Investigating The Roles of CREBBP/EP300 and BRD9 in Somatic Cell Reprogramming to Pluripotency

by

Kenan Sevinç

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

Doctor of Philosophy

in

Molecular Biology and Genetics

KOÇ ÜNİVERSİTESİ

August 11, 2021

Investigating The Roles of CREBBP/EP300 and BRD9 in Somatic Cell Reprogramming to Pluripotency

Koç University

Graduate School of Sciences and Engineering

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

Kenan Sevinç

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

Committee Members:

Prof. Tamer Önder (Advisor)

Assoc. Prof. Tuğba Bağcı-Önder

Assoc. Prof. Umut Şahin

Prof. Udo Oppermann

Assoc. Prof. Nurhan Özlü

Date: 11.08.2021

To my beloved wife and daughter…

ABSTRACT

Investigating Roles of CREBBP/EP300 and BRD9 in Somatic Cell Reprogramming to Pluripotency

Kenan Sevinç

Doctor of Philosophy in Molecular Biology and Genetics

August 11, 2021

Somatic cells can be reprogrammed to induced pluripotent stem cells (iPSC) by overexpressing four transcription factors, OCT4, SOX2, KLF4 and MYC (OSKM). Epigenetic pathways that safeguard somatic cell identity are barriers for somatic cell reprogramming. Chemical or genetic inhibition of such pathways increases the efficiency of reprogramming. In this thesis, I investigated the effects of three chromatin factors CREB (cyclic-AMP response element binding protein) binding protein (CREBBP), E1A binding protein of 300 kDa (EP300) and bromodomain containing 9 (BRD9) during reprogramming using a combination of chemical probes and genetic loss of function experiments.

In the first chapter, I investigated barrier function of bromodomain-mediated interaction of CREBBP/EP300 in reprogramming with specific bromodomain inhibitors. RNA-Sequencing showed that CREBBP/EP300 bromodomain inhibition in fibroblasts decreases the expression of genes that are highly expressed in the cells. However, inhibition of the acetyltransferase activity of CREBBP/EP300 by another small molecule, A485, decreased the reprogramming efficiency by downregulating pluripotency-associated genes. Assay for transposase-accessible chromatin (ATAC-Sequencing) and H3K4me1 and H3K27ac chromatin immune-precipitation (ChIP-Sequencing) results indicated that CREBBP/EP300 bromodomain inhibition decreases H3K27ac at accessible chromatin regions and at putative enhancers. Additionally, I functionally showed that continuous expression of *PRRX1*, a downregulated transcription factor upon CREBBP/EP300 bromodomain inhibition, impairs iPSC generation by suppressing pluripotency-related gene expression. These results uncovered a role for the interaction of CREBBP/EP300 with acetylated lysines in safeguarding somatic cell identity.

In the second chapter of my thesis, I performed a SWI/SNF-focused CRISPR-Cas9-mediated knockout screen to identify regulatory subunits of reprogramming. This screen identified non-canonical BRG1- or BRM-associated factors (ncBAF) specific subunits, BRD9 and BRD4 Interacting Chromatin Remodeling Complex Associated Protein (BICRA, GLTSCR1) as barriers to reprogramming. Additionally, I showed that three structurally distinct BRD9 bromodomain inhibitors and a PROteolysis Targeting Chimera (PROTAC) degrader of BRD9 increase reprogramming efficiency and generate iPSCs with only OCT4 and SOX2 expression. BRD9 knockout iPSCs expressed pluripotency markers and short term BRD9 inhibition did not decrease cell proliferation and OCT4-positive cells indicating that BRD9 is dispensable for pluripotency induction and maintenance. However, BRD9 depletion and inhibition impaired mesoderm differentiation of iPSCs. RNA-Seq was performed on fibroblasts and results showed that BRD9 inhibition downregulates highly expressed genes in fibroblasts. ATAC-Seq results showed that chromatin accessibility around putative active enhancers rather than promoters was decreased upon BRD9 inhibition. Additionally, overexpression of MN1 and ZBTB38, which were downregulated upon BRD9 perturbations, decreased reprogramming efficiency. These results indicate a role for BRD9 and ncBAF in maintaining somatic cell identity through maintenance of active enhancers. The findings presented in this thesis show that chromatin factors safeguarding somatic cell identity are barriers to cellular reprogramming and inhibition of such factors can be utilized to increase the low efficiency of human iPSC generation.

ÖZETÇE

CREBBP/EP300 ve BRD9'un Somatik Hücrelerin Pluripotensiye Yeniden Programlamasındaki Rollerinin Araştırılması

Kenan Sevinç
Moleküler Biyoloji ve Genetik, Doktora
August 11, 2021

Somatik hücreler, dört transkripsiyon faktörünün; OCT4, SOX2, KLF4 ve MYC (OSKM) anlatımıyla, uyarılmış pluripotent kök hücrelere (uPKH) yeniden programlanabilir. Somatik hücrelerin kimliğini muhafaza eden epigenetik yolaklar, somatik hücrelerin yeniden programlamasına engeldir. Bu tarz yolakların kimyasal veya genetik inhibisyonu yeniden programlama verimini arttırmaktadır. Bu tezde, üç kromatin faktörünün; CREBBP, EP300 ve BRD9'un yeniden programlamaya etkisini kimyasal ve genetik işlev kaybı deneyleriyle araştırdım.

Tezin ilk bölümde, spesifik inhibitörlerle CREBBP/EP300'ün bromodomain'inin yeniden programlamaya engel olma fonksiyonunu araştırdım. RNA-dizileme, CREBBP/EP300 bromodomain inhibisyonunun fibroblastlarda çok ifade edilen genlerin anlatımını azalttığını göstermiştir. Ancak, CREBBP/EP300'ün A485 kimyasalıyla asetiltransferaz aktivitesi inhibe edilince pluripotensi genlerinin anlatımı düşerek yeniden programlama verimi azalmıştır. Transpozaz erişebilir kromatin analizi sonrası dizileme (ATAC-dizileme), H3K4me1 ve H3K27ac kromatin immün-çöktürme sonrası dizileme (ChIP-dizileme) sonuçları H3K27ac seviyesinin, erişilebilir kromatin ve olası enhancer (arttırıcı) bölgelerinde CREBBP/EP300 bromodomain inhibisyonuyla azaldığını göstermiştir. Buna ek olarak, CREBBP/EP300 bromodomain inhibisyonuyla anlatımı azalan transkripsiyon faktörlerinden biri olan *PRRX1*'in devamlı anlatımının uPKH oluşumunu pluripotensiyle alakalı genlerin anlatımını baskılayarak engellediğini fonksiyonel olarak gösterdim. Bu sonuçlar CREBBP/EP300'ün asetillenmiş lizinlerle etkileşiminin, somatik hücrelerin kimliğini muhafaza ettiğini ortaya çıkarmıştır.

Tezimin ikinci bölümünde, yeniden programlamayı düzenleyen faktörleri bulmak için CRISPR-Cas9 aracılığıyla SWI/SNF komplekslerine odaklanmış bir tarama yaptım. Bu tarama, kanonik olmayan ncBAF kompleksinin spesifik elemanlarından BRD9 ve GLTSCR1'in yeniden programlamaya engel olduklarını buldu. Buna ek olarak, üç farklı BRD9 bromodomain inhibitörünün ve BRD9 yıkıcısının yeniden programlama verimini arttırdığını, sadece OCT4 ve SOX2 anlatımıyla uPKH oluşumunu sağladığını gösterdim. BRD9 nakavt uPKH'lar pluripotensi belirteçlerini üretmesi, kısa süreli BRD9 inhibisyonunun hücre bölünmesi ve OCT4 pozitif hücre sayısını azaltmaması BRD9'un pluripotensi oluşumu ve devamlılığında vazgeçilebilir olduğuna işaret etmektedir. Ancak, BRD9 kaybı ve inhibisyonu, uPKH'ların mezoderme farklılaşmasını azaltmıştır. Fibroblastlarda RNA-dizilemesi yapılmış ve sonuçlar BRD9 inhibisyonunun bu hücrelerde anlatımı fazla olan genleri baskıladığını göstermiştir. ATAC-dizilemesi, erişebilir kromatinin promotör yerine olası aktif enhancer (arttırıcı) bölgelerinde BRD9 inhibisyonuyla azaldığını göstermiştir. Buna ek olarak, BRD9 inhibisyonuyla anlatımı azalan transkripsiyon faktörlerinden ZBTB38 ve MN1'in anlatımını arttırmak yeniden programlama verimini azaltmıştır. Bu sonuçlar, BRD9 ve ncBAF kompleksinin enhancer'lar aracılığıyla somatik hücrelerin kimlikleri muhafaza ettiğine işaret etmektedir. Bu tezde sunulan bulgular, somatik hücre kimliğini muhafaza eden kromatin faktörlerinin yeniden programlamaya engel olduğunu ve bu tarz faktörlerin inhibisyonunun düşük verimli uPKH oluşumunu arttırmakta kullanılabileceği gösterilmiştir.

ACKNOWLEDGMENTS

I would like to thank my supervisor Prof. Tamer Önder for teaching me invaluable lessons on how to approach a scientific idea and data. He encouraged me to take a step back and think about the big picture rather than persuading small excitements or distractions on the road. I learned a lot from his neutral stand against science and patience for mistakes. At the moment when I first started to read his first-author papers, I realized that those were kind of science that I want to pursue and I never regret to join his group. Beyond his scientific influence on me, he was the one who introduced me with Gülben by saying "She is everything to our lab." and three years later he was my marriage witness.

I am grateful to Prof. Udo Oppermann who welcomed me twice in his lab at Oxford University. Assist. Prof. Can Aztekin taught me the essentiality of keeping up with the literature during my very first semester as a PhD student and I am thankful for his heads up. I would like to thank my thesis committee members Assoc. Prof. Tuğba Bağcı-Önder and Assoc. Prof. Umut Şahin for their constructive criticisms on my presentation skills and data. I am grateful that Assoc. Prof. Nurhan Özlü has accepted to be one of the thesis jury members.

I thank all previous and current Onder lab members for creating a nice working atmosphere. I was lucky to learn most of molecular biology and tissue culture techniques from experts such as Can, Gülben and Ayyub. Luckily, they remained as close friends if nothing more to me. I had a chance to supervise several undergraduate students during this thesis and I am thankful for Zeynep, Sidar, Hazal, Öznur, Gamze and Enes for showing how to be a better mentor for future. I spent more than 6 months as a visiting PhD student in Oxford and members of Oppermann never treated me like an outsider. Especially, I thank Dr. Martin Philpott, Dr. James Edward, Graham Wells and Dr. Adam Cribbs as they are the ones that guided and helped me through new working environment and unfamiliar experiments.

During the thesis, I was supported by TÜBİTAK-BİDEB 2211-A scholarship and I am grateful for their generous stipend. This research was supported by TUBITAK 1001 - 219Z209I grant. I thank Graduate School of Health Sciences and Science & Engineering programs for providing side benefits. Newton Advanced Fellowship granted for my supervisor Tamer Önder covered my whole stay at Oxford.

I had many colleagues but three of them became special to me. I thank Tunç Morova, he was the weird guy who made part of my PhD a sit-com show. Besides, he was cherishing

and motivating when I needed and I was always happy to discuss our projects that intersect computational and chromatin biology. Dr. Alişan Kayabölen was another mate that I enjoyed spending time together while we were brewing our own beers and honey mead. Together with Doğancan Özturan and Alişan, I came up with wonderful scientific projects that most of them were published at least 20 years ago. Despite this, I gained a lot of experience as being Reviewer #3 while evaluating their proposals.

I should thank Ali Burak Ünal for being my best friend at high school. It was one of my happiest moments when he decided to choose Bilkent University for undergraduate studies and we were roommates for many years. I am grateful to my high school math teacher, Onur Barut's guidance on acknowledging the importance of science and motivating me to become a scientist even unintentionally. During my bachelor's degree, I met Muzaffer Yıldırım and Alper Duranel and we became a group of people who can talk about anything especially while boozing.

With Gülben being an essential part of my life, her family and close friends became my family, too. I'd like to thank my mother-, father, sister-in-law, Nuray, Hayrettin and Zeycan for their faith in us and support. I am grateful to have joyful, mindful and supportive friends Özge and Eren Şahin and the recent member of Şahin family, Bilge.

I cannot thank enough my mother, Ayten, my father, Metin and my brother, Erdinç. They are always here to help me when I need most. I loved our chats on breakfast table on biological and medical questions while Erdinç and I were little teenagers. I think the reason why Erdinç is a medical doctor and I am a scientist is their dedication to open discussion. My parents sacrificed many things and came to Istanbul to care their grand-daugther, Maya. I owe them big time.

At last, I feel to be the luckiest guy on earth to marry Gülben, she is always the one for me. She is always there when I feel depressed, when I am too worried for something or even when I need a hand for my experiments. She is the equalizer of my life and she gave me the world's best gift by bearing our little girl, Maya, while she was still a PhD student. We tackled many obstacles together and with Maya we will tackle many more together.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

2C	2-cell stage embryos
ICM	inner cell mass
TE	trophoectoderm
PE	primitive endoderm
ESC	Embryonic stem cell
ECC	embryonic carcinoma cell
PGC	primordial germ cell
TF	transcription factor
OSKM	Oct4, Sox2, Klf4 and c-Myc
iPSC	induced pluripotent stem cells
NHEK	neonatal human epidermal keratinocyte
NHDF	Normal human dermal fibroblast
MEF	mouse embryonic fibroblast
MET	mesenchymal-to-epithelial transition
EMT	epithelial-to-mesenchymal transition
RPAP1	RNA polymerase II associated protein 1
DMSO	Dimethyl sulfoxide
BET	bromodomain and extra-terminal
CREB	cAMP-response element binding protein
CREBBP, CBP	cAMP-response element binding protein (CREB) binding protein

EP300, p300	E1A binding protein p300
BRD9	Bromodomain containing-9
BRD7	Bromodomain containing-7
BAF	BRG1- or BRM-associated factor
ncBAF	non-canonical BAF
PBAF	polybromo-associated BAF
SWI/SNF	SWItch/Sucrose Non–Fermentable
PRC	Polycomb repressive complex
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
GSEA	Gene set enrichment analysis
ChIP	chromatin immunoprecipitation
IGV	Integrative Genomics Viewer
GO	Gene Ontology
DNMT	DNA (cytosine-5)-methyltransferase
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
DMEM	Dulbecco's Modified Eagle Medium
PBS	Phosphate-buffered saline
KO	knock-out
RT qPCR	Real time quantitative polymerase chain reaction
Ct	threshold cycle
cDNA	complementary DNA

Chapter 1: INTRODUCTION

1.1.Mammalian pre-implantation embryo development

In mammalian organisms, zygote and cells from 2-cell stage embryos (2C) are totipotent, meaning that each of them has the potential to differentiate into all cells that make up body as well as into cells that form extra-embryonic lineages (Baker and Pera, 2018). Although individual cells from 4-cell and 8-cell stage are able to contribute to cells of both extraembryonic and embryonic lineages of chimeric organism, a single cell does not have the capacity to generate an organism by itself. Around 8-cell stage, embryo becomes compacted, then, reaches to morula-stage in which cleavages divide cells without an increase in size and observable cell lineage specifications (Płusa and Piliszek, 2020). However, even at 4-cell stage, each cell has a bias towards a specific lineage differentiation and the orientation of in and out during 8-cell stage allows a spatial decision following differences in external morphogens and relative physical locations indicating early cell-fate decisions starting from this stage (Leung and Zernicka-Goetz, 2015). The first visible lineage specification occurs at the blastula stage, when inner cells form inner cell mass (ICM) and outer cells form extra-embryonic lineage called trophoectoderm (TE) (Jedrusik, 2015). Several lineage-specific transcription factors and certain signaling pathways dictate the initial lineage specification and make sure the maintenance of lineage identity. Cdx2, Gata3 and Eomes plays key role in TE differentiation and inhibition of Hippo pathway upregulates Cdx2 to maintain TE lineage identity (Zhu and Zernicka-Goetz, 2020). The interplay between Oct4, Sox2 and Nanog in the inner cells define an ICM fate and Oct4-mediated downregulation of Cdx2 safeguards ICM lineage identity (Jedrusik, 2015).

