

**CRISPR-Cas9 based Investigation of Chromatin Modifiers in Human Somatic Cell
Reprogramming**

by

Can Aztekin

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

Master of Science

in

Biomedical Sciences and Engineering

Koç University

June 2016

Koç University
Graduate School of Sciences and Engineering

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

Can Aztekin

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

Committee Members:

Tamer Önder, Ph. D. (Advisor)

Nathan Lack, Ph. D.

Batu Erman, Ph. D.

Date

Abstract

Reprogramming of somatic cells to pluripotency via four transcription factors, Oct4, Sox2, Klf4, and c-Myc (OSKM) involves extensive remodeling of the epigenome which is carried out by a number of chromatin modifiers. A number of these genes have been studied in the context of reprogramming and defined as barriers or essential regulators of this process. However, established protocols limit a comprehensive examination due to the susceptibility of shRNA vectors to proviral silencing which accompanies reprogramming, and the unavailability of human knock-out cell lines. Therefore, we hypothesized that pooled CRISPR-Cas9 knock-out libraries can enable systematic investigation of chromatin modifiers and reveal their role in different steps of reprogramming by enabling permanent gene modifications. To this end, we constructed a CRISPR-Cas9 knock-out library targeting the 247 chromatin modifiers including chromatin remodelers, writers, readers and erasers of histone and DNA modifications. We utilized this library to generate heterogeneous cell populations harboring putative chromatin modifier knock-outs, which were then reprogrammed with the delivery of OSKM. FACS-based enrichment of TRA-1-60 positive cells, and low copy exogenous retroviral EGFP silencing was utilized to collect emerging iPSCs at different time points of reprogramming. The abundance of CRISPR guide RNAs in isolated cells were determined by next generation sequencing. Analysis of relative enrichment of specific guide RNAs confirmed the majority of the known reprogramming regulators and identified novel stage specific barriers and essential regulators for human somatic cell reprogramming. These include PAXIP1, BRD2 and USP22 as barriers, and CHAF1A, KDM8 and PRMT5 as essential regulators. In addition, functional analysis based on class of genes revealed that various MLL complexes act as barriers, and BAD, SWI/SNF, and MOZ/MORF complexes act as essential regulators for reprogramming. Based on the screen results, stage specific combinatorial use of small molecule inhibitors enabled accelerated and more synchronous efficient reprogramming. Taken together this study establishes roadblocks of reprogramming through chromatin landscape, and can guide more efficient reprogramming strategies to be formulated.

Özet

Somatik hücrelerin Oct4, Sox2, Klf4 ve c-Myc transkripsiyon faktörleri ile pluripotensiye yeniden programlanması, epigenomun kromatin modifiye edenler aracılığıyla yeniden biçimlenmesiyle gerçekleşir. Bu bağlamda çalışılan genlerin bazıları yeniden programlamanın bariyer ve gerekli genleri olarak adlandırılmıştır. Bununla birlikte gelişen protokollerle yapılan geniş kapsamlı çalışmaları, yeniden programlanmanın beraberinde gelen shRNA verktörlerinin susturulmaya duyarlılığı ve bu genlerin knockout olduğu insan hücre hatlarının bulunmaması sınırlandırıyor. Böylece, yeniden programlanmanın farklı evrelerinde bu genlerin kalıcı modifikasyonu CRISPR-Cas9 knockout kütüphanesi kullanarak kromatin modifiye edenlerin sistematik incelenmesi hipotezini belirledik. Bu amaçla, 247 histon ve DNA modifiye eden kromatin yeniden düzenleyiciler, yazıcılar, okuyucular ve siliciler içeren kromatin modifiye edenlerin CRISPR-Cas9 kütüphanesi oluşturuldu. Ardından bu kütüphane, heterojen hücre popülasyonu oluşturulması ve sonrasında OSKM ile yeniden programlanmaya tabi tutulmasında kullanıldı. Bunun için, FACS aracılığıyla ayrıştırılmış Tra-1-60 pozitif ve retroviral yolla eksojen olarak üretilen yeşil floresan proteinleri (EGFP) susturulmuş hücrelerden elde edilen uPKH'lerden farklı zaman aralıklarında toplanılanlar kullanıldı. İzole edilen hücrelerde CRISPR rehber RNA'ların miktarı ileri nesil sekanslama yöntemi ile belirlendi. Spesifik rehber RNA'ların çokluğunun analizi, önceden çalışılmış ve ispatlanan somatik hücrelerin yeniden programlanmasının düzenleyicilerini doğruladı ve farklı dönemlere özgü yeni bariyer ve gerekli düzenleyiciler saptandı. Bunlardan PAXIP1, BRD2 ve USP22 bariyer ve CHAF1A, KDM8 ve PRMT5 gerekli düzenleyiciler olarak belirlendi. Üstelik gen kümelerinin işleve dayalı analizi, MLL komplekslerinin yeniden programlanmanın muhtemel bariyeri, BAF, SWI/SNF ve MOZ/MORF komplekslerinin de önemli düzenleyiciler olarak görev aldıklarını açıkladı. Tarama sonuçlarına dayanarak, farklı dönemlere özgü inhibitörlerin kombinasyonlarının kullanılmasının, yeniden programlanmayı hızlı ve daha verimli hale getirdiği görüldü. Bu çalışma yeniden programlamanın kromatin bazlı engellerinin açıklanmasına ve daha verimli yeniden programlama yolları için rehber olacaktır.

Acknowledgements

First, I would like to thank my thesis supervisor, Dr. Tamer Önder. He had provided many different opportunities for me to learn and think about science in general. The independence granted by him contributed me to become the scientist I am today. Throughout my masters, I tried to benefit from his guidance in every aspect, and I hope that I was able to make most out of my experiences in his lab.

My sincerest thanks goes to Dr. Tuğba Bağcı Önder, Dr. Mehmet Gönen and especially, to Dr. Nathan Lack. They all helped this thesis to turn into something highly valuable and showed me different perspectives on science. Particularly, I am grateful to Dr. Nathan Lack both for his time to become one of the juries in this thesis and also for all the daily conversations, and discussions about science. I always tried my best to carry the mindset of my mentors with me while trying to become a better scientist, and I am glad that I found new mentors in this journey.

I appreciate Dr. Batu Erman for his time and evaluating my work by becoming one of the juries in this thesis. He was one of the truly inspirational instructors during my undergraduate education, and I used many different skills I gained from his lectures in this thesis.

I would like to thank the people from Koç University School of Medicine (KUSOM), specifically all the members of Lack Lab, Bağcı-Önder Lab, and Önder Lab. I believe that each one of them contributed to my growth as a scientist in different ways. Especially, I am thankful to Gülben Gürhan, not only because all the experiments she helped me to conduct, but also becoming like a big sister to me, Kenan Sevinç for his trolling trials in every opportunity he had and for endless repetitive scientific discussions that we had, Tunç Morova for both reminding me the existence and importance of mathematics, and for his energetic spirit that motivates for further work. I also have to thank the undergraduate students Cihan Civan Civas and Ekin Deniz Aksu as I sincerely appreciate them for showing me the long way I need to overcome to become a great mentor. In addition to the people from KUSOM, I have to thank Deniz Çonkar and Haroon Qureshi whom I shared many great discussions on science through our nerdy coffee breaks.

Particularly, I have to express my gratitude to Zeynep Kaya for all her support for *simply* everything. Without her, my negative energy would probably resonate through all the labs. I hope that I was also able to show her a tiny bit of patience that she showed me and support her in her journey. Hope to see her soon while chasing swans in England!

I want to thank my friends from my homeland whom I consider as my second family. Çağlar Nişancı, Engin Yazar, Bilgesu Parmaksız, Işıl Okur, Ege Kılınç, Saygın Sop, Melih Alan, and Leyla Sokullu, I am glad to have the chance to grow up with you. Also, thank you *once again* for showing me the distance and conditions are irrelevant for friendship.

For last and most, I have to thank my parents, Belgin and Metin Aztekin, and my sister, Hazal Aztekin. There is no way I can thank them enough for raising me with the personality I have now. I know that I still have a lot to learn from them and can not thank them enough for always supporting me and being there for me.

Table of Contents

1	Introduction	12
1.1.	Reprogramming and Induced Pluripotent Stem Cells	12
1.1.1.	Stages of inducing pluripotency from a somatic cell	13
1.1.2.	Distinguishing Features of Fully Reprogrammed Cells	16
1.1.3.	Regulators of Reprogramming: Barriers, Enhancers and Essential Regulators	17
1.2.	Chromatin Modifications.....	18
1.2.1.	Chromatin Modifications as Barriers, Enhancers and Essential Regulators of Reprogramming	20
1.3.	CRISPR Cas9 system	22
1.3.1.	Comparison of CRISPR-Cas9 against shRNA in Arrayed and Pooled Screens	24
1.4.	Aim	25
2	Materials and Methods	26
2.1	Cell Culture	26
2.2	Generation of Modified dH1fs	26
2.3	Reprogramming assays	26
2.4	TRA-1-60 HRP Immunostaining	27
2.5	TRA-1-60-PE staining for Flow Cytometry	28
2.6	DNA Constructs	28
2.7	Cloning.....	29
2.8	cDNA synthesis and qPCR.....	32
2.9	T7 Endonuclease Assay	33
2.10	Production of Viruses.....	34
2.11	Concentration of Viruses.....	35
2.12	Titering of Viruses	35
2.13	Flow Cytometry.....	35
2.14	Fluorescence Activated Cell Sorting (FACS)	36
2.15	Library Design.....	36
2.16	Oligo array synthesis and pooled library cloning.....	36
2.17	Genomic DNA Sequencing.....	37
2.18	Data Processing and analysis of Next-Generation Sequencing	38
2.19	TRA-1-60 ,Low Copy Exogenous EGFP Silencing, and Proliferation Screens	39
2.20	Differential Gene Expression Analysis	39

2.21	Statistical Analysis	39
3	Results	40
3.1	CRISPR-Cas9 system is suitable for investigations of regulators of reprogramming in human fibroblasts	40
3.2	Generation of a focused CRISPR-Cas9 Knock-out Library Targeting Chromatin Modifiers	43
3.3	TRA-1-60 Expression Based Enrichment Screen Results and Data Analysis	44
3.4	Characterization of low copy exogenous EGFP silencing, Screen Results and Data Analysis	59
3.5	Formulation of novel reprogramming strategies based on screen results	67
4	Discussion.....	71
4.1	Findings that validate and propose reprogramming regulator chromatin modifier genes, complexes, and functions	71
4.2	Precautions and observations about the experiments.....	74
4.3	Future directions.....	77
5	Appendix	78
6	Bibliography	131
7	Vita	145

List of Figures

Figure 1.1 Stages of inducing pluripotency.	15
Figure 3.1 Cas9 Activity Assay shows the suitability of CRISPR-Cas9 for dH1fs	40
Figure 3.2 CRISPR-Cas9 system is suitable for gene investigations in reprogramming experiment.....	42
Figure 3.3 Chromatin modifiers targeting CRISPR-Cas9 Knock-Out Library Generation.....	43
Figure 3.4 Proposed cell surface antigen enrichment based screen outline for investigating 247 chromatin modifiers role in reprogramming	44
Figure 3.5 TRA-1-60 staining to investigate stage-specific features of modifications validated with the effect of DOT1L inhibition on different stages	46
Figure 3.6 Correlation between two independent TRA-1-60 sorting screen.	47
Figure 3.7 Screen results validate most of the known human reprogramming regulators	48
Figure 3.8 NGS results from the screen analyzed and quantified with MTC and normalized to the NT scores	49
Figure 3.9 Screen results show consistency throughout reprogramming and identify stage- specific effects of modifications	50
Figure 3.10 Proliferation screen results and its effect of proliferation on successful iPSC generation.....	51
Figure 3.11 Individual gRNA testing for validation of screen	52
Figure 3.12 Screen results categorized for chromatin modifier complexes to identify reprogramming regulators at the protein complex level	53
Figure 3.13 Chromatin modifiers belonging to different functional categories of act differently through reprogramming	54
Figure 3.14 Chromatin modifiers represent as essential regulators of reprogramming.....	55
Figure 3.15 Specific histone marks and their reader, writer, erasers do not show any clustered phenotype.....	56
Figure 3.16 Differential gene expression between iPSC and fibroblasts might indicate their role in reprogramming	57
Figure 3.17 Hits identified in the initiation stage impairs somatic cell identity in fibroblasts without OSKM infection.....	58
Figure 3.18 Overall hits from the screen enhance expression of pluripotency-associated genes at Day 7 of reprogramming.....	59

Figure 3.19 EGFP loss is observed in the early stage of reprogramming is due to the OSKM infection, and this phenotype is mediated by proviral silencing mechanism	60
Figure 3.20 Proviral silencing functions in copy number dependent manner through reprogramming.....	62
Figure 3.21 Early stage exogenous EGFP silencing can be used to enrich emerging iPSC....	63
Figure 3.22 Negative selection strategy for validation of proviral silencing.....	64
Figure 3.23 Effect of CRISPR induced gene perturbations can be assessed with proviral silencing at early time points of reprogramming	65
Figure 3.24. Correlation between technical replicates of low copy exogenous EGFP screen.	66
Figure 3.25. Low copy exogenous EGFP silencing based screen revealed regulators of reprogramming.....	67
Figure 3.26 Validation of stage-specific screen results by small molecule inhibitors.....	68
Figure 3.27 Novel reprogramming strategies can be formulated with screen results.....	70

List of Tables

Table 1. Barriers, Enhancers, and Essential Regulators of human iPSC generation	22
Table 2. List of gRNAs used in the reprogramming experiments.	27
Table 3. List of used small molecule compounds.....	27
Table 4. List of Addgene plasmids used in this study	29
Table 5. List of gRNAs cloned into the lentiCRISPRv1	30
Table 6. List of gRNAs cloned into the lentiGuide	31
Table 7. List of gRNAs cloned into the lenti sgRNA(MS2)_zeo	31
Table 8. List of gRNAs cloned into the lentiCRISPRv2	32
Table 9. List of primers for qRT-PCR used in the study.	33
Table 10. List of Tested Surveyor Primers for gRNAs	34
Table 12. List of external and internal primers.....	38
Table 13. List of time points divisions to estimate stage scores.....	39

List of Appendix

APPENDIX 1. Library represented through different time points in the screen.	78
APPENDIX 2 Screen results showing consistency between overlapping time points in probe level.....	79
APPENDIX 3. Screen results identify stage-specific effects of modifications in probe level.	79
APPENDIX 4. Histone modifications for specific residues and their regulator writer, reader, erasers do not show specific clustering in a region for different stages of reprogramming. ...	80
APPENDIX 5. SAM system is functional in transient transfection experiments in 293ts.	80
APPENDIX 6. SAM system protocol for 293t transfection.	81
APPENDIX 7. Full list of gRNAs in EPIKOL Library.	81
APPENDIX 8. Full Enrichment Scores for an average of two independent TRA-1-60 screens.	122
APPENDIX 9. Full Enrichment Scores for average of three replicates of low copy exogenous EGFP silencing screen.	128

1 Introduction

1.1. Reprogramming and Induced Pluripotent Stem Cells

Upon fertilization, development starts and embryonic cells differentiate into all the different cell types present in the adult body. This process has been thought to be irreversible. However, the discovery of reprogramming challenged this notion. Somatic cell nuclear transfer (SCNT) in which a nucleus of a somatic cell is inserted into an enucleated egg can reprogram cells into pluripotent stem cells (PSCs). Therefore, SCNT shows that each somatic cell contains sufficient information for all the cell fates and developmental processes can be reversed. A different line of experimentation has shown that, a single transgene, MYOD overexpression enables fibroblasts to be reprogrammed to myoblast cells [1]. This transdifferentiation experiment also shows that the unidirectional path leading progenitor cells to become terminally differentiated cell types can be altered such that cell fate can even change between different tissue types [1]. This notion has led to the hypothesis that even fully differentiated somatic cells might be converted into pluripotent stem cells in the presence of certain transcription factors and culture conditions. Such cells would be valuable sources for both understanding developmental processes and their prospective applications in the regenerative medicine [2].

This hypothesis has been proven by the work of Takahashi and Yamanaka who showed that the ectopic expression of four transcription factors that are highly expressed in ESCs, namely OCT4, SOX2, KLF4 and C-MYC (OSKM) and specific culture conditions are sufficient to convert adult fibroblasts to induced pluripotent stem cells (iPSC) [3, 4]. Following this discovery, alternative transgene cocktails have also been utilized for reprogramming to pluripotency [5]. In 2013, transient exposure of small molecule compounds have been shown to convert somatic cells into the pluripotent state in mice [6]. iPSCs show morphological, molecular and functional characteristics of ESCs. Morphologically, mouse-derived iPSCs resemble the inner cell mass isolated embryonic stem cells which grow as small dome-shaped colonies [7, 8]. Human derived iPSCs resemble epiblast stem cells which grow as larger colonies with flattened morphology [3, 7]. At the molecular level, iPSC and ESCs show similar global gene expression levels, DNA methylation status, histone marks including establishing bivalency in lineage-specific gene promoters, active TERT function, and the ability to silence transgenes [3, 4]. Functionally, iPSCs have similar ability to self-renew and differentiate as ESCs. They are immortal *in vitro*, can form teratomas when transplanted into

an immunodeficient mouse, differentiate into three germ layers *in vitro*, and can give rise to chimeras if injected into the blastocyst [3, 4, 9].

The traditional method for generating human iPSCs involves overexpressing OSKM in fibroblasts, co-culturing with mouse embryonic fibroblasts, and supplementing them with embryonic stem cell medium. This method has been used to investigate the mechanism of reprogramming but has certain drawbacks. First, although the generated cells are pluripotent, the efficiency is very low. In initial experiments, retroviral vectors were used for transgene expression that led to a reprogramming efficiency of 0.001 - 0.1% [3, 4]. Subsequent methods that utilize lentiviral vectors, transfection of mRNA or proteins did not significantly alter the efficiency [10]. Low efficiency in the delivery of transgenes does not seem to be the reason for the low efficiency as evidenced from the results of mouse secondary iPSC systems. Where, fibroblasts were reprogrammed with tet-inducible OSKM expression vectors, and a single iPSC colony was picked for fibroblast differentiation. After this process, differentiated fibroblasts which have the same integrated OSKM transgene in their genome were tested once again for reprogramming with doxycycline addition [11]. Even with this system, reprogramming efficiency reached up to only 4% [10]. These results suggest that there are unknown transgene-independent factors in this type of cell fate conversion.

Contrary to the SCNT, generation of iPSCs requires continuous cell division and takes approximately 14-21 days. If cells are prone to senescence or have impaired proliferative capacity, generation efficiency severely decreases [11, 12]. On the other hand, in nuclear transfer, there is no requirement for proliferation as the enucleated egg contains the necessary information at mRNA or protein level to induce reprogramming [2]. In addition, the usage of oncogenic KLF4 and C-MYC as inducers create a potential risk for transforming the generated iPSCs to a cancerous phenotype, therefore alternative strategies to generate these cells have to be formulated to eliminate this problem [9].

1.1.1. Stages of inducing pluripotency from a somatic cell

Three main stages for inducing pluripotency have been identified: the initiation stage in which somatic cells lose their identity, the maturation stage where cells establish their pluripotency regulatory network, and the last, the maintenance or stabilization phase that enables sustaining pluripotency without differentiation [13, 14].

During the initiation phase, ectopic expression of OSKM leads to remodelling of heterochromatin regions through the joint binding of OCT4, SOX2, KLF4 to their targets to initiate the transcription of a limited set of early pluripotency and late embryogenesis genes [15]. Furthermore, cells initially undergo a large proliferation wave, lose their mesenchymal characteristics, start to express epithelial markers, and shift their metabolism towards glycolytic pathways instead of oxidative phosphorylation [16–18]. Moreover, the early phase of reprogramming is a stressful event in which reactive oxygen species levels and DNA damage lead cells to apoptosis [19, 20]. This latter response is partially blocked by the overexpression of the KLF4 and the C-MYC oncogenes [21].

In the maturation stage, cells enter into an intermediate state in which the endogenous pluripotency circuit has to be activated. The initial burst of OCT4 and SOX2 expression allow these factors to bind their shared promoters, however, successful transcription of target genes requires additional inputs. Accordingly, Klf4 acts as a trigger for paused Polymerase (Pol) II to enable successful expression of these genes [22]. In addition to its role in elongation, KLF4 expression also enables chromosomal looping necessary for activation of the endogenous OCT4 locus [23]. Within the maturation stage reprogrammed cells begin to express pluripotency markers, such as alkaline phosphatase, SSEA-3, TRA-1-60 and telomerase is activated to elongate telomeres with each cell division [24]. Recent evidence also suggests that these intermediate cells also have gene expression profiles observed during primitive streak formation, with endoderm and ectoderm genes being upregulated [17, 25]. These findings suggest that cells backtrace their developmental origins.

In the last maintenance stage, pluripotency is sustained and regulated solely by the endogenous pluripotency network, and differentiation is curtailed under normal culture conditions. First, ectopic transgenes are silenced, and the expression of pluripotency genes are sustained from their endogenous loci. In fact, sustained transgene expression in the pluripotent state results in cells losing their ESC-like morphology and competency for chimera formation [18]. Furthermore, in the absence of transgene silencing, cells are unable to differentiate effectively [18]. Maintenance stage is the most important stage for the comparison of iPSCs with ESCs. It is important to note that ESCs have been cultured for a long time *in vitro*, and iPSCs have to go through a similar process, maintaining stem cell identity for a certain period *in vitro* to be faithfully used in comparisons of gene expression profiles and global DNA methylation status [26].

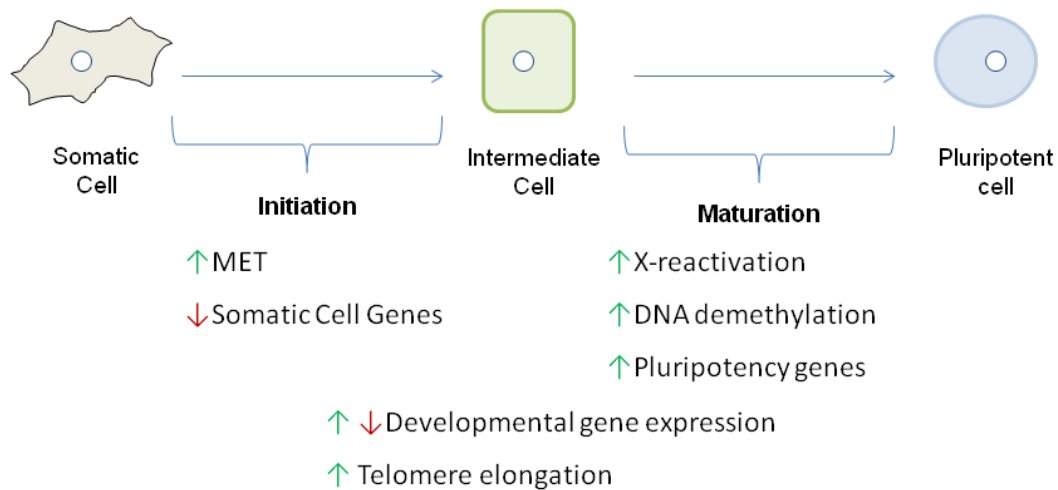


Figure 1.1 Stages of inducing pluripotency.

Two stages of reprogramming depicted with the major changes in gene expression and functions. Somatic cells initially upregulate proliferation genes and epithelial genes through activating mesenchymal to epithelial transition (MET) and downregulate somatic cell genes. In the maturation stage, cells upregulate pluripotency genes, and DNA demethylation and X-reactivation occurs to form pluripotent stem cells. Throughout the initiation and maturation stages, telomeres start to elongate, and developmental genes fluctuate as they are initially upregulated then become silenced towards the end of reprogramming [27]. (Figure adapted from Takahashi & Yamanaka, 2015).

Stages of reprogramming show similarities with developmental processes in reverse order. ESCs are derived from cells present in the blastocyst. In the implantation stage, pluripotent cells initiate a cascade of events that leads to the formation of the primitive streak [25, 27]. During this process, specification proceeds with the decrease in characteristic features of ESC and increases in DNA methylation levels and upregulation of tissue-specific genes. In addition, the gastrulation stage is primarily enabled by the activation of the epithelial to mesenchymal transition (EMT) which is reversed during the early stage of reprogramming, where cells undergo a mesenchymal to epithelial transition (MET). Consequently, reprogramming of somatic cells to the pluripotency follow the reverse order of developmental processes from the blastocyst stage to gastrulation [27].

1.1.2. Distinguishing Features of Fully Reprogrammed Cells

At the end of reprogramming, the majority of cells are morphologically similar to ESC-like cells but are functionally restricted intermediate cells [2]. Therefore, identifying markers to distinguish the bone fide pluripotent stem cells is critical.

Detection of fully potent ESC-like cells and intermediate cells has been studied [9, 24, 28]. The initial discovery of iPSCs achieved by selecting reprogrammed cells via stem cell specific FBX15 promoter activity linked with the Neomycin resistance gene [4]. Although this selection enabled isolation of ESC-like cells that show characteristics of ESCs, gene expression patterns of these cells were different than ESCs and they did not give rise to the chimeras when injected into the mouse blastocysts [4]. The use of an endogenous Nanog locus as a selection marker enabled chimera competent iPSC generation [9]. Subsequent studies indicated that multiple markers needs to be tested to distinguish fully reprogrammed cells. In reprogramming human cells, proviral silencing and expression of REX1, DNMT3B and TRA-1-60 are found to be suitable for selecting iPSCs equivalent to ESCs [24]. Moreover, proteomics studies identified CD73, CD49d, and CD200 as true indicators for reprogrammed cells in the intermediate stages [28]. Interestingly, although TRA-1-60 has been shown to identify fully reprogrammed cells, there is also evidence showing its expression can be observed in mid-stages of reprogramming, but is then lost in the same cells in the maintenance stage [29].

One distinguishing feature of fully reprogrammed cells is their ability to silence retroviral elements through epigenetic mechanisms [4]. Studies done in embryonic carcinoma cells and ESC showed that the pluripotent stem cells can recognize retroviral elements through their primer binding sites and methylate these regions resulting in heterochromatin [30–34]. Recognition of the primer binding site is regulated by ZFP809, ESET1, and TRIM28 and further methylation of histone H3 lysine 9 (H3K9) and recruitment of HP1 enable heterochromatin formation [32, 34, 35]. Moreover, in a genome-wide study that investigated regulators of proviral silencing, histone chaperones CHAF1A, CHAF1B, and ubiquitinase SUMO2, have been found to regulate this process [36]. Although PSCc can silence viral elements in general, there is some selectivity based on the type of promoters. Promoters of housekeeping genes used in lentiviral vectors are harder to be repressed as seen with PGK promoter [37, 38]. On the other hand, promoters of retroviruses such as LTRs are very

sensitive to silencing. Additionally, proviral silencing functions randomly and some of the elements might escape silencing with a low probability [39].

1.1.3. Regulators of Reprogramming: Barriers, Enhancers and Essential Regulators

While currently used reprogramming systems have a low iPSC generation efficiency, there are modifications that can increase or decrease this efficiency suggesting the presence of intrinsic barriers, enhancers and essential regulators for this process [21, 40]. The first such regulator to be identified was P53. KLF4 has been known to suppress P53, and P53 knock-down in addition to OSKM expression has been shown to increase reprogramming efficiency [41–43]. Thus, P53 is a barrier to reprogramming. Similarly, other tumor suppressor genes such as P21, P16, and PTEN have also been defined as barriers [44–47]. On the other hand, overexpression of several genes such as C/EBPa or FOXH1 in addition to OSKM can increase the efficiency of iPSC generation, indicating that these genes act as enhancers of reprogramming [25, 40, 48, 49]. For last, some genes such as UTX is required for successful iPSC generation. Its absence hampers reprogramming while its overexpression does not further enhance reprogramming. [50]. Therefore, it is an essential regulator of reprogramming.

In addition to the ectopic expression of certain genes, use of small molecule inhibitor compounds has been shown to increase reprogramming efficiency [51–53]. For example, histone deacetylase (HDAC) inhibitor valproic acid increases reprogramming efficiency, and can replace C-MYC in the OSKM cocktail [51]. This finding suggested that identifying small molecule inhibitors targeting reprogramming barriers can both increase the efficiency and can replace some of the exogenous factors. Therefore, such perturbations would both eliminate the low-efficiency and the potential problems associated with transformation due to the usage of transgenes. Further work in keratinocytes showed that combinatorial usage of small molecule inhibitors of NOTCH pathway and DOT1L shows a greater increase in reprogramming efficiency [53]. Therefore, manipulation of barriers with small molecule compounds can be used to formulate more efficient reprogramming strategies.

Large-scale studies have been conducted to discover gene sets or functional categories as barriers or essential regulators. Among those large scale efforts, a kinome-wide shRNA screen identified cytoskeletal remodeling as a regulator of reprogramming [54]. In another study, a genome-wide shRNA library identified the ubiquitin pathway, cell adhesion and motility genes and endocytosis were implicated as major barriers for reprogramming [55]. A

genome-wide transcriptome analysis and a pooled shRNA screen conducted in the mouse system revealed that non-differentially expressed genes are also important for cell fate decision [56]. However, although there are species difference within these screens, there was little to no overlap found.

The observed effects of perturbations during reprogramming depends on various parameters such as, species type, mode of inhibition, and starting cell type. Therefore, a gene may negatively or positively influence reprogramming in a context-dependent manner. For example, knock-down of EHMT1 hampers reprogramming of human cells [57] but increases reprogramming in mice [58]. Transient inhibition of EZH2 via small molecule inhibitors during the maturation stage increases iPSC formation while its stable knock-down with shRNAs decrease overall iPSC generation [6, 57]. Similarly, transient knockdown of MBD3 with siRNAs increases, but knockdown of MBD3 by integrating shRNAs decreases efficiency [59, 60]. The effect of cell type in reprogramming has been assessed by using blood progenitors and hepatoblasts as starting cell types [61]. Inhibition of Wnt pathway facilitated reprogramming of blood progenitors, but in hepatoblasts the same effect was observed with inhibition of TGF- β pathway. Moreover, inhibition of Wnt signaling in hepatoblasts and inhibition of TGF- β pathway in blood progenitors did not result in an increase in reprogramming efficiency. Lastly, barriers might only be effective when different perturbations are simultaneously present. For example, 4E-BP knock-down alone does not increase reprogramming efficiency, but when combined with the p53 knock-down, leads to more synchronous reprogramming than with p53 knock-down alone [62].

1.2. Chromatin Modifications

Genes have to be expressed or repressed in a controlled manner throughout developmental stages and during the formation of different types of tissues [63, 64]. This coordinated regulation of gene expression in different tissue types is in part organized by epigenetic modifications which occur on DNA and histone proteins. The combination of these two molecules is referred to as chromatin, and modifications to this structure are orchestrated by DNA and histone modifiers, and chromatin remodelers.

DNA is subjected to the methylation events as an epigenetic mechanism and results in gene silencing. The addition of methyl groups to the CpG dinucleotides present in promoters enables repression of gene expression by either recruiting repressing proteins to form heterochromatin or inhibiting the binding of transcription factors to the region which would

normally activate gene expression. DNA methylation is carried out by the activity of DNMT family of proteins. DNMT3A/B are required for *de novo* methylation, whereas DNMT1 acts to maintain the methylation status by recognizing newly replicated hemimethylated DNA strands[65]. Through development, as cells commit to certain lineages, genes specify alternative cell fates are subjected to the DNA methylation and silencing. Removal of DNA methylation is controlled both passively and actively. In its passive form, with each cell division, certain hemimethylated CpG islands may fail to be methylated and methylation in these regions are lost. In active DNA demethylation, TET proteins convert 5-methylcytosine to the 5-hydroxymethylcytosine which is then cleaved by repair mechanisms [65, 66].

Histone modifications are mainly found in the N-terminal tails of histone molecules. These tails do not contact the DNA and are subjected to various modifications. These modifications include addition or removal of methyl, acetyl, ubiquitin, and phosphate groups [67]. Each of these modifications occurs at specific histone residues and are associated with a certain expression status. Among them, acetylation of lysine residues is associated with open chromatin and active transcription. Other modifications such as methylation, ubiquitination and phosphorylation regulate gene expression depending on the modified residue. For example, H3K9me3 represses gene expression, but H3K36me3 correlates with transcriptional elongation [67–69].

There are three main categories of protein complexes associated with histone modifications; readers, writer, and erasers [70, 71]. Readers recognize histone marks and recruit additional proteins to the region to elicit a certain function. For example, proteins that recognize acetylated residues harbor bromodomains, and those that recognize methyl marks contain PHD domains. Writers have catalytic subunits for adding the marks to certain residues. For example, methyltransferases are writers that generally harbor SET domains. Conversely, erasers are proteins that catalyze the reactions for removal of histone marks. All of these proteins share similar structural domains but nevertheless, exhibit high specificity towards their target. Moreover, histone modifiers generally work in large complexes. For example, Polycomb Repressive Complex 2 (PRC2) consists of EED, EZH1, EZH2 and SUZ12 to deposit the H3K37me3 repressive mark. Similarly, the MLL complex, composed of KMT2A, WDR5, RbBP5, ASH2L, and KMT2A deposits the activation-associated H3K4me3 mark [72].

Chromatin remodelers enable association or dissociation of chromatin structure and interaction between DNA and histones. In addition, they regulate the placement of histones to regulate the dynamic nature of chromatin [63, 64]. These complexes carry out nucleosome sliding to facilitate elongation, unwrap nucleosome via recruiting histone acetyltransferases to open chromatin, and can exchange histone variants. There are mainly four chromatin remodeler families in mammals, SWI/SNF, ISWI, CHD, INO80. By regulating major changes in chromatin dynamics, these chromatin remodelers have been shown to play important roles in zygotic activation and differentiation of each terminally differentiated cell [63, 64].

1.2.1. Chromatin Modifications as Barriers, Enhancers and Essential Regulators of Reprogramming

Chromatin modifiers have been investigated for their ability to be either a barrier or essential regulator of reprogramming due to their role in developmental processes and gene expression regulation. Individual genes, protein complexes, and histone marks have been proposed as barriers or essential regulators of reprogramming by investigating their effects on sustaining pluripotency in ESCs, and by loss of function screens. An integrative study analyzed the cells in the intermediate stage of reprogramming and have shown that DNMT1 knockdown with lentiviral shRNAs, siRNA transfection, or with 5-Azacytine treatment can convert cells to the fully reprogrammed cells. Therefore, DNMT1 is classified as a barrier, and its loss of function enables more efficient reprogramming [73]. Gene perturbations that affect pluripotency have been tested with their role in reprogramming and found to be essential regulators or enhancers for reprogramming [74, 75]. For an example, BRD4 knockdown causes ESC to lose their pluripotency and is required for reprogramming by enabling Polymerase II release after transcription initiation [74, 76]. Thus, BRD4 is an essential regulator for reprogramming, and as its overexpression enhances reprogramming efficiency, it is also an enhancer. Screenings sets of chromatin modifiers identified the sole H3K79 methyltransferase DOT1L as a barrier in both human and mouse while histone chaperones CHAF1A and CHAF1B act as barriers only in mouse reprogramming [57, 77]. Moreover, knock-down of members of PRC1 and PRC2 complex decrease iPSC generation and thus, PRC1 and PRC2 complexes are classified as essential regulators [57, 78, 79]. In addition to the individual genes and protein complexes, specific histone marks have also been proposed to be barriers. In a study where H3K9 demethylases are knocked-down in mice, a general

trend of decreased iPSC generation has been observed, indicating that H3K9me3 is a barrier to the reprogramming [80].

Most of the studies to date have investigated the overall effect of chromatin modifiers on iPSC generation without focusing on their role at different stages of reprogramming. However, there is increasing evidence that chromatin modifiers carry out different functions in each of the different stages. For example, H3K27 demethylase UTX has been shown to promote the efficiency of reprogramming in the initiation phase by facilitating opening of chromatin for the pluripotency regulatory network [50]. However, during the maintenance stage, iPSC can not sustain their pluripotency as some other regions preserves H3K27 methyl marks. Therefore, UTX is one of the stage-specific barriers for reprogramming. On the other hand, HDAC2 has been shown to be a barrier at the maturation stage where its presence blocks TET1-associated DNA methylation of pluripotency-associated genes. HDAC2 inhibition enables demethylation of such gene promoters, and results in increased iPSC generation efficiency [81].

It is also important to note that while inhibition of certain chromatin modifiers has been shown to facilitate reprogramming, their complete knockout might result in differentiation defects. For example, transient DOT1L inhibition via small molecule inhibitors and shRNA enables increasing iPSC generation [57]. However, complete knockout of DOT1L in mice results in defects in developmental processes [82, 83].

Table 1. Barriers, Enhancers, and Essential Regulators of human iPSC generation

List of reported chromatin modifiers with their role in reprogramming is listed. Contradictory results were obtained for manipulations for MBD2 and MBD3. KD: knock-down, OE: overexpression. Table adapted from [21].

	Name	Phenotype	Classification
Chromatin Remodelers	MBD2	KD+/OE-	Barrier
	MBD2	OE+	Enhancer
	MBD3	KD+/OE-	Barrier
	MBD3	KD-/OE+	Enhancer
Histone Modifiers	BM11	KD-/OE+	Enhancer
	RING1	KD-	Positive Regulator
	LSD1	KD+	Barrier
	EHMT1	KD-	Positive Regulator
	SUV39H1	KD+	Barrier
	SUV39H2	KD+	Barrier
	JMJD1A	KD-	Positive Regulator
	JMJD1B	KD-	Positive Regulator
	JMJD2B	KD-/OE+	Enhancer
	JMJD2C	KD-	Positive Regulator
	EED	KD-	Positive Regulator
	EZH1	KD+	Barrier
	EZH2	KD-/OE+	Enhancer
	JMJD3	KD+/OE-	Barrier
	UTX	KD+	Barrier
	DOT1L	KD+	Barrier
	SIRT6	OE+	Enhancer
DNA modifiers	DNMT1	KD+	Barrier
	AID	KD-/OE+	Enhancer

1.3. CRISPR Cas9 system

Clustered regularly interspaced short palindromic repeats (CRISPR) system is a defense mechanism which protects bacteria and archea against viral infections [84, 85]. Upon infection, DNA elements, which are called spacers, from the invading virus are incorporated into the CRISPR array of the bacterial genome. The CRISPR system is orchestrated by two non-coding RNAs which are called pre-crRNA and trans-crRNA, and endonucleases which are called Cas proteins. Pre-crRNAs include spacers that guide Cas proteins to the viral DNA, and trans-crRNA enables stable binding of Cas proteins to viral DNA. Guided Cas proteins bind viral DNA upon the recognition of protospacer adjacent motifs (PAM). Binding is

followed by creation of double strand breaks (DSB) at three bp downstream from 3' end of the PAM sequences, which results in elimination of viral DNA.

This bacterial defense mechanisms has been adapted *in vitro* for creating mammalian knockout cell lines [86]. This system is modified by fusing trancrRNA and crRNA to create one chimeric guideRNA (gRNA), and using Cas9 gene of *Streptococcus pyogenes* as the endonuclease for inducing DNA cleavages in mammalian cells [86]. When Cas9 induces a DSB, mammalian cells trigger DNA repair responses namely, homology-directed repair or non-homologous end joining [85, 87]. Depending on the repair system triggered, different DNA modifications can occur on target sequence region. In the homology-directed repair, cells utilize a repair template to correct DSB cleavage which results in precise modification of DNA. On the other hand, in the non-homologous end joining, cells try to ligate two ends of the DSB without using a template leading to imperfect repair. During this process, insertion or deletion (indel) mutations can occur in DSB regions leading to change in the open reading frame which can result in unfolded or misfolded proteins or truncated proteins due to the presence of early stop codons. Thus, Cas9 induced DSBs can be utilized to generate knock-out alleles.

The genome editing power of CRISPR-Cas9 system is limited by its off-target activity [88–90]. GuideRNAs which are 20 bp in length need to be precisely selected to prevent non-specific cleavage. Although mismatches at 12 bp upstream region from 3' end of gRNAs is not tolerated by Cas9, those at 8 bp downstream region from 5' end of gRNAs can be tolerated and results in off-target cleavage [87–89]. Off-target activities of Cas9 is commonly studied and to map the Cas9 off-target activities, different methodologies have been used including targeted validation of off-target regions and CHIP-Seq experiments for Cas9 binding off-target regions[91]. Among those methods, GUIDE-Seq showed unbiased data towards identifying off-targets of this system [90]. In GUIDE-Seq cells are subjected to the CRISPR components to create knock-out in the presence of supplied DNA segments that can easily be incorporated into DSB regions generated by CRISPR system. Sequencing analysis of these incorporated regions showed that CRISPR system has high off-target activity all over the genome.

Although it has high off-target activity, methods have been developed to overcome this problem. First, mutating one catalytic domain of Cas9 to convert it to a nickase and using two paired guideRNAs to create DSBs decrease the off-target activity of the enzyme [92].

Second, Cas9 protein can be replaced with similarly working Cas proteins such as Cpf1 which has been previously shown to have greater specificity and cleavage efficiency [93]. Third, truncated versions of guideRNAs can be used to eliminate possible off-targets. These versions shorten the guideRNA sequence from 5' end. Therefore, there is less possibility for supporting Cas9 binding DNA for mismatch harboring regions [94].

In addition to creating DSB, Cas9 protein has been used for different purposes in molecular biology research. For example, fusion proteins of nuclease deficient Cas9 (dCas9) with GFP can be used as an *in situ* hybridization method [95, 96], or fusion proteins that have dCas9 and acetyltransferase domain of p300 have been proposed to acetylate histones at targeted areas [97]. CRISPR-Cas9 system is also shown to be useful for regulating gene expression in a system called CRISPR interference (CRISPRi). To knock down genes, dCas9 is targeted to the promoter regions of genes where successful binding of RNA Polymerase II is blocked, resulting in downregulation of the targeted gene. This downregulation can even further be enhanced by fusing dCas9 with repressive protein domains such as KRAB [98]. In contrast, CRISPRi can be used to overexpress genes by using dCas9 fused with transactivators to force gene expression from endogenous loci. Initial trials for such applications has been successful by using the fusion of VP64 with dCas9 and testing these constructs with reporter assays [98]. Further work showed that in addition to the dCas9, addition of MS2 aptamers to guideRNAs and MS2 aptamers to strong enhancers such as p65 and HSF1 enables robust gene overexpression from endogenous loci [99].

1.3.1. Comparison of CRISPR-Cas9 against shRNA in Arrayed and Pooled Screens

The applications of CRISPR-Cas9 system and shRNA approaches differ due to the differences in their function. First, CRISPR-Cas9 system knocks out genes by creating frame-shifts at DNA level while shRNA system interferes with translation of targeted mRNAs. Second, CRISPR-Cas9 can permanently and completely deplete protein expression whereas shRNA systems can only downregulate translated protein levels. Although the downregulation via shRNAs can deplete gene expression to a large extent, residual expression of genes might still be relevant inside the cell [100]. Furthermore, when a population of cells are infected with CRISPR-Cas9 system, a heterogeneous population is generated. Depending on the indel mutations introduced, the cells can either retain the wild-type phenotype or the gene can be knocked-out with heterogenous or homozygous modification. On the other hand, shRNAs lead to downregulation of genes in a homogenous

manner. Therefore, shRNAs enable more consistent phenotype to be seen for a targeted modification [84]. Consequently, targeted screens using these two methods might have their challenges.

The initial discovery of pooled barcoding screens using shRNA libraries enabled quantitative investigations for large gene sets [101]. In this approach, each cell is infected with viruses containing one copy of a library which creates a heterogeneous population putatively modified. Then a selective pressure is applied to this population of cells. Next-Generation-Sequencing quantifies the abundance of inserted library segments before and after the test conditions. Therefore, comparison between two samples enables quantitative understanding of gene modifications for the test. Such screens resulted in the identification of essential genes in cancer cell lines [102–104]. Moreover, pooled CRISPR-Cas9 libraries have been generated to serve in similar experimental setups [103–105]. Comparison screens have been performed to quantify specificity and accuracy of pooled shRNA and CRISPR-Cas9 libraries for identifying essential genes. Consequently, the CRISPR-Cas9 libraries for pooled screens have been shown to have less off-target activity, and less variability between biological replicates than shRNA libraries [106, 107].

1.4. Aim

In this study, my initial aim was to test whether the CRISPR-Cas9 system would be suitable for gene investigations for human somatic cell reprogramming. I utilized this system to investigate the role of 247 chromatin modifiers at different stages of human iPSC generation. To that end, a focused CRISPR-Cas9 knockout library was generated and enrichment screens were conducted using the TRA-1-60 cell surface antigen expression or exogenous EGFP silencing as markers of reprogramming. Next-Generation Sequencing of guideRNA populations in isolated emerging iPSCs was used to analyze the effects of gene perturbations on reprogramming efficiency. With this work, I anticipate to expand the knowledge of the effect of chromatin modifiers on reprogramming, fill the gaps in the literature about the stage-specific features of gene modifications for chromatin modifiers and find functional relations between chromatin modifications for generation of iPSCs. In addition to further understand reprogramming mechanism through the landscape of chromatin modifications, I aim to formulate new reprogramming strategies based on screen results by using small molecule compounds targeting chromatin modifiers and enable more deterministic and accelerated reprogramming strategies.

2 Materials and Methods

2.1 Cell Culture

HEK293T cells were maintained in D10 Medium (DMEM (Gibco) containing, 10% FBS (Gibco), and 1% Penicillin-Streptomycin (Gibco)) and passaged 1:6 in every 2 days. dH1f cells were maintained in either in D10 Medium or in α -MEM Medium (Alpha-MEM (Lonza) containing, 10% FBS (Gibco), 1% Penicillin-Streptomycin (Gibco) and 1% Non-essential amino acids (Gibco)) and passaged 1:2 every 4 days. Induced pluripotent stem cells were maintained in human ES medium (DMEM/F12 (Lonza), 20% KOSR (knock-out serum replacement) (Gibco), 10 ng/ml bFGF, 0.1 mM β -mercaptoethanol, 1% Non-essential amino acids, 1% Penicillin/Streptomycin). All cells were cultured at 37°C with 5% CO₂ in an incubator.

2.2 Generation of Modified dH1fs

PGK-EGFP positive dH1fs were generated by lentiviral infection with the PGK-EGFP-Puro vector at the indicated MOI values and selected for 4 days with 1 μ g/ml puromycin. LTR-EGFP positive dH1fs were generated by retroviral infection of pBabe-GFP-Puro or Hygro vectors at indicated MOI values and selected for 4 days or 7 days with 1 μ g/ml puromycin or 50 μ g/ml hygromycin, respectively. Cas9-expressing dH1fs were generated by lentiviral infection with lentiCas9Blast and selected for 5 days with 7 μ g/ml Blasticidin.

2.3 Reprogramming assays

50.000 dH1f cells were transduced with lentiCRISPR gRNA vectors in a medium containing 8 μ g/ml protamine sulfate (Sigma-Aldrich) at low MOI in 12-well plates. Transduced cells were selected with 1 μ g/ml puromycin for 4 days and transferred to a 6-well plate on Day 5. CRISPR-infected cells were reseeded on Day 10 and 50.000 dH1f cells infected with lentiviral reprogramming factors (pSIN4-CMV-K2M and pSIN4-EF2-O2S; 100-250 μ l supernatant/virus) in a medium containing 8 μ g/ml protamine sulfate. Six or seven days later, cells were trypsinized and re-plated at a 1/8 ratio on 75.000 mitomycin-treated mouse embryonic fibroblasts (mito-MEFs) in 12-well plates. The medium was switched to the hES media next day. Media was changed every other day until day fourteen of reprogramming after which media were changed daily until day 21. For proviral silencing experiments, α -MEM based fibroblast medium (as described in 2.1) was used. For negative selection with

puromycin experiment, cells were transferred onto vitronectin-coated plates instead of mito-MEFs, instead of hES media, E8 culture conditions used as described before [108]. gRNA sequences used in the reprogramming experiments are listed below along with their guideRNA Names.

Table 2. List of gRNAs used in the reprogramming experiments.

guideRNA_Gene Name	guideRNA Name	guideRNA Sequence	guideRNA_Gene Name	guideRNA Name	guideRNA Sequence
OCT4	OCT4 EXON4 gRNA L1	ACCCACCAATAGAACCCCC	g2_CARM1	CARM1_MTA_4	TGGAGCACGGAAAATCTACG
NT	NonTargetingControl_1	ACGGAGGCTAAGCGTCGCAA	g1_SETD2	SETD2_SET_3	ACCTTTGTCCTAGAATATTG
NT	NonTargetingControl_2	CGCTTCCGCGGCGGCTTCAA	g2_SETD2	SETD2_SET_4	GGACTGTGAACGGACAACGTG
LATS2	LATS2 gRNA #1	TCGGTTCAGGGGCTACCCGC	g1_SFMBT1	SFMBT1_1	GGTCCACAGCAGTTCCCTAT
NR2F2	NR2F2 gRNA #2	CGGGCGTCTGTGGCGTGTGC	g2_SFMBT1	SFMBT1_3	CATAGCCATCATAGCGGAGA
H2AFY	H2AFY gRNA #1	CGTGTACATGGCCGCGTCC	g1_PRMT6	PRMT6_SAM_2	CTCTACCGCGTACACGCGCC
H2AFY2	H2AFY2 gRNA #3	AATTGGCCGGCAATGCCGCG	g2_PRMT6	PRMT6_SAM_3	CCGGCGCGGTACGCGGTAG
YY1	YY1 gRNA L1	ATGTAGAGGGTGTGCCCCG	g1_KDM5C	KDM5C_JMJC_3	AAGTGAGGGC ACCCCATACC
DOT1L	DOT1L_HGLibA_13805	CTGAGCCCGCGCTCTACCCG	g2_KDM5C	KDM5C_PHD_3	TCTGCCGGATGTGTTCTCGA
P21	p21 gRNA #3	GAGAAGATCAGCCGCGT	g1_EP300	EP300_RING_1	CTGTAATAAGTGGCATC ACG
P53	p53 gRNA #3	CCATTGTTCAATATCGTCCG	g2_EP300	EP300_RING_3	CCACAGACACTGTGTTGCTA
P16	p16 gRNA #3	TCCCGGGCAGCGTCTGTCAC	g1_PBRM1	PBRM1_BAH_1	GGCCTGGTGCGTCTCTGCTG
g1_DNMT3A	DNMT3A_DMTA_2	AAAGAGAGACAGC ACCCGGA	g1_BRWD1	BRWD1_BROMO_2	ATTTAGACAACCTGTTGATT
g2_DNMT3A	DNMT3A_DMTA_5	GGACCGCTACATTGCCTCGG	g2_BRWD1	BRWD1_BROMO_4	CAACTGGAAGAAACAGTGTA
g1_HDAC10	HDAC10_HDAC_1	AGGACGCTCGATCTCGCACT	g1_TAF1	TAF1_BROMO_3	AACAGACGTTCAATCTCC
g2_HDAC10	HDAC10_HDAC_5	CGCGCTGCCGCGAGGCGATCC	g2_TAF1	TAF1_GECKO_1	AAGAATCACCTGAGTGGCAA
g1_PRDM3	PRDM3_SET_1	GAAAGACCCAGTTATGGAT	g1_SMARCA5	SMARCA5_DEXD_1	ACCAAAACCATATGAGGCC
g2_PRDM3	PRDM3_SET_2	GTCAACCAGATGTTGGAAGC	g2_SMARCA5	SMARCA5_HELIC_2	AAATTATGAGTACTGCAGGT
g1_KMT2A	KMT2A_BROMO_1	CCGGTAGCGTAGCAAATGGC	g1_SMYD5	SMYD5_SET2_5	CTGCATTGGGCACACAACGTG
g2_KMT2A	KMT2A_BROMO_5	CAGCAGCCTTTAGATCTAGA	g1_ACTR	ACTR_PAS_3	CGGATTATATCTTCAAAGCC
g1_DNMT3L	DNMT3L_3	ACACAGCACCCCTCTTTGA	g2_ACTR	ACTR_PAS_5	GTCTAAATGATGGGCAGTCA
g2_DNMT3L	DNMT3L_5	CAGATCCCTCCCTCAAACAG	g1_KDM8	KDM8_CUPIN_1	AQTGGTCAGCCACGCCTTTC
g1_CARM1	CARM1_MTA_3	CTTGGGCGGCAAAAACGAC	g2_KDM8	KDM8_CUPIN_5	CGTTGACCGTCATGAGGGTC

Small molecules were used to test the effects on fibroblast to iPSC conversion. Chemicals were added to the media according to the indicated time period.

Table 3. List of used small molecule compounds.

Small Molecule	Source	Concentration	Solvent	Target
EPZ 4777	Tocris Bioscience	3 μ M	DMSO	DOT1L inhibitor
EPZ 5676	Tocris Bioscience	3 μ M	DMSO	DOT1L inhibitor
EPZ 6438	Tocris Bioscience	3 μ M	DMSO	EZH2 inhibitor
MS-275	Cayman Chemicals	0.1 μ M	DMSO	HDAC1, 3 inhibitor
Bromosporine	Tocris Bioscience	0.5 μ M	DMSO	BRD1, 2, 4 inhibitor

2.4 TRA-1-60 HRP Immunostaining

Cells were washed once with PBS and fixed with 4% PFA for 10-15 min at room temperature followed by 5 washes with PBS. Biotin conjugated anti-human TRA-1-60 (BioLegend Catalog #330604) was diluted 1:250 in staining buffer and added onto wells (500 μ l for 6-well, 300 μ l for a 12-well plate, 100 μ l for a 24-well plate). Plates were incubated overnight at 4°C on an orbital shaker. Next day, samples were washed with 5 times with PBS and Streptavidin-HRP (BioLegend Catalog #405210) secondary antibody diluted 1:500 in staining buffer was added onto wells. Plates were incubated for 2 hours at room temperature.

Samples were washed again 5 times with PBS. The stain was developed with Vector Labs DAB kit for 10 min at RT (Development solution: 5ml dH₂O, add 2 drops of the buffer, mix, add 4 drops of DAB, mix, add 2 drops of hydrogen peroxide, mix, add 2 drops of nickel solution, mix). Development solution was washed away 2 times. with PBS. Stained cells were left in PBS at RT for at least one day. Cream was added to the wells to provide a contrast for scanning. Colony counts were quantified by ImageJ software.

