprogramming factors OCT4, SOX2, KLF4, and c-MYC, unconcentrated viral supernatants were prepared containing the plasmids pSIN4-EF2-O2S (Addgene #21162) and pSIN4-CMV-K2M (Addgene #21164). One day after seeding, the dH1f cells were transduced with a mix of the viruses carrying the reprogramming factors. After overnight viral incubation, the transduction medium was replaced with fresh D10 medium on the next day, considered as day 1 of reprogramming. The D10 medium was again refreshed on days 3 and 5 post-transduction. On day 6 of reprogramming, 1/8 of the dH1f cells were transferred on the vitronectin (Gibco A31804) coated plate. Beginning the day after transfer, the culture medium was switched to E7 medium. The E7 medium was formulated using all the components found in the E8 medium, with the only exception of TGF-β1. The medium was changed every 2 days until day 14, after which daily medium refreshment was carried out until the end of reprogramming on day 18. The cells were maintained at 5% CO2 at 37°C. At day 18 post-reprogramming, the induced pluripotent stem cell (iPSC) cultures were fixed and immunostained for the pluripotency marker TRA-1-60 to quantify reprogrammed colony numbers. To begin, the cells in tissue culture plates were fixed with 4% paraformaldehyde (PFA) for 15-30 minutes at room temperature and then washed

with PBS. For TRA-1-60 staining, the cells were incubated overnight at 4°C with biotinylated TRA-1-60 primary antibody (Table 2.6) diluted 1:250 in a staining solution composed of 3% fetal bovine serum (FBS) and 3% Triton X-100 in PBS. After primary antibody incubation, streptavidin linked HRP secondary antibody was diluted 1:500 in the same staining solution and applied to the cells for 2 hours at room temperature to detect the biotinylated primary antibody. TRA-1-60 positive colonies were then visualized by incubating with DAB substrate, which forms a brown precipitate. ImageJ software was utilized to quantify the DAB-stained TRA-1-60 positive colonies, indicating successful iPSC reprogramming.

2.8 Reprogramming dH1f Cells to Naive iPSCs with PGXL and HENSM Media

The same reprogramming procedure was followed for the first 6 days as described previously, involving retroviral transduction of dH1f cells with OCT4, SOX2, KLF4, c-MYC, and culture in D10 medium. On day 6 of reprogramming, 1/8 of the dH1f cells were transferred to irradiated mouse embryonic fibroblast (iMEF) feeder layers. Beginning on the day after MEF transfer, the culture medium was switched from D10 to either PGXL or HENSM medium to induce naive pluripotency. The HENSM medium contained a 1:1 ratio of Neurobasal and DMEM/F12 supplemented with 1% N2, 2% B27, 1% GlutaMAX, 1% non-essential amino acids (PAN-Biotech P08-32100), 1% penicillin-streptomycin, 50 μg/mL ascorbic acid 2-phosphate, 1 mM sodium pyruvate, 100 μM β-mercaptoethanol, 20 ng/mL human LIF, and the following small molecule inhibitors: 1 μM PD0325901, 2 μM XAV939, 2 μM Gö6983, 1.2 μM CGP77675, 0.8 μM BIRB0796, and 1.2 μM Y27632. The PGXL and HENSM media were refreshed every 2 days until day 14 of reprogramming. The cells were maintained at 5% O2 and 5% CO2 at 37°C. On day 14, the reprogrammed cells were fixed and stained using the same TRA-1-60

immunostaining protocol described previously. ImageJ quantification of TRA-1-60 positive colonies was carried out to assess reprogramming under the PGXL and HENSM conditions.

2.9 Conversion of primed iPSCs to Naive iPSCs

At stage 1 (days 1-4), iPSCs were plated on the irradiated mouse embryonic fibroblast (iMEF) feeder layer in mTeSR medium with 1 μM Y27632 ROCK inhibitor. On day 1, the medium was changed to mTeSR without Y27632. On day 2, the medium was transitioned to a 1:1 ratio of mTeSR and CRM1 medium (N2B27 basal medium: 1:1 mix of Neurobasal and DMEM/F-12 with 1% N2 and 2% B27 supplements, 20 ng/mL human LIF, 1% Glutamax, 1% penicillin-streptomycin, 100 μM β-mercaptoethanol supplemented with the following small molecule inhibitors: 1 μM PD0325901 MEK inhibitor, 1mM NaBut or 1mM VPA HDAC inhibitor) to start resetting. On day 3, only CRM1 medium was used until stage 2. At stage 2 (days 5-9), On day 5, CRM1 was replaced with CRM2 medium (N2B27 basal medium with 1 μM PD0325901, 2 μM Gö6983 PKC inhibitor, and 2 μM XAV939 Wnt inhibitor). At Stage 3 (from day 9), Naive colonies appeared around days 9-10. Cells were then passaged onto PGXL medium on the iMEF feeder layer. To dissociate cells during passaging, 0.5 mM EDTA was used, and cells were maintained at 5% O2 and 5% CO2 at 37°C.

Table 2. 2 Composition of defined stem cell culture media E8, HENSM, and PGXL.

	E8 **100 mL Media**	**HENSM** **100 mL Media**	**PGXL** **100 mL Media**
DMEM/F12		46.25 mL	47.50 mL

Neurobasal		46.25 mL	47.50 mL
Pen/Strep	1 mL (1X)	1 mL (1X)	1 mL (1X)
Glutamax	-	1 mL (1X)	1 mL (1X)
NEAA	-	1 mL (1X)	-
Sodium Pyruvate	-	1 mM	-
N2 Supplement	-	1 mL (1X)	1 mL (1X)
B27 Supplement	-	2 mL (2X)	2 mL (2X)
L-ascorbic acid	64 µg/mL	50 µg/mL	-
holo-Transferrin	10 µg/mL	-	-
Insulin	20 µg/mL	-	-
FGF2	100 ng/mL	-	-
TGFβ1	1 ng/mL	-	-
hLIF	-	20 ng/ml	20 ng/mL
XAV939	-	2 µM	2 µM
Gö6983	-	2 µM	2 µM
PD0325901	-	1 µM	1 µM
CGP77675	-	1.2 µM	-
Y27632	-	1.2 µM	-
BIRB 796	-	0.8 µM	-
Selenite	13.6 ng/mL	-	-

2.10 Transfection Techniques for iPSCs

To transfect induced pluripotent stem cells (iPSCs) cultured in 6-well plates, the iPSCs were seeded in E8 medium supplemented with 10 μM Y27632 ROCK inhibitor the day before transfection. This was done to achieve approximately 70% (800,000 cells) confluency for transfection. On the day of transfection, the plasmid DNA to be transfected was diluted in 125 μL of DMEM/F-12 medium along with 10 μL of p3000 reagent (Invitrogen L3000001) in the same tube. This DNA/p3000 mixture was incubated for 5 minutes at room temperature. Separately, 3.75 μL of Lipofectamine 3000 reagent (Invitrogen L3000001) was diluted in 125 μL of DMEM/F-12 in another tube and incubated for 5 minutes at room temperature. After the 5-minute incubations, the diluted DNA/p3000 solution was combined with the diluted Lipofectamine 3000 reagent solution, allowing DNA-lipid complexes to form. This DNA-lipid complex mixture was incubated for 15 minutes at room temperature. The complexes were then added dropwise to the iPSCs. The cells were incubated with the DNA-lipid complexes for 16 hours to enable uptake. After 16 hours, the transfection medium was replaced with a fresh, normal iPSC culture medium.

For nucleofecting iPSCs, first, the culture medium was removed, and the cells were harvested using EDTA treatment for an extended time (e.g., 10 minutes) to obtain a single-cell suspension. An aliquot of the detached cells was counted to determine cell density. Approximately 1 million cells per well were centrifuged for a 6-well plate format. The supernatant was completely discarded to ensure no residual medium remained on the cell pellet. The cell pellet was then resuspended in 110 μL of Nucleofector R buffer (Invitrogen MPK10096). For each sample, 110 μL of the cell suspension was combined with 1-5 μg of plasmid DNA to be transfected. 100 μL of the cell-DNA mixture was transferred using the specialized pipette and tips provided with the 4D-Nucleofector

system (Lonza) into a nucleofection cuvette that had been pre-filled with E2 Nucleofection buffer (Invitrogen MPK10096). Care was taken during the transfer to avoid introducing any bubbles into the cuvette. The sample was nucleofected using the 4D-Nucleofector system (Lonza) with settings of 1300V, 20 ms pulse length, and 1 pulse. Immediately after nucleofection, the sample was removed from the cuvette and gently transferred dropwise into the prepared 6-well plate. The transfected iPSCs were incubated at 5% CO2 at 37°C. The medium was changed after 16 hours post-nucleofection.

2.11 cDNA Synthesis and Real-Time Quantitative PCR (RT-qPCR)

To analyze gene expression, RNA was first isolated from the cells using the NucleoSpin RNA kit (MN 740955.50). For cDNA synthesis, 1000 ng or 500 ng of the purified RNA was used as the starting material. In the first step, the RNA was assembled in PCR tubes along with 2.5 µL of 2mM dNTP mix (Thermo Scientific R0192) and 2 µL of random hexamer primers (Invitrogen N8080127). Nuclease-free water was added to bring the volume to 16.5 µL. This mixture was incubated at 65°C for 5 minutes, then immediately placed on ice. Next, the reverse transcription mix was prepared by combining 5 µL 5X First Strand Buffer (Invitrogen 18057018), 2 µL 0.1M DTT (Invitrogen 18057018), and 0.5 µL Rnasin ribonuclease inhibitor (Promega N251B). This was added to the RNA mixture and incubated at room temperature for 10 minutes. After this, 1 µL of M-MLV reverse transcriptase (Invitrogen 28025013) was added to synthesize cDNA, and the reaction proceeded for 1 hour at 37°C. The reaction was inactivated by heating at 70°C for 15 minutes. The cDNA was diluted 1:4 with water and stored at -20°C. For qPCR, a SYBR green mixture was prepared with 2X SYBR Green master mix (Qiagen 208056), 2 µL of 2.5 µM forward and reverse primer mix, 6 µL water, and 2 µL cDNA. Reactions were performed with Roche LightCycler machine. Ct data

was analyzed using the second derivative maximum method relative to the β-Actin housekeeping gene for normalization. Primers are listed in Table 2.3.

Table 2.3 Forward and reverse primers list used for RT-qPCR in 5' to 3' direction.

EGFP RT qPCR Forward	AAGCAGAAGAACGGCATCAA
EGFP RT qPCR Reverse	GGGGGTGTTCTGCTGGTAGT
KLF17 Knock-in RT qPCR Forward	CCACAGGCCAACAACAACAA
EGFP Knock-in RT qPCR Reverse	GCTGAACTTGTGGCCGTTTA
NANOG RT Forward	TGATTTGTGGGCCTGAAGAAA
NANOG RT Reverse	TGGTGGTAGGAAGAGTAAAG
KLF17 Forward	GCTGCCCAGGATAACGAGAAC
KLF17 Reverse	ATCTCTGCGCTGTGAGGAAAG
LIN28A RT Forward	TGCGGGCATCTGTAAGTGG
LIN28A RT Reverse	GGAACCCTTCCATGTGCAG
SUSD2 Forward	CCTGGCATGTCAAGTCGCT
SUSD2 Reverse	CCTGTGAGTAGGGCATTCCTG
DNMT3L Forward	CAGTGCCTGCTCCTTATGGCT
DNMT3L Reverse	TGAACAAGGAAGACCTGGACG
β-Actin Forward	TGAAGTGTGACGTGGACATC
β-Actin Reverse	GGAGGAGCAATGATCTTGAT

2.12 T7 Endonuclease I Assay

The T7 endonuclease I assay was performed to analyze gene editing efficiency following CRISPR/Cas9 targeting. First, PCR was carried out to amplify the genomic region surrounding the CRISPR/Cas9 cleavage site. The PCR product was then annealed: 400 ng of the PCR amplified DNA was assembled in a reaction with 2 µL of 10X NEBuffer 2 (NEB B7002S) and nuclease-free water up to a final volume of 19 µL. This mixture was denatured at 95°C for 5 minutes, then slowly annealed from 95°C to 85°C at a ramp rate of -2°C/second, continuing from 85°C to 25°C at -0.1°C/second ramp rate, finishing with a hold at 4°C. After annealing, T7 endonuclease I enzyme was added to cleave mismatched DNA. 1 µL of T7 Endonuclease I (NEB M0302L) was added to the 19 µL annealed PCR products and incubated at 37°C for 45 minutes. This allows T7 endonuclease I to cleave at the mismatch sites. To stop the reaction, 1.5 µL of 0.25 M EDTA was added. Prior to gel electrophoresis, 6 µL of 30% glycerol loading dye was added to the reaction products, and the fragments were resolved on a 2% agarose gel.

2.13 Q5 Site-Directed Mutagenesis

The Q5 site-directed mutagenesis kit (NEB E0554) was used to introduce targeted mutations into the plasmid vector. Mutagenesis primers were designed with the desired nucleotide changes using NEBaseChanger online tool. For the mutagenesis PCR, 12.5 µL Q5 Hot Start High-Fidelity 2X Master Mix was assembled with 2.5 µL of forward and reverse mutagenic primers mix (10 µM), 10 ng of template plasmid, and water to 25 µL final volume. PCR was carried out with an initial 30 sec 98°C denaturation, followed by 25 cycles of 98°C for 10 sec, a 30-sec annealing step at 50-72°C depending on primer Tm, and 3 min extension at 72°C, ending with a final 5 min extension at 72°C and 4°C hold. After amplifying the mutant plasmid by PCR, 1 µL of the PCR product was treated

with the KLD enzyme mix to digest the parental template plasmid. The KLD mix was prepared by combining 5 µL of 2X KLD Reaction Buffer, 1 µL of 10X KLD Enzyme Mix (containing kinase, ligase, and DpnI enzymes), and 3 µL of nuclease-free water. This 1 µL PCR sample was incubated with the 9 µL KLD mix at room temperature for 1 hour. The kinase activity phosphorylates the linear mutagenized PCR product, the ligase circulates it into a closed circular plasmid, and DpnI digests the methylated parental template plasmid. After KLD treatment, 1 µL of the reaction was transformed into Stbl3-competent E. coli cells for plasmid propagation. Single colonies were picked from transformation plates to inoculate cultures for mini plasmid preps. The isolated mutant plasmid was then sequenced by Sanger sequencing.

2.14 The Polymerase Chain Reaction (PCR)

The polymerase chain reaction (PCR) amplified a specific DNA sequence in vitro using thermal cycling. Genomic DNA was first extracted and purified from a sample (NM 740952.50). Forward and reverse primers were designed for the target DNA region that needed to be amplified (Table 2.4). These primers initiated replication by annealing to the complementary DNA template strands. In the first PCR step, the DNA sample was heated to 96°C to denature the double-stranded DNA into single strands. The temperature was lowered to allow primer annealing to the single-stranded templates (55-72°C depending on primer Tm). Next, a heat-stable DNA polymerase (NEB E0553S) extended the primers by adding nucleotides complementary to the DNA templates, making a copy of the target region. These steps of denaturation, annealing, and extension were repeated for 30-40 cycles, exponentially amplifying the target sequence. A final extension at 72°C was performed. The cycles were completed by a final 4-10°C hold step. Different primer sets had optimized annealing and extension times (Table 2.5).

Table 2. 4 Forward and reverse primer sequences listed in the 5' to 3' direction.

KLF17 T7 Forward	CCCCGACTTGCCATACAGGA
KLF17 T7 Reverse	TACACCCACCAAAGAGAAGAACTG
HAL Forward with Sac1	TTAAT**GAGCT**CCTGGGCCACTGGTGGGTAATTG
HAL Reverse with Nhe1	ATTTAT**GCTAGC**AGGACCAGCAGCAGGAGGAC
HAR Forward with Asc1	TATTA**GGCGCGCC**GTCAGTCTCCTCTCCTCATTCTG
HAR Reverse with Fse1	TTATAT**GGCCGGCC**GTTAAGGGCATCCGAGACAGG
Random Integration Forward	TACCGCCTTTGAGTGAGCTG
Random Integration Reverse	GCTCAGGTAGTGGTTGTCGG
KLF17 Out PCR Forward	GCAACTACGAGAACTGCGGA
KLF17 Out PCR Reverse	TTTTAAGCTGGGGAGCGACAA
KLF17 (Puro) Forward	CATCGGCAAGGTGTGGGTC
KLF17 (EGFP) Reverse	TGGCCGTTTACGTCGCC
KLF17 5' In PCR Forward	AGAGTTAGACCCCTTCCTTCCT
KLF17 3' In PCR Reverse	ATCTAATCCTACACCCACCAAAGAG
Q5 Guide1 Forward	CCTGCTGCTGCTCCTGCTAGC
Q5 Guide1 Reverse	AGGACTGTCCTGCTCTCC
Q5 Guide3 Forward	TGAGGTCTGATCATCTCAAGCAACACC
Q5 Guide3 Reverse	TGAACTCCCGGCTGCACTG
HAL Q5 Forward	TTCCGGCTCGTATGTTGTGTGG

HAL Q5 with EcoRI Reverse	TTATA**GAATTC**GTTAGTGGCGCTAGCAGG
KLF5 KO T7 Forward	CGCCGATTGTTCGTTCTTCC
KLF5 KO T7 Reverse	TGGATGCGTCGTTTCTCCAA
LIN28A KO T7 Forward	TTCATAGGGCCGTCTCTGGG
LIN28A KO T7 Reverse	ATAATTGGTGCTGCCTCCCAC

Table 2. 5 Primer combinations with annealing temperatures.

Primer Mix	**Annealing Temperature**
KLF17 Exon 3 Forward + KLF17 Exon 3 Reverse	68 °C
HAL Forward with Sac1 + HAL Reverse with Nhe1	70 °C
HAR Forward with Asc1 + HAR Reverse with Fse1	66 °C
Random Integration Forward + Random Integration Reverse	65 °C
KLF17 Out PCR Forward + KLF17 Out PCR Reverse	68 °C
KLF17 (Puro) Forward + KLF17 Out PCR Reverse	70 °C
KLF17 Out PCR Forward + KLF17 (EGFP) Reverse	65 °C
KLF17 In PCR Forward + KLF17 In PCR Reverse	68 °C
Q5 Guide1 Forward + Q5 Guide1 Reverse	67 °C
Q5 Guide3 Forward + Q5 Guide3 Reverse	71 °C
HAL Q5 Forward + HAL Q5 with EcoR1 Reverse	68 °C
KLF5 KO T7 Forward + KLF5 KO T7 Reverse	68 °C
LIN28A KO T7 Forward + LIN28A KO T7 Reverse	72 °C

2.15 Immunostaining of Tissue Culture Plates

To perform immunofluorescence staining, cultured cells in tissue culture plates were first fixed with 4% paraformaldehyde (PFA) for 15-30 minutes at room temperature, followed by washing with PBS. For storage after fixation, plates can be kept at 4°C. Next, cells were permeabilized with 0.1% Triton X-100 in PBS solution for 10 minutes. Cells were then blocked using a blocking buffer composed of 0.5 g BSA, 2252 mg glycine, and 100 µL Tween-20 diluted in 100 mL PBS, incubating for 30 minutes. After blocking, primary antibodies were diluted in primary antibody dilution buffer comprised of 1 g BSA and 100 µL Tween-20 in 100 mL PBS, then added to the cells and incubated overnight at 4°C on a rotator. The primary antibodies used were anti-KLF17 (rabbit polyclonal) at 1:250 dilution and anti-Tra-1-60 (mouse monoclonal) at 1:250 dilution (Table 2.6). Following primary antibody incubation, cells were washed three times with PBS for 5 minutes per wash. Secondary antibodies were diluted in secondary antibody dilution buffer containing 1 g BSA in 100 mL PBS, then applied to the cells for 1 hour at room temperature, protected from light. The secondary antibodies used were goat anti-rabbit Alexa Fluor 594 for KLF17 at 1:1000 dilution and avidin-FITC for detecting biotinylated Tra-1-60 at 1:1000 dilution (Table 2.6). After secondary antibody incubation, cells were washed three times with PBS for 5 minutes per wash. Finally, nuclei were counterstained with DAPI diluted 1:5000 in PBS for 10 minutes, followed by one final PBS wash. Stained cell samples were imaged using a Cytation 5 automated fluorescence microscope. For a 12-well plate, 250 µL of solution was used per well for all antibody incubation steps.

Table 2. 6 List of primary and secondary antibodies.

Antibody	**Catalog Number**
Atlas Polyclonal Anti-KLF17 Antibody (Primary Antibody) (Rabbit)	HPA024629
Invitrogen Monoclonal TRA-1-60 Antibody, Biotin (Primary Antibody) (Mouse)	13-8863-82
BD Pharmingen Avidin FITC	544057
Invitrogen Goat anti-Rabbit Alexa Fluor 594 (Secondary Antibody)	R37117
BioLegend HRP Streptavidin	405210

Chapter 3:RESULTS

3.1 KLF17 Is Specifically Expressed in Naive but Not Primed Pluripotent Cells

Single-cell RNA sequencing data indicates that KLF17 expression is specific to pre-implantation epiblast cells, representing the naive pluripotent state (Yan et al., 2013). After implantation, post-implantation epiblast cells transition to a primed pluripotent identity and lose naive marker expression. As shown in Figure 3.1, the RNA-seq data reveals robust KLF17 expression in pre-implantation but not post-implantation epiblast cells and KLF17 is selectively expressed in the naive cells. To validate the naive-specific expression of KLF17, I performed quantitative PCR (qPCR) analysis on several cell types. I examined KLF17 mRNA levels in NO naive cells, NO fibroblasts, NO primed and dH1f cells, and KLF17-overexpressing derivatives of the NO fibroblasts, NO primed, and dH1f cells (Figure 3.1). The qPCR results confirmed that KLF17 expression is restricted to the naive cells. The somatic cells and primed stem cells showed no detectable KLF17 expression. In summary, these transcriptional datasets confirm that KLF17 specifically marks the naive pluripotent state and is not expressed in primed or differentiated cells.

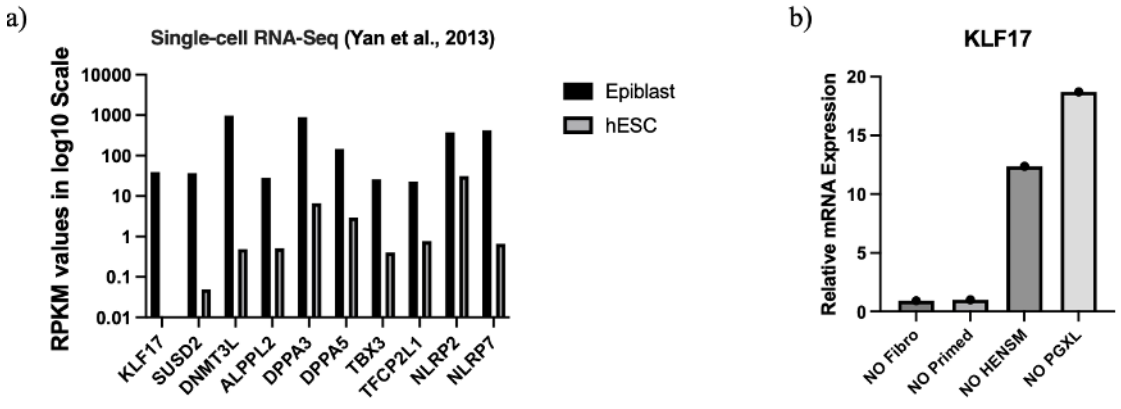


Figure 3. 1 Transcriptional datasets confirming naive-specific expression of KLF17. (a) Single-cell RNA sequencing data adapted from Yan et al. showing RPKM expression values (log10 scale) for key pluripotency factors in pre-implantation (naive) and human ESCs. (b) Quantitative PCR data displayed as relative mRNA expression levels on a log10 scale for KLF17 across several cell types with and without KLF17 overexpression.

3.2 Preparation of The KLF17 Knock-in Targeting Vector Expressing EGFP and Cloning of CRISPR Guide RNAs

The guide RNAs targeting KLF17 and the donor template containing the T2A-GFP cassette flanked by KLF17 homology arms were designed and cloned. The plasmids were nucleofected into iPSCs, and cells were single-cell seeded into 96-well plates post-transfection. Expanded clones were screened by PCR to identify targeted candidates. PCR-positive clones were validated by Sanger sequencing. Confirmed targeted clones were transfected with Cre recombinase to remove the selection cassette and optimize GFP expression. After single-cell seeding, PCR and sequencing verified the successful excision of the selection cassette. Validated KLF17-T2A-GFP knock-in clones were transitioned to naive pluripotency, and GFP reporter activity was examined to confirm functional editing (Figure 3.2).

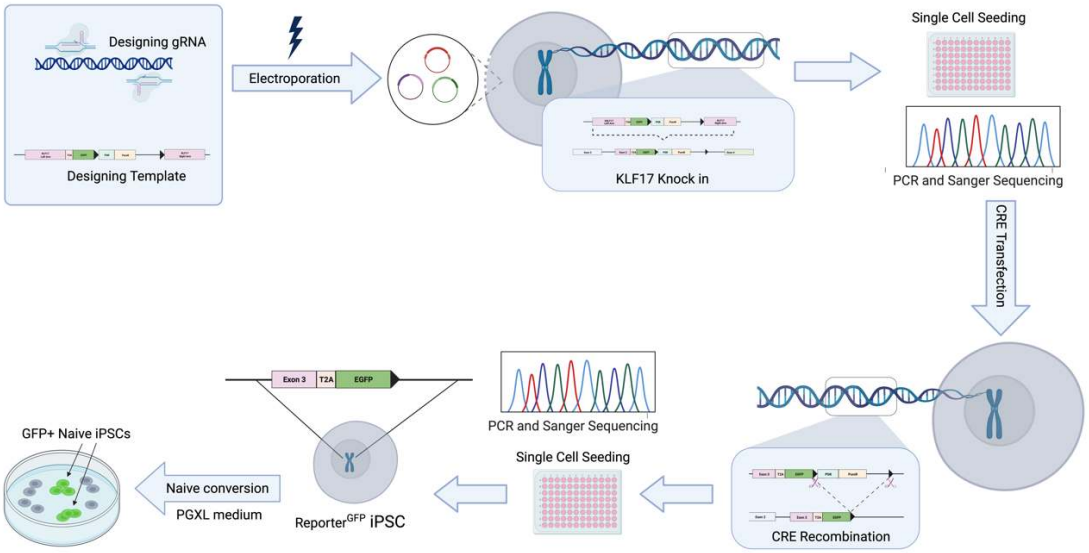


Figure 3. 2 Overview of experimental strategy for generating KLF17 reporter by CRISPR/Cas9n.