The second major cell fate specification event follows ICM and TE differentiation during pre-implantation embryo development (Thiery et al., 2009). Cells from ICM generate primitive endoderm (PE) and epiblast (EPI) at this stage depending on physical location, forces, signaling pathways and cell fate determining transcription factor expression. FGF/ERK signaling is critical for this cell fate determination as inhibiting this pathway directs ICM cells to EPI-cells whereas increase in FGF2/4 results in bias towards differentiating into PE-cells (Płusa and Piliszek, 2020). This signaling pathway downregulates OCT4 and NANOG expression while upregulates the expression of

SOX17 and GATA6 whose target genes, further, establish the PE fate determination and maintenance. On the contrary, other cells in ICM continue to express OCT4, SOX2 and NANOG all of which ensures a pluripotent fate maintenance (Molè et al., 2020).

1.2.Embryonic stem cells and pluripotency

Embryonic stem cells (ESC) can be derived from inner cell mass of a blastula and can be cultured *in vitro* (Cowan et al., 2004; Evans and Kaufman, 1981; Klimanskaya et al., 2006; Martin, 1981; Strelchenko et al., 2004; Thomson et al., 1998). ESCs self-renew indefinitely, are pluripotent and they can differentiate into three germ layers; ectoderm, endoderm and mesoderm upon exposure to external cues (Oh and Jang, 2019; Romito and Cobellis, 2016). ESCs were successfully derived from multiple species including human, which facilitated the study of different pluripotent states, species-specific characteristics and dependencies on signaling pathways involved to maintain this cell identity.

Mouse and human ESCs (mESC and hESC, respectively) differ from each other by pluripotency states and essential signaling pathways that maintain them. mESCs derived from pre-implantation ICM form tightly packed, round shaped colonies (Ohtsuka and Dalton, 2008). hESCs, similarly, self-renew indefinitely and differentiate into embryonic lineages *in vitro* and *in vivo* as in teratoma assays. Although the most stringent test for mouse pluripotency is contribution to chimeric organism, this test cannot be performed for human pluripotency for ethical concerns (Smith et al., 2009). Unlike mESC, hESCs grow as flat and compact colonies. LIF is dispensable for hESC maintenance and BMP4 induces trophoblast differentiation (Dahéron et al., 2004; Li et al., 2013). TGF-β members Activin A and Nodal, FGF2 and IGF are necessary for hESCs to self-renew without losing pluripotency (Vallier et al., 2005). In serum-free conditions, mESCs self-renew by retaining pluripotency under 2i conditions: MAPK inhibition by PD0325901 and GSK3 inhibition by CHIR99021 (Shimizu et al., 2012). This serum-free culture condition makes mESCs much closer to pluripotent cells from ICM which is called ground state or naïve pluripotency. These mESCs are morphologically, transcriptionally and epigenetically more homogeneous and their differentiation potential to embryonic lineages is highly preserved compared to mESCs cultured in serum-containing media (Ghimire et al., 2018). In female mESCs, there are two active X chromosomes whereas majority of female hESCs have one active, one silenced X chromosomes (Minkovsky et al., 2012). Mouse primed pluripotent stem cells, EpiSCs, derived from post-implantation epiblasts share previously mentioned morphology,

signaling pathway conditions, transcriptional and epigenetic programs of hESCs (Stadtfeld and Hochedlinger, 2010). These observations indicate that under standard culture conditions, hESCs are primed pluripotent stem cells (Chen and Lai, 2015). Overexpression of OCT4 and KLF4 or KLF4 and KLF2 along with 2i/LIF culture conditions can result in dome-shaped hESC colonies which re-activate silenced X chromosome, reminiscent of a naïve pluripotent state (Hanna et al., 2010). Several others established culture conditions including different inhibitors, growth factors and basal media to revert primed hESCs into naïve hESCs (Chan et al., 2013; Gafni et al., 2013; Takashima et al., 2014; Theunissen et al., 2014; Ware et al., 2014). These studies show that hESCs are inherently primed, resembling with mouse EpiSCs, however, naïve hESCs can be obtained under certain culture conditions.

1.3.Cellular reprogramming and ways to achieve it

It was proposed by Waddington that cell with greater potential to differentiate loses its potential during development (Waddington, 1956). Differentiation was presumed to be one-way and the cell identity was thought to be strictly conserved upon differentiation. However, this notion was challenged by several observations over the last seventy years. First, classical experiments by John Gurdon showed that when nucleus of an intestine cell is transferred into enucleated Xenopus oocyte, it generated a mature frog (Fishberg et al., 1958; Gurdon, 1960, 1962b, 1962a). Likewise, when the nucleus of a somatic cell is transferred into an enucleated egg, this cell becomes pluripotent (Byrne et al., 2003; Wilmut et al., 1997). Second, fully differentiated cells de-differentiate to gain a pluripotent character upon fusion with a pluripotent stem cell such as ESCs, embryonic carcinoma cells (ECCs) and primordial germ cells (PGCs) (Cowan et al., 2005; Do and Schöler, 2006; Tada et al., 1997, 2001). Third, cellular identity could be altered by overexpressing of lineage-specific transcription factors (TFs). For example, overexpressing MyoD, a transcription factor that is highly expressed in muscle cells, transdifferentiates mouse fibroblasts to myoblasts (Choi et al., 1990; Lattanzi et al., 1998). Further studies proved that forced expression of specific transcription factors can mediate lineage-conversions (Duran Alonso et al., 2018; Firas et al., 2015; Hu et al., 1995; Ieda, 2013; Pfisterer et al., 2011; Son et al., 2011; Vierbuchen et al., 2010; Wang et al., 2019b; Zhu et al., 2020). Together, these reprogramming approaches contradicted the Waddington's landscape and transformed this landscape with two-directional ways and cross-paths (Figure 1.1).

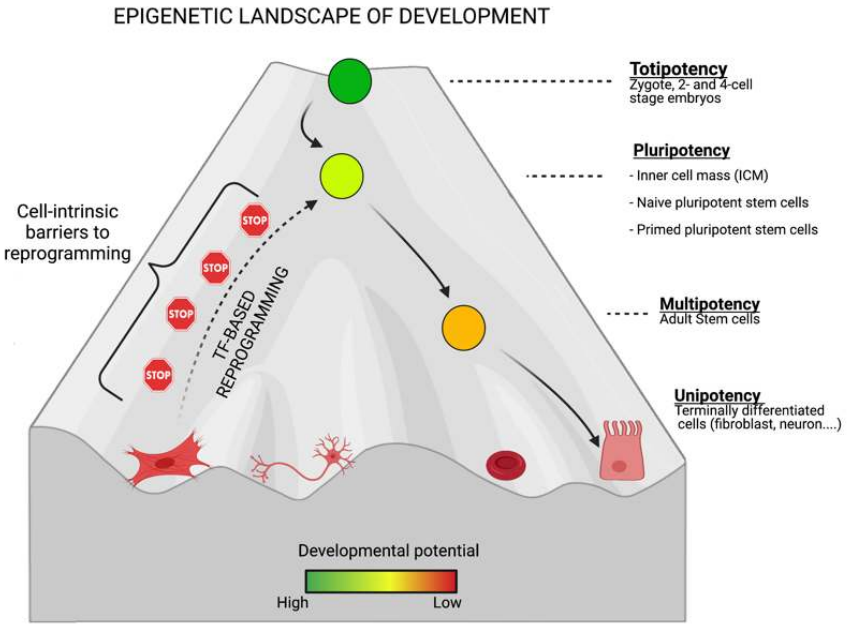


Figure 1.1. Epigenetic landscape during development and ways to reprogram cell fate.

1.4.Induced pluripotent stem cells (iPSC) and clinic implications

In 2006, Takahashi and Yamanaka discovered that overexpressing four transcription factors; Oct4, Sox2, Klf4 and c-Myc (OSKM) reprogrammed mouse embryonic fibroblasts into pluripotent-like cells which they called induced pluripotent stem cells (iPSC) (Takahashi and Yamanaka, 2006). iPSCs generated from MEFs by these 4 TFs expressed ESC markers, differentiated into three germ layers *in vivo* and contributed to chimeric mice when injected into blastocysts (Takahashi and Yamanaka, 2006). Subsequent studies by this group and others showed that the same factors can reprogram other types of somatic cells from different species including human into iPSCs (Aasen et al., 2008; Aoi et al., 2008; Lowry et al., 2008; Maherali and Hochedlinger, 2008; Park et al., 2008a, 2008b; Takahashi et al., 2007). These studies opened up great translational opportunities as iPSCs enable the autologous cells for individuals in need and offer disease modeling at high throughput (Robinton and Daley, 2012; Shi et al., 2017).

Several studies modeled complex neurologic diseases such as Parkinson and Alzheimer using iPSC derived neurons (Liao et al., 2016; Nguyen et al., 2011; Soldner et al., 2016; Yagi et al., 2011). In addition, iPSC derived three dimensional culture systems

such as organoids were abundantly utilized due to their easy manipulation and patient-specific character (Akbari et al., 2019; Camp et al., 2015; Dye et al., 2015; Garcez et al., 2016). Large scale drug screening platform provided by disease-specific iPSC derived cell types identified several compounds that are currently in trial to treat neurologic diseases including amyotrophic lateral sclerosis and spinal muscular atrophy (Bright et al., 2015; McNeish et al., 2015; Naryshkin et al., 2014; Shi et al., 2017). Excitingly, clinical trials started to test efficacy and safety of functional cells derived from iPSCs to treat neovascular age-related macular degeneration and heart diseases (Cyranoski, 2018; Mandai et al., 2017).

1.5.Roles of OSKM in reprogramming

Somatic cell reprogramming is thought to occur in three major stages. First stage is called initiation and corresponds to the time interval between OSKM overexpression and induction of mesenchymal-to-epithelial transition (MET) (Samavarchi-Tehrani et al., 2010). During this stage, OSK cooperatively binds their target DNA motifs and silences transcriptional program that maintains somatic cell identity. OSK occupy most of the genome as they have many co-binding sites and redistribute the occupancy of somatic cell-specific transcription factors such as Fra1 (Chronis et al., 2017). This redistribution results in a decrease in the expression of somatic cell-specific genes and a concomitant increase in pluripotency-associated gene expression. Generally and historically, the most studied initial somatic cell types during reprogramming are fibroblasts (Park et al., 2008a; Takahashi and Yamanaka, 2006; Takahashi et al., 2007). Silencing of the fibroblast transcriptional program during the initial stage of reprogramming is accompanied by a mesenchymal-to-epithelial transition (MET) (Pei et al., 2019).

The duration between MET and activation of the pluripotency circuit is called maturation stage of reprogramming (Samavarchi-Tehrani et al., 2010). In between initial and maturation stages, reprogramming cells become more prone to cellular senescence and proliferation slows down (Aarts et al., 2017). MYC overexpression along with OSK, provides reprogramming cells with enhanced potential to proliferate and overcome the senescence barrier (Aasen et al., 2008; Rand et al., 2018). Endogenous pluripotency circuit should be activated during maturation stage for successful pluripotency induction. OCT4, SOX2 and KLF4 cooperatively bind to their regulatory element targets to activate pluripotency-associated genes (Chronis et al., 2017). Additionally, KLF4 triggers paused polymerase II activity to express these genes and enables chromatin looping to initiate endogenous OCT4 expression (Nishimura et al., 2014; Wei et al., 2013). Reprogramming

cells start to express pluripotency markers such as alkaline phosphatase, TRA-1-60 and telomerase during maturation stage (Chan et al., 2009). Interestingly, a transient upregulation of genes related to embryonic lineages was observed which suggest that reprogramming to pluripotency traces back to development (Cacchiarelli et al., 2015; Takahashi et al., 2014).

The last stage of reprogramming is the maintenance stage where iPSC colonies emerge and grow in size. Exogenous transgenes are silenced and iPSCs should self-renew without OSKM transgenes (Papp and Plath, 2011). Rather than emergence of new colonies, during this stage, it is more essential for an already formed iPSC colony not to differentiate (Golipour et al., 2012). To achieve this goal, iPSCs need FGF2, feeder cells or a coating material such as matrigel to maintain the endogenous transcription program that keeps them in an undifferentiated state (Mossahebi-Mohammadi et al., 2020). Proper closure of maintenance stage requires extensive epigenetic resetting as failure in this stage would impair self-renewal and cause spontaneous differentiation (Buganim et al., 2013).

1.6.Cell cycle checkpoints and apoptosis as barriers to reprogramming

Reprogramming to pluripotency is a low efficiency, stochastic event, resulting a range of 0.01-1% iPSC colonies compared to the number of input somatic cells (Malik and Rao, 2013). There are cell intrinsic barriers to somatic cell reprogramming and overcoming these barriers increases the number of iPSC colonies generated (Table 1.1). During reprogramming, OSKM induces senescence transiently and re-gain of proliferation is necessary for iPSC generation. That's why, cell cycle checkpoints and senescence have been studied by using cellular reprogramming as a model approach. Inhibiting cell cycle checkpoints during somatic cell reprogramming increases reprogramming efficiency and their interaction with one another, transcription factors and apoptotic pathways have been studied (Kim et al., 2018; Rand et al., 2018; Wang et al., 2020; Zviran et al., 2019). To be more specific, Myc induces a transient apoptotic time-frame during reprogramming and genetic suppression of BAK and BAX increases OSKM-induced reprogramming efficiency through eliminating apoptosis, a major barrier to successful iPSC generation (Kim et al., 2018). However, when MYC is eliminated from reprogramming factors, deleting BAK and BAX did not increase reprogramming efficiency (Kim et al., 2018). Similarly, overexpressing a transcription factor, Tfap2c, rescued OSKM-induced reprogramming cells from apoptosis and, thereby, increased reprogramming efficiency (Wang et al., 2020). Although MYC has been shown to induce a transient apoptosis during reprogramming, its inclusion to reprogramming factors

increases reprogramming efficiency as it positively regulates biosynthesis, codon usage, alternative splicing and proliferation (Hirsch et al., 2015; Rand et al., 2018; Takahashi and Yamanaka, 2006; Zviran et al., 2019). During OSK-induced reprogramming, early-iPSC-track cells determined by expression of ESRG and TRA-1-60 undergo proliferation pause caused by p16 and p21 induced activation of RB. Importantly, OSKM-induced reprogramming downregulates p16 and p21 which results in reduced RB activity that rescues reprogramming trajectory cells from proliferation pause (Rand et al., 2018). Overall, these studies show that cell cycle checkpoints and MYC-induced apoptosis are barriers to reprogramming and their counter-interaction can be investigated by utilizing reprogramming to pluripotency as a model system.