2.5 TRA-1-60-PE staining for Flow Cytometry

Cells were harvested by using 0.05% trypsin which was inactivated with 1X PBS supplemented with 10% FBS and was collected in eppendorfs. Eppendorfs were spun washed twice in 1X PBS supplemented with 1% FBS at 1500 rpm for 10 minutes. TRA-1-60-PE antibody (BioLegend Catalog #330610) was added to the pellet in 100 μ l 1X PBS supplemented with 1% FBS depending on the reprogramming day analysis and colony formation efficiency different concentration of antibody used. For DMSO treated cells, 0.12, 0.25, 0.5 μ l antibody added for Day 7, 14, 21, respectively. For 3 μ M EPZ5676 treated cells, 0.25, 0.5, 1 μ l antibody added for Day 7, 14, 21, respectively. If wells are mostly consist of iPSCs 1.5 μ l of antibody added at Day 21. Samples were incubated on ice for 1 hour in a dark condition. Samples were flicked 1-2 times during incubation to prevent pellet formation. Cells were spun washed twice in 1X PBS supplemented with 1% FBS at 1500 rpm for 10 minutes. Cells were resuspended in at least 400 μ l 1X PBS supplemented with 1% FBS. Afterward, samples were passed through cell strainer flow tubes. BD Accuri C6 Flow Cytometer (BD Biosciences) used to calculate percentage of stained cells. Compensation applied if necessary.

2.6 DNA Constructs

pBabe-GFP-Puro was cloned by digesting pWZL-Blast-EGFP with EcoRI, gel purifying EGFP sequence (~800 bp) by Macherey-Nagel (MN) PCR-Clean Up kit, and ligating it to the EcoRI digested and Antarctic Phosphatase (AP) (NEB) treated pBabe-Puro. All other plasmids were purchased from Addgene.

Table 4. List of Addgene plasmids used in this study

Plasmid Name	Addgene Catalog No
lentiCRISPRv1	#49535 (not available in Addgene anymore)
lentiCRISPRv2	#52691
lenti sgRNA(MS2)_zeo	#61427
lentiGuide- Puro	#52693
lentiCas9_Blast	#52962
lentiCas9-VP64_Blast	#61425
lentiMS2-P65-HSF1_Hygro	#61426
pSIN4-CMV-K2M	#21164
pSIN4-EF2-O2S	#21162
pBabe GFP Hygro	#61215
lenti PGK GFP Puro	#19070
pWZL Blast EGFP	#12269
pBabe Puro	#1764
pUMVC	#8449
pCMV-dR8.2 dvpr	#8455
pCMV-VSV-G	#8454

2.7 Cloning

lentiCRISPRv1, lentiCRISPRv2, lenti sgRNA(MS2)_zeo, lentiGuide vectors were digested with BsmBI (NEB) and gel purified with MN PCR-Clean Up Kit. Extracted plasmids were treated with AP. GuideRNAs were designed either by <http://crispr.mit.edu/> [87] or taken from GeCKO and SAM Libraries. 1 µl of each oligo (100 uM) was annealed with T4 Polynucleotide Kinase (3' phosphatase minus, NEB) and T4 DNA Ligase Buffer (NEB) in 10 µl reaction in a thermal cycler with the following settings: 37°C for 30 min, 95°C for 4 min and then ramp down to 25°C at 5°C/min. Oligos were diluted to 1:200 were used as inserts for cloning. 20 µl ligation reactions were prepared with 50 ng BsmBI-digested, AP-treated vectors, 1 µl of inserts and Quick Ligase (NEB) or T4 DNA Ligase (NEB) in 2X Quick Ligase Buffer or 10X T4 DNA Ligase buffer (NEB), respectively. Ligation reaction were incubated for at least 10 minutes at RT. Ligation mix was vortexed and spun down. 50 µl Stbl3 bacteria was mixed with 5 µl of the ligation reaction and left on ice for at least 5 minutes (longer durations increased cloning efficiency). Afterward, heat shock was applied at 42°C for 30 seconds in water bath or 48°C for 40 seconds in heat-block and eppendorfs were placed on ice for 5 minutes. 150 µl LB was supplemented and cultured in 37°C at 225 rpm in a shaker for 1 hour. 100 µl of the reaction was spread on Ampicillin or Carbenicillin

resistance plates. U6 promoter sequencing (Forward sequencing primer for U6 promoter ACTATCATATGCTTACCGTAAC) was used to confirm cloned gRNAs listed in the below table.

Table 5. List of gRNAs cloned into the lentiCRISPRv1

gRNA Name	GuideRNA Sequence	gRNA Name	GuideRNA Sequence
RNF40 gRNA #1	AACAAACGCGCCGCCGCGCA	YY1_gRNA_L1	ATGTAGAGGGTGTCGCCCCG
RNF40 gRNA #2	CCGGCGGCGCGTTTGTGGCC	YY1_gRNA_L2	TCGGGTCGTCGGTGACCAG
RNF40 gRNA #3	TCTTGAAGTGTAGTACCTTC	YY1_gRNA_3	TGGGCGGCGACGACTCGGA
p53 gRNA #1	TCGACGCTAGGATCTGACTG	ZBTB8A_guide_1	AAATGGACTCTACTCCTGT
p53 gRNA #2	CTCAGAGGGGGCTCGACGCT	ZBTB8A_guide 2	ACTCCTGTTGGCTATCAGTA
p53 gRNA #3	CCATTGTCAATATCGTCCG	ZBTB8A_guide 3	TGACCGTACTGATAGCCAAC
NR2F2 gRNA #1	CAGCACGTGGCGCGACCCCC	NonTargetingControl_1	ACGGAGGCTAAGCGTCGCAA
NR2F2 gRNA #2	CGGGCGTCTGTGGCGTGTGC	NT_2	CGCTTCGCGGCCCGTTCAA
NR2F2 gRNA #3	TATATCCGGACAGGTACGAG	gRNA T1	TGAACCGCATCGAGCTGAA
LATS2 gRNA #1	TCGGTTCAGGGGCTACCCGC	gRNA T2	GAGCGCACCATCTTCTTCA
LATS2 gRNA #2	CGACTGCAAGAGATTCGTGA	MAPKBP1	CGTTGACGATCATGTATGC
LATS2 gRNA #3	TCATTAAGCAGACCTCCCC	CARM1	CCCGTACTCACGGCTGTAGA
H2AFY gRNA #1	CGTGACATGCCCGCGTCC	ZBTB8A	TGACCGTACTGATAGCCAAC
H2AFY gRNA #2	AGGACGGCGGCCATGTACAC	GPATCH1	CTCACTTTGCTGGCGTTATG
H2AFY gRNA #3	AGAAGGGACGGGTCACACCC	ANKRD40	CCCGGGGAAAGTCCCTAAC
H2AFY2 gRNA #1	TTCAAGTACCGGATCAGCG	CDKN2A	GACCCGTGCACGACGCTGCC
H2AFY2 gRNA #2	TTCAAGTACCGGATCAGCGT	NKG7	GACGCAGACCTTCAGCATT
H2AFY2 gRNA #3	AATTGGCCGGCAATGCCGCG	LYRM5	TACTTCTCCTCTTAAAGAAT
p16 gRNA #1	ACCGTAAGTATTCGGTGCCT	TRIM15	TCTGAGGAAGTCACCCGGCT
p16 gRNA #2	CACCGAATAGTTACGGTCGG	YY1	GATGTAGAGGGTGTCGCCCCG
p16 gRNA #3	TCCCGGGCAGCGTCGTGCAC	WIPF1	ACGCCCCCGTTACCTTCGCC
p21 gRNA #1	CCGCGACTGTGATGCGCTAA	BRD8	TCTTGCTTGACCGTCATTTT
p21 gRNA #2	CGACTGTGATGCGCTAATGG	hsa-mir-181b-1	CTGAACAATGAATGCAACTG
p21 gRNA #3	GAGAAGATCAGCCGGCGTT	DPH3P1	AGTCCATAACATGACGTATC
DOT1L,HGLibA_13805	CTGAGCCCCCGTCTACCCG	ZNF701	GTGCACATATTTACCCCGA
TRIM28_gRNA_L1	TACCAGTAGAGCGCACAGTA	SNX11	TTCATTGAGAAGCGACGACA
TRIM28_gRNA_L2	ACGCACCTGGTTCGCATCC	SEPTIN4	CTCTACCTCAGAGGAATCAT
TRIM28_gRNA_L3	ACGTTACCATCCCGAGACT	PRRC2A	TCCGGGCCCAGACGACTGTT
OCT4_EXON4_gRNA_L1	ACCCACCAAATAGAACCCCC	WNK3	ATCATAAAATCGAACTATAT
OCT4_EXON1_gRNA_L2	CCAAGCTCCTGAAGCAGAAG	hsa-mir-5684	ATGGTCTGTGTTTGTGCCCC
OCT4_EXON3_gRNA_3	CAAGTATCACCCCAAGTTTA	RNF31	TTGACACCACGCCAGTACCG
HEXIM1_gRNA_L1	CGGCGGGAAGTCGCCGCG	P2RX3	GCTCTTCAACAACCCGACT
HEXIM1_gRNA_L2	AATGGGGACGACTCGTCCGC	GLTSCR2	TATCCACAGACCCCTGGAC
HEXIM1_gRNA_3	AGAAGTGAACCGACGCCCG		

Table 6. List of gRNAs cloned into the lentiGuide

gRNA Name	GuideRNA Sequence
gRNA T1	TGAACCGCATCGAGCTGAA
gRNA T2	GAGCGCACCATCTTCTTCA

Table 7. List of gRNAs cloned into the lenti sgRNA(MS2)_zeo

gRNA Name	GuideRNA Sequence
KLF4_SAM1_T	GATGGAAGGGAGCCTCGGGG
KLF4_SAM2_T	GGCAGCTAAATCAACAAACT
SOX2_SAM1_T	GGCGGGGCCTCCCGCGCCGC
SOX2_SAM2_T	GCCGGGTTTTCATGAAAGG
NANOG_SAM1_T	CCCGTCTACCAGTCTCACCA
NANOG_SAM2_T	TGTCTTCAGGTTCTGTTGCT
LIN28A_SAM1_T	TTGTGGGGGAGGGCCGGAGC
LIN28A_SAM2_T	TCCGGTCTCTGACACCTCTG
MYC_SAM1_T	AGCTAGAGTGCTCGGCTGCC
MYC_SAM2_T	GAACCCGGGAGGGGCGCTTA
POU5F1_SAM1_T	GGAAAACCGGGAGACACAAC
POU5F1_SAM2_T	GGATGTTTGCTAATGGTGG
NonTargetingControl_1	ACGGAGGCTAAGCGTCGCAA
NonTargetingControl_2	CGCTTCCGCGGCCCGTTCAA

Table 8. List of gRNAs cloned into the lentiCRISPRv2

gRNA Name	GuideRNA Sequence	gRNA Name	GuideRNA Sequence
PAXIP1_BRCT_2	ATTTCTTATTGAGGGTTAGC	PBRM1_BAH_1	GGCCTGGTGCGTCCTCGTGT
PAXIP1_BRCT_4	AGAACGTAACCAAAGCCCAC	PBRM1_BAH_4	GTGTGCTGTGTTGTCATTCA
BRD2_BROMO_1	ATCAGTTCGCATGGCCATTC	PRMT6_SAM_2	CTCTACCGCGTACACGCGCC
BRD2_BROMO_3	CCTTGTGTAGGTATTGCAGC	PRMT6_SAM_3	CCGGCGCGTGTACGCGGTAG
USP22_PC19_2	AGTCCCGCAGAAGTGGCGTG	SFMBT1_1	GGTCCACAGCAGTTCCCTAT
USP22_PC19_5	GCAACCCGCTGTGAAGATC	SFMBT1_3	CATAGCCATCATAGCGGAGA
CREBBP_BROMO_3	GCCGGAAAGGTAATGACTCT	KDM8_CUPIN_1	AGTGGTCAGCCACGCCTTTC
CREBBP_BROMO_4	TGCCGGAAAGGTAATGACTC	KDM8_CUPIN_5	CGTTGACCGTCATGAGGGTC
CREBBP_KAT_5	ACAAAAACCTCCCCGGCTTC	ACTR_PAS_3	CGGATTATATCTTCAAAGCC
PRDM11_SET_1	CACATCTTCGGCCCCCTATGA	ACTR_PAS_5	GTCTAAATGATGGGCAGTCA
PRDM11_SET_3	GGTGGAGATCTGCCCCCTCAT	SMARCA5_DEXD_1	ACCAAAACCATATGAGGCCC
CARM1_MTA_3	CTTGGGCGGCAAAAAACGAC	SMARCA5_HELIC_2	AAATTATGAGTACTGCAGGT
CARM1_MTA_4	TGGAGCACGGAATCTACG	SMYD5_SET2_3	CTGCATTGGGCACACAACCTG
DNMT3A_DMTA_2	AAAGAGAGACAGCACCCGGA	SMYD5_SET2_5	TGGAAAGGAGGTCTCTGCAT
DNMT3A_DMTA_5	GGACCGCTACATTGCCTCGG	TAF1_BROMO_3	AACAGACGTTCAATCTCC
DNMT3L_3	ACACAGCACCTCTGTTTGA	TAF1_GECKO_1	AAGAATCACCTGAGTGGCAA
DNMT3L_5	CAGATCCCTCCCTCAAACAG	SETD2_SET_3	ACCTTTGTCCTAGAATATTG
EP300_RING_1	CTGTAATAAGTGGCATCACG	SETD2_SET_4	GGACTGTGAACGGACAACCTG
EP300_RING_3	CCACAGACACTGTGTTGCTA	PRDM3_SET_1	GAAAGACCCCACTTATGGAT
HDAC10_HDAC_1	AGGACGCTCGATCTCGCACT	PRDM3_SET_2	GTCAACCAGATGTTGGAAGC
HDAC10_HDAC_5	CGCGCTGCCGAGGCGATCC	BRWD1_BROMO_2	ATTTAGACAACCTGTTGATT
KDM5C_JMJC_3	AAGTGAGGGCACCCCATACC	BRWD1_BROMO_4	CAACTGGAAGAAACAGTGTA
KDM5C_PHD_3	TCTGCCGATGTGTTCTCGA	NonTargetingControl_1	ACGGAGGCTAAGCGTCGCAA
KMT2A_BROMO_1	CCGGTAGCGTAGCAAATGGC	NonTargetingControl_2	CGCTTCCGCGGCCCGTTCAA
KMT2A_BROMO_5	CAGCAGCCTTTAGATCTAGA		

2.8 cDNA synthesis and qPCR

MN NucleoSpin RNA isolation kit used to isolate RNA from collected cells. For cDNA synthesis, 16.5 µl reactions were prepared in a PCR tube by mixing 200 ng to 500 ng RNA, 2.5 µl 2 mM dNTP (Life Technologies) and 1 µl 50 µM random hexamer mix (Invitrogen). The mixture was incubated at 65°C for 5 min and briefly chilled on ice. Reverse transcriptase mix prepared by mixing 5 µl 5X First Strand Buffer (Invitrogen), 2 µl 0.1M DTT (Invitrogen) and 0.5 µl RNasin (Promega) and added to the samples, and incubated for 10 min at RT. 1 µl of MMLV-RT enzyme (Invitrogen) was added and cDNA was synthesized at 37°C for 1 hour and followed by the inactivation of the enzyme by incubating the samples at 70°C for 15 min. The synthesized cDNA from 250 ng or 500 ng was diluted to 50 µl or 100 µl with nuclease free water. 20 µl PCR reaction prepared in 96-well opaque plates (Roche).

10 µl of LightCycler 480 SYBR Green I Master Mix (Roche), 2 µl of cDNA, 2 µl of 2.5 mM forward and reverse primer mix, and 6 µl of nuclease free water. The threshold cycle (Ct) values determined by LightCycler 480 Instrument II (Roche) according to the manufacturer's instructions. All quantifications were normalized to the indicated internal controls. Differential gene expression was quantified by $2^{-(C_t - C_c)}$ (C_t and C_c represents the mean threshold cycle differences for target sequence and control sequence, respectively).

Table 9. List of primers for qRT-PCR used in the study.

Gene	Forward primer sequence	Reverse primer sequence
β-ACTIN	TGAAGTGTGACGTGGACATC	GGAGGAGCAATGATCTTGAT
E-CADHERIN	GAGTCACCCGGTTCATCTA	GGTGTGGGAGTGCAATTTCT
DNMT3B	ATGGCAAGTTCTGCGAGGTCT	CGCACCTTCGACTTATTAACCA
NANOG	TGATTTGTGGGCTGAAGAAA	TGGTGGTAGGAAGAGTAAAG
LIN28A	TGCGGGCATCTGTAAGTGG	GGAACCCCTCCATGTGCAG
ZEB1	GATGATGAATGCCAGTCAGATGC	ACAGCAGTGTCTTGTGTTGTAG
GREM1	AAGGGAAAAAGAAAGGGTCCCA	GTCTGCTCTGAGTCATTGTGC
COL3A1	TGGTCCCCAAGGTGTCAAAG	GGGGGTCTGGGTTACCATTA
POSTN	TAAGTTTGTTCTGTTAGCACC	GTGTGGGTCCTTCAGTTTGTATA
DKK1	ATAGCACCTTGATGGGTATTCC	CTGATGACCGAGACAAACAG
KLF4	GATGAACTGACCAGGCAGTA	GTGGGTCATATCCACTGTCT
C-MYC	TGCCTCAAATTGGACTTTGG	GATAATTCTGTGTAAGTGC

Thermal cycler condition for qRT-PCR was set as below:

Step	Temperature	Time
Initial Denaturation	95°C	3 min.
40x cycle	95°C	10 sec.
	60°C	30 sec.
	72°C	30 sec.
Final Extension	72°C	5 min.

2.9 T7 Endonuclease Assay

Genomic DNA (gDNA) was isolated with MN Nucleospin Tissue kit. PCR reaction was set with primers that amplify cleavage sequence introduced by CRISPR-Cas9 infections from genomic DNA. At least 600 ng PCR product was used to form heteroduplex according to the protocol of 5 minutes 95°C; ramp down to 85°C at -2°C/s; ramp down to 25°C at -0.1°C /s; hold at 4°C. 1 µl T7 Endonuclease (NEB) was added in the reaction and samples heated at 37

°C for 45 minutes. Reaction was inactivated by addition of 2 µl of 0.25 EDTA and visualized by gel electrophoresis.

Table 10. List of Tested Surveyor Primers for gRNAs

gRNA Name	Forward Surveyor Primer	Reverse Surveyor Primer	Expected PCR Product	Expected Cleavage Product 1	Expected Cleavage Product 2
RNF40 gRNA #1	TCTGTGATTGGTTGTGGGC	CAGCTGGAAGGAGAAAACATCA	324	206	118
RNF40 gRNA #2	TCTGTGATTGGTTGTGGGC	CAGCTGGAAGGAGAAAACATCA	324	216	108
RNF40 gRNA #3	AACGTGAGCCTCCTCACAAC	GCAGGGGAATCTGGGGATTG	385	236	149
p53 gRNA #1	TCAGACACTGGCATGGTGTT	GGTGAAAAGAGCAGTCAGAGG	405	160	245
p53 gRNA #2	TCAGACACTGGCATGGTGTT	GGTGAAAAGAGCAGTCAGAGG	405	166	239
p53 gRNA #3	TGGTCCTCTGACTGCTCTTTT	GCATTGAAGTCTCATGGAAGCC	312	208	104
NR2F2 gRNA #1	AACCAACAAAGCTGAGCCGA	CTGCTGCTGCTGCTTGTC	460	154	306
NR2F2 gRNA #2	CTAACCAAGCTCAACCAACAG	CACGCTGCGCTTGAAGAAG	456	298	158
NR2F2 gRNA #3	AAAGGAATGGTGCCAAGAGC	GCCAGTTCGCAATGTTCCTG	336	265	71
LATS2 gRNA #1	ACTGGACTAACAAATGAGGCCAA	CTGTACCCACCTGCTGCG	325	61	264
LATS2 gRNA #2	ACTGGACTAACAAATGAGGCCAA	CTGTACCCACCTGCTGCG	325	108	217
LATS2 gRNA #3	GCCCTGGAGTACATCAGCAA	AGACCATTTTTACGCCCGCT	376	71	305
H2AFY gRNA #1	CCTTCCAGGGGACTAGAGTGT	ATGCACACACACAATGAGGGG	459	271	188
H2AFY gRNA #2	CCTTCCAGGGGACTAGAGTGT	ATGCACACACACAATGAGGGG	459	281	178
H2AFY gRNA #3	TCCCAAAACAGGGCATTGGA	TTGGGCAGTATGACCACCTG	335	262	73
H2AFY2 gRNA #1	GCAGTCAGAGGTGTCACAGT	AATTTGAGCATCCCTCCCC	381	272	109
H2AFY2 gRNA #2	GCAGTCAGAGGTGTCACAGT	AATTTGAGCATCCCTCCCC	381	274	107
H2AFY2 gRNA #3	TTCTTTCCAGGGCCGAGACAT	TGAGATCTATGCCAGAGGCAAT	395	135	260
p16 gRNA #1	GTCCCTCCAGAGGATTGAGG	GGAAGCCTCCCTTTTTCGG	388	282	106
p16 gRNA #2	GTCCCTCCAGAGGATTGAGG	GGAAGCCTCCCTTTTTCGG	388	280	108
p16 gRNA #3	CCTGGCTCTGACCATCTCTGTT	GCTTTGGAAGCTCTCAGGGTA	345	128	217
p21 gRNA #1	GACCAAGGGCCTTCCTTGAT	TGTCGGCACCTGTATGCT	452	145	307
p21 gRNA #2	GACCAAGGGCCTTCCTTGAT	TGTCGGCACCTGTATGCT	452	140	312
p21 gRNA #3	TGGACGTTCCCGAGTTCTT	CAACCTCCGGGAGAGAGGAA	459	164	295
DOT1L_HGLbA_13805	CTCATTGTGCTCGCTTCACG	GAGCAGCGACAGTGGAGTC	831	317	514
OCT4_EXON4_gRNA_L1	GATGCTCTGCATCCCTTGGA	CATGGGTGAGGGTAGTCTGC	480	395	85
OCT4_EXON1_gRNA_L2	TCCCTTCGCAAGCCCTCATTT	TTGCCTCTCACTCGGTCTCTC	787	532	255
OCT4_EXON3_gRNA_3	CGCTGGCTTATAGAAGGTGCT	TGTGCATAGTCGCTGCTTGA	1000	350	650
HEXIM1_gRNA_L1	CCTGGAATCCCAACCACTC	CTTGAGTACAGGCGGGTTT	484	392	92
HEXIM1_gRNA_L2	CCTGGAATCCCAACCACTC	CTTGAGTACAGGCGGGTTT	484	379	105
HEXIM1_gRNA_3	CCTGGAATCCCAACCACTC	CTTGAGTACAGGCGGGTTT	484	335	149
YY1_gRNA_L1	CTCTACATGCCACGGACG	GAATCTGATCCTCGAAGCCGT	493	400	93
YY1_gRNA_L2	CTCGGGTTGAGGAGTTGGTAT	TGCTCCCACTTCTTTTGCC	376	270	106
YY1_gRNA_3	GGAGACCATCGAGACCACAG	AGCTTTGCCCACTTCTGCAC	500	260	240

2.10 Production of Viruses

1 x 10⁶, 2.5 x 10⁶, 7 x 10⁶ HEK 293T cells were seeded on 6 cm, 10 cm, 15 cm plates, respectively. One day later, cells were transfected with 100, 225, 750 ng plasmid encoding the envelope protein (VSVG), 1000 ng, 2250 ng, 6750 ng Gag-Pol (pUMVC for retroviruses and 8.2DeltaVPR for lentiviruses) and 1 µg, 2.5 µg, 7.5 µg viral vector in a total volume of 100, 250, 500 µl using 6, 15, 45 µl FuGENE® 6 Transfection Reagent (Promega) in a total volume of 100, 250, 500 µl serum-free DMEM or Opti-MEM according to the manufacturer's instructions. After at least 8 hours, the transfection medium was replaced with fresh 3, 8, 20 ml 293T media. Viruses were harvested at 48 and 72 hours after transfection and stored in 4°C. Viral supernatants from the two collections were combined and filtered through a 0.45 µm low protein binding syringe filter to clear viruses from cells and debris.

2.11 Concentration of Viruses

PEG-8000 (Sigma) was dissolved in 1x PBS to the final concentration of 10% (w/v). Collected and filtered viruses were mixed with 1/5 ratio of PEG and refrigerated overnight. Supernatant/PEG mixture was centrifuged at 1500 g for 30 minutes at 4°C. Supernatant was decanted to the bleached bottle. Virus containing pellets was spun down again at 1500 g for 5 minutes. Remaining supernatant was removed with pipettes to a bleached bottle. Pellet was resuspended with 1/10 or 1/100 of the original volume of viral supernatant by using cold 1X PBS. Viruses were aliquoted and stored at -80.

2.12 Titering of Viruses

For EGFP expressing vectors, 100.000 dH1fs were infected with different volume of viruses and EGFP percentage assessed after 48 hours with flow cytometry. The titer of the virus was calculated by multiplying EGFP positive cell percentage with the total number of initially infected cells and then dividing this number by the volume of the virus used. Same calculations done for at least two different volume of viruses that showed a linear relationship between the volume of virus used and EGFP positive cells.

For vectors expressing antibiotic resistance genes, 100.000 dH1fs were infected with different volume of viruses and selected with antibiotics. Two wells of dH1fs were left uninfected for further use. One of the uninfected wells was used as a positive control for antibiotic selection and when antibiotic kills all the cells in this well, selected cells in each infected well were counted by haemocytometer. Percentage of transduced and selected cell to uninfected and non-antibiotic selected cell counts estimated. This value was used to estimate initial MOI infection of each volume according to the Poisson distribution. The titer of the virus was calculated by multiplying estimated MOI values with the total number of initially infected cells and then dividing this number by the volume of the virus used. Same calculations done for at least two different volume of viruses that showed linear correlation in between volume of viruses and counted cells.

2.13 Flow Cytometry

For the EGFP analysis, cells were harvested by trypsinization and washed twice with 1X PBS supplemented with 1% FBS. Cells were centrifuged at 1500 rpm for 5 min and resuspended in 1x PBS supplemented with 1% FBS. BD Accuri C6 Flow Cytometer (BD Biosciences) was used to analyze the samples according to the manufacturer's instructions.

2.14 Fluorescence Activated Cell Sorting (FACS)

Cells were harvested by trypsinization and washed at least one time with 1x PBS supplemented with 1% FBS for least one time. Same protocols for TRA-1-60-PE staining or proviral silencing experiments conducted. BD FACSARIA III was used to sort cells according to manufacturer's instructions.

2.15 Library Design

gRNAs were designed to target DNA sequences encoding the catalytic or critical domains of 247 chromatin modifiers. For this, we manually collected DNA sequences of protein domains from NCBI Unigene software and used CCTop to design gRNAs[109]. 3 bp mismatches at off-target sequences were allowed in the 8 bp downstream from 5' end of guideRNA sequences. In cases where no gRNA with these criteria were found, gRNA sequences from GeCKO library that target 5' first coding exon of genes were taken [105]. For each gene, 5-10 gRNAs were designed and 80 non-targeting gRNAs from GeCKO Library was added to the library.

2.16 Oligo array synthesis and pooled library cloning

DNA oligonucleotide array was synthesized by the OligoMix (LC Sciences). Oligos harboring homologous arms to target vectors and gRNA sequences were amplified in a 50 µl PCR reaction by mixing 10 µl 5X HF Buffer, 1µl 10 mM dNTP, 2.5 µl 10 uM forward primer, 2.5 µl 10 uM reverse primer, 0.1 µl oligo mix template, 0.5 µl high-fidelity polymerases and filled with NF water. Thermal cycler condition for PCR was set as below:

Step	Temperature	Time
Initial Denaturation	98°C	30 seconds
20x cycle	98°C	10 seconds
	63°C	20 seconds
	72°C	15 seconds
Final Extension	72°C	3 minutes
Hold	4°C	

The lentiCRISPRv2 vector was digested with BsmBI, treated with AP and gel-purified. Inserts ligated to prepared lentiCRISPRv2 plasmids with Gibson Assembly (NEB) in 1:5 ratio and transformed with EnduraTM electrocompetent cells in seven parallel electroporation and plated on large bioassay petri plates (Corning). Total $\sim 1.4 \times 10^6$ colonies collected from

seven plates to have 800x library coverage, and pooled plasmids isolated using maxi-prep kit (Qiagen).

GuideRNA regions of isolated plasmids were amplified by PCR to confirm cloning by Illumina Sequencing. Sequencing was performed on a MiSeq at MIT BioMicro Center. PCR reaction prepared in 50 μ l reactions by mixing 10 μ l 5x HF Buffer, 1 μ l dNTP (10 mM), 1 μ l HF Polymerase, 2 μ l forward staggered primer mix, 2 μ l reverse primer with index, 10ng library plasmid pool template. Thermal cycler condition for PCR is set as below:

Step	Temperature	Time
Initial Denaturation	98°C	30 seconds
16x cycle	98°C	10 seconds
	63°C	15 seconds
	72°C	20 seconds
Final Extension	72°C	4 minutes
Hold	4°C	

2.17 Genomic DNA Sequencing

Genomic DNA of samples were isolated with MN Nucleospin Tissue kit by following regular protocol. 100 μ l external and internal PCR reactions were prepared by mixing 2 μ l dNTP (10 mM), 1 μ l Phusion Polymerase, 20 μ l 6X GC Buffer, 5 μ l Forward Primer (10 μ M), 5 μ l Reverse Primer (10 μ M), 1 μ g gDNA (to ensure at least 100x library coverage) and filled with NF water. For internal PCR reaction, primers were changed accordingly, and 2 μ l external PCR product without clean-up directly was added instead of 1 μ g gDNA. Thermal cycler condition for PCR is set as below:

Step	Temperature	Time
Initial Denaturation	95°C	3 minutes
17x cycle (external PCR) 23x cycle (internal PCR)	95°C	30 seconds
	65°C	20 seconds
	72°C	15 seconds
Final Extension	72°C	3 minutes
Hold	4°C	

Samples run on a gel and ~350 bp bands were extracted by MN PCR-Clean-up Kit. Samples sequenced with Illumina Sequencing, HiSeq 4000 to obtain maximum coverage at Berkeley Bioinformatic Core Center.

External and internal PCR primers are listed below. Forward staggered primers were pooled to serve as internal forward primers, and for each sample different reverse primer that has different indexes used.

Table 11. List of external and internal primers.

External Forward Primer	AATGGACTATCATATGCTTACCGTAACCTGAAAGTATTTTCG
External Reverse Primer	TCTACTATTCTTTCCCTGCACTGTGTGGCGATGTGGCTCTG
Forward_stag1	AATGATACGGCGACCAACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTCTGTTGAAAGGACGAAACACCG
Forward_stag2	AATGATACGGCGACCAACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTCTGTTGAAAGGACGAAACACCG
Forward_stag3	AATGATACGGCGACCAACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTCTGTTGAAAGGACGAAACACCG
Forward_stag4	AATGATACGGCGACCAACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTCTGTTGAAAGGACGAAACACCG
Forward_stag5	AATGATACGGCGACCAACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTCTGTTGAAAGGACGAAACACCG
Forward_stag6	AATGATACGGCGACCAACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTCTGTTGAAAGGACGAAACACCG
Forward_stag7	AATGATACGGCGACCAACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTCTGTTGAAAGGACGAAACACCG
Forward_stag8	AATGATACGGCGACCAACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTCTGTTGAAAGGACGAAACACCG
Forward_stag9	AATGATACGGCGACCAACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTCTGTTGAAAGGACGAAACACCG
rev_index1	CAAGCAGAAGACGGCATACGAGATCGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT
rev_index2	CAAGCAGAAGACGGCATACGAGATCATCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT
rev_index4	CAAGCAGAAGACGGCATACGAGATTGGTCAGTGAAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT
rev_index5	CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT
rev_index6	CAAGCAGAAGACGGCATACGAGATATTGGCGTGAAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT
rev_index7	CAAGCAGAAGACGGCATACGAGATGATCTGGTGAAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT
rev_index8	CAAGCAGAAGACGGCATACGAGATTCAAGTGTGAAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT
rev_index12	CAAGCAGAAGACGGCATACGAGATTACAAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT
rev_index9	CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT
rev_index10	CAAGCAGAAGACGGCATACGAGATAGCTAGTGAAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT
rev_index11	CAAGCAGAAGACGGCATACGAGATGTAGCCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT
rev_index3	CAAGCAGAAGACGGCATACGAGATGCTAAAGTGAAGTTCAGACGTGTGCTCTTCCGATCTTCTACTATTCTTTCCCTGCACTGT

2.18 Data Processing and analysis of Next-Generation Sequencing

The Illumina Sequencing data was trimmed only to keep the 20 bp gRNA sequence and aligned to a custom fasta library of the known library sequences using Bowtie. Picards tool was used to quantify relative count of each gRNA from the alignment (BAM) file. Resulting scores were divided into respective time points to obtain stage scores. Average of two independent screens were taken for divided read counts and converted to $\log_2$ scale for the identification of negative values. Afterward, gene score calculations were done by MTC algorithm in which per million normalized reads counts were divided into respective time points to obtain stage scores. These values were used to calculate average of two independent screens and converted to $\log_2$ scale for the identification of negative values. Afterward, non-parametric fitting was applied to the $\log_2$ data which resulted in a probability density function. $\log_2$ read counts were used to calculate probability values according to the fitted distribution. For each gene, the weighted sum of $\log_2$ scores were calculated using values of gRNA as weights and used as the gene scores. Identified gene scores were divided to gNT scores and termed as Enrichment Score (ES).

Table 12. List of time points divisions to calculate stage scores.

Stage Name	Division of NGS Reads Counts
Initiation	Day 7 / Day -1
Maturation	Day 14 / Day 7
Maintenance	Day 21 / Day 14
Overall	Day 21 / Day -1

2.19 TRA-1-60 ,Low Copy Exogenous EGFP Silencing, and Proliferation Screens

For all of the screens 20×10^6 dH1f were seeded as 800.000 cells/10 cm plate and library infected with low MOI (~ 0.3). Puromycin selection was applied on day 2, and cells split 1:2 onto 10 cm plates in every 4 days. For reprogramming experiments a total of 8×10^6 dH1f cells (~ 4750 cells/gRNA) were seeded as 600.000 cells/10 cm plate. For the TRA-1-60 screen, on day 7 of reprogramming, cells were trypsinized and 1/8th of the cells were transferred onto MEF-seeded 10 cm plates for sorting at day 14 and day 21. Rest of the cells were spun down at 1500 rpm for 5 minutes and washed twice with PBS supplemented with 1% FBS. Pellets were resuspended with 2 ml PBS supplemented with 1% FBS. Same procedure was performed on Day 14 and Day 21. For day 7, 14, 21 TRA-1-60-PE antibody was added at the following ratios 0.5:100, 1:100, 1.5:100, respectively. For the EGFP screen, cells were trypsinized on day 7, were spun down at 1500 rpm for 5 minutes, washed twice with PBS supplemented with 1% FBS, and all of the cells were used for cell sorting by FACS. For the proliferation screen, cells seeded in a 10 plate as 600.000 cells (~ 350 cells/guide) and passaged 1/3 in every 4 days until day 21.

2.20 Differential Gene Expression Analysis

Differential gene expressions between pluripotent stem cells and fibroblast cells are computed by affy and limma packages from R. Samples of dH1f and BJ fibroblasts are compared to samples of their relative iPSCs and embryonic stem cells from GEO data series GSE55679.

2.21 Statistical Analysis

Data were analyzed by Student's t-test for comparison of two groups. The Pearson correlation coefficient was used for the correlation analysis related to the screen.

3 Results

3.1 CRISPR-Cas9 system is suitable for investigations of regulators of reprogramming in human fibroblasts

To investigate whether the CRISPR-Cas9 system could be used in human fibroblasts, I performed Cas9 activity assay in the dH1f human fibroblast cell line. First, I generated Cas9 and GFP expressing dH1fs with stable infection of lentiCas9Blast and pBabe-GFP-Hygro, then, targeted EGFP sequences with guideRNA T1 and T2 in lentiGuide vectors (gT1 and gT2), and monitored EGFP levels with flow cytometry. Results demonstrated that CRISPR-Cas9 induced GFP loss was up to ~60% starting with Day 6 and the system became saturated around Day 10 for both gRNAs. Furthermore, gRNAs and their effect of loss of EGFP differs as T1 becomes saturated with Day 6 and T2 with Day 8 (Figure 3.1.A). In conclusion, dH1f cells are suitable for the CRISPR-Cas9 system, requires approximately ten days for Cas9 to work, and each guideRNA has different efficiency.

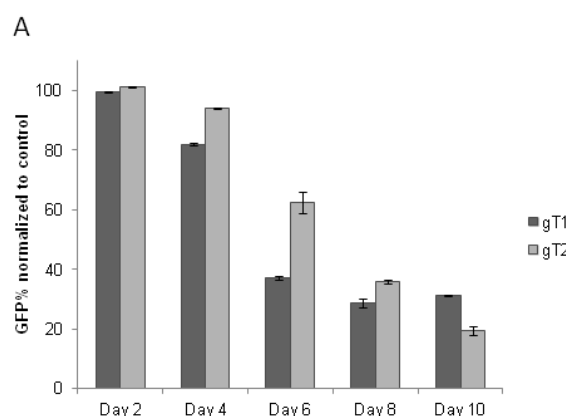


Figure 3.1 Cas9 Activity Assay shows the suitability of CRISPR-Cas9 for dH1fs

A) EGFP levels following infection of gRNAs are monitored with flow cytometry every two days until day 10. GFP values are normalized to the non-gRNA infected control (n=3; error bars, $\pm$ s.d.).

After establishing the applicability of CRISPR-Cas9 in dH1f cells, I evaluated its adaptability and limitations for investigating gene modifications during reprogramming. First, I targeted known reprogramming barriers and essential regulators (Table 2) with the CRISPR-Cas9 system (Figure 3.2.A), confirmed successful DNA cleavage for a number of genes by the T7E assay (Figure 3.2.B) and, induced cells to the pluripotency via OSKM infection. Quantification of TRA-1-60 positive colonies revealed parallel results with the literature with slightly increased fold changes than reported (Figure 3.2.C). For example, shRNA-mediated

knockdown of DOT1L and p53 have been shown to increase reprogramming efficiency 3-4 fold, but with the CRISPR-Cas9 system, the fold differences compared to control cells were 5-6 fold [45, 52]. Second, I evaluated multiplexing ability of the CRISPR-Cas9 system with small molecule treatment. To that end, I repeated the same reprogramming experiment with the addition of iDOT1L (EPZ5676) during the initial seven days of reprogramming (Figure 3.2.C). iDOT1L treatment has been shown to increase reprogramming efficiency by 3-4 fold [57]. Results show that for each gRNA except those targeting DOT1L, fold increases in reprogramming was further improved by 3-4 fold, indicating multiplexing with iDOT1L acts synergistically. Importantly, DOT1L targeting gRNAs did not have a significant effect on reprogramming efficiency in the presence of iDOT1L. This might be either due to off-targets of the small molecule of iDOT1L or the presence of cells with unmodified DOT1L alleles which were then effected by iDOT1L treatment. Third, I tested whether multiplexing gRNAs targeting different barrier combinations with different functions would yield combinatorial pairing (Figure 3.2.D). For this, I selected two gRNA pairs, targeted them sequentially before reprogramming and induced to the pluripotency. As a result quantified colony counts revealed that most of the pairing had additive effect rather than synergistic effect. Particularly, gNR2F2 pairings were not effective with gP53 or gYY1 but only effective with gLATS2 in terms of effecting emerging iPSC colony count. Moreover, gP53 and gLATS2 pairing did not further increased the emerging iPSC number with EPZ5676 treatment. The most significant enrichments were seen with gYY1 pairings with either gP53 or gLATS2, both results in 9 fold increase in colony count compared to the empty controls. In conclusion, the CRISPR-Cas9 system is suitable for gene investigations in reprogramming, can be multiplexed with different gRNAs or small molecule treatments, and multiplexing different barriers might encounter redundancy in function and can not further increase reprogramming efficiency.

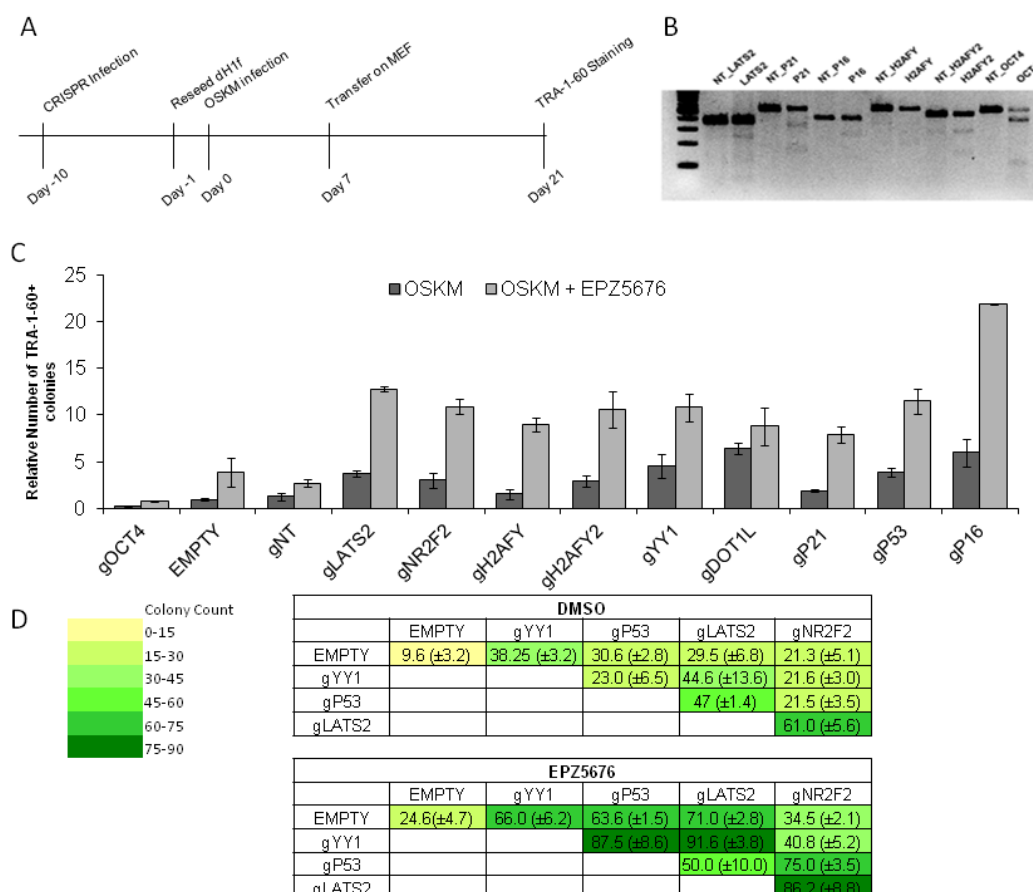


Figure 3.2 CRISPR-Cas9 system is suitable for gene investigations in reprogramming experiment

A) Timeline of CRISPR-Cas9 components infection and iPSC generation. B) CRISPR-Cas9 induced cleavage for a number of reprogramming barriers and essential regulators assessed by T7E assay. Samples were collected 10 days after gRNA transduction. C) A select number of known reprogramming barriers and essential regulators were targeted with guideRNAs using the lentiCRISPRv1 vector. 10 days after gRNA infection, pluripotency was induced by OSKM infection and cells were cultured with standard methods for 21 days. TRA-1-60-HRP staining was performed on Day 21 of reprogramming and colony counts were quantified with ImageJ software. Colony counts were normalized to the average of colony counts of untransduced controls. Same experimental setup was employed with 3 μ M EPZ5676 treatment for first six days of reprogramming to assess the effects of combining CRISPR and small molecule treatment (n=3; error bars, $\pm$ s.d.). D) A select number of known reprogramming barriers and essential regulators were targeted with guideRNAs and cultured for ten days. Later, second round gRNA infections were targeted for different genes. Generated dH1fs were infected with OSKM and cultured with DMSO or EPZ5676 for the first 6 days of reprogramming. iPSCs were stained with TRA-1-60-HRP at day 21 and colony

number was quantified with ImageJ software. Calculated colony counts were presented in boxes (n=3; error bars, $\pm$ s.d. shown in parentheses).

3.2 Generation of a focused CRISPR-Cas9 Knock-out Library Targeting Chromatin Modifiers

Having demonstrated the feasibility of using the CRISPR-Cas9 system to investigate barriers and essential regulators of reprogramming, a custom CRISPR-Cas9 knockout library targeting chromatin modifiers named EPIKOL (**Epigenetic Knock-Out Library**) was generated. 247 different chromatin modifiers compromising chromatin remodelers, DNA modifiers, and histone modifiers including readers, writers, and erasers (APPENDIX 7) were included in the library. GuideRNAs were computationally designed to target exons encoding the catalytic or critical domains of each chromatin modifier. This approach has previously been shown to increase the detection of loss-of-function phenotypes (Figure 3.3.A). For most of the genes, 5-10 gRNAs that were predicted to have minimal off-target activity were selected. 80 non-targeting guideRNAs from the GeCKO Library were included in the library to serve as internal controls [105]. Oligo array synthesized gRNAs were cloned into lentiCRISPRv2 vectors by Gibson Assembly (Figure 3.3.B). Confirmation of library cloning was validated by next-generation sequencing of pooled plasmids (Figure 3.3.C).

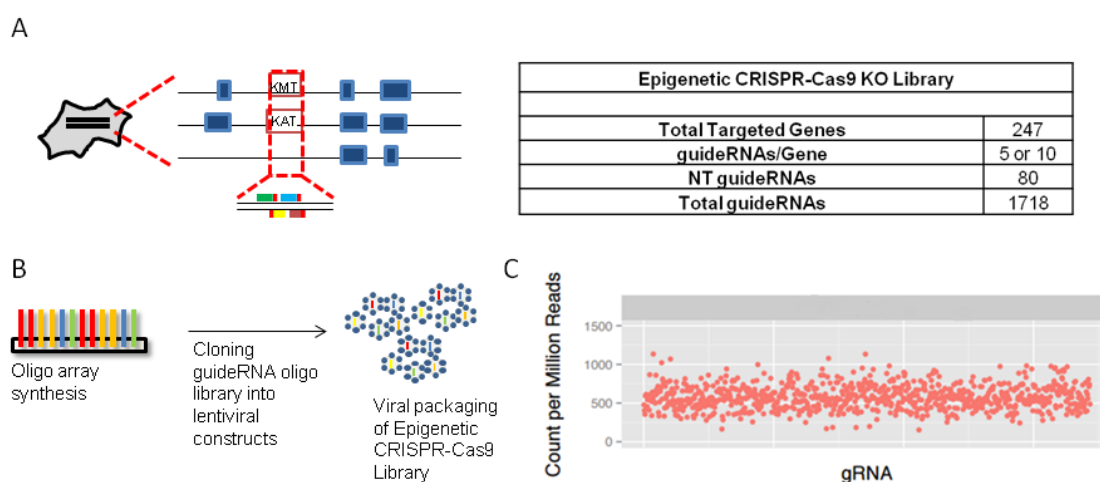


Figure 3.3 Chromatin modifiers targeting CRISPR-Cas9 Knock-Out Library Generation

A) GuideRNAs were computationally designed to target catalytic or critical domains of chromatin modifiers. B) GuideRNAs were synthesized with oligo arrays and cloned into the

lentiCRISPRv2 vector by Gibson Assembly. C) GuideRNA regions were amplified from cloned plasmids to confirm cloning and sequenced by next-generation sequencing.

3.3 TRA-1-60 Expression Based Enrichment Screen Results and Data Analysis

To investigate the overall and stage-specific effects of chromatin modifiers, I have designed a screen as follows: dH1f cells were infected with the EPIKOL library and induced to the reprogramming with OSKM infection. Emerging iPSCs were collected at different time points of reprogramming by sorting based on the presence of TRA-1-60 cell surface antigen. Collected cells were used to quantify integrated gRNA sequences by Next-Generation Sequencing (NGS), and quantified read counts were normalized the read counts of non-targeting gRNA then to the respective time points to identify stage-specific enrichment (Figure 3.4.A).

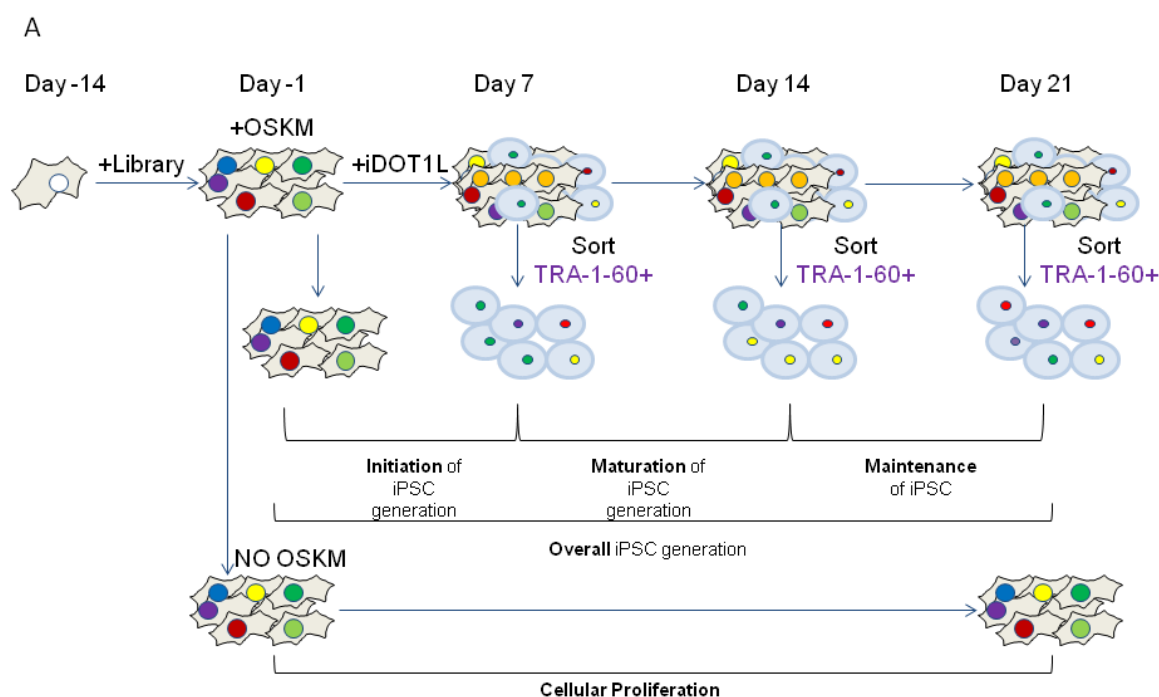


Figure 3.4 Proposed cell surface antigen enrichment based screen outline for investigating 247 chromatin modifiers role in reprogramming

A) dH1fs will be infected with the library at low MOI (~0.3). Cells will be cultured for at least ten days, and the equal portions of the population will be used to induce reprogramming with OSKM infection, cultured as the same duration of screen time to identify proliferation regulators, collected for gDNA isolation for Day 0 gRNA quantification. OSKM infected cells will be sorted according to their TRA-1-60 expression at different time points (Day 7, 14, 21). gDNAs of collected cells will be used to quantify guideRNA sequence abundance in

different days of reprogramming. Quantified read counts will be used for comparison between different time points to identify enrichment in initiation (Day 7/ Day -1), maturation (Day 14/ Day 7), maintenance stages (Day 21/ Day 14), and overall effect (Day 21/Day-1). To identify gene modifications that regulate proliferation, cultured cells without OSKM infection will be collected at Day 21 and read counts will be compared to the Day -1 read counts. To maintain library representation through reprogramming, and enable successful sorting of TRA-1-60 expressing cells at Day 7, EPZ5676 will be used to boost reprogramming efficiency.

Before conducting the screen, I tested the suitability of this screen strategy for identifying stage-specific effects of modifications. I hypothesized that the percentage of the cells stained with TRA-1-60 would represent guideRNA read counts in the screen; thus, investigation of stage-specific and the overall effect of gene manipulations would be possible by the use of fold increase in marker-positive cells between different time points. To test this strategy, I decided to assess the effect of a known reprogramming barrier DOT1L, which has been shown to increase reprogramming efficiency mostly at the initiation and maturation stages. Therefore, I induced dH1fs to pluripotency with OSKM infection, treated with either DMSO or EPZ5676 for the first seven days of reprogramming. On Day 7, 14, and 21 of the reprogramming I have quantified the percentage of TRA-1-60 expressing cells with flow cytometry. Results indicated that iDOT1L treatment increased the percentage of TRA-1-60 positive cells at each time point compared to the DMSO control (Figure 3.5.A). Moreover, analysis of percentage comparisons between time points showed that seven day iDOT1L treatment increases TRA-1-60 positive cells specifically in initiation and maturation stages as reported before, and there was no significant difference in maintenance stage (Figure 3.5.B). In conclusion, gRNA abundance as a read count instead of TRA-1-60 positive cell percentage, might be used to identify stage-specific effects of modifications of chromatin modifiers in the screen.

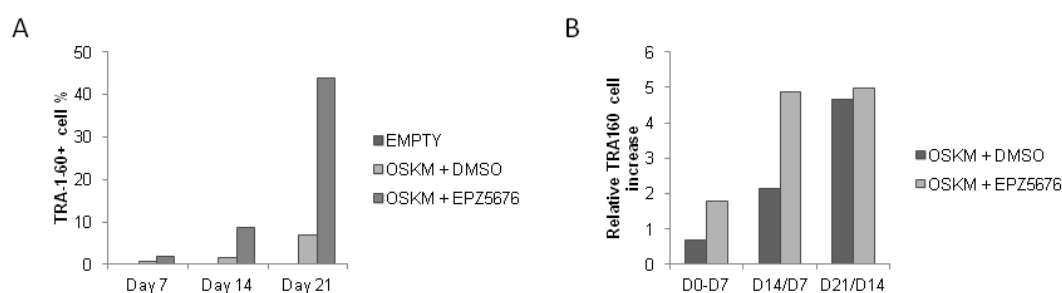


Figure 3.5 TRA-1-60 staining to investigate stage-specific features of modifications validated with the effect of DOT1L inhibition on different stages

A) dH1f cells were induced to pluripotency with OSKM and the percentage of TRA-1-60 positive cells were assessed by TRA-1-60-PE staining with flow cytometry on days 7, 14, and 21. dH1fs were treated with EPZ5676 or DMSO as a control. B) Fold increases between different time points were calculated. For maturation, Day 14 (D14) TRA-1-60 cell percentage was divided to Day 7 (D7) TRA-1-60 cell percentage, and for maintenance Day 21 (D21) TRA-1-60 cell percentage was divided to Day 14 (D14) TRA-1-60 cell percentage. For initiation, Day 7 (D7) TRA-1-60 cell percentage plotted as Day 0 (D0) TRA-1-60 cell percentage was zero.