KLF17 gene was selected as a candidate for a reporter system to label and sort the naive stem cells from heterogeneous cell populations. The human KLF17 gene is located on chromosome 1 in humans and contains four exons. The coding sequence spans part of the exons1 and exon2 and exon3. Exon 4 is a non-coding region. The start codon (ATG) is located in exon 1, and the stop codon (TAG) is located at the 3' end of exon 3 (Figure 3.3). The coding region of human KLF17 is 1170 base pairs and codes for a 389 amino acid protein of 43 kDa molecular weight.

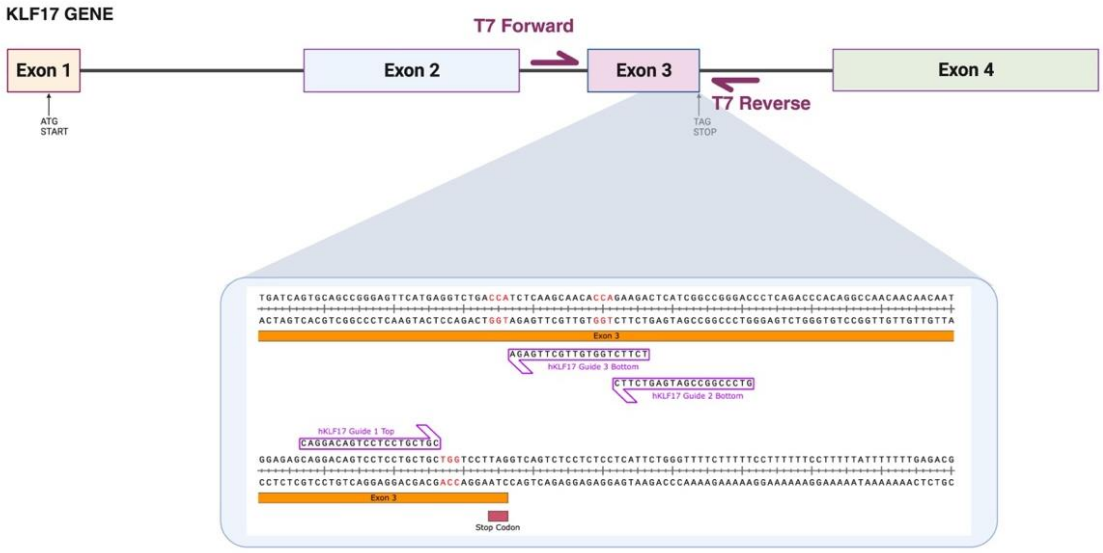


Figure 3. 3 Schematic representation of human KLF17 (Krüppel Like Factor 17) gene structure, and the positioning of T7 primers and guide RNA target site for knock-in strategy.

To generate a reporter, nickase mutant Cas9 was utilized. To edit the genome using Cas9n, the first step was to design single guide RNAs (sgRNAs) targeting the *KLF17* gene with the aim of inserting a T2A- EGFP cassette at the end 3' end of exon 3 of the gene while removing the stop codon. The target sequence in *KLF17* was identified, and three sgRNAs were designed to target 20-24 nucleotides adjacent to available NGG PAM motifs. sgRNAs were first cloned into WT Cas9 expressing lentiCRISPRv2 (Plasmid #52961) and were confirmed by Sanger sequencing (Figure 3.4).

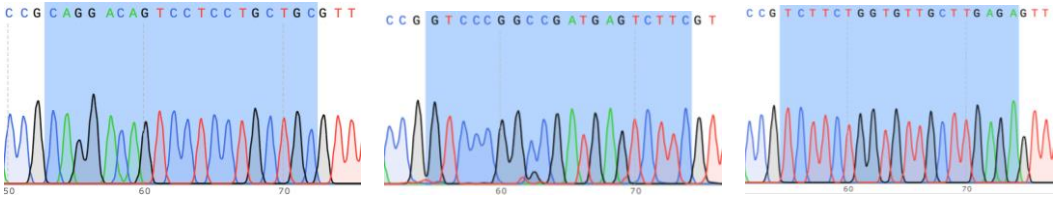


Figure 3. 4 The Sanger sequencing results for cloning the sgRNAs into LentiCRISPR v2 (Plasmid #52961), respectively hKLF17 Guide 1, hKLF17 Guide 2, and hKLF17 Guide 3.

To assess the efficiency of sgRNAs, LentiCRISPR v2 plasmids expressing *KLF17* targeting gRNAs or non-targeting control gRNA were transfected into HEK293T cells. After puromycin selection, genomic DNA was isolated from the transfected cells. PCR was performed to amplify the *KLF17* genomic regions targeted by the different gRNAs. The expected length for PCR products was 833 base pairs. T7 endonuclease assay was performed using the PCR products amplified from the target sites. The purpose of this assay was to determine which gRNAs were most efficient at introducing DNA mutations following double-strand breaks. Analysis of the T7 cleavage patterns revealed that guide 1 and guide 3 generated the most prominent bands (Figure 3.5). This indicated that these two gRNAs were able to cleave their target sites in the genome most effectively. Based on these results, guide 1 and guide 3 were selected for further experiments.

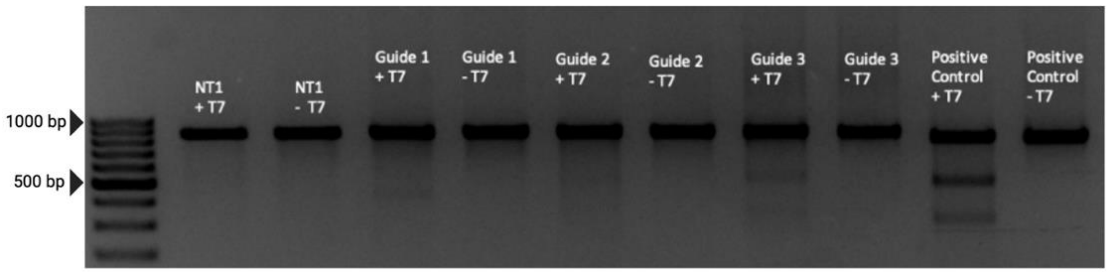


Figure 3. 5 T7 assay results for PCR products targeting the KLF17 gene using guide 1, guide 2, and guide 3. The expected band sizes for PCR products were 833 base pairs. The expected band sizes following T7 digestions were 418 and 415 base pairs for guide 1, 484 and 349 base pairs for guide 2, and 499 and 334 base pairs for guide 3.

On the basis of T7 assay results, guide 1 and guide 3 targeting exon 3 of *KLF17* were cloned into pSpCas9n(BB)-2A-Puro (PX462) V2.0 (Plasmid #62987), which expresses Cas9D10A nickase for gene editing via homology-directed repair (HDR) mechanism (Zhang et al., 2017). The resulting clones were verified by Sanger sequencing (Figure 3.6).

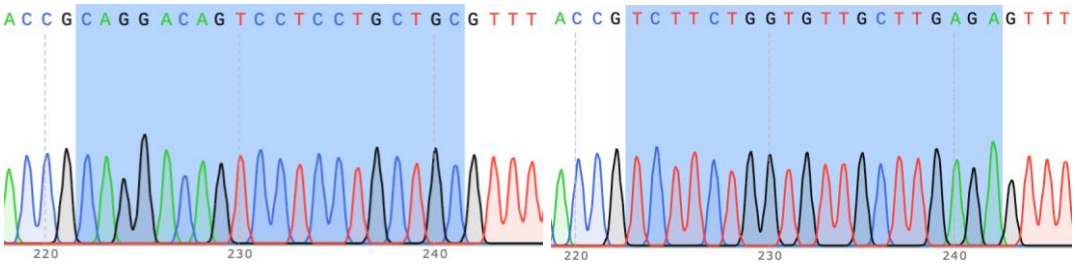


Figure 3. 6 The Sanger sequencing results for cloning the sgRNAs into pSpCas9n(BB)-2A-Puro (PX462) V2.0 (Plasmid #62987), respectively hKLF17 Guide 1, and hKLF17 Guide 3.

To generate the knock-in homology template, a T2A-EGFP-PGK-Puro (Plasmid #83344) containing plasmid was modified. The original plasmid has human Brachyury gene (hTBXT) knock-in homology arms (Figure 3.7). The first aim was to remove the Brachyury gene homology arms and insert human *KLF17* gene sequences into the same site. To insert *KLF17* gene, the first step was to design PCR primers that amplify the *KLF17* regions of the genome with cut sites. For the homology arm left, the PCR product had sacI and NheI restriction sites, and for the homology arm right, the PCR product had AscI and FseI restriction sites. These amplified *KLF17* regions were the homology arms of the targeting vector. Typically, the homology arms should be 200 - 800 base pairs in length (Zhang et al., 2017). Therefore, 500 base pairs upstream and downstream of the integration site were selected to amplify. The PCR products amplified from the genome were the left and right homology arms for cloning into the targeting vector. The right arm contained an additional SacI site, so a sequential cloning approach was used. First, the left homology arm was cloned, followed by the right arm cloning. Therefore, the KLF17 knock-in targeting vector was constructed by cloning the Left and right homology arms of the *KLF17* gene into the backbone targeting template vector. The final version of the plasmid containing both KLF17 homology arms was sequence verified by whole plasmid sequencing.

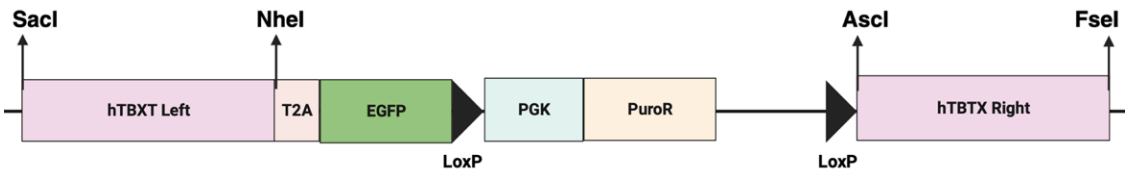


Figure 3. 7 Schematic representation of T-2A-EGFP-PGK-Puro (Plasmid #83344) showing the restriction sites for homology arm cloning.

Next, a previously generated control wildtype iPSC line was nucleofected using the 4D-Nucleofector system with the KLF17 targeting vector and Cas9D10A-gRNA vectors in order to deliver the constructs into the cells. The Cas9D10A enzyme generates single-strand nicks in the DNA at the target site specified by the guide RNAs. When the targeting vector is co-delivered, the HDR pathway repairs these nicks using the homology arms on the vector as a template for precise incorporation of the construct into the genomic *KLF17* locus (Figure 3.8). Following nucleofection, puromycin selection was applied to select cells that had integrated the targeting vector containing the puromycin resistance gene. After puromycin selection, iPSCs were resuspended into a single-cell suspension and seeded individually into wells of 96-well plates in order to generate single clones. The individual clones generated from single cells were expanded. Finally, genomic DNA was isolated from the iPSC clones, and PCR was performed with the primers specific to the homology arms in the targeting vector. This PCR validation enabled the identification of correctly targeted clones with the integration of the construct.

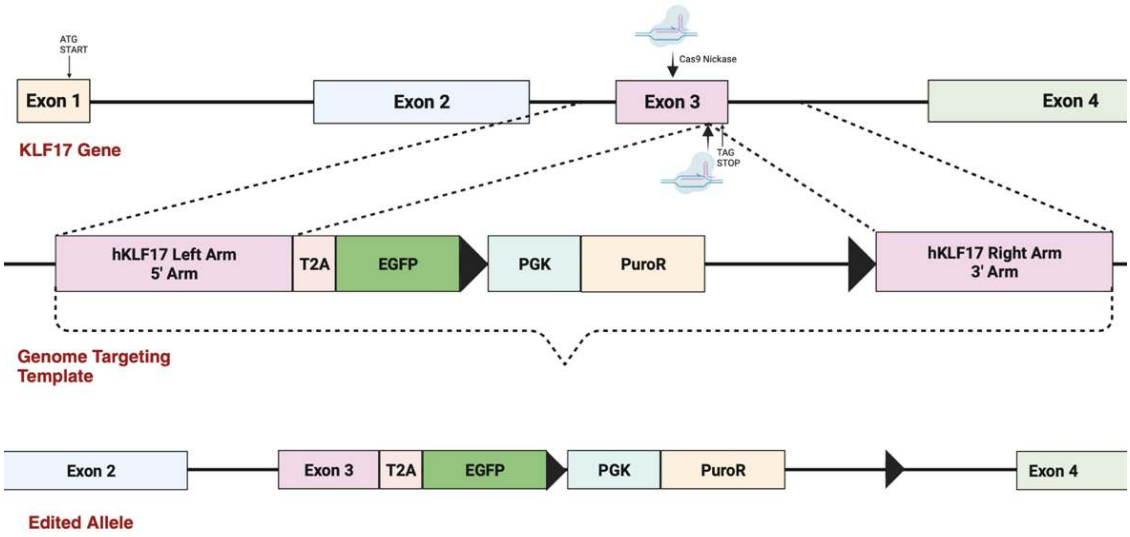


Figure 3. 8 Targeted knock-in of KLF17 using CRISPR/Cas9 nickase-mediated homology-directed repair. Schematic of the genomic KLF17 locus before and after

coupling with the targeting vector. The Cas9D10A nickase introduces single-strand breaks at the target site in exon 3 directed, which enables incorporation of the targeting vector containing left and right homology arms flanking the T-2A-EGFP-PGK-Puro cassette through homology-directed repair.

3.3 PCR Validation of KLF17 Knock-in Clones Generated by Genome Editing

PCR was performed on genomic DNA isolated from the expanded iPSC clones after single cell seeding and selection. PCR was utilized to validate proper targeted integration of the T2A-EGFP-PGK-Puro knock-in cassette at the *KLF17* locus in iPSC clones after CRISPR/Cas9 nickase genome editing. Four types of PCR validations were performed to characterize the clones: random integration PCR, OUT PCR, EGFP PCR, and PURO PCR (Table 3.1). Random integration PCR was designed with one primer binding to the plasmid backbone outside the region of homology arms. This enabled the detection of random integration events, as specific amplification would only occur if plasmid beyond the homology arms were incorporated. Clones showing amplification with the random integration PCR were determined to have random plasmid integration into the genome instead of correct targeted integration. These randomly integrated clones were not selected for further analysis since they do not have the desired targeted modification of the *KLF17* locus. Only clones with no amplification from the random integration PCR were considered candidates with properly targeted clones (Figure 3.9).

Table 3. 1 PCR screening of clones. This table shows results from four types of PCR validations performed: random integration, OUT, EGFP, and PURO PCR.

	EXPECTED BANDS			
PCR Name	**Homozygous Integration**	**Heterozygous Integration**	**No Integration**	**Random Integration**
Random Integration PCR	-	-	-	1503 bp
OUT PCR	3912 bp (Long extension)	3912 bp (Long extension) 1247 bp (Short extension)	1247 bp	-
EGFP PCR 5′ PCR	796 bp	796 bp	-	-
PURO PCR 3′ PCR	1658 bp	1658 bp	-	-

Figure 3. 9 Schematic showing the primer design for random integration PCR. One primer binds within the plasmid backbone outside the region spanned by the left homology arm. The other primer binds within the integration site.

The OUT PCR primers were designed to anneal just outside of the intended integration sites adjacent to the left and right homology arms. For homozygous targeted clones where both *KLF17* alleles are modified, the extension time of the PCR can determine the detection of amplification. With a short extension time, no band would be observed since the ~4kb integrated allele would not be amplified. However, using a long extension time would allow amplification of the full ~4kb integrated locus. For heterozygous clones where one *KLF17* allele is unmodified, a band is expected even with short extension times due to amplification of the wildtype 1247 bp locus. Therefore, by manipulating the OUT PCR extension time, homozygous and heterozygous targeted clones could be distinguished based on the presence or absence of an amplified product. (Figure 3.10). The OUT PCR primers flank the intended integration site, so amplification indicates that the unmodified *KLF17* locus is still present. However, a lack of OUT PCR amplification alone is not definitive evidence of correct targeting. The clones must also be screened with the EGFP PCR and PURO PCR to confirm integration of the 5′ and 3′ ends of the knock-in cassette.

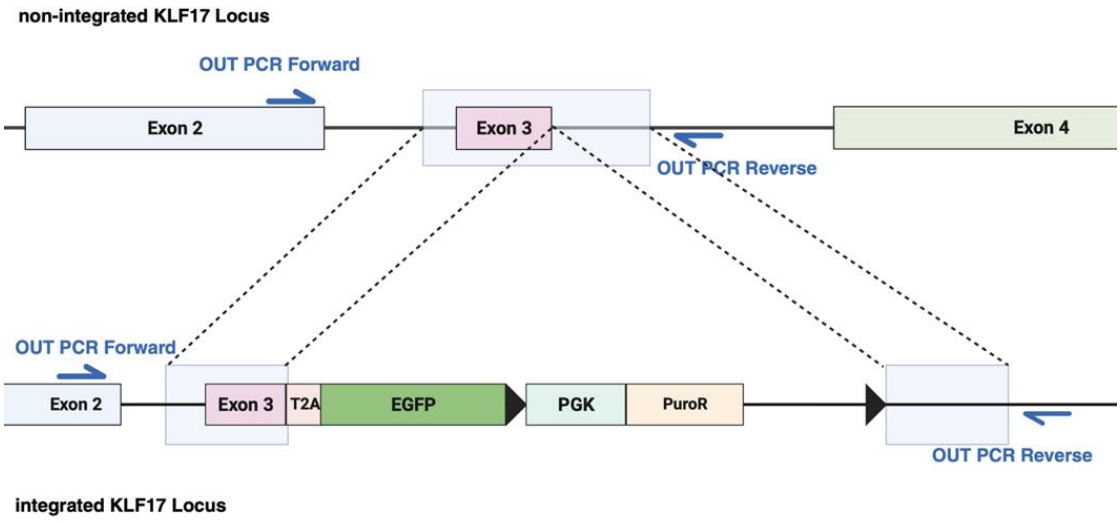


Figure 3. 10 Schematic of the non-integrated KLF17 locus (top) and the targeted integrated locus (bottom). The OUT PCR primer binding sites flanking the intended integration site are indicated. Expected out PCR results: The non-integrated locus with intact primer binding sites will amplify a 1247 bp. The targeted locus has the primer binding sites separated by 3912 bp.

The EGFP PCR was designed to detect the correct integration of the 5′ end of the knock-in construct by amplifying the junction between the EGFP cassette and the genomic region just outside the 5′ homology arm. The forward OUT primer binds within the genomic region, flanking the 5′ integration site outside of the homology arm. The reverse EGFP primer binds inside the EGFP cassette adjacent to the homology arm. This allowed amplification across the 5′ junction from the genome into the EGFP cassette, generating a 796 bp product. Properly targeted clones that have integrated the 5′ end of the construct containing the T2A-EGFP cassette showed specific amplification of the 796 bp band. Clones without targeted integration of the 5′ end did not amplify this junction. Along with the out PCR and Puro PCR results, the presence of the EGFP PCR product spanning into the EGFP cassette indicated correct targeting at the 5′ end of the knock-in construct (Figure 3.11).

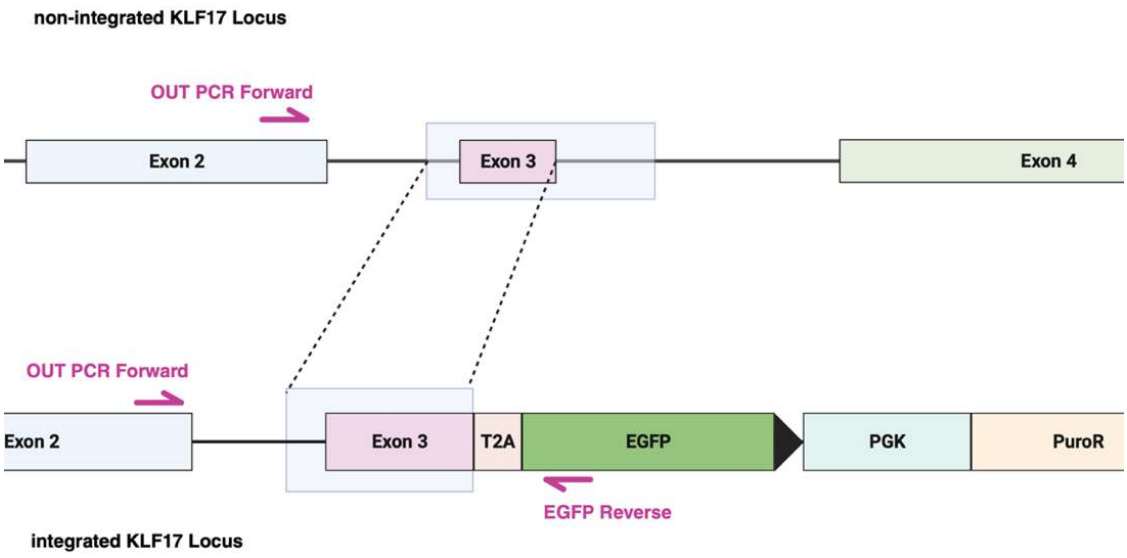


Figure 3. 11 Schematic of the non-integrated KLF17 locus (top) and the targeted locus with 5′ integration (bottom). The EGFP PCR primer binding sites adjacent to the 5′ homology arm are indicated. Expected EGFP PCR results: The non-integrated locus will not amplify. The targeted locus will amplify a 796 bp product spanning the 5′ junction into the EGFP cassette.

The Puro PCR was designed to validate proper integration of the 3' end of the knock-in construct by amplifying the region spanning from the puromycin resistance (PuroR) into the adjacent genomic DNA. PURO forward primer binds inside the PuroR cassette, and the OUT reverse primer binds to the genomic sequence outside the 3' homology arm. This primer combination enabled amplification across the junction between the PuroR and the genomic DNA just outside the 3' homology arm, generating a 1658 bp PCR product (Figure 3.12). Clones that had correctly integrated the 3' end of the construct showed the expected 1658 bp band, while clones without properly targeted integration at the 3' junction did not amplify. When combined with the OUT PCR and EGFP PCR results, the presence of the 1658 bp PuroR junctional PCR product provided evidence for proper targeting at the 3' arm of the integrated construct.

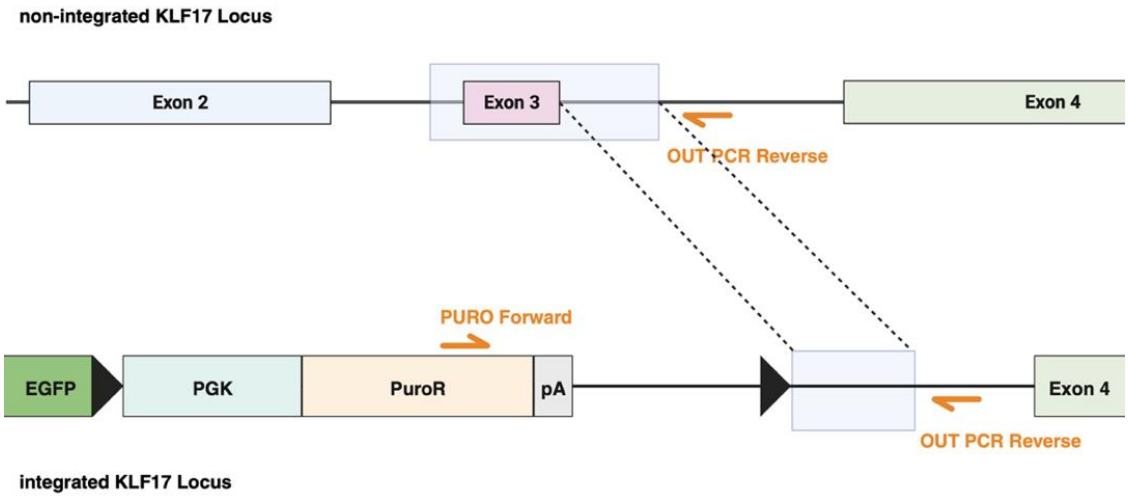


Figure 3. 12 Schematic of the non-integrated KLF17 locus (top) and the targeted locus with 3′ integration (bottom). The Puro PCR primer binding sites adjacent to the 3′ homology arm are indicated. Expected Puro PCR results: The non-integrated locus will not amplify. The targeted locus will amplify a 1658 bp product spanning the 3′ junction into the genome.