1.7.The role of pluripotency maintenance factors in reprogramming

Reprogramming enables identification of pluripotency-maintaining factors, as their inhibition and activation decreases and increases reprogramming efficiency, respectively. In the breakthrough discovery of Takahashi and Yamanaka, they selected a set of transcription factors to induce pluripotency based on whether they were required for either maintenance, proliferation of embryonic stem cells or ESC-specific expression pattern (Takahashi and Yamanaka, 2006). Subsequent work showed that an alternative highly ESC-specific isoform of FOXP1 regulates expression of pluripotency-specific genes. Inhibiting the formation of this isoform through siRNA mediated targeting of alternatively included exon impaired reprogramming (Gabut et al., 2011). Interestingly, overexpressing a pluripotency marker, DPPA5, has been shown to increase reprogramming efficiency by increasing NANOG protein level in a post-transcriptional manner (Qian et al., 2016). Similarly, overexpressing ESC-specific markers Dppa2 and Dppa4 along with OSKM efficiently downregulate somatic genes and activate pluripotency enhancers which make iPSC generation possible even at 2-4 days of reprogramming and overall efficiency about 80% (Hernandez et al., 2018). Taken together these studies show that somatic cell reprogramming to pluripotency is a useful model to identify pluripotency-maintaining factors' role in pluripotency induction.

1.8.Chromatin-based regulation of gene expression

Nucleosome is composed of histone octamers by double of hetero dimers; H2A-H2B, H3-H4 and 146-bp DNA wrapped around this octamer (Kornberg, 1977). Nucleosome units structure into euchromatin and heterochromatin, depending on loose or compact organization, respectively (Tamaru, 2010). Euchromatin is associated with active transcription, whereas, heterochromatin is associated with gene silencing

(Huisinga et al., 2006). General and tissue-specific transcription factors are recruited to regulatory cis-regulatory elements such as promoters and enhancers by interacting with specific histone readers, transcriptional co-activators, post-translational modifications of histone tails and DNA structure/motif/methylation in a hierarchical manner (Calo and Wysocka, 2013; Heinz et al., 2015; Jin et al., 2011; Zhang et al., 2015). This recruitment initiates and enables elongation of RNA polymerase-dependent transcription of house-keeping and cell type-specific genes to ensure cell viability and identity (Catarino and Stark, 2018; Haberle and Stark, 2018; Panigrahi and O'Malley, 2021).

Enhancers are cis-regulatory regions that loop into proximity of promoters to activate gene expression (Pennacchio et al., 2013). They can be identified by certain features such as H3K27ac, H3K4me1, DNase hypersensitivity and p300 binding. H3K27ac and bi-directional enhancer RNAs classify a region as super-enhancer, clustered enhancers that are responsible to regulate the expression of cell identity genes (Hamdan and Johnsen, 2018; HniszBrian et al., 2013; Spicuglia and Vanhille, 2012). However, these features alone do not guarantee a functional enhancer in a cell type, further confirmatory analysis including high-throughput identification by self-transcribing active regulatory region sequencing (STARR-Seq) or individual reporter assays should be performed to ensure putative enhancer's activity (Klein et al., 2020; Kvon, 2015).

Histone writers such as methyltransferases and acetyltransferases add methyl and acetyl marks, respectively, on histone tails (Bannister and Kouzarides, 2011; Miller and Grant, 2013). Histone erasers such as demethylases and deacetylases remove these marks and histone readers such as PHD, chromo, WD40, Tudor, MBT, PWWP domain and bromodomain containing proteins bind to these modified histone tails (Hyun et al., 2017; Musselman et al., 2012). Along with the aforementioned role of H3K4me1 and H3K27ac modifications on activating gene expression through enhancers, there are other histone modifications that are shown to be important in gene expression regulation or elongation (Dong and Weng, 2013; Karlić et al., 2010; Kimura, 2013). H3K4me3 and H3K27ac marks near TSS define active promoters whereas H3K4me3 and H3K27me3 near TSS define bivalent promoters (Bae and Lesch, 2020; Voigt et al., 2013). Methylations of H3K9 compacts nucleosomes to form heterochromatin that silences gene expression (Allshire and Madhani, 2018; Wang et al., 2016). On the other hand, methylations of H3K36 and H3K79 are observed along actively transcribed gene bodies and implicated in transcription-coupled RNA splicing and RNA-polymerase elongation (Kim et al.,

2012; Wagner and Carpenter, 2012; Woo et al., 2017; Wozniak and Strahl, 2014). Histone acetylations are associated with euchromatin and active gene expression as negative charge on phosphate groups of DNA repels acetylated histones to enable TF binding (Bannister and Kouzarides, 2011; Blum, 2015). These studies showed that a subset of post-translational histone modifications regulate gene expression.

1.9.Molecular mechanisms governing somatic cell reprogramming at chromatin level

Changes in transcriptome and epigenome make cell fate transition possible (Chronis et al., 2017; Hanna et al., 2008; Knaupp et al., 2017; Lowry et al., 2008; Vierbuchen et al., 2010; Zviran et al., 2019). During the early phase of reprogramming, genes responsible for somatic cell identity are silenced and mesenchymal to epithelial transition (MET) takes place (Hanna et al., 2008; Pei et al., 2019; Wang et al., 2020). Genes responsible for maintaining pluripotency get activated during the transition of maturation to maintenance stages. All these expression changes occur through alterations of post-translational modification state of histone proteins. While H3K4me2 mark decreased at EMT genes, H3K4me1, H3K4me2 and H3K4me3 marks are increased at proliferation, metabolism and pluripotency genes during the early phase of reprogramming (González and Huangfu, 2016; Koche et al., 2011). H3K27me3 levels retain globally but are depleted from sites where H3K4 is methylated (Koche et al., 2011). ATAC-Seq during reprogramming classified chromatin states as closed-to-open and open-to-closed, of which regulatory sites were associated with pluripotency and somatic genes, respectively (Li et al., 2017). Importantly, this study showed that closed-to-open chromatin state regulatory elements contain OSK motif whereas open-to-closed elements do not. Authors proposed Sin3A corepressor complex mediated activity in coordinating open-to-closed chromatin state during reprogramming by reduced H3K27ac levels at these loci of somatic genes (Li et al., 2017). These studies pinpoint extensive chromatin remodeling induced by OSKM expression during somatic cell reprogramming to achieve pluripotency induction.

Demethylation of repressive histone marks on regulatory elements of epithelial and pluripotency related genes is necessary for iPSC generation (Table 1.1). Repression of miR-132 and miR-212 increases reprogramming efficiency and the activity of p300 and KDM5A, two targets of these miRNAs, are necessary for iPSC generation (Pfaff et al., 2017). H3K27 demethylase Utx has been shown to cooperate with OSK to open pluripotency-circuit and its depletion resulted in inefficient reprogramming as a result of

inactive H3K27me3 deposition at regulatory regions of unsuccessfully derepressed pluripotency genes (Jiang et al., 2020; Mansour et al., 2012). Ascorbic acid facilitates reprogramming via activating Jhdm1a/1b induced H3K36me2/3 demethylation which results in subsequent changes to favor pluripotency (Wang et al., 2011). Similarly, demethylation of H3K9me3 and H3K36me3 by KDM4B has been shown to increase reprogramming efficiency (Wei et al., 2017). Another demethylase of H3K27me3, JMJD3, has been shown to positively regulate reprogramming in the same manner as previously discussed demethylases but at the same time, it impairs reprogramming by inducing Ink4a, a pro-senescence factor, expression and degrading PHF20, a pluripotency regulator (Huang et al., 2020).

Table 1.1. Examples of chromatin modifiers and mechanism of action to regulate reprogramming

Chromatin Modifier	Species	Method	Change in efficiency	Mechanism of action	Reference
KDM6A	mouse, human	KO, shRNA	decrease	Interact with OSK to open pluripotency-circuit	(Mansour et al., 2012)
KDM2A, KDM2B	mouse	OE	increase	Suppresses senescence, enhances proliferation, activate microRNA 302/367 cluster	(Wang et al., 2011)
KDM4B	mouse	OE	increase	Decreases H3K9me3 and H3K36me3 levels	(Wei et al., 2017)
macro-H2A	mouse	shRNA	increase	Occupies repressed pluripotency genes in somatic cells	(Pasque et al., 2012)
APLF	mouse	shRNA	increase	Suppresses MET by depositing macroH2A on *Cdh1* cis-regulatory regions	(Syed et al., 2016)
Chaf1a, Chaf1b	mouse	RNAi	increase	Impairs Sox2 binding to enhancers to open and	(Cheloufi et al., 2015)

				upregulate pluripotency-circuit	
SSRP1, SUPT16H	human	shRNA	increase	FACT subunits maintain somatic cell-specific gene expression and reprogramming barriers such as PRRX1	(Kolundzic et al., 2018)
DOT1L	human, mouse	shRNA, inhibitor, KO	increase	Maintains the expression of somatic cell-specific genes by methylation of H3K79	(Onder et al., 2012)
AF10	human	CRISPR-Cas9, shRNA	increase	Interacts with DOT1L to maintain H3K79 methylation and somatic cell-specific gene expression	(Uğurlu-Çimen et al., 2021)
RPAP1	mouse	shRNA	increase	Interacts with mediator and RNAPII to maintains somatic cell-specific transcription	(Lynch et al., 2018)
Ubc9, SUMO	human, mouse	shRNA, inhibitor	increase	Enables somatic TFs bind to somatic enhancers to maintain somatic cell-specific transcription	(Cossec et al., 2018)
BET domain	human	inhibitor	increase	Safeguards somatic cell identity by maintaining related transcription	(Shao et al., 2016)

Like H3K27me3, macro-H2A was shown to occupy cis-regulatory elements of pluripotency-specific genes in somatic cells and depleted from these regions in pluripotent stem cells. Removing this histone variant increases the efficiency of inducing

pluripotency immensely (Pasque et al., 2012). Suppressing the expression of a histone chaperone, APLF, increases reprogramming efficiency and facilitates MET by upregulating CDH1, NANOG and KLF4 through depleting macroH2A from promoters and depositing H3K4me2 to the promoters of associated genes (Syed et al., 2016). Depleting subunits of another histone chaperon complex CAF-1 increased reprogramming efficiency by enabling more chromatin accessibility at enhancers, increasing Sox2 binding to pluripotency genes to upregulate them (Cheloufi et al., 2015). These results showed that the interplay between histone variants and histone chaperones dictates chromatin accessibility, histone modifications, TF binding and gene expression to orchestrate pluripotency induction.

1.10. Epigenetic factors safeguarding somatic cell identity

Somatic cell reprogramming to pluripotency is a well-established model to study factors that are important in safeguarding somatic cell identity as their inhibition increases reprogramming efficiency (Arabacı et al., 2020; Brumbaugh et al., 2019; Nashun et al., 2015) (Table 1.1). Previous studies showed that inhibiting FACT, DOT1L, RPAP1, SUMO and BRD4 facilitates reprogramming by downregulating somatic cell-specific genes (Cheloufi et al., 2015; Cossec et al., 2018; Kolundzic et al., 2018; Lynch et al., 2018; Onder et al., 2012; Shao et al., 2016). Specifically, knocking down FACT subunits, SSRP1 and SUPT16H increased human somatic cell reprogramming to both iPSC and induced neurons through downregulating somatic genes and reprogramming barriers such as PRRX1 (Kolundzic et al., 2018; Yang et al., 2011). Importantly, the function of FACT to safeguard cell identity seems to be conserved from *C. Elegans* to human and reprogramming to pluripotency serves as a nice model to study cell identity.

Knocking down or inhibiting the catalytic activity of only H3K79 methyltransferase DOT1L increases reprogramming efficiency in a wide range of systems (Kim et al., 2021; Onder et al., 2012; Tao et al., 2017; Tran et al., 2019; Wang et al., 2019a). This enzyme maintains the expression of key somatic cell-specific transcription factors and continuous expression of these transcription factors impairs pluripotency induction. A follow-up study showed that interaction of DOT1L with AF10 mediates somatic cell-specific gene expression and inhibiting AF10 increases reprogramming efficiency (Uğurlu-Çimen et al., 2021). These studies highlight DOT1L and related H3K79 methylation as epigenetic regulators of maintaining somatic cell-specific gene expression by using reprogramming to pluripotency as an *in vitro* model.

Knocking down RNA polymerase II associated protein 1 (RPAP1) specifically

decreased murine somatic cell proliferation and led them to apoptosis (Lynch et al., 2018). RPAP1 knockdown downregulated genes belonging to EMT, development and morphogenesis in MEFs and increased various combinations of OSKM-induced reprogramming efficiency. Mechanistically, RPAP1 is in a functional complex with Mediator and RNA Pol II regulatory axis and depleting RPAP1 impairs interaction between Mediator and RNA Pol II to downregulate somatic cell-specific genes (Lynch et al., 2018). Interestingly, the de-differentiation achieved by RPAP1 loss was so intense that by overexpressing Klf4 and Myc, fibroblasts started to colonialize and express pre-iPSC markers which shows the efficiency of reprogramming to pluripotency is achievable with less reprogramming factors when somatic cell identity is already impaired.

Chromatin immunoprecipitation (ChIP) profiling of SUMO1 and SUMO2 exhibited differential binding of these histone writers on regulatory sites of somatic cells and pluripotent stem cells (Cossec et al., 2018). Knocking down SUMO E2 ligase, Ubc9, and transient inhibition of SUMO E1 ligase by highly specific inhibitor, ML-792, enhanced both mouse and human somatic cell reprogramming efficiency to pluripotency. Mechanistically, they showed that loss of histone sumoylation led to decrease in OSK and TF binding to somatic enhancers and downregulation of somatic-specific genes during reprogramming. Importantly, the role of SUMO in resisting cell-fate conversion was not limited to reprogramming to pluripotency but it additionally applied to conversions of fibroblasts to neurons, pre-B cells to macrophages, leukemic promyelocytes to granulocytes and ESCs to 2C-like cells (Cossec et al., 2018). This study shows that impairing initial cell identity enhances reprogramming potential and reprogramming of different lineages to others serves as a nice model to show generalized function of a cell identity maintaining factor's role in cell fate conversions.

Transiently inhibiting the bromodomain and extra-terminal (BET) domains by JQ1 (50 nM), iBET-151 (100 nM) and CPI203 (50 nM) during reprogramming greatly increased the number of iPSC colonies (Shao et al., 2016). The mechanism proposed by authors to explain the increase in reprogramming efficiency upon BET inhibition was based on both gene expression profile of fibroblasts and reprogramming cells treated with BET inhibitor and morphological assessment. 2 day BET inhibition downregulated somatic cell-specific genes in fibroblasts and reprogramming cells and 5 day BET inhibition of fibroblasts resulted in more round or polygonal shape cell phenotype rather than conventional stretched and long fibroblast morphology (Shao et al., 2016). This study, as well as others discussed above, showed that epigenetic factors that safeguards

cell identity are associated with active transcription in somatic cells and their inhibition results in more efficient reprogramming through facilitated closure of somatic transcriptional program (Table 1.1).