I next performed a series of experiments to ensure that the library representation would be maintained throughout the screen. First, I checked the possible loss of library representation during viral packaging and infection of dH1f cells. I infected dH1fs with the library at low MOI (<0.3), and two days after infection, I isolated genomic DNA from these samples. Next-generation sequencing of PCR-amplified gRNAs from the genomic DNA samples was used to quantify guideRNA abundance in the population. Integrated gRNAs showed that almost all of the library was represented in the infected cell population (APPENDIX 1.A). Second, I calculate the minimum number of cells needed to ensure maintenance of library representation during reprogramming. For the reprogramming part of the screen, each guideRNA has been calculated to be represented by ~ 4750 cells which would give rise to the 4-47 iPSC colonies assuming that reprogramming efficiency is 0.1-1%. However, this number is not sufficient for proper data analysis after next generation sequencing. Hence, instead of increasing the initial cell number, reprogramming efficiency was increased by 3-4 folds with iDOT1L treatment for the first 7 Days of reprogramming which is also shown to act synergistically with most of the known barriers.

Next, I conducted two independent screens and analyzed the data according to a custom algorithm (MTC) which uses gRNA read counts from different time points (APPENDIX 1.B), to calculate gene scores. These gRNA read counts were initially divided to their respective time points (Table 12) and converted to $\log_2$ scale. $\log_2$ scores applied to a non-parametric fitting which results in probabilistic density function (PDF). For each gene, the weighted sum of $\log_2$ scores were calculated using gRNA values and used assigned as gene scores. GuideRNAs that had similar to each other as high or low fold change deconvoluted in high or low gene scores, respectively. Moreover, for every targeted gene, outlier gRNAs

generated low PDF values. Therefore, the contribution of outlier gRNAs to the gene scores are minimized. The two independent experiments exhibited low correlation between them both at gene level (Figure 3.6.A) and gRNA read count level (Figure 3.6.B). By taking average of both read counts before MTC scoring was also able to mitigate the noise among the two screens. Calculated scores were then normalized to the non-targeting gRNA scores and termed as enrichment score (ES) (APPENDIX 8).

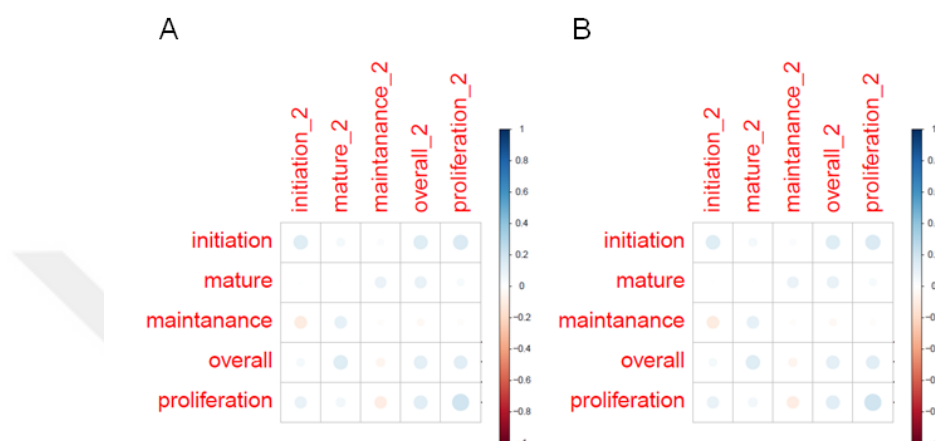


Figure 3.6 Correlation between two independent TRA-1-60 sorting screen.

A) Gene score correlations between two independent screens. The Pearson correlation coefficient values were used to estimate correlation. B) guideRNA read count division correlations between two independent screens. The Pearson correlation coefficient values were used to estimate correlation.

Screen results also represented for known reprogramming regulators (Figure 3.7.A) and identified novel barriers and essential regulators for both overall effects and stage-specific effects (Figure 3.8.A). Top five hits in overall were novel barrier genes PAXIP1, BRD2, USP22, HDAC10 and a known barrier DNMT3A and bottom five hits were CHAF1A, KDM8, PRMT5, SMYD3, ACTR as essential regulators for the overall effect (Figure 3.8.B) [52].

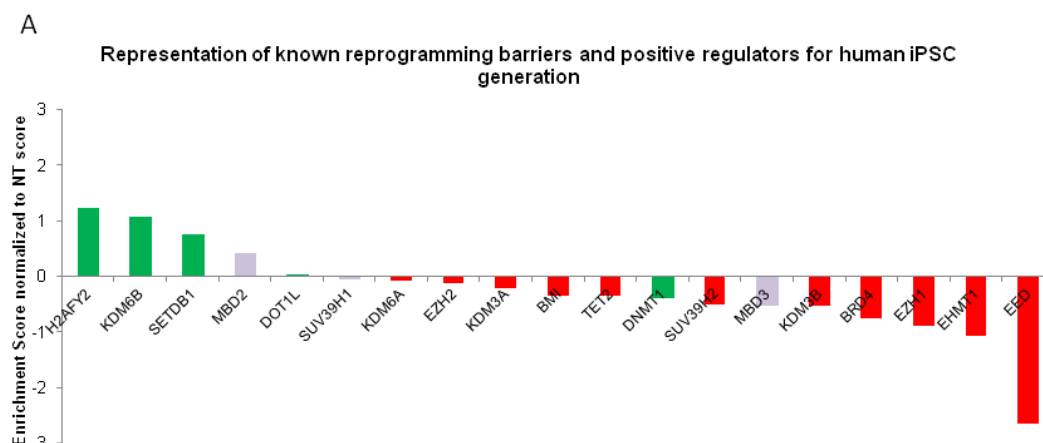


Figure 3.7 Screen results validate most of the known human reprogramming regulators

A) Overall enrichment scores of known human reprogramming regulators were bar graphed. Color code was applied to imply reported functions. Green ones were reported to be barrier genes, red ones were reported essential regulators, and purple ones had contradictory results in the literature. DNMT1 is a barrier to the reprogramming in the literature, however the screen results suggest that it has no effect [73]. Literature has contradictory results obtained with shRNA investigations for SUV39H1, MBD2, MBD3 genes [21]. DOT1L did not increase overall score as iDOT1L treatment was present in the screen.

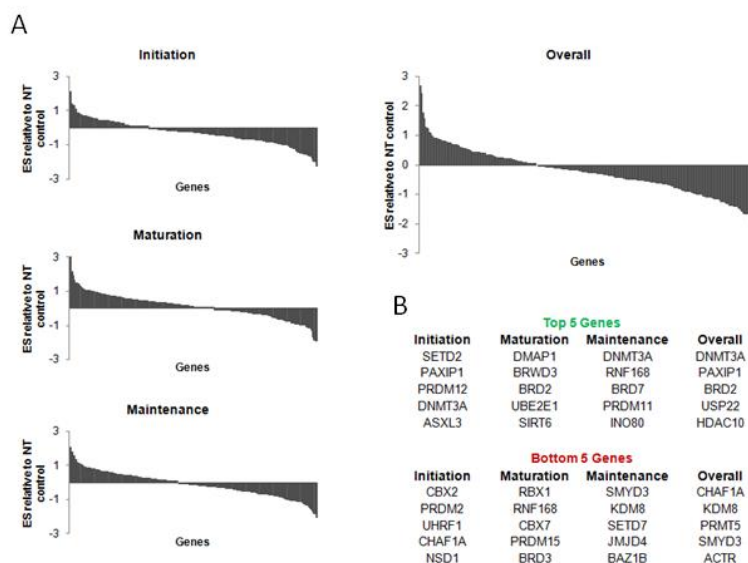


Figure 3.8 NGS results from the screen analyzed and quantified with MTC and normalized to the NT scores

A) Enrichment score for initiation, maturation, maintenance, and overall were analyzed and quantified with MTC algorithm and sorted scores are plotted on bar graphs. B) Top and bottom five genes were listed for each stage.

To assess the consistency of the screen results between different time points, correlation values between samples that have overlapping time points were calculated, namely, fold changes in Day 7/ Day-1, Day 14/Day -1, and Day 21/ Day -1. Correlations between enrichment scores indicated that obtained gene scores were consistent within themselves, and gene modifications had the same trend through different time points both at the gene level (Figure 3.9.A) and at the gRNA level. Therefore, although some of the genes have a high stage-specific effect, barriers and essential regulators generally follow the same effect through reprogramming.

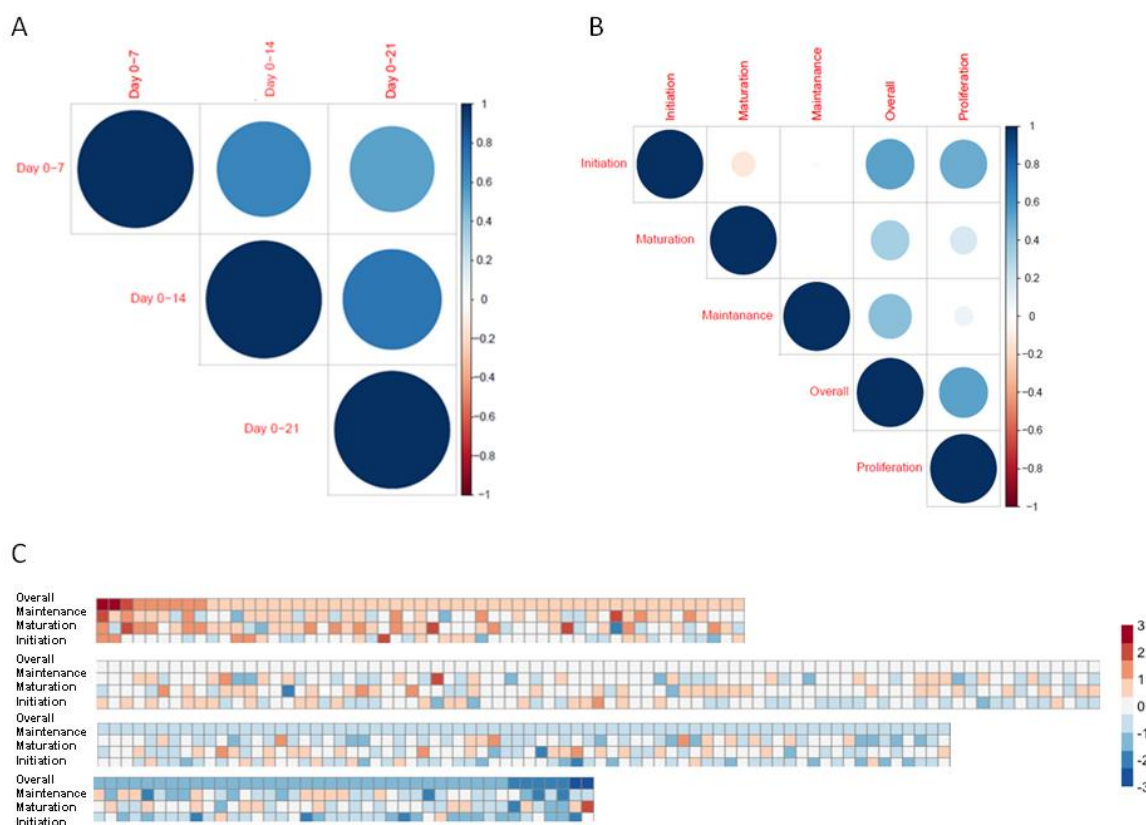


Figure 3.9 Screen results show consistency throughout reprogramming and identify stage-specific effects of modifications

A) Continuity of gene scores throughout reprogramming was assessed by gene scores from overlapping time periods (Day 7/Day 0, Day 14/Day 0, Day 21/Day 0). The Pearson correlation coefficient values were used to calculate correlation. B) Comparison of enrichment scores in different stages were compared. The Pearson correlation coefficient values were used to calculate correlation. C) Genes were sorted based on their overall score and scores of the each stage was compared to its overall score.

To systematically decipher the stage-specific effects of chromatin modifiers on reprogramming, I first compared the overall scores with different stages both at the gene (Figure 3.9.B) and at the gRNA level (APPENDIX 3.A). Correlations showed that overall score is mostly affected by the initiation stage, then maintenance stage, and least by the maturation stage. Moreover, for the majority of the genes, a similar trend is observed with overall scores and stage-specific scores (Figure 3.9.C).

To test whether our screen results were consistent with the known effect of proliferation on the iPSC generation, I conducted a screen in a similar setup except cells were not transduced

with OSKM and were collected upon day 21s of passage. Hits from this proliferation screen (Figure 3.10.A) validated reported tumor suppressor genes such as SETD2, DNMT3A, and TRIM24 in top five ranked genes (Figure 3.10.C) [110–112]. I then compared the reprogramming overall scores with proliferation scores. Results from proliferation screen was correlated with overall scores (Figure 3.10.B), also showed high correlation particularly with the initiation scores (Figure 3.9.B).

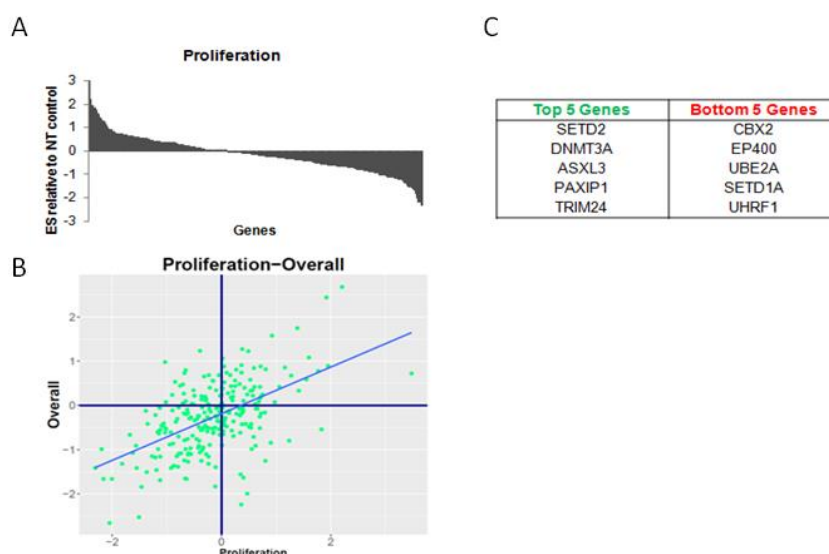


Figure 3.10 Proliferation screen results and its effect of proliferation on successful iPSC generation.

A) Enrichment scores for proliferation was analyzed and quantified with MTC algorithm B) The scatter plot for overall and proliferation scores used to calculate linear regression ($R^2 = 0.29$). C) Top and bottom five genes for proliferation are listed.

To validate the screen results, I cloned individual gRNAs of hits from the screen and conducted reprogramming experiments in the arrayed format with the CRISPR-Cas9 system. These hits were selected based on their placement in the overall scores and included those that might act as barriers, have no effect, or be essential regulators. The effect of each gene was tested by two different gRNAs. Quantification of TRA-1-60 positive colonies demonstrated that most of the effect is with only one gRNA, therefore, DNMT3A, HDAC10, KMT2A, CARM1, and SETD2 might act as barriers. On the other hand, TAF1, SMARCA5, SMYD5, ACTR, and KDM8 might be essential regulators of reprogramming. PRDM3 and DNMT3L were barrier hits in the screen but the independent experiments did not validate the screen results. Moreover, genes that are suggested to have no effect on reprogramming

showed varying TRA-1-60 colony counts. For example, gSFMBT1 showed both non significant enrichment and significant depletion in colony count (Figure 3.11.A).

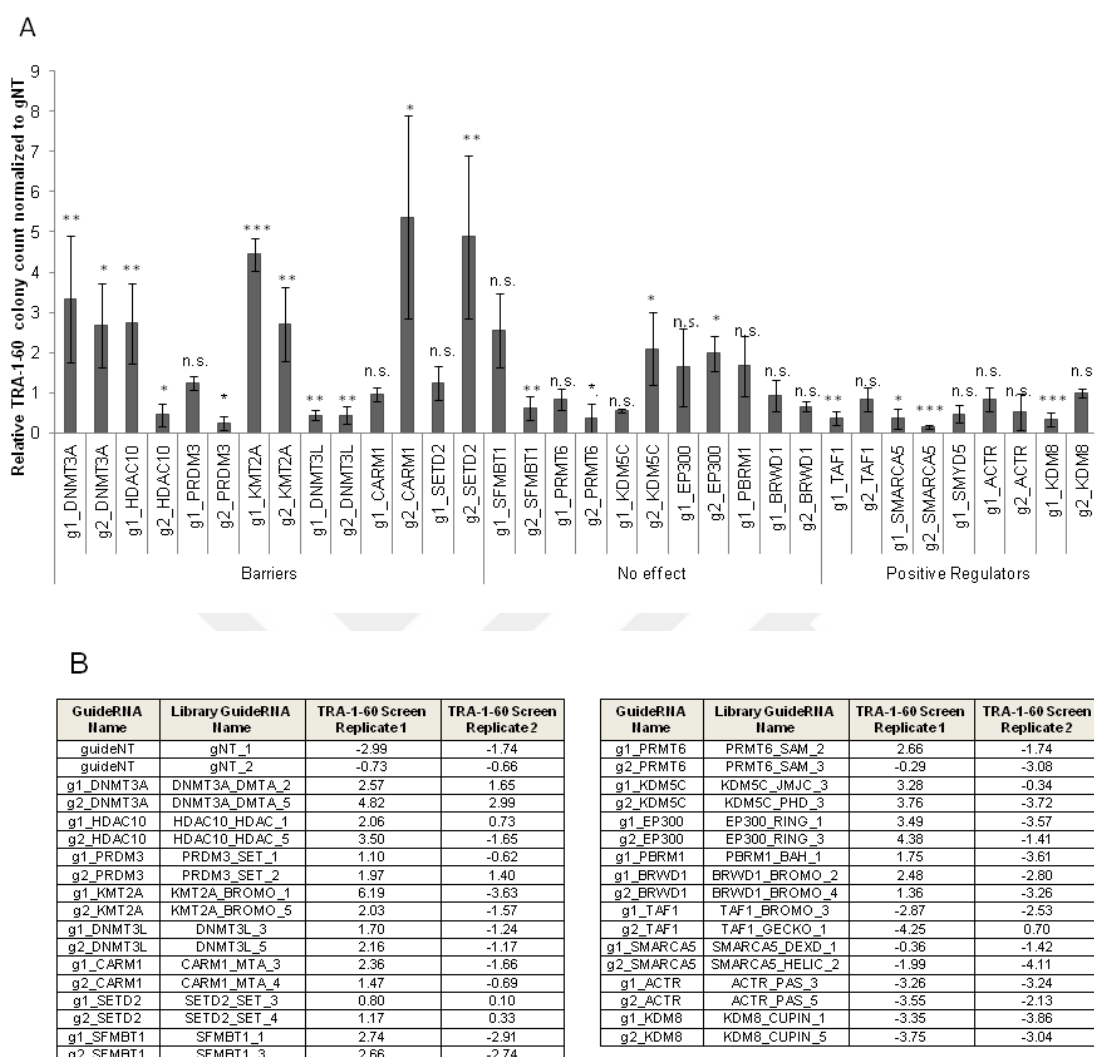


Figure 3.11 Individual gRNA testing for validation of screen

A) A select number of screen hits were individually targeted with the indicated gRNAs and the resulting cells were reprogrammed in the presence of EPZ4777 for first seven days. TRA-1-60-HRP staining was performed at Day 21 and emerging colony number quantified with ImageJ software. Colony counts were normalized to the average colony counts in four different gNT control. Results are the average of two independent experiments. Statistical significance evaluated for comparison of the gNT control with samples. (n.s: not significant, * $p < 0.05$, ** $p < 0.005$, *** $p < 0.001$, $n=4$; error bars, $\pm$ s.d.). B) GuideRNAs that had been used in validations experiments were listed with their overall log₂ scores for both screen replicates.

To identify regulators at the protein complex level, I sorted overall scores and categorized the genes according to their placement in the rankings (Figure 3.12.A). Ranking was done from 1 to 247 based on overall scores, indicating the most enriched and depleted scores as rank 1-25 and rank 201-247, respectively. NT control ES was ranked as 87. Genes were then categorized into protein complex groups based on EpiFactors Database [113]. Consistent with the literature, loss of function of members of PRC1 and PRC2 complexes had either no effect or negative effect on overall reprogramming [57, 78]. In addition members of MOZ/MORF, SWI/SNF, and BAF complexes were suggested to be essential regulators. Contrarily, loss of function of the members of the MLL1,2 and MLL3,4 complexes had a positive effect on reprogramming; suggesting that they may act as barriers. Members of NuRD, CHD8, and COMPASS protein complexes showed no consistent phenotype. In conclusion, the screen data both validated the role of known essential regulators, PRC1 and PRC2 complexes, and proposed novel complexes as regulators of iPSC generation.

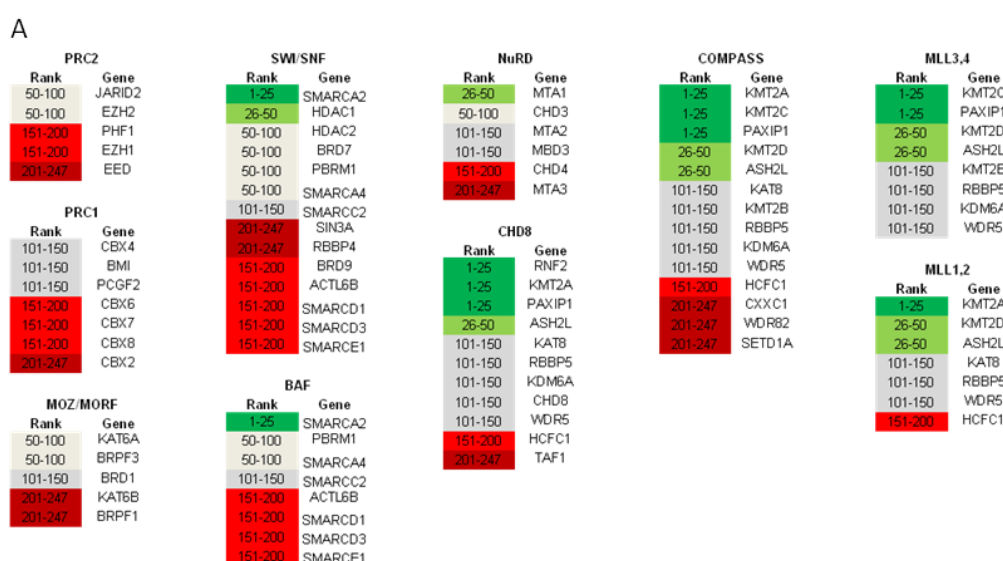


Figure 3.12 Ezh2 results categorized for chromatin modifier complexes to identify reprogramming regulators at the protein complex level

A) Gene scores were ranked for overall and categorized for different chromatin remodeling complexes. Rank 1-25 is the top scoring genes and rank 201-247 is the bottom of the list among this categorization. NT control value as 0 is placed at Rank 87. Highlighted genes as green were shown to be a possible barrier, gray ones had possible no effect, and red ones were possible essential regulators.

Previous investigations showed that specific functions can be important for the iPSC generation [55, 80]. Therefore, I evaluated the screen data to identify common functions

between chromatin modifiers and their effect on reprogramming. Genes in the library placed into three functional categories; namely chromatin remodelers, DNA modifiers, and histone modifiers (Figure 3.13.A) and the data was analyzed accordingly.



Figure 3.13 Chromatin modifiers belonging to different functional categories of act differently through reprogramming

A) Gene scores were ranked for overall and categorized based on functional differences. Rank 1-25 was the top scored genes and rank 201-247 was the bottom of the list based on this categorization. NT control value as 0 was placed at Rank 87. Highlighted genes as green were shown to be a possible barrier, gray ones had possible no effect, and red ones were possible essential regulators.

Chromatin remodelers were predicted to be mainly essential regulators and appeared in two main clusters. Genes in the first cluster generally had low overall and proliferation scores, while those in the second cluster, had generally higher overall and proliferation scores. Next I questioned whether these genes have differential expression in iPSCs as compared to fibroblasts (Figure 3.14.A). Results suggested that low scored genes were highly expressed in iPSCs while high scorers did not show significant expression difference between iPSCs and fibroblasts. This finding might suggest that chromatin remodelers that are expressed in iPSCs might be important in pluripotency. CHAF1A is an example for these genes; it has been previously reported to have an essential role in heterochromatin organization in pluripotent

stem cells and it had low scores in screen and higher expression in iPSCs compared to the fibroblasts [114].

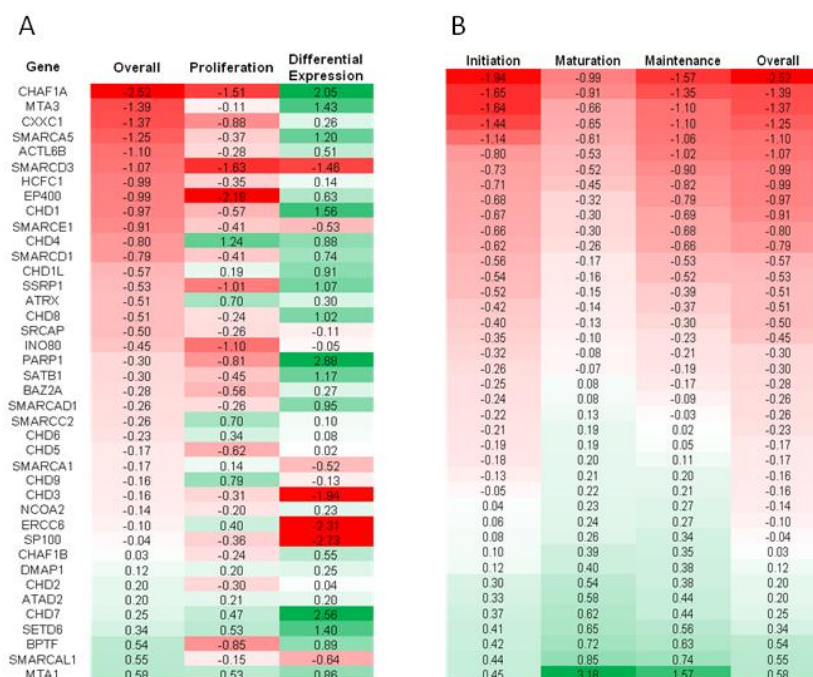


Figure 3.14 Chromatin modifiers represent as essential regulators of reprogramming

A) The overall enrichment scores of all chromatin remodelers in the library were sorted and respective proliferation scores and differential expression values were presented in color codes. B) Chromatin remodelers in the library were sorted within itself for each stage and shown to decrease reprogramming efficiency in an increasing trend starting with the maturation stage.

No specific clustering for either enriched or depleted genes was observed among DNA modifiers. DNA modifications peaked for DNMT3A and MBD4 genes, but there was not increasing trend for the rest of the genes rest (Figure 3.13.A). Contrarily, DNMT3B ranked in the bottom for DNA modifications. Moreover, DNMT3B had a very high gene expression value and used as a marker for pluripotency.

When histone modifiers were clustered as reader, writer, and eraser, no common trend was observed in overall scores among any classes. For this reason, selected histone marks associated with reader, writer, eraser classes were analyzed independently. Four major histone marks, H3K4, H3K27, H3K9, H3K36, that are regulated by different chromatin modifiers were selected. Except for the decreasing trend in H3K27 erasers, no clustering of

genes was observed (Figure 3.15.A & APPENDIX 4APPENDIX 4.A). Furthermore, contrary to previous reports, the screen results did not identify H3K9 methylation as a barrier [80].

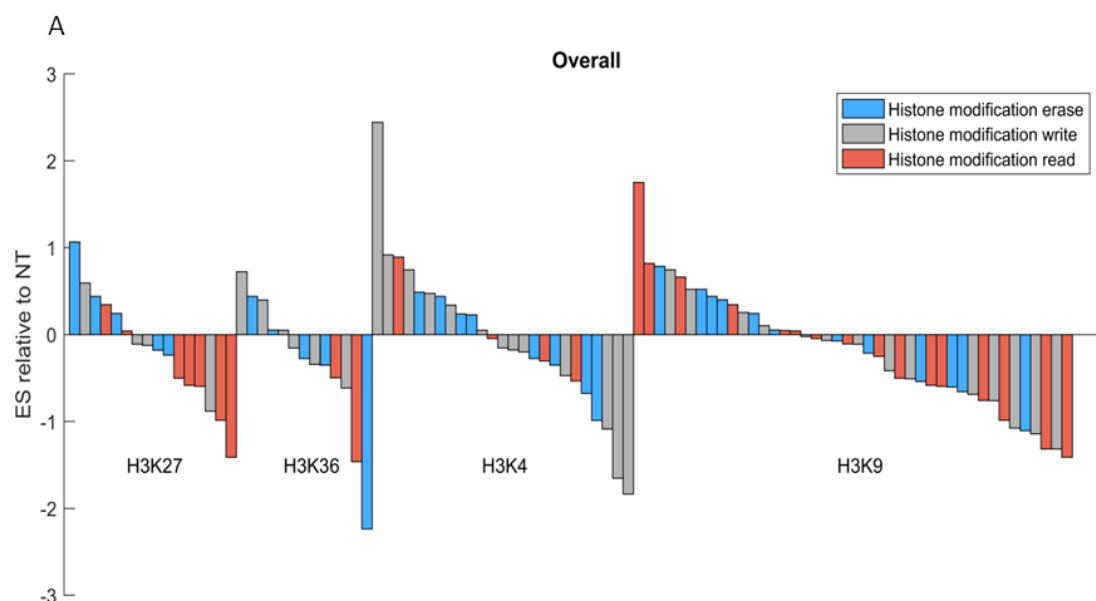


Figure 3.15 Specific histone marks and their reader, writer, erasers do not show any clustered phenotype

A) Histone modifiers were color coded for their function as reader, writer, eraser and were indicated for main histone marks with their sorted overall scores. Genes were sorted based on the presented histone marks.

Differential expression can be used to test if being a barrier or essential regulator is correlated with gene expression levels in iPSCs and fibroblasts,. Therefore, I compared differential expression values (Fc) with overall enrichment scores (Figure 3.16.A). Fc values above 0 indicate relatively higher expression in iPSCs compared to fibroblasts, and values below 0 indicate the vice-versa. Although not significant ($p > 0.05$), barriers have low or no differential gene expression in iPSCs. In contrast, essential regulators on average have higher gene expression in iPSC (Figure 3.16.B).

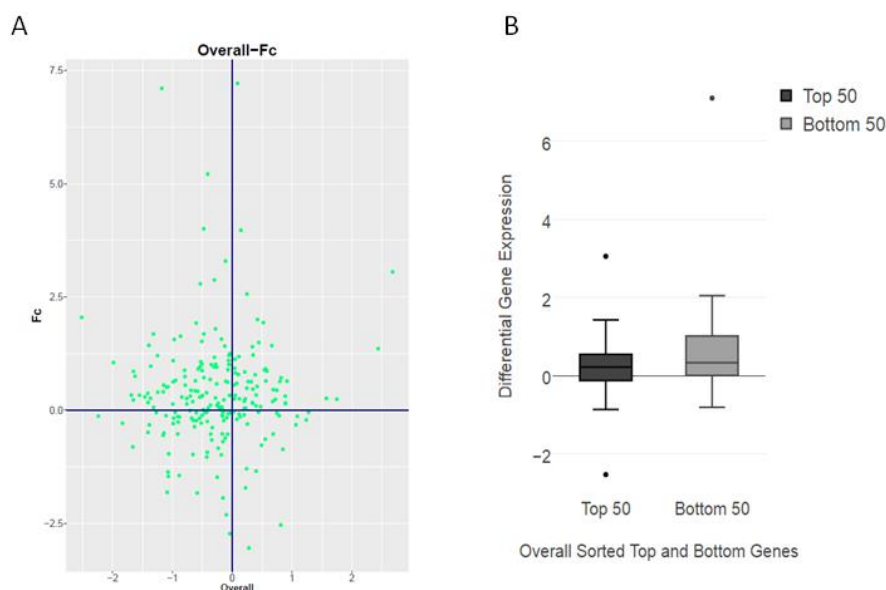


Figure 3.16 Differential gene expression between iPSC and fibroblasts might indicate their role in reprogramming

A) Scatter plot was used to compare Fc and overall scores. B) Fc values of top 50 and bottom 50 genes in overall sorted scores were demonstrated as a box plot (the difference between samples was not significant ($p > 0.05$) as assessed by two tail Student's T test assuming unequal variance).

Knockout of the barriers of reprogramming can increase the efficiency of induced pluripotency by facilitating loss of somatic cell identity in the initiation stage. Hence, I hypothesized that the genes in the initiation stage are involved in the maintenance of the somatic cell identity. To test this hypothesis, I infected dH1fs with gRNAs hits from the initiation stage (SETD2 and DNMT3A which had highest scores, and PRMT6 was one of the top ranked gene but had a lower score) and treated them with EPZ5676 for six days. At the end of the sixth day, fibroblast marker gene expressions (COL3A, ZEB1, POSTN, DKK1, GREM) decreased by half compared to the gNT control particularly in gSETD2 and gDNMT3A samples. Hence, these genes are required for maintaining somatic cell identity (Figure 3.17.A).

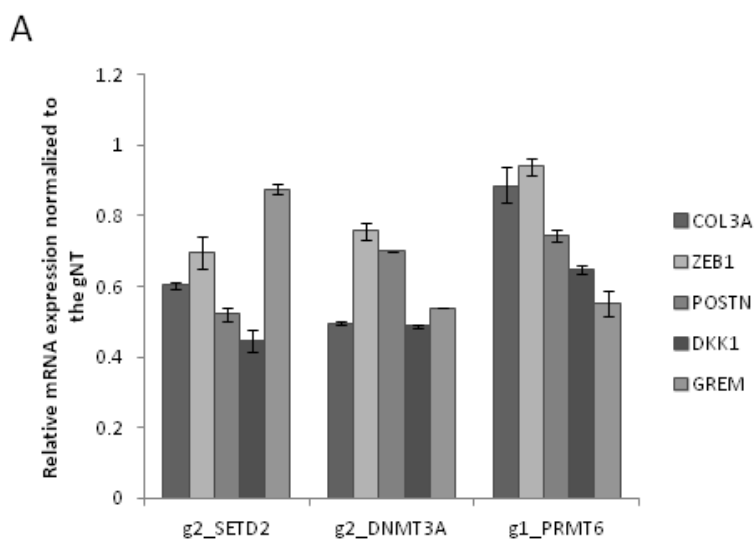


Figure 3.17 Hits identified in the initiation stage impairs somatic cell identity in fibroblasts without OSKM infection

A) dH1fs infected with guideRNAs were cultured for ten days to enable CRISPR work. To mimic the screen conditions the cells were treated with EPZ5676 for following 6 days. The decrease in mRNA levels of fibroblast marker genes was assessed by qRT-PCR. β -actin was used as an internal control. (n=2; error bars, $\pm$ s.d.).

To assess whether validated hits from overall scores (HDAC10, KMT2A, SETD2, CARM1) also facilitate reactivation of pluripotency regulatory network, I investigated the effect of these modifications on NANOG and LIN28A gene expressions at Day 7 of reprogramming. Results demonstrated that knockout of these genes increased NANOG and LIN28A expression compared to the gNT control (Figure 3.18.A).

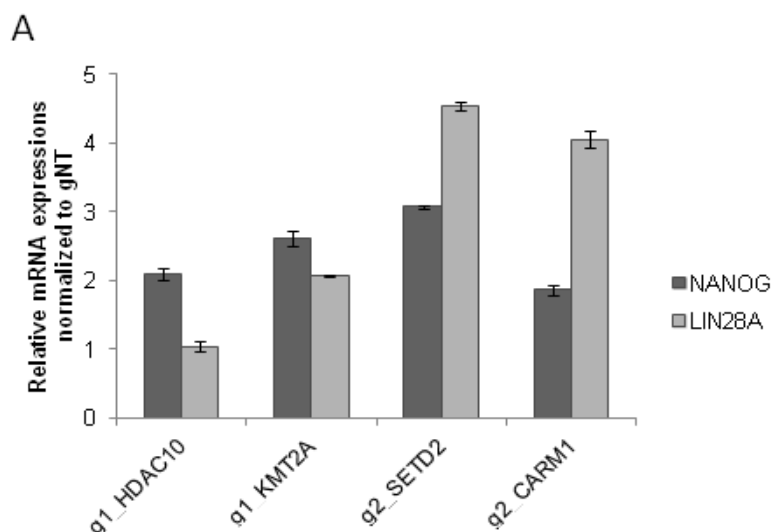


Figure 3.18 Overall hits from the screen enhance expression of pluripotency-associated genes at Day 7 of reprogramming

A) dH1fs infected with guideRNAs were cultured for ten days to enable CRISPR work. Reprogramming was induced with OSKM and cells were treated with EPZ4777 for 7 Days to mimic the screen conditions. Samples were collected at Day 7 and expression of pluripotency genes, NANOG and LIN28A expression were assessed by qRT-PCR. β -actin was used as an internal control. (n=2; error bars, $\pm$ s.d.).

3.4 Characterization of low copy exogenous EGFP silencing, Screen Results and Data Analysis

In parallel to the TRA-1-60 based sorting screen, I designed another strategy to isolate emerging iPSCs. For this, I decided to use a more stringent marker for sorting of iPSCs which is the proviral silencing mechanism that is recognized as a late stage functional marker of (Figure 3.19.A). To test this idea, I have generated LTR-driven EGFP positive dH1fs by stable transduction with pBabe-GFP-hygro and mimicked screen conditions by treating cells with EPZ5676 for the first six days of reprogramming and monitored EGFP levels relative to the non-OSKM infected cells until Day 18 of reprogramming. Unexpectedly, even on Day 6, a significant portion (~10%) of cells had lost their EGFP expression. To focus on this observation of EGFP loss in early time points, I monitored the EGFP loss kinetics for the first six days of reprogramming, and to ensure that loss of EGFP was not the effect of EPZ5676, I included DMSO-treated cells in the experiment. Results showed that EGFP loss was not due to the presence of EPZ5676, and can also be observed in the samples treated with DMSO

(Figure 3.19.B). Moreover, loss of EGFP signal could be observed on day four for both conditions, and this percentage was slightly increased with EPZ5676 treatment. As DOT1L inhibition increases reprogramming, I hypothesized that this loss of EGFP levels is related to the reprogramming efficiency. To validate that the loss of EGFP is due to proviral silencing, I decided to continue with functional assays involving perturbations of proviral silencing regulators. Knock-out or knock-down of regulators of proviral silencing such as TRIM28 might be required for successful iPSC generation, and it could be hard to interpret the results for functional characterization of proposed mechanism [30, 115]. Therefore, I have generated two EGFP cell lines; one resistant to silencing, and the other one sensitive, namely PGK-EGFP dH1fs and LTR-EGFP dH1fs, respectively [38]. When I repeated the experiments for both of these cell lines with DMSO or EPZ5676, I observed lower EGFP loss in the resistant to silencing PGK-EGFP dH1fs than in the sensitive to silencing LTR-EGFP dH1fs (Figure 3.19.C). Moreover, EGFP loss was further increased with EPZ5676 treatment in both cell lines. These results suggested that the phenotype of EGFP loss is due to proviral silencing, and reprogramming accelerator iDOT1L treatment increases proviral silencing dependent EGFP loss.

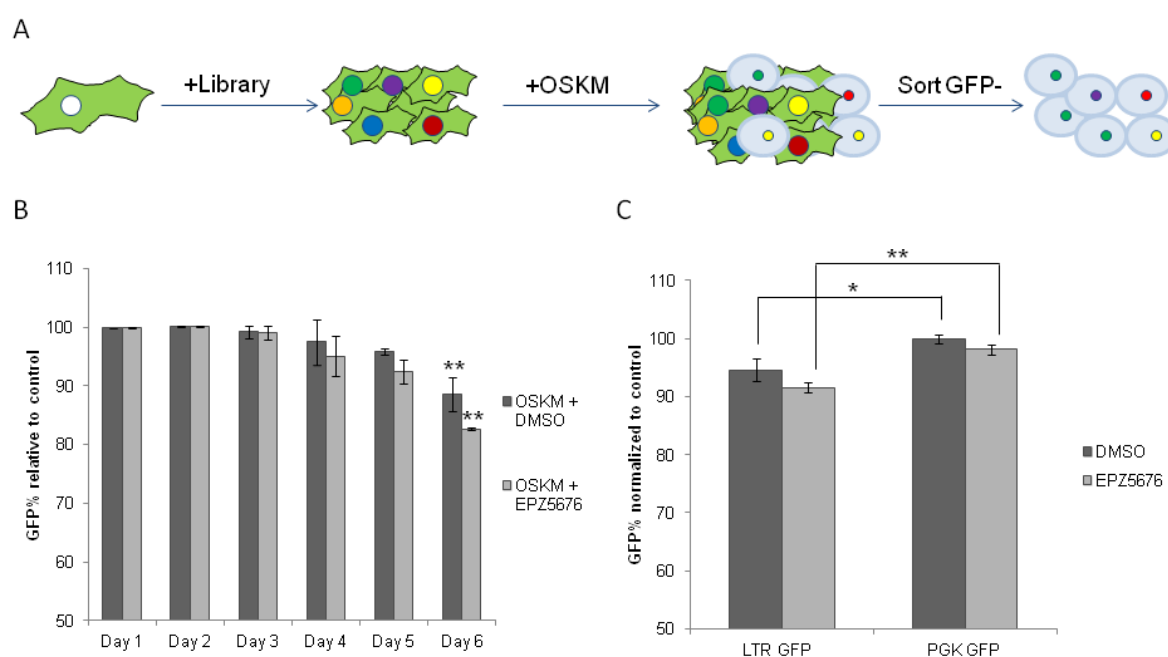


Figure 3.19 EGFP loss is observed in the early stage of reprogramming is due to the OSKM infection, and this phenotype is mediated by proviral silencing mechanism

A) Proposed screen scheme to identify regulators of reprogramming by sorting cells with low copy exogenous EGFP silencing. B) EGFP dH1fs were infected with OSKM and treated with DMSO or EPZ5676 for the first six days of reprogramming. The levels of EGFP were

monitored by flow cytometry on a daily basis. dH1fs which were not infected with OSKM used as controls (n=3; error bars, $\pm$ s.d.). C) LTR EGFP and PGK EGFP dH1fs were infected with OSKM and cultured with DMSO or EPZ5676 treatment for the first six days of reprogramming. EGFP levels were determined on day 6 of reprogramming. dH1fs which were not infected with OSKM were used as controls (* $p < 0.05$, ** $p < 0.001$, n=3; error bars, $\pm$ s.d.).

The results indicated that proviral silencing occurs in the early stages of reprogramming contrarily to the reported studies in the literature stating that is a late stage event [2, 3, 16, 24]. Findings in the literature had shown that EGFP silencing is a late stage event based on the observations in cells that co-express EGFP and OSKM. As each of the OSKM viruses were infected with MOI 3-4 in those studies, each cell had integrated more than ten copies of EGFP vectors. On the other hand, our cells were generated by transduction at low MOI of EGFP viruses. I hypothesized that this discrepancy with previous reports was due to the increased integrated copy number of EGFP vectors. To test this hypothesis, I have generated two separate cell lines having high (MOI 3), or low copy (MOI 0.3) of LTR EGFP integrated into their genome and reprogrammed the resulting cell lines. As a result, I observed decreased silencing of EGFP in the cells with higher copy number and more silencing in low copy EGFP cell line (Figure 3.20.A). Moreover, EPZ5676 treatment increased silencing in both conditions. These results indicated that proviral silencing is a copy number dependent mechanism through reprogramming, and happens in the early stage of reprogramming.

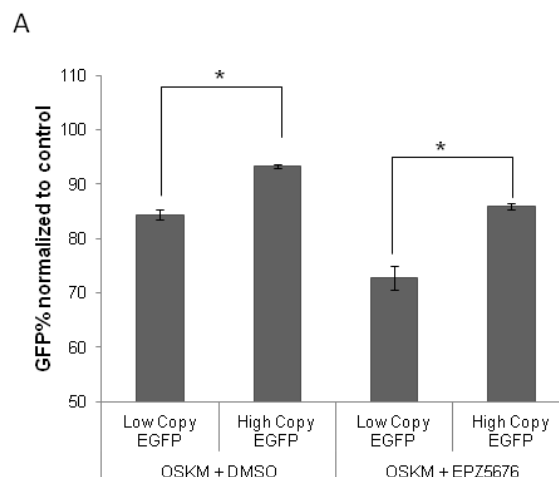


Figure 3.20 Proviral silencing functions in copy number dependent manner through reprogramming

A) Low copy and high copy LTR EGFP harboring dH1fs were infected with OSKM and cultured with DMSO or EPZ5676 treatment for the first six days of reprogramming. EGFP levels were determined at Day 6 of reprogramming. dH1fs which were not infected with OSKM was used as controls (* $p < 0.05$, $n=3$; error bars, $\pm$ s.d.).

Next, I questioned whether the cells that can silence transgenes in early stages are initially reprogrammed cells. For this, I conducted both positive and negative selection experiments. For positive selection experiment, I repeated the previous experiment with LTR-EGFP cells and sorted EGFP-negative, and positive, and population of cells at day 7 and cultured them in separate wells as a regular reprogramming experiment (Figure 3.21.A). At the end of the Day 21, I stained cells against TRA-1-60 and observed that sorted EGFP-negative cells gave rise to the increased number of colonies in which were also bigger in size than the both EGFP-positive and EGFP positive and negative mixed population sorted cells (Figure 3.21.B). Moreover, I quantified the pluripotency-associated gene expression in the sorted cells on day 7 by qRT-PCR and found that EGFP negative cells had higher expression of pluripotency genes, such as E-CADHERIN, LIN28A, DNMT3B, NANOG (Figure 3.21.C). These results suggested that early stage proviral silencing represents initially reprogrammed cells and can be used for the enrichment of reprogrammed cells.

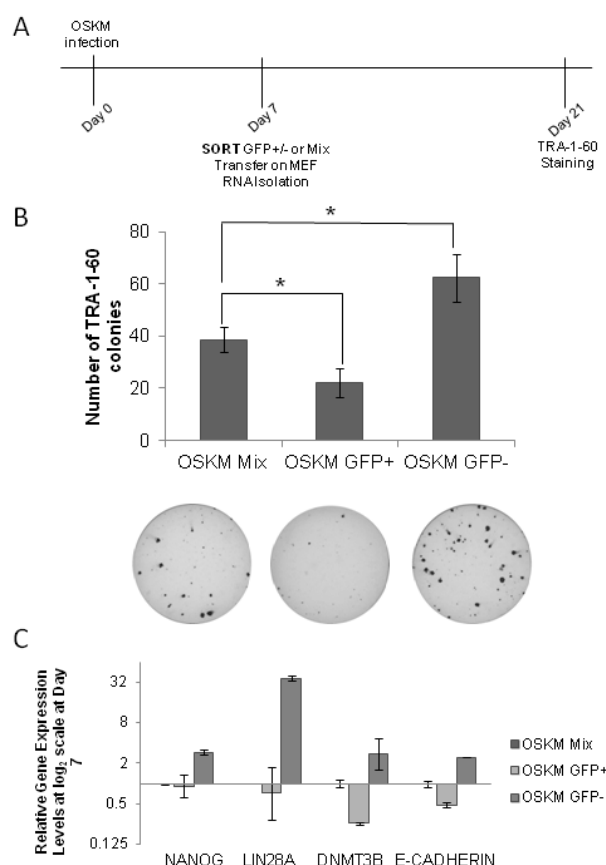


Figure 3.21 Early stage exogenous EGFP silencing can be used to enrich emerging iPSC

A) Experimental setup for positive selection of EGFP silenced cells to assess their reprogramming ability. B) EGFP dH1fs were infected with OSKM and cells sorted at Day 7 based on their EGFP levels. OSKM mix population was sorted to serve as a control. Sorted cells were cultured for 21 Days. Samples were stained with TRA-1-60 at Day 21 and analyzed with ImageJ software to quantify emerged colony count. Plates represent OSKM Mix, OSKM GFP+, OSKM GFP-, from left to right. ($p^* < 0.05$ $n=3$; error bars, $\pm$ s.d.). C) EGFP dH1fs were infected with OSKM and collected based on their EGFP levels by sorting on day 7 of reprogramming. Collected cells were used for RNA isolation, and pluripotency-associated gene expressions were assessed by qRT-PCR. β -actin was used as an internal control. ($n=2$; error bars, $\pm$ s.d.).

For the negative selection experiment, I decided to further benefit from the PGK and LTR EGFP cell lines and took the advantage of the presence of a puromycin resistance gene on these vectors. The same mechanism leading to silencing of EGFP also has to silence puromycin resistance gene. Reprogramming experiments were conducted with PGK and LTR cells and the cells were cultured with or without puromycin between Day 7-10 and stained at

Day 21 (Figure 3.22.A). Results demonstrated that puromycin treated samples had less iPSC generation for both cell lines (Figure 3.22.B). Moreover, the number of emerging iPSC colonies was decreased in LTR cells as compared to the PGK cells; validating proviral silencing at early stages of reprogramming. Overall, both of these experimental setups suggested that low copy transgene silencing can be used to enrich emerging iPSCs.

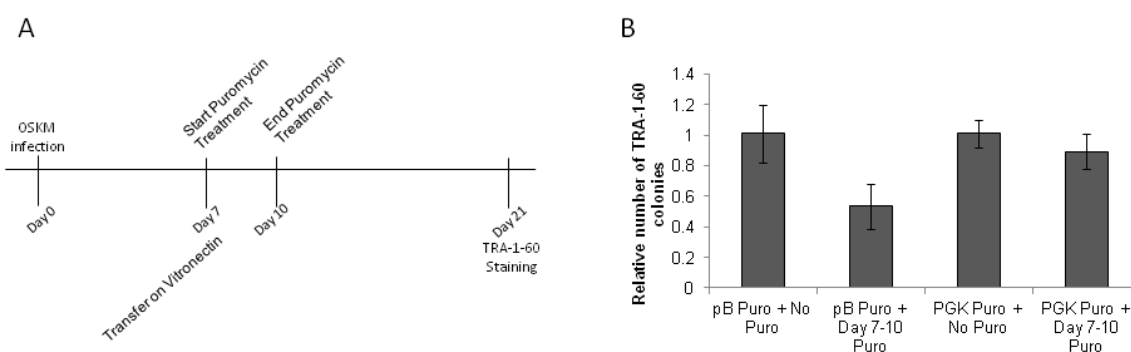


Figure 3.22 Negative selection strategy for validation of proviral silencing.

A) LTR and PGK Puro dH1fs were infected with OSKM and treated with or without 1 μ g/mL puromycin between Days 7 to 10. iPSC cells were stained with TRA-1-60-HRP at Day 21 and quantified by ImageJ software. Values are normalized to the no puro conditions (n=4; error bars, $\pm$ s.d.).

After establishing that proviral silencing works in the early stage of reprogramming, I re-evaluated our screen outline and decided to end the screen at Day 7 to only capture regulators of reprogramming in the initiation stage. I already saw a further decrease in EGFP levels when cells were treated with EPZ5676 which inhibits a known barrier of reprogramming. This result indicated that the effects of barriers can also be captured based on EGFP silencing. To test whether CRISPR-Cas9 induced gene loss was able to show the same phenotype; i.e., EGFP silencing, I infected LTR-EGFP positive cells with gRNAs that target known reprogramming barriers, and then I induced reprogramming by OSKM infection. I observed lower EGFP levels at Day 6 of reprogramming for each targeted known barrier genes compared to no gRNA control (Figure 3.23.A). Moreover, the levels of silencing in these cells were in parallel with their TRA-1-60-HRP stained experiments. Therefore, I ensured that proviral silencing at early time points of reprogramming can also be used for the screen.

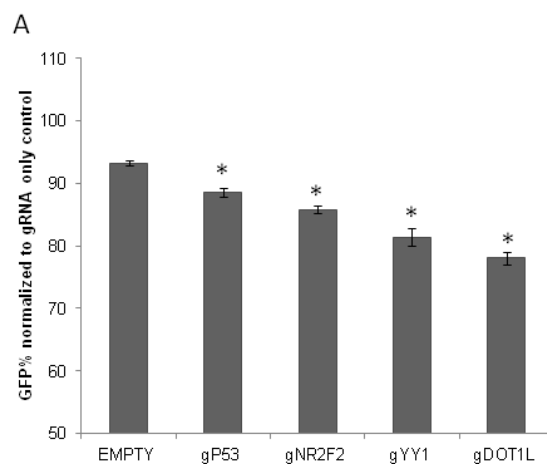


Figure 3.23 Effect of CRISPR induced gene perturbations can be assessed with proviral silencing at early time points of reprogramming

A) EGFP dH1Fs were infected with guideRNAs targeting known reprogramming barriers and cultured for at least ten days followed by infection with OSKM. Proviral silencing was assessed at Day 6 of reprogramming by flow cytometry. EGFP levels were normalized to the only gRNA infected controls. Statistical significance was evaluated by comparison of the empty control with samples. (* $p < 0.05$, $n=3$; error bars, $\pm$ s.d.).

After establishing that proviral silencing can be used to identify early regulators of reprogramming, I conducted the exogenous EGFP silencing based screen with three technical replicates in a similar setup with previous screens, using low copy LTR-EGFP-dH1fs. On day 7 of reprogramming, cells which did not have EGFP expression were sorted and used for quantification of integrated gRNA sequences by next-generation-sequencing. Three replicates showed high correlation values for read counts for both day 0 (Figure 3.24.A) and day 7 samples (Figure 3.24.B).

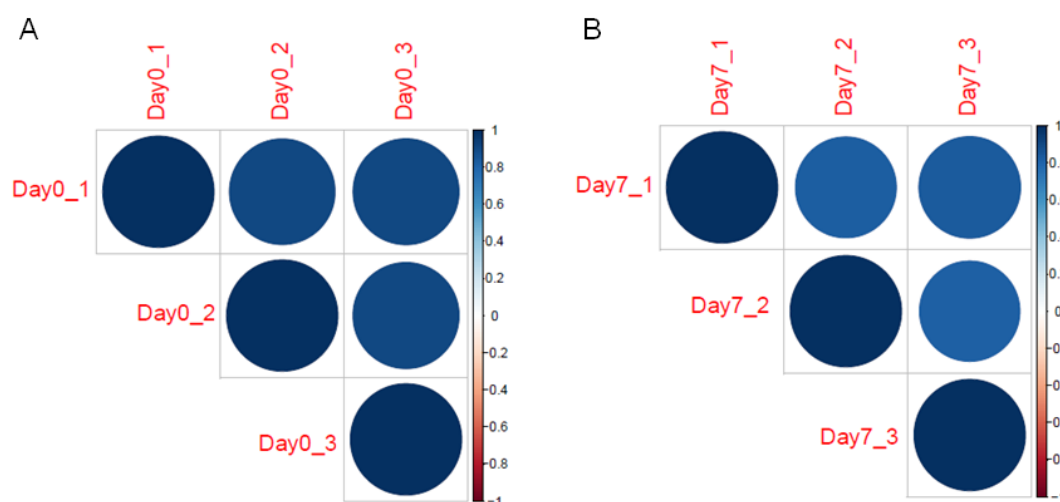


Figure 3.24. Correlation between technical replicates of low copy exogenous EGFP screen.