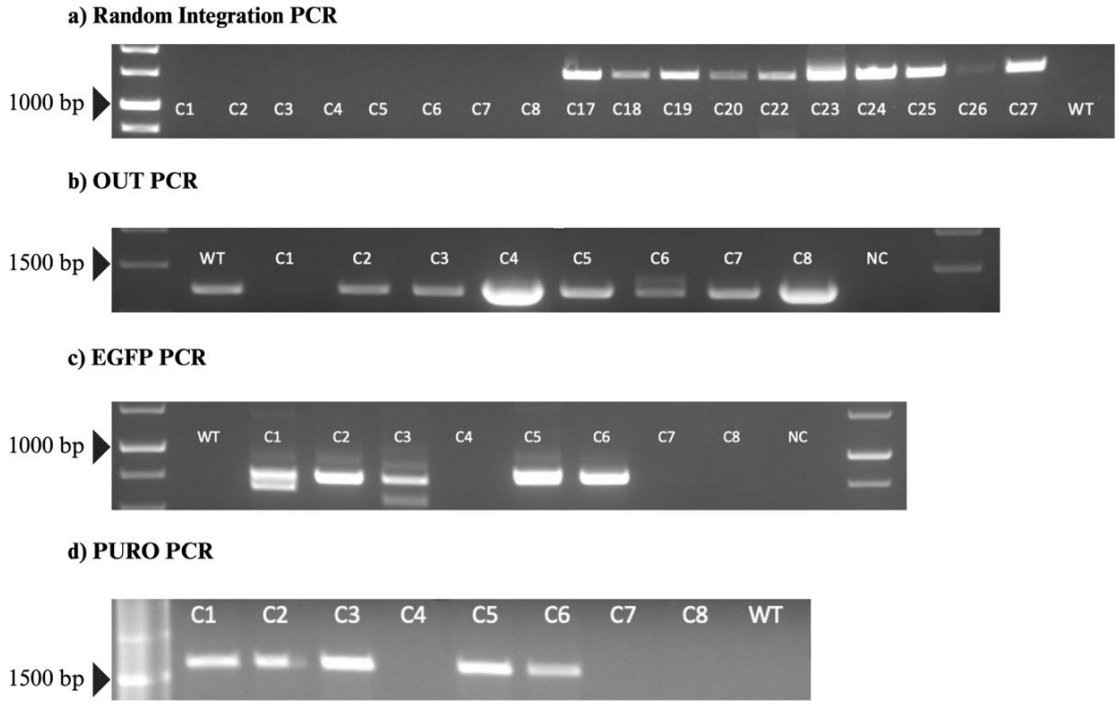


Figure 3. 13 PCR screening of clones. Agarose gel electrophoresis results for the four PCR validations performed: (a) Random integration PCR expected band size 1503 bp. (b) OUT PCR expected band size 1247. (c) EGFP PCR expected band size 796 bp. (d) PURO PCR expected band size 1658 bp.

The initial random integration PCR screening showed no nonspecific amplification for clones 1 through 8, indicating these clones did not have plasmid insertion in the genome. The remaining clones screened exhibited bands suggesting random integration, so they were excluded from further analysis (Figure 3.13). The out PCR was then used to check for potential homozygous versus heterozygous targeting in clones 1-8, which did not have random integration. Clone 1 had faint out PCR bands, meaning the results were ambiguous for this clone. It could still be a candidate for potential homozygous targeting, but it was not definitive.

On the other hand, clones 2, 3, 5, 6, and 7 all had clear OUT PCR bands of the expected 1247 bp size. This specific amplification of the unmodified genomic region in these clones demonstrated they were heterozygotes, with only one *KLF17* allele targeted and the other allele remaining intact. The OUT PCR screening showed ambiguously strong bands for clones 4 and 8 compared to other clones. While band presence can imply an unmodified genomic region, suggesting potential heterozygous targeting, it could also come from entirely non-targeted clones with two unmodified alleles (Figure 3.13).

The EGFP and PURO PCR screens were used to validate proper targeted integration at the 5′ and 3′ junctions, respectively. Clones 4, 7, and 8 did not show the expected bands for either the EGFP or PURO PCRs. The lack of amplification indicated that these three clones did not harbor the integrated construct at either junction site. In contrast, clones 1, 2, 3, 5, and 6 exhibited the anticipated bands in the EGFP and PURO PCRs. For EGFP PCR, these clones showed the 796 bp band spanning the 5′ integration site. For Puro PCR, they showed the 1658 bp band spanning the 3′ site. The presence of bands in both confirmation PCRs provided evidence that these five clones had the complete integrated construct (Figure 3.13)

The EGFP PCR screened for proper targeted integration at the 5′ junction did not show clear, defined bands for most clones except clone 6. Clones 1, 2, 3, and 5 had faint, smeared, and inconsistent amplification in the EGFP PCR rather than the expected discrete 796 bp band. Several factors could contribute to the ambiguous EGFP PCR results, such as polyclonal populations with mixtures of targeted and non-targeted cells after single-cell seeding or continued Cas9 activity after single-cell seeding introducing mutations at the integration site.

To investigate the ambiguous EGFP PCR results and confirm proper integration, the EGFP and Puro PCR amplicons from clones 1, 2, 3, 5, and 6 were purified and

subjected to standard Sanger sequencing analysis using the forward and reverse PCR primers. Sequencing was performed on the 796 bp EGFP junctional PCR product to analyze the 5′ integration site. The 1658 bp Puro PCR product was sequenced to analyze the 3′ targeted locus.

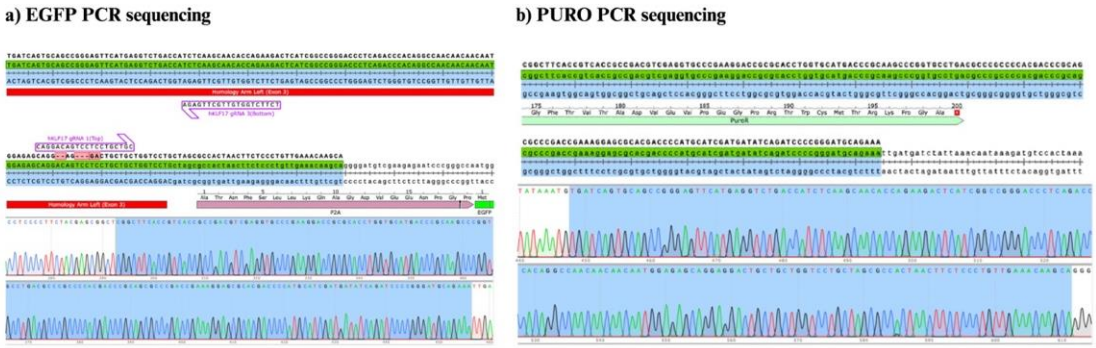


Figure 3. 14 Sanger sequencing results of clone 6 PCR products. (a) EGFP PCR product sequenced with forward primer, showing the expected sequence spanning the 5′ integration site. (b) Puro PCR product sequenced with reverse primer, showing the expected sequence spanning the 3′ integration site.

Sanger sequencing results of the junctional PCR products showed that all clones, including clone 6, contained variations from the expected targeted sequence. Point mutations, deletions, and insertions were observed at the 5′ integration sites in the sequenced clones. These mutations likely arose due to unwanted Cas9n enzyme activity after the initial targeted integration occurred. Even following single-cell seeding, the Cas9n present in the cells could remain active and continuously generate double-strand breaks at the target locus. This post-integration cleavage by Cas9n could have led to the accumulation of mutations introduced during the error-prone non-homologous end-joining repair process. However, the Puro PCR analysis, which detected integration at the

3′ end of the construct, showed the expected 1658 bp band for all clones screened. This included clones 1, 2, 3, 5, and 6, which were analyzed by Sanger sequencing. The consistent Puro PCR results, generating the anticipated band, demonstrated successful and precise integration at the intended 3′ locus in the genome across all clones tested. This verified that the 3' integration junction was intact in the targeted clones despite additional Cas9-mediated mutations arising after the initial proper targeting at 5′ site (Figure 3.14).

The identification of mutations at 5′ integration site in all sequenced clones demonstrated that Cas9n activity persisted after clonal expansion. This introduced undesired variations from the intended knock-in sequence despite the clones containing the correctly integrated construct. To prevent the undesirable off-target effects, Cas9n activity needed to be eliminated after the initial targeting. One strategy to achieve this is modifying the PAM sequence in the targeting construct. Cas9n needs an appropriate PAM motif adjacent to the target site to be able to bind and cut the DNA (Ran et al., 2013). To prevent residual Cas9n activity after the desired integration occurs, the PAM sequence can be mutated in the initial targeting construct. By altering the PAM sequence in the targeting vector to an incompatible variant before performing the targeting, Cas9n could not recognize or cut at the integrated locus. This would prevent additional off-target mutations from accumulating later due to ongoing Cas9n function.

3.4 Introducing PAM Mutations into The KLF17 Targeting Cassette

To mutate the PAM site, PCR primers were designed that contain the desired nucleotide substitutions to make the PAM incompatible with Cas9n binding. PCR was performed with Q5 polymerase and the targeting vector template to amplify the full plasmid. Because of the large size of the original KLF17 targeting vector, Q5 polymerase

could not amplify the entire construct for site-directed mutagenesis successfully. To overcome this, just the target region containing the left homology arm was subcloned into the smaller pUC19 vector. Primers were designed to PCR amplify the left homology arm region from the original targeting vector and clone it into pUC19. After the cloning step, primers containing the desired point mutations to alter the PAM sequence were designed to introduce the PAM mutations into the left homology arm. These primers changed the PAM sequence from TGG to TGC and from TGG to TGA. PCR amplification of the PAM mutated left homology arm from the pUC19 backbone was performed, followed by cloning back into the KLF17 targeting vector, resulting in the incorporation of the PAM mutations in the targeting vector. The TGG to TGC change converted glycine to alanine, while the TGG to TGA change maintained the same amino acid. This approach allowed the introduction of the required PAM mutations into the KLF17 targeting vector (Figure 3.15).

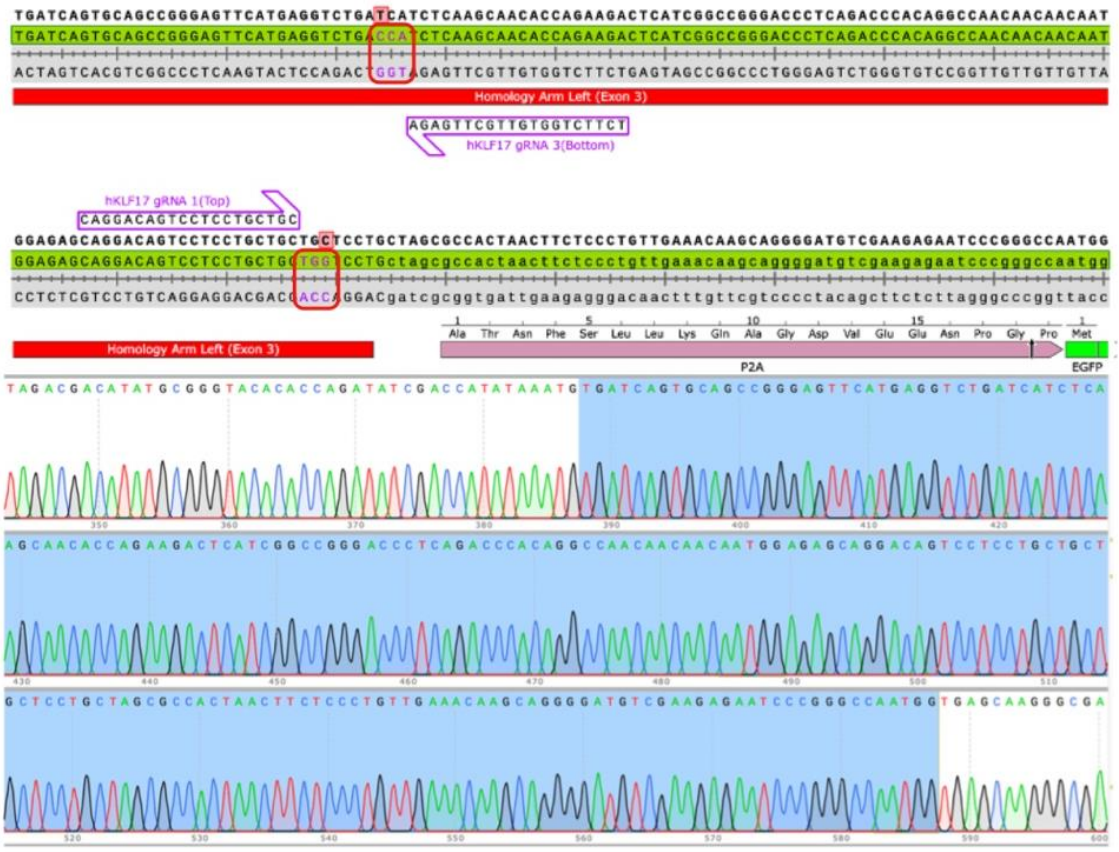


Figure 3. 15 Sanger sequencing of PAM mutation incorporation into KLF17 targeting vector. Incorporation of the intended PAM mutations, TGG to TGC, and TGG to TGA into the left homology arm are indicated. The PAM sequence is indicated by the red squares.

3.5 PCR Screening of PAM-Mutated KLF17 Clones

A second round of nucleofection using the PAM-mutated homology template was performed, and genomic DNA was isolated from the expanded single-cell clones for additional PCR screening. EGFP PCR was performed on 93 clones to detect targeted integration at the 5′ end. 36 clones showed the expected 796 bp EGFP amplification band, indicating successful 5′ integration (Figure 3.16). These 36 potential positive clones were then analyzed by random integration PCR. Only 8 clones lacked amplification in the random integration assay, suggesting they did not harbor random plasmid insertion

(Figure 3.17). The 8 candidate clones were further analyzed by PURO PCR to detect 3′ integration and OUT PCR to screen for heterozygous vs. homozygous targeting.

PURO PCR showed the expected 1658 bp band for 6 of the 8 clones, verifying the correct 3′ integration in these clones. However, the two clones did not amplify a Puro band, suggesting a lack of proper 3′ targeting (Figure 3.17). OUT PCR screening found that seven clones exhibited amplification of the unmodified locus, indicating they were heterozygously targeted. Clone 2E1 uniquely lacked an OUT PCR band, implying potential homozygous targeting with the integration of both *KLF17* alleles (Figure 3.17).

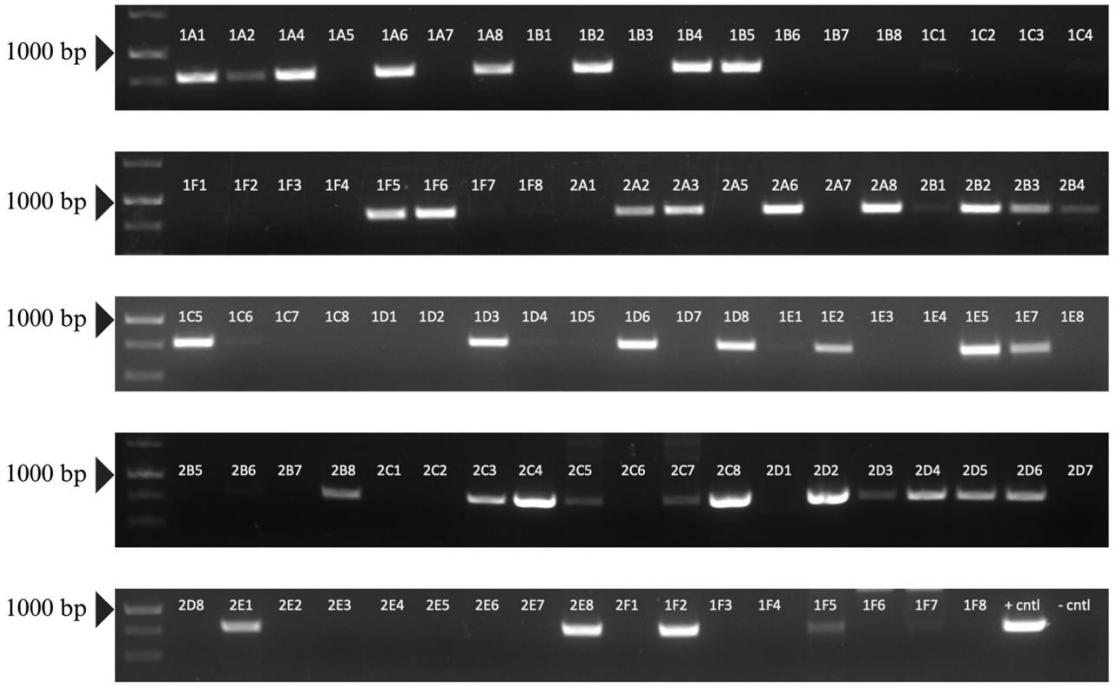


Figure 3. 16 Screening of 93 clones after introducing PAM mutations by EGFP PCR. Clones showing the expected 796 bp band contain the correctly integrated EGFP cassette.

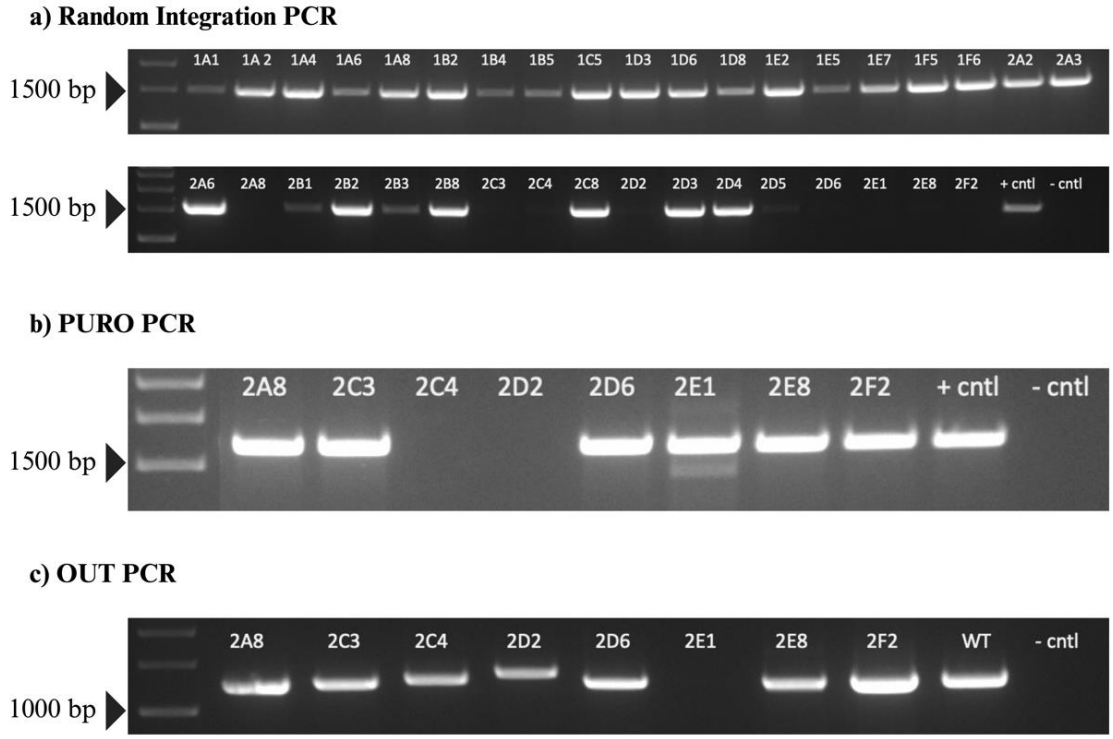


Figure 3. 17 PCR validations of EGFP integrated clones. (a) Random integration PCR of 36 clones. The expected band size is 1503 bp. The amplification indicates uncontrolled plasmid insertion. (b) Puro PCR results for 8 clones. The 1658 bp band validates the correct 3′ integration. (c) Out PCR analysis of 8 clones. Amplification signifies an unmodified locus, while no band suggests a potential homozygote clone.

The clones that were positive for both 5′ and 3′ integration by EGFP and PURO PCR were further analyzed by Sanger sequencing of the junctional amplicons. Sequencing was performed on the 796 bp 5′ EGFP product to validate integration at this end, as well as the 1658 bp 3′ PURO product to confirm proper 3′ targeting. Sequence analysis of the 5′ and 3′ PCR products for the 6 clones verified successful and precise integration at the intended genomic loci without random insertions or mutations. All clones exhibited the expected sequences spanning the 5′ and 3′ junctions between the targeting vector and adjacent genomic regions (Figure 3.18).

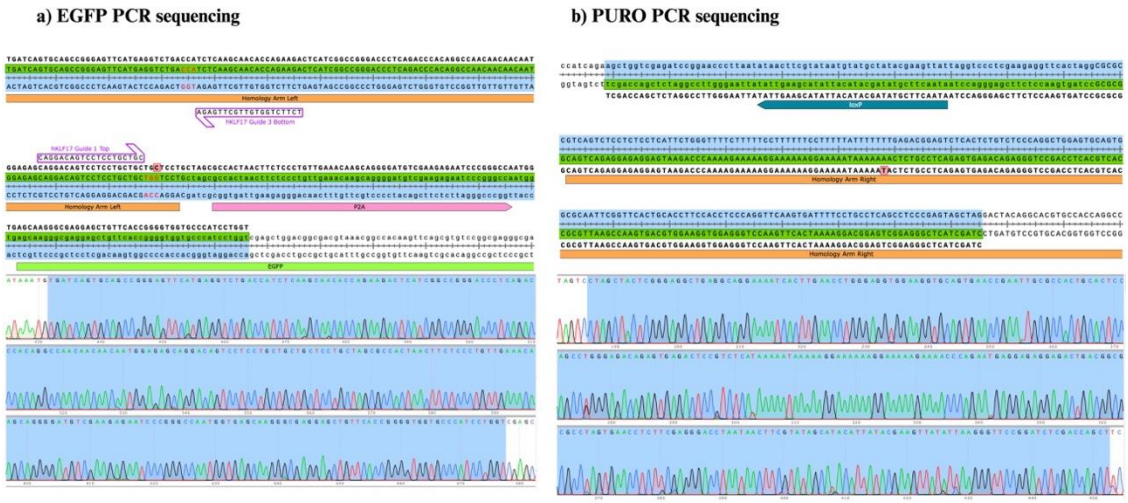


Figure 3. 18 Sanger sequencing results of clone 2E1 PCR products. (a) 796 bp EGFP PCR product sequenced, showing the expected targeted sequence spanning the 5′ integration site. (b) 1658 bp Puro PCR product sequenced, displaying the expected targeted sequence across the 3′ integration site.

3.6 PCR Validation of PGK-PuroR Excision by Cre Recombination in Clone 2E1

The initial targeting vector integrated into clone 2E1 contained a PGK-PuroR site flanked by loxP sites downstream of the EGFP knock-in sequence. After identifying the correctly targeted clone, Cre recombinase (Addgene #140284) was transfected into clone 2E1 cells to remove the PGK-PuroR site flanked by loxP sites in the integrated construct. This leaves a single loxP site downstream of the EGFP cassette and restores the polyA signal adjacent to EGFP, which is important for stabilizing EGFP mRNA (Figure 3.19). The Cre recombinase plasmid contained a Blastidicin resistance gene for selection. After Cre transfection, Blastidicin selection was applied to select cells that had taken up the Cre plasmid and excised PGK-PuroR. Surviving cells were seeded into 96-well plates as single cells and expanded into clonal populations. Genomic DNA was extracted from the clones, and PCR screening was performed again using OUT PCR and IN PCR primers

spanning the recombined junction (Table 3.2). This step validated the successful removal of PGK-PuroR in clone 2E1 and the restoration of the EGFP polyA tail configuration needed for optimal expression.

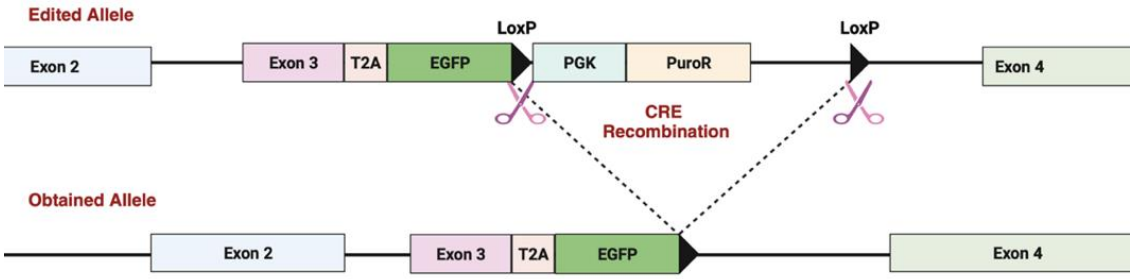


Figure 3. 19 *KLF17* locus after Cre recombinase treatment. Schematic shows the integrated construct before Cre recombination (edited allele) with PGK-PuroR cassette. After Cre recombination, PGK-PuroR is excised leaving, one loxP site.

Table 3. 2 Expected PCR products for OUT and IN PCR. It shows expected PCR bands with or without Cre recombination in clone 2E1.

	EXPECTED BANDS		
PCR Name	**W/OUT CRE**	**W/ CRE**	**WT NO iPSC**
IN PCR	3382 bp	1593 bp	717 bp
OUT PCR	3912 bp	2123 bp	1247 bp

Out PCR was performed on the 24 clones after Cre treatment to check for excision of PGK-PuroR. Four clones - C10, C11, C15, and C21 - showed a band at 2123 bp, reduced from the 3912 bp size expected before Cre recombination. This suggests successful excision of the PGK-PuroR cassette in these four clones (Figure 3.20). These four clones were then analyzed by IN PCR using primers inside the homology arms (Figure 3.21). All four clones exhibited a 1593 bp band, smaller than the 3382 bp product expected before Cre recombination. Detection of this reduced 1593 bp in PCR product in the four clones provided further evidence that PGK-PuroR was correctly excised by Cre recombinase (Figure 3.22).