1.11. Genome-wide and focused screens to identify reprogramming barriers

Several screening approaches are carried out to identify barriers to reprogramming in mammalian cells. Genome-wide RNAi approach identified genes related to ubiquitination (UBE2E, UBE2D and RNF40), ADAM family (ADAM7, ADAM21 and ADAM29), endocytosis (DRAM1 and SLC17A5), chromatin regulators (ATF7IP and MED19) and transcription factors (TTF2 and TMF1) as reprogramming barriers in human fibroblasts (Qin et al., 2014). Another genome-wide shRNA screen revealed Tfdp1, Gtf2e1, Nfe2, Foxn3, Erf, Cdkn2aip, Msx3, Ssbp3, Dbx1, Hoxd4, Lzts1, Arx, Hoxd12, Gtf2i, Ankrd22 and Hoxc10 as novel barriers to reprogramming at its early stage in mouse (Yang et al., 2014). The protein modifier SUMO2 is another validated top hit that blocks reprogramming from a serial genome-wide shRNA screen in mouse (Borkent et al., 2016). SMAD3 (TGF-β pathway), ZMYM2 (epigenetic modifier), SFRS11 (splicing factor), SAE1 (Sumo-activating enzyme) and ESET (H3K9 methyltransferase) are barriers to human reprogramming uncovered by a genome-wide siRNA screen (Toh et al., 2016). A genome-wide CRISPR-UMI screen and revealed novel human somatic cell reprogramming repressors; Pias1, an E3 SUMO-protein ligase and Men1, transcriptional cofactor (Michlits et al., 2017). These genome-wide screens enabled the identification of various biological processes as barriers to reprogramming.

Various screens have focused on epigenetics rather than genome-wide approach to understand the mechanism of reprogramming through RNA regulation or chromatin modifications. miRNA-212/132, identified by the miRNA screen, is a barrier to reprogramming in mouse (Pfaff et al., 2017). A focused shRNA screen targeting DNA or histone methylation pathways identified H3K79 methyltransferase DOT1L, H3K9 methyltransferase SUV39H1 and context-dependent transcriptional activator or repressor YY1 (Onder et al., 2012). A broader shRNA screen targeting both known and predicted chromatin regulators revealed histone chaperone CAF-1 as a barrier to reprogramming by safeguarding somatic cell identity (Cheloufi et al., 2015). Another shRNA screen targeting epigenetic modifiers identified TRIM28 and SETDB1 which establish H3K9me3 to maintain somatic cell identity as novel barriers to reprogramming (Miles et al., 2017). shRNA screen targeting mouse USP family genes uncovered USP26 as negative regulator of somatic cell reprogramming (Ning et al., 2017). These epigenetics-

focused screens have highlighted the importance of chromatin based regulation of somatic cell reprogramming to pluripotency.

In addition to genetic screens, a number of chemical compound screens have been carried out to identify molecules that increase reprogramming efficiency (Hou et al., 2013; Zhao et al., 2015). 5′-azacytidine (DNA methyltransferase inhibitor) and HDAC inhibitors; suberoylanilide hydroxamic acid (SAHA), trichostatin A (TSA) and VPA were found to greatly enhance mouse reprogramming efficiency through a chromatin focused compound screening using *Oct4*-GFP transgenic reporter (Huangfu et al., 2008). A kinase inhibitor screen identified compounds that inhibit p38, inositol trisphosphate 3-kinase and Aurora A kinase as enhancers of reprogramming efficiency (Li and Rana, 2012). Epiblastin A was found as an enhancer of mouse EpiSC reprogramming to naïve pluripotent stem cells by screening LOPAC library (Ursu et al., 2016). These inhibitor screens identified several reprogramming enhancing compounds.

Compound screening approach enabled identification of compounds that could enable somatic cell reprogramming to iPSCs with less OSKM factors. For example, screening drugs during OK overexpression-mediated reprogramming of MEFs identified BIX-01294 (G9a histone methyltransferase inhibitor) and Bayk8644 (L-channel calcium agonist) as substitutes for Sox2 expression (Shi et al., 2008). Similarly, a compound screen identified PS48 (phosphoinositide dependent kinase 1 (PDK1) activator) and sodium butyrate (histone deacetylase inhibitor) as boosters of neonatal human epidermal keratinocyte (NHEK) reprogramming to iPSCs by OK overexpression (Zhu et al., 2010). Importantly, combining these small molecules with MAPK/ERK and TGF-β inhibitors enabled reprogramming without SKM (Zhu et al., 2010). A small molecule screen performed during *Oct4*-GFP MEF reprogramming to iPSCs found two TGF-β receptor 1 kinase inhibitors, E-616452 (RepSox) and E-616451 as functional substitutes of Sox2 (Ichida et al., 2009). Subsequent work found that iPYrazine replaced Sox2 from a high-throughput compound screen during reprogramming MEF harboring the firefly luciferase gene in the Nanog locus to iPSCs by overexpressing OKM (Staerk et al., 2011). Further work identified iPYrazine as a pan-Src family kinase inhibitor and other pan-Src family kinase inhibitors dasatinib and PP1 could replace Sox2 as well (Staerk et al., 2011). The same compound library was used to reveal Kenpaullone as a functional replacement of Klf4 during OSM-mediated reprogramming of Nanog-luciferase MEFs to iPSCs (Lyssiotis et al., 2009). These screens enabled identification of small molecules that could functionally replace some of OSKM factors and paved the way to achieve chemically

induced pluripotent stem cells.

1.12. Species-specific regulation of pluripotency induction.

As OSKM could reprogram both murine and human somatic cells to iPSCs, it was long believed that mechanisms governing this process is similar in both species. However, the resulting mouse and human iPSCs differ in terms of their pluripotency states (Takahashi et al., 2018; Weinberger et al., 2016). Mouse iPSCs have been shown to be closer to naïve state as they form round colonies, contribute to chimeric embryos and reactivate the inactive X chromosome. However, human iPSCs with conventional reprogramming resemble primed pluripotent stem cells; they form flat colonies, contribute little to chimeric embryo when transplanted to other mammalian blastocysts and do not reactivate their inactive X chromosomes (Kilens et al., 2018). However, recent advances in culture conditions enable naïve iPSC line derivation from human somatic cells (Bredenkamp et al., 2019; Kilens et al., 2018; Lynch et al., 2020). There are key differences observed at the chromatin level during reprogramming between mouse and human (Wang et al., 2018b). Reprogramming to pluripotency is shorter in mice (14 days) compared to human (21 days) and during reprogramming, OSK factors bind to more closed regions in human somatic cells compared to mouse (Fu et al., 2018). Furthermore, POU5F1 (OCT4) cannot be substituted with other POU family proteins in mouse somatic cell reprogramming, however, in human somatic cell reprogramming, OCT6 could replace OCT4 (Kim et al., 2021). In the presence of a cocktail of small molecules that include SGC0946 (DOT1L inhibitor), SB431542 (TGF-β inhibitor) and RN-1 (LSD1 inhibitor), most of POU families and a set of TFs enable human iPSC generation along with retroviral SKM overexpression (Kim et al., 2021). Removing Oct4 from murine somatic cell reprogramming factors and even knocking out Oct4 does not prevent iPSC generation but increases developmental potential of mouse iPSCs, revealing another species-specific reprogramming route as human iPSC generation is not possible with only lentiviral SKM overexpression (Velychko et al., 2019). These studies highlight differences in chromatin-based regulation between mouse and human during pluripotency acquisition probably owing to the different pluripotent states achieved by reprogramming species' somatic cells.

1.13. Mouse chemically induced pluripotent stem cells

In 2013, mouse embryonic fibroblasts (MEFs) were successfully reprogrammed to chemically induced pluripotent stem cells, albeit with a lower efficiency than conventional OSKM reprogramming (Hou et al., 2013). In this study, Hou and colleagues

first eliminated the need for SKM in reprogramming with VC6T (VPA, CHIR99021, 616452 and Tranylcypromine). They screened for chemicals that can substitute Oct4 in reprogramming by using Oct4 promoter-driven GFP expressing MEFs. Forskolin (F) is identified as one of the substitutes for transgene Oct4 from the screen and added to the chemical cocktail. Treatment of MEFs with VC6TF for the first 20 days and adding DZNep to the cocktail until 36 days formed epithelioid and Oct4-GFP+ colonies. After switching to dual inhibition (2i) medium to inhibit mitogen-activated protein kinase (MAPK) signaling and glycogen synthase kinase-3 (GSK3), some Oct4-GFP+ colonies developed embryonic stem cell (ESC)-like colonies (Hou et al., 2013).

Table 1.2 Small molecules and their targets at different stages of reprogramming to mouse chemically induced pluripotent stem cells (Zhao et al., 2015).

Stages and duration (cocktail)	**Compound**	**Target**
Stage 1 - 16 days (VC6TFEA)	Valproic Acid (V)	HDAC
	CHIR99021 (C)	GSK3β
	616452 (6)	ALK5
	Tranylcypromine (T)	LSD1
	Forskolin (F)	adenylate cyclase (activator)
	EPZ004777 (E)	DOT1L
	AM580 (A)	RAR (agonist)
Stage 2 – 12 days (VC6TFZAS5)	DZNep (Z)	S-adenosylhomocysteine hydrolase
	SGC0946 (S)	DOT1L
	5-aza-dC	DNMTs
Stage 3 – 12 days (N2B27-2i-LIF)	PD0325901 CHIR99021 (2i)	ERK GSK3β

To increase the efficiency of chemically induced pluripotent stem cell (ciPSC) generation, another study came up with three different small molecule cocktails to treat MEFs at three different stages of reprogramming (Zhao et al., 2015). Interestingly, they added two different DOT1L inhibitors to two cocktails (Table 1.2). However, these small molecules did not induce pluripotency from human somatic cells indicating different mechanisms taking place during cell type conversion between these two species.

1.14. Chromatin focused small molecule screen during human somatic cell

reprogramming to pluripotency

To identify additional chromatin-associated reprogramming barriers, Dr. Ayyub Ebrahimi performed a small molecule screen during the initial stage of human somatic cell reprogramming (Ebrahimi et al., 2019). This screen was performed in DOT1L inhibitor background with inhibitors targeting histone writers, readers and erasers (Figure 1.2). Choosing the initial stage of reprogramming for treatments enabled us to study chromatin factors that are important in maintaining somatic cell identity because inhibitors that increase reprogramming efficiency may facilitate the silencing of somatic cell-specific gene expression. Importantly, this screen identified several known reprogramming enhancing small molecules such as CHIR99021, valproic acid and MS-275. Several novel small molecules from this screen such as SGC-CBP30, I-CBP112 and LP99 were further validated as barriers to human somatic cell reprogramming to pluripotency (Figure 1.3). SGC-CBP30 and I-CBP112 are structurally different bromodomain inhibitors of EP300 and CREBBP and LP99 is BRD9 bromodomain inhibitor.

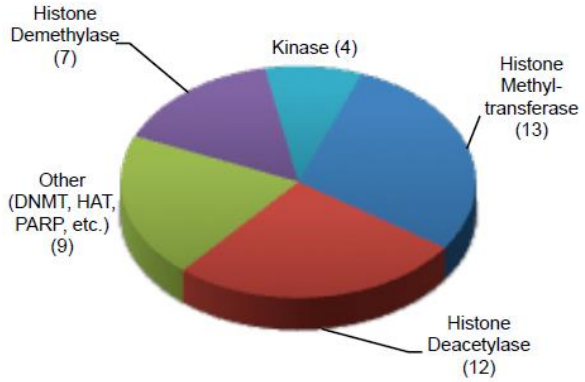


Figure 1.2. Targets of small molecules used to screen during reprogramming

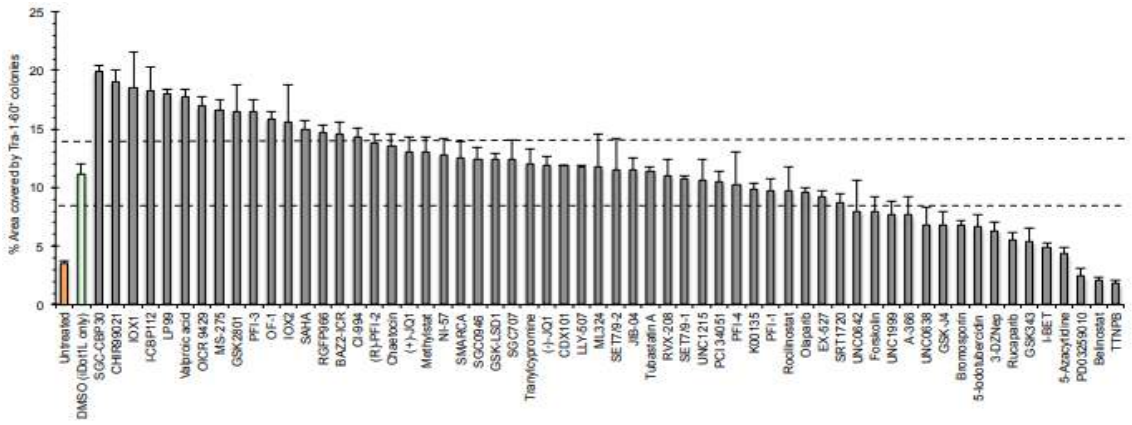


Figure 1.3. Small molecule screen identified CREBBP/EP300 and BRD9 bromodoamin inhibitors as barriers to reprogramming, adapted from (Ebrahimi et al., 2019).

1.15. cAMP-response element binding protein (CREB) binding protein (CREBBP), E1A binding protein p300 (EP300) and their roles in gene expression

EP300 and *CREBBP* genes code for transcriptional coactivators that can both read and write H3K27ac mark (Attar and Kurdistani, 2017; Bedford and Brindle, 2012). These two large size paralog proteins belong to KAT3 family and share 63% homology with matching conserved functional domains (Iyer et al., 2004; Wang et al., 2013). They contain a central chromatin association and modification module and transactivation domains (Wang et al., 2013). Bromodomain enables these proteins to bind to acetylated histones and transcription factors to tether them on specific regulatory loci. PHD finger and KAT11 domains enable CREBBP/EP300 to recognize and acetylate target substrates including themselves (Rack et al., 2014). Transcriptional-adaptor zinc-finger (TAZ) and kinase inducible domain of CREB interacting (KIX) domains are important in CREBBP/EP300-protein interactions including CREB (Bedford and Brindle, 2012).

EP300 and CREBBP are shown to bind lineage-specific enhancers, thus, enhancing the expression of genes that are responsible for cell identity including pluripotency (Fauquier et al., 2018; Garcia-Carpizo et al., 2018; Hennig et al., 2013; Lai et al., 2017; Lim et al., 2021; Martire et al., 2020; Wang et al., 2021; Zhang et al., 2021b). During reprogramming of mouse embryonic fibroblast (MEF) into iPSCs, Chronis et al. observed that there is no change in EP300 expression level but binding of EP300 is lower on somatic enhancers dependent on less binding of somatic TFs on these regions and independent of OSK binding as early as 48 hours (Chronis et al., 2017). Surprisingly, Pfaff et al. observed that knocking down EP300 expression decreases reprogramming efficiency of MEFs (Pfaff et al., 2017). However, the role of these chromatin factors was not investigated in the context of human somatic cell reprogramming until recently, which is in part included in this thesis (Ebrahimi et al., 2019).