A) Three technical replicates of Day 0 samples were compared for their correlation values. The Pearson correlation coefficient values were used to calculate correlation. B) Three technical replicates of Day 7 EGFP negative sorted samples were compared for their correlation values. The Pearson correlation coefficient values were used to calculate correlation.

Read counts from NGS results were analyzed by MTC algorithm (

APPENDIX 9). Distribution of gene scores revealed a similar trend with the TRA-1-60 based screen for initiation results (Figure 3.25.A). Moreover, gNT scores placed in similar rankings (being at rank 87 in TRA-1-60 screens, and 78 in exogenous EGFP screen). However, results demonstrated that the differences between gRNA enrichments were less in exogenous EGFP screen compared to the TRA-1-60 screens. To eliminate possible perturbations leading to EGFP loss due to the interference to the transcription and translation, I compared hits from two different screen strategies. Among the top 25 enriched genes in transgene silencing screen and the TRA-1-60 based screen, there were 6 common genes. These were, DNMT3A, DNMT3L, PAXIP1, PRDM4, SETD2, and SMARCA4 (Figure 3.25.C). For the validation of these hits, I conducted independent reprogramming experiments with low copy LTR-EGFP dH1fs and tested the effect of gDNMT3A and gSETD2 infected cells on proviral silencing on day 7 by using two different gRNAs for each. Loss of both of the genes were able to further increase EGFP loss compared to the gNT; confirming the screen result (Figure 3.25.D).

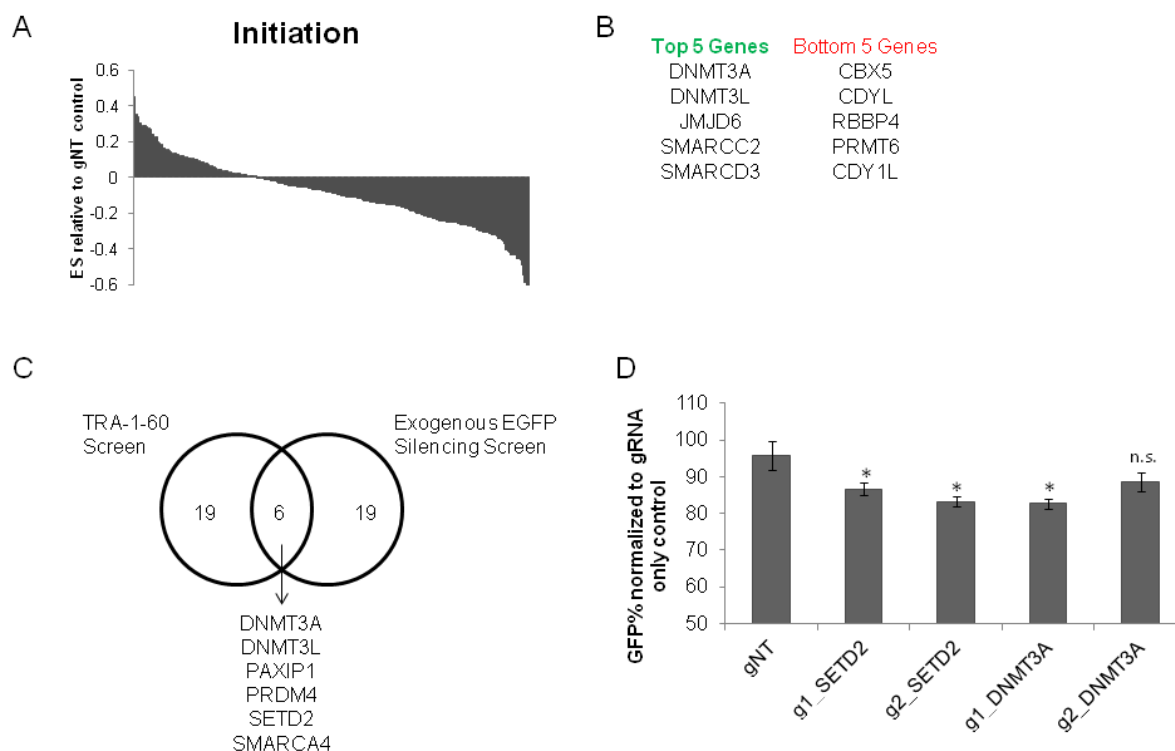


Figure 3.25. Low copy exogenous EGFP silencing based screen revealed regulators of reprogramming

A) Enrichment scores for low copy exogenous EGFP silencing screen were analyzed and quantified with MTC algorithm and sorted scores are plotted on bar graphs. B) Top and bottom five genes were listed. C) Top 25 genes in initiation stage from both screen strategies were overlapped to identify common regulators. D) Low copy LTR-EGFP-dH1fs were infected with gRNAs that target DNMT3A and SETD2 and cultured for at least 10 days followed by induction of pluripotency by OSKM infection. EGFP levels were assessed on Day 7 of reprogramming by flow cytometry. EGFP levels were normalized to the only gRNA infected controls. Statistical significance was evaluated by comparison of the the gNT control with samples. (* $p < 0.05$, n.s: not significant, $n=3$; error bars, $\pm$ s.d.).

3.5 Formulation of novel reprogramming strategies based on screen results

OSKM induced reprogramming has certain limitations due to its low efficiency and requires 21 days to generate iPSCs. Therefore, based on screen results, one can formulate novel reprogramming strategies to overcome such limitations by transiently using small molecule compounds targeting specific chromatin modifiers. For this, first, I tested whether the screen results can be reproduced by transient inhibition of single or multiple genes by small molecule compounds. I decided to use BRD1, 2, 4 inhibitor Bromosporine, EZH2 inhibitor

EPZ6438 (EZH2i), and HDAC1, 3 inhibitor MS-275 based on screen result, and treated dH1fs with these drugs only for the duration of a specific stage and also treated them throughout reprogramming (Figure 3.26.A). Treatments were mostly able to validate screen results either as increasing colony count or as increasing total stained area (Figure 3.26.B). For example, the screen identified EZH2 as one of the stage-specific barriers for the iPSC generation. Its initiation enrichment score is negative but its maturation enrichment score is positive. The small molecule experiment validated this result based on stained area percentage but not for emerging colony count. Area percentage in the stainings represent the total number of iPSC contrarily to the colony count. Moreover, the read counts from NGS results are quantified by cell count rather than colony count. Therefore, I speculated that high or low enrichment scores in the screen might be correlated to either emerging colony count or the average size of the colonies.

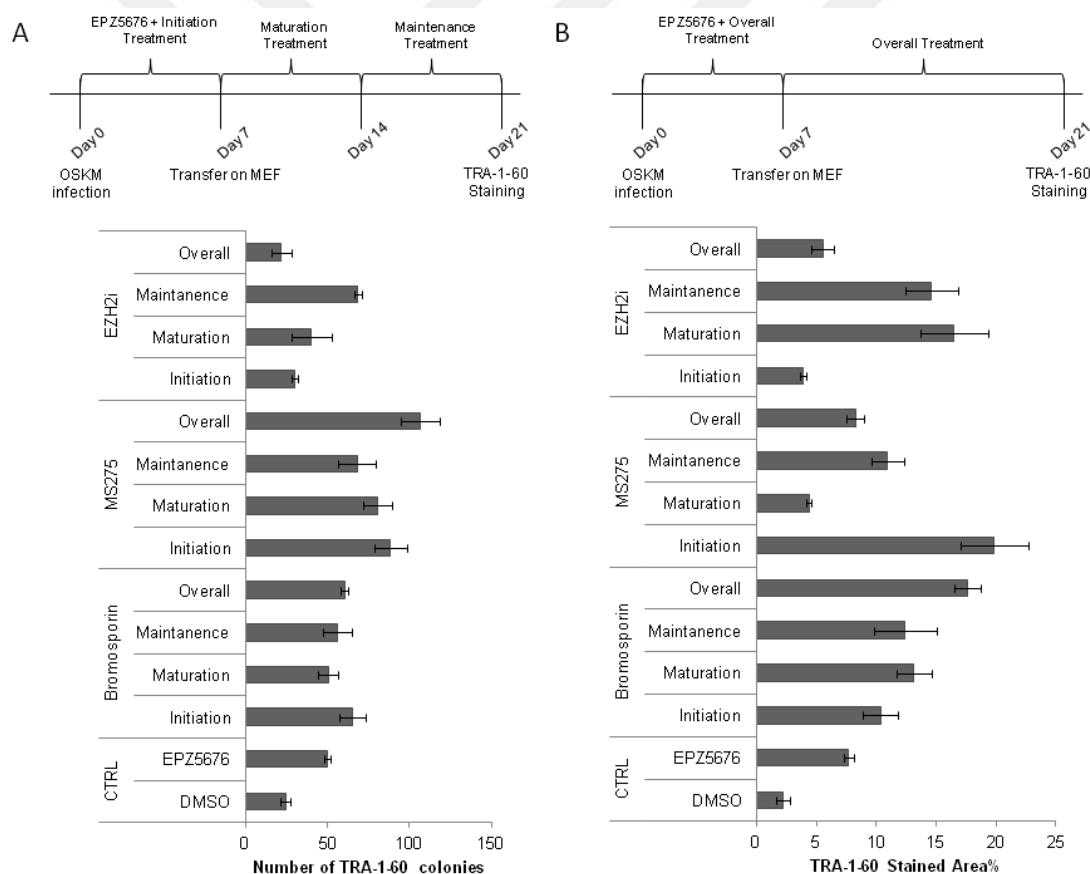


Figure 3.26 Validation of stage-specific screen results by small molecule inhibitors

A) Treatment and reprogramming timeline for testing stage-specific and overall screen results with small molecule inhibitors. Different drugs were used for each stage of reprogramming based on screen results. Treatment durations are indicated in the timeline. The EPZ5676 was

included for the first seven days in all conditions except for DMSO to mimic screen conditions. B) dh1fs were infected with OSKM and were treated with different small molecules for different stages and throughout reprogramming. Emerging iPSC colony count was determined by TRA-1-60 staining and quantified by ImageJ software. (n=3; error bars, $\pm$ s.d.).

Taken together, I devised a combination treatment to increase the efficiency of iPSC generation with the usage of small molecules. For that purpose, I treated dh1fs with Bromosporine and EPZ5676 during initiation, with EPZ6438 during maturation, and with MS-275 during maintenance stages (Figure 3.27.A). Emerging iPSC colonies were increased in number compared to the DMSO control (Figure 3.27..B). Therefore, I hypothesized that reprogramming can be accelerated to reduce 21-day requirement. For this, I repeated the same experiment with giving less duration for small molecule treatments and ended the experiment at Day 14 of reprogramming (Figure 3.27.C). Results showed that formulated combination even with 14 days of reprogramming can give rise to more iPSC cells, approximately 4 fold compared to the DMSO control sample (Figure 3.27.D).

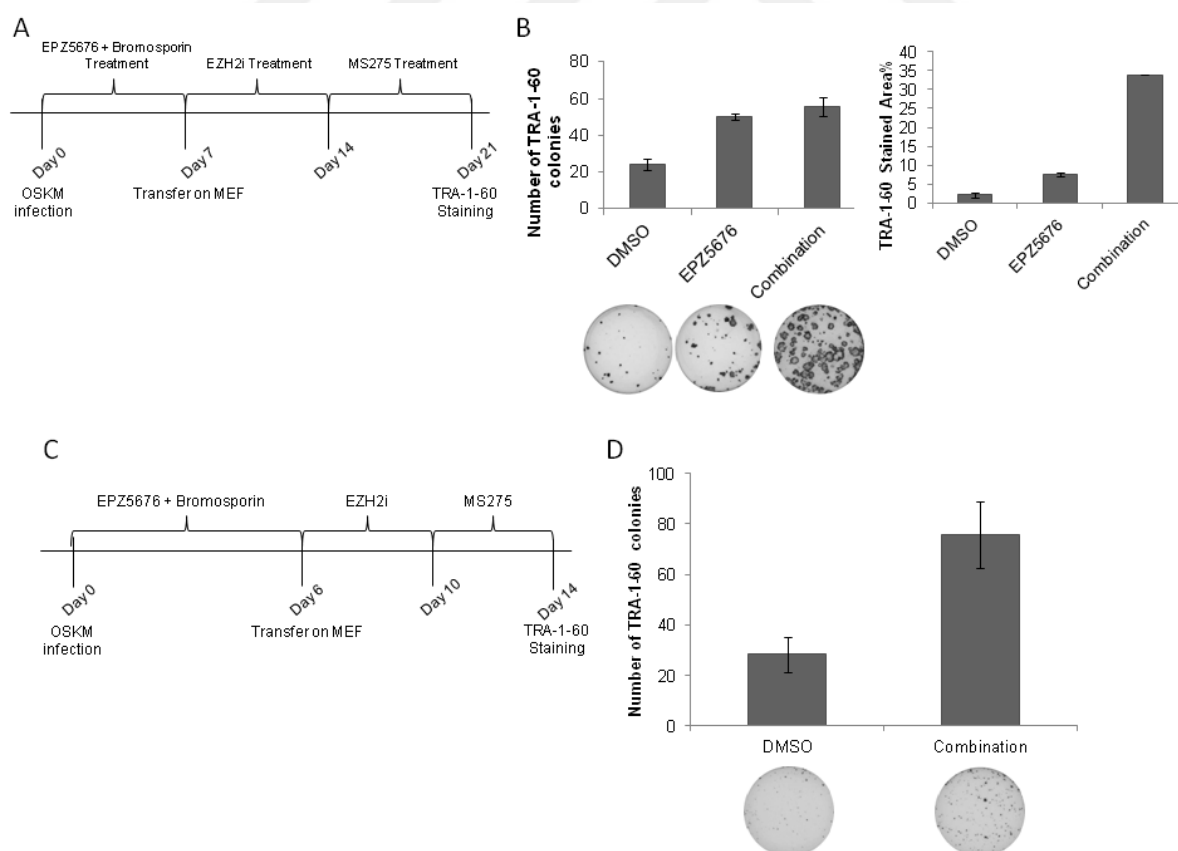


Figure 3.27 Novel reprogramming strategies can be formulated with screen results

A) Experimental plan for combination treatment. B) Based on the screen results, combination treatment was formulated. Bromosporine and EPZ5676 were treated only for between Day 0-7, EPZ6438 for Day 7-14, and MS275 for Day 14-21. All cells were cultured for 21 Days. TRA-1-60-HRP staining was performed at Day 21 and colony number and total stained area was quantified with ImageJ software (colony count difference $p < 0.05$, $n=2$; error bars, $\pm$ s.d.). C) Experimental plan for the combination treatment for 14 days reprogramming. D) Cells were treated with Bromosporine and EPZ5676 only for between Day 0-6, EPZ6438 for Day 6-10, and MS-275 for Day 10-14. All cells were cultured in total for 14 Days. TRA-1-60-HRP staining was performed at Day 14 and emerging colony number was quantified with ImageJ software ($n=3$; error bars, $\pm$ s.d.).

Acknowledgements

I would like to thank Zeynep Kaya, Gülben Gürhan, Kenan Sevinç, Tunç Morova, Dr. Nathan Lack, Dr. Mehmet Gönen and members of Önder Lab for their comments and discussions on these experiments. I also thank Gülben Gürhan, Kenan Sevinç, Ekin Deniz Aksu, and Cihan Civan Civas for their help with conducting experiments. Nathan Lack performed the quantifications of read counts from NGS results. Dr. Mehmet Gönen advised for screen data analysis and Tunç Morova implemented the algorithms computationally and helped with the further analysis of the data. Hazal Ayhan helped with the visualization of histone marks and overall ES for hits. Deniz Uğurlu helped with the puromycin resistance proviral silencing experiment. Zeynep Karakan helped with the sorting experiments. Dr. Tamer Onder, Dr. Tuğba Bağcı-Önder, Dr. Nathan Lack, Fidan Şeker, Ezgi Özyerli and the author of this thesis designed EPIKOL library and Dr. Tamer Önder cloned the library.

4 Discussion

In this discussion, I will initially point out the findings in this thesis that both validate known reports and provide novel information related to reprogramming, then discuss the crucial observations about the experiments, and suggest further work.

4.1 Findings that validate and propose reprogramming regulator chromatin modifier genes, complexes, and functions

Screen results were consistent with previously published reports. At the individual gene level, 18/19 genes that are known to be regulators of human reprogramming had significant scores in the reported function for reprogramming. PRC1 and PRC2 complexes were suggested as essential regulators for reprogramming both in previous reports and in the screen [57, 78]. However, the screen was not able to show a trend in H3K9 methylation as a barrier as previously reported [80]. Conceptually, proliferation as a requirement for the iPSC generation, and the link between knock-down of tumor suppressors and accelerated initiation stage were also demonstrated with our screen data [11, 41, 45].

Screen results also yielded novel information about regulatory genes, complexes, and conceptual information about the reprogramming. First, screen and further validation identified novel barriers as HDAC10, KMT2A, SETD2, and CARM1. There is little information about the function of HDAC10 in general. However, it has been shown to be upregulated during differentiation of ESC [116]. Therefore, its knock-out might be able to sustain proliferation further by inhibiting differentiation. KMT2A has been recently shown to promote naïve pluripotent state conversion from epiblast stem cells in mice. KMT2A absence has been found to facilitate activation of primed state genes [117]. Genes that have a similar function for such prior state conversions previously noted to be barriers for reprogramming. These genes include MBD3 and CHAF1 [60, 77, 118]. In addition to KMT2A, MLL1,2 complex has been suggested as a barrier. Therefore, KMT2A and MLL1,2 members might be working through a similar mechanism as a barrier to the reprogramming. CARM1 has been shown to be associated with sustaining pluripotency [119]. Therefore, increased iPSC generation in its absence is unexpected. Moreover, it has been shown to modulate SOX2 in mice [120]. It also has been shown to be associated with CBP, which is a part of a complex associated with depositing super-enhancer associated acetylation marks on histones [121]. Therefore, it might be interesting to investigate this gene as its loss might facilitate the loss of

somatic cell identity, but may, also hamper pluripotency at maintenance stage. Moreover, as small molecule inhibitors of CARM1 are currently available, its transient inhibition might also be useful to formulate novel reprogramming strategies [122, 123]. Finally, SETD2 is known with depositing H3K36me3 mark enabling successful elongation for genes [68]. DOT1L is another gene that deposits histone mark that allows successful elongation. Therefore, these two genes might be a barrier through maintaining similar mechanism. In the results, SETD2 and EPZ5676 treatment without OSKM infection was also found to downregulate fibroblast marker expression. Their absence might facilitate losing somatic cell identity as successful gene expression for maintaining cell identity would be impaired. Thus, it might be interesting to further validate the effect of SETD2 in preserving somatic cell identity also in different tissue types.

Screen results also identified essential regulators for reprogramming both in individual gene level and in complex level. We have found SMARCA5, ACTR, KDM8, and TAF1 are required for successful iPSC generation. SMARCA5 is a chromatin remodeler and have been shown to interact with core pluripotency factors such as OCT4 and SOX2 [124]. Another found essential regulator from the screen is ACTR, which is also shown to be a coactivator of an essential pluripotency gene, ESRRB [125]. Moreover, ACTR enables CARM1 and CBP to activate Nanog locus [126]. Therefore, it might be interesting to decipher CARM1 and ACTR axis to identify reverse functions of the absence of these genes for reprogramming. KDM8 is an H3K36me2 demethylase and shown to regulate cell cycle of ESC. Also, its depletion causes accumulation of P21 linked to G1 arrest which afterward dissociates the pluripotent stem cell state in human ESC [127]. Loss of function of KDM8 might be working in reverse function of the SETD2 absence in reprogramming. Although both of these genes regulate contrary function for H3K36me mark, we did not observe a significant clustering for erasers and writers for H3K36 mark. Therefore, it would be interesting to investigate targets of SETD2 and KDM8. Recently, another validated hit from the screen, TAF1 is demonstrated to be associated with BRD4 [128], a known essential regulator for reprogramming by enabling release of paused Pol II at promoters of pluripotency-associated genes. Therefore, TAF1 might be involved in the same process, and its absence might be causing similar problems as the BRD4 loss. In addition to these individual genes, major chromatin remodelers SWI/SNF and BAF are also identified as essential regulators of reprogramming. Both of these complexes are also known to regulate the self-renewal ability of ESC and absence of either one of them shown to cause loss of self-renewal ability [129–132]. Moreover, histone

acetyltransferase MOZ/MORF complex members are also proposed to be essential regulators for reprogramming. It has been well known, most of the HDAC inhibitors work as facilitators of reprogramming, and it might also be interesting to identify which regions needed to be acetylated by MOZ/MORF complex to induce pluripotency and whether same mechanisms are required for sustaining pluripotency in ESC.

For functional aspects, the screen results were consistent with the findings suggesting the proliferation is a requirement for successful iPSC generation. Moreover, increased proliferation rates are not directly linked to higher reprogramming efficiency. Gene expression levels of initial cell type and iPSC cells might give information about the barriers and essential regulators. As seen on the screen, the majority of the barriers have less expression in iPSCs compared to the fibroblasts. It might be interesting to investigate this feature of barriers in general scheme rather than an only limited set of chromatin modifiers.

Results about proviral silencing showed that it is a copy number dependent early stage event. Moreover, it is both affected by barrier genes and also passage number of dH1fs which are both linked to reprogramming efficiency. Recently, it has been found that endogenous retroviruses play crucial roles through developmental stages, particularly during preimplantation stage [133]. Therefore, it might be interesting to characterize proviral silencing mechanism further through reprogramming. Moreover, as demonstrated as an early-stage event in the results, it also suggests that currently widely used shRNAs in retroviral vectors for gene investigations are also prone to the silencing. Therefore, their silencing might be occurring in the early stage of reprogramming particularly and hinder similar results to be obtained. In this respect, using CRISPR-Cas9 system might be beneficial as it creates permanent DNA modification leading to the loss of function phenotype and can not be affected by the proviral silencing mechanism. Moreover, screen results based on exogenous EGFP silencing revealed similar hits as seen with TRA-1-60 screen in Day 7 results such as DNMT3A and SETD2. Additionally, genes that might have lethal effect could cause no EGFP phenotype for these hits. Therefore, comparison of these two screen could be used to identify true regulators of reprogramming.

Screen results also complement with the missing information about different genes. Recently, CHAF1A and CHAF1B has been shown to increase reprogramming efficiency in mice, contrarily, without showing any data, it has also been indicated that their knock-down does not follow the same trend in human reprogramming [77]. In the screen, these genes showed

decreased overall and proliferation ES. Secondly, assay types to estimate fold changes can influence different results to be obtained. For example, EZH2 inhibition has been both speculated about its ability to increase and decrease reprogramming in a different context [6, 57, 134, 135]. In the data, EZH2 inhibition with small molecule compound treatment was shown to not increase reprogramming efficiency by enabling more iPSC colony generation, rather, it accelerates the system and enables bigger colonies to be generated. Moreover, iDOT1L treatment has been shown to increase emerging colony count 3-4 fold [57]. Contrarily, emerging iPSC count as individual cells assessed by TRA-1-60-PE staining at Day 21 and found to have 8 fold difference. In addition, without comparing to the controls, iDOT1L can even be proposed as a chemical that can enable ~45% iPSC generation based on flow cytometry data. Therefore, assays that take into account of cell count such as using flow cytometry and assays that take into account of emerged iPSC colony count might obtain different results.

There are multiple pooled shRNA screens reported to identify regulators of reprogramming. Four of them were genome-wide, and two of them used shRNA libraries targeting chromatin modifiers [17, 55, 56, 77, 136, 137]. However, these screens were not able to identify similar barriers and essential regulators. Therefore, I did not compare results from those screens with screens in this thesis. Differences between these pooled screens might be due to the pooled screen conditions, species difference, or mode of inhibition type. The majority of these screens infected their pooled library with above MOI 1, indicating multiple modifications occur in a single cell. Therefore, their results might be harder to be interpreted as in the multiplexing CRISPR experiment for barrier genes. Moreover, some of these screens conducted in mice secondary reprogramming system. Thus, different genes might be identified in a context-dependent manner. Last, all of these screens utilized shRNAs contrarily, the CRISPR-Cas9 system had been used in this screen, and knock-down of genes might not be able to give similar phenotype when genes are knocked out.

4.2 Precautions and observations about the experiments

Throughout this thesis, my primary aim was to identify the role of chromatin modifiers during human iPSC generation by using the CRISPR-Cas9 system. The use of CRISPR-Cas9 to study reprogramming has several important drawbacks and the following have to be taken into account.

First, reprogramming experiments that include the CRISPR-Cas9 system in arrayed assays have high variability. This problem might be due to the technical errors, the low cleavage efficiency of Cas9, or non-deterministic nature of CRISPR-Cas9 in mutation generation. The low cleavage efficiency of Cas9 as a proposed mechanism for the reason of low reproducibility is ruled out by the fact that the cells were selected with selection marker after infection of the CRISPR components ensuring that all the cells in the population has Cas9 enough for endonuclease activity. Low reproducibility might also be due to the stochastic nature of both CRISPR-Cas9 system and iPSC generation process. Particularly, fitting one of this stochastic process into more deterministic state can help for obtaining more reproducible data. In the literature, there is currently only one report using CRISPR to investigate gene modification in the context of reprogramming [77]. In this report, secondary mice iPSC system was used to induce reprogramming. Therefore, the only stochastic method was CRISPR-Cas9. The gold standard method for CRISPR would be picking single cell colonies that have the same mutation resulting knock-out phenotype in all of the cells and then inducing reprogramming. However, with limited division rates, human fibroblasts are not suitable for such procedures. The reproducibility problem holds for the validation of the selected gRNAs experiments as gRNAs targeting same genes did not show the consistent phenotype. In the experiments for individual testing of hits, the gRNAs that decreases reprogramming efficiency or do not seem to be working show consistent phenotype, while those that cause fold increase have variation in fold changes. Besides, although screen data indicates DNMT3L and PRDM3 as barriers, I could not validate these hits in individual experiments. Furthermore, the genes that ranked close to gNT in the screen either subtly increased or decreased reprogramming efficiency in validation experiments. Scores of these genes was ranked in the noise area rather than a specific phenotype area as a barrier or essential regulator and this explanation might be the reason for the problem for reproducibility of the screen result.

Second, for the data analysis of screen results, I used a custom algorithm rather than a previously published algorithm. My initial trial with RIGER algorithm was not able to capture essential regulators [101]. I thought this might be due to several reasons: RIGER algorithm scores the highest two ranked gRNAs. If gRNAs of possible essential regulators might have the same scores as gNT in case the gRNAs were non-functional, RIGER was not able to identify essential regulators of reprogramming, leading to false-negative conclusion. Moreover, as genes in the EPIKOL library were not targeted by equal number of gRNAs, it is

highly probable for a gene which were targeted by more gRNAs to be identified as a hit because of their highly scored outlier gRNAs. I next used MaGECK algorithm instead [138]. In this analysis type, suggested hits from MaGECK such as BRD4 was not showing consistent enrichment in individual gRNA level, moreover, genes that have off-target linked enrichments at gRNA levels were placed in top hits. Therefore, this algorithm was not employed in the analysis. There are other algorithms that have been adapted for pooled screens [139–141]. However, all of these algorithms were tested for the screening data favoring drop-out phenotype. More specifically, these algorithms were devised to identify hits with the highest confidence from binary events. RIGER and MaGECK have been shown to be useful for such screens. On the contrary, in my screen, I wanted to determine gene scores relative to the non-targeting ones in both directions as enriched or depleted and therefore used a custom algorithm.

Third, there was a low reproducibility in read counts between TRA-1-60 screen replicates. Although the screens have a relatively high library representation (4750 cells/guide), initial cell number for the screens might not be enough due to the low reprogramming efficiency ranging in between 0.1-1%. However, this idea was disproved by the similar inconsistency between the proliferation screen results. This problem might be due to the technical errors such as reprogramming efficiencies of both screens were different, or unknown elements in the experiments. Particularly, in the second screen, gDOT1L was in higher scores in the initiation stage. The high gDOT1L score might be due to the possible use of non-functional EPZ5676 in the second screen. I would expect to see similar enrichment scores for gDOT1L in the initiation stage and the gNT, as observed in the first screen. Taken together, to limit increasing levels of false positives due to the differences between two biological replicates, I took the average of fold changes for gRNAs in both screens to be analyzed in MTC. Therefore, deconvolution of gene scores enabled ranking of those genes near the gNT area and preventing their use for further analysis.

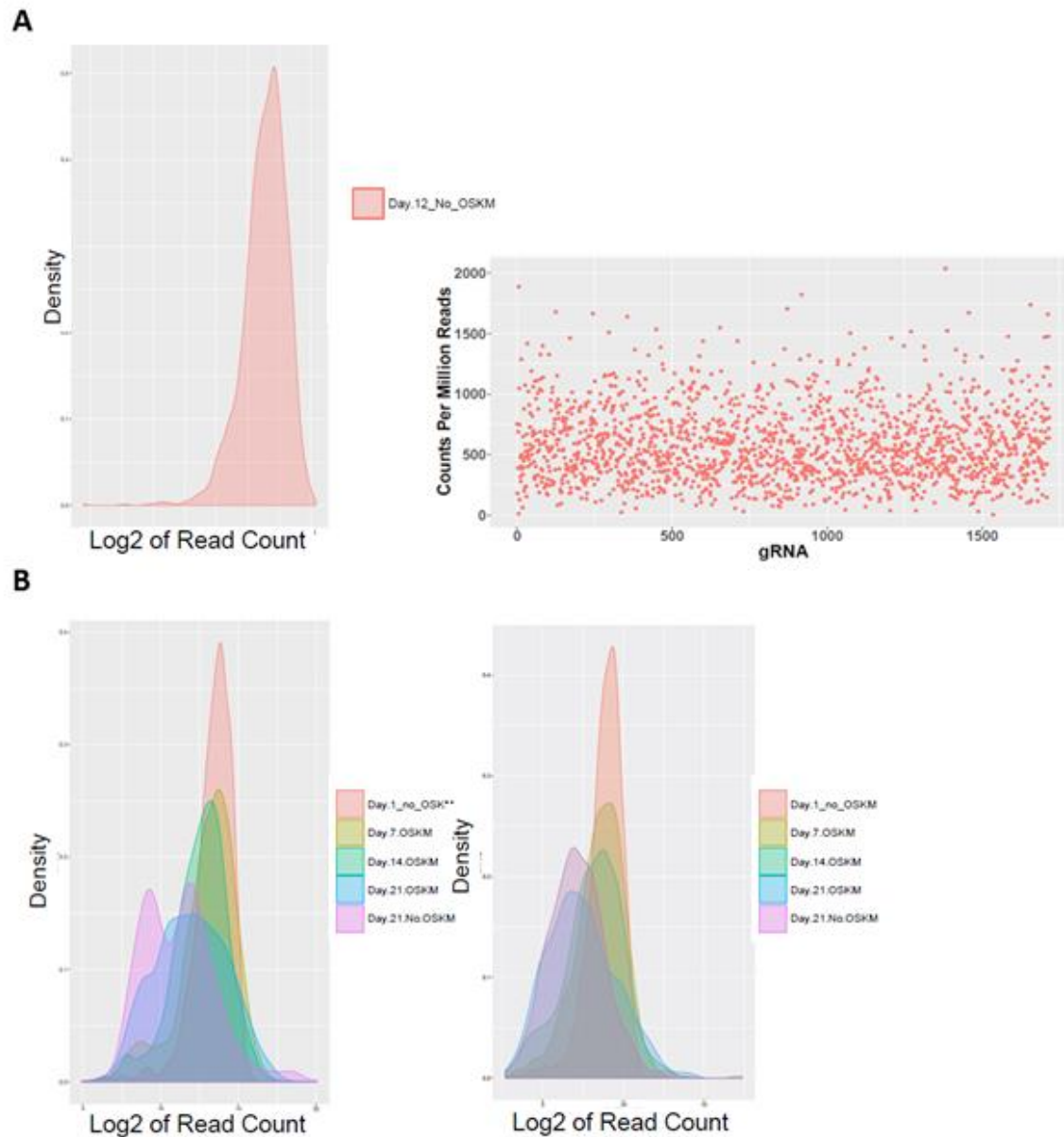
Fourth, although most of the gNT in the screens were clustered around same $\log_2$ score when their distribution is plotted, there was still a significant variance in the gNT distribution (data not shown). This variance might be due to the off-targets of the CRISPR-Cas9 system. In the screen results, MTC was able to convert this gRNA distribution to a gene score within the clustered range, but for arrayed experiments, multiple gNT controls should be included.

4.3 Future directions

First of all, mechanisms of any validated novel reprogramming regulator can be studied both in the context of their mechanism in reprogramming and their ability to sustain pluripotency in ESCs. Secondly, experimental strategy followed in this thesis was limited to OSKM induced pluripotency. It would be interesting to see whether same genes would be barriers or essential regulators if induction type is changed to OCT4, SOX2, NANOG, LIN28A or any other induction type. Additionally, changing the starting cell type to blood cells or any other tissue type and repeating the same screen setup might give different results and enable understanding of fibroblast specific barriers and reactivation of pluripotency-associated barriers.

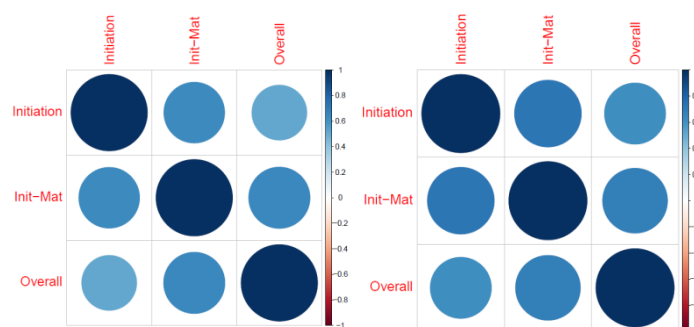
I initially planned to conduct an overexpression screen for the same set of genes with CRISPR-Cas9 synergistic activator mediators (SAM) system to identify barriers and enhancers of reprogramming [99]. However, I could not implement this system to dH1f cells due to the technical problems. For this approach, I first reproduced the data regarding the transient transfection of NANOG, LIN28A, KLF4, and C-MYC gRNAs and overexpression vectors, and successfully observed the overexpression of target genes by this system in 293t cell line as previously reported (APPENDIX 5.A). Next, gRNAs that target promoter sites for genes in the same set of chromatin modifiers were taken from a genome-wide SAM library and cloned into lenti sgRNA(MS2)_zeo vectors. As a future plan, conducting a screen with this system will enable the comparison of the hits to be verified from knockout screen results, leading to the identification of the true enhancers and barriers of the reprogramming.

5 Appendix



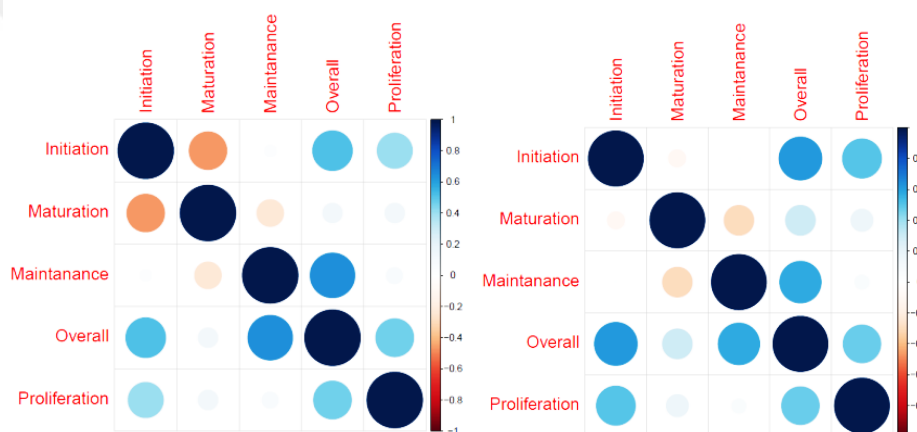
APPENDIX 1. Library represented through different time points in the screen.

A) Read count distribution for Day -12 is plotted to validate no loss of library representation during viral packaging and infection. B) Read count distribution for screen time points is plotted. (Left: First screen replicate, Right: Second screen replicate)



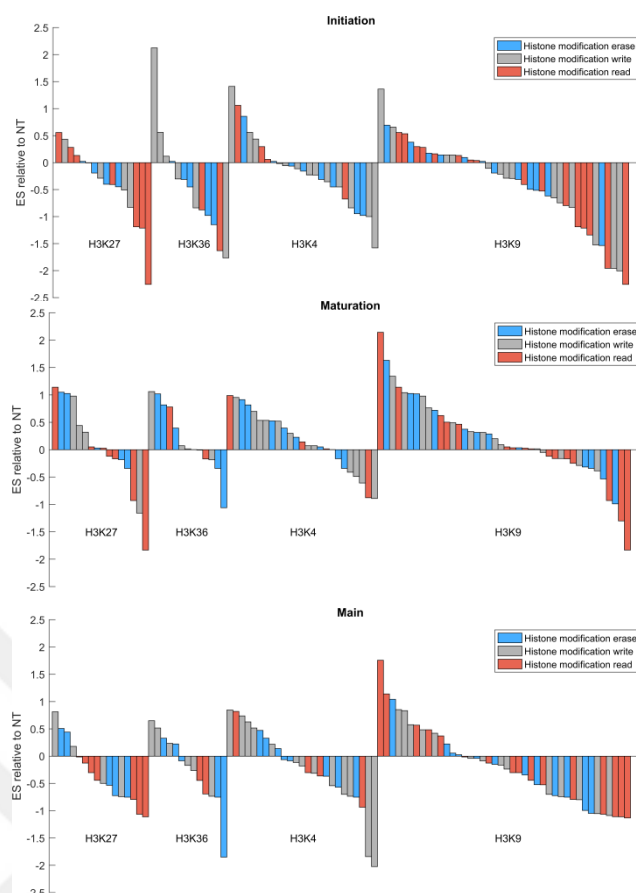
APPENDIX 2 Screen results showing consistency between overlapping time points in probe level.

A) Probe correlations to assess continuity of gene scores throughout reprogramming were visualized. Scores from overlapping time periods were Day 7/Day 0, Day 14/Day 0, Day 21/Day 0 (Left: First screen replicate, Right: Second screen replicate).



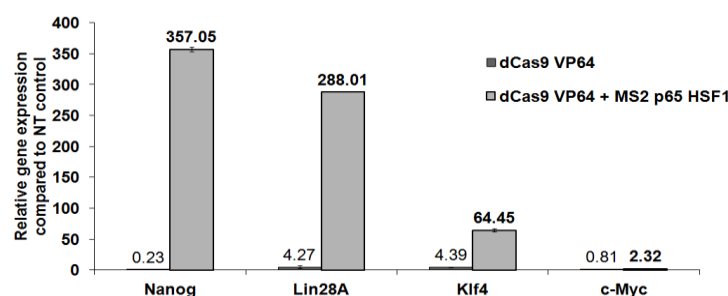
APPENDIX 3. Screen results identify stage-specific effects of modifications in probe level.

A) Probe correlations to assess stage-specific features of reprogramming regulators were visualized. Gene scores compared with different stages, overall and proliferation data (Left: First screen replicate, Right: Second screen replicate).



APPENDIX 4. Histone modifications for specific residues and their regulator writer, reader, erasers do not show specific clustering in a region for different stages of reprogramming.

A) Histone modifiers were color coded for their function as readers, writers, erasers and shown for main histone marks with their sorted initiation, maturation and maintenance scores. Genes are sorted among the presented histone marks.



APPENDIX 5. SAM system is functional in transient transfection experiments in 293ts.

A) 293ts were transfected with different SAM components and gRNAs targeting different genes. Cells were collected two days after transfection. Gene overexpression levels were quantified by qRT-PCR and normalized to gNT transfected cells (n=2; error bars, $\pm$ s.d.).

APPENDIX 6. SAM system protocol for 293t transfection.

20.000 293t were seeded in a 24-well plate and transfected with total 300 ng DNA compromising 1:1:1 mass ratio of gRNA plasmid with gene specific targeting sequence or pBS control plasmid:MS2-effector plasmid or pBS:dCas9 plasmid, dCas9-effector plasmid or pBS. 16-24 hours after transfection, GFP positive cell percentile was checked in RLL GFP transfected wells. If transfection efficiencies in 100 ng and 300 ng RLL GFP wells were both above >90%, the experiment was continued by changing the media. 48 hours after transfection, cells were harvested and total RNA was isolated for qRT-PCR.

APPENDIX 7. Full list of gRNAs in EPIKOL Library.

Library ID	Gene ID	Sequence
epikol_00001	ACAT1_AT_1	GCTTGGTTCATTGCAATTC
epikol_00002	ACAT1_AT_2	AGTGCTACAAGAACACCCAT
epikol_00003	ACAT1_AT_3	TGCAATGGAACCAAGCTTAG
epikol_00004	ACAT1_AT_4	ATGGAACCAAGCTTAGTGGC
epikol_00005	ACAT1_AT_5	CTGCCTAAAAAAGATCCAAT
epikol_00006	ACTL6B_1	TCAGTCCGCGCTGGGTACGC
epikol_00007	ACTL6B_2	CAGTCCGCGCTGGGTACGCT
epikol_00008	ACTL6B_3	GGCTCCTTCTCAGTCCGCGC
epikol_00009	ACTL6B_4	GCTCCTTCTCAGTCCGCGCT
epikol_00010	ACTL6B_5	AGTCCGCGCTGGGTACGCTG
epikol_00011	ACTR_PAS_1	TGTCTGGTAATAAAGCTCTC
epikol_00012	ACTR_PAS_2	GCTTTGAAGATATAATCCGA
epikol_00013	ACTR_PAS_3	CGGATTATATCTTCAAAGCC
epikol_00014	ACTR_PAS_4	TTCTTGATAGTGACGTTTCT
epikol_00015	ACTR_PAS_5	GTCTAAATGATGGGCAGTCA
epikol_00016	ASH1_BROMO_1	CTTTGAAGATCTGGGCTAGG
epikol_00017	ASH1_BROMO_2	TTTCTTTGAAGATCTGGGCT
epikol_00018	ASH1_BROMO_3	CACAAATTTCTTTGAAGATC
epikol_00019	ASH1_BROMO_4	ACAAATTTCTTTGAAGATCT
epikol_00020	ASH1_BROMO_5	TTTGAAGATCTGGGCTAGGC
epikol_00021	ASH1_SET_1	TCTAGAACGATTTGAGCTG
epikol_00022	ASH1_SET_2	AGGGGAGGTCGTCAGTGAAC
epikol_00023	ASH1_SET_3	GTTCACTGACGACCTCCCT
epikol_00024	ASH1_SET_4	CATCATTGAATACCTAGGGG
epikol_00025	ASH1_SET_5	GTTTCATCATTGAATACCTAG
epikol_00026	ASH2L_PHD_1	TGGTTCACGGCTGACACATT
epikol_00027	ASH2L_PHD_2	GAAGAGAATGGCCGACAGTT
epikol_00028	ASH2L_PHD_3	TGAAGAGAATGGCCGACAGT

epikol_00029	ASH2L_PHD_4	GAATGGCCGACAGTTGGGTG
epikol_00030	ASH2L_PHD_5	GCTCTACCTCACCCAAGTGT
epikol_00031	ASH2L_SPRY_1	GCCTCTCATGGAGTACGGAA
epikol_00032	ASH2L_SPRY_2	CCCAGTCTGGCAGCGGTATC
epikol_00033	ASH2L_SPRY_3	ACCTTTCCGTACTCCATGAG
epikol_00034	ASH2L_SPRY_4	CTCAGATGACCGGCTGACTG
epikol_00035	ASH2L_SPRY_5	AGTCTGGCAGCGGTATCTGG
epikol_00036	ASXL1_PHD_1	CTTTCAGGCTGCACGCACAC
epikol_00037	ASXL1_PHD_2	TTCTGTACAGATGACTGTAT
epikol_00038	ASXL1_PHD_3	GTTCACTGACAGCAGCACGG
epikol_00039	ASXL1_PHD_4	GACAGAACGCACCGCAGCCT
epikol_00040	ASXL1_PHD_5	AAAGCCATGATCATGTGCCA
epikol_00041	ASXL2_PHD_1	TCTTCGGAGAACGCCTGAAC
epikol_00042	ASXL2_PHD_2	ACGCCTGAACAGGAATGGTC
epikol_00043	ASXL2_PHD_3	AGGGGCCGATGCAATCATCA
epikol_00044	ASXL2_PHD_4	GCAGGAGACGCACAGTTTGG
epikol_00045	ASXL2_PHD_5	AAGCGGCAGTAACATTTCTGA
epikol_00046	ASXL3_PHD_1	AAGGACCTATGCAGTCGTCA
epikol_00047	ASXL3_PHD_2	TTCTGCCATGACGACTGCAT
epikol_00048	ASXL3_PHD_3	CTGTCCGCAAAGTTCTGAAC
epikol_00049	ASXL3_PHD_4	TATGCAGTCGTCATGGCAGA
epikol_00050	ASXL3_PHD_5	AATCCTGCCACAGGCTTGTC
epikol_00051	ATAD2_BROMO_1	CAACAGGCTTAGTAAACACT
epikol_00052	ATAD2_BROMO_2	TCTTAAGAAATGTTACACAT
epikol_00053	ATAD2_BROMO_3	TACAATCCAGATAGAGATCC
epikol_00054	ATAD2_BROMO_4	ATTACAGATGAAAGGTCCAT
epikol_00055	ATAD2_BROMO_5	TCTATCTGGATTGTATTCTA
epikol_00056	ATAD2B_BROMO_1	TAACATCTTCAGCAAACCGG
epikol_00057	ATAD2B_BROMO_2	GAGAGTTGCGGTTGTTTCTC
epikol_00058	ATAD2B_BROMO_3	TTCTCAGGGATGTAACCAAG
epikol_00059	ATAD2B_BROMO_4	CAGGGATGTAACCAAGAGGC
epikol_00060	ATAD2B_BROMO_5	GATGTTAAAGCGTTTATCTG
epikol_00061	ATRX_HELIC_1	TGAAATTCTTCGAATGGCAG
epikol_00062	ATRX_HELIC_2	TTCGAATGGCAGAGGAAATT
epikol_00063	ATRX_HELIC_3	CTTCAATGGCAGAGGAAAT
epikol_00064	ATRX_HELIC_4	TCTCTTTGAAATTCTTCGAA
epikol_00065	ATRX_HELIC_5	TCGAATGGCAGAGGAAATTG
epikol_00066	BAZ1A_BROMO_1	CGAAGGAGTTCTGGCCGACA
epikol_00067	BAZ1A_BROMO_2	TGGTACGACATGATGACAGC
epikol_00068	BAZ1A_BROMO_3	AGGAGTTCTGGCCGACAGGG
epikol_00069	BAZ1A_BROMO_4	ATTCATGAACTCCTCCCTGT
epikol_00070	BAZ1A_BROMO_5	TGAACAACCTGTTGTAGAAT

epikol_00071	BAZ1A_PHD_1	GAGCTTTGGTCGAACACAGT
epikol_00072	BAZ1A_PHD_2	TGCAAGATATGTGCAAAGAA
epikol_00073	BAZ1A_PHD_3	CTTTGTGATGGCTGTGATAG
epikol_00074	BAZ1A_PHD_4	TTCTTTGTGATGGCTGTGAT
epikol_00075	BAZ1A_PHD_5	GAAAACATGGTTCTTTGTGA
epikol_00076	BAZ1B_BROMO_1	AGCGGTACTTCACGATCTTG
epikol_00077	BAZ1B_BROMO_2	TCGTGAAGTACCGCTTCAGC
epikol_00078	BAZ1B_BROMO_3	GGTACTTCACGATCTTGTGG
epikol_00079	BAZ1B_BROMO_4	GCTGAGGTTTACAACGTCCG
epikol_00080	BAZ1B_BROMO_5	TGATGTGATCACGCACCCCA
epikol_00081	BAZ2A_BROMO_1	CTCACCAAACGTGGGTTCAC
epikol_00082	BAZ2A_BROMO_2	CGGTACCCACTCACCAAACG
epikol_00083	BAZ2A_BROMO_3	TGAACCCACGTTTGGTGAGT
epikol_00084	BAZ2A_BROMO_4	GGTACCCACTCACCAAACGT
epikol_00085	BAZ2A_BROMO_5	GTGAACCCACGTTTGGTGAG
epikol_00086	BAZ2A_MBD_1	CCTGGTATTATGGCCCCTGT
epikol_00087	BAZ2A_MBD_2	CCCACAGGGGCCATAATACC
epikol_00088	BAZ2A_MBD_3	AGGTCTCCCCCTGCCATCGG
epikol_00089	BAZ2A_MBD_4	GAGAGAGGTGCGCATCAAGA
epikol_00090	BAZ2A_MBD_5	GAAGGGCAGCCACCGATGGC
epikol_00091	BAZ2B_BROMO_1	GTAAACTTGAACTTGTTCC
epikol_00092	BAZ2B_BROMO_2	TGGAAACTCATGAGGATGCA
epikol_00093	BAZ2B_BROMO_3	GGAACAAGTTTCAAGTTTAC
epikol_00094	BAZ2B_BROMO_4	CTAATTGTGAAAAATCCAT
epikol_00095	BAZ2B_BROMO_5	ACTTAGTTTCTCTCTAATTG
epikol_00096	BAZ2B_HMBD_1	AACTTTGGAGGGCGCCTTCA
epikol_00097	BAZ2B_HMBD_2	GTAGCATATTATGCTCCATG
epikol_00098	BAZ2B_HMBD_3	AATATGCTACTTCTCCTTGA
epikol_00099	BAZ2B_HMBD_4	CTCCATGTGGAAGAAACTT
epikol_00100	BAZ2B_HMBD_5	TGCCTAAGTTTCTTTCCACA
epikol_00101	BMI_RING_1	TAGACATTCTATTATGGTTG
epikol_00102	BMI_RING_2	GGAATGTAGACATTCTATTA
epikol_00103	BMI_RING_3	AACGTGTATTGTTTCGTTACC
epikol_00104	BMI_RING_4	GACAATACTTGCTGGTCTCC
epikol_00105	BMI_RING_5	ACAAATAGGACAATACTTGC
epikol_00106	BPTF_BROMO_1	GGATTATGAGGGGTTGAAGA
epikol_00107	BPTF_BROMO_2	AGGATTATGAGGGGTTGAAG
epikol_00108	BPTF_BROMO_3	GGTGCATCATTAGGGTCTAC
epikol_00109	BPTF_BROMO_4	CATTAGGGTCTACTGGTTCA
epikol_00110	BPTF_BROMO_5	GAAAGGCCAGGCCATCTTAT
epikol_00111	BPTF_GECKO_3	GTTCCGCAGTACCTCGTAAA
epikol_00112	BPTF_GECKO_4	ACGCTATCTTTCAGATCAGC

epikol_00113	BPTF_GECKO_5	ACCATTAGCTGCCGCCAGAA
epikol_00114	BPTF_PHD_1	TGGTACCATGGGCGCTGCGT
epikol_00115	BPTF_PHD_2	AGATGCCAACGCAGCGCCCA
epikol_00116	BRD1_BROMO_1	GCTGCGCAAATATCCTGGCG
epikol_00117	BRD1_BROMO_2	TGCGCAGCCCCGTGAGTCTGA
epikol_00118	BRD1_BROMO_3	GGGCTGCGCAAATATCCTGG
epikol_00119	BRD1_BROMO_4	TGCAAGACAAGGACCCCGCC
epikol_00120	BRD1_BROMO_5	GAGCGCAGCAGCACCGTCAG
epikol_00121	BRD1_PHD_1	AAAAGACAGATGACGACCGC
epikol_00122	BRD1_PHD_2	GTGTTTCATCGAGCCCATCGA
epikol_00123	BRD1_PHD_3	GGGCTCGATGAACACCGTGT
epikol_00124	BRD1_PHD_4	AAAGACAGATGACGACCGCT
epikol_00125	BRD1_PHD_5	TGTTCATCGAGCCCATCGAT
epikol_00126	BRD2_BROMO_1	ATCAGTTCGCATGGCCATTC
epikol_00127	BRD2_BROMO_2	TGTGGAAACATCAGTTCGCA
epikol_00128	BRD2_BROMO_3	CCTTGTGTAGGTATTGCAGC
epikol_00129	BRD2_BROMO_4	GCCTGTGGATGCTGTCAAAC
epikol_00130	BRD2_BROMO_5	CCTGTGGATGCTGTCAAAC
epikol_00131	BRD3_BROMO_1	ATTTGATTGCGTCCACGGGC
epikol_00132	BRD3_BROMO_2	TTCAATTTGATTGCGTCCAC
epikol_00133	BRD3_BROMO_3	GTTCAATTTGATTGCGTCCA
epikol_00134	BRD3_BROMO_4	TCTGGAAACACCGAGTTCGCC
epikol_00135	BRD3_BROMO_5	TGCGTCCACGGGCTGGTAGA
epikol_00136	BRD4_BROMO_1	TTCAGCTTGACGGCATCCAC
epikol_00137	BRD4_BROMO_2	GAGTGGTGCTCAAGACACTA
epikol_00138	BRD4_BROMO_3	GCTGGAAAGGCCATGCCAAAC
epikol_00139	BRD4_BROMO_4	CTCTGAGCAGGTATTGCAGT
epikol_00140	BRD4_BROMO_5	GAGCAGGTATTGCAGTTGGT
epikol_00141	BRD7_GECKO_1	AACTGATGAGACAATTGCAG
epikol_00142	BRD7_GECKO_2	GTTATTCCAACCTTCTTGTT
epikol_00143	BRD7_GECKO_3	AGCTCTTTAGCCAAACAAGA
epikol_00144	BRD7_GECKO_4	AGTTACCTTTAGTTCTTCTA
epikol_00145	BRD7_GECKO_5	GTTGATTCAAAGCTTCTTGA
epikol_00146	BRD8_BROMO_1	AGCTCTCCATACAAGCATGA
epikol_00147	BRD8_BROMO_2	GTTACAGATGACATAGCACC
epikol_00148	BRD8_BROMO_3	GGTGCTATGTCATCTGTAAC
epikol_00149	BRD8_BROMO_4	TGTCATCTGTAACAGGCTGC
epikol_00150	BRD8_BROMO_5	TGCACAATGCTGTGGTAGCC
epikol_00151	BRD9_BROMO_1	GCTGGCGGAGGAAGTGTTC
epikol_00152	BRD9_BROMO_2	TGTTCCAGGAGTTGCTGAAT
epikol_00153	BRD9_BROMO_3	GGAGCAATTGCATCCGTGAC
epikol_00154	BRD9_BROMO_4	GTCACGGATGCAATTGCTCC