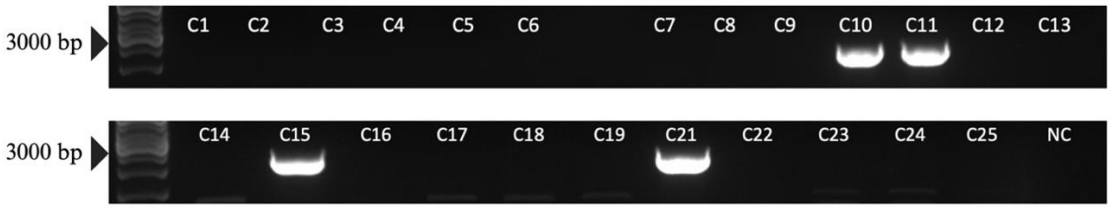


Figure 3. 20 Screening of 24 clones after CRE recombination by OUT PCR. Clones showing the expected 2123 bp band indicating proper PGK-PuroR excision.

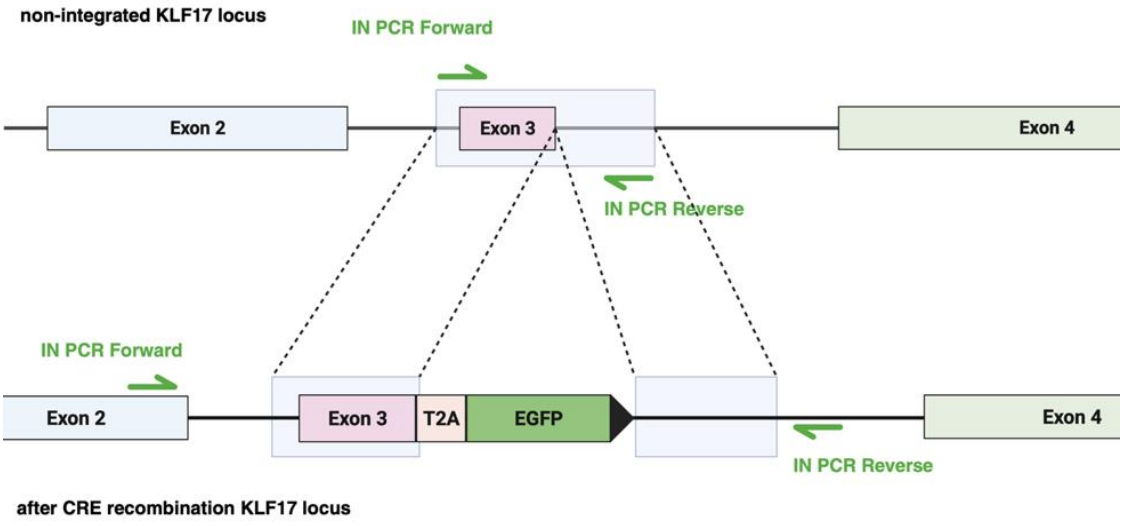


Figure 3. 21 IN PCR primer locations (arrows) inside the 5' and 3' homology arms. Expected PCR products are 717 bp for the non-integrated *KLF17* locus and 1593 bp for the *KLF17* locus after CRE recombination.

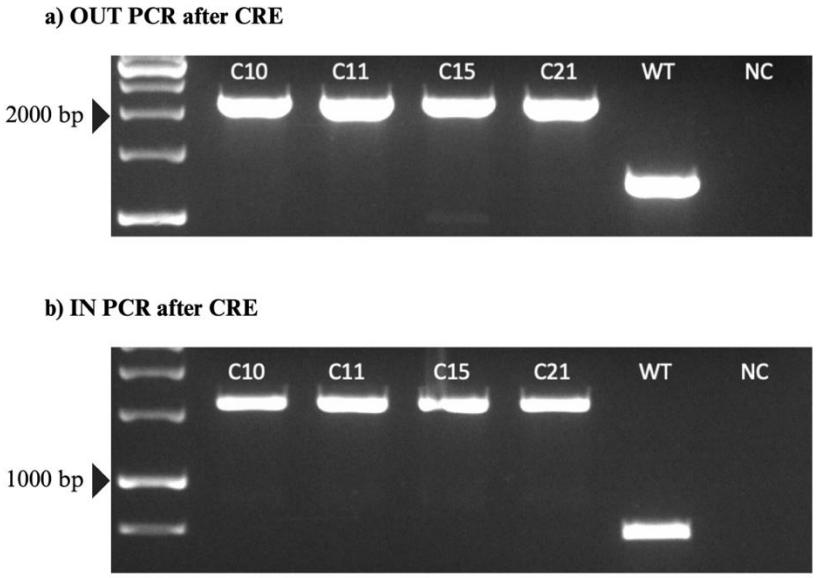


Figure 3. 22 PCR Validation of CRE recombination. (a) OUT PCR expected band size 2123 bp for integrated *KLF17* locus after Cre recombination. (b) IN PCR, expected band size 1593 bp integrated *KLF17* locus after Cre recombination. For wild-type locus, OUT PCR gives a 1247 bp band, and IN PCR gives a 717 bp band.

Next, IN PCR products of the four clones positive for PGK-PuroR excision (C10, C11, C15, and C21) were sequenced by Sanger sequencing for further validation. Sequencing of the 1593 bp IN PCR products using forward and reverse primers confirmed precise Cre recombination in these clones. Sequence alignment showed the expected junction between the 3' homology arm and EGFP after removal of PGK-PuroR. No mutations were observed compared to the expected sequence following clean excision (Figure 3.23). These 2E1 C10, C11, C15, and C21 clonal cell lines obtained were suitable for further experiments and characterization.

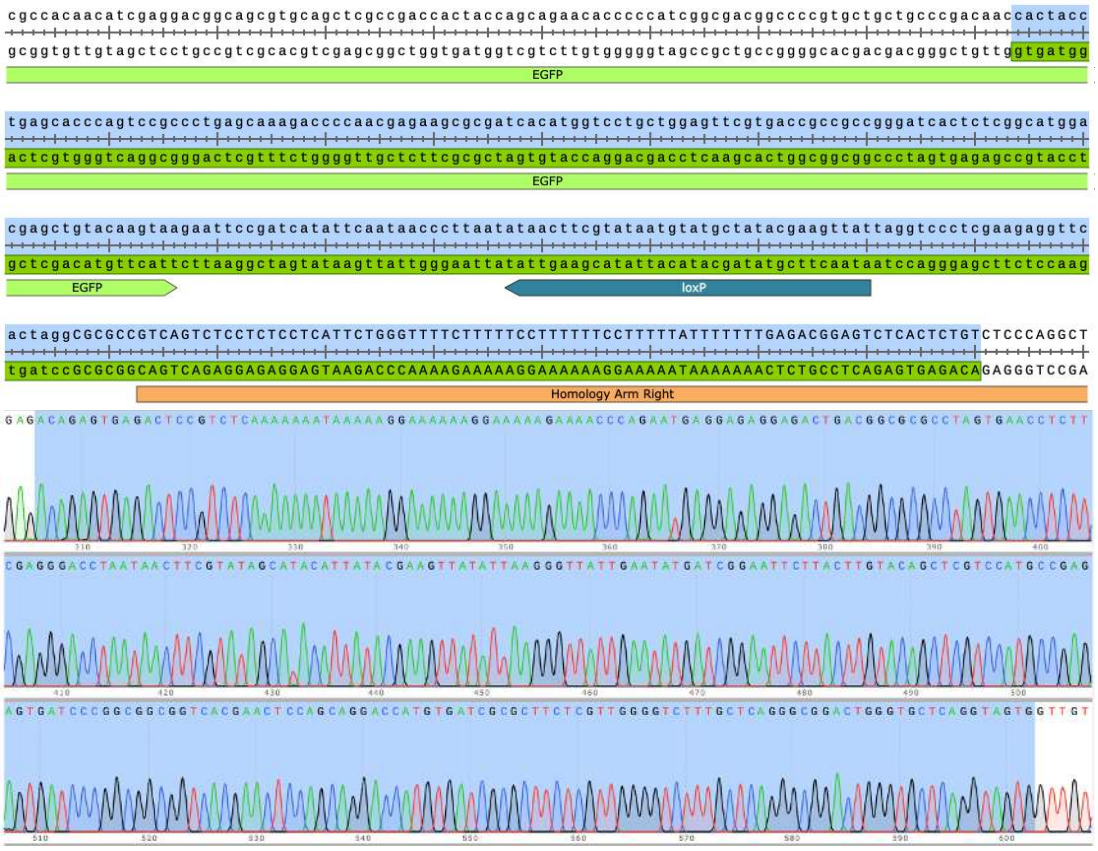


Figure 3. 23 Sanger sequencing results of 2E1 C10 PCR product. 1593 bp IN PCR product sequenced, showing the expected PGK-PuroR removal at the 3′ integration site.

3.7 Assessing GFP Reporter Expression in the 2E1 c10 Knock-in Cell Line

The 2E1 c10 KLF17 knock-in reporter line was converted to naive pluripotency to examine GFP activity under KLF17 expression in the naive state. iPSCs were plated on irradiated MEFs in mTeSR medium. On day 2, the medium was switched to a 1:1 ratio of mTeSR and CRM1. On day 3, only CRM1 was used. On day 5, the medium was changed to CRM2. Around days 9-10, naive colonies appeared and were passaged onto PGXL medium on iMEFs. The naive cells were passaged at a 1:3 ratio every 3-4 days (Figure 3.24).

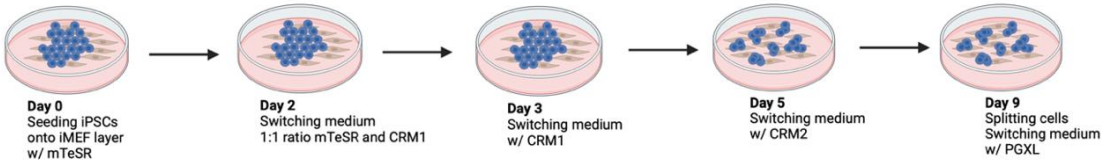
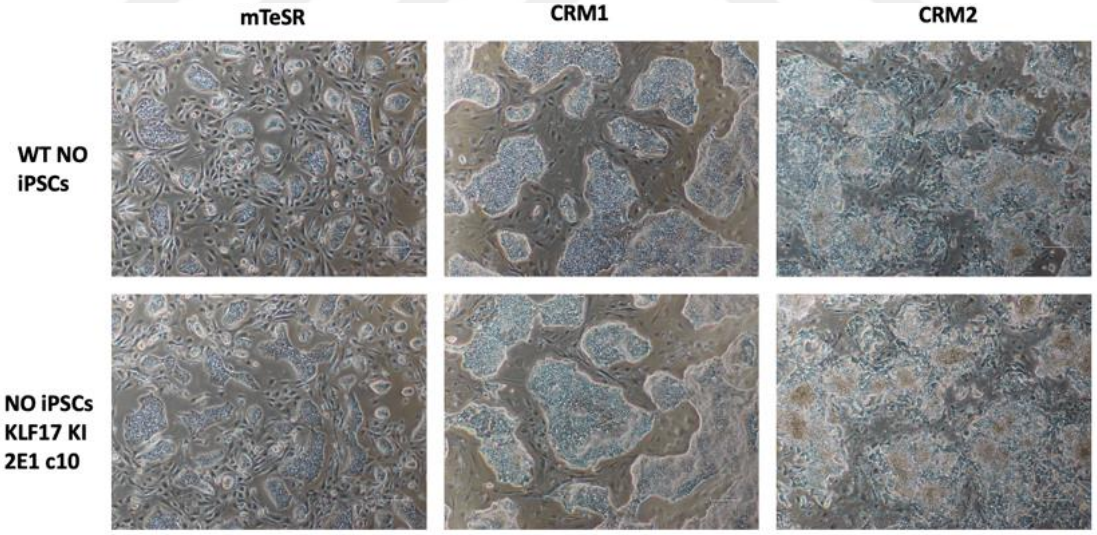


Figure 3. 24 Overview of the stage-wise protocol for resetting primed iPSCs to naive state.

During the transition from primed to naive pluripotency, the iPSCs underwent noticeable morphological changes. In the primed state, the cells exhibited a flattened, epithelial-like morphology. However, after undergoing metabolic and signaling resetting, the emerging naive cells took on a more domed, compact morphology with clustered growth (Figure 3.25).



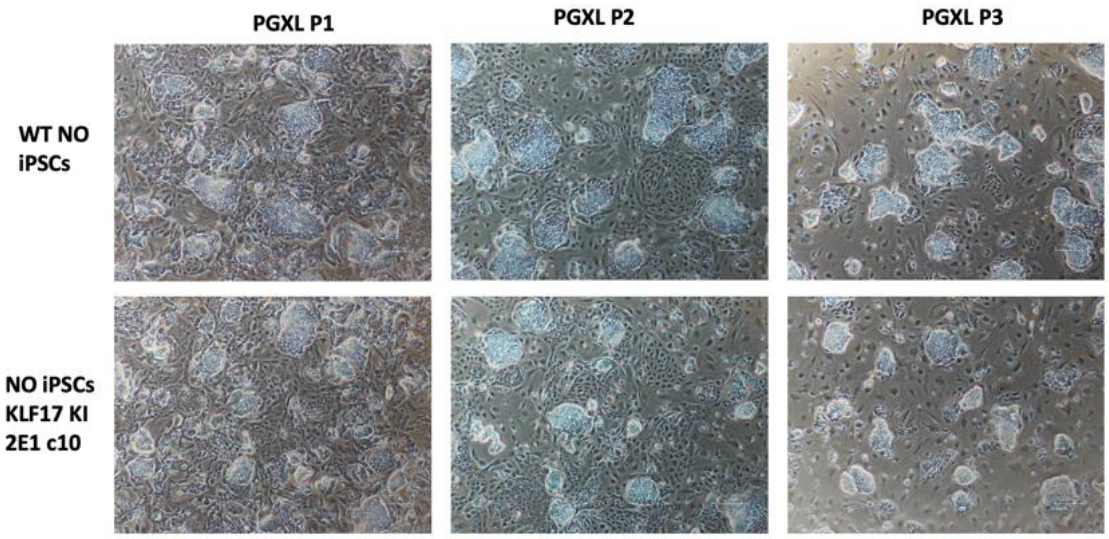


Figure 3. 25 Phase contrast microscopy of iPSCs during transition from primed to naive pluripotency.

At each stage of the naive conversion protocol, RNA samples were collected from the iPSC cultures for analysis by quantitative PCR (qPCR). The expression levels of KLF17, SUSD2, and DNMT3L were quantified by qPCR in order to assess the upregulation of these naive markers during the timeline of resetting primed iPSCs into the naive state. The qPCR data showed that expression of the naive markers KLF17, SUSD2, and DNMT3L increased in CRM1 and CRM2 compared to the original primed iPSCs. This indicates an upregulation of the naive transcriptional program during the resetting steps. Interestingly, the expression of these markers decreased again once the cells passaged in PGXL. However, their expression remained higher compared to the original primed cells (Figure 3.26). The transient further increase in naive marker expression in CRM1 and CRM2 is likely due to the inclusion of the histone deacetylase inhibitor in CR1 media. HDAC inhibition induces widespread chromatin accessibility, allowing activation of previously silenced pluripotency genes.

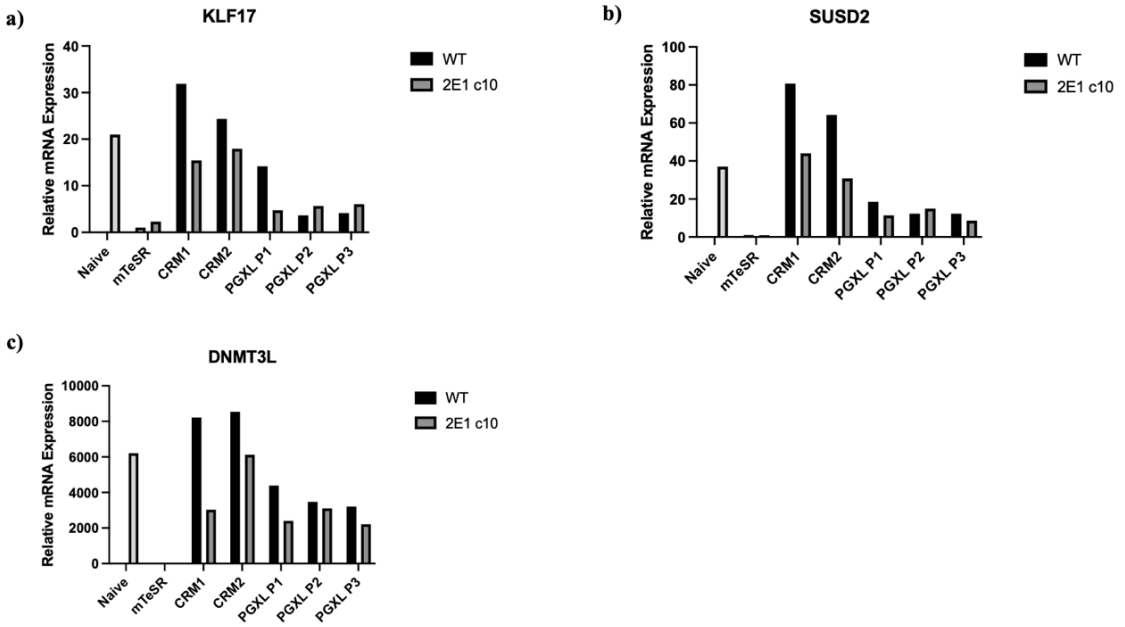


Figure 3. 26 Quantitative PCR data displayed as relative mRNA expression levels for KLF17, SUSD2 and DNMT3L.

To confirm the functional expression of the integrated EGFP reporter in the 2E1 c10 knock-in line, qPCR primers were designed to anneal to the EGFP and the KLF17 exon 3 regions (Figure 3.27). qPCR was performed using these primers on both wildtype control and the 2E1 c10 knock-in cells. The results showed significant amplification of EGFP only in the 2E1 c10 knock-in cells but not in the wildtype control. This demonstrates the specific expression of the targeted EGFP reporter from the engineered KLF17 locus in the genome-edited 2E1 c10 cells (Figure 3.28). In summary, qPCR analysis successfully validated the expression of the knocked-in reporter in the 2E1 c10 cell line, confirming functional integration of the KLF17-T2A-EGFP allele.

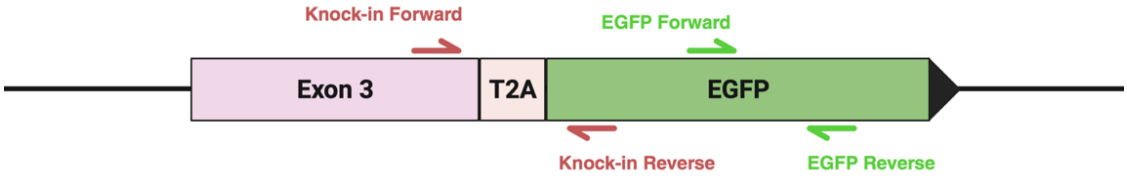


Figure 3. 27 Schematic showing the EGFP and Knock-in qPCR primer design.

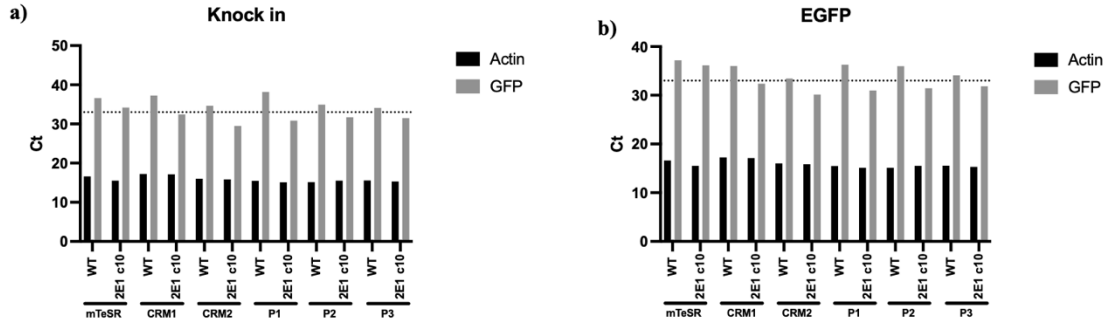


Figure 3. 28 qPCR detection of EGFP and knock-in alleles. qPCR analysis using primers spanning the 5' integration junction (a) or targeting EGFP (b) specifically amplifies the knock-in allele in the genome. Ct values for amplifying each allele are plotted, with a dashed line indicating the detection threshold at a Ct of 33.

3.8 KLF17, NANOG, and LIN28A Overexpression Effect on Naive State Induction

Lentiviral constructs expressing KLF17, mCherry, and NANOG/LIN28A were individually packaged into viral particles (using plasmids from Addgene #120503, #135003, #21163). Next, the primed iPSCs were infected with these viruses separately. This allowed overexpression of the KLF17, NANOG/LIN28A, and mCherry (control). 72 hours after infection, the primed cells were split onto MEF feeders to start the naive induction stages. RNA samples were collected at multiple time points during the naive conversion protocol to enable tracking of gene expression changes.

For the KLF17 and mCherry overexpression experiment set, qPCR analysis showed upregulation of the naive markers KLF17, SUSD2, and DNMT3L in both conditions after naive conversion relative to the original primed cells. However, overexpression of KLF17 resulted in significantly higher levels of SUSD2 and DNMT3L activation compared to mCherry control cells during the naive resetting process. The enhanced induction of SUSD2 and DNMT3L markers with KLF17 overexpression indicates exogenous expression of KLF17, which further stimulates the naive conversion steps (Figure 3.29).

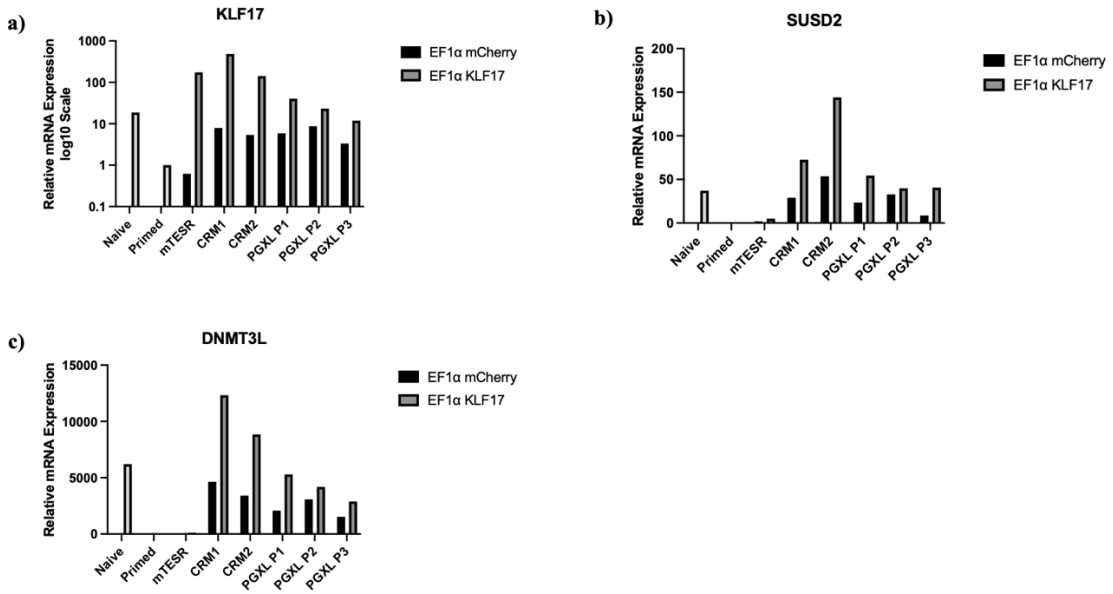


Figure 3. 29 qPCR analysis of naive marker expression during naive conversion after KLF17 overexpression. Expression of KLF17, SUSD2, and DNMT3L during naive conversion of KLF17 and mCherry overexpressed cells.

For the NANOG/LIN28A overexpression experiments, qPCR analysis showed high induction of NANOG in the cells transduced with NANOG lentivirus, as expected. However, LIN28A expression was only mildly increased with LIN28A lentiviral

transduction because LIN28A is already abundantly expressed in primed iPSCs (Figure 3.30). Overexpression of NANOG during naive resetting similarly enhanced activation of the naive markers KLF17, SUSD2, and DNMT3L compared to uninfected wildtype control (Figure 3.31). This suggests that NANOG overexpression further stimulates the naive conversion steps. In summary, overexpression of KLF17 or NANOG helps enhance resetting to the naive pluripotent state.

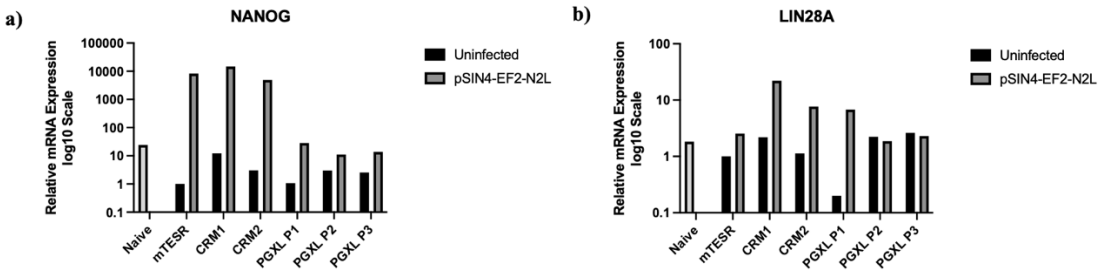


Figure 3. 30 qPCR analysis of NANOG and LIN28A expression during naive conversion after NANOG/LIN28A overexpression. Expression of NANOG (a) and LIN28A (b) during naive conversion of NANOG/LIN28A overexpressed and uninfected cells.