1.16. Bromodomain containing-9 (BRD9)

Bromodomain Containing 9 (BRD9) is the specific subunit of non-canonical BRG1-associated factor (ncBAF) complex and it can directly bind to acetyl-, propionyl- or butyryl-lysines of histones (Flynn et al., 2015). It belongs to the family IV bromodomain-containing proteins. Although less is known about the biological ligands of BRD9 bromodomain, it has been shown to recognize the doubly acetylated histone

H4K5acK8ac and butyryllated H4K5buK8bu ligand with a similar affinity (Lloyd and Glass, 2018). Its close homolog BRD7 is a subunit of polybromo-associated BAF (PBAF) complex. BRD7 has a weak binding affinity to five acetylated histone ligands including; H3K9ac, H3K14ac, H4K8ac, H4K12ac and H4K16ac15 (Sun et al., 2007). BRD9 and BRD7 are able to recognize the di-propionylated ligand H4K5prK8pr and butyryllated lysine side chain in histone H4K5buK8bu (Flynn et al., 2015). In addition, BRD7 binds and regulates non-histone proteins like BRCA1 and p53 through its bromodomain. BRD9 and BRD7 have been shown to bind to K91ac of VDR (vitamin D receptor) and their differential binding brings ncBAF or PBAF, respectively, to promote β cell survival for the latter case (Huddy et al., 2018). BRD9 is shown to co-localize with H3K27ac, H3K4me1, H3K4me3 and CTCFs in a sarcoma cell line indicating that ncBAF complex occupy active enhancer and promoters (Michel et al., 2018).

1.17. SWItch/Sucrose Non–Fermentable (SWI/SNF) complexes

BAF and PBAF are two broad mammalian counterparts of SWI/SNF (SWItch/Sucrose Non-Fermentable) complexes discovered through genetic studies in yeast. SWI/SNF complexes are composed of 15 subunits encoded by 29 genes (Alfert et al., 2019). They can remodel chromatin by hydrolyzing ATP. Unlike other chromatin remodeling complexes such as ISWI, NuRD/Mi-2/CHD, INO80 and SWR1, SWI/SNF complexes are well-characterized due to their earlier discovery and association with cancer (Hargreaves and Crabtree, 2011). SWI/SNF shows tissue-specific composition of subunits and this feature has been shown to be important in development, tissue-type specific cancers and somatic cell reprogramming (Alver et al., 2017; Innis and Cabot, 2020; Masliah-Planchon et al., 2015).

In 2018, by combining gradient centrifugation and immunoprecipitation, a study revealed another SWI/SNF complex, GBAF (GLTSCR1 containing BAF) or ncBAF (Alpsoy and Dykhuizen, 2018). All SWI/SNF complexes contain a core and an ATPase module (Figure 1.4). To form the SWI/SNF core complexes, firstly SMARCC dimer is formed (Mashtalir et al., 2018). There are two proteins to occupy this space: SMARCC1 and SMARCC2. Initial BAF core is formed by the addition of one of SMARCD1/2/3. SMARCE1 and SMARCB1 join the initial BAF core to result in either BAF or PBAF complex. For ncBAF, these positions are occupied by GLTSCR1/L protein. BRD9 is another specific subunit of ncBAF to join the core module. ARID1A/B and DPF1/2/3 are the specific subunits of core module of BAF complex whereas ARID2, BRD7 and PH10 are the specific subunits of the PBAF core module. ATPase module is formed by

ACTL6A, β-actin, BCL7A/B/C and SMARCA2/4. For BAF and ncBAF complex formation, SS18/L1 also joins the ATPase module. This position is occupied by PBRM1 in PBAF complexes (Figure 1.4). The resulting SWI/SNF complexes differ in their molecular weights. ncBAF is the lightest with 0.87 MDa, PBAF is the heaviest with 1.41 MDa (Mashtalir et al., 2018).

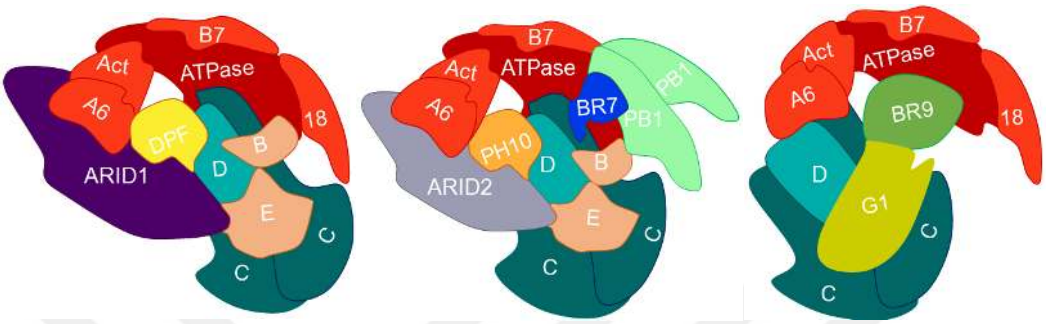


Figure 1.4. Differential organizations of SWI/SNF complex members generate BAF (left), PBAF (middle) and ncBAF (right) complexes (adapted from (Mashtalir et al., 2018)). Abbrevations: C: SMARCC1/C2, D: SMARCD1/D2/D3, B: SMARCB1, E: SMARCE1, ARID1: ARID1A/B, ARID2: ARID2, DPF: DPF1/2/3, PH10:PHF10, BR7: BRD7, ATPase: SMARCA2/4, G1: GLTSCR1/1L, BR9: BRD9, B7: BCL7A/B/C, 18: SS18/L1, A6: ACTL6A, Act: ß Actin, PB1: PBRM1

1.18. The role of SWI/SNF complexes in reprogramming and pluripotency

SWI/SNF complexes have been studied in mouse embryonic stem cells. Embronic stem cell specific BAF (esBAF) contains Smarca4, Smarcc1 and Smarcd1 and it lacks Smarca2, Smarcc2 and Smarcd3 (Alfert et al., 2019; Ho et al., 2009, 2011). Conditional deletion or knockdown of Smarca4 or Smarcc1 from mESCs significantly impairs proliferation and differentiation. Knocking out Arid1a or Arid1b impaired ES cell proliferation (Ho et al., 2009, 2011; Zhang et al., 2014). Smarcc1 regulates lineage specification of the mouse blastocyst (Panamarova et al., 2016). LIF/STAT signaling and Smarca4 are important for maintaining the pluripotency of mESCs by regulating common set of genes (Ohtsuka et al., 2015). Deleting Smarca4 closes previously accessible chromatin for Stat3 and mESCs lose their pluripotency (Ho et al., 2011). BAF complex can interact with p300 preferentially at distal lineage-specific enhancers rather than super-enhancers to regulate H3K27 acetylation. These distal enhancers are important in differentiation during development (Alver et al., 2017). These previous studies identified a role for a specific BAF complex, esBAF, in maintenance of pluripotency and

differentiation.

The role of ncBAF complex in mouse embryonic cell identity was investigated in a recent study (Gatchalian et al., 2018). shRNA-mediated BRD9 knockdown and BRD9 bromodomain inhibition by I-BRD9 and BI-9564 decreased naïve mESC cell survival by downregulating pluripotency-associated gene expression. This study revealed that BRD9 occupies distal sites and promoters with BRG1 and BRD9 is found on TAD (topologically associated domain) boundaries (Gatchalian et al., 2018). TADs are contact domains of chromatin formed by loop-extrusion mechanism of cohesin and CTCF interplay (Beagan and Phillips-Cremins, 2020). TAD boundaries act as insulators to maintain active/repressing histone marks within compartment and are generally associated with active gene expression (Szabo et al., 2019). BRD9 bromodomain inhibitors I-BRD9 and BI-9564 decreased the clonogenicity of mouse naïve embryonic stem cells and BRD9 has been shown to transiently interact with BRD4 (Gatchalian et al., 2018). However, I-BRD9 treatment reduced this interaction globally. JQ1 (BET inhibitor) treatment decreased BRD9 ChIP signals genome-wide whereas I-BRD9 treatment did not result in a significant change in BRD4 ChIP signals. These results identified a role for ncBAF complex in maintaining mouse naïve pluripotency through BRD4 bromodomain dependent localization of BRD9 and ncBAF complex on regulatory sites (Gatchalian et al., 2018).

First study investigating BAF complex in the context of reprogramming reported that Smarca4 and Smarcc1 overexpression can increase mouse reprogramming by facilitating enhanced Oct4 binding to target promoters (Singhal et al., 2010). However, a subsequent study showed that knocking down Smarca2 after OSKM induction increases murine somatic cell reprogramming efficiency (Jiang et al., 2015). The discrepancy between two studies may arise from the fact that Smarca4 is expressed in mESCs whereas Smarca2 is absent. If Smarca4 is overexpressed during reprogramming, BAF complex specific to mESC identity can be formed earlier. On the other hand, suppressing the expression of Smarca2 that is specific to somatic cells increases reprogramming (Jiang et al., 2015). However, the role of ncBAF complex in mouse and human somatic cell reprogramming and human pluripotency was not investigated until the recent preprint published as part of this thesis (Sevinç et al., 2021).

Chapter 2. MATERIAL & METHOD

2.1.Cloning

Specific shRNA oligonucleotides were designed to target genes as listed in Table 2.1. Each shRNA oligo was added to a 50 µL PCR reaction mixture (1X ThermoPol buffer, 2 mM Mg_2SO_4, 5% DMSO, 1 M betaine, 200 µM dNTPs, 0.1 µM template oligo, 0.5 µM common primer mix (Table 2.1), 1 unit VENT Polymerase) on ice. PCR was performed in a thermocycler with following steps: 94 ^{0}C for 5 min, 25 cycles of 94 ^{0}C for 30 sec, 54 ^{0}C for 30 sec, 75 ^{0}C for 30 sec and a final extension step at 75 ^{0}C for 2 min. Amplified DNA was purified by MN PCR-Clean up kit. pSMP vector (Addgene plasmid no. #36394) and amplified shRNA oligos were digested with XhoI and EcoRI at 37 ^{0}C for 2 hours. Digested products were run on 0.8% agarose gel, 0.1 and 6.5 kb bands corresponding to amplified shRNA oligo and digested pSMP vector, respectively, were purified from agarose gel using MN PCR-Clean up kit. Concentrations of purified DNAs were measured on nanodrop and set up in a 20 µL ligation mixture (50 ng digested pSMP vector, 2.3 ng digested shRNA oligo, 1 µL T4 ligase, 1X T4 ligase buffer). This ligation mixture was incubated at room temperature for 2 hours and 5 µL of it was used to transform chemically competent bacteria.

Table 2.1. List of shRNAs and sequences

shRNA	Sequence
shPRRX1 1	TGCTGTTGACAGTGAGCGCCCACTTGTTTCTGTGTGTGTTTAGTGAAGCCACAGATGTAAAC ACACACAGAAACAAGTGGTTGCCTACTGCCTCGGA
shPRRX1 2	TGCTGTTGACAGTGAGCGAGCTGCGTCTGGCTAATTCTAATAGTGAAGCCACAGATGTATT AGAATTAGCCAGACGCAGCCTGCCTACTGCCTCGGA
shPRRX1 3	TGCTGTTGACAGTGAGCGCGGGCTTTAAATGAATTAGAAATAGTGAAGCCACAGATGTATT TCTAATTCATTTAAAGCCCTTGCCTACTGCCTCGGA
shPRRX1 5	TGCTGTTGACAGTGAGCGAGCAGCGAAGGAATAGGACAATAGTGAAGCCACAGATGTATT GTCCTATTCCTTCGCTGCTTGCCTACTGCCTCGGA
shPRRX1 7	TGCTGTTGACAGTGAGCGGCTGATACATCTGATCTATCATAGTGAAGCCACAGATGTATGA TAGATCAGATGTATCAGCTGCCTACTGCCTCGGA
shPRRX1 8	TGCTGTTGACAGTGAGCGGAGAGCCATGCTAGCCAATAATAGTGAAGCCACAGATGTATTA TTGGCTAGCATGGCTCTCTGCCTACTGCCTCGGA
ACLY sh1	TGCTGTTGACAGTGAGCGCCCTCGAAGAAGCCAAATCTTATAGTGAAGCCACAGATGTATA AGATTTGGCTTCTTCGAGGTTGCCTACTGCCTCGGA
ACLY sh2	TGCTGTTGACAGTGAGCGAGGGAGGAAGCTGATGAATATATAGTGAAGCCACAGATGTAT ATATTCATCAGCTTCCTCCCGTGCCTACTGCCTCGGA
ACLY sh3	TGCTGTTGACAGTGAGCGCCAGCCAGAACTTGGTAGTCAATAGTGAAGCCACAGATGTATT GACTACCAAGTTCTGGCTGATGCCTACTGCCTCGGA

ACSS2 sh1	TGCTGTTGACAGTGAGCGCACGCTTTGAGACAACCTACTTTAGTGAAGCCACAGATGTAAA GTAGGTTGTCTCAAAGCGTTTGCCTACTGCCTCGGA
ACSS2 sh2	TGCTGTTGACAGTGAGCGACAGGATTGATGACATGCTCAATAGTGAAGCCACAGATGTATT GAGCATGTCATCAATCCTGCTGCCTACTGCCTCGGA
ACSS2 sh4	TGCTGTTGACAGTGAGCGCCCAAGTGTGTCAGTTCAGCAATAGTGAAGCCACAGATGTATT GCTGAACTGACACACTTGGATGCCTACTGCCTCGGA
BRD9 sh510	TGCTGTTGACAGTGAGCGCCCTGGATATTCAATGATAATATAGTGAAGCCACAGATGTATA TTATCATTGAATATCCAGGATGCCTACTGCCTCGGA
BRD9 sh561	TGCTGTTGACAGTGAGCGCGACAAAATTGTAGCTAATGAATAGTGAAGCCACAGATGTATT CATTAGCTACAATTTTGTCTTGCCTACTGCCTCGGA
BRD7 sh1	TGCTGTTGACAGTGAGCGGTACTAATGCCATGATTTATAGTGAAGCCACAGATGTATAAAT CATGGCATTAGTACTGCCTACTGCCTCGGA
BRD7 sh3	TGCTGTTGACAGTGAGCGCCAGAGACCATTTATTATATAGTGAAGCCACAGATGTATATAA TAAATGGTCTCTGGTGCCTACTGCCTCGGA
BRD7 sh4	TGCTGTTGACAGTGAGCGGCACGTATGGAGTTCGAAATAGTGAAGCCACAGATGTATTTCG AACTCCATACGTGCTGCCTACTGCCTCGGA
shFF	TGCTGTTGACAGTGAGCGCCCGCCTGAAGTCTCTGATTAATAGTGAAGCCACAGATGTATT AATCAGAGACTTCAGGCGGTTGCCTACTGCCTCGGA
Common-fw	GATGGCTGCTCGAGAAGGTATATTGCTGTTGACAGTGAGCG
Common-rev	GTCTAGAGGAATTCCGAGGCAGTAGGCA

For sgRNA cloning, lentiCRISPR v2 (Addgene plasmid no. #52961) vector was digested with BsmBI enzyme and gel purified with MN PCR-Clean up Kit. Top and bottom guide RNAs (Table 2.2) designed to specifically target genes were annealed with T4 Polynucleotide Kinase (3' phosphatase minus, NEB) and T4 DNA Ligase Buffer (NEB) 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 0.1°C/second. Annealed sgRNA oligos were diluted 1:200 and were used as inserts for ligating to 50 ng digested lentiCRISPR v2. 5 µL of 20 µL ligation reaction mix was used to transform chemically competent bacteria.