epikol_00155	BRD9_BROMO_5	TTCATGGTGCCAAAATCCAT
epikol_00156	BRDT_BROMO_1	ATGGCCCTTTCAACGTCCTG
epikol_00157	BRDT_BROMO_2	GCATCCACAGGACGTTGAAA
epikol_00158	BRDT_BROMO_3	AGCATCCACAGGACGTTGAA
epikol_00159	BRDT_BROMO_4	TGTAGTTTCACAGCATCCAC
epikol_00160	BRDT_BROMO_5	AAGTTGTCCTAAAGGATTTA
epikol_00161	BRPF1_BROMO_1	CTGCCTCAAGTATAACGCCA
epikol_00162	BRPF1_BROMO_2	TCTGTCTGAGGTAACCGAAT
epikol_00163	BRPF1_BROMO_3	TTGATGTGGTCTAGGTAGTC
epikol_00164	BRPF1_BROMO_4	AGGCAACATCTTCAGCGAGC
epikol_00165	BRPF1_BROMO_5	CATCCTCCTTCGAAAACCT
epikol_00166	BRPF1_PHD_1	AGCAGACAGATGACGGGCGC
epikol_00167	BRPF1_PHD_2	GCAGACAGATGACGGGCGCT
epikol_00168	BRPF1_PHD_3	ACCGCCCTTGTTGGGGCACA
epikol_00169	BRPF1_PHD_4	TTTGCAAACAACGGGGCTCA
epikol_00170	BRPF1_PHD_5	ATTTGCAAACAACGGGGCTC
epikol_00171	BRPF3_BROMO_1	TGTTCTGTTGAGGACAACAC
epikol_00172	BRPF3_BROMO_2	GACAACACTGGACCTGCTGC
epikol_00173	BRPF3_BROMO_3	GTCCTCAACAGAACATTGAA
epikol_00174	BRPF3_BROMO_4	CAGGATCCTTCTCCTGCAGC
epikol_00175	BRPF3_BROMO_5	ACTGGACCTGCTGCAGGAGA
epikol_00176	BRPF3_PHD_1	TTGAGGCTGGTTTCGCGCAT
epikol_00177	BRPF3_PHD_2	ATTGAGGCTGGTTTCGCGCA
epikol_00178	BRPF3_PHD_3	TTGAAGGCGCCACCCCTATT
epikol_00179	BRPF3_PHD_4	ACAATATCCCCTGCCCCGC
epikol_00180	BRPF3_PHD_5	CGCCACCCTTATTGGGGCAA
epikol_00181	BRWD1_BROMO_1	GGATATTCAACCAAATCAAC
epikol_00182	BRWD1_BROMO_2	ATTTAGACAACCTGTTGATT
epikol_00183	BRWD1_BROMO_3	AAATCAACAGGTTGTCTAAA
epikol_00184	BRWD1_BROMO_4	CAACTGGAAGAAACAGTGTA
epikol_00185	BRWD1_BROMO_5	GTAAGGGAAACTCTAGATGC
epikol_00186	BRWD3_BROMO_1	AGAATGTGAACGGGTTATTC
epikol_00187	BRWD3_BROMO_2	CCGTTACATTCTTCGTCTC
epikol_00188	BRWD3_BROMO_3	GAATGTGAACGGGTTATTCA
epikol_00189	BRWD3_BROMO_4	CAGAGACGAAGAATGTGAAC
epikol_00190	BRWD3_BROMO_5	CAAACCCAGGAAGGAGAGT
epikol_00939	C14orf169_CUPIN_1	TCGTCCGTTGATGTAGCGAG
epikol_00940	C14orf169_CUPIN_2	GACGCCGCTCGCTACATCAA
epikol_00941	C14orf169_CUPIN_3	CCCCCACTACGACGACATCG
epikol_00942	C14orf169_CUPIN_4	GCCTCGATGTCGTCGTAGTG
epikol_00943	C14orf169_CUPIN_5	TCCATGAAAGCGGTTCTTCG
epikol_00191	CARM1_MTA_1	AGCACGGAAAATCTACGCGG

epikol_00192	CARM1_MTA_2	ACGGAAAATCTACGCGGTGG
epikol_00193	CARM1_MTA_3	CTTGGGCGGCAAAAAACGAC
epikol_00194	CARM1_MTA_4	TGGAGCACGAAAAATCTACG
epikol_00195	CARM1_MTA_5	TGACCACGATGCGGTCCGTC
epikol_00196	CBX1_GECKO_1	GGCCGACTCACCATTTCATC
epikol_00197	CBX1_GECKO_2	AAAAGTTCTCGACCGTCGAG
epikol_00198	CBX1_GECKO_3	CTTTGCCCTTTACCACTCGA
epikol_00199	CBX1_GECKO_4	CTGCAGGAAAACTCTGATG
epikol_00200	CBX1_GECKO_5	TTACAGTCAGAAAAGCCACG
epikol_00201	CBX2_CHROM_1	ACCTGGTCAAGTGGCGCGGC
epikol_00202	CBX2_CHROM_2	GAGTACCTGGTCAAGTGGCG
epikol_00203	CBX2_CHROM_3	ACCAGCCGCGCCACTTGACC
epikol_00204	CBX2_CHROM_4	GGAAGGCCAGGAGCAGCCTC
epikol_00205	CBX2_CHROM_5	TGGAAGGCCAGGAGCAGCCT
epikol_00206	CBX3_GECKO_1	AACACAGTGCTGACAATACT
epikol_00207	CBX3_GECKO_2	AGACTTGGTGCTGGCGAAAG
epikol_00208	CBX3_GECKO_3	AAACAGGCTGACAAACCAAG
epikol_00209	CBX3_GECKO_4	AGAAACGCTTCAATCAATTC
epikol_00210	CBX3_GECKO_5	AAACAGGAAAGATTCAAGATG
epikol_00211	CBX4_CHROM_1	GAGTATCTGGTGAAATGGAG
epikol_00212	CBX4_CHROM_2	ATCTGGTGAAATGGAGAGGC
epikol_00213	CBX4_CHROM_3	TGCTGATCGCTTCCAGAAC
epikol_00214	CBX4_CHROM_4	TGGAAGGCGATCAGCAGCCT
epikol_00215	CBX4_CHROM_5	GGAAGGCGATCAGCAGCCTG
epikol_00216	CBX5_CHROM_1	CTAGACAGGCGCGTGGTTAA
epikol_00217	CBX5_CHROM_2	GCTAGACAGGCGCGTGGTTA
epikol_00218	CBX5_CHROM_3	GAAGGTGCTAGACAGGCGCG
epikol_00219	CBX5_CHROM_4	GCGCGTGGTTAAGGGACAAG
epikol_00220	CBX5_CHROM_5	AAGTGGAATATCTACTGAAG
epikol_00221	CBX6_CHROM_1	TCGGATCCGCCGTTTGATGA
epikol_00222	CBX6_CHROM_2	CCGCCGTTTGATGATGGATT
epikol_00223	CBX6_CHROM_3	CCGAATCCATCATCAAACGG
epikol_00224	CBX6_CHROM_4	CGGCCGAATCCATCATCAAA
epikol_00225	CBX6_CHROM_5	GCATCGAGTACCTGGTGAAA
epikol_00226	CBX7_CHROM_1	GCATCCGGAAGAAGCGCGTG
epikol_00227	CBX7_CHROM_2	AAGTCGAGTATCTGGTGAAG
epikol_00228	CBX7_CHROM_3	ATCTGGTGAAGTGAAAGGA
epikol_00229	CBX7_CHROM_4	TCGTAGGCCATGACGAGGCG
epikol_00230	CBX7_CHROM_5	TCTCCTCGTAGGCCATGACG
epikol_00231	CBX8_CHROM_1	CCCTCCTGAAGCGGCGCATA
epikol_00232	CBX8_CHROM_2	TCCGTATGCGCCGCTTCAGG
epikol_00233	CBX8_CHROM_3	CCGTATGCGCCGCTTCAGGA

epikol_00234	CBX8_CHROM_4	CGGCCGAAGCCCTCCTGAAG
epikol_00235	CBX8_CHROM_5	GCGCCGCTTCAGGAGGGCTT
epikol_00236	CDY1_GECKO_1	TTAACATGGATTCTATCCCT
epikol_00237	CDY1_GECKO_2	GTCTTTCCTCATGGCTTCCC
epikol_00238	CDY1_GECKO_3	CTGATCATTCTGTGTCTTTC
epikol_00239	CDY1_GECKO_4	ATATCCTAGAATACCTCTGT
epikol_00240	CDY1_GECKO_5	CTCTGTCCAAAGGTCGTATA
epikol_00241	CDY1L_CHROM_1	GTGGAAAGGCTATGACACGG
epikol_00242	CDY1L_CHROM_2	ACAGCACCTCGTGAACCTGTG
epikol_00243	CDY1L_CHROM_3	CGAGGACGACACTTGGGAGC
epikol_00244	CDYL_GECKO_4	AAAGCCGGTCGGAGCTTTAT
epikol_00245	CDYL_GECKO_5	TACGACGTCTGACAGATGAC
epikol_00246	CHAF1A_1	TCTTCCACCAGCTCGCCCGA
epikol_00247	CHAF1A_2	CTTCCACCAGCTCGCCCGAG
epikol_00248	CHAF1A_3	CTCTTCCACCAGCTCGCCCG
epikol_00249	CHAF1A_4	TGAGTACAGAGTCTTCTTCG
epikol_00250	CHAF1A_5	AGGCGGCCCTCGGGCGAGC
epikol_00251	CHAF1B_1	ACACATCGAGGGTCTTCGCC
epikol_00252	CHAF1B_2	CCTGAGGGGGAGTCGTGCCG
epikol_00253	CHAF1B_3	AGGGGTTCCTCTACCGGTC
epikol_00254	CHAF1B_4	GGGGTTCCTCTACCGGTCT
epikol_00255	CHAF1B_5	GCCTGAGGGGGAGTCGTGCC
epikol_00256	CHD1_CHROM_1	GGTCCCATATCCACAACACT
epikol_00257	CHD1_CHROM_2	CTCCCAAGTGTGTGGATAT
epikol_00258	CHD1_CHROM_3	CTCAAGCAGCAGAATGTTAG
epikol_00259	CHD1_CHROM_4	GTCCCATATCCACAACACTT
epikol_00260	CHD1_CHROM_5	TCTAACATTCTGCTGCTTGA
epikol_00261	CHD1_HELIC_1	AGGCTAGAGCCCATCGAATT
epikol_00262	CHD1_HELIC_2	TTGTTATATTTGATTCCGAT
epikol_00263	CHD1_HELIC_3	ACAGAATGATCTTCAGGCAC
epikol_00264	CHD1_HELIC_4	CCCTAGACCTCCAGCTCTTG
epikol_00265	CHD1_HELIC_5	GATCAATAAAAGGAGAACTG
epikol_00266	CHD1L_HELIC_1	ATGCTAGTAGCTTATCCAGC
epikol_00267	CHD1L_HELIC_2	CTACTAGCATTCTGTATTCT
epikol_00268	CHD1L_HELIC_3	TACTAGCATTCCTGTATTCT
epikol_00269	CHD1L_HELIC_4	CTTGAGAATATCCAACATC
epikol_00270	CHD1L_HELIC_5	TATCCAACATCTGGGTCATT
epikol_00271	CHD2_CHROM_1	GAAGCTAATGGCGACCCTAG
epikol_00272	CHD2_CHROM_2	CCTTCCACTTGATGAGGTAC
epikol_00273	CHD2_CHROM_3	GTATATGCGATTGAAGCTAA
epikol_00274	CHD2_CHROM_4	ACCAACCCTTCCACTTGATG
epikol_00275	CHD2_CHROM_5	CAGTGTCAAAGTCACCACTA

epikol_00276	CHD2_HELIC_1	GGTGAGAATGTTGGATATCC
epikol_00277	CHD2_HELIC_2	GTTGACAAGACTTCGAGAAA
epikol_00278	CHD2_HELIC_3	TAATAGTTAGGTATTTCAGCC
epikol_00279	CHD2_HELIC_4	TTGACAAGACTTCGAGAAAG
epikol_00280	CHD2_HELIC_5	TGACAAGACTTCGAGAAAGG
epikol_00281	CHD3_CHROM_1	TACCATATGACCAGTCCACG
epikol_00282	CHD3_CHROM_2	ACCATATGACCAGTCCACGT
epikol_00283	CHD3_CHROM_3	TCCCACGTGGACTGGTCATA
epikol_00284	CHD3_CHROM_4	ATATGACCAGTCCACGTGGG
epikol_00285	CHD3_CHROM_5	CATCTTCCTCCCACGTGGAC
epikol_00286	CHD3_HELIC_1	TCAGCTTTCGCAGCATCTTC
epikol_00287	CHD3_HELIC_2	GAAGATGCTGCGAAAGCTGA
epikol_00288	CHD3_HELIC_3	CTGCGAAAGCTGAAGGAGCA
epikol_00289	CHD3_HELIC_4	TACAAGTATGAGCGCATCGA
epikol_00290	CHD3_HELIC_5	GCGCATCGATGGTGGTATCA
epikol_00291	CHD4_PHD_1	GGATGATCTCACCGCCTTGC
epikol_00292	CHD4_PHD_2	CACGGGGACAGGTATCACAC
epikol_00293	CHD4_PHD_3	CCTTCTCCATGTCTGGGATCC
epikol_00294	CHD4_PHD_4	CGACATGGAGAAGGCTCCCG
epikol_00295	CHD4_PHD_5	CAGACCATGTGGTAAGCACG
epikol_00296	CHD5_CHROM_1	GTCGATGTCATCGATCTCCC
epikol_00297	CHD5_CHROM_2	TGCTTGAGGTGTCTGTAGTA
epikol_00298	CHD5_CHROM_3	TCCCAGGTGCACTGGTCGTA
epikol_00299	CHD5_CHROM_4	CTACTACGACAACCTCAAGC
epikol_00300	CHD5_CHROM_5	AGGTGCACTGGTCTGAGGGC
epikol_00301	CHD5_PHD_1	GGCCTACCATCTCGTATGCC
epikol_00302	CHD5_PHD_2	ATACGAGATGGTAGGCCCTC
epikol_00303	CHD5_PHD_3	CCCTCGGGCAGGTGTCTGCAC
epikol_00304	CHD5_PHD_4	CCTGTGCGACACCTGCCCGA
epikol_00305	CHD5_PHD_5	GTCCAGGCATACGAGATGGT
epikol_00306	CHD6_CHROM_1	CTTGAGGTGGCCACACCA
epikol_00307	CHD6_CHROM_2	CTGTTTCTGCATCCTTGGTG
epikol_00308	CHD6_CHROM_3	TGTTTCTGCATCCTTGGTGT
epikol_00309	CHD6_CHROM_4	CACACCAAGGATGCAGAAAC
epikol_00310	CHD6_CHROM_5	ACACCAAGGATGCAGAAACA
epikol_00311	CHD6_HELIC_1	GGTAATCTTCTAGGATGTCTG
epikol_00312	CHD6_HELIC_2	CTACTCCCTAAGCTGATTGC
epikol_00313	CHD6_HELIC_3	TGGCCACCTGCAATCAGCTT
epikol_00314	CHD6_HELIC_4	CTCCCTAAGCTGATTGCAGG
epikol_00315	CHD6_HELIC_5	GGCCACCTGCAATCAGCTTA
epikol_00316	CHD7_CHROM_1	GCACGTAGCACAGATGACCG
epikol_00317	CHD7_CHROM_2	TACGTGCAAAGTCCATTATC

epikol_00318	CHD7_CHROM_3	TGCACGTAGCACAGATGACC
epikol_00319	CHD7_CHROM_4	TTGCACGTAGCACAGATGAC
epikol_00320	CHD7_CHROM_5	CATAGATCAAGCAAAGATCG
epikol_00321	CHD7_HELIC_1	GGTGCGCTGCTTGGACATAC
epikol_00322	CHD7_HELIC_2	TTCCCAGATGGTGCCTGCT
epikol_00323	CHD7_HELIC_3	AGGATCGACGGCCGAGTAAG
epikol_00324	CHD7_HELIC_4	CAAGCTGCTGCCAAAAGTGA
epikol_00325	CHD7_HELIC_5	CTGCTGCCAAAAGTGAAGGC
epikol_00326	CHD8_CHROM_1	GCTTCAGTATATTGTCCAGA
epikol_00327	CHD8_CHROM_2	CTTCAGTATATTGTCCAGAA
epikol_00328	CHD8_CHROM_3	TACTATCTCCCAACTAGAGA
epikol_00329	CHD8_CHROM_4	AGCCCATTCACAATGCAGAT
epikol_00330	CHD8_CHROM_5	CCTCTTATCCTTCTCTAGTT
epikol_00331	CHD8_HELIC_1	GGTTCGTTAGCCGGCAAAC
epikol_00332	CHD8_HELIC_2	TGGCCACCAGCTTTAAGCTT
epikol_00333	CHD8_HELIC_3	GTCATAAGAACCAGTTTGC
epikol_00334	CHD8_HELIC_4	AATAATCCTCTAGGATGTCT
epikol_00335	CHD8_HELIC_5	TTGCTTCCAAAGCTTAAAGC
epikol_00338	CHD9_CHROM_3	GGCCCACTCACAGTGAAGAT
epikol_00339	CHD9_CHROM_5	TTTCAAAAGCTGCTCTTCTG
epikol_00336	CHD9_GECKO_1	TATCTAAAGCAAGAAGTTAT
epikol_00337	CHD9_GECKO_2	CTCTATATCATCTTTAAAGA
epikol_00340	CHD9_GECKO_4	CATCTTTAAAGATGATATAG
epikol_00341	CHD9_HELIC_1	GGTTCGTTGCCTTGACATTC
epikol_00342	CHD9_HELIC_2	CTTCCCAAAATGAAAGCCGG
epikol_00343	CHD9_HELIC_3	GATGAGCACTTTATGACCTC
epikol_00344	CHD9_HELIC_4	ATGACCTCCGGCTTTTCATT
epikol_00345	CHD9_HELIC_5	GATAGTCCTCCAGAATGTCA
epikol_00346	CLOCK_PAS_1	CATCTAGACATAGTTTAGAA
epikol_00347	CLOCK_PAS_2	GGCAAATACCCTATTATGGG
epikol_00348	CLOCK_PAS_3	AGGCTATGATTACTATCATG
epikol_00349	CLOCK_PAS_4	AATGGCAAATACCCTATTAT
epikol_00350	CLOCK_PAS_5	AAATGGCAAATACCCTATTA
epikol_00351	CREBBP_BROMO_1	CCTAGAAGCACTGTATCGAC
epikol_00352	CREBBP_BROMO_2	AGGGCCTGGCGTAACTCCTC
epikol_00353	CREBBP_BROMO_3	GCCGGAAAGGTAATGACTCT
epikol_00354	CREBBP_BROMO_4	TGCCGGAAAGGTAATGACTC
epikol_00355	CREBBP_BROMO_5	CTGTCGATACAGTGCTTCTA
epikol_00356	CREBBP_KAT_1	CTCCACCGTCTTGTCTGAGC
epikol_00357	CREBBP_KAT_2	CAAAAACCTCCCCGGCTTCA
epikol_00358	CREBBP_KAT_3	CCCCGGCTTCAGGGTGATTC
epikol_00359	CREBBP_KAT_4	GAATCACCTGAAGCCGGGG

epikol_00360	CREBBP_KAT_5	ACAAAAACCTCCCCGGCTTC
epikol_00361	CXXC1_PHD_1	AGCAGTTGATGTCCGGTTTG
epikol_00362	CXXC1_PHD_2	CCGACAGTACCACTCCCGGA
epikol_00363	CXXC1_PHD_3	CATCCGGGAGTGGTACTGTC
epikol_00364	CXXC1_PHD_4	GTACCACTCCCGGATGGCCT
epikol_00365	CXXC1_PHD_5	CCATCCGGGAGTGGTACTGT
epikol_00366	DMAP1_DMTA_1	GCAGCGGCGCACGGAACGCA
epikol_00367	DMAP1_DMTA_2	ACAGGAGCTGCGCAAGATTG
epikol_00368	DMAP1_DMTA_3	CACTGCAGAGCAGCGGCGCA
epikol_00369	DMAP1_DMTA_4	TCTTGCGCAGCTCCTGTAGC
epikol_00370	DMAP1_DMTA_5	CGTGCGCCGCTGCTCTGCAG
epikol_00371	DNMT1_DMTA_1	GCAGCCAGAAAACACATCCA
epikol_00372	DNMT1_DMTA_2	TTCTGGCTGCGGGGGGTTGT
epikol_00373	DNMT1_DMTA_3	GGCTGCGGGGGGTTGTGCGGA
epikol_00374	DNMT1_DMTA_4	AGTAGGTGCGGAATTGAAG
epikol_00375	DNMT1_DMTA_5	CGCAGCCAGAAAACACATCC
epikol_00376	DNMT3A_DMTA_1	AAGAGAGACAGCACCCGGAT
epikol_00377	DNMT3A_DMTA_2	AAAGAGAGACAGCACCCGGA
epikol_00378	DNMT3A_DMTA_3	CGGGTGCTGTCTCTCTTTGA
epikol_00379	DNMT3A_DMTA_4	GGGGAAGATCATGTACGTCG
epikol_00380	DNMT3A_DMTA_5	GGACCGCTACATTGCCTCGG
epikol_00381	DNMT3B_DMTA_1	AACAATGACAGGACTCGAAT
epikol_00382	DNMT3B_DMTA_2	CGATGCCATCAACAATGAC
epikol_00383	DNMT3B_DMTA_3	CGAGTCCTGTCATTGTTTGA
epikol_00384	DNMT3B_DMTA_4	AAACAATGACAGGACTCGAA
epikol_00385	DNMT3B_DMTA_5	TCAAATACGTGAACGACGTG
epikol_00386	DNMT3L_1	CATCTGCTGCGGAAGTCTCC
epikol_00387	DNMT3L_2	AGATCCCTCCCTCAAACAGA
epikol_00388	DNMT3L_3	ACACAGCACCCCTCTGTTTGA
epikol_00389	DNMT3L_4	AGCACCTCTGTTTGAGGGA
epikol_00390	DNMT3L_5	CAGATCCCTCCCTCAAACAG
epikol_00391	DOT1L_MTA_1	AGACATCCCGGCCAAGTATG
epikol_00392	DOT1L_MTA_2	CAGCAACCTGGAGCACGACC
epikol_00393	DOT1L_MTA_3	CGTCGAGAAAGCAGACATCC
epikol_00394	DOT1L_MTA_4	TGTGGGCCAGGTCGTGCTCC
epikol_00395	DOT1L_MTA_5	GCCATAGTGATGTTTGCAGT
epikol_00396	DTX3L_RING_1	CATGGCTTTGTTGATACAAG
epikol_00397	DTX3L_RING_2	GACATGGCTTTGTTGATACA
epikol_00398	DTX3L_RING_3	GTCTGGCATGTGGGACAGAT
epikol_00399	DTX3L_RING_4	ACATGGCTTTGTTGATACAA
epikol_00400	DTX3L_RING_5	CAGAATTCATGCTTGCACTT
epikol_00401	EED_WD40_1	TTTAACTGGCACAGTAAAGA

epikol_00402	EED_WD40_2	CCGGTTGTTGCAATCTTACG
epikol_00403	EED_WD40_3	CCACGTAAGATTGCAACAAC
epikol_00404	EED_WD40_4	ATGGCTCGTATTGCTATCAT
epikol_00405	EED_WD40_5	CAATACGAGCCATCCTCTGC
epikol_00406	EHMT1_SET_1	CTCTACCGGACGCGGGACAT
epikol_00407	EHMT1_SET_2	GCTCTACCGGACGCGGGACA
epikol_00408	EHMT1_SET_3	GCTGCAGCTCTACCGGACGC
epikol_00409	EHMT1_SET_4	GGCTGCAGCTCTACCGGACG
epikol_00410	EHMT1_SET_5	CCCAGCCCCATGTCCCGCGTC
epikol_00411	EHMT2_SET_1	CGCAGATGAAGGTCCCCTGT
epikol_00412	EHMT2_SET_2	CTCTACCGAACAGCCAAGAT
epikol_00413	EHMT2_SET_3	GCTCTACCGAACAGCCAAGA
epikol_00414	EHMT2_SET_4	CCGAACAGCCAAGATGGGCT
epikol_00415	EHMT2_SET_5	TCGCAGATGAAGGTCCCCTG
epikol_00416	ELP3_1	ATCAAGACCTTGCGATACTG
epikol_00417	ELP3_2	GCAATGATATCCACCAGGCG
epikol_00418	ELP3_3	CGGCAGCAATGATATCCACC
epikol_00419	ELP3_4	ATGGGTTTCGCCTTTAACTT
epikol_00420	ELP3_5	TGATATCCACCAGGCGGGGC
epikol_00421	EP300_BROMO_1	TTTGGAGGCACTTTACCGTC
epikol_00422	EP300_BROMO_2	GGGTCCACAGGTTGACGAAA
epikol_00423	EP300_BROMO_3	AGGGTCCACAGGTTGACGAA
epikol_00424	EP300_BROMO_4	CCTTCCCTTTTCGTCAACCTG
epikol_00425	EP300_BROMO_5	CCTAAAAGCTGAGGGTCCAC
epikol_00426	EP300_RING_1	CTGTAATAAGTGGCATCACG
epikol_00427	EP300_RING_2	CCGTAGCAACACAGTGTCTG
epikol_00428	EP300_RING_3	CCACAGACACTGTGTTGCTA
epikol_00429	EP300_RING_4	GTTCTGGTAACTGTAATAAG
epikol_00430	EP300_RING_5	CCAAAGAAACGCTCTCCCTT
epikol_00431	EP400_HELIC_1	CATCGATTCTTACATAGGTG
epikol_00432	EP400_HELIC_2	ATTGAAATCTGAAGGACGTC
epikol_00433	EP400_HELIC_3	AGGAGTGGTGCATAGGATC
epikol_00434	EP400_HELIC_4	AATTGAAATCTGAAGGACGT
epikol_00435	EP400_HELIC_5	TTCTTACATAGGTGAGGTAA
epikol_00436	ERCC6_HELIC_1	TTGGGTACTGGAACGTTCT
epikol_00437	ERCC6_HELIC_2	GTTGAAAATATGGCACAAGC
epikol_00438	ERCC6_HELIC_3	TTGAAAATATGGCACAAGCA
epikol_00439	ERCC6_HELIC_4	TCATTGTATCTCGTAATCAG
epikol_00440	ERCC6_HELIC_5	TTGAGTCTTTGTTGAAAATA
epikol_00441	EZH1_SET_1	CGGATGGGGCACCTTCATAA
epikol_00442	EZH1_SET_2	CCCCCTCTGATGTGGCCGGAT
epikol_00443	EZH1_SET_3	CCCCCTCTGATGTGGCCGGA

epikol_00444	EZH1_SET_4	CTTTATGAAGGTGCCCCATC
epikol_00445	EZH1_SET_5	GCCCCATCCGGCCACATCAG
epikol_00446	EZH2_SET_1	CTGGCACCATCTGACGTGGC
epikol_00447	EZH2_SET_2	CACCATCTGACGTGGCAGGC
epikol_00448	EZH2_SET_3	ACCATCTGACGTGGCAGGCT
epikol_00449	EZH2_SET_4	CCCCAGCCTGCCACGTCAGA
epikol_00450	EZH2_SET_5	CATCTGACGTGGCAGGCTGG
epikol_00461	GTF3C4_1	CCGAGGTGCGGTGGATAACC
epikol_00462	GTF3C4_2	CATCTGCGACGTGCACAACC
epikol_00463	GTF3C4_3	ATCTGCGACGTGCACAACCC
epikol_00464	GTF3C4_4	TGCGGTGGATAACCAGGTCC
epikol_00465	GTF3C4_5	CGACGTGCACAACCCGGGCC
epikol_00466	H2AFY2_1	GTTCAAGTACCGGATCAGCG
epikol_00467	H2AFY2_2	TTCAAGTACCGGATCAGCGT
epikol_00468	H2AFY2_3	CAAGCTGTCCCGTTCAGCTA
epikol_00469	H2AFY2_4	GGCGGCAGTCATTGAGTACC
epikol_00470	H2AFY2_5	CAGGGGCGCCACGCTGATC
epikol_00471	H2AFZ_1	CTTATCAAGGCTACCATAGC
epikol_00472	H2AFZ_2	TTATCAAGGCTACCATAGCT
epikol_00473	H2AFZ_GECKO_1	ACCCATTTACCTCTGCGGTG
epikol_00474	H2AFZ_GECKO_2	CGCGCCGCAGGCTGGCGGTA
epikol_00475	H2AFZ_GECKO_3	TTCATCGACACCTAAAATCT
epikol_00476	HAT1_1	AAGAGTTGATGGGTATACTC
epikol_00477	HAT1_2	GAGTATACCCATCAACTCTT
epikol_00478	HAT1_3	GTATACCCATCAACTCTTTG
epikol_00479	HAT1_4	GTATGAAGTAAGGTTCCGAA
epikol_00480	HAT1_5	CAACACGGAACATTGTTGAC
epikol_00486	HCFC1_1	GGGACTCGCCTCCTGGTGT
epikol_00487	HCFC1_2	CTCGCCTCCTGGTGTTTGGT
epikol_00488	HCFC1_3	ACTCGCCTCCTGGTGTTTGG
epikol_00489	HCFC1_4	TTGGTGGGATGGTGGAGTAT
epikol_00490	HCFC1_5	CCTCCTGGTGTTTGGTGGGA
epikol_00491	HDAC1_HDAC_1	AGTAGTAACAGACTTTCCTC
epikol_00492	HDAC1_HDAC_2	TGAGTCATGCGGATTCGGTG
epikol_00493	HDAC1_HDAC_3	CTATGGTCTCTACCGAAAAA
epikol_00494	HDAC1_HDAC_4	TTCGGTGAGGCTTCATTGGG
epikol_00495	HDAC1_HDAC_5	GGATTTCGGTGAGGCTTCATT
epikol_00496	HDAC10_HDAC_1	AGGACGCTCGATCTCGCACT
epikol_00497	HDAC10_HDAC_2	TCTGCGGTGTGCAGCCCGCG
epikol_00498	HDAC10_HDAC_3	GGACGCTCGATCTCGCACTC
epikol_00499	HDAC10_HDAC_4	GTTGTCAGCCCGCGAGGCCT
epikol_00500	HDAC10_HDAC_5	CGCGCTGCCGAGGCGATCC

epikol_00501	HDAC11_HDAC_1	AGGCCCCCTTCGGACCCAGAC
epikol_00502	HDAC11_HDAC_2	TTCCTCCTGTCTGGGTCCGA
epikol_00503	HDAC11_HDAC_3	ATCCCTTTGATGCCGGA AAAA
epikol_00504	HDAC11_HDAC_4	CCCTTTGATGCCGGA AAAATG
epikol_00505	HDAC11_HDAC_5	TCCCTTTGATGCCGGA AAAAT
epikol_00506	HDAC2_HDAC_1	TGGGTCATGCGGATTCTATG
epikol_00507	HDAC2_HDAC_2	GCGGATTCTATGAGGCTTCA
epikol_00508	HDAC2_HDAC_3	ACAGCAAGTTATGGGTCATG
epikol_00509	HDAC2_HDAC_4	CATAACTTGCTGTAAATTA
epikol_00510	HDAC2_HDAC_5	CATAATTTAACAGCAAGTTA
epikol_00511	HDAC3_HDAC_1	TATTTCTACGACCCCGACGT
epikol_00512	HDAC3_HDAC_2	CTATTTCTACGACCCCGACG
epikol_00513	HDAC3_HDAC_3	CACGTCGGGGTCGTAGAAAT
epikol_00514	HDAC3_HDAC_4	AGTGGAAGTTGCCACGTCG
epikol_00515	HDAC3_HDAC_5	TAGTGGAAGTTGCCACGTC
epikol_00516	HDAC4_HDAC_1	GCATCAGCGTGCATACACG
epikol_00517	HDAC4_HDAC_2	TGGATCCTCCCGCGTGCTC
epikol_00518	HDAC4_HDAC_3	TCCTCCCGCGTGCTCGGGG
epikol_00519	HDAC4_HDAC_4	CTGCAGGAGACGGGCCTCCG
epikol_00520	HDAC4_HDAC_5	TGCTACTCCCGCAGGTGCAC
epikol_00521	HDAC5_HDAC_1	GTGTGTTCCCGCACATGCAC
epikol_00522	HDAC5_HDAC_2	TGGATCCGGCCAGCATGCTC
epikol_00523	HDAC5_HDAC_3	GGATCCGGCCAGCATGCTCA
epikol_00524	HDAC5_HDAC_4	TAAAGCACCAGTGCATGTGC
epikol_00525	HDAC5_HDAC_5	CACGTGCACCCTGAGCATGC
epikol_00526	Ino80_HELIC_3	GCTTTGTCTGCCCTAAGCGG
epikol_00527	Ino80_HELIC_4	TGGACAGGGCCCACCGCTTA
epikol_00528	Ino80_HELIC_5	CTTTGTCTGCCCTAAGCGGT
epikol_00569	Jarid2_JMJC_1	TACATTGACTACTTACACAC
epikol_00570	Jarid2_JMJC_2	GTGTGTAAGTAGTCAATGTA
epikol_00571	Jarid2_JMJC_3	AAGAGGGGATCAAGGTGCAC
epikol_00572	Jarid2_JMJC_4	CCGGGAAGCAGACGACAAAC
epikol_00573	Jarid2_JMJC_5	CTCTTTGCACAGCACCTCCG
epikol_00574	Jarid2_JMJCN_1	CATCGAGTCGGTCCGCGCTC
epikol_00575	Jarid2_JMJCN_2	TCCGCTCATCTACATCGAGT
epikol_00576	Jarid2_JMJCN_3	CGGATCGTGGAACCTCTTGG
epikol_00577	Jarid2_JMJCN_4	ATCGTGGAACCTCTTGCGG
epikol_00578	Jarid2_JMJCN_5	GGGTGATCCCCCTCCGGAC
epikol_00599	Jmjd1c_JMJCSF_1	TGAACAAAAAGCTCCGTC AA
epikol_00600	Jmjd1c_JMJCSF_2	TCACGTATTGGATCATGTTT
epikol_00601	Jmjd1c_JMJCSF_3	CATGTTCTGGTAGAACTTCA
epikol_00602	Jmjd1c_JMJCSF_4	ATCCAATACGTGACCAAAGT

epikol_00603	Jmjd1c_JMJCSF_5	AAGGAAGTGAATAAGAGTAC
epikol_00649	Jmjd4_JMJC_1	TGGACCCACAGTTTGAAC
epikol_00650	Jmjd4_JMJC_2	TCAGAGACTACATCACCTAC
epikol_00651	Jmjd4_JMJC_3	AGGCGGGCTACTCCTCTCCC
epikol_00652	Jmjd4_JMJC_4	GCTGTCTCTACCTCAAAGAC
epikol_00653	Jmjd4_JMJC_5	AGTTGCAAAGTGTGGGGTCC
epikol_00659	Jmjd6_JMJC_1	ATCAAAGTGACCCGAGACGA
epikol_00660	Jmjd6_JMJC_2	GGGCCACCACGCTCCGGAAC
epikol_00661	Jmjd6_JMJC_3	TGGGATTACATCGACCCTC
epikol_00662	Jmjd6_JMJC_4	TAGTTCAGGGCCACAAGCGC
epikol_00663	Jmjd6_JMJC_5	AAAGTGACCCGAGACGAAGG
epikol_00664	Jmjd7_CUPIN_1	GCGGATAATGCACGGCCTGT
epikol_00665	Jmjd7_CUPIN_2	CGGATAATGCACGGCCTGTT
epikol_00666	Jmjd7_CUPIN_3	AGAGCGTTGCGGATAATGCA
epikol_00667	Jmjd7_CUPIN_4	CGCTCCACTTCTACCGGGAC
epikol_00668	Jmjd7_CUPIN_5	GGATAATGCACGGCCTGTTG
epikol_00669	Jmjd8_JMJC_1	CAGGATGACGGGCCTGACGA
epikol_00670	Jmjd8_JMJC_2	CGTCAGGCCCGTCATCCTGC
epikol_00671	Jmjd8_JMJC_3	GCAACCTGTGCGGGGAGCAC
epikol_00672	Jmjd8_JMJC_4	CGAAGCCAGCAACCTGTGCG
epikol_00673	Jmjd8_JMJC_5	GGGCCCTGTGCTCCCGCGAC
epikol_00674	Kat2a_BROMO_1	GATGACCTCGTAGTAGTCAG
epikol_00675	Kat2a_BROMO_2	GCGGCAGTACTCGCTGTCCG
epikol_00676	Kat2a_BROMO_3	CTGTGCGGAGTACAACCCCC
epikol_00677	Kat2a_BROMO_4	GAAGCCGCTACTACGTGACC
epikol_00678	Kat2a_BROMO_5	CGGCAGTACTCGCTGTCCGG
epikol_00451	KAT2A_HELIC_1	AACAGAGTATGGAGGTTATC
epikol_00452	KAT2A_HELIC_2	TGGGGAGGTAACAAAGATAT
epikol_00453	KAT2A_HELIC_3	ACAGAGTATGGAGGTTATCA
epikol_00454	KAT2A_HELIC_4	CAAATTCAAACAGAGTATGG
epikol_00455	KAT2A_HELIC_5	TATGGAGGTTATCAGGGCTT
epikol_00456	KAT2A_NAT_1	CCAGACAGACCACCAAAGTC
epikol_00679	Kat2a_NAT_1	GACAATCTCCGTGAAGCCCT
epikol_00457	KAT2A_NAT_2	GCTGTTACCCAGATTATCA
epikol_00680	Kat2a_NAT_2	GGCCTTGATCAAGGATGGGC
epikol_00458	KAT2A_NAT_3	CCAGACTTTGGTGGTCTGTC
epikol_00681	Kat2a_NAT_3	GCTTGATGTGATACTCCTTC
epikol_00459	KAT2A_NAT_4	CTGTTACCCAGATTATCAA
epikol_00682	Kat2a_NAT_4	TGATGTGATACTCCTTCAGG
epikol_00460	KAT2A_NAT_5	GACTTTGGTGGTCTGTCTGG
epikol_00683	Kat2a_NAT_5	CCTTCAGGTGGTTCATCAGG
epikol_00684	Kat2b_BROMO_1	TGAAGGGCCAAGCGCTTTGA

epikol_00999	KAT2B_BROMO_1	TGAAGGGCCAAGCGCTTTGA
epikol_00685	Kat2b_BROMO_2	CAGGATATTATGAAGTTATA
epikol_01000	KAT2B_BROMO_2	CAGGATATTATGAAGTTATA
epikol_00686	Kat2b_BROMO_3	ATTCTTGAGGCGTTCACCTCA
epikol_01001	KAT2B_BROMO_3	ATTCTTGAGGCGTTCACCTCA
epikol_00687	Kat2b_BROMO_4	ACACGTAGTACCTATTCTTG
epikol_01002	KAT2B_BROMO_4	ACACGTAGTACCTATTCTTG
epikol_00688	Kat2b_BROMO_5	GGGGTTGTACTCTTTGCAAT
epikol_01003	KAT2B_BROMO_5	GGGGTTGTACTCTTTGCAAT
epikol_00689	Kat2b_NAT_1	ATTAAAGATGGCCGTGTTAT
epikol_01004	KAT2B_NAT_1	ATTAAAGATGGCCGTGTTAT
epikol_00690	Kat2b_NAT_2	AAAGATGGCCGTGTTATTGG
epikol_01005	KAT2B_NAT_2	TCTCTGTGAATCCTTGAGAT
epikol_00691	Kat2b_NAT_3	TCTCTGTGAATCCTTGAGAT
epikol_01006	KAT2B_NAT_3	TTCCGTATGTTCCCATCTCA
epikol_00692	Kat2b_NAT_4	TTCCGTATGTTCCCATCTCA
epikol_01007	KAT2B_NAT_4	AACAGATACCACCAATAACA
epikol_00693	Kat2b_NAT_5	AACAGATACCACCAATAACA
epikol_01008	KAT2B_NAT_5	ATCTCTGTGAATCCTTGAGA
epikol_00694	Kat5_CHROM_1	ATGGGTGACGCATGAGCGGC
epikol_01524	KAT5_CHROM_1	GGCCGAGATCCTGAGCGTGA
epikol_00695	Kat5_CHROM_2	ATGAATGGGTGACGCATGAG
epikol_01525	KAT5_CHROM_2	GCGTGAAGGACATCACTGGC
epikol_00696	Kat5_CHROM_3	GTCCTTCACGCTCAGGATCT
epikol_01526	KAT5_CHROM_3	ATGGGTGACGCATGAGCGGC
epikol_00697	Kat5_CHROM_4	GGCCGAGATCCTGAGCGTGA
epikol_01527	KAT5_CHROM_4	ATGAATGGGTGACGCATGAG
epikol_00698	Kat5_CHROM_5	GCGTGAAGGACATCACTGGC
epikol_01528	KAT5_CHROM_5	CACTGATGTCCTTCACGCTC
epikol_00699	Kat5_NAT_1	ATCAACGGAAGACTACAATG
epikol_01529	KAT5_NAT_1	GCTTGCCGTAGCCCCGGCGC
epikol_00700	Kat5_NAT_2	GCTTGCCGTAGCCCCGGCGC
epikol_01530	KAT5_NAT_2	CTGCCTCCCTACCAGCGCCG
epikol_00701	Kat5_NAT_3	CTGCCTCCCTACCAGCGCCG
epikol_01531	KAT5_NAT_3	TCAGCAGCTTGCCGTAGCCC
epikol_00702	Kat5_NAT_4	TCAGCAGCTTGCCGTAGCCC
epikol_01532	KAT5_NAT_4	ATCAACGGAAGACTACAATG
epikol_00703	Kat5_NAT_5	GCCGTAGCCCCGGCGCTGGT
epikol_01533	KAT5_NAT_5	CCGGCGCTGGTAGGGAGGCA
epikol_00864	KAT6A_HAT_1	TGGCTCCACATCGTAATAGA
epikol_00865	KAT6A_HAT_2	ATGGCTCCACATCGTAATAG
epikol_00866	KAT6A_HAT_3	TCTTCCTCAATACCAGCGTA

epikol_00867	KAT6A_HAT_4	CTAACACAGAATGATGTCAA
epikol_00868	KAT6A_HAT_5	TAGCCCTTACGCTGGTATTG
epikol_00869	KAT6A_PHD_1	ACTGTTGCCACAGTCGGCAC
epikol_00870	KAT6A_PHD_2	CAGAAACTACAGATGGGGAT
epikol_00871	KAT6A_PHD_3	TTCACCTCGAACCGTTAGTTC
epikol_00872	KAT6A_PHD_4	TCACTCGAACCGTTAGTTCA
epikol_00873	KAT6A_PHD_5	CACTCGAACCGTTAGTTCAG
epikol_00874	KAT6B_HAT_1	CAGCAAATGAAATTTACCGA
epikol_00875	KAT6B_HAT_2	CTTCGGTAAATTTTCATTGTC
epikol_00876	KAT6B_HAT_3	TGATGAAAAGGGCTGTCATC
epikol_00877	KAT6B_HAT_4	ACCGTCCAAATCCTTGCCTT
epikol_00878	KAT6B_HAT_5	ACCAAAGGCAAGGATTTGGA
epikol_00879	KAT6B_PHD_1	CCCAAACAGAAGCTACATAT
epikol_00880	KAT6B_PHD_2	ACAGAAGCTACATATTGGAA
epikol_00881	KAT6B_PHD_3	GCACAAGAGAGGAGTCTTTC
epikol_00882	KAT6B_PHD_4	TGCCACAATCTGCACAAGAG
epikol_00883	KAT6B_PHD_5	TCCAATATGTAGCTTCTGTT
epikol_00481	KAT7_1	ACCAGGTATCAAGCTCATAG
epikol_00482	KAT7_2	AGGATATGGAGAATGATACC
epikol_00483	KAT7_3	CGATATATCTCATCACCAGG
epikol_00484	KAT7_4	CTGCCAAAACCTGTGCCTGT
epikol_00485	KAT7_5	CTTGAAAATTATGCAGGTAG
epikol_00859	KAT7_GECKO_1	GGGTGACTCGAGCAGATCGT
epikol_00860	KAT7_GECKO_2	GACAACTCACCATGTGCCGG
epikol_00861	KAT7_GECKO_3	TTCAATCTCTGTGTTTGAAG
epikol_00862	KAT7_GECKO_4	CTCGAGATCTGTAGAAAAAT
epikol_00863	KAT7_GECKO_5	TCTTGCAGGCCAAATGTGTG
epikol_00849	KAT8_CHROM_1	GGAATTCTATGTACACTACG
epikol_00850	KAT8_CHROM_2	CCTGGTCGTTCACTCGAGAC
epikol_00851	KAT8_CHROM_3	GAATTCTATGTACACTACGT
epikol_00852	KAT8_CHROM_4	GTGGGTAGACAAGAACCGGC
epikol_00853	KAT8_CHROM_5	TACAGCATCCTTCACTGTCT
epikol_00854	KAT8_HAT_1	CAATTTCTGTAGTTCCCGATG
epikol_00855	KAT8_HAT_2	CACCATTCCCCGAAGACTAT
epikol_00856	KAT8_HAT_3	TCACCATTCCCCGAAGACTA
epikol_00857	KAT8_HAT_4	GGGGAATGGTGAGAAATACC
epikol_00858	KAT8_HAT_5	TTCCCATAGTCTTCGGGGAA
epikol_00709	KDM1B_AMOX_1	CAACGGTGGATTAAACTGAA
epikol_00710	KDM1B_AMOX_2	TCAATTATCACCGAGTACCC
epikol_00711	KDM1B_AMOX_3	TTCAATTATCACCGAGTACC
epikol_00712	KDM1B_AMOX_4	TCTGCTAACTCCCGGGTACT
epikol_00713	KDM1B_AMOX_5	GAGTGTGGTCACCAGCAAAC

epikol_00584	KDM3A_JMJC_1	GCCATCTCGCCTTGTGTACT
epikol_00585	KDM3A_JMJC_2	CGCCTTGTGTACTCGGGCAG
epikol_00586	KDM3A_JMJC_3	CCCGAGTACACAAGGCGAGA
epikol_00587	KDM3A_JMJC_4	GTTCCATATTTCCGATCTTC
epikol_00588	KDM3A_JMJC_5	GCTAATGTCATGGTCTATGT
epikol_00589	KDM3B_JMJC_1	ATTGAGCCTGCCATCTCGTT
epikol_00590	KDM3B_JMJC_2	GGCTACCTAGCTACTTTGTA
epikol_00591	KDM3B_JMJC_3	CTACTTTGTAAGGCCTGATC
epikol_00592	KDM3B_JMJC_4	AATATACCAAACGAGATGGC
epikol_00593	KDM3B_JMJC_5	CCATCTCGTTTGGTATATTC
epikol_00594	KDM3B_JMJCSF_1	ACGAAGCAGAGGATTTCATGA
epikol_00595	KDM3B_JMJCSF_2	CCTCATAGAGTCGCTTACGG
epikol_00596	KDM3B_JMJCSF_3	CTCATAGAGTCGCTTACGGA
epikol_00597	KDM3B_JMJCSF_4	ACTCCTCATAGAGTCGCTTA
epikol_00598	KDM3B_JMJCSF_5	CCTCCGTAAGCGACTCTATG
epikol_00604	KDM4A_JMJC_1	TACTCTGTTCCACCTGAGCA
epikol_00605	KDM4A_JMJC_2	CAGCATTAACGGGGAAATCA
epikol_00606	KDM4A_JMJC_3	ATATTTCTTCAGCATTAACG
epikol_00607	KDM4A_JMJC_4	AGCCGGCATGGTAACCATAA
epikol_00608	KDM4A_JMJC_5	ACTCCGCACAGTTAAACCA
epikol_00609	KDM4A_JMJCN_1	TCCCAAGGAGCTCATCGGGC
epikol_00610	KDM4A_JMJCN_2	TCATCGGGCAGGGCTAGCCA
epikol_00611	KDM4A_JMJCN_3	TTGAATCCCAAGGAGCTCAT
epikol_00612	KDM4A_JMJCN_4	GCCCTGCCCCGATGAGCTCCT
epikol_00613	KDM4A_JMJCN_5	TTTCGGAATCTTCCATAGT
epikol_00614	KDM4B_JMJC_1	ACCAGAGCACGGCAAGCGCC
epikol_00615	KDM4B_JMJC_2	TACGCCATCCCACCAGAGCA
epikol_00616	KDM4B_JMJC_3	AGGGGATCCCGTACTTCTTC
epikol_00617	KDM4B_JMJC_4	CGCAGCCCTGCGAGCTCCCG
epikol_00618	KDM4B_JMJC_5	ACCACGCCGGCTTCAATCAC
epikol_00619	KDM4B_JMJCN_1	CGTGGCCTACATAGAGTCGC
epikol_00620	KDM4B_JMJCN_2	GTGGCCTACATAGAGTCGCA
epikol_00621	KDM4B_JMJCN_3	GGCTCCCTGCGACTCTATGT
epikol_00622	KDM4B_JMJCN_4	CTTGGCCAGGCCCGCCCGGT
epikol_00623	KDM4B_JMJCN_5	GGGAGCCCACCGGGCGGGCC
epikol_00624	KDM4C_JMJC_1	GGAGATAATTAATGCTATAG
epikol_00625	KDM4C_JMJC_2	AGCATTAATTATCTCCACTT
epikol_00626	KDM4C_JMJC_3	TATGCTATACCTCCGGAGCA
epikol_00627	KDM4C_JMJC_4	AACCAGCATGGTAGCCATAT
epikol_00628	KDM4C_JMJC_5	TCATGAATTCTCCAGCCTCC
epikol_00629	KDM4C_JMJCN_1	CCTTTGCAAGACCCGCACGA
epikol_00630	KDM4C_JMJCN_2	CCATCGTGGGGTCTTGCAA

epikol_00631	KDM4C_JMJCN_3	TCTAAAGGAGCCCATCGTGC
epikol_00632	KDM4C_JMJCN_4	CTTTAGACTCCATGTATGCA
epikol_00633	KDM4C_JMJCN_5	CAAGGTATTTGTTGAACTCC
epikol_00634	KDM4D_GECKO_1	AACCTGAACGCTATGACCTG
epikol_00635	KDM4D_GECKO_2	GCTCCAAATCTTCGAAATTC
epikol_00636	KDM4D_GECKO_3	CAGATTATCCACCCGTCAAA
epikol_00637	KDM4D_JMJC_1	TGGCCAGGCGTTCCAGGCGC
epikol_00638	KDM4D_JMJC_2	CCATGGTTTCAACTGCGCAG
epikol_00639	KDM4D_JMJC_3	CCACAACCCCGGGAAGTCC
epikol_00640	KDM4D_JMJC_4	CCTCTGCGCAGTTGAAACCA
epikol_00641	KDM4D_JMJCN_1	AGCCAGCTCTGTGTGCACCT
epikol_00642	KDM4D_JMJCN_2	TCATTAAACTCTTCTTTGGT
epikol_00643	KDM4D_JMJCN_3	AGGTGCACACAGAGCTGGCT
epikol_00529	KDM5A_GECKO_5	ACGACTGACCCAAGTGTCTAT
epikol_00530	KDM5A_JMJC_1	CCATGCTCCATTAGCACGTT
epikol_00531	KDM5A_JMJC_2	CATGCTCCATTAGCACGTTG
epikol_00532	KDM5A_JMJC_3	TAGCACGTTGGGGTTCATGA
epikol_00533	KDM5A_JMJC_4	GGCTGGGATTCAAATAACTC
epikol_00534	KDM5A_JMJC_5	TTTCCTCGTGCCTATCACTC
epikol_00535	KDM5A_PHD_1	CCACTACCTGATGTGCCCAA
epikol_00536	KDM5A_PHD_2	CCTTTGGGCACATCAGGTAG
epikol_00537	KDM5A_PHD_3	TTGGGCACATCAGGTAGTGG
epikol_00538	KDM5A_PHD_4	CTGATGTGCCCAAAGGAGAC
epikol_00539	KDM5B_JMJC_1	ATACCCTGGGACTCCATACC
epikol_00540	KDM5B_JMJC_2	CCCCAGCACACTGATTAGTT
epikol_00541	KDM5B_JMJC_3	CCTCAGCAAAATTAACACCC
epikol_00542	KDM5B_JMJC_4	TACCGAACTAATCAGTGTGC
epikol_00543	KDM5B_JMJC_5	TTTCCAAGAGCCTACCACAG
epikol_00544	KDM5B_PHD_1	GACCGGCTACTGTTGTGTGA
epikol_00545	KDM5B_PHD_2	AGCCATCACACAACAGTAGC
epikol_00546	KDM5B_PHD_3	CCTCTCCATGATGTTCCCAA
epikol_00547	KDM5B_PHD_4	CCCTTGGAACATCATGGAG
epikol_00548	KDM5B_PHD_5	AGTCTCCCTTGGAACATCA
epikol_00549	KDM5C_JMJC_1	ACTCCATTAACCTCCAC
epikol_00550	KDM5C_JMJC_2	GTGAGCCGAAGACCTGGTAT
epikol_00551	KDM5C_JMJC_3	AAGTGAGGGCACCCCATACC
epikol_00552	KDM5C_JMJC_4	CCATGGGACATGAGGGTGTT
epikol_00553	KDM5C_JMJC_5	AGCCGCTGTGGTAAGCACGG
epikol_00554	KDM5C_PHD_1	GGCACCGCCAGACACCTTG
epikol_00555	KDM5C_PHD_2	GATGTGTTCTCGAGGGGATG
epikol_00556	KDM5C_PHD_3	TCTGCCGGATGTGTTCTCGA
epikol_00557	KDM5C_PHD_4	CATCCCCTCGAGAACACATC