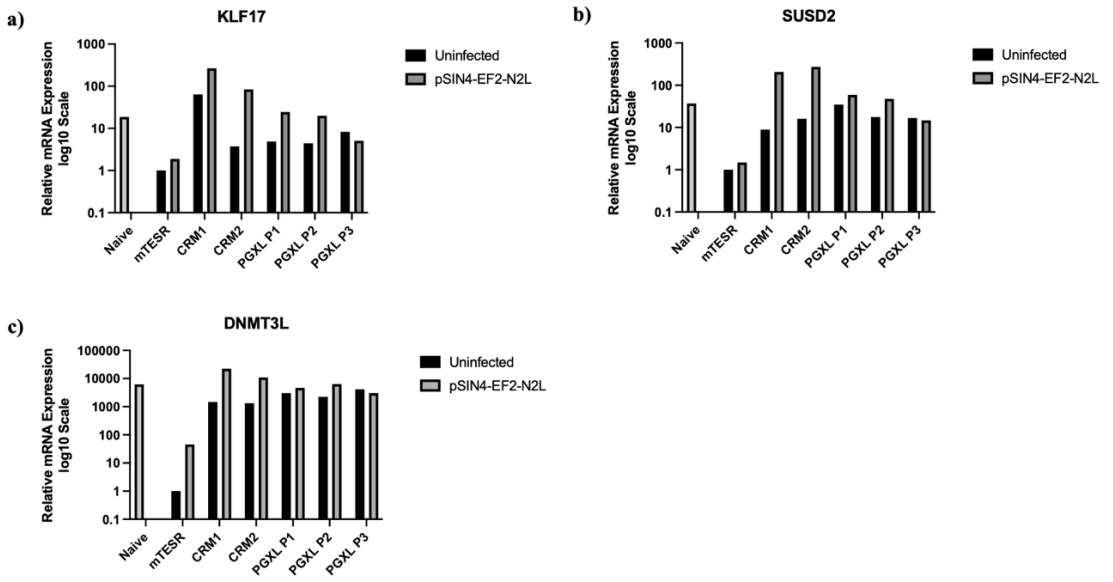


Figure 3. 31 qPCR analysis of naive marker expression during naive conversion after NANOG/LIN28A overexpression. Expression of KLF17, SUSD2, and DNMT3L during naive conversion of NANOG/LIN28A overexpressed and uninfected cells.

3.9 Investigating NANOG and KLF17 Overexpression Effect on Reprogramming Efficiency

KLF17 is a naive state-specific gene that is expressed in naive but not primed stem cells. KLF17 promotes the naive gene regulatory network by activating key naive transcription factors (Bayerl et al., 2021). In addition, NANOG is a core transcription factor required for sustaining pluripotency features for both primed and naive stem cells (Boyer et al., 2005). As a first step, I generated viruses from the existing NANOG (Addgene #16578) and KLF17 (Addgene #120503) overexpression vectors. I then transduced dH1f cells with the NANOG and KLF17 lentiviruses separately. After 72 hours, I isolated RNA from the transduced dH1f cells and analyzed NANOG and KLF17 overexpression by quantitative PCR (qPCR) relative to control cells. The qPCR results

validated significant overexpression of NANOG and KLF17 from their respective vectors (Figure 3.32).

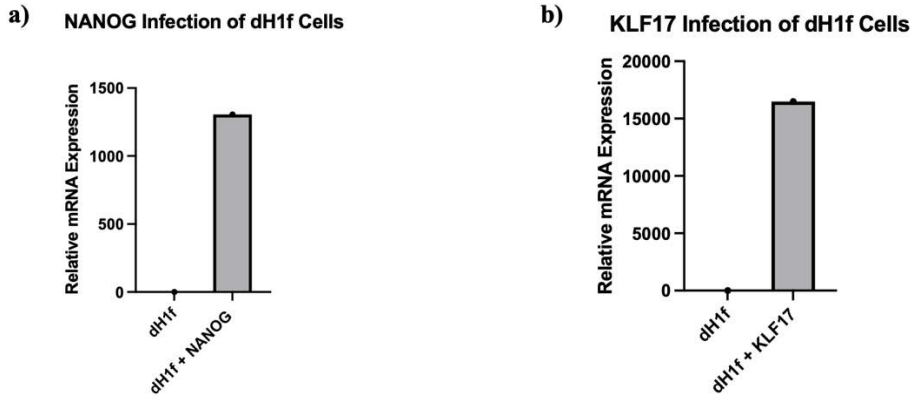


Figure 3. 32 Relative mRNA expression levels in dH1f cells after lentiviral infection. Quantitative PCR analysis of NANOG (a) and KLF17 (b) mRNA levels.

I conducted reprogramming experiments in dH1f cells following the previously mentioned protocol while simultaneously introducing NANOG and KLF17 through co-infection. At day 6 post-transduction, the cells were split at a 1:8 ratio onto MEF feeder layers for the PGXL and HENSM conditions to induce naive pluripotency. For E7 condition, Cells were seeded onto vitronectin-coated plates. Cells under PGXL and HENSM were fixed and analyzed on day 14, while cells in E7 were fixed and analyzed on day 18 (Figure 3.33). To assess reprogramming efficiency, I quantified Tra-1-60 positive colonies. By comparing the number of Tra-1-60 colonies in the NANOG and KLF17 overexpression conditions versus control (only OSKM), I evaluated the impact of their overexpression on primed and naive reprogramming.

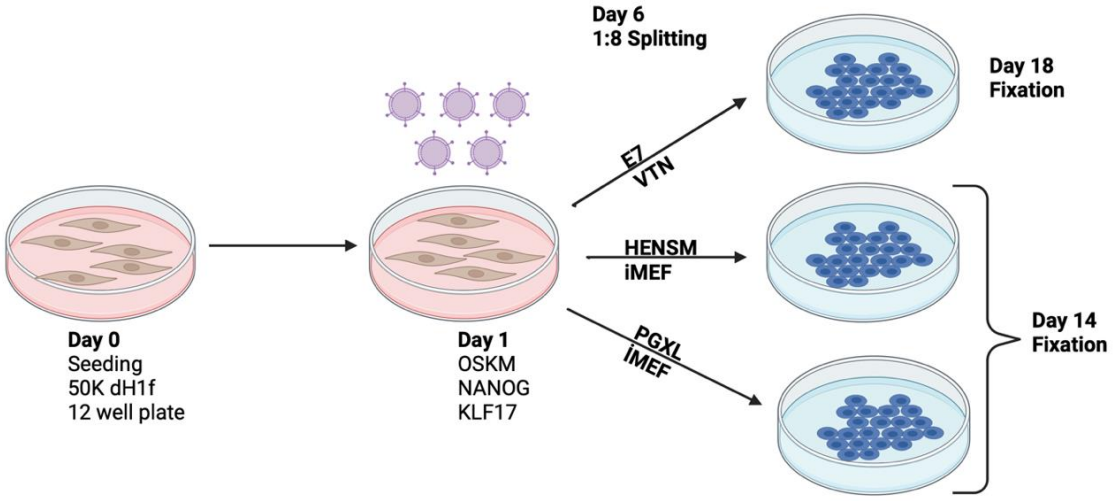


Figure 3. 33 Experimental workflow for reprogramming dH1f cells with OSKM and NANOG/KLF17 overexpression.

Under E7 and PGXL conditions, NANOG overexpression increased the reprogramming efficiency to both naive and primed pluripotent states compared to OSKM alone. This indicates that adding Nanog on top of OSKM enhances reprogramming under these conditions. KLF17 overexpression led to a decrease in reprogramming efficiency compared to OSKM alone, although this decrease was not statistically significant. This suggests that KLF17 does not enhance reprogramming under E7 and PGXL conditions and may even have slightly inhibitory effects. However, under HENSM condition, Neither NANOG nor KLF17 overexpression impacted reprogramming efficiency compared to OSKM alone. Therefore, neither transcription factor enhanced reprogramming under HENSM conditions (Figure 3.34). The statistical significance of variations in reprogramming efficiency between the NANOG and KLF17 overexpression (OE) conditions and the control with OSKM alone was evaluated through two-tailed t-tests. Each specific condition was compared separately to the OSKM alone control. In conclusion, the effects of overexpressing NANOG and KLF17 are dependent

on the culture conditions. NANOG increased Tra-1-60 positive colonies under E7 and PGXL conditions, while KLF17 did not. Meanwhile, both factors were ineffective at improving reprogramming under HENSM conditions.

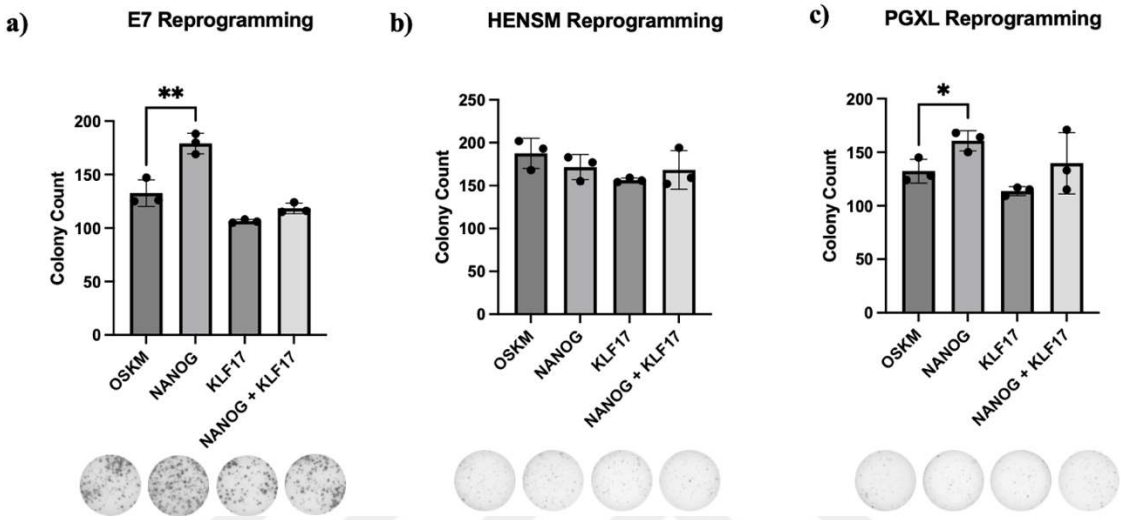


Figure 3. 34 Number of Tra-1-60 positive colonies of reprogrammed dH1f cells with NANOG and KLF17 overexpression. (a) Tra-1-60 staining under primed condition (E7 media). (b) Tra-1-60 staining under naive condition (HENSM media). (c) Tra-1-60 staining under naive condition (PGXL media). The bar graphs display the average values from three biological replicates, and the error bars depict the standard deviation. Corresponding images of individual wells for each condition are provided below the graph. The two-tailed Student's t-test was the specific statistical test used to compute p-values. * Denotes $p<0.05$, ** denotes $p<0.005$.

3.10 Assessing the Effect of KLF5, LIN28A, and KLF17 Knockouts on Naive Reprogramming

KLF5 and LIN28A are important transcription factors known to help maintain pluripotency in embryonic stem cells (ESCs) (Ema et al., 2008; Zhang et al., 2016).

KLF17 also plays a key role in regulating the naive pluripotent state in ESCs (Bayerl et al., 2021). To better understand the functions of LIN28A, KLF5, and KLF17 in establishing naive pluripotency, we attempted to generate knockout of these genes in fibroblasts (dH1f) and then assess reprogramming efficiency. To generate the knockouts, single guide RNAs (sgRNAs) targeting LIN28A, KLF5, and KLF17 were designed and cloned into the LentiCRISPR v2 vector. Sanger sequencing was performed to validate the successful cloning of the sgRNAs into the vector. Correctly cloned sgRNA-LentiCRISPR v2 vectors were then packaged into lentiviral particles using HEK293T cells. Next, dH1f cells were infected with the lentiviral particles containing the sgRNAs and selected for stably infected cells using puromycin. This generated populations of LIN28A, KLF5, and KLF17 knockout dH1f cells (Figure 3.35). To validate the knockout of the target genes, genomic DNA was extracted from the infected cells, and T7 assays were performed. (Figure 3.36). The T7 assay results revealed that the sgRNAs targeting KLF5 (guides KLF5 2_1 and KLF5 2_3) and LIN28A (guide LIN28A 2_2) did not show expected cleavage bands indicating indel formation. The lack of cleaved bands for these three guides suggests they were ineffective at inducing insertions or deletions at the target sites in the genomic DNA. This could be due to the poor specificity of these guides, resulting in inefficient Cas9 cutting. However, all the remaining sgRNAs were observed to be functional in generating DNA mutations.

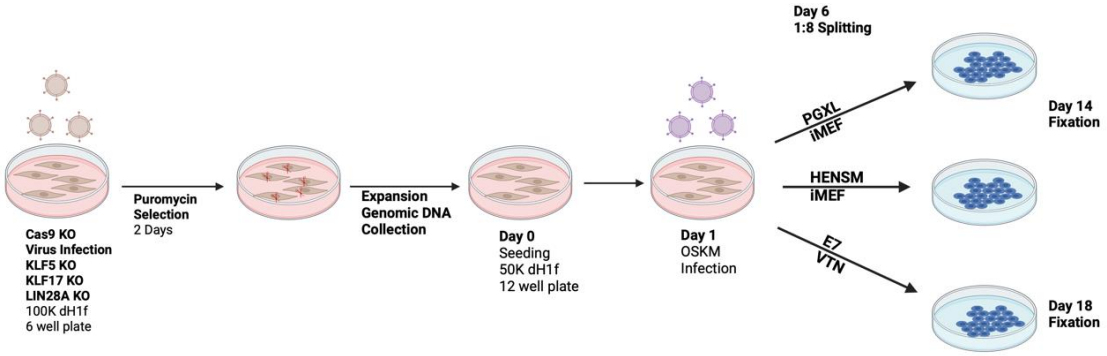


Figure 3. 35 Experimental workflow for generating KLF5, KLF17, and LIN28A knockout dH1f cells and reprogramming knockout dH1f cells with OSKM.

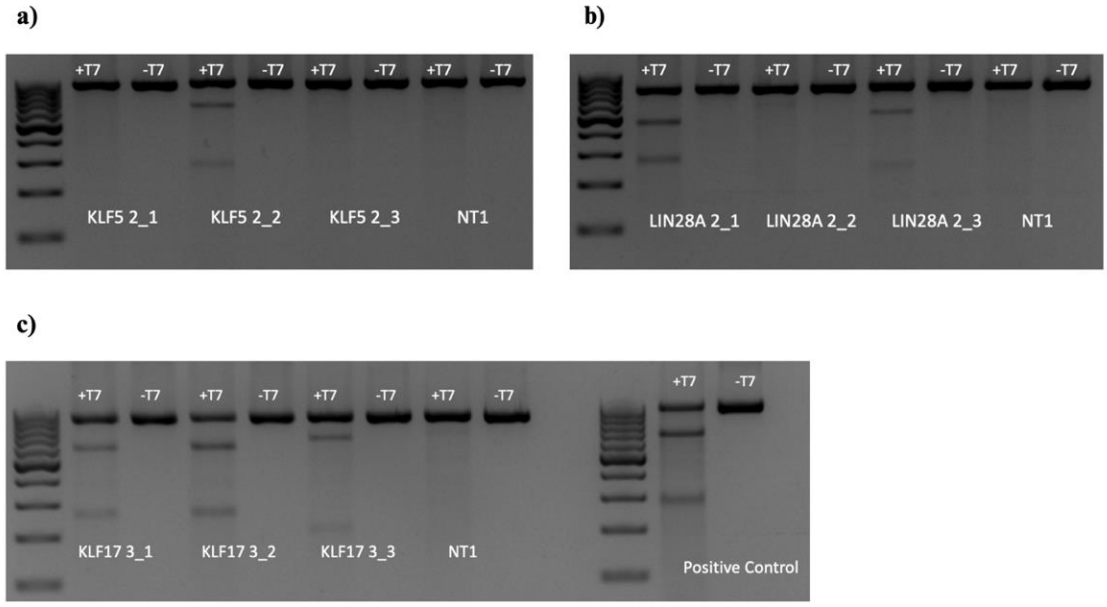


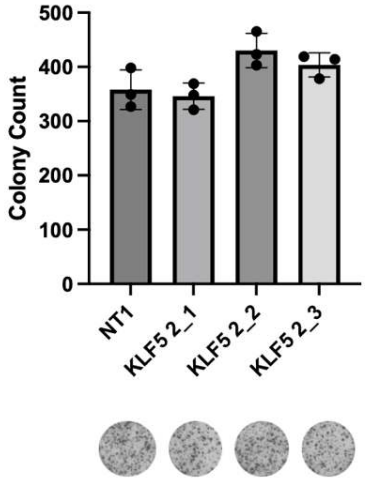
Figure 3. 36 T7 assay results for Knockout dH1f cells. T7 assay was performed to validate knockout of KLF5 (a), LIN28A (b), and KLF17 (c) in dH1f cells. Cleaved bands indicate successful Cas9-mediated cutting at the desired locus.

To assess the roles of KLF5, KLF17, and LIN28A in reprogramming to naive pluripotency, the knockout dH1F cells were reprogrammed by using the OSKM reprogramming factors (OCT4, SOX2, KLF4, c-MYC) which were delivered via

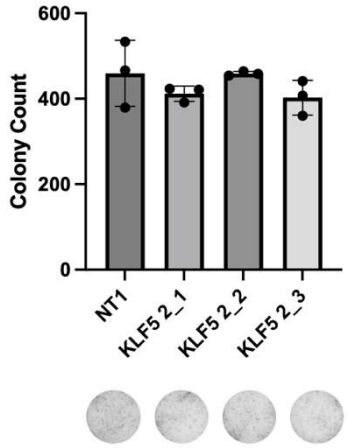
lentiviral infection. After infection, cells were cultured in fibroblast media for 6 days. On day 6 post-infection, the cells were split into 3 conditions: E7 media (Primed), HENSM media(naïve), and PGXL media (naïve) at a 1:8 ratio. Cells in HENSM and PGXL conditions were cultured until day 14 before fixation. Cells in E7 media were cultured until day 16 before fixation. Tra-1-60 staining was performed on the fixed cells to quantify reprogramming efficiency. The experiments were conducted in biological triplicate for each condition to control for variability.

The Tra-1-60 staining results revealed no significant differences in reprogramming efficiency between the KLF5 knockout lines and NT1-infected dH1f cells under the E7, HENSM, or PGXL conditions. Specifically, the KLF5 knockouts generated using guides KLF5 2_1 and KLF5 2_3 showed similar Tra-1-60 induction to control dH1f cells. This was expected, given that the T7 assay showed no evidence of successful cutting and knockout with these two guides. However, even the KLF5 2_2 knockouts made using guides that cut efficiently (shown by T7) did not exhibit a meaningful effect on reprogramming. This suggests that the loss of KLF5 does not affect the conversion of somatic cells into a pluripotent state. KLF5 does not appear to be required for establishing naive or primed pluripotency (Figure 3.37).

a) **KLF5 Knockout Effect on E7 Reprogramming**



b) **KLF5 Knockout Effect on HENSM Reprogramming**



c) **KLF5 Knockout Effect on PGXL Reprogramming**

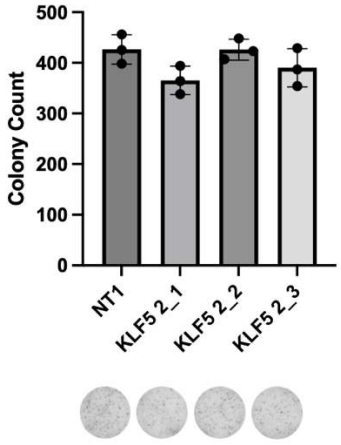
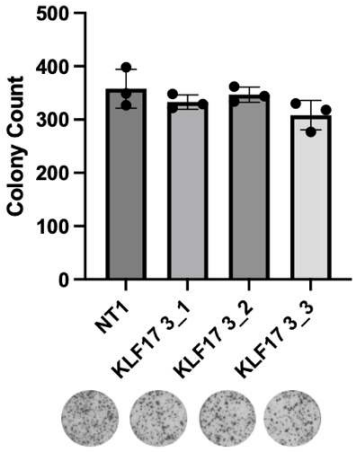


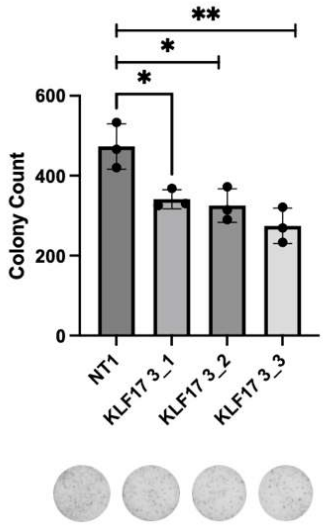
Figure 3. 37 Number of Tra-1-60 positive colonies of KLF5 knockout dH1f cells after reprogramming. (a) Tra-1-60 staining under primed condition (E7 media). (b) Tra-1-60 staining under naive condition (HENSM media). (c) Tra-1-60 staining under naive condition (PGXL media). The bar graphs display the average values from three biological replicates, and the error bars depict the standard deviation. Corresponding images of individual wells for each condition are provided below the graph.

Targeting of KLF17 resulted in no significant differences in reprogramming efficiency under primed cultured conditions using E7 media. However, under the HENSM and PGXL conditions for naive reprogramming, the KLF17 knockouts exhibited reduced Tra-1-60 positive colony numbers compared to controls. These findings indicate that while loss of KLF17 does not affect the induction of a primed pluripotent state, it impairs the reprogramming to naive pluripotency. The requirement for KLF17 appears specific to establishing the naive gene regulatory network. The KLF17 knockout cells showed an even more significant decrease in TRA-1-60 positive colonies compared to NT1 cells when reprogrammed under the HENSM naive conditions (Figure 3.38).

a) KLF17 Knockout Effect on E7 Reprogramming



b) KLF17 Knockout Effect on HENSM Reprogramming



c) KLF17 Knockout Effect on PGXL Reprogramming

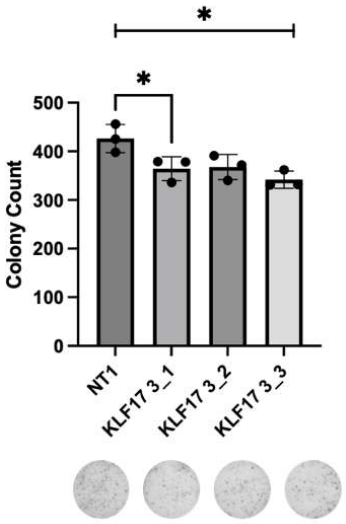


Figure 3. 38 Number of Tra-1-60 positive colonies of KLF17 knockout dH1f cells after reprogramming. (a) Tra-1-60 staining under primed condition (E7 media). (b) Tra-1-60 staining under naive condition (HENSM media). (c) Tra-1-60 staining under naive condition (PGXL media). The bar graphs display the average values from three biological replicates, and the error bars depict the standard deviation. Corresponding images of individual wells for each condition are provided below the graph. The two-tailed Student's t-test was the specific statistical test used to compute p-values. * Denotes $p<0.05$, ** denotes $p<0.005$.

In addition to TRA-1-60 staining, I also performed co-immunostaining for TRA-1-60 and KLF17 in the HENSM and PGXL reprogramming conditions using the KLF17 knockout lines. The quantification of colony numbers showed a reduced number of colonies in comparison to NT1 cells in both HENSM and PGXL culture conditions. This confirms impaired naive reprogramming in the absence of KLF17 (Figure 3.39). Interestingly, not all TRA-1-60 positive colonies were also KLF17 positive, even in wildtype NT1 cells under HENSM and PGXL naive conditions. This indicates that KLF17 expression is not absolutely required for the emergence of some TRA-1-60 positive naive colonies. Additionally, representative immunofluorescence images under primed (E7) and naive (HENSM, PGXL) conditions are shown in Figure 3.40. These images visualize the TRA-1-60 and KLF17 staining patterns for reprogramming cells. In summary, while KLF17 is not essential for the formation of any TRA-1-60+ naive colonies, it decreases naive reprogramming efficiency.

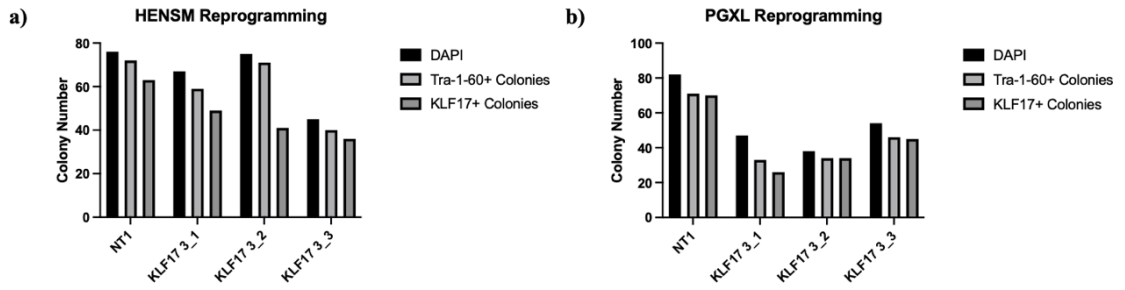


Figure 3. 39 Quantification of TRA-1-60 and KLF17 positive colonies after reprogramming of KLF17 knockout dH1f cells under naive conditions. Immunofluorescence images (a) HENSM and (b) PGXL reprogramming were analyzed with Image J.