Table 2.2. List of sgRNAs and sequences

sgRNA	Sequence
BRD7 sg801 top	CACCGAACTGATGAGACAATTGCAG
BRD7 sg801 bottom	AAACCTGCAATTGTCTCATCAGTTC
BRD7 sg802 top	CACCGAGCTCTTTAGCCAAACAAGA
BRD7 sg802 bottom	AAACTCTTGTTTGGCTAAAGAGCTC
BRD7 sg804 top	CACCGAGGCAAGTCTAATCTCACAG
BRD7 sg804 bottom	AAACCTGTGAGATTAGACTTGCCTC
BRD9 sg821 top	CACCGACTCCAGTTACTATGATGAC
BRD9 sg821 bottom	AAACGTCATCATAGTAACTGGAGTC
BRD9 sg822 top	CACCGAGAGAGGGAGCACTGTGACA
BRD9 sg822 bottom	AAACTGTCACAGTGCTCCCTCTCTC
BRD9 sg823 top	CACCGAGATACCGTGTACTACAAGT
BRD9 sg823 bottom	AAACACTTGTAGTACACGGTATCTC

SMARCC1 sg5791 top	CACCGAGATATCATCAAACGACATC
SMARCC1 sg5791 bottom	AAACGATGTCGTTTGATGATATCTC
SMARCC1 sg5792 top	CACCGAGCCCCAAGAATGTGACATA
SMARCC1 sg5792 bottom	AAACTATGTCACATTCTTGGGGCTC
GLTSCR1 sg1 top	CACCGGGTTGGGCTCTGGGTTCAGG
GLTSCR1 sg1 bottom	AAACCCTGAACCCAGAGCCCAACCC
GLTSCR1 sg2 top	CACCGCAACCAGCCGGCCCCCAGTG
GLTSCR1 sg2 bottom	AAACCACTGGGGGCCGGCTGGTTGC
GLTSCR1 sg3 top	CACCGGGACCAGTGGCGACGAGCCG
GLTSCR1 sg3 bottom	AAACCGGCTCGTCGCCACTGGTCCC
GLTSCR1 sg4 top	CACCGGCAGAACCTGACGTTCATGG
GLTSCR1 sg4 bottom	AAACCCATGAACGTCAGGTTCTGCCC
GLTSCR1L sg1 top	CACCGGTAAGCAACCAGCTTGGAGA
GLTSCR1L sg1 bottom	AAACTCTCCAAGCTGGTTGCTTACC
GLTSCR1L sg2 top	CACCGGAAAGTGGCAGTCCATCACT
GLTSCR1L sg2 bottom	AAACAGTGATGGACTGCCACTTTCC
GLTSCR1L sg3 top	CACCGGTTCATCTTCAAGAAACTGG
GLTSCR1L sg3 bottom	AAACCCAGTTTCTTGAAGATGAACC
ACTB 41 top	CACCGAAGAGCTACGAGCTGCCTGA
ACTB 41 bottom	AAACTCAGGCAGCTCGTAGCTCTTC
ACTB 42 top	CACCGAGCCATGTACGTTGCTATCC
ACTB 42 bottom	AAACGGATAGCAACGTACATGGCTC
ACTL6A 51 top	CACCGACCACCATACCAATAGCTGT
ACTL6A 51 bottom	AAACACAGCTATTGGTATGGTGGTC
ACTL6A 52 top	CACCGACTGCAATTCCAGTCCACGA
ACTL6A 52 bottom	AAACTCGTGGACTGGAATTGCAGTC
ARID1A 321 top	CACCGAATACTCACAGGCAAGCTGG
ARID1A 321 bottom	AAACCCAGCTTGCCTGTGAGTATTC
ARID1A 322 top	CACCGATGGTCATCGGGTACCGCTG
ARID1A 322 bottom	AAACCAGCGGTACCCGATGACCATC
ARID1B 331 top	CACCGAAGTTGCTTCCGTTCCCGTG
ARID1B 331 bottom	AAACCACGGGAACGGAAGCAACTTC
ARID1B 332 top	CACCGCAAAGTTGCTTCCGTTCCCG
ARID1B 332 bottom	AAACCGGGAACGGAAGCAACTTTGC
ARID2 341 top	CACCGAGAAGTTGTTACATACTGTG
ARID2 341 bottom	AAACCACAGTATGTAACAACTTCTC
ARID2 342 top	CACCGAGACACATGAAGTTCTGTCC
ARID2 342 bottom	AAACGGACAGAACTTCATGTGTCTC
DPF1 1771 top	CACCGACACGTACCCCGCCCGCTGT
DPF1 1771 bottom	AAACACAGCGGGCGGGGTACGTGTC
DPF1 1772 top	CACCGACGCCAGGACACCGCTTCCC
DPF1 1772 bottom	AAACGGGAAGCGGTGTCCTGGCGTC
DPF2 1781 top	CACCGAAGAAGATACTCCCAAGCGT
DPF2 1781 bottom	AAACACGCTTGGGAGTATCTTCTTC
DPF2 1782 top	CACCGAAGGAAAGTCGTGGATCCTC

DPF2 1782 bottom	AAACGAGGATCCACGACTTTCCTTC
DPF3 1792 top	CACCGAAGCGAAAGAACAGGACTAG
DPF3 1792 bottom	AAACCTAGTCCTGTTCTTTCGCTTC
DPF3 1794 top	CACCGAGCGCTACAAGAACCGACCG
DPF3 1794 bottom	AAACCGGTCGGTTCTTGTAGCGCTC
PBRM1 4323 top	CACCGATGTGCGATGTAAGCCTGAG
PBRM1 4323 bottom	AAACCTCAGGCTTACATCGCACATC
PBRM1 4325 top	CACCGCAAATCCCAGAGTTTGCAAG
PBRM1 4325 bottom	AAACCTTGCAAACTCTGGGATTTGC
PHF10 4454 top	CACCGCACCATCACTGTCTAGAGCA
PHF10 4454 bottom	AAACTGCTCTAGACAGTGATGGTGC
PHF10 4460 top	CACCGTTACGGAATACTCTTTATAA
PHF10 4460 bottom	AAACTTATAAAGAGTATTCCGTAAC
SMARCA2 5738 top	CACCGGGTCCAGGCCTTGCGGTTTC
SMARCA2 5738 bottom	AAACGAAACCGCAAGGCCTGGACCC
SMARCA2 5732 top	CACCGACACCTAGGCTATTCAAATG
SMARCA2 5732 bottom	AAACCATTTGAATAGCCTAGGTGTC
SMARCA4 5741 top	CACCGCATCCCGGGGGGCACGCCCG
SMARCA4 5741 bottom	AAACCGGGCGTGCCCCCCGGGATGC
SMARCA4 5742 top	CACCGCCTGTTGCGGACACCGAGGG
SMARCA4 5742 bottom	AAACCCCTCGGTGTCCGCAACAGGC
SMARCB1 5781 top	CACCGAACTACCTCCGTATGTTCCG
SMARCB1 5781 bottom	AAACCGGAACATACGGAGGTAGTTC
SMARCB1 5782 top	CACCGAGGCGAAATGCGGACCGGGC
SMARCB1 5782 bottom	AAACGCCCGGTCCGCATTTCGCCTC
SMARCC2 5801 top	CACCGATACCTATTGTAGTCAAACC
SMARCC2 5801 bottom	AAACGGTTTGACTACAATAGGTATC
SMARCC2 5802 top	CACCGATTGTAGCAACTGTACAACC
SMARCC2 5802 bottom	AAACGGTTGTACAGTTGCTACAATC
SMARCD1 5811 top	CACCGAATAGAGAGGCCGCGTGAGC
SMARCD1 5811 bottom	AAACGCTCACGCGGCCTCTCTATTC
SMARCD1 5812 top	CACCGAGCTACTCACTTCTACCAGA
SMARCD1 5812 bottom	AAACTCTGGTAGAAGTGAGTAGCTC
SMARCD2 5821 top	CACCGACCAGACCATTGCTCGCAAG
SMARCD2 5821 bottom	AAACCTTGCGAGCAATGGTCTGGTC
SMARCD2 5822 top	CACCGAGCCTGTTACGACATCGATG
SMARCD2 5822 bottom	AAACCATCGATGTCGTAACAGGCTC
SMARCD3 5831 top	CACCGATGGCCGCGGACGAAGTTGC
SMARCD3 5831 bottom	AAACGCAACTTCGTCCGCGGCCATC
SMARCD3 5832 top	CACCGCAGCGCCCCGGGATGCCGTC
SMARCD3 5832 bottom	AAACGACGGCATCCCGGGGCGCTGC
SMARCE1 5841 top	CACCGAAACTTACTGTTGCAGGAGC
SMARCE1 5841 bottom	AAACGCTCCTGCAACAGTAAGTTTC
SMARCE1 5842 top	CACCGAATGCAGGTCCTCAAACGGC
SMARCE1 5842 bottom	AAACGCCGTTTGAGGACCTGCATTC

SS18L1 6006 top	CACCGGAGCTACGACCGGTCCTTCG
SS18L1 6006 bottom	AAACCGAAGGACCGGTCGTAGCTCC
SS18L1 6010 top	CACCGTCTTGGCCGGGCAGACGCGA
SS18L1 6010 bottom	AAACTCGCGTCTGCCCGGCCAAGAC
ACTL6B 622 top	CACCGACTATGAGCGGGGGCGTCTA
ACTL6B 622 bottom	AAACTAGACGCCCCCGCTCATAGTC
ACTL6B 623 top	CACCGATGACCTCCGCTCCATCCCG
ACTL6B 623 bottom	AAACCGGGATGGAGCGGAGGTCATC
BCL7A 4337 top	CACCGGTTCATCGAATGTTCCATCG
BCL7A 4337 bottom	AAACCGATGGAACATTCGATGAACC
BCL7A 4338 top	CACCGGGACCCATTTGTAGATTCGT
BCL7A 4338 bottom	AAACACGAATCTACAAATGGGTCCC
BCL7B 4340 top	CACCGTCAGCAGCCCGAGAACCTAA
BCL7B 4340 bottom	AAACTTAGGTTCTCGGGCTGCTGAC
BCL7B 4341 top	CACCGACACACCTCCGACTTCCGCA
BCL7B 4341 bottom	AAACTGCGGAAGTCGGAGGTGTGTC
BCL7C 4343 top	CACCGGCACCCACTTGAAGATACGA
BCL7C 4343 bottom	AAACTCGTATCTTCAAGTGGGTGCC
BCL7C 4344 top	CACCGGATCCCGGGGGCATAACTGC
BCL7C 4344 bottom	AAACGCAGTTATGCCCCCGGGATCC
BCL11A 4289 top	CACCGGATAAACAATCGTCATCCTC
BCL11A 4289 bottom	AAACGAGGATGACGATTGTTTATCC
BCL11A 4290 top	CACCGGGATACCAACCCGCGGGGTC
BCL11A 4290 bottom	AAACGACCCCGCGGGTTGGTATCCC
BCL11B 4292 top	CACCGCTAGGCGCACTGCCGCAGTA
BCL11B 4292 bottom	AAACTACTGCGGCAGTGCGCCTAGC
BCL11B 4293 top	CACCGGTCGTGCGTCCGTGAAGCCC
BCL11B 4293 bottom	AAACGGGCTTCACGGACGCACGACC
SS18 47101 top	CACCGCAGGCATCTGGCCGTTCATC
SS18 47101 bottom	AAACGATGAACGGCCAGATGCCTGC
SS18 47102 top	CACCGGCCTAACCATATGCCTATGC
SS18 47102 bottom	AAACGCATAGGCATATGGTTAGGCC

2.2.Preparation of chemically competent bacteria

To generate chemically competent bacteria, Stbl3 bacteria (NEB, Catalog: C737303) was streaked onto an agar plate without any antibiotics and incubated at 37 °C overnight. A single colony was picked and cultured in a 2 mL LB broth without antibiotics overnight at 37°C, 225 rpm. 0.5 mL of bacteria culture was added in 100 mL LB broth without antibiotics for 3 hour incubation at 37°C, 225 rpm. When OD600 value reached 0.4-0.6, as measured by a spectrophotometer, culture was incubated on ice for 15 minutes. Bacteria was pelleted by centrifugation at 3000 rpm, 4°C for 15 minutes and the pellet

was resuspended in 33 mL RF1 buffer (100mM RbCl, 50 mM $MnCl_2.4H_2O$, 30mM potassium acetate, 10 mM $CaCl_2$, 15% glycerol, pH 5.8). This resuspension was incubated on ice for 15 minutes and centrifuged as before. Pellet was resuspended in 8 mL RF2 buffer (10 mM MOPS, 10mM RbCl, 75mM $CaCl_2$, 15% glycerol, pH6.8) and aliquoted into tubes on ice. Aliquots were quickly frozen in liquid nitrogen and stored at -80°C freezers until use.

2.3.Virus production

2.5×10^6 293T cells were seeded per 10-cm dish in D10 media (10% FBS, 1% pen-strep in DMEM). The next day, cells were transfected with 2.5 µg viral vector, 2.25 µg psPAX2 (Addgene plasmid no. 12260) for lentiviruses or pUMVC (Addgene plasmid no. 8449) for retroviruses and 0.25 µg pCMV-VSV-G (Addgene plasmid no. 8454) using 20 µl FuGENE 6 (Promega) in 400 µl DMEM per plate. Supernatants were collected 48 h and 72 h post-transfection and filtered through 45-µm pore size filters. To concentrate the viruses, viral supernatants were mixed with PEG3350 solution (Sigma P3640, dissolved in PBS, 10% final concentration) and left overnight at 4 °C. The next day, supernatants were centrifuged at 2000 rpm for 20 min, and the pellets were re-suspended in PBS. Viral transductions were carried out overnight in the presence of 8 µg/ml protamine sulfate (Sigma). GFP-expressing viruses were titered on 293T cells. PRRX1 and ZBTB38 expression vector in pLenti6.2/V5-DEST was obtained from DNASU Plasmid Repository (Clone IDs: HsCD00330006 and HsCD00436332, respectively). MSCV-MN1 plasmid was a gift from Kathrin Bernt (Riedel et al., 2016). Transduced cells were selected with 7 µg/ml blasticidin for PRRX1 and ZBTB38 expressions and 1 µg/ml puromycin for MSCV-MN1 expression. Human fibroblasts were doubly transduced with either shRNA or CRISPR vectors in consecutive days in the presence of 8 µg/ml protamine sulfate (Sigma). Transduced cells were selected by puromycin at 1 µg/ml.

2.4.Preparation of inactivated mouse embryonic fibroblast for stem cell culture

Pregnant Balb/c mice were sacrificed at day 13 pc by cervical dislocation. The uterine horns were dissected, briefly rinsed by 70% ethanol and placed into a petri dish containing PBS. Embryos were separated from placenta, surrounding membranes, brain and dark red organs. They were washed with PBS to remove blood. Embryos were finely minced with razor blades till they become pipettable. They were resuspended in 1-2 mL 0.05% trypsin-EDTA per each embryo and incubated at 37°C for 15 minutes. 2-4 mL

D10 media was added to inactivate trypsin and any remaining large pieces were carefully removed. Cells were centrifuged at 1500 rpm for 5 minutes and pellet was resuspended in 5 mL D10 media per each embryo. 2 embryo equivalents were plated in 0.1% gelatin coated 10 cm dishes at passage 0 and next day, culture media were replaced with fresh D10 media. When MEFs reached confluency, they were trypsinized and resuspended in 1 mL freezing media (90% FBS, 10% DMSO) for each plate. Resuspended cells were aliquoted in cryovials and stored at -80°C freezers in styrofoam. To mitotically inactivate MEFs used as feeder cells during reprogramming, 40-60 of confluent 15 cm dishes of MEFs at passages 4-5 were washed once with PBS. 10 mL D10 supplemented mitomycin C at a final concentration of 10 µg/mL per each 15 cm dish was added on MEF culture and cells were incubated at 37°C for 2 hours. Cells were washed twice with PBS and 5 mL trypsin per each dish was added. After 5 minutes incubation at 37°C, trypsin was inactivated with 5-10 mL D10 media. Cells were counted with hematocytometer and resuspended in freezing media to correspond to a specific number of cells at each cryovial.