epikol_00558	KDM5C_PHD_5	CTGAGATCCCCAAGGGTGTC
epikol_00559	KDM5D_JMJC_1	ACTCTATTAAGTATCTGCAT
epikol_00560	KDM5D_JMJC_2	CTGGTATGGTGTACCCTCCC
epikol_00561	KDM5D_JMJC_3	CCCAACACTTTGATGTCCCA
epikol_00562	KDM5D_JMJC_4	TCCGCACAAACCAGTGTGCA
epikol_00563	KDM5D_JMJC_5	GGCTGGCTATCAAACAGCTC
epikol_00564	KDM5D_PHD_1	CATCCCCACGGGAGCATACT
epikol_00565	KDM5D_PHD_2	TTGCCAAGTATGCTCCCGTG
epikol_00566	KDM5D_PHD_3	ATTTGCCAAGTATGCTCCCG
epikol_00567	KDM5D_PHD_4	TTTGCCAAGTATGCTCCCGT
epikol_00568	KDM5D_PHD_5	CTGAAATCCCCAGAGGCATC
epikol_01609	KDM6A_JMJC_1	CAAACCATTACAGTCACCT
epikol_01610	KDM6A_JMJC_2	GAATGGTTTGTGTTCTCTGA
epikol_01611	KDM6A_JMJC_3	GTTGTTCTCTGAAGGTTACTG
epikol_01612	KDM6A_JMJC_4	GTTCAATTGGGTTTACAGCTAT
epikol_01613	KDM6A_JMJC_5	TGGGCCACCAAGAACCCATT
epikol_00644	KDM6B_JMJC_1	GCGAACCACCTCGCAGTCGCC
epikol_00645	KDM6B_JMJC_2	TCGCGGTGCACGAGCACTAC
epikol_00646	KDM6B_JMJC_3	CGCGGTGCACGAGCACTACT
epikol_00647	KDM6B_JMJC_4	TGCTCCGTCAACATCAACAT
epikol_00648	KDM6B_JMJC_5	ACCTCGTGTGGATTAATGCG
epikol_00579	KDM7A_JMJC_1	ACTTCGGTGGAAGTTTCACTC
epikol_00580	KDM7A_JMJC_2	ATTTGGCACGTTATGAATCT
epikol_00581	KDM7A_JMJC_3	CTCCAAAGAACACCTCACTC
epikol_00582	KDM7A_JMJC_4	CCATGCTGTGCTCACTTCTC
epikol_00583	KDM7A_JMJC_5	CCTGAGAAGTGAGCACAGCA
epikol_00654	KDM8_CUPIN_1	AGTGGTCAGCCACGCCTTTC
epikol_00655	KDM8_CUPIN_2	GCCACGCCTTTCAGGATCAC
epikol_00656	KDM8_CUPIN_3	ACTGGCCGTGCATGCAGAAG
epikol_00657	KDM8_CUPIN_4	GAACTCGTTGACCGTCATGA
epikol_00658	KDM8_CUPIN_5	CGTTGACCGTCATGAGGGTC
epikol_00749	KMT2A_BROMO_1	CCGGTAGCGTAGCAAATGGC
epikol_00750	KMT2A_BROMO_2	TCTGTCTCGGGATTTAAGTC
epikol_00751	KMT2A_BROMO_3	GTCTCGGGATTTAAGTCTGG
epikol_00752	KMT2A_BROMO_4	AGGTCCTTCGGGGGAGCTGC
epikol_00753	KMT2A_BROMO_5	CAGCAGCCTTTAGATCTAGA
epikol_00754	KMT2A_PHD_1	GGAAGGGCTCACAACAGACT
epikol_00755	KMT2A_PHD_2	TGGAGGACCAGCTGAAAAAT
epikol_00756	KMT2A_PHD_3	AGAGGAGAACGAGCGCCCTC
epikol_00757	KMT2A_PHD_4	TCTCCTCTAAACAAAACCTTG
epikol_00758	KMT2A_PHD_5	CTCTAAACAAAACCTTGTTGA
epikol_00779	KMT2B_PHD_1	TGCCCCAGCATCACGACACC

epikol_00780	KMT2B_PHD_2	AGGTGTCGTGATGCTGGGGC
epikol_00781	KMT2B_PHD_3	AGCACCAGGTGTCGTGATGC
epikol_00782	KMT2B_PHD_4	GCACCAGGTGTCGTGATGCT
epikol_00783	KMT2B_PHD_5	CACCAGGTGTCGTGATGCTG
epikol_00784	KMT2B_SET_1	AGCGGGAGAAGTTCTACGAT
epikol_00785	KMT2B_SET_2	CGATGTTGCGCTTACAGAAC
epikol_00786	KMT2B_SET_3	AGATCAGCCATCCACGGGCG
epikol_00787	KMT2B_SET_4	CAACATCGACGCGGGGAGA
epikol_00788	KMT2B_SET_5	AAGCGGGAGAAGTTCTACGA
epikol_00769	KMT2C_GECKO_1	GACACAGATCGCTGAAGAGT
epikol_00770	KMT2C_GECKO_2	AAACCAACCTTCTAACAAGA
epikol_00771	KMT2C_GECKO_3	TCTTCAGCGATCTGTGTCTG
epikol_00772	KMT2C_GECKO_4	CTCCATGGCAAGATAAATCC
epikol_00773	KMT2C_GECKO_5	CTTGATCATCTGAAGCTGTC
epikol_00774	KMT2C_SET_1	CAATGTCTCGAGCAGCATAC
epikol_00775	KMT2C_SET_2	GTGATTGACGCGACGCTCAC
epikol_00776	KMT2C_SET_3	TGAGCGTCGCGTCAATCACA
epikol_00777	KMT2C_SET_4	TTGACGCGACGCTCACAGGA
epikol_00778	KMT2C_SET_5	ATTGACGCGACGCTCACAGG
epikol_00759	KMT2D_PHD_1	GCACTGAGAGCACTGCGAAC
epikol_00760	KMT2D_PHD_2	GAGAGCACTGCGAACAGGCA
epikol_00761	KMT2D_PHD_3	TGGTATGTGGCAGCTTTGGC
epikol_00762	KMT2D_PHD_4	ACTGCGAACAGGCAAGGAGG
epikol_00763	KMT2D_PHD_5	ACCAAGGTGATGCTGCTCAA
epikol_00764	KMT2D_SET_1	ATTGATGCTACGTTGACCGG
epikol_00765	KMT2D_SET_2	CACCATCATTCGGAACGAGG
epikol_00766	KMT2D_SET_3	CTGGCTCGCTCCCGTATCCA
epikol_00767	KMT2D_SET_4	CCTTGGCTGCATAGAGCCCC
epikol_00768	KMT2D_SET_5	GTGATTGATGCTACGTTGAC
epikol_00789	KMT2E_PHD_1	CGTTCACATAGATATGTATC
epikol_00790	KMT2E_PHD_2	TTGACTGCATGGGGATTGAT
epikol_00791	KMT2E_PHD_3	GGCAACATATTGACTGCATG
epikol_00792	KMT2E_PHD_4	TGGCAACATATTGACTGCAT
epikol_00793	KMT2E_PHD_5	TTGGCAACATATTGACTGCA
epikol_00794	KMT2E_SET_1	TATTCAATGATAAGTGCATC
epikol_00795	KMT2E_SET_2	TCAATGATAAGTGCATCAGG
epikol_00796	KMT2E_SET_3	GGAATGAGGCTCGATTTCATC
epikol_00797	KMT2E_SET_4	ATGAGGCTCGATTTCATCAGG
epikol_00798	KMT2E_SET_5	GGAATGAGGCTCGATTTCATC
epikol_00704	L3mbtl1_MBT_1	AGTACATGGACGGGTGTTGA
epikol_00705	L3mbtl1_MBT_2	AAGTACATGGACGGGTGTTG
epikol_00706	L3mbtl1_MBT_3	GCAGTCACTCACAACAAGAA

epikol_00707	L3mbtl1_MBT_4	CATGTACTTCATCCTCACCG
epikol_00708	L3mbtl1_MBT_5	TCTTGTGTGAGTGACTGCC
epikol_00714	Mbd1_MEC_1	CGCGAAGTCTTTCGCAAGTC
epikol_00715	Mbd1_MEC_2	ATAGGTGTCTGAGCGTCCAC
epikol_00716	Mbd1_MEC_3	CTTCGCGGCGCTTCCAGCCA
epikol_00717	Mbd1_MEC_4	GTTGAGCTGACTCGATACCT
epikol_00718	Mbd1_MEC_5	GCCTTGTTGAAGTCGAAGA
epikol_00719	Mbd2_MEC_1	CCTTCTTCCATCCGGGGGGG
epikol_00720	Mbd2_MEC_2	GAGGAAGTGATCCGAAAATC
epikol_00721	Mbd2_MEC_3	CCTCCCCCGGATGGAAGA
epikol_00722	Mbd2_MEC_4	AGGAAGTGATCCGAAAATCT
epikol_00723	Mbd2_MEC_5	CCCCCGGATGGAAGAAGG
epikol_00724	Mbd3_MEC_1	GCCGACAGCCCCGACCTTCT
epikol_00725	Mbd3_MEC_2	CGGGGCTGTCGGCCGGCCAC
epikol_00726	Mbd3_MEC_3	GGCCGACAGCCCCGACCTTC
epikol_00727	Mbd3_MEC_4	GGAGTGCCCGGCGCTCCCGC
epikol_00728	Mbd3_MEC_5	AGAAGGTGCGGGCTGTCGGC
epikol_00729	Mbd4_MEC_1	GATCTGAAGTTCAGTCTTG
epikol_00730	Mbd4_MEC_2	GTTGTGAAGCAAAGGTTATT
epikol_00731	Mbd4_MEC_3	GCTAATTATCTTCACAAAAA
epikol_00732	Mbd4_MEC_4	TTTACTGTACTTTCTAAAAG
epikol_00733	Mbd4_MEC_5	ACAGTAAAATCAAAATCTTC
epikol_00734	Mecp2_MEC_1	GGACACGGAAGCTTAAGCAA
epikol_00735	Mecp2_MEC_2	TCCGTGTCCAGCCTTCAGGC
epikol_00736	Mecp2_MEC_3	GACTTCACGGTAAGTGGGAG
epikol_00737	Mecp2_MEC_4	ATTGCGTACTTCGAAAAGGT
epikol_00738	Mecp2_MEC_5	AAAAGCCTTTCGCTCTAAAG
epikol_00739	MgeA5_NAG_1	GCGAGAGATGTATTCAGTGG
epikol_00740	MgeA5_NAG_2	GGTGTGGATAGCAACGTAGT
epikol_00741	MgeA5_NAG_3	ACCTCTTCGATGGACTCTAC
epikol_00742	MgeA5_NAG_4	CTATGCGATCTCACCTGGAT
epikol_00743	MgeA5_NAG_5	TATGTGTGCAGCAGACAAAG
epikol_00744	Mina_CUPIN_1	AATGTCTGCCGGTGTGTCAA
epikol_00745	Mina_CUPIN_2	CTCTTCCATAGTACATCCCC
epikol_00746	Mina_CUPIN_3	CTGCACTGGCCACATACTAT
epikol_00747	Mina_CUPIN_4	AATGTGTACATAACTCCCGC
epikol_00748	Mina_CUPIN_5	CATCATCATAATGGGGCGGC
epikol_00799	MMP8_GECKO_1	ATTCCATTGGGTCCATCAAA
epikol_00800	MMP8_GECKO_2	TGCAACACTCCAGAGTTCAA
epikol_00801	MMP8_GECKO_3	CTATACCCACAGCTGTCAG
epikol_00802	MMP8_GECKO_4	AACGCACTAACTTGACCTAC
epikol_00803	MMP8_GECKO_5	GCAACCAGTATCAGTCTACA

epikol_00804	MORF4L1_GECKO_1	GCCTCTTCTTTATGAAGCAA
epikol_00805	MORF4L1_GECKO_2	ACATATTACCAATAATCTC
epikol_00806	MORF4L1_GECKO_3	ACCTTTGCTTCATAAAGAAG
epikol_00807	MORF4L1_GECKO_4	CCACATCTCCTGAGATTATT
epikol_00808	MORF4L1_GECKO_5	CATGCATTACCTAGAGGTC
epikol_00809	MORF4L1_GECKO_6	TGTACTGAGGGAGCTCCTTC
epikol_00810	MORF4L1_TUDOR_1	CTCCTTCAGGAACAGGAAAA
epikol_00811	MORF4L1_TUDOR_2	GAGGGAGCTCCTTCAGGAAC
epikol_00812	MORF4L1_TUDOR_3	AGGAGGGGGTGGTGTACTGA
epikol_00813	MORF4L1_TUDOR_4	CAGGAGGGGGTGGTGTACTG
epikol_00814	MSL3_GECKO_1	GTTCAATTAGCTGGGATAGA
epikol_00815	MSL3_GECKO_2	TCGTAGATTACAGCGTAAAT
epikol_00816	MSL3_GECKO_3	GTGCCTGACAATTACCCCC
epikol_00817	MSL3_GECKO_4	TGAGCAAATGAGCGCGAGCG
epikol_00818	MSL3_GECKO_5	ATCGTACAGCACTCGCGCCT
epikol_00819	MTA1_BAH_1	ATACCTGATCCGGAGAATCG
epikol_00820	MTA1_BAH_2	CGATTCTCCGGATCAGGTAT
epikol_00821	MTA1_BAH_3	TCGATTCTCCGGATCAGGTA
epikol_00822	MTA1_BAH_4	CCAAAGTGGTGTGCTTCTAC
epikol_00823	MTA1_BAH_5	CTATAAGGCCGGACCGGGG
epikol_00824	MTA1_SANT_1	GAGGAAGCCCTGAAAAATA
epikol_00825	MTA1_SANT_2	CAGGGCTTCCTCGAAAAGGT
epikol_00826	MTA1_SANT_3	TCATTGAGTACTACTACATG
epikol_00827	MTA1_SANT_4	GTAGTAGTACTCAATGATGC
epikol_00828	MTA1_SANT_5	ATGCTGGTCAGCGACTTCCA
epikol_00829	MTA2_BAH_1	TTACCTGGTTAGACGGATTG
epikol_00830	MTA2_BAH_2	TCTAACCAGGTAAGGATTGC
epikol_00831	MTA2_BAH_3	GCAATCCTTACCTGGTTAGA
epikol_00832	MTA2_BAH_4	GTTCCCGGTGCTTCAGTTGA
epikol_00833	MTA2_BAH_5	AGATATCTTGAGCCAGTACC
epikol_00834	MTA2_SANT_1	GGACTTCAATGATATTCCGC
epikol_00835	MTA2_SANT_2	TAGGGCCTCCTCAAAATAGCA
epikol_00836	MTA2_SANT_3	CTCAGAGGCCATGCTATTTG
epikol_00837	MTA2_SANT_4	GGCCCTAGAGAAGTATGGGA
epikol_00838	MTA2_SANT_5	AGAGGCCATGCTATTTGAGG
epikol_00839	MTA3_BAH_1	TCTTATTAGGTATGGGTTGC
epikol_00840	MTA3_BAH_2	GCAACCCATACCTAATAAGA
epikol_00841	MTA3_BAH_3	TCTATCCTTCTTATTAGGTA
epikol_00842	MTA3_BAH_4	AGTATTGTCATATCTTGATA
epikol_00843	MTA3_BAH_5	TGATTCTGTCTCATTCAGAA
epikol_00844	MTA3_GECKO_1	GCAACCCATACCTAATAAGA
epikol_00845	MTA3_GECKO_2	AAAACACTATTAGCTGACAA

epikol_00846	MTA3_GECKO_3	GTTCTTCTATCCTTCTTATT
epikol_00847	MTA3_GECKO_4	TCAATCTGTCGATCCGTAAG
epikol_00848	MTA3_GECKO_5	ACCGTGCTACAACATAAAAC
epikol_00884	NAP1L1_GECKO_1	CATTACATTCTTCTTCCGT
epikol_00885	NAP1L1_GECKO_2	TACCTTTCAATGTATCCTGT
epikol_00886	NAP1L1_GECKO_3	TAATGCAATTTATGAACCTA
epikol_00887	NAP1L1_GECKO_4	CTATCAGCCTCTATTTGATA
epikol_00888	NAP1L1_GECKO_5	CACCAACAGGATACATTGAA
epikol_00889	NAP1L2_NAP_1	CGCAGTTCCCTATAGGGAT
epikol_00890	NAP1L2_NAP_2	CTCAATCGCAGTTCCTTAT
epikol_00891	NAP1L2_NAP_3	TATGATCCCCATCCCTATAG
epikol_00892	NAP1L2_NAP_4	ACTGCGATTGAGTATTCCAC
epikol_00893	NAP1L2_NAP_5	TTATGATCCCCATCCCTATA
epikol_00894	NAP1L3_NAP_1	CAGTAGCAGTTGTGCGACTC
epikol_00895	NAP1L3_NAP_2	CAGAGTCGCACAACCTGCTAC
epikol_00896	NAP1L3_NAP_3	TGAAACCGCCTATCATACAG
epikol_00897	NAP1L3_NAP_4	AGATCCTAAAGAAGTCCCCC
epikol_00898	NAP1L3_NAP_5	GATTCTATGATTGGGAAGC
epikol_00899	NAT10_HELIC_1	ACAGAGTATGGAGGTTATCA
epikol_00900	NAT10_HELIC_2	GGGCGTATCGGATTGACTCC
epikol_00901	NAT10_HELIC_3	ATGCCATGAAAACAAGGTAG
epikol_00902	NAT10_HELIC_4	CGTGGTCTTATTCTCAGCAG
epikol_00903	NAT10_HELIC_5	TTAGCTTGAGGGACAGTGAC
epikol_00904	NAT10_NAT_1	CCAGACAGACCACCAAAGTC
epikol_00905	NAT10_NAT_2	GCTGTTACCCAGATTATCA
epikol_00906	NAT10_NAT_3	CCAGACTTTGGTGGTCTGTC
epikol_00907	NAT10_NAT_4	CTGTTACCCAGATTATCAA
epikol_00908	NAT10_NAT_5	GACTTTGGTGGTCTGTCTGG
epikol_00909	NCOA1_PAS_1	CGTCTACAGCATACTGCACG
epikol_00910	NCOA1_PAS_2	CAGCTACTTAGGTTACAATC
epikol_00911	NCOA1_PAS_3	GAGAATGTAACCAGCTACTT
epikol_00912	NCOA1_PAS_4	GCATCCTGCAGTTAAAGGTA
epikol_00913	NCOA1_PAS_5	TTTCGTCGTGTTGCCTCTTG
epikol_00914	NCOA1_SRC_1	TCTACACCGGCTCTTACAGG
epikol_00915	NCOA1_SRC_2	CAACTGCCGAACAGCAGTTA
epikol_00916	NCOA1_SRC_3	GCCTCTGCTAACTCTTCAGG
epikol_00917	NCOA1_SRC_4	ACAAAGTGGTGATATCTGAG
epikol_00918	NCOA1_SRC_5	TCTGCCTCTGCTAACTCTTC
epikol_00919	NCOA2_PAS_1	AAGGTATGGCTGTTCCGCCT
epikol_00920	NCOA2_PAS_2	CTTGGTCTGGCGAACCTCCG
epikol_00921	NCOA2_PAS_3	GTATGGCTGTTCCGCCTCGG
epikol_00922	NCOA2_PAS_4	TACCTTCAATTGTCGGATGC

epikol_00923	NCOA2_PAS_5	GCATCCGACAATTGAAGGTA
epikol_00924	NCOA2_SRC_1	GTTTGTATCCGACAAAGAGC
epikol_00925	NCOA2_SRC_2	TACACATGGAACCTCGCTCA
epikol_00926	NCOA2_SRC_3	GGGCTCCATCTGATCAGATT
epikol_00927	NCOA2_SRC_4	GTTAACCTGGCCAAGTCCAC
epikol_00928	NCOA2_SRC_5	GACCACCAAATCTGATCAGA
epikol_00929	NCOA3_PAS_1	CGTCTCGATTACCCACAAAT
epikol_00930	NCOA3_PAS_2	CTATTTGTGGTGAATCGAGA
epikol_00931	NCOA3_PAS_3	GCAATATAAGCAAGAGGACC
epikol_00932	NCOA3_PAS_4	GGTCCTCTTGCTTATATTGC
epikol_00933	NCOA3_PAS_5	TGTAAACACTTGTGTTAACC
epikol_00934	NCOA3_SRC_1	ACCTGTTCTTCTGATGACCG
epikol_00935	NCOA3_SRC_2	CCTCTACATCCAATATGCAT
epikol_00936	NCOA3_SRC_3	TAACAGTGACCCATGCATAT
epikol_00937	NCOA3_SRC_4	ACCCCGGTCATCAGAAGAAC
epikol_00938	NCOA3_SRC_5	CCCATGCATATTGGATGTAG
epikol_00944	NSD1_PHD_1	CAGTCAATCCAAGGCACTCC
epikol_00945	NSD1_PHD_2	TTGGCATCTCAGTCAATCCA
epikol_00946	NSD1_PHD_3	GCTTTCCACCTGGAGTGCCT
epikol_00947	NSD1_PHD_4	GGGTGAGCTGCTGTTATGTG
epikol_00948	NSD1_PHD_5	TATGTGAGGCTCAGTGCTGT
epikol_00949	NSD1_SET_1	CCGCACATTACAGCGGGGTT
epikol_00950	NSD1_SET_2	CCCAACCCCGCTGTAATGTG
epikol_00951	NSD1_SET_3	ACACAGAAGTGGTCTGTGAA
epikol_00952	NSD1_SET_4	GTGAATGGAGATACCCGTGT
epikol_00953	NSD1_SET_5	TAGTGCAAAAAGGCTACAC
epikol_00974	PADI4_PAD_1	CACGCGTACACCTCCTGCGG
epikol_00975	PADI4_PAD_2	TGTGGTGTTCGAAGACAGCG
epikol_00976	PADI4_PAD_3	CTTGGAACACCACAGCCTCG
epikol_00977	PADI4_PAD_4	GGGAGGCCCTGCTGTTTCGAA
epikol_00978	PADI4_PAD_5	CAGGGGCAAGGAATACCCGC
epikol_00979	PARP1_PARP_1	CGCATACTCCATCCTCAGTG
epikol_00980	PARP1_PARP_2	CCCAAAGTCGTGGGGGATCA
epikol_00981	PARP1_PARP_3	ATCTTTAAGATAGAGCGTGA
epikol_00982	PARP1_PARP_4	CTAATGTTAGCTGAAGGATC
epikol_00983	PARP1_PARP_5	TCCAGACTAATGTTAGCTGA
epikol_00984	PAXIP1_BRCT_1	GGGCTTTGGTTACGTTCTAT
epikol_00985	PAXIP1_BRCT_2	ATTTCTTATTGAGGGTTAGC
epikol_00986	PAXIP1_BRCT_3	GGCTTTGGTTACGTTCTATG
epikol_00987	PAXIP1_BRCT_4	AGAACGTAACCAAAGCCAC
epikol_00988	PAXIP1_BRCT_5	TTTGATTGTTCCAGAGCCAA
epikol_00989	PBRM1_BAH_1	GGCCTGGTGCGTCCTCGTGT

epikol_00990	PBRM1_BAH_2	TCTTCTGGGTGAATGAAGAT
epikol_00991	PBRM1_BAH_3	TGGTATTTCAGTTGGCCTGC
epikol_00992	PBRM1_BAH_4	GTGTGCTGTGTTGTCATTCA
epikol_00993	PBRM1_BAH_5	GAAAAAGTATGGGTTTCGAGA
epikol_00994	PBRM1_BROMO_1	TAATACCATCCGAGACTATA
epikol_00995	PBRM1_BROMO_2	CCGAGACTATAAGGATGAAC
epikol_00996	PBRM1_BROMO_3	TCTCTGTGAGCTCTTCATTA
epikol_00997	PBRM1_BROMO_4	ATTATAGAGTTCATGGCACA
epikol_00998	PBRM1_BROMO_5	CAAACCTCATTTCTTGTTCGA
epikol_01009	PCGF2_GECKO_1	CTGCCCCATGTGTGACGTGC
epikol_01010	PCGF2_GECKO_2	TCAAGACATTGTCTACAAAT
epikol_01011	PCGF2_GECKO_3	CGATGAAGTACCCCCGCAG
epikol_01012	PCGF2_GECKO_4	CATCGACGCCACCACTATCG
epikol_01013	PCGF2_GECKO_5	AACCTGCATCGTGCCTACC
epikol_01014	PCGF2_RING_1	CATCGACGCCACCACTATCG
epikol_01015	PCGF2_RING_2	CGATGAAGTACCCCCGCAG
epikol_01016	PCGF2_RING_3	CAGGCACTCCACGATAGTGG
epikol_01017	PCGF2_RING_4	AACCTGCATCGTGCCTACC
epikol_01018	PCGF2_RING_5	GGCAGTATTGTGTTGGTCTCC
epikol_01019	PHF1_PHD_1	CAGCTGACCAGCCGGTTCCC
epikol_01020	PHF1_PHD_2	ACTTCTCACAGCTGACCAGC
epikol_01021	PHF1_PHD_3	CGCTCTGAGACTGTGGTCCC
epikol_01022	PHF1_PHD_4	TGTCTGTCGCTCTGAGACTG
epikol_01023	PHF1_PHD_5	GCTCTGAGACTGTGGTCCCT
epikol_01024	PHF1_TUDOR_1	GACTGATGGGCTGCTATACT
epikol_01025	PHF1_TUDOR_2	TGCTGGCCAGATGGACTGAT
epikol_01026	PHF1_TUDOR_3	GTGCTGGCCAGATGGACTGA
epikol_01027	PHF1_TUDOR_4	GTCAAGATGTGCTGGCCAGA
epikol_01028	PHF1_TUDOR_5	ATTCGCAGTTTCTGTTCTA
epikol_01029	PHF2_GECKO_1	GACGTCCTGACATAGAACG
epikol_01030	PHF2_GECKO_2	CCAAACTGTGAGAAAACCCA
epikol_01031	PHF2_JMJC_1	TGTCATGGGTAGAAAACCTAC
epikol_01032	PHF2_JMJC_2	AGTTTCTTTACAATGTCAGG
epikol_01033	PHF2_JMJC_3	GTA CTGCCTAATCTGCGTGA
epikol_01034	PHF2_JMJC_4	AAC TGTCCTTCACGCAGATT
epikol_01035	PHF2_JMJC_5	CACGCAGATTAGGCAGTACT
epikol_01036	PHF2_PHD_1	TGAAGCGGGTAACGTCGTAG
epikol_01037	PHF2_PHD_2	GATCGAGTGCGACGCCTGCA
epikol_01038	PHF2_PHD_3	AGTGCGACGCCTGCAAGGAC
epikol_01039	PHF8_JMJC_1	AGCTATCTCGCACACTCATG
epikol_01040	PHF8_JMJC_2	GACAGCTTTCGAACAATCTT
epikol_01041	PHF8_JMJC_3	GATTGTTTCAAAGCTGTCAT

epikol_01042	PHF8_JMJC_4	CAATCTTCGGTGTCTCCACA
epikol_01043	PHF8_JMJC_5	TGTCATGGGTCGAAAACTTG
epikol_01044	PHF8_PHD_1	GGGTCACATCGTAAGGCAGC
epikol_01045	PHF8_PHD_2	GATCGAGTGTGACATGTGCC
epikol_01046	PHF8_PHD_3	CATCGTAAGGCAGCCGGCAG
epikol_01047	PHF8_PHD_4	ATGTGCCAGGACTGGTTTCA
epikol_01048	PHF8_PHD_5	TGTTGGTGTGAAGAGGAGA
epikol_01049	PPhLN1_1	TGCTATTGCTTTAGAGCGAG
epikol_01050	PPhLN1_2	CTGTGAAACTTTCGGGATGG
epikol_01051	PPhLN1_3	AGACTGTGAAACTTTCGGGA
epikol_01052	PPhLN1_4	CGAAAGTTTCACAGTCTTGT
epikol_01053	PPhLN1_5	GACAAGACTGTGAAACTTTC
epikol_01054	PRDM1_SET_1	CAGTTCCTAAGAACGCCAAC
epikol_01055	PRDM1_SET_2	CATTAAAGCCGTCATGAAG
epikol_01056	PRDM1_SET_3	GGTCCAAAACGTGTGCCCTT
epikol_01057	PRDM1_SET_4	GGTGTAGATTTCACCTATTA
epikol_01058	PRDM1_SET_5	TGGTGTAGATTTCACCTATT
epikol_01059	PRDM10_SET_1	GCGCTTGGGGATGCGCCGCT
epikol_01060	PRDM10_SET_2	GGGGCCAAACTGGGTGCGCT
epikol_01061	PRDM10_SET_3	GGCCATACTGGTAAGCCACC
epikol_01062	PRDM10_SET_4	AATACACATGGTGGCCATAC
epikol_01063	PRDM10_SET_5	GGGATTTACATGAAGACCTA
epikol_01064	PRDM11_SET_1	CACATCTTCGGCCCCTATGA
epikol_01065	PRDM11_SET_2	TGTAAACGAGGTCATCCCA
epikol_01066	PRDM11_SET_3	GGTGGAGATCTGCCCTCAT
epikol_01067	PRDM11_SET_4	GGGCCGAAGATGTGGCCCTT
epikol_01068	PRDM11_SET_5	GCCAGCTGATTGTCTCTGGG
epikol_01069	PRDM12_SET_1	CGTGCTCCGGGCGATCACG
epikol_01070	PRDM12_SET_2	CACCGGCCGCGTGATCGCCC
epikol_01071	PRDM12_SET_3	CTCCGGGGCGATCACGCGGC
epikol_01072	PRDM12_SET_4	TCCCGCCTTGATCCACGTCT
epikol_01073	PRDM12_SET_5	GATCAAGGCGGAACCGAGA
epikol_01074	PRDM13_SET_1	GGCCGGCTTGCGCCTCGGAC
epikol_01075	PRDM13_SET_2	TTGCGCCTCGGACCGGTGCC
epikol_01076	PRDM13_SET_3	CACCGGTCCGAGGCGCAAGC
epikol_01077	PRDM13_SET_4	CAAGTACCTGTCAGACCGCA
epikol_01078	PRDM13_SET_5	GCTGGACGTCTCGCAATGCT
epikol_01079	PRDM14_SET_1	CATCACAGAATTGTCTCCGT
epikol_01080	PRDM14_SET_2	TCCGTAGGTCTTCACTTCAC
epikol_01081	PRDM14_SET_3	GAAGCGGGCACAGTTGACAT
epikol_01082	PRDM14_SET_4	GGGCCCAAACCTGACTCCTT
epikol_01083	PRDM14_SET_5	TCGCCAAAGGAGTCAGGTTT

epikol_01084	PRDM15_SET_1	GTCCAGGAGGGTCGCCAAAT
epikol_01085	PRDM15_SET_2	AGTCCAGGAGGGTCGCCAAA
epikol_01086	PRDM15_SET_3	TTCCCATTTGGCGACCCTCC
epikol_01087	PRDM15_SET_4	CCGCTTGACGAGCTGAGTGA
epikol_01088	PRDM15_SET_5	CGACTGGAAGATGGAGCCGA
epikol_01089	PRDM16_SET_1	AAAGGAGACAGACTTCGGAT
epikol_01090	PRDM16_SET_2	GTGGAAGTGTCGCCCCAGGA
epikol_01091	PRDM16_SET_3	GCGTGGATGCAAATCAGGCG
epikol_01092	PRDM16_SET_4	GTTCTGCGTGGATGCAAATC
epikol_01093	PRDM16_SET_5	GTCATTAAGGACATTGAGCC
epikol_01094	PRDM2_SET_1	TATGTGAATTGGGCTTGCTC
epikol_01095	PRDM2_SET_2	GGGTCTTGCAACAGCAGAA
epikol_01096	PRDM2_SET_3	AAAAAATTTGGGCCATTTGT
epikol_01097	PRDM2_SET_4	ACTGGCTGCGATATGTGAAT
epikol_01098	PRDM2_SET_5	CTGGCTGCGATATGTGAATT
epikol_01099	PRDM3_SET_1	GAAAGACCCCAGTTATGGAT
epikol_01100	PRDM3_SET_2	GTCAACCAGATGTTGGAAGC
epikol_01101	PRDM3_SET_3	GATGCCAGTCAACCAGATGT
epikol_01102	PRDM3_SET_4	CTCAAGTACATTAGATTCGC
epikol_01103	PRDM3_SET_5	GTAGTTGCAGACATTGCGCC
epikol_01104	PRDM4_SET_1	GGTCCAAAGCAAGTCCGCAC
epikol_01105	PRDM4_SET_2	ATTCTGTGCGGACTTGCTT
epikol_01106	PRDM4_SET_3	ACTTGCTTTGGACCTCTAAT
epikol_01107	PRDM4_SET_4	ACAAGGCAGTTAACCATATC
epikol_01108	PRDM4_SET_5	TGACTCTGCTGGCCAATTAG
epikol_01109	PRDM5_SET_1	TGGATGAAAATATGGATTAC
epikol_01110	PRDM5_SET_2	GAAGCCAGTTGGAGTGCCGT
epikol_01111	PRDM5_SET_3	CTGGCTTCGCTTCGTTTCATG
epikol_01112	PRDM5_SET_4	ATGAACGAAGCGAAGCCAGT
epikol_01113	PRDM5_SET_5	ACAGACACGGAGCTTCTGAT
epikol_01114	PRDM6_SET_1	GTCCAATCCAGGTGCCTTGC
epikol_01115	PRDM6_SET_2	ACCTGGATTGGACCTTTCCA
epikol_01116	PRDM6_SET_3	GGGAACCTAGTAAGTCGAGC
epikol_01117	PRDM6_SET_4	CTAGTAAGTCGAGCTGGATG
epikol_01118	PRDM6_SET_5	CTACAGCACTTTATTGATGG
epikol_01119	PRDM8_SET_1	AGCATTCTCAGGGATGTCGC
epikol_01120	PRDM8_SET_2	GAACAGTACCGTATATCTTT
epikol_01121	PRDM8_SET_3	CATATAGGGAAGTATGGCTC
epikol_01122	PRDM8_SET_4	TGTTCTACCGCTCTCTCCGC
epikol_01123	PRDM8_SET_5	AACTAGTAACCTCTCGTCTT
epikol_01124	PRDM9_GECKO_1	CAGAAGATTCTGATGAAGAA
epikol_01125	PRDM9_GECKO_2	GTTCTCTCTGTGTCTTCTTC

epikol_01126	PRDM9_GECKO_3	TCCCTCTTACCTTGCTGCCT
epikol_01127	PRDM9_SET_1	CAGCCAGGAGTATCCATTGT
epikol_01128	PRDM9_SET_2	TGGATGGAAAAGATAAAATCC
epikol_01129	PRMT1_SAM_1	CCGGGGCCCGCAAGGTCATC
epikol_01130	PRMT1_SAM_2	CCCGATGACCTTGCGGGCCC
epikol_01131	PRMT1_SAM_3	TGCTCTATGCCCCGGGACAAG
epikol_01132	PRMT1_SAM_4	CTTGTCGCCGGGCATAGAGCA
epikol_01133	PRMT1_SAM_5	CAACACCGTGCTCTATGCCC
epikol_01134	PRMT2_SAM_1	TCTTCTGTGCACACTATGCG
epikol_01135	PRMT2_SAM_2	TGTCAGCAAAGCCGTTCTGC
epikol_01136	PRMT2_SAM_3	GCCCCGAGAAGGTGGACGTGC
epikol_01137	PRMT2_SAM_4	CGAGTCCATCCTGTATGCCC
epikol_01138	PRMT2_SAM_5	TGAAGGAGGACGGGGTCATT
epikol_01139	PRMT5_SAM_1	ATACAGCTTTATCCGCCGCT
epikol_01140	PRMT5_SAM_2	CAGCATAACAGCTTTATCCGC
epikol_01141	PRMT5_SAM_3	CCGCAGGGAAGCGTTCACCA
epikol_01142	PRMT5_SAM_4	CTGATGAGACTACGGTCACT
epikol_01143	PRMT5_SAM_5	CGTAGTCTCATCAGACATGA
epikol_01144	PRMT6_SAM_1	TCAGCCACTTGTTTCGCGCG
epikol_01145	PRMT6_SAM_2	CTCTACCGCGTACACGCGCC
epikol_01146	PRMT6_SAM_3	CCGGCGCGTGTACGCGGTAG
epikol_01147	PRMT6_SAM_4	CCTCTACCGCGTACACGCGC
epikol_01148	PRMT6_SAM_5	GGTGGTGCGGTTCACGGGC
epikol_01149	PRMT7_SAM_1	CACGGGACTCTTGTCATGA
epikol_01150	PRMT7_SAM_2	CTCGATGGCATAGCAGAAGT
epikol_01151	PRMT7_SAM_3	TGGCTTTAGTGATAAGATTA
epikol_01152	PRMT7_SAM_4	TTCACAGCAGCATCAGCCAT
epikol_01153	PRMT7_SAM_5	TCAATGATGGCGGTACAGC
epikol_01154	PRMT8_SAM_1	CCTTTCCATGTTTCGCTGCCA
epikol_01155	PRMT8_SAM_2	CCCTGCCTTGGCAGCGAACA
epikol_01156	PRMT8_SAM_3	CCATGTTTCGCTGCCAAGGCA
epikol_01157	PRMT8_SAM_4	TCAACACGGTGATCTTTGCC
epikol_01158	PRMT8_SAM_5	GATCATTAAGGCCAACCACT
epikol_01159	RBBP4_GECKO_1	CTACTCCATGAGTCTCTGTT
epikol_01160	RBBP4_GECKO_2	AGTAAGTGCCCACTGAGATT
epikol_01161	RBBP4_GECKO_3	AACCCAGACTTGCGTCTCCG
epikol_01162	RBBP4_GECKO_4	GACCTTGTCAGCTGATCCTG
epikol_01163	RBBP4_GECKO_5	CATTTAAGACCATCTGCCTG
epikol_01164	RBBP5_WD40_1	ATGATGGCCGAATTGTCATC
epikol_01165	RBBP5_WD40_2	TTTCAGGCGACTGTGACCAG
epikol_01166	RBBP5_WD40_3	AAATCGGAGTCATCGTCCAC
epikol_01167	RBBP5_WD40_4	GTCAATTGAGTTTGCCCGGA

epikol_01168	RBBP5_WD40_5	GATGGGGAATACATCGTGCC
epikol_01169	RBX1_RING_1	CTTCTGAAGTAGCGGACGCC
epikol_01170	RBX1_RING_2	AGAAGAGTGTACTGTGCGAT
epikol_01171	RBX1_RING_3	TCGCTGGCTCAAAACACGAC
epikol_01172	RBX1_RING_4	AGTACACTCTTCTGAAGTAG
epikol_01173	RBX1_RING_5	ACTTCCACTGCATCTCTCGC
epikol_01174	RNF168_RING_1	TCGACGGTCGACTGGAAGCA
epikol_01175	RNF168_RING_2	CTTCCAGTCGACCGTCGAAA
epikol_01176	RNF168_RING_3	CAGCGTGTGGTTACACGGGA
epikol_01177	RNF168_RING_4	AGCACGGTTTACACAGCGTG
epikol_01178	RNF168_RING_5	GATCTGCATGGAAATCCTCG
epikol_01179	RNF2_GECKO_2	AATTCAGTGTGTAGACTTCG
epikol_01180	RNF2_GECKO_3	ATTATTGTGCTTGTGATCC
epikol_01181	RNF2_GECKO_4	TTTCAACATACCCACTTCTA
epikol_01182	RNF2_GECKO_5	TGAGTTACAACGAACACCTC
epikol_01183	RNF2_RING_1	TCATCACAGCCCTTAGAAGT
epikol_01184	RNF20_RING_1	ACGCATGTTACAGCACGGAC
epikol_01185	RNF20_RING_2	TACACTTGGGACATTTGCGC
epikol_01186	RNF20_RING_3	CTTGGGACATTTGCGCTGGC
epikol_01187	RNF20_RING_4	GTGCTGTAACATGCGTAAAA
epikol_01188	RNF20_RING_5	ACTTGGGACATTTGCGCTGG
epikol_01189	RNF8_RING_1	AGAATGCCCCATTGTGCGGA
epikol_01190	RNF8_RING_2	GCTCCTACTGTATCAATGAA
epikol_01191	RNF8_RING_3	GAAACTGTGGGCACAGTTCA
epikol_01192	RNF8_RING_4	ACAGTAGGAGCAGAAACTGT
epikol_01193	RNF8_RING_5	AGATAGAATGCCCCATTGT
epikol_01194	SATB1_1	GCAATGCCATTTTCGATCAGC
epikol_01195	SATB1_2	CGCCATTGAATATGATTGCA
epikol_01196	SATB1_3	GGAGCATGCAGAATTTGTGC
epikol_01197	SATB1_4	TGTTCCACCACACAGAAAAC
epikol_01198	SATB1_5	AGAATTTGTGCTGGTGAGAA
epikol_01199	SETD1A_SET_1	CCGGAGCCGGATCCACGAGT
epikol_01200	SETD1A_SET_2	GGATCCGGCTCCGGCCAAAT
epikol_01201	SETD1A_SET_3	TGTCGTGGTCCACCCGGAAC
epikol_01202	SETD1A_SET_4	GCGGGAGAAGCGCTACGTGC
epikol_01203	SETD1A_SET_5	ATCTCCTCGTCCACGCCAAT
epikol_01204	SETD2_SET_1	GTCATACTCACAGAAAAGAA
epikol_01205	SETD2_SET_2	AGAGTTTAAAGCTCGAGTGA
epikol_01206	SETD2_SET_3	ACCTTTGTCCTAGAATATTG
epikol_01207	SETD2_SET_4	GGACTGTGAACGGACAACCTG
epikol_01208	SETD2_SET_5	AACGTAACTCTGAGCCTGA
epikol_01209	SETD3_SET1_1	GGCCAGTGCGATGTTTCCCA

epikol_01210	SETD3_SET1_2	AGCCATGGGAAACATCGCAC
epikol_01211	SETD3_SET1_3	GACCGAATCCTTCAAGCCAT
epikol_01212	SETD3_SET1_4	CGATGTTTCCCATGGCTTGA
epikol_01213	SETD3_SET1_5	GAGCCAGCCCTAACTCCTTC
epikol_01214	SETD3_SET2_1	GAAGGATTCTTTCACTTACG
epikol_01215	SETD3_SET2_2	TGCCAACAACTACCCCTTGA
epikol_01216	SETD3_SET2_3	GGATGGTTCCTCGGTGACCC
epikol_01217	SETD3_SET2_4	CAGTCTCTTCTGTTATGACG
epikol_01218	SETD3_SET2_5	ACGCGGGAACCATCCTCTGT
epikol_01219	SETD4_SET_1	CTGTGCACTCGCTCCGTACC
epikol_01220	SETD4_SET_2	GGGCCTCAGGTACACGGCTC
epikol_01221	SETD4_SET_3	GAAAGGCATTCCCCTGCCT
epikol_01222	SETD4_SET_4	AAATTAGAACGACTTCACGT
epikol_01223	SETD4_SET_5	GAAGAGGTATTCATCTGTTA
epikol_01224	SETD6_SET_1	GTTCTAGGGAGCGAACCCCTG
epikol_01225	SETD6_SET_2	TGAAGGGCAGCACGATGGAC
epikol_01226	SETD6_SET_3	TTCTAGATTGGCGTTGTGAT
epikol_01227	SETD6_SET_4	GGCACCATCACGGGGGAGTT
epikol_01228	SETD6_SET_5	ATCTCATGGCCTTTAGGAAT
epikol_01229	SETD7_SET_1	GATACGTGGTTATAGGGCTC
epikol_01230	SETD7_SET_2	AGGCACAGTACTTGGATACG
epikol_01231	SETD7_SET_3	TCCCAAGGAGGCACAGTACT
epikol_01232	SETD7_SET_4	CATTTGATGGGCCCAAACG
epikol_01233	SETD7_SET_5	GAAGAGCTCACCGTTGCCTA
epikol_01234	SETD8_GECKO_1	TGGGCAAATACCTCTAGCCA
epikol_01235	SETD8_GECKO_2	AGGTACCAGGGGACCTTTCC
epikol_01236	SETD8_GECKO_3	TAAGTGAGTCTCTTCCTGA
epikol_01237	SETD8_GECKO_4	ACGGAGCGCCATGAAGTCCG
epikol_01238	SETD8_GECKO_5	GCCTGCTTACCCCGTCGGTG
epikol_01239	SETDB1_SET1_1	GGGCTGGGTATCCGCTGCT
epikol_01240	SETDB1_SET1_2	TGCTTGGATGACATTGCCAA
epikol_01241	SETDB1_SET1_3	TCTGGACCATATCGAGAGCG
epikol_01242	SETDB1_SET1_4	AGTTCTCCACGCTCTCGATA
epikol_01243	SETDB1_SET1_5	CTTTGCAGACAAGGAGGGTC
epikol_01244	SETDB1_SET2_1	ATCATTGATGCCAAGCTTGA
epikol_01245	SETDB1_SET2_2	GGGTATCCACGAAGACATTC
epikol_01246	SETDB1_SET2_3	GTTTGTCCAGAATGTCTTCG
epikol_01247	SETDB1_SET2_4	GGGCTGGGACAGAACTTACT
epikol_01248	SETDB1_SET2_5	GCCAAGCTTGAAGGCAACCT
epikol_01249	SETDB2_SET1_1	TGTCTAGATGACATTGACAG
epikol_01250	SETDB2_SET1_2	GGTGTTCAAACTGAGCAGA
epikol_01251	SETDB2_SET1_3	CTTATGGTATTGATGAAAAC

epikol_01252	SETDB2_SET1_4	AAAACTGAGCAGAAGGGATG
epikol_01253	SETDB2_SET1_5	GCTAACACTGAAAAATCTTA
epikol_01254	SETDB2_SET2_1	TTATTGGATGCCACAAAAGA
epikol_01255	SETDB2_SET2_2	TTGGTGAAGAATGCCACCAA
epikol_01256	SETDB2_SET2_3	ACATTCTGTACCAAGAGATT
epikol_01257	SETDB2_SET2_4	AAGCAAGAACAGAGCTAACA
epikol_01258	SETDB2_SET2_5	ACAGAGCTAACATGGGATTA
epikol_01259	SETMAR_SET_1	ACCATTGAGTCAATTCGGAC
epikol_01260	SETMAR_SET_2	AGGCTGGGGACTTCGTACCT
epikol_01261	SETMAR_SET_3	TGACTCAATGGTACCTAAGT
epikol_01262	SETMAR_SET_4	GCGTCTTGAACACTTGAAG
epikol_01263	SETMAR_SET_5	CCAATTACATTATAGCCATC
epikol_01264	SFMBT1_1	GGTCCACAGCAGTTCCCTAT
epikol_01265	SFMBT1_2	CTCACAGGTAGTGATAACGG
epikol_01266	SFMBT1_3	CATAGCCATCATAGCGGAGA
epikol_01267	SFMBT1_4	AGGCTGATCTCTACCCCAT
epikol_01268	SFMBT1_5	GTGGACACACGTTTGCAAAA
epikol_01269	SFMBT2_1	TCTGCGCTACTGCGGTTACG
epikol_01270	SFMBT2_2	CTGGTGTGACGTAGTCATCG
epikol_01271	SFMBT2_3	CCCGCACGTGGTAATGATCG
epikol_01272	SFMBT2_4	CATCGCGGATTTCACCCCG
epikol_01273	SFMBT2_5	GAACAACCCGGACACGTACT
epikol_01274	SHPRH_H15_1	GATACGATGTTCAACGGAAC
epikol_01275	SHPRH_H15_2	GGTCTCGTGAAACAGATCAA
epikol_01276	SHPRH_H15_3	AGAACTGACATACTTATAGA
epikol_01277	SHPRH_H15_4	ATACTTATAGATGGAAAGGA
epikol_01278	SHPRH_H15_5	GTGAAACAGATCAAAGGCCA
epikol_01279	SHPRH_HELIC_1	CGTTGAGAAAACGAGTGCTT
epikol_01280	SHPRH_HELIC_2	AGCTCTTACTGACAACAACA
epikol_01281	SHPRH_HELIC_3	ATAGGGAGGGTGCACCGAAT
epikol_01282	SHPRH_HELIC_4	TCATGGGCAGGGTTCAATAT
epikol_01283	SHPRH_HELIC_5	ATCCATTAGAACCTGTGTGC
epikol_01284	SIN3A_1	CTTTGGTAAGGCTCGATAGC
epikol_01285	SIN3A_2	AGGAGTCCGTCCTGTACACT
epikol_01286	SIN3A_3	TCTGGCAACAATCCGGGTTC
epikol_01287	SIN3A_4	AGACCAATCTGGCAACAATC
epikol_01288	SIN3A_5	CTGCTGGTAACTCTTTGGTA
epikol_01289	SIN3B_1	GCTGGATCCTATGCGCTTGC
epikol_01290	SIN3B_2	TTTGGGGAGTGCCCGGTAGC
epikol_01291	SIN3B_3	AATCGTCTGTCCACGAACCC
epikol_01292	SIN3B_4	TGGATTCCAGGATATAGTGC
epikol_01293	SIN3B_5	CTTCTGGAGCTGACGAACG

epikol_01294	SIRT1_SIRT_1	TCAAACATCGCTTGAGGATC
epikol_01295	SIRT1_SIRT_2	CCATCCCTTGACCTGAAGTC
epikol_01296	SIRT1_SIRT_3	TACCCAGAACATAGACACGC
epikol_01297	SIRT1_SIRT_4	AACAGGTTGCGGGAATCCAA
epikol_01298	SIRT1_SIRT_5	ATAGCAAGCGGTTTCATCAGC
epikol_01299	SIRT2_SIRT_1	TAGGTTGTCATAGAGGCCGG
epikol_01300	SIRT2_SIRT_2	TTCCTTGCGGAGGGCGAAGA
epikol_01301	SIRT2_SIRT_3	CTACTTCATGCGCCTGTGTA
epikol_01302	SIRT2_SIRT_4	CCAGCCCGGCTATTCGCTCC
epikol_01303	SIRT2_SIRT_5	GGCACGAATACCCGCTAAGC
epikol_01304	SIRT3_SIRT_1	GGTACGGGAGATCGTACTGC
epikol_01305	SIRT3_SIRT_2	CTCTACACGCAGAACATCGA
epikol_01306	SIRT3_SIRT_3	CAAAAATGGCCTCGGGGTAC
epikol_01307	SIRT3_SIRT_4	ATGAGCTTCAACCAGCTTTG
epikol_01308	SIRT3_SIRT_5	GGGTCTTTGCGAGACTGTGC
epikol_01309	SIRT4_SIRT_1	GTCTGGTATCCCCGATTCGG
epikol_01310	SIRT4_SIRT_2	AGTTTCTCGCCAGTACCGC
epikol_01311	SIRT4_SIRT_3	TTTATGCCCGCACTGACCGC
epikol_01312	SIRT4_SIRT_4	TACTGGGCGAGAACTTCGT
epikol_01313	SIRT4_SIRT_5	GTCCCAACCTGCGTTCAATG
epikol_01314	SIRT5_SIRT_1	AGTGGTGTCCGACCTTCAG
epikol_01315	SIRT5_SIRT_2	GGGTGATGACCACGACTCGC
epikol_01316	SIRT5_SIRT_3	GTTCTACCACTACCGCGGG
epikol_01317	SIRT5_SIRT_4	TATGGCGCGGTGCCCGGCGT
epikol_01318	SIRT5_SIRT_5	GGGAGTTCTACCACTACCGG
epikol_01319	SIRT6_SIRT_1	GTGGTGTTCACACGGGTGC
epikol_01320	SIRT6_SIRT_2	CGTGGGCCGCGGCTCTCAA
epikol_01321	SIRT6_SIRT_3	CTCAAAGGTGGTGTGCAACT
epikol_01322	SIRT6_SIRT_4	CATGGAGCCCGTCCACGTTT
epikol_01323	SIRT6_SIRT_5	CTTCCACAAACATGTTCCCG
epikol_01324	SIRT7_SIRT_1	GCGTCTATCCCAGACTACCG
epikol_01325	SIRT7_SIRT_2	GGGTGGCTCGGCCTCGCTC
epikol_01326	SIRT7_SIRT_3	CAGACGGGTGATGCTCATGT
epikol_01327	SIRT7_SIRT_4	GTACATGTTCCCGTGGAGCT
epikol_01328	SIRT7_SIRT_5	AGCTCGGAGATGGCCGTGCG
epikol_01329	SMARCA1_DEXD_1	GAATATTTCCGGTAGTGTTTC
epikol_01330	SMARCA1_DEXD_2	AAAGTTTCACTGGCGATACC
epikol_01331	SMARCA1_DEXD_3	ATTCGTGATGAAATGATGCC
epikol_01332	SMARCA1_DEXD_4	ACGCAAACATCCCCTCTCC
epikol_01333	SMARCA1_DEXD_5	ACTAACCGCTTGCTCCTAAC
epikol_01334	SMARCA1_HELIC_1	GGTAGTTCTGGATAAACTAT
epikol_01335	SMARCA1_HELIC_2	TGGTTATGAGTATTGTGCAC

epikol_01336	SMARCA1_HELIC_3	CAGCCAGATGACTCGCTTGC
epikol_01337	SMARCA1_HELIC_4	GAAGATTATTGCATGTGGCG
epikol_01338	SMARCA1_HELIC_5	TGGAGGTCTCGGAATTAACC
epikol_01339	SMARCA2_DEXD_1	CGGAATCTTAGCCGATGAAA
epikol_01340	SMARCA2_DEXD_2	GATGAGTGCAATGGTCTGTA
epikol_01341	SMARCA2_DEXD_3	TGACAAATGGGCTCCTTCTG
epikol_01342	SMARCA2_DEXD_4	ACAAGGGAGCGACGCATGGC
epikol_01343	SMARCA2_DEXD_5	TGAATTTGCCACTCCGTAGC
epikol_01344	SMARCA2_HELIC_1	GCTGAACTGTATCGGGCCTC
epikol_01345	SMARCA2_HELIC_2	AGCAGCACGATCTTCAGACT
epikol_01346	SMARCA2_HELIC_3	TGGCCTGGGCTTAAATCTTC
epikol_01347	SMARCA2_HELIC_4	CGATGCGGTGAGCTCGGTCT
epikol_01348	SMARCA2_HELIC_5	TTCTTGCTGAGCACAAGAGC
epikol_01349	SMARCA4_DEXD_1	GAGGTACGTGATGAGCGCGA
epikol_01350	SMARCA4_DEXD_2	GGCATCCTGGCCGACGAGAT
epikol_01351	SMARCA4_DEXD_3	GTCAAACTCGTACGCCCAGT
epikol_01352	SMARCA4_DEXD_4	TGACAAGTGGGCCCCCTCCG
epikol_01353	SMARCA4_DEXD_5	ACAAAGGCCCGTCTTGCTGC
epikol_01354	SMARCA4_HELIC_1	GAAGATTACTTTGCGTATCG
epikol_01355	SMARCA4_HELIC_2	ATCGCGGCTTTAAATACCTC
epikol_01356	SMARCA4_HELIC_3	GTGGTTGGTTGCTCGGAGTT
epikol_01357	SMARCA4_HELIC_4	CTCAGAGCCGGGCTCGTTGA
epikol_01358	SMARCA4_HELIC_5	ACCACGAAGGCGGAGGACCG
epikol_01359	SMARCA5_DEXD_1	ACCAAAACCATATGAGGCCC
epikol_01360	SMARCA5_DEXD_2	TAGTAATAGATGAAGCTCAC
epikol_01361	SMARCA5_DEXD_3	TTAGGAACCAAAACCATATG
epikol_01362	SMARCA5_DEXD_4	ACACATACATCCCATTCTCC
epikol_01363	SMARCA5_DEXD_5	GATGAGTGAATTCAAGAGAT
epikol_01364	SMARCA5_HELIC_1	ATGCATCTAGTAACCAACAG
epikol_01365	SMARCA5_HELIC_2	AAATTATGAGTACTGCAGGT
epikol_01366	SMARCA5_HELIC_3	CAGTCAAATGACAAGGGTAT
epikol_01367	SMARCA5_HELIC_4	TTCATGTTAAGCACGCGTGC
epikol_01368	SMARCA5_HELIC_5	TCTGCCCCAATTCTATGTGCT
epikol_01369	SMARCAD1_DEXD_1	CTGGCATACTCTATCAGGA
epikol_01370	SMARCAD1_DEXD_2	CATTATTACCTCCTGATAG
epikol_01371	SMARCAD1_DEXD_3	TTTATGGTGCCCTACTTTGA
epikol_01372	SMARCAD1_DEXD_4	ATGACCGTAGTCTGTTTCGA
epikol_01373	SMARCAD1_DEXD_5	CATAAGGTGCTGGTAGCGAA
epikol_01374	SMARCAD1_HELIC_1	GAGTCTTTCCATCTAATCTG
epikol_01375	SMARCAD1_HELIC_2	TAGCCAATTTACCATGATGC
epikol_01376	SMARCAD1_HELIC_3	CCATAGAGTAGGCCAGACTA
epikol_01377	SMARCAD1_HELIC_4	TGATGAGTTTAATACCGATA