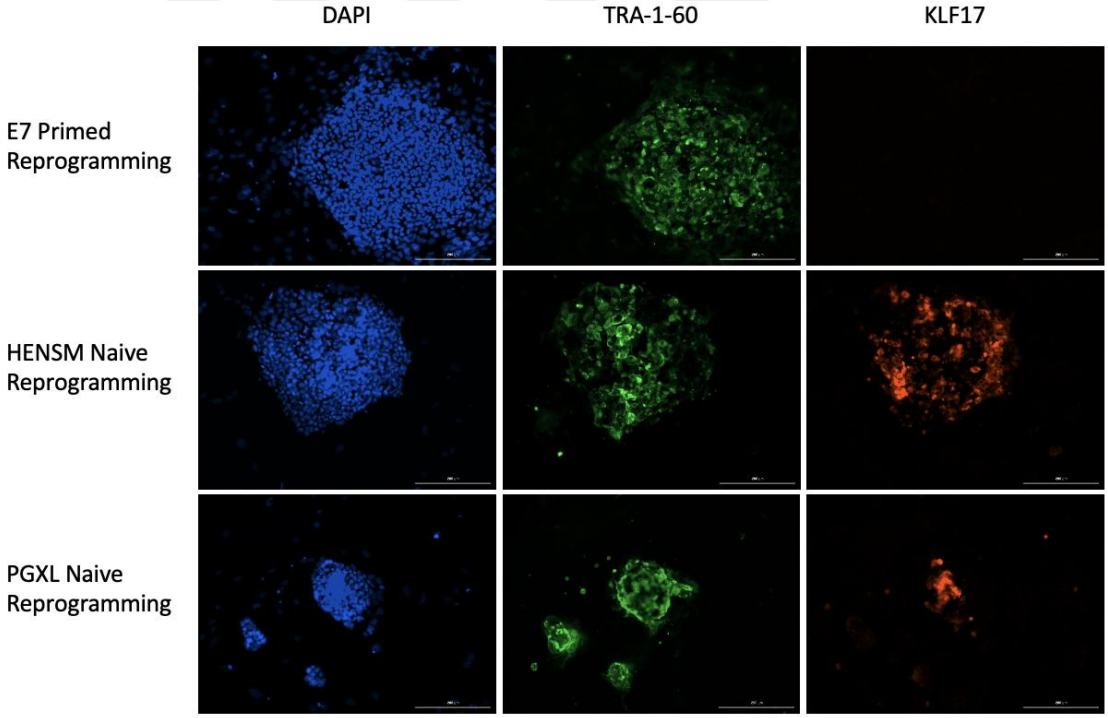
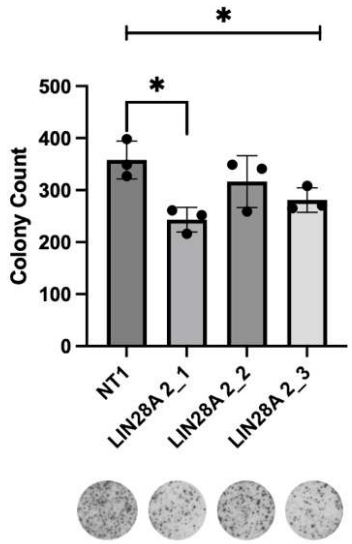


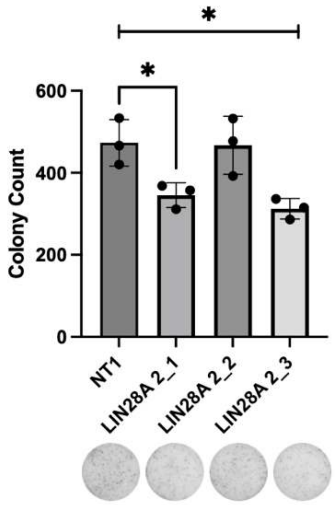
Figure 3. 40 Representative immunofluorescence images. Immunofluorescent images show TRA-1-60 (green) and KLF17 (red) staining after reprogramming of dH1f cells with E7, HENSM, and PGXL. Nuclei counterstained with DAPI (blue).

In LIN28A knockout dH1f cells, reprogramming results revealed decreased staining compared to controls under all conditions tested - E7, HENSM, and PGXL. The reduction in Tra-1-60 positive colonies was more explicit in the LIN28A knockouts generated using guides LIN28A 2_1 and LIN28A 2_3, which showed clear Cas9-mediated cutting in the T7 assay, confirming efficient knockout (Figure 3.41). This suggests that loss of LIN28A expression may impair the transcription factor network changes and epigenetic remodeling needed to establish both primed and naive pluripotency. This indicates that loss of LIN28A has a negative effect on the induction of both primed and naive pluripotent states.

a) **LIN28A Knockout Effect on E7 Reprogramming**



b) **LIN28A Knockout Effect on HENSM Reprogramming**



c) **LIN28A Knockout Effect on PGXL Reprogramming**

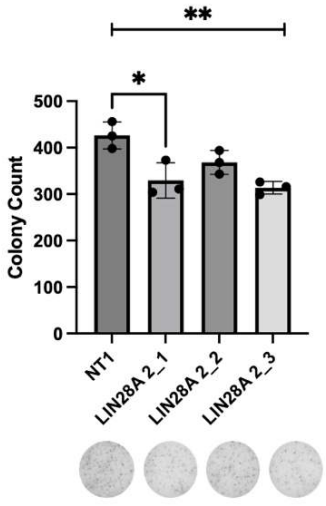


Figure 3. 41 Number of Tra-1-60 positive colonies of LIN28A knockout dH1f cells after reprogramming. (a) Tra-1-60 staining under primed condition (E7 media). (b) Tra-1-60 staining under naive condition (HENSM media). (c) Tra-1-60 staining under naive condition (PGXL media). The bar graphs display the average values from three biological replicates, and the error bars depict the standard deviation. Corresponding images of individual wells for each condition are provided below the graph. The two-tailed Student's t-test was the specific statistical test used to compute p-values. * Denotes $p<0.05$, ** denotes $p<0.005$.

3.11 Enhancing Naive Reprogramming Efficiency with Inhibitors of Epigenetic Regulators

To evaluate the effects of epigenetic inhibitors on naive and primed reprogramming, human fibroblasts were reprogrammed using OSKM while treated with various epigenetic modulators separately. The inhibitors, as outlined in Table 3.3, were administered at the beginning of the reprogramming on day 1 and refreshed every 2 days during media changes. On day 6, cells were split at a 1:8 ratio. On day 7, culture conditions were switched to E7, HENSM, or PGXL media. Inhibitors were continually administered until day 14. Cells in PGXL and HENSM conditions were fixed and analyzed on day 14. For E7 primed reprogramming, cells were maintained through day 18 before fixing and staining. TRA-1-60 staining was used to quantify reprogramming efficiency under E7, HENSM, or PGXL conditions (Figure 3.42).

Table 3. 3 List of inhibitors used during reprogramming with their functions.

Inhibitor	Function
A485	CBP/p300 (histone acetyltransferase) catalytic inhibitor
iCBP112	CBP/p300 bromodomain inhibitor
iCBP30	CBP/p300 bromodomain inhibitor
iBRD9	BRD9 (Bromodomain-containing protein 9) inhibitor
dBRD9	BRD9 degrader
RN1	LSD1 (Lysine-specific demethylase 1) inhibitor
A83-01	Activin/NODAL/TGF-β receptor inhibitor
EPZ5676	DOT1L (histone methyltransferase) inhibitor
EPZ719-2	SETD2 (histone methyltransferase) inhibitor

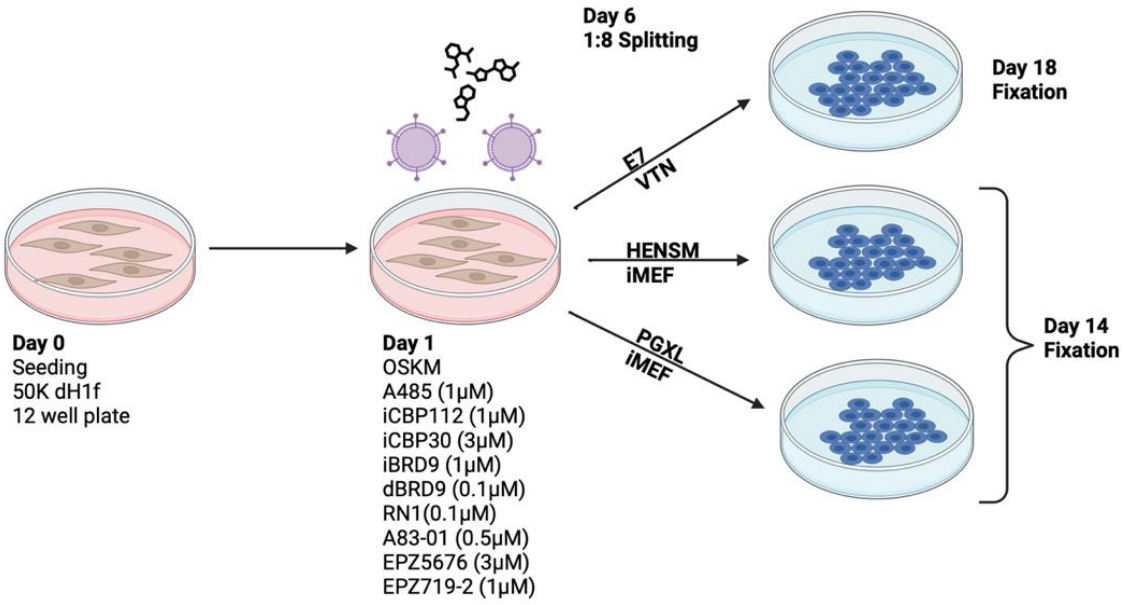
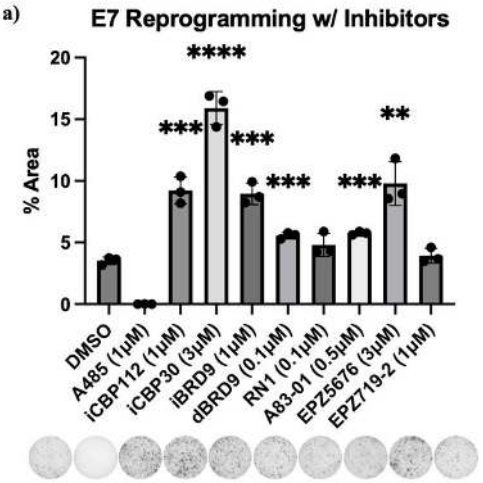


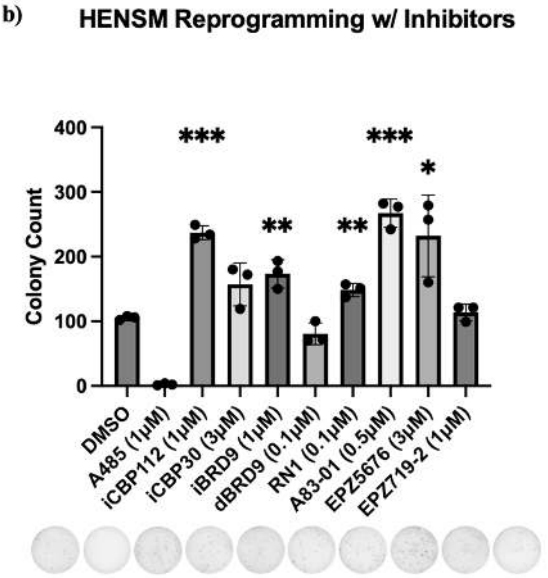
Figure 3. 42 Experimental workflow for reprogramming dH1f cells with OSKM and epigenetic inhibitors.

It was found that the inhibitors iCBP112, iBRD9, RN1, A83-01, and EPZ5676 all significantly increased the reprogramming efficiency of both the primed (E7) and naive (HENSM, PGXL) reprogramming compared to OSKM alone. Although these epigenetic inhibitors enhanced reprogramming under all conditions, the degree of enhancement varied. Under the E7 and HENSM conditions, iCBP112 and A83-01 caused the largest increases in TRA-1-60 positive colonies. In the PGXL naive reprogramming, iCBP112 and EPZ5676 led to the most pronounced enhancements in reprogramming efficiency (Figure 3.43). The statistical significance of differences in reprogramming efficiency between inhibitor conditions and the DMSO control was assessed using two-tailed t-tests. Each inhibitor condition was compared independently to the DMSO vehicle control. In summary, inhibition of certain epigenetic regulators improved both primed and naive reprogramming when combined with OSKM, but the extent of enhancement depended on the specific media conditions.

a)


E7 Reprogramming w/ Inhibitors
% Area
20
15
10
5
0
****
***
***
***
***
**
DMSO
A485 (1μM)
iCBP112 (1μM)
iCBP30 (3μM)
iBRD9 (1μM)
dBRD9 (0.1μM)
RN1 (0.1μM)
A83-01 (0.5μM)
EPZ5676 (3μM)
EPZ719-2 (1μM)

b)


HENSM Reprogramming w/ Inhibitors
Colony Count
400
300
200
100
0
***
**
**
***
*
DMSO
A485 (1μM)
iCBP112 (1μM)
iCBP30 (3μM)
iBRD9 (1μM)
dBRD9 (0.1μM)
RN1 (0.1μM)
A83-01 (0.5μM)
EPZ5676 (3μM)
EPZ719-2 (1μM)

c)

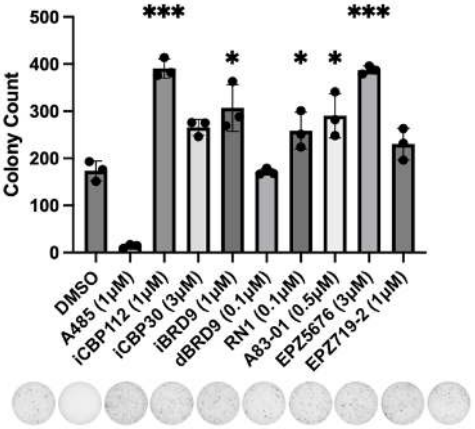

PGXL Reprogramming w/ Inhibitors
Colony Count
500
400
300
200
100
0
***
*
*
*
***
DMSO
A485 (1μM)
iCBP112 (1μM)
iCBP30 (3μM)
iBRD9 (1μM)
dBRD9 (0.1μM)
RN1 (0.1μM)
A83-01 (0.5μM)
EPZ5676 (3μM)
EPZ719-2 (1μM)

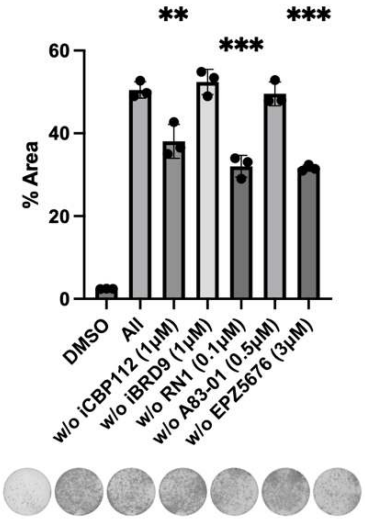
Figure 3. 43 Number of Tra-1-60 positive colonies dH1f cells after reprogramming with inhibitors. (a) Tra-1-60 staining under primed condition (E7 media). (b) Tra-1-60 staining under naive condition (HENSM media). (c) Tra-1-60 staining under naive condition (PGXL media). The bar graphs display the average values from three biological replicates, and the error bars depict the standard deviation. Corresponding images of individual wells for each condition are provided below the graph. * Denotes $p<0.05$, ** denotes $p<0.005$, *** denotes $p<0.0005$.

Since the inhibitors iCBP112, iBRD9, RN1, A83-01, and EPZ5676 were all found to individually enhance the reprogramming efficiency to both primed and naive states compared to OSKM alone, the next step was to evaluate combinations of these inhibitors. The aim was to determine if using these inhibitors together could show additive or synergistic effects to further increase reprogramming efficiency. To assess the contribution of each inhibitor to the reprogramming enhancement, a cocktail containing all 5 inhibitors was tested, and each inhibitor was individually removed from the full cocktail to determine the particular epigenetic inhibitor impact on the overall reprogramming efficiency under primed and naive conditions, and same experimental workflow was followed for reprogramming.

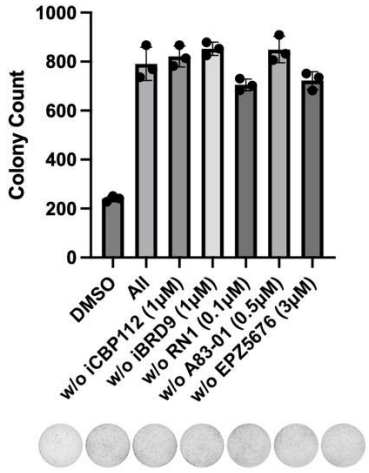
Under E7 primed conditions, removing iCBP112, RN1, or A83-01 from the full inhibitor cocktail led to significant decreases in reprogramming efficiency compared to the complete inhibitor set. For PGXL naive reprogramming, removal of RN1, A83-01, or EPZ5676 significantly reduced reprogramming efficiency versus the full cocktail, indicating these inhibitors cooperate to increase naive induction. In the context of HENSM naive conditions, the removal of RN1 or EPZ5676 led to a slight reduction in reprogramming efficiency when compared to the use of the complete inhibitor set.

However, this decrease was not statistically significant. This suggests that these two inhibitors have relatively minor additional effects under HENSM conditions. (Figure 3.44). Two-tailed t-tests were used to determine if differences in reprogramming efficiency between inhibitor combinations and the full cocktail were statistically significant. Each inhibitor combination was independently compared to the all-inhibitor control using separate t-tests. In summary, iCBP112, RN1, and A83-01 have functions in enhancing primed reprogramming, while RN1, A83-01, and EPZ5676 improve naive induction under PGXL conditions.

a) **E7 Reprogramming w/ Combination of Inhibitors**



b) **HENSM Reprogramming w/ Combination of Inhibitors**



c) **PGXL Reprogramming w/ Combination of Inhibitors**

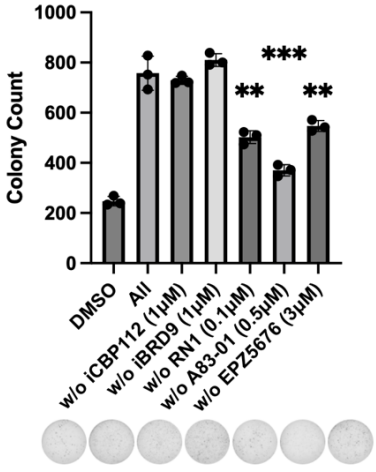


Figure 3. 44 Tra-1-60 staining of dH1f cells after reprogramming with combination of inhibitors. (a) % Area of Tra-1-60 positive colonies under primed condition (E7 media). (b) Number of Tra-1-60 positive colonies under naive condition (HENSM media). (c) Number of Tra-1-60 positive colonies under naive condition (PGXL media). The bar graphs display the average values from three biological replicates, and the error bars depict the standard deviation. Corresponding images of individual wells for each condition are provided below the graph. * Denotes $p<0.05$, ** denotes $p<0.005$, *** denotes $p<0.0005$.

Chapter 4:DISCUSSION

In this thesis, I focused on generating a reporter to isolate human naive stem cells and uncovering genetic and epigenetic modulators that may impede or enhance the establishment of naive pluripotency. A KLF17-T2A-EGFP fluorescent reporter was established in human iPSCs using CRISPR-Cas9n to enable specific labeling and isolation of naive cells from heterogeneous cultures. The construction of the reporter was validated by Sanger sequencing and tested during naive conversion stages. Overexpression of transcription factors KLF17 and NANOG/LIN28A during naive resetting increased naive marker activation in the first stages versus controls, suggesting these factors may improve induction. Additionally, I tested the effects of overexpressing KLF17/NANOG on both primed and reprogramming and found that NANOG increased naive and primed reprogramming efficiency. I also assessed the effect of knocking out KLF5, KLF17, and LIN28A. While KLF5 knockout had no effect, KLF17 knockout reduced only naive reprogramming. LIN28A knockout decreased both naive and primed reprogramming efficiency. Finally, inhibiting certain epigenetic regulators was found to increase reprogramming into both E7 primed programming and HENSM/PGXL naive reprogramming. Using iCBP112, iBRD9, RN1, A83-01, and EPZ5676 significantly increased primed and naive reprogramming.

4.1 Preventing Excess Cas9 Activity Following Desired On-Target Genome Editing

In the initial attempt to generate the reporter construct, sanger sequencing revealed all initial clones contained mutations at the 5' integration junction, indicating persistent Cas9 nickase activity after clonal expansion led to off-target effects. To prevent these undesirable mutations, the Cas9n function needed to be eliminated following the on-target integration event. Several approaches could potentially be utilized to address this issue.

One strategy is introducing point mutations to the PAM site of the integrated donor construct, and it can be altered such that Cas9 no longer recognizes it after the desired edit occurs (Paquet et al., 2016). This is accomplished by designing the homology-directed repair donor template to contain mutations in the PAM sequence within the homology arms. Once the donor integrates into the genome correctly, the target locus will harbor a modified PAM sequence. Since the genomic locus now contains an altered PAM after integration of the donor template, any remaining Cas9 will no longer recognize and re-cleave the target site. This prevents further on-target mutagenesis. Another potential strategy is using small molecule compounds that inhibit Cas9 activity through allosteric regulation without needing to genetically alter Cas9 itself (Maji et al., 2019). It was shown that some molecules bind to a regulatory allosteric site on Cas9, separate from its catalytic domain. These compounds induce a conformational change in Cas9 that reconfigures key active site residues into an inactive arrangement unable to cleave DNA. This enables reversible chemical control of Cas9 function by simply adding the small molecule inhibitors. The approaches of mutating the PAM sequence and using small molecule allosteric regulators can potentially both be implemented to inhibit off-target effects by preventing continued Cas9 activity after the desired on-target modification has been achieved.

4.2 Strategies for GFP Signal Detection in CRISPR-Engineered Cells

The KLF17-T-2A-EGFP reporter failed to exhibit bright GFP fluorescence during naive conversion when examined by microscopy and flow cytometry. However, qPCR analysis using primers spanning the KLF17-EGFP junction demonstrated that the EGFP transcript was expressed and upregulated, especially during the initial stages, as expected based on KLF17 induction. While qPCR analysis verified successful transcription of the

KLF17-EGFP reporter mRNA, several factors in post-transcriptional processing may have prevented efficient translation into functional GFP protein. Intact mRNA production alone does not guarantee successful translation if mechanisms like mRNA splicing, nuclear export, localization, or stability are disrupted. To investigate whether mRNA processing or transport defects impede the protein production, further analysis of both transcript and protein levels is needed. Performing targeted western blots using specific antibodies against KLF17 and EGFP could help determine if translation is being disrupted. Additionally, immunofluorescence staining of the cells that have been converted to a naive state using antibodies against KLF17 and EGFP can be performed to determine if there is a problem with protein production. Examining both mRNA and protein expression will help to identify the issue responsible for the lack of intense GFP fluorescence.

During the process of converting primed cells into a naive state and while sustaining naive stem cells in culture, the compound Gö6983 is used. Gö6983 is a small molecule inhibitor of protein kinase C (PKC) that is often used during naive iPSC culture to help maintain the naive state (Theunissen et al., 2016). However, Gö6983 exhibits autofluorescence when excited by wavelengths used to detect GFP. Since the KLF17 reporter expresses low levels of GFP, the autofluorescence background from Gö6983 treatment could obscure or interfere with any GFP signal. The low KLF17 expression may be insufficient to produce enough GFP fluorescence above the Gö6983 background. To address this, an alternative PKC inhibitor without autofluorescence could be used. Removing the auto-fluorescent compound would clarify any specific GFP fluorescence. In addition to changing the PKC inhibitor, the reporter could be redesigned to use a more strongly expressed naive marker instead of KLF17. Pastor et al. identified the transcription factors DPPA3 and TFCP2L1 as being more highly expressed in naive

human PSCs compared to primed, making them more suitable candidates to replace KLF17 (Pastor et al., 2016). Moreover, surface markers that are upregulated in the naive state could be utilized to design a reporter system. Trusler et al. performed transcriptome analysis comparing naive and primed human PSCs and found a set of surface markers enriched in naive cells (Trusler et al., 2018). These included CD7, CD75, and CD77. Flow cytometry confirmed their elevated protein levels on naive vs. primed. Therefore, targeting any of these surface proteins could produce a reporter that labels the plasma membrane of naive cells. An advantage of using surface markers compared to transcription factors is that surface reporters do not require events downstream of transcription like splicing, nuclear export, and translation. Surface proteins are directly visible on the membrane immediately after expression. In summary, the Gö6983 green autofluorescence may obscure the detection of the weak GFP reporter signal. Switching to a non-fluorescent PKC inhibitor and/or using a more strongly expressed naive marker could help isolate specific reporter fluorescence.

4.3 Enhancing Naive Conversion through NANOG and KLF17 Overexpression

Overexpression of NANOG and KLF17 during naive conversion results in increased expression of the naive markers SUDS2 and DNMT3L compared to control cells, potentially promoting naive state conversion. NANOG is a core transcription factor that establishes and sustains naive pluripotency, as demonstrated in mouse ESCs. Reducing NANOG caused mouse ESCs in ground state 2i/LIF conditions to convert to a primed-like state while restoring NANOG enabled reversion to the 2i/LIF naive state. This indicates that NANOG has an essential role in maintaining and inducing mouse ground state pluripotency (Silva et al., 2009). NANOG is a highly conserved central regulator of pluripotency. Findings in mouse ESCs indicate that NANOG controls entry

and maintenance of the naive state. Our data showing enhanced naive conversion with NANOG overexpression suggest it likely plays a similar role in governing human naive status. Therefore, modulating NANOG may induce the conversion to the naive state in human cells.

Similarly, KLF17 is a pro-naive transcription factor that primes the naive gene regulatory network (Wang et al., 2021). Overexpression of KLF17 in human PSCs has been shown to promote conversion to a naive state, indicating it can enhance resetting (Lea et al., 2021). Consistent with these published findings, I also found that KLF17 overexpression led to increased expression of other naive markers, such as SUSD2 and DNMT3L. Taken together, these data support KLF17 functioning as an important driver of human naive pluripotency, similar to NANOG. Both my results and previous studies demonstrate that forced expression of KLF17 can induce naive conversion in human PSCs. In summary, the established roles of NANOG and KLF17 as central naive transcription factors explain their capability to facilitate and induce resetting to a naive state when overexpressed.