2.5.Reprogramming MEFs to iPSCs

For murine reprogramming, $2.5X10^4$ mouse embryonic fibroblasts (MEF) cells were seeded at each well of 0.1% gelatin coated 12-well plate. The next day, they were transduced by lentiviruses containing a single stem cell cassette (Sommer et al., 2009). Next day, culture media was replenished with small molecules I-BRD9 (1 µM), LP99 (1 µM) and dBRD9 (0.3 µM). On day 5 of reprogramming, cells were transferred onto inactivated MEFs. Next day and every other day until 14th day of reprogramming, culture media were replenished with mouse embryonic stem cell media (20% FBS, 1% NEAA, 1% Pen-Strep, 1% L-glutamine, 55µM 2-Mercaptoethanol, 1000 U/mL Leukemia Inhibitory Factor (LIF) (ESGRO, Catalog #ESG1107) in Knockout DMEM (Gibco, Catalog #10829-018)). iPSC colonies emerged between 12^{th} and 14^{th} days of reprogramming, after which cells were fixed with 4% paraformaldehyde and SSEA1 staining was performed to quantify reprogramming efficiency.

2.6.Human somatic cell reprogramming to iPSC

Human fibroblast cells (dH1f) were differentiated from H1 ESCs by passaging 8-10 times for a month using D10 media (10% FBS, 1% pen-step in DMEM). dH1f cells at passages 13-15 were seeded at a density of $5X10^4$ cells per well of a 12-well plate and transduced overnight with pSIN4-EF2-O2S and pSIN4-CMV-K2M viruses (Addgene, plasmids no. 21162 and no. 21164, respectively). Compounds were added the next day and replenished

every two days. I-BRD9, LP99, BI-7273 and dBRD9 were used at final concentrations of 1 μM, 3 μM, 1 μM and 0.3 μM, respectively, for the first two weeks of reprogramming. EPZ-004777, CBP30, I-CBP112 and A485 were used at final concentrations of 3 μM, 0.5 μM, 1 μM and 1 μM, respectively, for the first week of reprogramming. On day 6 after OSKM transduction, fibroblasts were passaged onto gelatin coated mitomycin-C-treated MEF seeded 12-well plates at a 1/8 ratio. Medium was changed to hESC medium (20% KnockOut Serum Replacement (Gibco, cat no: 10828028), 1% NEAA, 1% Pen-Strep, 1% L-glutamine, 55μM 2-Mercaptoethanol, 10 ng/mL b-FGF in DMEM/F12) the next day. Cells were fixed with 4% paraformaldehyde between 18 and 21 days of reprogramming and TRA-1-60 staining was performed to quantify reprogramming efficiency.

HDF-A cells (ScienCell Research Laboratories, Catalog #2320) were reprogrammed with the following episomal vectors pCXLE-hOCT3/4-shp53-F, Addgene plasmid no. 27077; pCXLE-hUL, Addgene plasmid no. 27080; pCXLE-hSK, Addgene plasmid no. 27078. Each plasmid (1 μg) was electroporated into $1x10^6$ million fibroblasts using a 4D-Nucleofector system (Lonza) with NHDF, human, neonatal program and DT 130 pulse code. Fibroblasts were cultured in DMEM medium for six days after which they were passaged onto mitotically arrested MEFs and cultured in hESC medium. Upon emergence of iPSC colonies (21-25 days), cells were fixed with 4% paraformaldehyde and TRA-1-60 staining was performed to quantify reprogramming efficiency.

For two-factor (OS) and three-factor (OSK) reprogramming, $5x10^4$ dH1f cells were infected with pMIG-OCT4 (Addgene plasmid no. 17225), pMIG-SOX2 (Addgene plasmid no. 17226) or pMIG-KLF4 (Addgene plasmid no. 17227) retroviruses at a multiplicity of infection of 2.5. The same protocol was applied for the rest of reprogramming as lentiviral and episomal reprogramming. Between 18^{th} and 21^{st} day of reprogramming, cells were fixed with 4% paraformaldehyde and TRA-1-60 staining was performed to quantify reprogramming efficiency.

2.7.SWI/SNF-focused CRISPR-Cas9 knockout screen for fibroblast proliferation rate and reprogramming to pluripotency

$5x10^4$ dH1f cells were seeded in each well of a 12-well plate by D10 media. Cells were transduced twice with lentiviral CRISPR-Cas9 constructs targeting SWI/SNF-related proteins. Next day, cells were passaged onto 6-cm dishes and puromycin selection (1 μg/mL) was started. Upon completion of selection, cells were split into two wells of a

6-well plate at a 1:5 (used as first replicate) and 1:10 (used as second replicate) ratios. At the same time, 1 in 25th of each replicate was seeded in 2 black 96-well plates in triplicates for cell proliferation assay. Cell proliferation was measured with the CellTiter-Glo Luminescent Cell Viability Assay (Promega) according to manufacturer's instructions using a plate reader (Synergy H1 Reader, BioTek). The last read of CTG was used to normalize cell counts in 6-well plates instead of counting each condition. 5x10^4 fibroblasts were seeded in technical duplicates for lentiviral OSKM reprogramming as described in section 2.6

2.8.iPSC culture and differentiation assays

Individual iPSC colonies generated from BRD9 sgRNA and Cas9-expressing fibroblasts or iPSCs that were generated by OS/OSKM-reprogramming with different treatments were manually picked and cultured on inactivated MEFs with hESC media and ROCK inhibitor Y-27632 at a final concentration of 10 μM. After a few passages clones were transitioned to feeder-free conditions on matrigel (Corning, Catalog #354277)-coated plates with MEF-conditioned hESC media. iPSCs used for CHIR99021-induced mesendodermal differentiation experiments were picked from adult fibroblasts, electroporated with non-integrative episomal plasmids and cultured on feeder-free conditions on matrigel-coated plates with mTESR1 media (Alici-Garipcan et al., 2020; Fidan et al., 2015). When the cells reached 40-60% confluency, they were treated with DMSO, 1 μM I-BRD9 or 0.1 μM dBRD9 in 5 μM CHIR99021 containing D10 media for 48 hours.

2.9.Flow cytometry

GFP positive iPSC and 293T cells were identified by BD Accuri c6 using negative/positive controls to set gates. Cell surface TRA-1-60 expression was analyzed at 6th day of reprogramming. Cells were detached with 0.3 mL 0.05% trypsin-EDTA solution from each well of a 12-well plate and incubated at 37°C for 5 minutes. 0.3 mL 10% FPS in PBS was added to each well and centrifuged at 1500 rpm for 5 minutes. Pellets were resuspended in 1% FPS containing PBS and centrifuged as before. Pellets were resuspended in 0.25 μL PE-conjugated anti-human TRA-1-60-R antibody (Biolegend, catalog no. 330610) in 1% FBS containing PBS solution and incubated on ice for 1 hour in dark. Cells were washed twice with 1% FBS containing PBS with the same centrifuge conditions as before. Pellets were resuspended in 100 μL 1% FBS

containing PBS solution and an Accuri C6 flow cytometer (BD) was used to assess reprogramming efficiency by setting gates on GFP and PE channels depending on negative control.

2.10. Cell proliferation assay

$2x10^4$ fibroblasts were seeded at each well of 96-well plates in triplicates. On the following day, cells were either infected with lentiviral OSKM vectors or left uninfected. The uninfected group was treated with compounds on the same day and the OSKM-transduced group was treated with compounds the day after. Cell proliferation was measured on the next day after seeding as day 0 and every other day with the CellTiter-Glo Luminescent Cell Viability Assay (Promega) according to manufacturer's instructions using luminometric measurements with a plate reader (Synergy H1 Reader, BioTek).

OCT4-GFP iPSCs were seeded on matrigel-coated 96-well plates in mTESR1 media with ROCK inhibitor Y-27632 (10 μM). Following day, media was refreshed with mTESR1 media without ROCK inhibitor. When cells reached 40-60% confluency, they were treated with DMSO, 1 μM I-BRD9 or 0.1 μM dBRD9 in mTESR1 media for 48 hours. Cell proliferation was measured with the CellTiter-Glo Luminescent Cell Viability Assay (Promega) according to manufacturer's instructions using luminometric measurements with a plate reader (Synergy H1 Reader, BioTek).

2.11. β-galactosidase staining

Fibroblasts on the 3rd and 6th days of DMSO, I-BRD9 and dBRD9 treatments and either induced by OSKM or left uninfected were washed with PBS. Cells were fixed in 0.2% glutaraldehyde for 5 minutes at room temperature. Cells were immediately washed once with PBS and stained with X-gal solution (1 mg/mL X-gal, 150 mM NaCl, 2 mM $MgCl_2$, 5 mM $K_3Fe(CN)_6$, 5 mM $K_3Fe(CN)_6$, 5 mM $K_4Fe(CN)_6$, 40 mM NaP_i) for 3 hours at 37°C. Ten spots per well were randomly captured and average β-gal positive cells were calculated.

2.12. Immunostaining

To mouse and human iPSC colonies were quantified via SSEA1 and TRA-1-60 immunostaining, respectively. Paraformaldehyde-fixed cells were washed twice with PBS and incubated with biotin-anti-TRA-1-60 (eBioscience, catalog no. 13-8863-82) or anti-SSEA1 antibody (Biolegend, catalog no. 125604) at 1:200 dilution in staining

solution (3% FBS, 0.3% TritonX-100 in PBS) overnight at 4°C. Next day, cells were washed at least twice with PBS and incubated with streptavidin horseradish peroxidase (Biolegend, catalog no. 405210) at 1:500 dilution in staining solution for 2 hours at room temperature. A master mix of DAB solution was prepared by adding and vortexing following in PBS: 0.1% 3,3'-diaminobenzidine (Sigma D8001), 0.05% nickel ammonium sulfate and 0.015% H_2O_2. 15 minutes of incubation of cells with DAB solution at room temperature resulted in brown color development and cells were washed at least twice with PBS. Single cream or diluted double cream was used to cover the surface of wells to create contrast during scanning of plates. Plates were scanned in JPEG format by at least 600 dpi resolution and ImageJ software (https://imagej.nih.gov/ij/) was used to quantify reprogramming efficiency either by calculating percent area covered by stained iPSC colonies or the number of iPSC colonies emerged.

For immunofluorescence-based characterization of iPSCs, cells were passaged onto coverslips with mitomycin-C-treated MEFs in hESC medium. Cells were fixed with 4% paraformaldehyde for 20 minutes at room temperature and washed at least twice with PBS. Cells were incubated overnight at 4 °C with primary antibodies: OCT4 (Abcam, catalog no. ab19857), SSEA4/A647 (BD, catalog no. 560218) and NANOG (Abcam, catalog no. ab21624) diluted 1:100 in 3% donkey serum, 3% BSA, 0.01% Triton X-100 in PBS. For NANOG and OCT4, Alexa-488-conjugated secondary antibodies (Molecular Probes) at 1:500 dilution ratio in 3% donkey serum, 3% BSA, 0.01% Triton X-100 in PBS were used. Nuclei were stained with VECTASHIELD Antifade mounting medium with DAPI (H-1200-10). Images were acquired using a Nikon 90i confocal microscope.

2.13. Teratoma assays

All experiments were carried out under a protocol approved by Koç University Animal Experiments Ethics Committee and all relevant ethical regulations were complied with. iPSCs from a confluent 6-well dish were collected using 1 mg/mL collagenase IV in DMEM/F12 and resuspended in 100 µL ice-cold 1:1 mixture of Matrigel (Corning) and DMEM. Intramuscular injections were performed in severe combined immunodeficiency mice. Teratomas were collected 8–12 weeks after injection and analyzed histologically.

2.14. Protein extraction

Nuclear lysates were isolated by resuspending cell pellets in cytosolic lysis buffer (10mM HEPES pH 7.9, 10mM KCl, 0.1mM EDTA, 0.4% NP-40, Protease Inhibitor (1X) (Roche)) and shaken on ice for 15 minutes. After centrifugation at 3000g at 4°C for 3

minutes, pellets were washed with cytosolic buffer again. Pellets were resuspended in nuclear lysis buffer (20mM HEPES pH 7.9, 0.4M NaCl, 1mM EDTA, 10% Glycerol, Protease Inhibitor (1X) (Roche)) and sonicated in a Bioruptor Pico soniacator (Diagenode) by high amplitude for 10 seconds four times settings. Supernatant containing nuclear fragments was collected after centrifugation at 15000g at 40C for 5 minutes. Lysates were incubated at 95°C for 15 minutes with 4X laemmli buffer (BioRad) containing β-mercaptoethanol (BioRad).

Whole cell lysates were isolated by resuspending cell pellets in lysis buffer (50 mM Tris, pH 7.4, 250 mM NaCl, 5 mM EDTA, 50 mM NaF, 1% Noidet P40, 1 mM PMSF, 1X protease inhibitor and 0.025% NaN_3). Resuspended cells were incubated on ice for 30 minutes with 10 minute vortexing intervals. They were centrifuged at highest speed for 10 minutes at 4°C. Supernatant containing whole cell lysate was transferred to a fresh tube and protein concentrations were quantified using BCA protein assay kit. Equal amount of lysates (50 µg) were incubated at 95°C for 15 minutes with 4X laemmli buffer (BioRad) containing β-mercaptoethanol (BioRad).

Histone extraction was performed by resuspending cell pellets in triton extraction buffer (PBS containing 0.5% Triton X-100, 2 mM phenylmethylsulfonyl fluoride, 0.02% NaN_3) and suspension was incubated on ice for 10 minutes with gentle stirring. To pellet nuclei, suspension was centrifuged at 6500g for 10 minutes at 4°C and supernatant was discarded. Nuclei were resuspended in half volume of initial triton extraction buffer and centrifuge step was repeated. Pellet was resuspended in 0.2 M HCl for acid extraction of histones overnight at 4°C on a rotator. Next day, samples were centrifuged as before and supernatant containing histones was transferred to a fresh tube. Supernatant was neutralized with NaOH with 1/10 volume of HCl volume used for histone extraction. BCA protein assay kit was used to quantify protein concentrations following user manual. Equal amount of lysates (5 µg) were incubated at 95°C for 15 minutes with 4X laemmli buffer (BioRad) containing β-mercaptoethanol (BioRad).