epikol_01378	SMARCAD1_HELIC_5	TCAACAAAAGCTGGTGGATT
epikol_01379	SMARCAL1_DEXD_1	CGGCCACTCCTTCCGGTAAA
epikol_01380	SMARCAL1_DEXD_2	CTCCCAGGTGAAGCGCACGG
epikol_01381	SMARCAL1_DEXD_3	CCCAGATTGCATCAACGTCG
epikol_01382	SMARCAL1_DEXD_4	TGCATCAACGTCGTGGTGAC
epikol_01383	SMARCAL1_DEXD_5	GGGAAGGACCGCCTGACAGC
epikol_01384	SMARCAL1_HELIC_1	GCTCTTGGCGTAATTGCGTCC
epikol_01385	SMARCAL1_HELIC_2	TGCACACCATAAGGTGGTCC
epikol_01386	SMARCAL1_HELIC_3	GTGCAGCACATCCGCATCGA
epikol_01387	SMARCAL1_HELIC_4	TCCATCACCGCTGCCAATAT
epikol_01388	SMARCAL1_HELIC_5	AGCAGTTCCAACTGTGCGGAG
epikol_01389	SMARCC2_CHROM_1	GGGTACGACCAGTCATGAAG
epikol_01390	SMARCC2_CHROM_2	ATAAGCAGGTTCTTCTGCAC
epikol_01391	SMARCC2_CHROM_3	GGTACGACCAGTCATGAAGA
epikol_01392	SMARCC2_CHROM_4	TGCTTATCCCTCTTCATGAC
epikol_01393	SMARCC2_CHROM_5	GATCCCAGCGAGTGAAATTG
epikol_01394	SMARCD1_RSC6_1	GATGGTTGTCTGGCCCATAC
epikol_01395	SMARCD1_RSC6_2	GTATGGGCCAGACAACCATC
epikol_01396	SMARCD1_RSC6_3	GGCCGGGAGACGTGAATGTA
epikol_01397	SMARCD1_RSC6_4	GCACAGGACCGCCACTACCC
epikol_01398	SMARCD1_RSC6_5	GTGTACTGTCCTACTGATGC
epikol_01399	SMARCD3_RSC6_1	TGCCTCCTGGGAGCTACGGG
epikol_01400	SMARCD3_RSC6_2	CAAGAGTTTGGTCATCGAGC
epikol_01401	SMARCD3_RSC6_3	GCATCGGACACCCACGACCC
epikol_01402	SMARCD3_RSC6_4	GAGCCGCTCAGCCATTGTCC
epikol_01403	SMARCD3_RSC6_5	GGCGGGGATCCAGTTGAAC
epikol_01404	SMARCE1_GECKO_1	AAACTTACTGTTGCAGGAGC
epikol_01405	SMARCE1_GECKO_2	TTACCATCTGGATCTTCAGC
epikol_01406	SMARCE1_GECKO_3	GTGGGAGGTGGGGCATAAGA
epikol_01407	SMARCE1_GECKO_4	TCGACAGAGACAATCTCGCA
epikol_01408	SMARCE1_GECKO_5	AGACGACGAGAACATTCCGA
epikol_01409	SMYD1_SET_1	GCCAACAGTCATGGTTCACC
epikol_01410	SMYD1_SET_2	AGAATTGAGCTCCGGGCCCT
epikol_01411	SMYD1_SET_3	GCCCTAGGCAAGATCTCAGA
epikol_01412	SMYD1_SET_4	TACAGTTGGGCCAACAGTCA
epikol_01413	SMYD1_SET_5	CGTAGGCATCTTCCCAACC
epikol_01414	SMYD2_SET_1	AGATGAAGAACTTTCTCATT
epikol_01415	SMYD2_SET_2	TGTGACCTACAAAGGGACCC
epikol_01416	SMYD2_SET_3	TTCTGCCAGGGTCCCTTTGT
epikol_01417	SMYD2_SET_4	GCTGTACAGGAAATCAAGCC
epikol_01418	SMYD2_SET_5	GGCAGAAGTCAGAGCTGTAC
epikol_01419	SMYD3_SET_1	GCTTCAAAAAGGTCAAAGGC

epikol_01420	SMYD3_SET_2	GCGAGCAGTCCGAGACATCG
epikol_01421	SMYD3_SET_3	GACTGCTCGCAGTAAGAGGT
epikol_01422	SMYD3_SET_4	TGAACACAATCGAACAGTTG
epikol_01423	SMYD3_SET_5	ATTGAACACAATCGAACAGT
epikol_01424	SMYD4_SET_1	ACAGCTTCAGTGTAACGCTC
epikol_01425	SMYD4_SET_2	TTCTAATCCGCTGTGACGCC
epikol_01426	SMYD4_SET_3	GGAGCATCGTTACCGACAGC
epikol_01427	SMYD4_SET_4	AATCCGCTGTGACGCCCGGA
epikol_01428	SMYD4_SET_5	AGTGCTAATGAAGGACACGC
epikol_01429	SMYD5_SET1_1	CAGGGGCCGTTCTACGAAGA
epikol_01430	SMYD5_SET1_2	GGGAGACCATCTTCGTAGAA
epikol_01431	SMYD5_SET1_3	TGCCACACAGCTCATCCGGA
epikol_01432	SMYD5_SET1_4	CATCTTCGTAGAACGGCCCC
epikol_01433	SMYD5_SET1_5	CGAAGATGGTCTCCCCCTTC
epikol_01434	SMYD5_SET2_1	TTTCTTAAGTGTGAAGGATC
epikol_01435	SMYD5_SET2_2	GCTCTGGAGGATATTAAGCC
epikol_01436	SMYD5_SET2_3	CTGCATTGGGCACACAAGTG
epikol_01437	SMYD5_SET2_4	GGAAAGGAGGTCTCTGCATT
epikol_01438	SMYD5_SET2_5	TGGAAAGGAGGTCTCTGCAT
epikol_01439	SP100_BROMO_1	AATCACAGTAGACCTTCAAG
epikol_01440	SP100_BROMO_2	CAGATGTACACCCGAGTAGA
epikol_01441	SP100_BROMO_3	CTGCACAAACCTTCTACTC
epikol_01442	SP100_BROMO_4	TTAACCACATGGGCTTCTGT
epikol_01443	SP100_BROMO_5	TTGACTTTGTTTAACCACAT
epikol_01444	SRCAP_DEXD_1	CATCTCTCTGCTTGCCCACT
epikol_01445	SRCAP_DEXD_2	CAGTTCAACATCACGCTGGT
epikol_01446	SRCAP_DEXD_3	CCACCAGCGTGATGTTGAAC
epikol_01447	SRCAP_DEXD_4	GCGCCAGTTCTTGCGACGGA
epikol_01448	SRCAP_DEXD_5	GATAGCGCCAGTTCTTGCGA
epikol_01449	SRCAP_HELIC_1	CATCCAGCATTGCGGTCATC
epikol_01450	SRCAP_HELIC_2	TAGTAGATCCATCCAGGCGC
epikol_01451	SRCAP_HELIC_3	CACCCAGATGACCCGAATGC
epikol_01452	SRCAP_HELIC_4	CAGGACCGCTGTACCGAAT
epikol_01453	SRCAP_HELIC_5	TCATCCTTTCAACTCGGAGT
epikol_01454	SSRP1_PHD1_1	CGAATGTCATAACGACCACG
epikol_01455	SSRP1_PHD1_2	GACGCAGTACTGTGGTGTAG
epikol_01456	SSRP1_PHD1_3	CTCGTGGTCGTTATGACATT
epikol_01457	SSRP1_PHD1_4	TCAGGAAGTGGTAGCGAGTT
epikol_01458	SSRP1_PHD1_5	CGAGTTTGGCCTTGCTTGAT
epikol_01459	SSRP1_PHD2_1	CTCAGGACTGCTTACCCGC
epikol_01460	SSRP1_PHD2_2	TCATCGAAGCGGATGTGCAC
epikol_01461	SSRP1_PHD2_3	GTCAAAGGAACGAGTAGTAG

epikol_01462	SSRP1_PHD2_4	GACTGCTCTACCCGCTGGAG
epikol_01463	SSRP1_PHD2_5	TCCTTTGTCAACTTTGCTCG
epikol_01464	SUPT16H_1	GAAACTGGACCCGCTCAAAG
epikol_01465	SUPT16H_2	TACAGGAATGGCGTTGATCA
epikol_01466	SUPT16H_3	ATGGGGTCAAGAGAGGCTAC
epikol_01467	SUPT16H_4	CTCTTGACCCCATCAAGGAA
epikol_01468	SUPT16H_5	CGCTGGTAAATGCTACGGAA
epikol_01469	SUPT16H_FACT_1	TCCTTGACAAAAGTCGCTTC
epikol_01470	SUPT16H_FACT_2	GCATCAAATATTAAGGCACC
epikol_01471	SUPT16H_FACT_3	GGCTGGTACTGTCTGTTCTC
epikol_01472	SUPT16H_FACT_4	ATACCGAGCATCAAATATTA
epikol_01473	SUPT16H_FACT_5	GCTGGTACTGTCTGTTCTCC
epikol_01474	SUV39H1_CHROM_1	GGATCTTCTTGTAATCGCAC
epikol_01475	SUV39H1_CHROM_2	ACACACTTGAGATTCTGCCG
epikol_01476	SUV39H1_CHROM_3	TATTACCTGGTGAAATGGCG
epikol_01477	SUV39H1_CHROM_4	GATATCCACGCCATTTACC
epikol_01478	SUV39H1_CHROM_5	GAGATTCTGCCGTGGCTCCC
epikol_01479	SUV39H1_SET_1	AGCTTCGTCATGGAGTACGT
epikol_01480	SUV39H1_SET_2	TTCCGCACGGATGATGGGCG
epikol_01481	SUV39H1_SET_3	CGTGAGGACGTGTACACCG
epikol_01482	SUV39H1_SET_4	GGGCCAGATCTACGACCGTC
epikol_01483	SUV39H1_SET_5	CGCCCTGACGGTCGTAGATC
epikol_01484	SUV39H2_CHROM_1	GGAATACTTGTGTGACTACA
epikol_01485	SUV39H2_CHROM_2	GGCCAGATTCTACAAATACT
epikol_01486	SUV39H2_CHROM_3	ATCTTGTAATAATGGAAAGGA
epikol_01487	SUV39H2_CHROM_4	TGGAATATTATCTTGTAATA
epikol_01488	SUV39H2_CHROM_5	TATTATCTTGTAATAATGGAA
epikol_01489	SUV39H2_SET_1	TTTCGAACTAGCAATGGACG
epikol_01490	SUV39H2_SET_2	AACTAGCAATGGACGTGGCT
epikol_01491	SUV39H2_SET_3	TGCATCTTTCGAACTAGCAA
epikol_01492	SUV39H2_SET_4	ACAGTGGATGCGGCTCGATA
epikol_01493	SUV39H2_SET_5	AATCACGTATCTCTTTGATC
epikol_01494	SUV420H1_SET_1	TCAGTGTCTGTAATCCACA
epikol_01495	SUV420H1_SET_2	GAGAACATGCTACTTAGACA
epikol_01496	SUV420H1_SET_3	AAACTGTGCTCAACTCTGGC
epikol_01497	SUV420H1_SET_4	GCTCTAAGAGACATTGAACC
epikol_01498	SUV420H1_SET_5	TCGAGATACAGCATGTGTGA
epikol_01499	SUV420H2_SET_1	TCATGTACTCAACCCGCAAG
epikol_01500	SUV420H2_SET_2	CAGCTGAGCACTCCGCTTGC
epikol_01501	SUV420H2_SET_3	GCAAGCGGAGTGCTCAGCTG
epikol_01502	SUV420H2_SET_4	GTGACATGCTTCTACGGCGA
epikol_01503	SUV420H2_SET_5	GTCCCGGAGCACCTTCACGC

epikol_01504	TAF1_BROMO_1	CAAAGTGATTGTCAATCCAA
epikol_01505	TAF1_BROMO_2	GGATGATGTAAACCTTATTC
epikol_01506	TAF1_BROMO_3	AACAGACGTTCCACAATCTCC
epikol_01507	TAF1_GECKO_1	AAGAATCACCTGAGTGGCAA
epikol_01508	TAF1_GECKO_2	GATAAAATCTAGAATGATAG
epikol_01509	TET1_OXY_1	GGAGCATACTGCTTATAAAT
epikol_01510	TET1_OXY_2	AACTCTGTAAGTTATCTTCA
epikol_01511	TET1_OXY_3	GTTGCCCGAGAATGTGCGCT
epikol_01512	TET1_OXY_4	GCGGTTATCTTCTCGAGTTA
epikol_01513	TET1_OXY_5	GCTTATAAAGAGGTAGCACA
epikol_01514	TET2_OXY_1	CCAAGGAAGTTTAAGCTGCT
epikol_01515	TET2_OXY_2	TTGGTGCCATAAGAGTGGAC
epikol_01516	TET2_OXY_3	GCAAAACCTGTCCACTCTTA
epikol_01517	TET2_OXY_4	AGAGCACCAGAGTGCCGTCT
epikol_01518	TET2_OXY_5	TTCAGACCCAGACGGCACTC
epikol_01519	TET3_OXY_1	CTCGCAAGTTCCGCTCGCA
epikol_01520	TET3_OXY_2	TCAACGGCTGCAAGTATGCT
epikol_01521	TET3_OXY_3	GCCGTTGAAGTACATGCTCC
epikol_01522	TET3_OXY_4	CGCTCCCCTGTACAAGCGAC
epikol_01523	TET3_OXY_5	GTACAGGGGAGCGACTTCGG
epikol_01534	TRIM24_BROMO_1	TTACTGCCATGAAATGAGCC
epikol_01535	TRIM24_BROMO_2	GAACAGGGTCTTGAAAAGCC
epikol_01536	TRIM24_BROMO_3	AAATCAGCTACAAAATCTTC
epikol_01537	TRIM24_BROMO_4	GCATTGGCTACTTCTGAATC
epikol_01538	TRIM24_BROMO_5	GATTCAGAAGTAGCCAATGC
epikol_01539	TRIM33_BROMO_1	TACTTAATTCATGGCAATAG
epikol_01540	TRIM33_BROMO_2	CCACAAAGTCATCCGGGATT
epikol_01541	TRIM33_BROMO_3	CATCCGGGATTGAGTAGTGT
epikol_01542	TRIM33_BROMO_4	CTTGAAGATCAAACGGACAT
epikol_01543	TRIM33_BROMO_5	TGGACCTGTCAATCATCCGG
epikol_01544	TRIM66_BROMO_1	GGGCTGACAGGTTTCATGGAA
epikol_01545	TRIM66_BROMO_2	GCTTCCTCCGGATGATTGAC
epikol_01546	TRIM66_BROMO_3	CGGATGATTGACAGGTCCAT
epikol_01547	TRIM66_BROMO_4	CCATGGGCCTCTTGATAATC
epikol_01548	TRIM66_BROMO_5	GCCTCTGCAACCTCGGAGTC
epikol_01549	TTF2_DEXD_1	GGATGAGCGCAATCATTGTC
epikol_01550	TTF2_DEXD_2	AGAAAAGCACAGCTTTGACG
epikol_01551	TTF2_DEXD_3	CTGGCACGTGAATCCCGGTT
epikol_01552	TTF2_DEXD_4	ATCTCTACCATGGGCCAAAC
epikol_01553	TTF2_DEXD_5	GATCACTACCTATAGCCTCG
epikol_01554	TTF2_HELIC_1	ATCTCTGTTGGCAGAATTGG
epikol_01555	TTF2_HELIC_2	CTGACTTATGCCACCATCGA

epikol_01556	TTF2_HELIC_3	GGGATTGACAGAGCCATCGA
epikol_01557	TTF2_HELIC_4	GAGGCATTTAACCCTCCAG
epikol_01558	TTF2_HELIC_5	GGTGTGGTCTAAACCTGAC
epikol_01559	UBE2A_UBC_1	ACCATTATGTTGTTCTCGGA
epikol_01560	UBE2A_UBC_2	CATCATAGGTTGGACTCCAA
epikol_01561	UBE2A_UBC_3	TCCGTCCGAGAACAACATAA
epikol_01562	UBE2A_UBC_4	TCCTCCAGCCGGAGTCAGCG
epikol_01563	UBE2A_UBC_5	ATGGAAGACACATCATAGGT
epikol_01564	UBE2B_UBC_1	GGTGCGCCACTGACACCCAC
epikol_01565	UBE2B_UBC_2	TTACAAGAGGACCCACCTGT
epikol_01566	UBE2B_UBC_3	GATAAAAACCTAACAGTTGG
epikol_01567	UBE2B_UBC_4	TTGGACTCCATCGATTCTGA
epikol_01568	UBE2B_UBC_5	TTGGCTGGACTGTTAGGATT
epikol_01569	UBE2E1_UBC_1	GCCTCCAGGATCCGTGTATG
epikol_01570	UBE2E1_UBC_2	CCACCCTCATACACGGATCC
epikol_01571	UBE2E1_UBC_3	ACTTTAGAAATGGTTAGTGC
epikol_01572	UBE2E1_UBC_4	CAGTCAAGGTGTTATTTGCT
epikol_01573	UBE2E1_UBC_5	CTTCCCACCAAGGGGTCGGC
epikol_01574	UBE2I_UBC_1	TCGCCCAGGAGAGGAAAGCA
epikol_01575	UBE2I_UBC_2	AACTGGGAGTGCGCCATTCC
epikol_01576	UBE2I_UBC_3	ATGAGGTTTCATCGTGCCATC
epikol_01577	UBE2I_UBC_4	GCACGATGAACCTCATGAAC
epikol_01578	UBE2I_UBC_5	ACCCGAATGTGTACCCTTCG
epikol_01579	UBR2_GECKO_1	ACTTAACACCTCTGAAATTG
epikol_01580	UBR2_GECKO_2	TGCATAACTTGAACCTTGAG
epikol_01581	UBR2_GECKO_3	CTCTTCTTCCTCAATTTAG
epikol_01582	UBR2_GECKO_4	TACACCATCTCTAAATCTGC
epikol_01583	UBR2_GECKO_5	ATTGTCGCACATCAGAATTT
epikol_01584	UHRF1_1	TCCCGTCCATGGTCCGAACC
epikol_01585	UHRF1_2	GGACAGCGAGTCCACCGTGT
epikol_01586	UHRF1_3	GTGGATCCAGGTTCCGACCA
epikol_01587	UHRF1_4	TGGACAGCGAGTCCACCGTG
epikol_01588	UHRF1_5	GGCGGACCTCGTAGTCGAAG
epikol_01589	USP22_PC19_1	CTTCATGAACTGCATCGTGC
epikol_01590	USP22_PC19_2	AGTCCCGCAGAAGTGGCGTG
epikol_01591	USP22_PC19_3	GCCTGCTAGGTGCCTCGCGT
epikol_01592	USP22_PC19_4	ACCTGGTGTGGACCCACGCG
epikol_01593	USP22_PC19_5	GCAACCCGCCTGTGAAGATC
epikol_01594	USP22_UHYD_1	CTTCATGAACTGCATCGTGC
epikol_01595	USP22_UHYD_2	CTGCGTGGGCTGATCAACCT
epikol_01596	USP22_UHYD_3	GCCTGCTAGGTGCCTCGCGT
epikol_01597	USP22_UHYD_4	ACCTGGTGTGGACCCACGCG

epikol_01598	USP22_UHYD_5	GCAACCCGCCTGTGAAGATC
epikol_01599	USP44_PC19_1	TTGTTGGGCGTAACCACGAA
epikol_01600	USP44_PC19_2	TTACGCCCAACAAGACGCTC
epikol_01601	USP44_PC19_3	GCTTGATCTGAACCAATGGC
epikol_01602	USP44_PC19_4	CCACTGCATTGATACCTTTC
epikol_01603	USP44_PC19_5	TGGGCTTCTGTGAGTACAAC
epikol_01604	USP44_UHYD_1	TTGTTGGGCGTAACCACGAA
epikol_01605	USP44_UHYD_2	TTACGCCCAACAAGACGCTC
epikol_01606	USP44_UHYD_3	AAATTCCTGAGCGTCTTGTT
epikol_01607	USP44_UHYD_4	CCACTGCATTGATACCTTTC
epikol_01608	USP44_UHYD_5	TATCTATGACTTGTCCGCGG
epikol_01614	UTY_JMJC_1	GTTGTACCTGAAGATTATTG
epikol_01615	UTY_JMJC_2	TGTTGTACCTGAAGATTATT
epikol_01616	UTY_JMJC_3	GTGCATTGGGTCAAGCTGT
epikol_01617	UTY_JMJC_4	TTGCTTCATAAAGATCTCA
epikol_01618	UTY_JMJC_5	TTGTTGTACCTGAAGATTAT
epikol_01619	WDR5_WD40_1	AAATTTGGGGCGCGTATGAT
epikol_01620	WDR5_WD40_2	ATCTGACGACCAGGCTACAT
epikol_01621	WDR5_WD40_3	CAAGACTTTGCCAGCTCACT
epikol_01622	WDR5_WD40_4	ATAGCTACTTGAACTATCA
epikol_01623	WDR5_WD40_5	CAGGATGTATTTGCCGTTCTG
epikol_01624	WDR82_WD40_1	TTTCAGCCCCAACGGCGAGA
epikol_01625	WDR82_WD40_2	ACGAATGAAGCTGCCGTTGG
epikol_01626	WDR82_WD40_3	CGTCTGATTGATGCATTCAA
epikol_01627	WDR82_WD40_4	CGAGATGACCGTCTCGCCGT
epikol_01628	WDR82_WD40_5	CTGCATGAGTGATCTGATG
epikol_00954	WHSC1_PHD_1	AGGTGGAAAGCTCCGCAGCA
epikol_00955	WHSC1_PHD_2	CTGCAGGTGAACCTCCCTTC
epikol_00956	WHSC1_PHD_3	GGCTTTCCCGGAGGCCAGAA
epikol_00957	WHSC1_PHD_4	CCTGCCTTGGGCTTTCCCGG
epikol_00958	WHSC1_PHD_5	GGGCTTTCCCGGAGGCCAGA
epikol_00959	WHSC1_SET_1	TGATGTCCCTCTTGCGACC
epikol_00960	WHSC1_SET_2	GAATTTGTTAACGAGTACGT
epikol_00961	WHSC1_SET_3	ATTTGTTAACGAGTACGTTG
epikol_00962	WHSC1_SET_4	AATTTGTTAACGAGTACGTT
epikol_00963	WHSC1_SET_5	TGGGGCCAGCGTCTATTATA
epikol_00964	WHSC1L1_PHD_1	ATGCAGATGAACCTTGCTATC
epikol_00965	WHSC1L1_PHD_2	ATGCCAATCCCAGGCACTCC
epikol_00966	WHSC1L1_PHD_3	CAGGAAGTGATGCCAATCCC
epikol_00967	WHSC1L1_PHD_4	GTGCTGCAAACACTTTCACC
epikol_00968	WHSC1L1_PHD_5	GACTCTCTGATTCCTTGTA
epikol_00969	WHSC1L1_SET_1	CAAAACGGAGCGGAGAGGCT

epikol_00970	WHSC1L1_SET_2	ATTAGTTACTGTTCTCGT
epikol_00971	WHSC1L1_SET_3	ATAATTGATGCCGGCCCAAA
epikol_00972	WHSC1L1_SET_4	TAATTATAACCTAGATTGTC
epikol_00973	WHSC1L1_SET_5	TTGGGCCGGCATCAATTATA
epikol_01629	ZMYND11_BROMO_1	ATTCATTGTCTCCCGCATGA
epikol_01630	ZMYND11_BROMO_2	ACAATAAACACCCGATGTAC
epikol_01631	ZMYND11_BROMO_3	ACACCCGATGTACAGGAGGC
epikol_01632	ZMYND11_BROMO_4	CGGTATTGTGAGAAGCAAT
epikol_01633	ZMYND11_BROMO_5	GTGAGCAAGCTGACATTGCG
epikol_01634	ZMYND8_BROMO_1	AATGGCAAACCTTGAGCAGGT
epikol_01635	ZMYND8_BROMO_2	ATTCCAGAAGCCCGTTCCAT
epikol_01636	ZMYND8_BROMO_3	GGGTGCTGTTCCAATGGAAC
epikol_01637	ZMYND8_BROMO_4	AAATCCACTTTGCATCAGCC
epikol_01638	ZMYND8_BROMO_5	TGGCTGCACAGAAGCCTTCC
epikol_01639	NonTargetingControlGuideForHuman_0001	ACGGAGGCTAAGCGTCGCAA
epikol_01640	NonTargetingControlGuideForHuman_0002	CGCTTCCGCGGCCCGTTCAA
epikol_01641	NonTargetingControlGuideForHuman_0003	ATCGTTTCCGCTTAACGGCG
epikol_01642	NonTargetingControlGuideForHuman_0004	GTAGGCGCGCCGCTCTCTAC
epikol_01643	NonTargetingControlGuideForHuman_0005	CCATATCGGGGCGAGACATG
epikol_01644	NonTargetingControlGuideForHuman_0006	TACTAACGCCGCTCCTACAG
epikol_01645	NonTargetingControlGuideForHuman_0007	TGAGGATCATGTGAGCGCC
epikol_01646	NonTargetingControlGuideForHuman_0008	GGGCCCCGCATAGGATATCGC
epikol_01647	NonTargetingControlGuideForHuman_0009	TAGACAACCGCGGAGAATGC
epikol_01648	NonTargetingControlGuideForHuman_0010	ACGGGCGGCTATCGCTGACT
epikol_01649	NonTargetingControlGuideForHuman_0011	CGCGGAAATTTTACCGACGA
epikol_01650	NonTargetingControlGuideForHuman_0012	CTTACAATCGTCGGTCCAAT
epikol_01651	NonTargetingControlGuideForHuman_0013	GCGTGCGTCCCGGGTTACCC
epikol_01652	NonTargetingControlGuideForHuman_0014	CGGAGTAACAAGCGGACGGA
epikol_01653	NonTargetingControlGuideForHuman_0015	CGAGTGTTATACGCACCGTT
epikol_01654	NonTargetingControlGuideForHuman_0016	CGACTAACCGGAACTTTTT
epikol_01655	NonTargetingControlGuideForHuman_0017	CAACGGGTTCTCCCGGCTAC
epikol_01656	NonTargetingControlGuideForHuman_0018	CAGGAGTCGCCGATACGCGT
epikol_01657	NonTargetingControlGuideForHuman_0019	TTCACGTCGTCTCGCGACCA
epikol_01658	NonTargetingControlGuideForHuman_0020	GTGTCCGATTCCGCCGCTTA
epikol_01659	NonTargetingControlGuideForHuman_0021	CACGAACTCACACCGCGCGA
epikol_01660	NonTargetingControlGuideForHuman_0022	CGCTAGTACGCTCCTCTATA
epikol_01661	NonTargetingControlGuideForHuman_0023	TCGCGCTTGGGTTATACGCT
epikol_01662	NonTargetingControlGuideForHuman_0024	CTATCTCGAGTGGAATGCG
epikol_01663	NonTargetingControlGuideForHuman_0025	AATCGACTCGAACTTCGTGT
epikol_01664	NonTargetingControlGuideForHuman_0026	CCCGATGGACTATACCGAAC
epikol_01665	NonTargetingControlGuideForHuman_0027	ACGTTCCGAGTACGACCAGCT
epikol_01666	NonTargetingControlGuideForHuman_0028	CGCGACGACTCAACCTAGTC

epikol_01667	NonTargetingControlGuideForHuman_0029	GGTCACCGATCGAGAGCTAG
epikol_01668	NonTargetingControlGuideForHuman_0030	CTCAACCGACCGTATGGTCA
epikol_01669	NonTargetingControlGuideForHuman_0031	CGTATTCGACTCTCAACGCG
epikol_01670	NonTargetingControlGuideForHuman_0032	CTAGCCGCCCAGATCGAGCC
epikol_01671	NonTargetingControlGuideForHuman_0033	GAATCGACCGACACTAATGT
epikol_01672	NonTargetingControlGuideForHuman_0034	ACTTCAGTTCGGCGTAGTCA
epikol_01673	NonTargetingControlGuideForHuman_0035	GTGCGATGTCGCTTCAACGT
epikol_01674	NonTargetingControlGuideForHuman_0036	CGCCTAATTTCGGATCAAT
epikol_01675	NonTargetingControlGuideForHuman_0037	CGTGCCCGGAACCGTCATAG
epikol_01676	NonTargetingControlGuideForHuman_0038	ACCCTCCGAATCGTAACGGA
epikol_01677	NonTargetingControlGuideForHuman_0039	AAACGGTACGACAGCGTGTG
epikol_01678	NonTargetingControlGuideForHuman_0040	ACATAGTCGACGGCTCGATT
epikol_01679	NonTargetingControlGuideForHuman_0041	GATGGCGCTTCAGTCGTCGG
epikol_01680	NonTargetingControlGuideForHuman_0042	ATAATCCGGAACGCTCGAC
epikol_01681	NonTargetingControlGuideForHuman_0043	CGCCGGGCTGACAATTAACG
epikol_01682	NonTargetingControlGuideForHuman_0044	CGTCGCCATATGCCGGTGGC
epikol_01683	NonTargetingControlGuideForHuman_0045	CGGGCCTATAACACCATCGA
epikol_01684	NonTargetingControlGuideForHuman_0046	CGCCGTTCCGAGATACTTGA
epikol_01685	NonTargetingControlGuideForHuman_0047	CGGGACGTCGCGAAAATGTA
epikol_01686	NonTargetingControlGuideForHuman_0048	TCGGCATAACGGACACACGC
epikol_01687	NonTargetingControlGuideForHuman_0049	AGCTCCATCGCCGCGATAAT
epikol_01688	NonTargetingControlGuideForHuman_0050	ATCGTATCATCAGCTAGCGC
epikol_01689	NonTargetingControlGuideForHuman_0051	TCGATCGAGGTTGCATTCCG
epikol_01690	NonTargetingControlGuideForHuman_0052	CTCGACAGTTCGTCCCGAGC
epikol_01691	NonTargetingControlGuideForHuman_0053	CGGTAGTATTAATCGCTGAC
epikol_01692	NonTargetingControlGuideForHuman_0054	TGAACGCGTGTTTCCTTGCA
epikol_01693	NonTargetingControlGuideForHuman_0055	CGACGCTAGGTAACGTAGAG
epikol_01694	NonTargetingControlGuideForHuman_0056	CATTGTTGAGCGGGCGCGCT
epikol_01695	NonTargetingControlGuideForHuman_0057	CCGCTATTGAAACCGCCAC
epikol_01696	NonTargetingControlGuideForHuman_0058	AGACACGTCACCGGTCAAAA
epikol_01697	NonTargetingControlGuideForHuman_0059	TTTACGATCTAGCGGCGTAG
epikol_01698	NonTargetingControlGuideForHuman_0060	TTCGCACGATTGCACCTTGG
epikol_01699	NonTargetingControlGuideForHuman_0061	GGTTAGAGACTAGGCGCGCG
epikol_01700	NonTargetingControlGuideForHuman_0062	CCTCCGTGCTAACGCGGACG
epikol_01701	NonTargetingControlGuideForHuman_0063	TTATCGCGTAGTGCTGACGT
epikol_01702	NonTargetingControlGuideForHuman_0064	TACGCTTGCGTTTAGCGTCC
epikol_01703	NonTargetingControlGuideForHuman_0065	CGCGGCCACGCGTCATCGC
epikol_01704	NonTargetingControlGuideForHuman_0066	AGCTCGCCATGTCGGTTCTC
epikol_01705	NonTargetingControlGuideForHuman_0067	AACTAGCCCGAGCAGCTTCG
epikol_01706	NonTargetingControlGuideForHuman_0068	CGCAAGGTGTCGGTAACCTT
epikol_01707	NonTargetingControlGuideForHuman_0069	CTTCGACGCCATCGTGCTCA
epikol_01708	NonTargetingControlGuideForHuman_0070	TCCTGGATACCGCGTGGTTA

epikol_01709	NonTargetingControlGuideForHuman_0071	ATAGCCGCCGCTCATTACTT
epikol_01710	NonTargetingControlGuideForHuman_0072	GTCGTCCGGGATTACAAAAT
epikol_01711	NonTargetingControlGuideForHuman_0073	TAATGCTGCACACGCCGAAT
epikol_01712	NonTargetingControlGuideForHuman_0074	TATCGCTTCCGATTAGTCCG
epikol_01713	NonTargetingControlGuideForHuman_0075	GTACCATACCGCGTACCCTT
epikol_01714	NonTargetingControlGuideForHuman_0076	TAAGATCCGCGGGTGGAAC
epikol_01715	NonTargetingControlGuideForHuman_0077	GTAGACGTCGTGAGCTTCAC
epikol_01716	NonTargetingControlGuideForHuman_0078	TCGCGGACATAGGCTCTAA
epikol_01717	NonTargetingControlGuideForHuman_0079	AGCGCAGATAGCGCGTATCA
epikol_01718	NonTargetingControlGuideForHuman_0080	GTTTCGCTTCGTAACGAGGAA

APPENDIX 8. Full Enrichment Scores for an average of two independent TRA-1-60 screens.

Genes	Initiation ES(Day 7/Day-1)	Maturation ES(Day14/Day7)	Maintenance ES(Day21/Day14)	Overall ES(Day21/Day-1)	Proliferation ES(Day21/day-1)	ES(Day 14/Day -1)
ACAT1	-0.695541373	0.445427411	-0.235045746	-0.136337481	0.108502543	-0.082702272
ACTL6B	-0.655265614	-0.25563104	-1.0614716	-1.101858825	-0.282945816	-0.524879985
ACTR	-0.671443983	0.425901168	-1.305852717	-1.826943921	-0.114056218	-0.381252296
ASH1	-0.301730583	0.015470317	-0.169019529	-0.342744333	-0.319611021	-0.140336049
ASH2L	-1.068969898	0.495474578	1.353388104	0.562386114	-0.163529433	-1.007489601
ASXL1	-0.030069655	0.349698541	-0.022613491	-0.183597874	0.271912393	0.141447465
ASXL2	0.295623285	0.015911421	-0.303227418	-0.045941717	-0.785160977	0.031466726
ASXL3	1.27391541	0.467087664	-1.167670919	0.890920443	1.952216702	1.575464327
ATAD2	0.055051547	0.578027403	-1.349149248	0.202594255	0.208976857	0.583549258
ATAD2B	-1.238010996	0.384146577	-0.286400843	-1.625019377	0.40173158	-1.446133841
ATRX	0.374382947	-0.528425997	-0.678787224	-0.507677834	0.701722338	-0.136867737
BAZ1A	-0.122243077	-0.856474902	-0.143537062	-0.334746317	-0.492682035	-0.676428462
BAZ1B	0.529627355	0.68766125	-1.596621449	0.200510394	0.717603664	1.199761414
BAZ2A	-0.624559137	-0.0971814	0.441757981	-0.280716224	-0.564103338	-0.309809638
BAZ2B	-0.229660745	0.682410556	-1.473042589	-0.620224134	-0.152929487	-0.016627809
BMI	-0.084185918	0.018636737	0.283561293	-0.344657812	0.054602361	0.133061087
BPTF	-0.248197017	0.205625895	0.560524505	0.538749582	-0.849149125	-0.131886378
BRD1	-0.876565127	0.779100952	-0.444202222	-0.498671323	-1.036707837	-0.510532637
BRD2	0.038215941	2.142924397	1.13912383	1.748930207	1.388910102	1.133469115
BRD3	0.161593838	-1.299949848	0.565952506	-0.108080441	-0.213726726	-0.491985675
BRD4	-0.796266003	0.463419156	0.369778228	-0.75927776	-0.458988092	-0.512383277
BRD7	-1.341090428	0.033693825	1.756056813	0.045127344	0.013106555	-0.450582446
BRD8	-0.213030322	-0.030097997	0.263624455	-0.090886348	-0.628619697	-0.1544591
BRD9	0.08496646	-0.890361204	-0.478778379	-0.577297379	0.029301417	-0.345449432
BRDT	-1.000884558	0.227379912	-0.146850952	-0.405590359	-1.059566157	-0.382482031
BRPF1	-1.631341637	-0.170530671	-0.693624902	-1.461564558	-1.19278987	-1.296314933
BRPF3	0.283833961	-0.225457875	-0.239887883	-0.151068019	0.386634214	0.353915274
BRWD1	-0.313268616	-0.684465887	0.165063617	-0.072520407	0.328576855	-0.761455583

BRWD3	-0.153772375	2.153647523	-0.413726008	0.445070764	-0.757261458	1.547496488
CARM1	-0.120269299	0.531230176	0.368058506	0.860763879	1.161279504	-0.051695062
CBX1	-1.212478697	-0.16707999	-1.110346943	-0.502854732	-0.114288277	0.125441685
CBX2	-2.253753367	0.049360327	-0.793767117	-1.412656943	-2.313122525	-1.119417737
CBX3	-0.527391877	0.50113543	-0.524031979	0.817402331	0.805920594	0.718290318
CBX4	0.048293252	0.622696154	-1.132808127	-0.251580788	-0.083741804	0.36610854
CBX5	0.556844693	-0.928218367	-0.014835327	0.040047735	-0.542817903	-0.475736886
CBX6	0.281355942	-0.120260068	-0.443796823	-0.585300176	-0.260473643	-0.132298352
CBX7	0.132095125	-1.834922288	-0.12758351	-0.596731118	-1.16954049	-0.437679237
CBX8	-0.405813933	0.026170934	-1.064950645	-0.987103202	-0.656903022	-0.642425215
CDY1	0.849125434	-0.267149598	-0.372404079	0.11346	-0.28457358	0.421335781
CDYL	-1.573032135	0.569018739	-0.240888613	-1.651869755	-0.667846346	-0.822267779
CHAF1A	-1.94438087	0.850830222	-0.820371625	-2.517814326	-1.512242729	-2.431931611
CHAF1B	-1.646220112	0.403103011	0.381193979	0.027799261	-0.242471327	-0.484727977
CHD1	-1.142547141	-0.162652333	-0.391701862	-0.969925091	-0.567263638	-0.770131551
CHD1L	-0.323821105	-0.450313017	-0.658414967	-0.568862639	0.187606825	-0.325268585
CHD2	-0.404348411	-0.074717643	0.743776838	0.200133098	-0.300321928	-0.597638218
CHD3	-0.524050578	0.128605395	-0.517938827	-0.155561378	-0.309470934	-0.12720517
CHD4	-0.420081181	0.258567799	0.3815201	-0.797161741	1.238955141	-0.463368548
CHD5	-0.535404187	-0.304864173	-0.18908875	-0.166882056	-0.615590456	-0.943743831
CHD6	-0.194527218	-0.32014452	0.045866803	-0.228495694	0.340718388	-0.557702513
CHD7	0.441570617	-0.661663657	0.629973199	0.245110585	0.470951685	-0.184249471
CHD8	-0.799125169	0.239512216	-0.212503851	-0.50708908	-0.238322551	-0.78177098
CHD9	-0.181590437	0.197681933	-0.170807575	-0.162283102	0.794084246	0.2801703
CLOCK	-0.81513411	-0.865881279	-0.954166447	-1.387275083	-0.79814814	-1.666051509
CREBBP	0.24915172	-0.002125936	0.883541793	1.238254255	-0.39474139	0.077668329
CXXC1	-0.730161936	-0.150703619	0.343025802	-1.366606466	-0.879789362	-1.682528626
DMAP1	-0.714722675	3.182679329	0.202979231	0.120151112	0.197422297	0.542357662
DNMT1	-0.936756055	-0.958321604	-0.432244412	-0.370082134	-0.03767125	0.123578184
DNMT3A	1.31661573	1.26871582	2.070098979	2.681199036	2.205925569	2.72013731
DNMT3B	-0.8415751	0.788105622	0.012181311	-1.177185911	-1.210538999	-0.004086476
DNMT3L	1.06099976	-0.877642566	0.820283798	0.889895427	0.046400213	0.049313016
DOT1L	0.703233237	-0.629862036	-0.573538345	0.033426414	0.614927953	0.42459567
DTX3L	0.374791693	-0.105519256	-0.139747049	0.222936443	0.573280997	0.240766492
EHMT1	-1.526266892	-0.386575426	0.852474728	-1.075151997	-0.740376229	-1.073629981
EHMT2	-0.827680348	0.090687598	-0.16466444	-0.760755655	-0.621144777	-0.832340019
ELP3	-0.254566867	-0.667204546	1.05032213	-0.700524783	-0.78979952	-0.687000063
EP300	-0.837514567	0.466239056	1.024307763	0.211134637	-0.367680834	-0.261308844
EP400	-1.435678164	-0.125705176	-0.094851859	-0.985628999	-2.194833943	-1.469806183
ERCC6	0.450358978	0.078863214	-1.096331332	-0.098590634	0.397399119	0.440666939
EZH1	-0.507590318	-1.158752969	0.812028035	-0.883147516	-0.946954909	-1.936696111
EZH2	-0.827505317	0.43953109	-0.49842664	-0.12400985	-1.385988184	-0.590801601
KAT2A	-0.317268229	-0.205152542	0.125704526	-0.077102479	-0.220609698	-0.401419456
GTF3C4	0.430246746	0.551700367	0.807343872	0.684446424	-0.232156639	0.204553897
H2AFY2	-0.075945371	0.88007825	-0.552092111	1.231496803	0.490071955	0.824708367

Appendix

H2AFZ	-0.477059969	0.36386334	-0.064106404	-0.298739924	-0.207932605	-0.519282244
HAT1	-0.286538567	-0.367424933	-0.238877207	-0.660814977	-0.033887618	-0.555793066
KAT7	-0.740730952	0.128496081	0.41157733	-0.054748334	-1.130868314	-0.967811423
HCFC1	0.299310452	0.617218918	0.114954553	-0.994471551	-0.346799941	-0.416042501
HDAC1	-0.258936304	0.958851325	0.087198984	0.383842377	0.020455297	0.536335788
HDAC10	-0.220391254	1.478348484	0.745065696	1.275645316	0.378521174	0.458301707
HDAC11	-0.661570925	0.919468074	0.168372671	0.68558062	-0.049858676	0.431028054
HDAC2	-0.664389426	0.769759099	0.136705658	0.110896748	0.400705858	0.119260693
HDAC3	-0.629321285	0.064649806	0.254438441	-1.14747217	-1.271871391	-0.134206383
HDAC4	0.306866669	-0.767155678	-1.073711566	-0.980459386	0.551439937	-0.463623912
HDAC5	-0.990754424	0.059332139	-0.31720239	-0.909173971	-1.5683776	-0.470433903
INO80	-0.128871399	-0.300663402	1.565374535	-0.454418664	-1.097347547	0.6276898
KDM5A	-0.311852402	0.525919337	-0.367184537	-0.678264988	-0.529381819	-0.503887059
KDM5B	-0.945719368	-0.164927359	-0.56898297	-0.987655839	-1.162566509	-0.665434841
KDM5C	0.857862996	0.912788792	0.138755113	0.224760697	-0.026676624	1.216319201
KDM5D	-0.064764269	0.225234379	-0.063292869	0.23780446	0.000225101	0.114650086
JARID2	-0.289747718	0.979159912	-0.742816127	-0.111741506	-0.898998536	-0.280740712
KDM7A	-0.193846525	1.024006173	-0.722643196	0.243331618	-0.903973731	0.186688001
KDM3A	0.095432067	-0.316085539	-0.344099309	-0.215496117	-1.209493046	-0.54585185
KDM3B	-0.49408507	0.032972475	0.025666947	-0.539800004	-0.0420853	-0.661061817
JMJD1C	-1.535085127	0.373432227	-0.523316334	-0.603091092	-0.926711241	-1.4269171
KDM4A	-0.449549798	0.817635096	0.327634975	-0.351455325	-0.078459546	0.250991974
KDM4B	-0.513873719	0.285245521	-1.047524683	-1.107708486	-0.61252109	-0.81725444
KDM4C	-0.311620954	1.019053329	0.22382617	0.052547487	0.002885595	0.236648431
KDM4D	0.376669526	-0.534859072	1.040555328	0.399744495	-0.217506333	-0.808727994
KDM6B	-0.005920467	1.047552475	-0.534468582	1.065007178	0.018473233	0.819401832
JMJD4	-0.221768067	-0.936734513	-1.726277691	-1.510535861	-1.370006385	-0.582747667
KDM8	-1.150015082	-1.058911027	-1.850010934	-2.23799823	0.358429559	-1.438513059
JMJD6	-1.230665459	-0.186493521	0.002497762	-0.631137038	-1.147734198	-1.094655577
JMJD7	-0.36405472	0.037338034	-0.917145324	-0.616613488	-1.032687915	-0.721916791
JMJD8	-0.992795138	-0.023233534	-0.381809444	-1.096072341	-0.56227076	-0.811559169
KAT2B	-0.231432939	0.29237468	0.221387896	0.276369536	0.11851576	-0.502199623
KAT5	-1.512251373	0.097785588	0.164924104	-1.409726227	-0.924538901	-1.487783161
L3MBTL1	-1.387760012	0.195402729	0.612348848	-0.785726512	-0.678542925	-1.273392753
KDM1B	-0.158249898	0.049973886	0.470170921	0.48786262	0.3311969	-0.033414508
MBD1	-0.280025391	0.349606836	-0.658300745	-0.65447362	0.030951734	-0.618750332
MBD2	-0.499623248	0.447347535	-0.086312991	0.421498739	0.703318505	-0.288669074
MBD3	-0.623830096	1.072839775	-0.767473205	-0.519057002	-0.168294979	0.263032494
MBD4	0.683676215	-1.052022061	0.727927467	0.887284532	0.714557839	0.327367477
MECP2	-0.186215528	0.348250965	0.862796771	-0.234473417	-0.109557794	0.118880982
MGEA5	-1.469199943	-0.417324176	-1.440382397	-0.946337521	-0.334974842	-0.803934952
MINA	-0.617205415	-0.988798893	0.054943423	-0.657528438	-1.678455241	-1.039407164
KMT2A	-0.224571838	0.697689979	-0.180763951	0.9163992	0.363600542	0.679598694
KMT2D	-0.449730467	0.071947824	-0.314267528	-0.177574853	0.724389272	-0.765673055
KMT2C	-0.058204663	0.51973705	0.737463078	0.746104583	0.092598964	0.214404724

KMT2B	0.43628433	-0.413228609	0.219642748	0.336705021	1.412786043	0.013266021
KMT2E	-0.112351566	0.952116655	0.628541831	0.4733307	-0.709637377	0.267468656
MMP8	-0.600427726	0.858212918	-0.568370652	-0.287979302	-0.589409143	-0.020006446
MORF4L1	0.365606157	0.543512664	-0.139268456	0.131851881	-0.364650513	0.087485967
MSL3	-1.084512125	0.30980183	-0.225719698	-0.903443689	0.009606862	-1.060950005
MTA1	0.406006211	0.19115626	-0.232210602	0.579763551	0.527016079	0.576009753
MTA2	-0.787691272	-0.19456902	0.218694263	-0.485859446	-0.420370856	-1.01763825
MTA3	-0.350474137	0.081826868	-0.896577109	-1.394338837	-0.10638419	-0.030238814
KAT8	-0.740116437	0.316455044	-0.108592626	-0.429308415	0.045276183	-0.417051958
KAT6A	-0.118490059	0.034316627	0.125246092	0.020252225	-0.511205778	0.078310378
KAT6B	-0.148789684	-0.202651555	-0.700816063	-1.270122288	-0.615078858	-0.440988005
NAP1L1	-0.156086509	1.454851045	-0.207128817	0.559119122	0.248773131	0.989002868
NAP1L2	-0.831423815	0.69085779	-0.858082601	-0.410872534	-1.353227479	-0.240426052
NAP1L3	-0.161184864	0.831221141	-0.999975144	-1.237605074	-0.680785756	0.824430562
NAT10	-0.720898214	0.459448185	-0.290566638	-0.475281041	-0.002236976	-0.638249898
NCOA1	-0.235366171	0.404792869	0.036976575	-0.141227375	-0.038146277	0.120977151
NCOA2	-0.208646935	0.392698445	0.273434483	-0.138145577	-0.199659727	-0.319634442
NCOA3	0.023042405	0.406945376	-0.526752089	-0.116299007	0.138610596	0.486846205
C14ORF169	-0.977883175	0.39369902	-0.085776554	-0.275733091	0.015030615	-0.675174276
NSD1	-1.764413104	1.060098006	-0.26203807	-0.61533412	0.343881699	-0.606684842
NSD2	0.434736666	0.320116496	0.179663415	0.59301931	1.551269928	0.43699391
NSD3	-0.853546868	0.707018663	0.594532727	-0.231691305	-0.434295808	-0.478699399
PADI4	-0.457438458	0.526367192	-0.235103035	-0.600212331	-0.148267497	-1.726297777
PARP1	0.044739284	-0.517201281	-0.371160603	-0.297824418	-0.812095503	-0.409975217
PAXIP1	1.409915871	-0.608007322	0.844985906	2.440449099	1.919897348	1.465273867
PBRM1	0.091627208	0.128429849	-0.178122526	0.078478594	0.972525355	0.317844306
PCGF2	-0.371142641	-0.363806416	0.469169508	-0.364490836	-0.836540937	-0.738360297
PHF1	-1.441825123	0.317083279	0.417616102	-0.545089512	-0.349302588	-1.419260917
PHF2	0.694228105	0.718897181	-0.150509361	-0.075590898	0.655853706	0.518794389
PHF8	0.02225551	-0.341534459	-0.752649451	0.439519792	0.250052918	0.303177335
PPHLN1	0.540860681	1.2276091	0.449253639	0.415826188	0.601025128	0.581109778
PRDM1	-0.105425369	-0.050869769	0.834285629	0.25326575	0.751075721	-0.123534629
PRDM10	0.115074849	-0.74385527	-0.270038605	-0.903768229	0.623915159	-0.708716453
PRDM11	-0.666397824	1.100628166	1.570534468	1.088486022	1.599530881	-0.077993692
PRDM12	1.362208663	0.762894593	-0.087191825	-0.069087849	0.629361004	0.598747689
PRDM13	-0.469350968	0.478118372	0.41376229	-0.543884064	-1.123887939	-0.496549967
PRDM14	0.547621909	-0.032642486	-0.444162915	-0.476585317	0.505030191	0.385711887
PRDM15	-0.080302871	-1.672210277	0.60117745	0.169334946	0.710500259	-0.39326692
PRDM16	-0.293249567	0.198833203	-0.797121116	-0.416774861	0.44281273	0.063427495
PRDM2	-2.004649653	0.49196925	0.577333002	-1.142564698	-0.820912486	-1.117160288
PRDM3	-0.882490755	0.8770613	0.284855903	0.976492016	-1.033946425	-0.072598453
PRDM4	0.695825988	-1.110845495	-0.740478201	-1.154948082	-1.122081691	-0.483398585
PRDM5	-0.743587705	-0.28912266	-1.052814949	-0.68889307	0.216986142	-0.904559958
PRDM6	-0.596227218	0.173315089	0.196748502	-0.169232031	0.374197937	0.498453596

Appendix

PRDM8	0.42915343	-0.315924252	-1.338716478	-0.70586625	0.587764478	0.219686846
PRDM9	0.10122072	0.377665606	-1.044474976	-0.424336102	0.374036063	-0.315579571
PRMT1	-0.937861018	-0.950739202	0.420480258	-1.363118934	-0.86793189	-1.377721739
PRMT2	-0.330054798	-0.0403767	-0.357717267	-0.885980056	-0.216173915	0.079351891
PRMT5	-1.53153235	-1.173569744	-0.8618401	-1.986540106	0.47059851	-0.764765438
PRMT6	0.4269097	-0.08268256	-0.046349863	0.265524851	-0.692257366	1.003650608
PRMT7	0.614843269	-0.768239793	-0.153588748	0.316181252	0.211458308	0.369995916
PRMT8	-0.483559738	0.668567612	-0.772927556	-0.842715899	0.90168908	0.353527926
RBBP4	-1.529097	0.919021074	0.683675858	-1.318412675	-0.239093942	-0.852475994
RBBP5	-0.229289468	0.534790391	-0.700557441	-0.473258261	0.091264874	-0.417017289
RBX1	-1.008059489	-1.917794631	0.14026163	-1.687384738	-1.124579817	-2.21773446
RNF168	0.234076262	-1.894127226	1.783836377	0.416065691	1.059053992	-0.693763728
RNF2	0.048357998	0.292031216	0.919216969	0.796675246	-0.657907659	0.231049399
RNF20	0.38972378	-0.295792073	-1.360494699	-0.954286202	-0.428322176	-0.075262892
RNF8	0.469077873	-0.359895964	-0.660209423	-1.550329022	0.349215331	-0.306678794
SATB1	0.417112768	0.21843146	0.020413567	-0.296602572	-0.445184775	0.418594332
SETD1A	-1.57956058	-0.892108618	-0.540802745	-1.654029293	-2.005866194	-2.174674955
SETD2	2.122595155	-0.185928505	0.233828648	0.721670572	3.477893093	1.123187976
SETD3	0.119198827	-0.008064649	0.64905055	0.397110023	0.008675739	0.151739116
SETD4	0.011576159	0.182687335	0.630038693	0.076273229	-1.020907842	-0.163529652
SETD6	-0.259005611	0.535940639	-0.298687915	0.336903139	0.531812118	0.350841643
SETD7	-1.000230766	0.535291131	-1.839768794	-1.08803913	-0.783847521	-0.350074505
SETD8	-0.359786417	-0.043028512	0.497169083	0.399068656	0.41993822	0.795801634
SETDB1	0.658107341	1.041580646	0.418725229	0.743459398	0.364686466	0.270194584
SETDB2	-0.212765065	0.327380827	-1.087723214	-0.02250462	0.381445336	0.264655016
SETMAR	0.560975285	0.000523128	-0.734081315	0.049652116	-0.709257076	0.537127934
SFMBT1	-0.411363929	-0.628476583	1.197165777	0.347672421	-0.099418409	-0.122042087
SFMBT2	-1.187881512	1.138911485	-0.301991895	0.342372326	0.563882562	0.189420143
SHPRH	0.87277479	0.25337109	0.242365687	-0.061153509	-0.096268831	-0.175941587
SIN3A	-0.867957179	-0.620616804	-0.487159573	-1.245820232	0.80030476	-0.965947001
SIN3B	-0.643630153	-0.05384267	0.071967081	-0.259568177	-0.560027276	-0.391128267
SIRT1	0.140503005	0.315353734	-0.041418529	0.520742772	0.050491093	-0.128089458
SIRT2	0.401446289	-0.352667198	-0.819769436	-0.426181426	0.378140848	0.238920748
SIRT3	-0.103861191	1.069260327	1.06170947	0.686500767	0.158590491	0.98714744
SIRT4	0.585943757	-0.008787728	-1.208516585	-0.499906853	0.53929401	0.44276441
SIRT5	-0.433953428	0.714034124	-0.614331762	0.737038623	-0.678490405	0.549246552
SIRT6	0.173542927	1.630057114	-0.995563608	0.785158584	1.769685375	1.261518696
SIRT7	-0.684515518	1.08493085	-0.253940778	-0.418029693	0.541922388	-0.469502527
SMARCA1	0.09572763	-0.076440221	-1.099985176	-0.16670696	0.139380298	0.379790174
SMARCA2	-0.198219461	1.484997698	0.607217309	0.810107313	0.26766577	0.658530982
SMARCA4	0.767222534	-0.004684479	0.247020761	-0.038836146	-0.042448315	0.096929309
SMARCA5	-0.675035577	-0.172138027	-1.016259696	-1.25077006	-0.370360063	-1.160568528
SMARCA D1	-1.63655025	0.232213474	0.441140527	-0.258816167	-0.261227345	-0.595709415
SMARCA L1	0.331312182	-0.649827181	0.268249271	0.552694794	-0.152701353	-0.775009311

Appendix

SMARCC2	0.118582313	-0.60607536	-0.792707182	-0.255125353	0.700091325	0.108862507
SMARCD1	0.080202238	0.186006847	-0.526064239	-0.791164435	-0.410369421	-0.595125838
SMARCD3	-0.050440304	0.645905912	-1.569990572	-1.069977586	-1.625422177	0.050784613
SMARCE1	-0.238453706	-0.993697482	0.353353234	-0.91146788	-0.414582258	-1.484410161
SMYD1	-0.353403592	0.299182829	-0.114837692	-0.200970775	-0.6491139	0.172175581
SMYD2	-0.837109561	0.071687606	0.516405889	-0.154552193	-0.014966129	-0.788964205
SMYD3	-0.01726216	-0.489478365	-2.022579884	-1.835636304	-1.465546097	-0.890013576
SMYD4	0.811935546	-0.320781228	0.113187101	-0.055098115	0.018526115	-0.625663657
SMYD5	-0.447212133	-0.946194343	-0.422741218	-1.655920066	-0.641846939	-1.058681049
SP100	-0.560683045	0.722061338	-0.692099472	-0.039938827	-0.355501376	0.197668054
SRCAP	-0.222080489	-0.90573515	-0.027614903	-0.503940959	-0.258830309	-0.353623387
SSRP1	-0.665654511	-0.137276783	0.209006064	-0.527700576	-1.013737045	-0.873630271
SUPT16H	-0.917360743	-0.247180373	0.259566072	-0.864165319	-0.888429964	-1.592030034
SUV39H1	-0.653790927	1.341639003	-0.69920919	0.102339526	-0.656661491	0.674734649
SUV39H2	0.140450221	0.014602553	-0.237591887	-0.509162904	-0.339668534	-0.102401509
SUV420H1	-0.671267954	-0.879543284	0.909299051	-1.076155705	-0.605274043	-1.208252834
SUV420H2	-0.716126529	0.273688508	0.454988171	0.345101803	-0.514078722	-0.629717431
TAF1	-0.64256588	-0.278476565	-1.274354909	-1.155742878	-0.238811681	-0.923653549
TET1	0.422819208	0.32592424	-0.549554833	0.144357376	-0.282099007	0.639050448
TET2	0.721284124	-0.687017932	-0.568779852	-0.359405106	0.218094783	-0.048238605
TET3	0.660122141	-0.39326788	0.606598271	0.218070154	-0.498017425	-0.006826924
TRIM24	0.060382303	0.141496299	-0.936207704	-0.534242821	1.830018709	-0.207143487
TRIM33	0.533849449	-0.249951244	-1.110200433	0.660879315	-0.318184951	1.106016513
TRIM66	-0.154540705	0.331306005	0.081895964	0.185832751	0.174173161	0.01640817
TTF2	-0.15496845	-1.155680542	1.108155883	-0.561283851	-0.377386417	-0.638967784
UBE2A	-1.506236049	-0.94528862	0.297570031	-1.661780379	-2.161337735	-2.612900148
UBE2B	-0.430236924	0.743480745	0.439733988	0.846807372	0.593425302	0.292793757
UBE2E1	0.050223826	1.935813517	-0.25758279	0.67844888	1.27285895	1.078385246
UBE2I	-0.900304773	0.168965711	0.686319604	0.435443344	0.172880875	-0.177341545
UBR2	0.737598491	0.605954599	-0.339261649	-0.44994278	-0.298275431	0.247043459
UHRF1	-1.95703249	-0.163489379	0.48356229	-1.315621476	-1.820467456	-1.762445285
USP22	0.287987231	1.425309254	0.816655542	1.573962951	0.924044788	1.149120783
USP44	-0.407174895	0.742496227	-0.688758814	0.089179157	0.681053048	0.206443037
KDM6A	-0.448517595	-0.182516846	0.505872386	-0.180674421	-0.689370746	-0.349166056
UTY	-0.401454173	0.029289139	0.441110984	-0.235969214	-0.414563623	-1.043915891
WDR5	-0.675485663	0.988143078	-0.363096448	-0.303452548	0.057280867	-0.302097285
WDR82	-0.512699593	-0.775792294	0.824349028	-1.406948738	-1.554161321	-1.373720509
ZMYND11	-0.456789159	-0.314942686	0.490082788	-1.062236815	-0.395382468	-1.460099605
ZMYND8	0.265419195	-0.723885584	-0.6235283	-0.279287272	0.071454403	0.121353709

APPENDIX 9. Full Enrichment Scores for average of three replicates of low copy exogenous EGFP silencing screen.