4.4 Different Effects of NANOG and KLF17 Overexpression on Reprogramming

My results showed that overexpression of NANOG significantly increased the efficiency of both primed and naive induced pluripotent stem cell (iPSC) generation from somatic cells. Multiple reports have also demonstrated that NANOG is an inducer of pluripotency and somatic cell reprogramming. Yu et al. showed that exogenous expression of NANOG together with OCT4, SOX2, and KLF4 was sufficient to reprogram human fibroblasts into induced pluripotent stem cells. NANOG is one of the core factors capable of reverting somatic cells into pluripotency (Yu et al., 2007). Furthermore, Theunissen et al. found that NANOG alone could overcome barriers and

induce pluripotency through direct activation of the endogenous circuitry even without OCT4 and SOX2 (Theunissen et al., 2011). My experimental data validates prior reports that showed overexpression of NANOG increases cellular reprogramming.

My experimental results showed that KLF17 overexpression did not increase the efficiency of primed or naive iPSC generation, unlike NANOG overexpression. There could be several underlying factors that might explain this phenomenon. First, KLF17 regulates a restricted set of target genes in ESCs, primarily related to metabolism rather than core transcriptional regulators. This suggests that KLF17 does not directly activate endogenous pluripotency genes to the same extent as NANOG. Additionally, KLF17 appears to act later during the naive resetting process by reinforcing self-renewal and repressing residual primed identity. Therefore, KLF17 may not drive the initial induction of pluripotency like NANOG but rather stabilizes the naive state after initial conversion (Bayerl et al., 2021; Lea et al., 2021).

4.5 KLF5 and KLF17 Knockout Effect on Reprograming

Knockout of KLF5 did not have any significant effect on the efficiency of reprogramming under either naive or primed conditions. I did not observe any changes in reprogramming upon loss of KLF5. This lack of impact is likely due to redundancy between KLF family members. Specifically, KLF2, KLF4, and KLF5 have been shown to possess functional redundancy in their ability to activate core pluripotency network components (Jiang et al., 2008; Bourillot & Savatier, 2010). The overlapping capabilities of these KLF proteins result in compensation when one member is lost. Therefore, due to this redundancy, knockout of KLF5 did not significantly disrupt the induction of pluripotency during reprogramming.

My results indicated that KLF17 knockout selectively reduced the efficiency of naive reprogramming but did not impact primed reprogramming. The lack of effect on primed reprogramming was expected since KLF17 expression is enriched in the naive state. However, there are discrepancies in published studies regarding the effects of KLF17 loss on the induction of naive stem cells. Lea et al. reported that KLF17 knockout human iPSCs could still be reset into naive iPSCs with slight transcriptomic changes (Lea et al., 2021). This suggests KLF17 is not absolutely required. In contrast, Bayerl et al. found that KLF17 knockdown prevents naive resetting, indicating it is important for efficient conversion (Bayerl et al., 2021). The differences could potentially be explained by variations in knockout strategies, cell lines, or naive induction protocols. While the discussed studies examined the effects of KLF17 loss during resetting primed PSCs into naive iPSCs, there is currently no published data assessing the impact of KLF17 knockout, specifically during direct reprogramming of somatic cells into naive iPSCs. Therefore, further studies are needed to corroborate this finding in direct somatic cell reprogramming systems.

4.6 LIN28A Knockout Reduces Both Naive and Primed Cellular Reprogramming

Multiple studies have shown that LIN28A acts as a key factor that enhances somatic cell reprogramming by promoting the induction of pluripotency. Yu et al. demonstrated that LIN28A overexpression can enhance the generation of human iPSCs (Yu et al., 2007). Additionally, Darr and Benvenisty found that the knockdown of endogenous LIN28A reduces reprogramming efficiency, indicating its importance in acquiring pluripotency (Darr & Benvenisty, 2009). Moreover, Qiu et al. revealed that LIN28A regulates the protein levels of core pluripotency factors like OCT4, and the loss of LIN28A decreases OCT4 protein levels and hinders reprogramming (Qiu et al., 2010).

Taken together, these published findings show that LIN28A promotes the induction of pluripotency, which supports my results indicating that LIN28A knockout reduces both naive and primed reprogramming.

4.7 Improving Naive Reprogramming Efficiency with Small Epigenetic Inhibitors

Prior studies demonstrated that small molecule inhibitors, including EPZ5676, RN1, iCBP112, and iBRD9, can enhance the reprogramming of human fibroblasts into primed iPSCs (Onder et al., 2012; Cacchiarelli et al., 2015; Ebrahimi et al., 2019; Sevinç et al., 2022). These compounds are proposed to increase reprogramming efficiency by promoting chromatin accessibility at pluripotency genes or repressing the expression of somatic cell genes. My experimental results showed that these epigenetic inhibitors also increased the efficiency of naive reprogramming. The improved naive induction indicates these compounds likely facilitate the opening of naive-specific genes and networks to increase the activation of the naive pluripotent state from fibroblasts. In addition, it was previously shown that inhibiting TGFβ signaling using the small molecule SB431542 could induce NANOG expression, and TGFβ pathway inhibition increased the efficiency of reprogramming of human fibroblasts into iPSCs (Ichida et al., 2009). Similarly, I observed that treating cells with the TGFβ inhibitor A83-01 increased the efficiency of both primed and naive iPSC reprogramming. This aligns with and supports the prior report that blocking TGFβ signaling enhances iPSC generation through the upregulation of NANOG. In summary, the manipulation of epigenetic and transcriptional regulators using small molecules, which work together to activate pluripotency networks while suppressing somatic identity, seems to enhance the generation of both primed and naive iPSCs from human fibroblasts.

In conclusion, this thesis describes the generation and validation of a KLF17-T-2A-EGFP fluorescent reporter that enables specific isolation of human naive pluripotent stem cells from mixed cultures. I gained insights into genetic and epigenetic factors that may act as barriers or facilitators during the establishment of naive pluripotency. Overexpression of KLF17 and NANOG during naive resetting was found to enhance the activation of naive markers and conversion efficiency. Furthermore, NANOG overexpression increased both naive and primed reprogramming. Transcription factor knockouts experiment revealed LIN28A and KLF17 knockouts selectively impaired primed and naive reprogramming. Testing of epigenetic inhibitors identified compounds like iCBP112, iBRD9, RN1, A83-01, and EPZ5676 that preferentially increased primed or naive reprogramming. Overall, this work provided a valuable new tool for isolating human naive stem cells and elucidated genetic and epigenetic regulators governing the acquisition and interconversion of pluripotent states.

BIBLIOGRAPHY

Acampora, D., Di Giovannantonio, L. G., & Simeone, A. (2013). Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition. *Development*, *140*(1), 43–55. https://doi.org/10.1242/dev.085290

Adewumi, O., Aflatoonian, B., Ahrlund-Richter, L., Amit, M., Andrews, P.W., Beighton, G., Bello, P.A., Benvenisty, N., Berry, L.S., Bevan, S. and Blum, B. (2007). Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. Nature Biotechnology, 25(7), 803–816. https://doi.org/10.1038/nbt1318

Apostolou, E., & Hochedlinger, K. (2013). Chromatin dynamics during cellular reprogramming. *Nature*, *502*(7472), 462–471. https://doi.org/10.1038/nature12749

Atlasi, Y., & Stunnenberg, H. G. (2017). The interplay of epigenetic marks during stem cell differentiation and development. *Nature Reviews Genetics*, *18*(11), 643–658. https://doi.org/10.1038/nrg.2017.57

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jørgensen, H. F., John, R. M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M., & Fisher, A. G. (2006). Chromatin signatures of pluripotent cell lines. *Nature Cell Biology*, *8*(5), 532–538. https://doi.org/10.1038/ncb1403

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S. L., & Lander, E. S. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*, *125*(2), 315–326. https://doi.org/10.1016/j.cell.2006.02.041

Blakeley, P., Fogarty, N. M. E., del Valle, I., Wamaitha, S. E., Hu, T. X., Elder, K., Snell, P., Christie, L., Robson, P., & Niakan, K. K. (2015). Defining the three cell lineages

of the human blastocyst by single-cell RNA-seq. *Development*. https://doi.org/10.1242/dev.123547

Boroviak, T., Loos, R., Lombard, P., Okahara, J., Behr, R., Sasaki, E., Nichols, J., Smith, A., & Bertone, P. (2015). Lineage-specific profiling delineates the emergence and progression of naive pluripotency in mammalian embryogenesis. *Developmental Cell*, *35*(3), 366–382. https://doi.org/10.1016/j.devcel.2015.10.011

Bourillot, P.-Y., & Savatier, P. (2010). Krüppel-like transcription factors and control of pluripotency. *BMC Biology*, *8*(1). https://doi.org/10.1186/1741-7007-8-125

Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., Gifford, D. K., Melton, D. A., Jaenisch, R., & Young, R. A. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, *122*(6), 947–956. https://doi.org/10.1016/j.cell.2005.08.020

Bredenkamp, N., Yang, J., Clarke, J., Stirparo, G. G., von Meyenn, F., Dietmann, S., Baker, D., Drummond, R., Ren, Y., Li, D., Wu, C., Rostovskaya, M., Eminli-Meissner, S., Smith, A., & Guo, G. (2019). WNT inhibition facilitates RNA-mediated reprogramming of human somatic cells to naive pluripotency. *Stem Cell Reports*, *13*(6), 1083–1098. https://doi.org/10.1016/j.stemcr.2019.10.009

Brons, I. G., Smithers, L. E., Trotter, M. W., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S. M., Howlett, S. K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R. A., & Vallier, L. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature*, *448*(7150), 191–195. https://doi.org/10.1038/nature05950

Buecker, C., Srinivasan, R., Wu, Z., Calo, E., Acampora, D., Faial, T., Simeone, A., Tan, M., Swigut, T., & Wysocka, J. (2014). Reorganization of enhancer patterns in

transition from naive to primed pluripotency. *Cell Stem Cell*, *14*(6), 838–853. https://doi.org/10.1016/j.stem.2014.04.003

Cacchiarelli, D., Trapnell, C., Ziller, M. J., Soumillon, M., Cesana, M., Karnik, R., Donaghey, J., Smith, Z. D., Ratanasirintrawoot, S., Zhang, X., Ho Sui, S. J., Wu, Z., Akopian, V., Gifford, C. A., Doench, J., Rinn, J. L., Daley, G. Q., Meissner, A., Lander, E. S., & Mikkelsen, T. S. (2015). Integrative analyses of human reprogramming reveal dynamic nature of induced pluripotency. *Cell*, *162*(2), 412–424. https://doi.org/10.1016/j.cell.2015.06.016

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., & Smith, A. (2003). Functional expression cloning of NANOG, a pluripotency sustaining factor in embryonic stem cells. *Cell*, *113*(5), 643–655. https://doi.org/10.1016/s0092-8674(03)00392-1

Chan, Y.-S., Göke, J., Ng, J.-H., Lu, X., Gonzales, K. A. U., Tan, C.-P., Tng, W.-Q., Hong, Z.-Z., Lim, Y.-S., & Ng, H.-H. (2013). Induction of a human pluripotent state with distinct regulatory circuitry that resembles Preimplantation Epiblast. *Cell Stem Cell*, *13*(6), 663–675. https://doi.org/10.1016/j.stem.2013.11.015

Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. (2007a). *Nature Biotechnology*, *25*(7), 803–816. https://doi.org/10.1038/nbt1318

Collier, A. J., Panula, S. P., Schell, J. P., Chovanec, P., Plaza Reyes, A., Petropoulos, S., Corcoran, A. E., Walker, R., Douagi, I., Lanner, F., & Rugg-Gunn, P. J. (2017). Comprehensive cell surface protein profiling identifies specific markers of human naive and primed pluripotent states. *Cell Stem Cell*, *20*(6). https://doi.org/10.1016/j.stem.2017.02.014

Collier, A. J., & Rugg-Gunn, P. J. (2018). Identifying human naïve pluripotent stem cells − evaluating state-specific reporter lines and cell-surface markers. *BioEssays*, *40*(5), 1700239. https://doi.org/10.1002/bies.201700239

Darr, H., & Benvenisty, N. (2009). Genetic analysis of the role of the reprogramming gene LIN28 in human embryonic stem cells. *Stem Cells*, *27*(2), 352–362. https://doi.org/10.1634/stemcells.2008-0720

Duncan, A. W., Rattis, F. M., DiMascio, L. N., Congdon, K. L., Pazianos, G., Zhao, C., Yoon, K., Cook, J. M., Willert, K., Gaiano, N., & Reya, T. (2005). Integration of Notch and Wnt signaling in hematopoietic stem cell maintenance. *Nature Immunology*, *6*(3), 314–322. https://doi.org/10.1038/ni1164

Dunn, S.-J., Martello, G., Yordanov, B., Emmott, S., & Smith, A. G. (2014). Defining an essential transcription factor program for naïve pluripotency. *Science*, *344*(6188), 1156–1160. https://doi.org/10.1126/science.1248882

Ebrahimi, A., Sevinç, K., Gürhan Sevinç, 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*(5), 519–528. https://doi.org/10.1038/s41589-019-0264-z

Ema, M., Mori, D., Niwa, H., Hasegawa, Y., Yamanaka, Y., Hitoshi, S., Mimura, J., Kawabe, Y., Hosoya, T., Morita, M., Shimosato, D., Uchida, K., Suzuki, N., Yanagisawa, J., Sogawa, K., Rossant, J., Yamamoto, M., Takahashi, S., & Fujii-Kuriyama, Y. (2008). Krüppel-like factor 5 is essential for blastocyst development and the normal self-renewal of mouse escs. *Cell Stem Cell*, *3*(5), 555–567. https://doi.org/10.1016/j.stem.2008.09.003

Evans, M. J., & Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature*, *292*(5819), 154–156. https://doi.org/10.1038/292154a0

Festuccia, N., Osorno, R., Halbritter, F., Karwacki-Neisius, V., Navarro, P., Colby, D., Wong, F., Yates, A., Tomlinson, S. R., & Chambers, I. (2012). ESRRB is a direct NANOG target gene that can substitute for NANOG function in pluripotent cells. *Cell Stem Cell*, *11*(4), 477–490. https://doi.org/10.1016/j.stem.2012.08.002

Fujikura, J., Yamato, E., Yonemura, S., Hosoda, K., Masui, S., Nakao, K., Miyazaki, J., & Niwa, H. (2002). Differentiation of embryonic stem cells is induced by GATA factors. *Genes & Development*, *16*(7), 784–789. https://doi.org/10.1101/gad.968802

Gafni, O., Weinberger, L., Mansour, A. A., Manor, Y. S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A., Rais, Y., Shipony, Z., Mukamel, Z., Krupalnik, V., Zerbib, M., Geula, S., Caspi, I., Schneir, D., Shwartz, T., … Hanna, J. H. (2013). Derivation of novel human ground state naive pluripotent stem cells. *Nature*, *504*(7479), 282–286. https://doi.org/10.1038/nature12745

Gao, Y., Chen, J., Li, K., Wu, T., Huang, B., Liu, W., Kou, X., Zhang, Y., Huang, H., Jiang, Y., Yao, C., Liu, X., Lu, Z., Xu, Z., Kang, L., Chen, J., Wang, H., Cai, T., & Gao, S. (2013). Replacement of oct4 by tet1 during IPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. *Cell Stem Cell*, *12*(4), 453–469. https://doi.org/10.1016/j.stem.2013.02.005

Gao, X., Nowak-Imialek, M., Chen, X., Chen, D., Herrmann, D., Ruan, D., Chen, A. C., Eckersley-Maslin, M. A., Ahmad, S., Lee, Y. L., Kobayashi, T., Ryan, D., Zhong, J., Zhu, J., Wu, J., Lan, G., Petkov, S., Yang, J., Antunes, L., … Liu, P. (2019).

Establishment of porcine and human expanded potential stem cells. *Nature Cell Biology*, *21*(6), 687–699. https://doi.org/10.1038/s41556-019-0333-2

Goodwin, M., Lee, E., Lakshmanan, U., Shipp, S., Froessl, L., Barzaghi, F., Passerini, L., Narula, M., Sheikali, A., Lee, C. M., Bao, G., Bauer, C. S., Miller, H. K., Garcia-Lloret, M., Butte, M. J., Bertaina, A., Shah, A., Pavel-Dinu, M., Hendel, A., … Bacchetta, R. (2020). CRISPR-based gene editing enables FOXP3 gene repair in IPEX patient cells. *Science Advances*, *6*(19). https://doi.org/10.1126/sciadv.aaz0571

Greber, B., Wu, G., Bernemann, C., Joo, J. Y., Han, D. W., Ko, K., Tapia, N., Sabour, D., Sterneckert, J., Tesar, P., & Schöler, H. R. (2010). Conserved and divergent roles of FGF signaling in mouse epiblast stem cells and human embryonic stem cells. *Cell Stem Cell*, *6*(3), 215–226. https://doi.org/10.1016/j.stem.2010.01.003

Guo, G., Von Meyenn, F., Rostovskaya, M., Clarke, J., Dietmann, S., Baker, D., Sahakyan, A., Myers, S., Bertone, P., Reik, W., Plath, K., & Smith, A. (2017). Epigenetic resetting of human pluripotency. *Development*, *144*(15), 2748–2763. https://doi.org/10.1242/dev.146811

Guo, G., von Meyenn, F., Santos, F., Chen, Y., Reik, W., Bertone, P., Smith, A., & Nichols, J. (2016). Naive pluripotent stem cells derived directly from isolated cells of the human inner cell mass. *Stem Cell Reports*, *6*(4), 437–446. https://doi.org/10.1016/j.stemcr.2016.02.005

Gurdon, J. B., Elsdale, T. R., & Fischberg, M. (1958). Sexually mature individuals of xenopus laevis from the transplantation of single somatic nuclei. *Nature*, *182*(4627), 64–65. https://doi.org/10.1038/182064a0

Hackett, J. A., & Surani, M. A. (2014). Regulatory principles of pluripotency: From the ground state up. *Cell Stem Cell*, *15*(4), 416–430. https://doi.org/10.1016/j.stem.2014.09.015

Hall, J., Guo, G., Wray, J., Eyres, I., Nichols, J., Grotewold, L., Morfopoulou, S., Humphreys, P., Mansfield, W., Walker, R., Tomlinson, S., & Smith, A. (2009). Oct4 and Lif/STAT3 additively induce KRÜPPEL factors to sustain embryonic stem cell self-renewal. *Cell Stem Cell*, *5*(6), 597–609. https://doi.org/10.1016/j.stem.2009.11.003

Hamilton, W. B., Mosesson, Y., Monteiro, R. S., Emdal, K. B., Knudsen, T. E., Francavilla, C., Barkai, N., Olsen, J. V., & Brickman, J. M. (2019). Dynamic lineage priming is driven via direct enhancer regulation by ERK. *Nature*, *575*(7782), 355–360. https://doi.org/10.1038/s41586-019-1732-z

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

Hanna, J., Cheng, A. W., Saha, K., Kim, J., Lengner, C. J., Soldner, F., Cassady, J. P., Muffat, J., Carey, B. W., & Jaenisch, R. (2010). Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proceedings of the National Academy of Sciences*, *107*(20), 9222–9227. https://doi.org/10.1073/pnas.1004584107

Hawkins, R. D., Hon, G. C., Lee, L. K., Ngo, Q., Lister, R., Pelizzola, M., Edsall, L. E., Kuan, S., Luu, Y., Klugman, S., Antosiewicz-Bourget, J., Ye, Z., Espinoza, C., Agarwahl, S., Shen, L., Ruotti, V., Wang, W., Stewart, R., Thomson, J. A., … Ren, B. (2010). Distinct epigenomic landscapes of pluripotent and lineage-committed

human cells. *Cell Stem Cell*, *6*(5), 479–491. https://doi.org/10.1016/j.stem.2010.03.018

Hou, P., Li, Y., Zhang, X., Liu, C., Guan, J., Li, H., Zhao, T., Ye, J., Yang, W., Liu, K., Ge, J., Xu, J., Zhang, Q., Zhao, Y., & Deng, H. (2013). Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science*, *341*(6146), 651–654. https://doi.org/10.1126/science.1239278

Ichida, J. K., Blanchard, J., Lam, K., Son, E. Y., Chung, J. E., Egli, D., Loh, K. M., Carter, A. C., Di Giorgio, F. P., Koszka, K., Huangfu, D., Akutsu, H., Liu, D. R., Rubin, L. L., & Eggan, K. (2009). A small-molecule inhibitor of TGF-β signaling replaces Sox2 in reprogramming by inducing NANOG. *Cell Stem Cell*, *5*(5), 491–503. https://doi.org/10.1016/j.stem.2009.09.012

Jaenisch, R., & Young, R. (2008). Stem Cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell*, *132*(4), 567–582. https://doi.org/10.1016/j.cell.2008.01.015

James, D., Levine, A. J., Besser, D., & Hemmati-Brivanlou, A. (2005). TGFB/Activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development*, *132*(6), 1273–1282. https://doi.org/10.1242/dev.01706

Jiang, J., Chan, Y.-S., Loh, Y.-H., Cai, J., Tong, G.-Q., Lim, C.-A., Robson, P., Zhong, S., & Ng, H.-H. (2008). A Core KLF circuitry regulates self-renewal of embryonic stem cells. *Nature Cell Biology*, *10*(3), 353–360. https://doi.org/10.1038/ncb1698

Kilens, S., Meistermann, D., Moreno, D., Chariau, C., Gaignerie, A., Reignier, A., Lelièvre, Y., Casanova, M., Vallot, C., Nedellec, S., Flippe, L., Firmin, J., Song, J., Charpentier, E., Lammers, J., Donnart, A., Marec, N., Deb, W., Bihouée, A., … Albert, M. L. (2018). Parallel derivation of isogenic human primed and naive

induced pluripotent stem cells. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-017-02107-w

Kim, J., Chu, J., Shen, X., Wang, J., & Orkin, S. H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. *Cell*, *132*(6), 1049–1061. https://doi.org/10.1016/j.cell.2008.02.039

Kunath, T., Saba-El-Leil, M. K., Almousailleakh, M., Wray, J., Meloche, S., & Smith, A. (2007). FGF stimulation of the ERK1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development*, *134*(16), 2895–2902. https://doi.org/10.1242/dev.02880

Ladewig, J., Koch, P., & Brüstle, O. (2013). Leveling waddington: The emergence of direct programming and the loss of Cell Fate Hierarchies. *Nature Reviews Molecular Cell Biology*, *14*(4), 225–236. https://doi.org/10.1038/nrm3543

Lea, R. A., McCarthy, A., Boeing, S., Fallesen, T., Elder, K., Snell, P., Christie, L., Adkins, S., Shaikly, V., Taranissi, M., & Niakan, K. K. (2021). KLF17 promotes human naïve pluripotency but is not required for its establishment. *Development*, *148*(22). https://doi.org/10.1242/dev.199378

Leitch, H. G., McEwen, K. R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W., Nashun, B., Knezovich, J. G., Smith, A., Surani, M. A., & Hajkova, P. (2013). Naive pluripotency is associated with global DNA hypomethylation. *Nature Structural & Molecular Biology*, *20*(3), 311–316. https://doi.org/10.1038/nsmb.2510

Lian, I., Kim, J., Okazawa, H., Zhao, J., Zhao, B., Yu, J., Chinnaiyan, A., Israel, M. A., Goldstein, L. S. B., Abujarour, R., Ding, S., & Guan, K.-L. (2010). The role of yap transcription coactivator in regulating stem cell self-renewal and differentiation.