Genes	Initiation ES (Day 7/ Day -1)
ACAT1	-0.328020479
ACTL6B	0.066187947
ACTR	-0.271463974
ASH1	0.006029976
ASH2L	-0.242951198
ASXL1	-0.151596874
ASXL2	-0.196017564
ASXL3	-0.21483566
ATAD2	0.291015146
ATAD2B	0.070827609
ATRX	0.015954609
BAZ1A	0.125599981
BAZ1B	-0.043812241
BAZ2A	-0.103705844
BAZ2B	-0.016627803
BMI	-0.408344538
BPTF	0.056645921
BRD1	-0.299914767
BRD2	0.158861634
BRD3	0.084738397
BRD4	-0.019941518
BRD7	0.142950154
BRD8	0.121873666
BRD9	-0.266886221
BRDT	0.237348923
BRPF1	-0.132006438
BRPF3	-0.295589707
BRWD1	0.194414585
BRWD3	0.164493025
CARM1	-0.229181411
CBX1	0.093905962
CBX2	0.135448561
CBX3	-0.40732449
CBX4	-0.461354121
CBX5	-0.490131955
CBX6	-0.343951494
CBX7	-0.054837708

Genes	Initiation ES (Day 7/ Day -1)
CBX8	-0.243579587
CDY1	0.008332206
CDY1L	-1.120878346
CDYL	-0.543129911
CHAF1A	-0.046704267
CHAF1B	-0.323602607
CHD1	-0.155199814
CHD1L	-0.142308514
CHD2	-0.256528194
CHD3	-0.27925128
CHD4	-0.134664374
CHD5	-0.051649321
CHD6	-0.011808527
CHD7	0.08781505
CHD8	-0.304417265
CHD9	0.131315689
CLOCK	0.083061627
CREBBP	0.100363412
CXXC1	-0.350465121
DMAP1	0.034895036
DNMT1	-0.148026458
DNMT3A	0.452398624
DNMT3B	-0.034552335
DNMT3L	0.349979275
DOT1L	-0.209084304
DTX3L	-0.033274418
EED	-0.038779001
EHMT1	-0.063226871
EHMT2	-0.056443353
ELP3	-0.31058202
EP300	-0.149805576
EP400	-0.017956239
ERCC6	0.051869366
EZH1	0.061925969
EZH2	-0.41866742
GCN5	-0.134648425
GTF3C4	-0.454741165

Genes	Initiation ES (Day 7/ Day -1)
H2AFY2	0.037934504
H2AFZ	-0.031300543
HAT1	-0.13634382
HBO1	-0.252944701
HCFC1	-0.457193354
HDAC1	-0.305089293
HDAC10	-0.020152171
HDAC11	-0.137168876
HDAC2	0.228297863
HDAC3	-0.266680604
HDAC4	-0.146505573
HDAC5	-0.353069828
INO80	-0.237800552
JARID1A	-0.14181418
JARID1B	0.030568805
JARID1C	-0.204110358
JARID1D	-0.112425072
JARID2	-0.087819429
JHDM1D	-0.08737639
JMJD1A	0.106885191
JMJD1B	-0.142906293
JMJD1C	-0.219016501
JMJD2A	-0.215583895
JMJD2B	-0.43036131
JMJD2C	-0.071044898
JMJD2D	-0.04834399
JMJD3	-0.250204205
JMJD4	0.097307675
JMJD5	-0.242858626
JMJD6	0.335322743
JMJD7	-0.154004772
JMJD8	-0.158821353
KAT2A	-0.162098714
KAT2B	-0.045604514
KAT5	-0.43294318
L3MBTL1	0.024676029
LSD2	-0.056442416

Genes	Initiation ES (Day 7/ Day -1)
MBD1	0.017851669
MBD2	-0.115155546
MBD3	-0.118337632
MBD4	-0.244310007
MECP2	-0.246643988
MGEA5	-0.103510241
MINA	-0.32262431
MLL1	0.013493956
MLL2	-0.058395315
MLL3	-0.251916686
MLL4	0.155565113
MLL5	-0.108597673
MMP8	-0.072020955
MORF4L1	-0.070140945
MSL3	0.112948362
MTA1	0.122074391
MTA2	-0.043540964
MTA3	-0.168483222
MYST1	-0.056198807
MYST2	0.275043711
MYST3	-0.031373056
MYST4	0.076277845
NAP1L1	-0.250682175
NAP1L2	-0.064757816
NAP1L3	-0.006105596
NAT10	-0.265873655
NCOA1	-0.052570555
NCOA2	0.08245453
NCOA3	-0.0064346
NO66	-0.266810498
NSD1	0.265624497
NSD2	0.292634442
NSD3	0.041935039
PADI4	-0.186248037
PARP1	0.121408519
PAXIP1	0.228873926
PBRM1	0.20190514
PCAF	-0.028849791
PCGF2	-0.106677805
PHF1	-0.261875533

Genes	Initiation ES (Day 7/ Day -1)
PHF2	0.059001155
PHF8	-0.17806745
PPHLN1	-0.280504873
PRDM1	0.287486079
PRDM10	-0.228408627
PRDM11	-0.068796478
PRDM12	-0.098744186
PRDM13	0.042922529
PRDM14	-0.013076458
PRDM15	-0.135510474
PRDM16	0.137531873
PRDM2	-0.321372741
PRDM3	-0.054690811
PRDM4	0.183592877
PRDM5	0.152215021
PRDM6	0.046641195
PRDM8	-0.011072589
PRDM9	-0.013814782
PRMT1	-0.252249406
PRMT2	-0.295617253
PRMT5	-0.370562901
PRMT6	-0.677269585
PRMT7	-0.10858088
PRMT8	-0.188722663
RBBP4	-0.590011771
RBBP5	-0.307870117
RBX1	-0.143483056
RNF168	-0.141122045
RNF2	0.035138284
RNF20	-0.084772237
RNF8	0.032284846
SATB1	0.010729026
SETD1A	-0.200459017
SETD2	0.162014787
SETD3	-0.081098241
SETD4	-0.132154718
SETD6	-0.067935746
SETD7	-0.205493569
SETD8	-0.076078842
SETDB1	-0.167459959

Genes	Initiation ES (Day 7/ Day -1)
SETDB2	-0.006484324
SETMAR	-0.431994061
SFMBT1	0.024927277
SFMBT2	-0.159794706
SHPRH	-0.145696974
SIN3A	0.135564632
SIN3B	-0.269802494
SIRT1	0.025636357
SIRT2	0.237153259
SIRT3	0.034396577
SIRT4	-0.189017716
SIRT5	-0.000705542
SIRT6	0.022801133
SIRT7	0.024575454
SMARCA1	-0.119019614
SMARCA2	-0.001094385
SMARCA4	0.285740047
SMARCA5	-0.251904788
SMARCA1	-0.202655428
SMARCA1	-0.113002702
SMARCC2	0.305084401
SMARCD1	0.003099835
SMARCD3	0.301224304
SMARCE1	-0.226841152
SMYD1	-0.093882125
SMYD2	0.003811897
SMYD3	-0.067834222
SMYD4	0.116530436
SMYD5	-0.036516798
SP100	-0.153499376
SRCAP	-0.219386406
SSRP1	-0.250663956
SUPT16H	-0.303327148
SUV39H1	-0.222910862
SUV39H2	0.10986046
SUV420H1	-0.148959856
SUV420H2	0.110602546
TAF1	-0.163864157
TET1	0.017102348
TET2	-0.048712535

Genes	Initiation ES (Day 7/ Day -1)
TET3	-0.170193039
TIP60	-0.312677808
TRIM24	-0.108683767
TRIM33	0.102467456
TRIM66	-0.089676885
TTF2	-0.255586753
UBE2A	-0.101785634
UBE2B	-0.0803303
UBE2E1	-0.125069563
UBE2I	-0.340822549
UBR2	-0.301550299
UHRF1	-0.084711741
USP22	0.026379742
USP44	-0.099378862
UTX	0.105767615
UTY	-0.170605413
WDR5	-0.070992056
WDR82	-0.270045043
ZMYND11	-0.431955745
ZMYND8	-0.10833249

6 Bibliography

1. Weintraub, H., Tapscott, S.J., Davis, R.L., Thayer, M.J., Adam, M.A., Lassar, A.B., Miller, A.D.: Activation of muscle-specific genes in pigment, nerve, fat, liver, and fibroblast cell lines by forced expression of MyoD. *Proc. Natl. Acad. Sci. U. S. A.* 86, 5434–8 (1989).
2. Takahashi, K., Yamanaka, S.: A decade of transcription factor-mediated reprogramming to pluripotency. *Nat. Rev. Mol. Cell Biol.* 17, 183–93 (2016).
3. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., Yamanaka, S.: Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell.* 131, 861–872 (2007).
4. Takahashi, K., Yamanaka, S.: Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell.* 126, 663–76 (2006).
5. Theunissen, T.W., Jaenisch, R.: Molecular control of induced pluripotency. *Cell Stem Cell.* 14, 720–734 (2014).
6. 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.: Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science (80-.).* 341, 651–654 (2013).
7. Romito, A., Cobellis, G.: Pluripotent stem cells: Current understanding and future directions. *Stem Cells Int.* 2016, (2016).
8. Weinberger, L., Ayyash, M., Novershtern, N., Hanna, J.H.: Dynamic stem cell states: naive to primed pluripotency in rodents and humans. *Nat. Rev. Mol. Cell Biol.* 030676 (2016).
9. Okita, K., Ichisaka, T., Yamanaka, S.: Generation of germline-competent induced pluripotent stem cells. *Nature.* 448, 313–317 (2007).
10. Robinton, D.A. a., Daley, G.G.Q.: The promise of induced pluripotent stem cells in research and therapy. *Nature.* 481, 295–305 (2012).
11. Hanna, J., Saha, K., Pando, B., van Zon, J., Lengner, C.J., Creighton, M.P., van Oudenaarden, A., Jaenisch, R.: Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature.* 462, 595–601 (2009).
12. Banito, A., Gil, J.: Induced pluripotent stem cells and senescence: learning the biology to improve the technology. *EMBO Rep.* 11, 353–9 (2010).

13. Polo, J.M.: Phases of reprogramming. *Stem Cell Res.* 12, 754–761 (2014).
14. Buganim, Y., Faddah, D.A., Jaenisch, R.: Mechanisms and models of somatic cell reprogramming. *Nat. Rev. Genet.* 14, 427–39 (2013).
15. Soufi, A., Donahue, G., Zaret, K.S.: Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell.* 151, 994–1004 (2012).
16. Apostolou, E., Hochedlinger, K.: Chromatin dynamics during cellular reprogramming. *Nature.* 502, 462–71 (2013).
17. Cacchiarelli, D., Trapnell, C., Ziller, M.J., Soumillon, M., Cesana, M., Karnik, R., Donaghey, J., Smith, Z.D., Ratanasirintrao, 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.: Integrative Analyses of Human Reprogramming Reveal Dynamic Nature of Induced Pluripotency. *Cell.* 162, 412–424 (2015).
18. Hussein, S.M.I., Puri, M.C., Tonge, P.D., Benevento, M., Corso, A.J., Clancy, J.L., Mosbergen, R., Li, M., Lee, D.-S., Cloonan, N., Wood, D.L.A., Munoz, J., Middleton, R., Korn, O., Patel, H.R., White, C.A., Shin, J.-Y., Gauthier, M.E., Lê Cao, K.-A., Kim, J.-I., Mar, J.C., Shakiba, N., Ritchie, W., Rasko, J.E.J., Grimmond, S.M., Zandstra, P.W., Wells, C.A., Preiss, T., Seo, J.-S., Heck, A.J.R., Rogers, I.M., Nagy, A.: Genome-wide characterization of the routes to pluripotency. *Nature.* 516, 198–206 (2014).
19. Ruiz, S., Panopoulos, A.D., Herreras, A., Bissig, K.-D., Lutz, M., Berggren, W.T., Verma, I.M., Izpisua Belmonte, J.C.: A high proliferation rate is required for cell reprogramming and maintenance of human embryonic stem cell identity. *Curr. Biol.* 21, 45–52 (2011).
20. Li, F., He, Z., Shen, J., Huang, Q., Li, W., Liu, X., He, Y., Wolf, F., Li, C.Y.: Apoptotic caspases regulate induction of iPSCs from human fibroblasts. *Cell Stem Cell.* 7, 508–520 (2010).
21. Xu, Y., Zhang, M., Li, W., Zhu, X., Bao, X., Qin, B., Hutchins, A.P., Esteban, M.A.: Transcriptional Control of Somatic Cell Reprogramming. *Trends Cell Biol.* xx, 1–17 (2016).
22. Nishimura, K., Kato, T., Chen, C., Oinam, L., Shiomitsu, E., Ayakawa, D., Ohtaka, M., Fukuda, A., Nakanishi, M., Hisatake, K.: Manipulation of KLF4 expression generates iPSCs paused at successive stages of reprogramming. *Stem Cell Reports.* 3, 915–929 (2014).
23. Wei, Z., Gao, F., Kim, S., Yang, H., Lyu, J., An, W., Wang, K., Lu, W.: Klf4 organizes long-range chromosomal interactions with the OCT4 locus in reprogramming and pluripotency. *Cell*

- Stem Cell. 13, 36–47 (2013).
24. Chan, E.M., Ratanasirintrawoot, S., Park, I.-H., Manos, P.D., Loh, Y.-H., Huo, H., Miller, J.D., Hartung, O., Rho, J., Ince, T. a, Daley, G.Q., Schlaeger, T.M.: Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. *Nat. Biotechnol.* 27, 1033–7 (2009).
 25. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Sasaki, A., Yamamoto, M., Nakamura, M., Sutou, K., Osafune, K., Yamanaka, S.: Induction of pluripotency in human somatic cells via a transient state resembling primitive streak-like mesendoderm. *Nat. Commun.* 5, 3678 (2014).
 26. Nishino, K., Toyoda, M., Yamazaki-Inoue, M., Fukawatase, Y., Chikazawa, E., Sakaguchi, H., Akutsu, H., Umezawa, A.: DNA methylation dynamics in human induced pluripotent stem cells over time. *PLoS Genet.* 7, (2011).
 27. Takahashi, K., Yamanaka, S.: A developmental framework for induced pluripotency. *Development.* 142, 3274–85 (2015).
 28. Lujan, E., Zunder, E.R., Ng, Y.H., Goronzy, I.N., Nolan, G.P., Wernig, M.: Early reprogramming regulators identified by prospective isolation and mass cytometry. *Nature.* 521, 352–356 (2015).
 29. Tanabe, K., Nakamura, M., Narita, M., Takahashi, K., Yamanaka, S.: Maturation, not initiation, is the major roadblock during reprogramming toward pluripotency from human fibroblasts. *Proc. Natl. Acad. Sci.* 110, 12172–12179 (2013).
 30. Wolf, D., Goff, S.P.: TRIM28 Mediates Primer Binding Site-Targeted Silencing of Murine Leukemia Virus in Embryonic Cells. *Cell.* 131, 46–57 (2007).
 31. Wolf, D., Hug, K., Goff, S.P.: TRIM28 mediates primer binding site-targeted silencing of Lys1,2 tRNA-utilizing retroviruses in embryonic cells. *Proc. Natl. Acad. Sci. U. S. A.* 105, 12521–12526 (2008).
 32. Wolf, D., Goff, S.P.: Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature.* 458, 1201–1204 (2009).
 33. Schlesinger, S., Goff, S.P.: Silencing of proviruses in embryonic cells: efficiency, stability and chromatin modifications. *EMBO Rep.* 14, 73–9 (2013).
 34. Ellis, J., Hotta, A., Rastegar, M.: Retrovirus Silencing by an Epigenetic TRIM. *Cell.* 131, 13–14 (2007).

35. Matsui, T., Leung, D., Miyashita, H., Maksakova, I. a, Miyachi, H., Kimura, H., Tachibana, M., Lorincz, M.C., Shinkai, Y.: Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature*. 464, 927–931 (2010).
36. Yang, B.X., El Farran, C.A., Guo, H.C., Yu, T., Fang, H.T., Wang, H.F., Schlesinger, S., Seah, Y.F.S., Goh, G.Y.L., Neo, S.P., Li, Y., Lorincz, M.C., Tergaonkar, V., Lim, T.M., Chen, L., Gunaratne, J., Collins, J.J., Goff, S.P., Daley, G.Q., Li, H., Bard, F.A., Loh, Y.H.: Systematic Identification of Factors for Provirus Silencing in Embryonic Stem Cells. *Cell*. 163, 230–245 (2015).
37. Hong, S., Hwang, D.-Y., Yoon, S., Isacson, O., Ramezani, A., Hawley, R.G., Kim, K.-S.: Functional analysis of various promoters in lentiviral vectors at different stages of in vitro differentiation of mouse embryonic stem cells. *Mol. Ther.* 15, 1630–9 (2007).
38. Xia, X., Zhang, Y., Zieth, C.R., Zhang, S.-C.: Transgenes delivered by lentiviral vector are suppressed in human embryonic stem cells in a promoter-dependent manner. *Stem Cells Dev.* 16, 167–76 (2007).
39. Schlesinger, S., Meshorer, E., Goff, S.P.: Asynchronous transcriptional silencing of individual retroviral genomes in embryonic cells. *Retrovirology*. 11, 31 (2014).
40. Ebrahimi, B.: Reprogramming barriers and enhancers: strategies to enhance the efficiency and kinetics of induced pluripotency. *Cell Regen. (London, England)*. 4, 10 (2015).
41. Kawamura, T., Suzuki, J., Wang, Y. V, Menendez, S., Morera, L.B., Raya, A., Wahl, G.M., Izpisu, J.C.: Linking the p53 tumour suppressor pathway to somatic cell reprogramming. 460, 2–7 (2009).
42. Zhao, Y., Yin, X., Qin, H., Zhu, F., Liu, H., Yang, W., Zhang, Q., Xiang, C., Hou, P., Song, Z., Liu, Y., Yong, J., Zhang, P., Cai, J., Liu, M., Li, H., Li, Y., Qu, X., Cui, K., Zhang, W., Xiang, T., Wu, Y., Zhao, Y., Liu, C., Yu, C., Yuan, K., Lou, J., Ding, M., Deng, H.: Two Supporting Factors Greatly Improve the Efficiency of Human iPSC Generation. *Cell Stem Cell*. 3, 475–479 (2008).
43. Hong, H., Takahashi, K., Ichisaka, T., Aoi, T., Kanagawa, O., Nakagawa, M., Okita, K., Yamanaka, S.: Suppression of induced pluripotent stem cell generation by the p53 – p21 pathway. *Nature*. 460, 1132–1135 (2009).
44. Kareta, M.S., Gorges, L.L., Hafeez, S., Benayoun, B.A., Marro, S., Zmoos, A.F., Cecchini, M.J., Spacek, D., Batista, L.F.Z., O'Brien, M., Ng, Y.H., Ang, C.E., Vaka, D., Artandi, S.E., Dick, F.A., Brunet, A., Sage, J., Wernig, M.: Inhibition of pluripotency networks by the Rb

- tumor suppressor restricts reprogramming and tumorigenesis. *Cell Stem Cell*. 16, 39–50 (2015).
45. Utikal, J., Polo, J.M., Stadtfeld, M., Maherali, N., Kulalert, W., Walsh, R.M., Khalil, A., Rheinwald, J.G., Hochedlinger, K.: Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. *Nature*. 460, 1145–8 (2009).
 46. Li, H., Collado, M., Villasante, A., Strati, K., Ortega, S., Cañamero, M., Blasco, M. a, Serrano, M.: The Ink4/Arf locus is a barrier for iPS cell reprogramming. *Nature*. 460, 1136–1139 (2009).
 47. Liao, J., Marumoto, T., Yamaguchi, S., Okano, S., Takeda, N., Sakamoto, C., Kawano, H., Nii, T., Miyamoto, S., Nagai, Y., Okada, M., Inoue, H., Kawahara, K., Suzuki, A., Miura, Y., Tani, K.: Inhibition of PTEN tumor suppressor promotes the generation of induced pluripotent stem cells. *Mol. Ther.* 21, 1242–50 (2013).
 48. Di Stefano, B., Collombet, S., Jakobsen, J.S., Wierer, M., Sardina, J.L., Lackner, A., Stadhouders, R., Segura-Morales, C., Francesconi, M., Limone, F., Mann, M., Porse, B., Thieffry, D., Graf, T.: C/EBP α creates elite cells for iPSC reprogramming by upregulating Klf4 and increasing the levels of Lsd1 and Brd4. *Nat Cell Biol.* 18, 371–381 (2016).
 49. Di Stefano, B., Sardina, J.L., van Oevelen, C., Collombet, S., Kallin, E.M., Vicent, G.P., Lu, J., Thieffry, D., Beato, M., Graf, T.: C/EBP α poises B cells for rapid reprogramming into induced pluripotent stem cells. *Nature*. 506, 235–9 (2014).
 50. Mansour, A.A., Gafni, O., Weinberger, L., Zviran, A., Ayyash, M., Rais, Y., Krupalnik, V., Zerbib, M., Amann-Zalcenstein, D., Maza, I., Geula, S., Viukov, S., Holtzman, L., Pribluda, A., Canaani, E., Horn-Saban, S., Amit, I., Novershtern, N., Hanna, J.H.: The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. *Nature*. 488, 409–13 (2012).
 51. Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., Melton, D. a: Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat. Biotechnol.* 26, 795–7 (2008).
 52. Onder, T.T., Kara, N., Cherry, A., Sinha, A.U., Zhu, N., Bernt, K.M., Cahan, P., Marcarci, B.O., Unternaehrer, J., Gupta, P.B., Lander, E.S., Armstrong, S. a., Daley, G.Q.: Chromatin-modifying enzymes as modulators of reprogramming. *Nature*. 483, 598–602 (2012).
 53. Ichida, J.K., T C W, J., Williams, L. a, Carter, A.C., Shi, Y., Moura, M.T., Ziller, M., Singh, S., Amabile, G., Bock, C., Umezawa, A., Rubin, L.L., Bradner, J.E., Akutsu, H., Meissner, A.,

- Eggan, K.: Notch inhibition allows oncogene-independent generation of iPS cells. *Nat. Chem. Biol.* 10, 1–9 (2014).
54. Sakurai, K., Talukdar, I., Patil, V.S., Dang, J., Li, Z., Chang, K.Y., Lu, C.C., Delorme-Walker, V., Dermardirossian, C., Anderson, K., Hanein, D., Yang, C.S., Wu, D., Liu, Y., Rana, T.M.: Kinome-wide functional analysis highlights the role of cytoskeletal remodeling in somatic cell reprogramming. *Cell Stem Cell.* 14, 523–534 (2014).
 55. Qin, H., Diaz, A., Blouin, L., Lebbink, R.J., Patena, W., Tanbun, P., Leproust, E.M., McManus, M.T., Song, J.S., Ramalho-Santos, M.: Systematic identification of barriers to human Ipsc generation. *Cell.* 158, 449–461 (2014).
 56. Yang, C.-S., Chang, K.-Y., Rana, T.M.: Genome-wide functional analysis reveals factors needed at the transition steps of induced reprogramming. *Cell Rep.* 8, 327–37 (2014).
 57. Onder, T.T., Kara, N., Cherry, A., Sinha, A.U., Zhu, N., Bernt, K.M., Cahan, P., Marcarci, B.O., Unternaehrer, J., Gupta, P.B., Lander, E.S., Armstrong, S. a, Daley, G.Q.: Chromatin-modifying enzymes as modulators of reprogramming. *Nature.* 483, 598–602 (2012).
 58. Sridharan, R., Gonzales-Cope, M., Chronis, C., Bonora, G., McKee, R., Huang, C., Patel, S., Lopez, D., Mishra, N., Pellegrini, M., Carey, M., Garcia, B. a, Plath, K.: Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1 γ in reprogramming to pluripotency. *Nat. Cell Biol.* 15, 872–82 (2013).
 59. Dos Santos, R.L., Tosti, L., Radzishewska, A., Caballero, I.M., Kaji, K., Hendrich, B., Silva, J.C.R.: MBD3/NuRD facilitates induction of pluripotency in a context-dependent manner. *Cell Stem Cell.* 15, 102–110 (2014).
 60. Rais, Y., Zviran, A., Geula, S., Gafni, O., Chomsky, E., Viukov, S., Mansour, A.A., Caspi, I., Krupalnik, V., Zerbib, M., Maza, I., Mor, N., Baran, D., Weinberger, L., Jaitin, D. a, Lara-Astiaso, D., Blecher-Gonen, R., Shipony, Z., Mukamel, Z., Hagai, T., Gilad, S., Amann-Zalcenstein, D., Tanay, A., Amit, I., Novershtern, N., Hanna, J.H.: Deterministic direct reprogramming of somatic cells to pluripotency. *Nature.* 502, 65–70 (2013).
 61. Vidal, S.E., Amlani, B., Chen, T., Tsigos, A., Stadtfeld, M.: Combinatorial modulation of signaling pathways reveals cell-type-specific requirements for highly efficient and synchronous iPSC reprogramming. *Stem Cell Reports.* 3, 574–584 (2014).
 62. Tahmasebi, S., Alain, T., Rajasekhar, V.K., Zhang, J., Prager-khoutorsky, M., Khoutorsky, A., Dogan, Y., Gkogkas, C.G., Petroulakis, E., Sylvestre, A., Ghorbani, M., Assadian, S., Yamanaka, Y., Vinagolu-baur, J.R., Teodoro, J.G., Kim, K.: Article Multifaceted Regulation

- of Somatic Cell Reprogramming by mRNA Translational Control. *Stem Cell*. 14, 606–616 (2013).
63. Chen, T., Dent, S.Y.R.: Chromatin modifiers and remodellers: regulators of cellular differentiation. *Nat. Rev. Genet.* 15, 93–106 (2014).
 64. Ho, L., Crabtree, G.R.: Chromatin remodelling during development. *Nature*. 463, 474–484 (2010).
 65. Bestor, T.H.: The DNA methyltransferases of mammals. *Hum. Mol. Genet.* 9, 2395–2402 (2000).
 66. Auclair, G., Weber, M.: Mechanisms of DNA methylation and demethylation in mammals. *Biochimie*. 94, 2202–2211 (2012).
 67. Bannister, A.J., Kouzarides, T.: Regulation of chromatin by histone modifications. *Cell Res.* 21, 381–395 (2011).
 68. Li, B., Carey, M., Workman, J.L.: The role of chromatin during transcription. *Cell*. 128, 707–19 (2007).
 69. Kouzarides, T.: Chromatin Modifications and Their Function. *Cell*. 128, 693–705 (2007).
 70. Torres, I.O., Fujimori, D.G.: Functional coupling between writers, erasers and readers of histone and DNA methylation. *Curr. Opin. Struct. Biol.* 35, 68–75 (2015).
 71. Musselman, C.A., Lalonde, M.-E., Côté, J., Kutateladze, T.G.: Perceiving the epigenetic landscape through histone readers. *Nat. Struct. Mol. Biol.* 19, 1218–27 (2012).
 72. Schuettengruber, B., Martinez, A.-M., Iovino, N., Cavalli, G.: Trithorax group proteins: switching genes on and keeping them active. *Nat. Rev. Mol. Cell Biol.* 12, 799–814 (2011).
 73. Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., Meissner, A.: Dissecting direct reprogramming through integrative genomic analysis. *Nature*. 454, 49–55 (2008).
 74. Di Micco, R., Fontanals-Cirera, B., Low, V., Ntziachristos, P., Yuen, S.K., Lovell, C.D., Dolgalev, I., Yonekubo, Y., Zhang, G., Rusinova, E., Gerona-Navarro, G., Ca??amero, M., Ohlmeyer, M., Aifantis, I., Zhou, M.M., Tsiganos, A., Hernando, E.: Control of embryonic stem cell identity by brd4-dependent transcriptional elongation of super-enhancer-associated pluripotency genes. *Cell Rep.* 9, 234–247 (2014).
 75. Wang, L., Du, Y., Ward, J.M., Shimbo, T., Lackford, B., Zheng, X., Miao, Y.L., Zhou, B.,

- Han, L., Fargo, D.C., Jothi, R., Williams, C.J., Wade, P.A., Hu, G.: INO80 facilitates pluripotency gene activation in embryonic stem cell self-renewal, reprogramming, and blastocyst development. *Cell Stem Cell*. 14, 575–591 (2014).
76. Liu, L., Xu, Y., He, M., Zhang, M., Cui, F., Lu, L., Yao, M., Tian, W., Benda, C., Zhuang, Q., Huang, Z., Li, W., Li, X., Zhao, P., Fan, W., Luo, Z., Li, Y., Wu, Y., Hutchins, A.P., Wang, D., Tse, H.F., Schambach, A., Frampton, J., Qin, B., Bao, X., Yao, H., Zhang, B., Sun, H., Pei, D., Wang, H., Wang, J., Esteban, M.A.: Transcriptional pause release is a rate-limiting step for somatic cell reprogramming. *Cell Stem Cell*. 15, 574–588 (2014).
 77. Cheloufi, S., Elling, U., Hopfgartner, B., Jung, Y.L., Murn, J., Ninova, M., Hubmann, M., Badeaux, A.I., Euong Ang, C., Tenen, D., Wesche, D.J., Abazova, N., Hogue, M., Tasdemir, N., Brumbaugh, J., Rathert, P., Jude, J., Ferrari, F., Blanco, A., Fellner, M., Wenzel, D., Zinner, M., Vidal, S.E., Bell, O., Stadtfeld, M., Chang, H.Y., Almouzni, G., Lowe, S.W., Rinn, J., Wernig, M., Aravin, A., Shi, Y., Park, P.J., Penninger, J.M., Zuber, J., Hochedlinger, K.: The histone chaperone CAF-1 safeguards somatic cell identity. *Nature*. 528, 218–224 (2015).
 78. Pereira, C.F., Piccolo, F.M., Tsubouchi, T., Sauer, S., Ryan, N.K., Bruno, L., Landeira, D., Santos, J., Banito, A., Gil, J., Koseki, H., Merckenschlager, M., Fisher, A.G.: ESCs require PRC2 to direct the successful reprogramming of differentiated cells toward pluripotency. *Cell Stem Cell*. 6, 547–556 (2010).
 79. Aloia, L., Di Stefano, B., Di Croce, L.: Polycomb complexes in stem cells and embryonic development. *Development*. 140, 2525–2534 (2013).
 80. Chen, J.J., Liu, H., Liu, J., Qi, J., Wei, B., Yang, J., Liang, H., Chen, Y., Chen, J.J., Wu, Y., Guo, L., Zhu, J., Zhao, X., Peng, T., Zhang, Y., Chen, S., Li, X., Li, D., Wang, T., Pei, D.: H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nat. Genet.* 45, 34–42 (2013).
 81. Wei, T., Chen, W., Wang, X., Zhang, M., Chen, J., Zhu, S., Chen, L., Yang, D., Wang, G., Jia, W., Yu, Y., Duan, T., Wu, M., Liu, H., Gao, S., Kang, J.: An HDAC2-TET1 switch at distinct chromatin regions significantly promotes the maturation of pre-iPS to iPS cells. *Nucleic Acids Res.* 43, 5409–5422 (2015).
 82. Jones, B., Su, H., Bhat, A., Lei, H., Bajko, J., Hevi, S., Baltus, G.A., Kadam, S., Zhai, H., Valdez, R., Gonzalo, S., Zhang, Y., Li, E., Chen, T.: The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure. *PLoS Genet.* 4, (2008).

83. Feng, Y., Yang, Y., Ortega, M.M., Copeland, J.N., Zhang, M., Jacob, J.B., Fields, T.A., Vivian, J.L., Fields, P.E.: Early mammalian erythropoiesis requires the Dot1L methyltransferase. *Blood*. 116, 4483–4491 (2010).
84. Shalem, O., Sanjana, N.E., Zhang, F.: High-throughput functional genomics using CRISPR-Cas9. *Nat. Rev. Genet.* 16, 299–311 (2015).
85. Hsu, P.D., Lander, E.S., Zhang, F.: Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell*. 157, 1262–1278 (2014).
86. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., Zhang, F.: Multiplex Genome Engineering Using CRISPR/Cas System. *Science* (80-.). 339, 819–824 (2013).
87. Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., Cradick, T.J., Marraffini, L.A., Bao, G., Zhang, F.: DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31, 827–32 (2013).
88. Doench, J.G., Hartenian, E., Graham, D.B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B.L., Xavier, R.J., Root, D.E.: Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat. Biotechnol.* 32, 1262–7 (2014).
89. Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., Virgin, H.W., Listgarten, J., Root, D.E.: Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* 1–12 (2016).
90. Tsai, S.Q., Zheng, Z., Nguyen, N.T., Liebers, M., Topkar, V. V, Thapar, V., Wyvekens, N., Khayter, C., Iafrate, A.J., Le, L.P., Aryee, M.J., Joung, J.K.: GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol.* 33, 187–197 (2015).
91. O’Geen, H., Henry, I.M., Bhakta, M.S., Meckler, J.F., Segal, D.J.: A genome-wide analysis of Cas9 binding specificity using ChIP-seq and targeted sequence capture. *Nucleic Acids Res.* 43, 3389–3404 (2015).
92. 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.: Double nicking by RNA-guided CRISPR cas9 for enhanced genome editing specificity. *Cell*. 154, 1380–1389 (2013).
93. Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S.,

- Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., Koonin, E.V., Zhang, F.: Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell*. 163, 759–771 (2015).
94. Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M., Joung, J.K.: Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol*. 32, 279–284 (2014).
 95. Malina, A., Mills, J.R., Cencic, R., Malina, A., Mills, J.R., Cencic, R., Yan, Y., Fraser, J., Schippers, L.M.: Repurposing CRISPR / Cas9 for in situ functional assays Repurposing CRISPR / Cas9 for in situ functional assays. 2602–2614 (2013).
 96. Nelles, D.A., Fang, M.Y., O’Connell, M.R., Xu, J.L., Markmiller, S.J., Doudna, J.A., Yeo, G.W.: Programmable RNA Tracking in Live Cells with CRISPR/Cas9. *Cell*. 165, 488–496 (2016).
 97. Hilton, I.B., D’Ippolito, A.M., Vockley, C.M., Thakore, P.I., Crawford, G.E., Reddy, T.E., Gersbach, C.A.: Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol*. 33, 510–7 (2015).
 98. Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., Lim, W.A., Weissman, J.S., Qi, L.S.: XCRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell*. 154, 442–451 (2013).
 99. Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., Nureki, O., Zhang, F.: Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*. 517, 583–588 (2014).
 100. Barrangou, R., Birmingham, A., Wiemann, S., Beijersbergen, R.L., Hornung, V., Smith, A.V.B.: Advances in CRISPR-Cas9 genome engineering: Lessons learned from RNA interference. *Nucleic Acids Res*. 43, 3407–3419 (2015).
 101. Luo, B., Cheung, H.W., Subramanian, A., Sharifnia, T., Okamoto, M., Yang, X., Hinkle, G., Boehm, J.S., Beroukhi, R., Weir, B.A., Mermel, C., Barbie, D.A., Awad, T., Zhou, X., Nguyen, T., Piquani, B., Li, C., Golub, T.R., Meyerson, M., Hacohen, N., Hahn, W.C., Lander, E.S., Sabatini, D.M., Root, D.E.: Highly parallel identification of essential genes in cancer cells. *Proc. Natl. Acad. Sci. U. S. A*. 105, 20380–5 (2008).
 102. Shi, J., Wang, E., Milazzo, J.P., Wang, Z., Kinney, J.B., Vakoc, C.R.: Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. *Nat. Biotechnol*. 1, (2015).

103. Wang, T., Wei, J.J., Sabatini, D.M., Lander, E.S.: Genetic screens in human cells using the CRISPR-Cas9 system. *Science*. 343, 80–4 (2014).
104. Wang, T., Birsoy, K., Hughes, N.W., Krupczak, K.M., Post, Y., Wei, J.J., Lander, E.S., Sabatini, D.M.: Identification and characterization of essential genes in the human genome. *Science* (80-.). 350, 1096–101 (2015).
105. Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelsen, T.S., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., Zhang, F.: Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. *Science* (80-.). 343, 84–87 (2014).
106. Evers, B., Jastrzebski, K., Heijmans, J.P.M., Grenrum, W., Beijersbergen, R.L., Bernards, R.: CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nat. Biotechnol.* 11–14 (2016).
107. Morgens, D.W., Deans, R.M., Li, A., Bassik, M.C.: Systematic comparison of CRISPR / Cas9 and RNAi screens for essential genes. *Nat. Biotechnol.* 1–4 (2016).
108. Beers, J., Gulbranson, D.R., George, N., Siniscalchi, L.I., Jones, J., Thomson, J.A., Chen, G.: Passaging and colony expansion of human pluripotent stem cells by enzyme-free dissociation in chemically defined culture conditions. *Nat. Protoc.* 7, 2029–40 (2012).
109. Stemmer, M., Thumberger, T., Del Sol Keyer, M., Wittbrodt, J., Mateo, J.L.: CCTop: An intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. *PLoS One*. 10, 1–11 (2015).
110. Duns, G., van den Berg, E., van Duivenbode, I., Osinga, J., Hollema, H., Hofstra, R.M.W., Kok, K.: Histone methyltransferase gene SETD2 is a novel tumor suppressor gene in clear cell renal cell carcinoma. *Cancer Res.* 70, 4287–4291 (2010).
111. Peters, S.L., Hlady, R. a, Opavska, J., Klinkebiel, D., Pirruccello, S.J., Talmon, G. a, Sharp, J.G., Wu, L., Jaenisch, R., Simpson, M. a, Karpf, a R., Opavsky, R.: Tumor suppressor functions of Dnmt3a and Dnmt3b in the prevention of malignant mouse lymphopoiesis. *Leukemia*. 28, 1138–42 (2014).
112. Herquel, B., Ouarrhni, K., Davidson, I.: The TIF1 α -related TRIM cofactors couple chromatin modifications to transcriptional regulation, signaling and tumor suppression. *Transcription*. 2, 231–236 (2011).
113. Medvedeva, Y.A., Lennartsson, A., Ehsani, R., Kulakovskiy, I. V., Vorontsov, I.E., Panahandeh, P., Khimulya, G., Kasukawa, T., Drabl??s, F.: EpiFactors: A comprehensive

- database of human epigenetic factors and complexes. Database. 2015, 1–10 (2015).
114. Houlard, M., Berlivet, S., Probst, A. V., Quivy, J.P., Héry, P., Almouzni, G., Gérard, M.: CAF-1 is essential for heterochromatin organization in pluripotent embryonic cells. *PLoS Genet.* 2, 1686–1696 (2006).
 115. Fasching, L., Kapopoulou, A., Sachdeva, R., Petri, R., Jönsson, M.E., Mönne, C., Turelli, P., Jern, P., Cammas, F., Trono, D., Jakobsson, J.: TRIM28 represses transcription of endogenous retroviruses in neural progenitor cells. *Cell Rep.* 10, 20–28 (2015).
 116. Luzzani, C., Solari, C., Losino, N., Ariel, W., Romorini, L., Blughermann, C., Seveler, G., Baraño, L., Miriuka, S., Guberman, A.: Modulation of chromatin modifying factors' gene expression in embryonic and induced pluripotent stem cells. *Biochem. Biophys. Res. Commun.* 410, 816–822 (2011).
 117. Zhang, H., Gayen, S., Xiong, J., Zhou, B., Shanmugam, A.K., Sun, Y., Karatas, H., Liu, L., Rao, R.C., Wang, S., Nesvizhskii, A.I., Kalantry, S., Dou, Y.: MLL1 Inhibition Reprograms Epiblast Stem Cells to Naïve Pluripotency. *Cell Stem Cell.* 18, 481–494 (2016).
 118. Ishiuchi, T., Enriquez-Gasca, R., Mizutani, E., Bošković, A., Ziegler-Birling, C., Rodriguez-Terrones, D., Wakayama, T., Vaquerizas, J.M., Torres-Padilla, M.-E.: Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. *Nat. Struct. Mol. Biol.* 22, 662–671 (2015).
 119. Wu, Q., Bruce, A.W., Jedrusik, A., Ellis, P.D., Andrews, R.M., Langford, C.F., Glover, D.M., Zernicka-Goetz, M.: CARM1 is required in embryonic stem cells to maintain pluripotency and resist differentiation. *Stem Cells.* 27, 2637–2645 (2009).
 120. Zhao, H. yong, Zhang, Y. jun, Dai, H., Zhang, Y., Shen, Y. fei: CARM1 mediates modulation of Sox2. *PLoS One.* 6, (2011).
 121. Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A., Young, R.A.: Super-enhancers in the control of cell identity and disease. *Cell.* 155, 934–47 (2013).
 122. Selvi, B.R., Batta, K., Kishore, a H., Mantelingu, K., Varier, R. a, Balasubramanyam, K., Pradhan, S.K., Dasgupta, D., Sriram, S., Agrawal, S., Kundu, T.K.: Identification of a novel inhibitor of coactivator-associated arginine methyltransferase 1 (CARM1)-mediated methylation of histone H3 Arg-17. *J. Biol. Chem.* 285, 7143–7152 (2010).
 123. Copeland, R. a, Solomon, M.E., Richon, V.M.: Protein methyltransferases as a target class for

- drug discovery. *Nat. Rev. Drug Discov.* 8, 724–732 (2009).
124. Gao, Z., Cox, J.L., Gilmore, J.M., Ormsbee, B.D., Mallanna, S.K., Washburn, M.P., Rizzino, A.: Determination of protein interactome of transcription factor Sox2 in embryonic stem cells engineered for inducible expression of four reprogramming factors. *J. Biol. Chem.* 287, 11384–11397 (2012).
 125. Percharde, M., Laval, F., Ng, J.H., Kumar, V., Tomaz, R.A., Martin, N., Yeo, J.C., Gil, J., Prabhakar, S., Ng, H.H., Parker, M.G., Azuara, V.: Ncoa3 functions as an essential Esrrb coactivator to sustain embryonic stem cell self-renewal and reprogramming. *Genes Dev.* 26, 2286–2298 (2012).
 126. Wu, Z., Yang, M., Liu, H., Guo, H., Wang, Y., Cheng, H., Chen, L.: Role of nuclear receptor coactivator 3 (Ncoa3) in pluripotency maintenance. *J. Biol. Chem.* 287, 38295–38304 (2012).
 127. Ishimura, a., Minehata, K. -i., Terashima, M., Kondoh, G., Hara, T., Suzuki, T.: Jmjd5, an H3K36me2 histone demethylase, modulates embryonic cell proliferation through the regulation of Cdkn1a expression. *Development.* 139, 749–759 (2012).
 128. Sdelci, S., Lardeau, C.-H., Tallant, C., Klepsch, F., Klaiber, B., Bennett, J., Rathert, P., Schuster, M., Penz, T., Fedorov, O., Superti-Furga, G., Bock, C., Zuber, J., Huber, K.V.M., Knapp, S., Müller, S., Kubicek, S.: Mapping the chemical chromatin reactivation landscape identifies BRD4-TAF1 cross-talk. *Nat. Chem. Biol.* (2016).
 129. Kidder, B.L., Palmer, S., Knott, J.G.: SWI/SNF-Brg1 regulates self-renewal and occupies core pluripotency-related genes in embryonic stem cells. *Stem Cells.* 27, 317–28 (2009).
 130. Zhang, X., Li, B., Li, W., Ma, L., Zheng, D., Li, L., Yang, W., Chu, M., Chen, W., Mailman, R.B., Zhu, J., Fan, G., Archer, T.K., Wang, Y.: Transcriptional repression by the BRG1-SWI/SNF complex affects the pluripotency of human embryonic stem cells. *Stem Cell Reports.* 3, 460–474 (2014).
 131. Ho, L., Ronan, J.L., Wu, J., Staahl, B.T., Chen, L., Kuo, A., Lessard, J., Nesvizhskii, A.I., Ranish, J., Crabtree, G.R.: An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proc. Natl. Acad. Sci. U. S. A.* 106, 5181–6 (2009).
 132. Panamarova, M., Cox, A., Wicher, K., Butler, R., Bulgakova, N., Jeon, S., Rosen, B., Seong, R.H., Skarnes, W., Crabtree, G., Zernicka-Goetz, M.: BAF chromatin remodelling complex is an epigenetic regulator of lineage specification in the early mouse embryo. *Development.* 155, dev.131961– (2016).

133. Grow, E.J., Flynn, R. a., Chavez, S.L., Bayless, N.L., Wossidlo, M., Wesche, D.J., Martin, L., Ware, C.B., Blish, C. a., Chang, H.Y., Pera, R.A.R., Wysocka, J.: Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. *Nature*. 522, 221–5 (2015).
134. Ding, X., Wang, X., Sontag, S., Qin, J., Wanek, P., Lin, Q., Zenke, M.: The Polycomb Protein Ezh2 Impacts on Induced Pluripotent Stem Cell Generation. *Stem Cells Dev.* 00, 19–24 (2014).
135. Rao, R.A., Dhele, N., Cheemadan, S., Ketkar, A., Jayandharan, G.R., Palakodeti, D., Rampalli, S.: Ezh2 mediated H3K27me3 activity facilitates somatic transition during human pluripotent reprogramming. *Sci. Rep.* 5, 8229 (2015).
136. Borkent, M., Bennett, B.D., Lackford, B., Bar-Nur, O., Brumbaugh, J., Wang, L., Du, Y., Fargo, D.C., Apostolou, E., Cheloufi, S., Maherali, N., Elledge, S.J., Hu, G., Hochedlinger, K.: A Serial shRNA Screen for Roadblocks to Reprogramming Identifies the Protein Modifier SUMO2. *Stem Cell Reports*. 6, 1–13 (2016).
137. Toh, C.-X.D., Chan, J.-W., Chong, Z.-S., Wang, H.F., Guo, H.C., Satapathy, S., Ma, D., Goh, G.Y.L., Khattar, E., Yang, L., Tergaonkar, V., Chang, Y.-T., Collins, J.J., Daley, G.Q., Wee, K.B., Farran, C.A.E., Li, H., Lim, Y.-P., Bard, F.A., Loh, Y.-H.: RNAi Reveals Phase-Specific Global Regulators of Human Somatic Cell Reprogramming. *Cell Rep.* 1–11 (2016).
138. Li, W., Xu, H., Xiao, T., Cong, L., Love, M.I., Zhang, F., Irizarry, R. a, Liu, J.S., Brown, M., Liu, X.S.: MAGECK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* 15, 554 (2014).
139. Diaz, A.A., Qin, H., Ramalho-Santos, M., Song, J.S.: HiTSelect: A comprehensive tool for high-complexity-pooled screen analysis. *Nucleic Acids Res.* 43, e16 (2015).
140. Yu, J., Silva, J., Califano, A.: ScreenBEAM: a novel meta-analysis algorithm for functional genomics screens via Bayesian hierarchical modeling. *Bioinformatics*. btv556 (2015).
141. Winter, J., Breinig, M., Heigwer, F., Bru, D., Leible, S., Pelz, O., Zhan, T., Boutros, M.: Systems biology caR pools : an R package for exploratory data analysis and documentation of pooled CRISPR / Cas9 screens. 1–3 (2015).

7 Vita

Can Aztekin was born in İzmir, Turkey on February 28, 1992. He has completed his Bachelor of Science degree in Biological Sciences and Bioengineering in Sabancı University in 2014, in İstanbul. He is working towards his Master of Science degree in Biomedical Sciences and Engineering at Koç University, İstanbul. His thesis is on CRISPR-Cas9 based investigation of chromatin modifiers in human somatic cell reprogramming.