Genes & Development, *24*(11), 1106–1118. https://doi.org/10.1101/gad.1903310

Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., Nery, J. R., Lee, L., Ye, Z., Ngo, Q.-M., Edsall, L., Antosiewicz-Bourget, J., Stewart, R., Ruotti, V., Millar, A. H., Thomson, J. A., Ren, B., & Ecker, J. R. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, *462*(7271), 315–322. https://doi.org/10.1038/nature08514

Liu, X., Nefzger, C. M., Rossello, F. J., Chen, J., Knaupp, A. S., Firas, J., Ford, E., Pflueger, J., Paynter, J. M., Chy, H. S., O'Brien, C. M., Huang, C., Mishra, K., Hodgson-Garms, M., Jansz, N., Williams, S. M., Blewitt, M. E., Nilsson, S. K., Schittenhelm, R. B., … Polo, J. M. (2017). Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming. *Nature Methods*, *14*(11), 1055–1062. https://doi.org/10.1038/nmeth.4436

Loh, Y.-H., Wu, Q., Chew, J.-L., Vega, V. B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., Wong, K.-Y., Sung, K. W., Lee, C. W., Zhao, X.-D., Chiu, K.-P., Lipovich, L., Kuznetsov, V. A., Robson, P., Stanton, L. W., … Ng, H.-H. (2006). The Oct4 and NANOG transcription network regulates pluripotency in mouse embryonic stem cells. *Nature Genetics*, *38*(4), 431–440. https://doi.org/10.1038/ng1760

Maji, B., Gangopadhyay, S. A., Lee, M., Shi, M., Wu, P., Heler, R., Mok, B., Lim, D., Siriwardena, S. U., Paul, B., Dančík, V., Vetere, A., Mesleh, M. F., Marraffini, L. A., Liu, D. R., Clemons, P. A., Wagner, B. K., & Choudhary, A. (2019). A high-throughput platform to identify small-molecule inhibitors of CRISPR-Cas9. *Cell*, *177*(4). https://doi.org/10.1016/j.cell.2019.04.009

Mak, W., Nesterova, T. B., de Napoles, M., Appanah, R., Yamanaka, S., Otte, A. P., & Brockdorff, N. (2004). Reactivation of the paternal x chromosome in early mouse embryos. *Science*, *303*(5658), 666–669. https://doi.org/10.1126/science.1092674

Malik, N., & Rao, M. S. (2013). A review of the methods for human IPSC derivation. *Methods in Molecular Biology*, 23–33. https://doi.org/10.1007/978-1-62703-348-0_3

Marks, H., Kalkan, T., Menafra, R., Denissov, S., Jones, K., Hofemeister, H., Nichols, J., Kranz, A., Francis Stewart, A., Smith, A., & Stunnenberg, H. G. (2012). The transcriptional and epigenomic foundations of ground state pluripotency. *Cell*, *149*(3), 590–604. https://doi.org/10.1016/j.cell.2012.03.026

Martello, G., Bertone, P., & Smith, A. (2013). Identification of the missing pluripotency mediator downstream of leukaemia inhibitory factor. *The EMBO Journal*, *32*(19), 2561–2574. https://doi.org/10.1038/emboj.2013.177

Martello, G., & Smith, A. (2014). The nature of embryonic stem cells. *Annual Review of Cell and Developmental Biology*, *30*(1), 647–675. https://doi.org/10.1146/annurev-cellbio-100913-013116

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A. A., Ko, M. S., & Niwa, H. (2007). Pluripotency governed by Sox2 via regulation of oct3/4 expression in mouse embryonic stem cells. *Nature Cell Biology*, *9*(6), 625–635. https://doi.org/10.1038/ncb1589

Mathieu, J., & Ruohola-Baker, H. (2017). Metabolic remodeling during the loss and acquisition of pluripotency. *Development*, *144*(4), 541–551. https://doi.org/10.1242/dev.128389

Matsuda, T., Nakamura, T., Nakao, K., Arai, T., Katsuki, M., Heike, T., & Yokota, T. (1999). STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. *The EMBO Journal*, *18*(15), 4261–4269. https://doi.org/10.1093/emboj/18.15.4261

McConnell, B. B., & Yang, V. W. (2010). Mammalian krüppel-like factors in health and diseases. *Physiological Reviews*, *90*(4), 1337–1381. https://doi.org/10.1152/physrev.00058.2009

Messmer, T., von Meyenn, F., Savino, A., Santos, F., Mohammed, H., Lun, A. T., Marioni, J. C., & Reik, W. (2019). Transcriptional heterogeneity in naive and primed human pluripotent stem cells at single-cell resolution. *Cell Reports*, *26*(4). https://doi.org/10.1016/j.celrep.2018.12.099

Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R. P., Lee, W., Mendenhall, E., O'Donovan, A., Presser, A., Russ, C., Xie, X., Meissner, A., Wernig, M., Jaenisch, R., … Bernstein, B. E. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, *448*(7153), 553–560. https://doi.org/10.1038/nature06008

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., & Yamanaka, S. (2003). The homeoprotein NANOG is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell*, *113*(5), 631–642. https://doi.org/10.1016/s0092-8674(03)00393-3

Mullen, A. C., Orlando, D. A., Newman, J. J., Lovén, J., Kumar, R. M., Bilodeau, S., Reddy, J., Guenther, M. G., DeKoter, R. P., & Young, R. A. (2011). Master transcription factors determine cell-type-specific responses to TGF-β Signaling. *Cell*, *147*(3), 565–576. https://doi.org/10.1016/j.cell.2011.08.050

Murry, C. E., & Keller, G. (2008). Differentiation of embryonic stem cells to clinically relevant populations: Lessons from embryonic development. *Cell*, *132*(4), 661–680. https://doi.org/10.1016/j.cell.2008.02.008

Navarro, P., Chambers, I., Karwacki-Neisius, V., Chureau, C., Morey, C., Rougeulle, C., & Avner, P. (2008). Molecular coupling of Xist regulation and pluripotency. *Science*, *321*(5896), 1693–1695. https://doi.org/10.1126/science.1160952

Nichols, J., Silva, J., Roode, M., & Smith, A. (2009). Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. *Development*, *136*(19), 3215–3222. https://doi.org/10.1242/dev.038893

Nichols, J., & Smith, A. (2012). Pluripotency in the embryo and in culture. *Cold Spring Harbor Perspectives in Biology*, *4*(8). https://doi.org/10.1101/cshperspect.a008128

Nichols, J., & Smith, A. (2009). Naive and primed pluripotent states. *Cell Stem Cell*, *4*(6), 487–492. https://doi.org/10.1016/j.stem.2009.05.015

Niwa, H., Burdon, T., Chambers, I., & Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes & Development*, *12*(13), 2048–2060. https://doi.org/10.1101/gad.12.13.2048

Niwa, H., Miyazaki, J., & Smith, A. G. (2000). Quantitative expression of oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature Genetics*, *24*(4), 372–376. https://doi.org/10.1038/74199

Niwa, H., Ogawa, K., Shimosato, D., & Adachi, K. (2009). A parallel circuit of Lif signalling pathways maintains pluripotency of mouse ES cells. *Nature*, *460*(7251), 118–122. https://doi.org/10.1038/nature08113

Nazor, K. L., Altun, G., Lynch, C., Tran, H., Harness, J. V., Slavin, I., Garitaonandia, I., Müller, F.-J., Wang, Y.-C., Boscolo, F. S., Fakunle, E., Dumevska, B., Lee, S.,

Park, H. S., Olee, T., D'Lima, D. D., Semechkin, R., Parast, M. M., Galat, V., … Laurent, L. C. (2012). Recurrent variations in DNA methylation in human pluripotent stem cells and their differentiated derivatives. *Cell Stem Cell*, *10*(5), 620–634. https://doi.org/10.1016/j.stem.2012.02.013

Onder, T. T., Kara, N., Cherry, A., Sinha, A. U., Zhu, N., Bernt, K. M., Cahan, P., Mancarci, B. O., Unternaehrer, J., Gupta, P. B., Lander, E. S., Armstrong, S. A., & Daley, G. Q. (2012). Chromatin-modifying enzymes as modulators of reprogramming. *Nature*, *483*(7391), 598–602. https://doi.org/10.1038/nature10953

Oshimori, N., & Fuchs, E. (2012). The harmonies played by TGF-β in Stem Cell Biology. *Cell Stem Cell*, *11*(6), 751–764. https://doi.org/10.1016/j.stem.2012.11.001

Paquet, D., Kwart, D., Chen, A., Sproul, A., Jacob, S., Teo, S., Olsen, K. M., Gregg, A., Noggle, S., & Tessier-Lavigne, M. (2016). Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/CAS9. *Nature*, *533*(7601), 125–129. https://doi.org/10.1038/nature17664

Pastor, W. A., Chen, D., Liu, W., Kim, R., Sahakyan, A., Lukianchikov, A., Plath, K., Jacobsen, S. E., & Clark, A. T. (2016). Naive human pluripotent cells feature a methylation landscape devoid of blastocyst or germline memory. *Cell Stem Cell*, *18*(3), 323–329. https://doi.org/10.1016/j.stem.2016.01.019

Qi, X., Li, T.-G., Hao, J., Hu, J., Wang, J., Simmons, H., Miura, S., Mishina, Y., & Zhao, G.-Q. (2004). BMP4 supports self-renewal of embryonic stem cells by inhibiting mitogen-activated protein kinase pathways. *Proceedings of the National Academy of Sciences*, *101*(16), 6027–6032. https://doi.org/10.1073/pnas.0401367101

Ran, F. A., Hsu, P. D., Lin, C.-Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., Scott, D. A., Inoue, A., Matoba, S., Zhang, Y., & Zhang, F. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*, *155*(2), 479–480. https://doi.org/10.1016/j.cell.2013.09.040

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-cas9 system. *Nature Protocols*, *8*(11), 2281–2308. https://doi.org/10.1038/nprot.2013.143

Reya, T., Duncan, A. W., Ailles, L., Domen, J., Scherer, D. C., Willert, K., Hintz, L., Nusse, R., & Weissman, I. L. (2003). A role for Wnt signalling in self-renewal of Haematopoietic Stem Cells. *Nature*, *423*(6938), 409–414. https://doi.org/10.1038/nature01593

Ryall, J. G., Cliff, T., Dalton, S., & Sartorelli, V. (2015). Metabolic reprogramming of stem cell epigenetics. *Cell Stem Cell*, *17*(6), 651–662. https://doi.org/10.1016/j.stem.2015.11.012

Qin, H., Hejna, M., Liu, Y., Percharde, M., Wossidlo, M., Blouin, L., Durruthy-Durruthy, J., Wong, P., Qi, Z., Yu, J., Qi, L. S., Sebastiano, V., Song, J. S., & Ramalho-Santos, M. (2016). Yap induces human naive pluripotency. *Cell Reports*, *14*(10), 2301–2312. https://doi.org/10.1016/j.celrep.2016.02.036

Qiu, C., Ma, Y., Wang, J., Peng, S., & Huang, Y. (2010). Lin28-mediated post-transcriptional regulation of OCT4 expression in human embryonic stem cells. *Nucleic Acids Research*, *38*(4), 1240–1248. https://doi.org/10.1093/nar/gkp1071

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S. A., Flynn, R. A., & Wysocka, J. (2010). A unique chromatin signature uncovers early developmental enhancers in humans. *Nature*, *470*(7333), 279–283. https://doi.org/10.1038/nature09692

Reubinoff, B. E., Pera, M. F., Fong, C.-Y., Trounson, A., & Bongso, A. (2000). Embryonic stem cell lines from human blastocysts: Somatic differentiation in vitro. *Nature Biotechnology*, *18*(4), 399–404. https://doi.org/10.1038/74447

Rodda, D. J., Chew, J.-L., Lim, L.-H., Loh, Y.-H., Wang, B., Ng, H.-H., & Robson, P. (2005). Transcriptional regulation of NANOG by Oct4 and Sox2. *Journal of Biological Chemistry*, *280*(26), 24731–24737. https://doi.org/10.1074/jbc.m502573200

Rossant, J., & Tam, P. P. L. (2017). New insights into early human development: Lessons for stem cell derivation and differentiation. *Cell Stem Cell*, *20*(1), 18–28. https://doi.org/10.1016/j.stem.2016.12.004

Sahakyan, A., Kim, R., Chronis, C., Sabri, S., Bonora, G., Theunissen, T. W., Kuoy, E., Langerman, J., Clark, A. T., Jaenisch, R., & Plath, K. (2017). Human naive pluripotent stem cells model X chromosome dampening and X inactivation. *Cell Stem Cell*, *20*(1), 87–101. https://doi.org/10.1016/j.stem.2016.10.006

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

Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T. W., & Smith, A. (2008). Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS Biology*, *6*(10). https://doi.org/10.1371/journal.pbio.0060253

Silva, J., Nichols, J., Theunissen, T. W., Guo, G., van Oosten, A. L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., & Smith, A. (2009). NANOG is the gateway to the

Pluripotent Ground State. *Cell*, *138*(4), 722–737. https://doi.org/10.1016/j.cell.2009.07.039

Simonsson, S., & Gurdon, J. (2004). DNA demethylation is necessary for the epigenetic reprogramming of somatic cell nuclei. *Nature Cell Biology*, *6*(10), 984–990. https://doi.org/10.1038/ncb1176

Smith, A. G., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M., & Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature*, *336*(6200), 688–690. https://doi.org/10.1038/336688a0

Sokol, S. Y. (2011). Maintaining embryonic stem cell pluripotency with Wnt Signaling. *Development*, *138*(20), 4341–4350. https://doi.org/10.1242/dev.066209

Sperber, H., Mathieu, J., Wang, Y., Ferreccio, A., Hesson, J., Xu, Z., Fischer, K. A., Devi, A., Detraux, D., Gu, H., Battle, S. L., Showalter, M., Valensisi, C., Bielas, J. H., Ericson, N. G., Margaretha, L., Robitaille, A. M., Margineantu, D., Fiehn, O., … Ruohola-Baker, H. (2015). The metabolome regulates the epigenetic landscape during naive-to-primed human embryonic stem cell transition. *Nature Cell Biology*, *17*(12), 1523–1535. https://doi.org/10.1038/ncb3264

Sridharan, R., Tchieu, J., Mason, M. J., Yachechko, R., Kuoy, E., Horvath, S., Zhou, Q., & Plath, K. (2009). Role of the murine reprogramming factors in the induction of pluripotency. *Cell*, *136*(2), 364–377. https://doi.org/10.1016/j.cell.2009.01.001

Surani, M. A., Kothary, R., Allen, N. D., Singh, P. B., Fundele, R., Ferguson-Smith, A. C., & Barton, S. C. (1990). Genome imprinting and development in the mouse. *Development*, *108*(Supplement), 89–98. https://doi.org/10.1242/dev.108.supplement.89

Séguin, C. A., Draper, J. S., Nagy, A., & Rossant, J. (2008). Establishment of endoderm progenitors by Sox transcription factor expression in human embryonic stem cells. *Cell Stem Cell*, *3*(2), 182–195. https://doi.org/10.1016/j.stem.2008.06.018

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

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., & Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, *131*(5), 861–872. https://doi.org/10.1016/j.cell.2007.11.019

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W., Reik, W., Bertone, P., & Smith, A. (2014). Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell*, *158*(6), 1254–1269. https://doi.org/10.1016/j.cell.2014.08.029

Ten Berge, D., Kurek, D., Blauwkamp, T., Koole, W., Maas, A., Eroglu, E., Siu, R. K., & Nusse, R. (2011). Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. *Nature Cell Biology*, *13*(9), 1070–1075. https://doi.org/10.1038/ncb2314

Teo, A. K., Arnold, S. J., Trotter, M. W. B., Brown, S., Ang, L. T., Chng, Z., Robertson, E. J., Dunn, N. R., & Vallier, L. (2011). Pluripotency factors regulate definitive endoderm specification through eomesodermin. *Genes & Development*, *25*(3), 238–250. https://doi.org/10.1101/gad.607311

Tesar, P. J., Chenoweth, J. G., Brook, F. A., Davies, T. J., Evans, E. P., Mack, D. L., Gardner, R. L., & McKay, R. D. (2007). New cell lines from mouse epiblast share

defining features with human embryonic stem cells. *Nature*, *448*(7150), 196–199. https://doi.org/10.1038/nature05972

Theunissen, T. W., Powell, B. E., Wang, H., Mitalipova, M., Faddah, D. A., Reddy, J., Fan, Z. P., Maetzel, D., Ganz, K., Shi, L., Lungjangwa, T., Imsoonthornruksa, S., Stelzer, Y., Rangarajan, S., D'Alessio, A., Zhang, J., Gao, Q., Dawlaty, M. M., Young, R. A., … Jaenisch, R. (2014). Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell*, *15*(4), 471–487. https://doi.org/10.1016/j.stem.2014.07.002

Theunissen, T. W., van Oosten, A. L., Castelo-Branco, G., Hall, J., Smith, A., & Silva, J. C. R. (2011a). Nanog overcomes reprogramming barriers and induces pluripotency in minimal conditions. *Current Biology*, *21*(1), 65–71. https://doi.org/10.1016/j.cub.2010.11.074

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., & Jones, J. M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science*, *282*(5391), 1145–1147.https://doi.org/10.1126/science.282.5391.1145

Trusler, O., Huang, Z., Goodwin, J., & Laslett, A. L. (2018). Cell surface markers for the identification and study of human naive pluripotent stem cells. *Stem Cell Research*, *26*, 36–43. https://doi.org/10.1016/j.scr.2017.11.017

Vallier, L., Alexander, M., & Pedersen, R. A. (2005). Activin/nodal and FGF pathways cooperate to maintain pluripotency of Human Embryonic Stem Cells. *Journal of Cell Science*, *118*(19), 4495–4509. https://doi.org/10.1242/jcs.02553

Vallier, L., Mendjan, S., Brown, S., Chng, Z., Teo, A., Smithers, L. E., Trotter, M. W., Cho, C. H.-H., Martinez, A., Rugg-Gunn, P., Brons, G., & Pedersen, R. A. (2009).

Activin/nodal signalling maintains pluripotency by controlling NANOG expression. *Development*, *136*(8), 1339–1349. https://doi.org/10.1242/dev.033951

Veillard, A.-C., Marks, H., Bernardo, A. S., Jouneau, L., Laloë, D., Boulanger, L., Kaan, A., Brochard, V., Tosolini, M., Pedersen, R., Stunnenberg, H., & Jouneau, A. (2014). Stable methylation at promoters distinguishes epiblast stem cells from embryonic stem cells and the in vivo epiblasts. *Stem Cells and Development*, *23*(17), 2014–2029. https://doi.org/10.1089/scd.2013.0639

Waddington, C.H. (1957) The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology. George Allen and Unwin, London.

Wakayama, T., Perry, A. C., Zuccotti, M., Johnson, K. R., & Yanagimachi, R. (1998). Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. *Nature*, *394*(6691), 369–374. https://doi.org/10.1038/28615

Wang, S.-H., Hao, J., Zhang, C., Duan, F.-F., Chiu, Y.-T., Shi, M., Huang, X., Yang, J., Cao, H., & Wang, Y. (2022). KLF17 promotes human naive pluripotency through repressing MAPK3 and ZIC2. *Science China Life Sciences*, *65*(10), 1985–1997. https://doi.org/10.1007/s11427-021-2076-x

Watabe, T., & Miyazono, K. (2008). Roles of TGF-β family signaling in stem cell renewal and differentiation. *Cell Research*, *19*(1), 103–115. https://doi.org/10.1038/cr.2008.323

Weinberger, L., Ayyash, M., Novershtern, N., & Hanna, J. H. (2016). Dynamic stem cell states: Naive to primed pluripotency in rodents and humans. *Nature Reviews Molecular Cell Biology*, *17*(3), 155–169. https://doi.org/10.1038/nrm.2015.28

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B. E., & Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into

a pluripotent es-cell-like state. *Nature*, *448*(7151), 318–324. https://doi.org/10.1038/nature05944

Williams, R. L., Hilton, D. J., Pease, S., Willson, T. A., Stewart, C. L., Gearing, D. P., Wagner, E. F., Metcalf, D., Nicola, N. A., & Gough, N. M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature*, *336*(6200), 684–687. https://doi.org/10.1038/336684a0

Wilmut, I., Schnieke, A. E., McWhir, J., Kind, A. J., & Campbell, K. H. S. (2007). Viable offspring derived from fetal and adult mammalian cells. *Cloning and Stem Cells*, *9*(1), 3–7. https://doi.org/10.1089/clo.2006.0002

Xu, R.-H., Sampsell-Barron, T. L., Gu, F., Root, S., Peck, R. M., Pan, G., Yu, J., Antosiewicz-Bourget, J., Tian, S., Stewart, R., & Thomson, J. A. (2008). NANOG is a direct target of TGFB/activin-mediated SMAD signaling in human escs. *Cell Stem Cell*, *3*(2), 196–206. https://doi.org/10.1016/j.stem.2008.07.001

Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., Liu, P., Lian, Y., Zheng, X., Yan, J., Huang, J., Li, M., Wu, X., Wen, L., Lao, K., Li, R., Qiao, J., & Tang, F. (2013). Single-cell RNA-seq profiling of human preimplantation embryos and Embryonic Stem Cells. *Nature Structural & Molecular Biology*, *20*(9), 1131–1139. https://doi.org/10.1038/nsmb.2660

Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., & Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature*, *453*(7194), 519–523. https://doi.org/10.1038/nature06968

Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., Nie, J., Jonsdottir, G. A., Ruotti, V., Stewart, R., Slukvin, I. I., & Thomson, J. A. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science*, *318*(5858), 1917–1920. https://doi.org/10.1126/science.1151526

Zhang, J., Khvorostov, I., Hong, J. S., Oktay, Y., Vergnes, L., Nuebel, E., Wahjudi, P. N., Setoguchi, K., Wang, G., Do, A., Jung, H.-J., McCaffery, J. M., Kurland, I. J., Reue, K., Lee, W.-N. P., Koehler, C. M., & Teitell, M. A. (2011). UCP2 regulates energy metabolism and differentiation potential of human pluripotent stem cells. *The EMBO Journal*, *30*(24), 4860–4873. https://doi.org/10.1038/emboj.2011.401

Zhang, J.-P., Li, X.-L., Li, G.-H., Chen, W., Arakaki, C., Botimer, G. D., Baylink, D., Zhang, L., Wen, W., Fu, Y.-W., Xu, J., Chun, N., Yuan, W., Cheng, T., & Zhang, X.-B. (2017). Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biology*, *18*(1). https://doi.org/10.1186/s13059-017-1164-8

Zhang, J., Ratanasirintrawoot, S., Chandrasekaran, S., Wu, Z., Ficarro, S. B., Yu, C., Ross, C. A., Cacchiarelli, D., Xia, Q., Seligson, M., Shinoda, G., Xie, W., Cahan, P., Wang, L., Ng, S.-C., Tintara, S., Trapnell, C., Onder, T., Loh, Y.-H., … Daley, G. Q. (2016). Lin28 regulates stem cell metabolism and conversion to primed pluripotency. *Cell Stem Cell*, *19*(1), 66–80. https://doi.org/10.1016/j.stem.2016.05.009

Zhang, X., Huang, C. T., Chen, J., Pankratz, M. T., Xi, J., Li, J., Yang, Y., LaVaute, T. M., Li, X.-J., Ayala, M., Bondarenko, G. I., Du, Z.-W., Jin, Y., Golos, T. G., & Zhang, S.-C. (2010). PAX6 is a human neuroectoderm cell fate determinant. *Cell Stem Cell*, *7*(1), 90–100. https://doi.org/10.1016/j.stem.2010.04.017

Zhong, X., & Jin, Y. (2009). Critical roles of coactivator p300 in mouse embryonic stem cell differentiation and NANOG expression. *Journal of Biological Chemistry*, *284*(14), 9168–9175. https://doi.org/10.1074/jbc.m805562200

Zhou, W., Choi, M., Margineantu, D., Margaretha, L., Hesson, J., Cavanaugh, C., Blau, C. A., Horwitz, M. S., Hockenbery, D., Ware, C., & Ruohola-Baker, H. (2012). HIF1A induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-episc/hesc transition. *The EMBO Journal*, *31*(9), 2103–2116. https://doi.org/10.1038/emboj.2012.71

APPENDIX A: LIST OF PLASMIDS

Appendix A - Table 1. 1 List of plasmids constructed and utilized in experimental procedures.

No	Name	Addgene No	Notes
1	KLF17 Guide 1 cloned LentiCRISPR v2 puro	52961	KLF17 Guide 1: **CACCG**CAGGACAGTCCTCCTGCTGC Used for T7 Assay
2	KLF17 Guide 2 cloned LentiCRISPR v2 puro	52961	KLF17 Guide 2: **CACCG**GTCCCGGCCGATGAGTCTTC Used for T7 Assay
3	KLF17 Guide 3 cloned LentiCRISPR v2 puro	52961	KLF17 Guide 3: **CACCG**TCTTCTGGTGTTGCTTGAGA Used for T7 Assay
4	KLF17 Guide 1 cloned Cas9n	62987	KLF17 Guide 1: **CACCG**CAGGACAGTCCTCCTGCTGC Used for KLF17 Reporter Construction
5	KLF17 Guide 3 cloned Cas9n	62987	KLF17 Guide 3: **CACCG**TCTTCTGGTGTTGCTTGAGA Used for KLF17 Reporter Construction
6	Homology Arm Edited T2A GFP	83344	KLF17 Exon 3 Locus Cloned Used for KLF17 Reporter Construction
7	Homology Arm Edited Mutated T2A GFP	83344	Q5 Site-Directed Mutagenesis for PAM Mutation

			Used for KLF17 Reporter Construction
8	KLF17 OE Plasmid	120503	NO iPSCs Infection dH1f Infection
9	N2L OE Plasmid	21163	NO iPSCs Infection
10	EF1α mCherry	135003	NO iPSCs Infection (Same Backbone with KLF17 OE Plasmid)
11	Nanog OE	16578	dH1f Infection
12	psPAX2	12260	Virus Packaging Plasmids
13	pCMV-VSV-G	8454	Virus Packaging Plasmids
14	KLF5 2_1 Cloned LentiCRISPR v2 puro	52961	KLF5 (2_1) KO gRNA: **CACCG**CAGTCGTAGACCAGTTCTTC Used for dH1f Infection - Reprogramming
15	KLF5 2_2 Cloned LentiCRISPR v2 puro	52961	KLF5 (2_2) KO gRNA: **CACCG**ACTGCAGTGAAACAATTCCA Used for dH1f Infection - Reprogramming
16	KLF5 2_3 Cloned LentiCRISPR v2 puro	52961	KLF5 (2_3) KO gRNA: **CACCG**TCCCTGAGTTCACCAGTATA Used for dH1f Infection - Reprogramming
17	KLF17 3_1 Cloned LentiCRISPR v2 puro	52961	KLF17 (3_1) KO gRNA: **CACCG**TGAGCTTAGACGACATATGC Used for dH1f Infection - Reprogramming
18	KLF17 3_2 Cloned LentiCRISPR v2 puro	52961	KLF17 (3_2) KO gRNA: **CACCG**ATGAGCTTAGACGACATATG Used for dH1f Infection - Reprogramming

19	KLF17 3_3 Cloned LentiCRISPR v2 puro	52961	KLF17 (3_3) KO gRNA: **CACCG**GAGGCCATATTCTTGCAACT Used for dH1f Infection - Reprogramming
20	LIN28A 2_1 Cloned LentiCRISPR v2 puro	52961	LIN28A (2_1) KO gRNA: **CACCG**AGTGGTTCAACGTGCGCATG Used for dH1f Infection - Reprogramming
21	LIN28A 2_2 Cloned LentiCRISPR v2 puro	52961	LIN28A (2_2) KO gRNA: **CACCG**CAGTGGATGTCTTTGTGCAC Used for dH1f Infection - Reprogramming
22	LIN28A 2_3 Cloned LentiCRISPR v2 puro	52961	LIN28A (2_3) KO gRNA: **CACCG**CTGTCCATGACCGCCCGCGC Used for dH1f Infection - Reprogramming