2.15. Western blotting

Boiled samples and protein marker (Bio-rad, Catalog: 161-0374) were loaded onto pre-cast SDS-PAGE gels (BioRad, Catalog: 456-1084). Proteins were transferred onto PVDF membrane (BioRad, Catalog: 1620177) using Bio-Rad tans-blot turbo transfer system at mixed weight transfer setting (whole cell and nuclear lysis) or low molecular weight transfer setting (histone extracts) followed by incubation in 5% blotting grade

blocker (BioRad Catalog: 1706404) solution for 1 hour. Membranes were incubated with antibodies against H3K27ac (Active Motif, catalog no. 39133) at 1:1000 ratio, H3K18ac (Abcam, ab1191) at 1:1000 ratio, H3 Total (Cell Signaling Technologies, catalog no. 4499S) at 1:1000 ratio, BRD9 (Active Motif, Catalog: 61537) at 1:1000 ratio, BRD7 (Cell Signaling, Catalog: D9K2T) at 1:1000 ratio and HDAC1 (Santa Cruz, Catalog: sc-7872) at 1:500 ratio overnight at 4°C. Next day, membranes were washed with TBS-T for 15 minutes three times and incubated with secondary antibodies (Abcam, ab97051 or ab97023) at 1:5000 ratio in blocking solution for 1 hour at room temperature. Membranes were washed with TBS-T for 15 minutes 3 times and incubated shortly with ECL western blotting substrate (ThermoFisher) before imaging using an Odyssey Imaging System (LI-COR Biosciences).

2.16. Genomic DNA PCR

Genomic DNA was extracted from cells with tissue isolation kit (Nucleospin catalog no. 740952.50) following manufacturer's instructions. 200 ng of genomic DNA was amplified with DreamTaq DNA Polymerase (catalog no. K1081) using indicated primer pairs (Table 2.3). OCT4, SOX2, KLF4 and c-MYC transgene-containing amplicons were amplified with 30 cycles of 30 s at 95 °C, 30 s at 54.5 °C and 60 s at 72 °C. Amplicons containing GLTSCR1L sgRNA-induced indels were amplified using a BIO-RAD T100 thermocycler with 30 cycles of 30 s at 95 °C, 30 s at 66.7 °C and 60 s at 72 °C.

Table 2.3. List of primers used to amplify genomic DNA

gDNA PCR Primer	Sequence
GLTSCR1L-T7- Forward	GGCTCACTCCTGTAATCCCA
GLTSCR1L-T7- Reverse	ACATGCGTCACCCCTATAGG
OCT4-transgene Forward	CCTCACTTCACTGCACTGTA
OCT4,SOX2,KLF4-transgene Reverse	CCTTGAGGTACCAGAGATCT
SOX2-transgene Forward	CCCAGCAGACTTCACATGT
KLF4-transgene Forward	GATGAACTGACCAGGCACTA
MYC-transgene Forward	TGCCTCAAATTGGACTTTGG
MYC-transgene Reverse	CGCTCGAGGTTAACGAATT

2.17. T7 endonuclease assay

PCR products containing GLTSCR1L-induced indels were purified by MN PCR-

Clean up Kit. 400 ng of purified DNA was denatured and heteroduplexed using NEBuffer 2 (catalog no. B7002S) in a total volume of 19 µL with the following settings in a thermocycler: 5 minutes at 95 °C, ramping down to 85 °C with 2 °C/s, ramping down to 25 °C with 0.1 °C/s. 1 µL T7 endonuclease (catalog no. M0302L) was added into the heteroduplexed DNA reaction and mixed thoroughly. The reaction was incubated at 37 °C for 1 hour and visualized in BIO-RAD Gel Doc XR+ after running on agarose gels.

2.18. RT-qPCR

Total RNA was extracted using NucleoSpin RNA kit (Macherey Nagel). Equal amounts of total RNAs (1 µg) from each sample were mixed with 2.5 µL of 2.5 mM dNTPs, 1 µL of 50 µM random hexamers up to a total volume of 16.5 µL in a PCR tube. This mixture was incubated at 65 °C for 5 minutes in a thermocycler and immediately placed on ice. 5 µL 5X First Strand Buffer, 2 µL of 0.1 M DTT and 0.5 µL Rnasin (Promega, catalog no. N2111) was added into each tube and mixed thoroughly. After 10 minutes of incubation at room temperature, 1 µL of MMLV RT (Invitrogen, catalog no. 28025-013) enzyme was added to each tube and mixed thoroughly. Mixture was incubated at 37 °C for 1 hour and 70 °C for 15 minutes in a thermocycler. 75 µL nuclease free water was added to each tube. The resulting complementary DNAs were used for PCR using SYBR-Green Master PCR mix (Roche) and run on a LightCycler 480 Instrument II (Roche) with 40 cycles of 30 s at 95 °C, 30 s at 60 °C and 30 s at 72 °C. All quantifications were normalized to an endogenous β-actin control. The relative quantification value for each target gene compared to the calibrator for that target is expressed as $2^{-(Ct-Cc)}$ (Ct and Cc are the mean threshold cycle differences after normalizing to β-actin). List of primers used in RT-qPCR is listed in Table 2.4.

Table 2.4. List of primers used in RT-qPCR

Primers	Sequence
ACSS2 Forward	AACTTGACGTGATGGGGCTT
ACSS2 Reverse	GTCTCCCCAGAATTCCCGC
ACLY Forward	GACTTCGGCAGAGGTAGAGC
ACLY Reverse	TCAGGAGTGACCCGAGCATA
PRRX1 Forward	ACACCCTGCAGGCGAAA
PRRX1 Reverse	CTTCTGAGTTCAGCTGGTCA
ACTB Forward	TGAAGTGTGACGTGGACATC
ACTB Reverse	GGAGGAGCAATGATCTTGAT
BRD7 Forward	CTGGAGATGCCGAAGCACAC
BRD7 Reverse	TGGGATCCACAGGATGGAGA

BRD9 Forward	GGTGCCAGCTGTTCCCAG
BRD9 Reverse	GCTTGTCGGCATAATCCTCG
ZBTB38 Forward	GGGAATGACCCAGAACCCAG
ZBTB38 Reverse	TGGTCACTTGCGTCATCGAA
NANOG Forward	TGATTTGTGGGCCTGAAGAAA
NANOG Reverse	TGGTGGTAGGAAGAGTAAAG
LIN28A Forward	TGCGGGCATCTGTAAGTGG
LIN28A Reverse	GGAACCCTTCCATGTGCAG
LEFTY2 Forward	TGGACCTCAGGGACTATGGAG
LEFTY2 Reverse	CCGAGGCGATACACTGTCG
TBXT Forward	ACGCAGTTCATAGCGGTGAC
TBXT Reverse	CAATTGTCATGGGATTGCAG
EOMES Forward	GTGCCCACGTCTACCTGTG
EOMES Reverse	TGGTGGCGGTGGAATTTGAG
MN1 Forward	GAAGGCCAAACCCCAGAAC
MN1 Reverse	GATGCTGAGGCCTTGTTTGC
ZBTB38 Forward	GGGAATGACCCAGAACCCAG
ZBTB38 Reverse	TGGTCACTTGCGTCATCGAA

2.19. Gene expression analysis by Microarray and RNA-Sequencing

Differential gene expressions between pluripotent stem cells and fibroblast cells were computed by affy and limma packages from R by Tunç Morova. Samples of dH1f and BJ fibroblasts are compared to their respective iPSCs and embryonic stem cells from GEO data series GSE55679. Genes that have a $\log_2$ fold change value of 3 or more in all fibroblasts compared with pluripotent cells were categorized as the fibroblast-related gene set and -3 and less were categorized as the pluripotency-related gene set (Table 2.5 and Table 2.6). Rank-ordered gene lists were used for gene-set enrichment analysis (Subramanian et al., 2005).

Table 2.5. Pluripotency-related genes

AASS	ESRP1	LIN28B	RPS6KA1
ACTA1	F11R	LINGO1	RRAGD
ADAMTS19	FABP5	LMNB1	SALL1
ADD2	FAM117A	LPAR4	SALL2
APOC1	FAM169A	LRAT	SALL3
ARID3B	FAM221A	LRRC8B	SALL4
ATP1A2	FGF13	LRRN1	SBK1
B4GALT6	FLVCR1	LSR	SCG3
BCL11A	FOS	MAL2	SCG5
BEND3	FOXA3	MAP7	SCGB3A2

BEX2	FOXH1	MARVELD3	SEMA6A
BMP2	FOXO1	MND1	SERPINB9
BST2	FRAT2	MRAP2	SFRP2
C9orf135	FREM2	MTF2	SLAIN1
CA14	FZD3	MYO5C	SLC16A1
CA2	FZD5	NANOG	SLC16A9
CACNA2D2	GABRB3	NAP1L2	SLC2A3
CADPS2	GAL	NAP1L3	SLC35F2
CALB1	GALNT12	NELL2	SLC7A3
CD200	GCA	NFE2L3	SNX10
CDC25A	GCNT2	NLGN4X	SORL1
CDC7	GLDC	NMU	SOX11
CDH1	GNAL	NPPB	SOX2
CDH3	GPC4	NPTX2	SPINT2
CENPV	GPM6B	NRARP	STOX2
CGN	GPR143	ORC1	STRBP
CHST6	GPR160	PAIP2B	SYNE4
CLDN10	GPR19	PALD1	SYNGR3
CLDN6	GPR27	PAX6	SYT6
CLGN	GRHL2	PDPN	TAF4B
CMTM8	GRTP1	PELI1	TDRP
CNTNAP2	GUCA1A	PHC1	TERF1
COBL	GYG2	PIM2	TET1
COCH	HELLS	PIPOX	TMEM125
COL9A3	HERC5	PITX2	TNFRSF21
CPVL	HOMER1	PKIB	TNNI3
CRISPLD1	HOOK1	PLBD1	TPD52
CTSV	HPGD	PLEKHA7	TRIM71
CXADR	HS6ST2	PLEKHH1	TSTD1
CXCL5	IDO1	PLP1	TUBB2B
CYP26A1	IGDCC3	PLS1	UGT8
CYP2S1	IGF2BP3	PMAIP1	UNC5D
CYP4X1	IGFBPL1	PMEL	USP44
CYYR1	IQGAP2	PODXL	VAMP8
DAZL	JARID2	POU5F1B	VANGL2
DDX25	KCND2	PPAT	VASH2
DEPDC1B	KCNK12	PPM1E	VAT1L
DMKN	KCNK5	PRKCQ	VSNL1
DNMT3A	KCNN2	PRKCZ	VWDE
DNMT3B	KCNS3	PRKX	WDHD1
DPPA2	KDR	PROM1	WNK3
DPPA4	KIAA1551	PRSS16	ZDHHC23
DSG2	KIF21A	PRSS8	ZFP42
EDNRB	KIF5C	PRTG	ZIC2
EGLN3	KRTCAP3	PTPRZ1	ZNF114

ELAVL2	L1TD1	RAB39B	ZNF165
ELOVL7	LAPTM4B	RARRES2	ZNF462
ENPP5	LCK	RASGEF1A	ZNF711
EPCAM	LEFTY1	RBPMS2	ZNF738
ERVMER34-1	LIN28A	RNF125	ZNF850

Table 2.6. Fibroblast-related genes

ACKR4	COMP	HEG1	NDRG1	SLFN11
ACTA2	COPZ2	HMCN1	NEGR1	SLFN5
ADAM12	CORIN	HTRA1	NEK7	SNAI2
ADAMTS1	COX7A1	IFI16	NEXN	SPHK1
ADAMTS2	CPQ	IFI6	NFIC	SPOCD1
ADAMTS5	CRIM1	IFIT1	NID2	SPOCK1
AHNAK2	CRIP1	IFIT3	NNMT	SRGN
AHR	CRLF1	IFIT5	NOV	SRPX
AHRR	CRYAB	IGFBP3	NPR3	SSPN
ALCAM	CSPG4	IGFBP4	NR2F2	SULF1
ALPK2	CSRP1	IGFBP5	NR3C1	SYNC
AMIGO2	CYBRD1	IGFBP6	NT5E	SYNPO2
ANGPT1	CYP1B1	IGFBP7	NTN4	TBC1D2
ANO3	DAAM2	IL10RB	NUPR1	TCEAL1
ANXA1	DAB2	IL13RA1	OMD	TFPI2
ARSJ	DCN	IL15	OSR1	TGFB1I1
ASPN	DDR2	IL1R1	OSR2	TGFBI
ATP10D	DKK1	IL6ST	OXTR	THBS1
ATP2B4	DKK3	INHBA	PALMD	THBS2
B3GALT2	DRAM1	IRF9	PARVA	TIMP1
BGN	EBF1	ITGA4	PCGF5	TIMP2
BICC1	ECM1	ITGBL1	PCOLCE	TIMP3
BST1	ECM2	KAT2B	PDE4DIP	TLCD2
C11orf87	EFEMP1	KCNE4	PDGFRB	TMEM117
C12orf75	EFEMP2	KDELR3	PEAR1	TMEM119
C1orf198	EHD2	KRTAP1-5	PLAGL1	TMEM173
C1R	ELN	LGALS1	PLAT	TNFRSF11B
C1S	EMILIN1	LIMS2	PLP2	TRAM2
CALHM2	EMP3	LOX	PLSCR4	TSHZ1
CASP4	ENPP2	LOXL1	POSTN	TSPO
CAV1	EPS8	LPAR1	PPP1R3C	TWIST1
CAV2	EVA1A	LTBP2	PRR16	ULBP2
CBR3	F2RL2	LUM	PRRX1	VASN
CCDC80	FAM114A1	LURAP1L	PRRX2	VEGFC
CCDC92	FAP	LXN	PRSS12	VGLL3
CCPG1	FAS	LY96	PRSS23	VIT

CD248	FBLN2	MAMLD1	PTGS1	WFDC1
CD44	FBLN5	MAP3K7CL	PTX3	WNT5A
CD59	FBN1	MARCH4	RARRES1	ZBTB20
CDC42EP5	FHL1	MEGF6	RCAN2	ZEB1
CDH11	FIBIN	MEIS2	RGL1	ZHX1
CDH13	FILIP1L	MFAP4	RGS4	ZNF25
CDKN2C	FKBP9	MFAP5	RHOJ	
CFH	FOSL2	MFSD1	RRAS	
CFL2	FOXC1	MGARP	S100A4	
CHSY3	FOXC2	MGLL	SAMD9	
CLIP4	FOXF2	MGP	SAMD9L	
CNRIP1	FRMD6	MKX	SCN9A	
COL12A1	GBE1	MMP1	SEMA3C	
COL15A1	GBP1	MMP3	SEPT11	
COL16A1	GBP3	MORC4	SERPINB2	
COL1A2	GEM	MRGPRF	SERPINE1	
COL3A1	GLIPR1	MRVI1	SFXN3	
COL5A1	GLIS3	MSRB3	SGIP1	
COL5A2	GLT8D2	MVP	SH3RF3	
COL5A3	GNG11	MXRA5	SIX1	
COL6A2	GNPDA2	MXRA7	SLC16A4	
COL6A3	GPNMB	MYOCD	SLC16A7	
COL8A1	GREM1	MYOF	SLC1A4	
COMMD3	HCFC2	NABP1	SLC22A4	

Control or OSKM-transduced fibroblasts were treated for 5 days with DMSO, CBP30 (0.5 μM), I-CBP112 (1 μM) or A485 (1 μM). Uninfected and OSKM-transduced fibroblasts were treated for 5 days with DMSO, BI-7273 (1 μM), I-BRD9 (1 μM) and dBRD9 (0.3 μM). In parallel, control gRNA (gNT1) and BRD9 g823-expressing cells were used. Total RNA was prepared using Direct-zol kit according to manufacturer's instructions (Zymo Research) in triplicates per sample. Libraries were prepared using NEBNext Poly(A) mRNA Magnetic Isolation Module from the NEBNext ultra-directional RNA kit to create a first stranded library. 20-25 million paired-reads were targeted per each replicate of a sample by sequencing libraries in Next-Seq500 (Illumina). Reads were mapped to hg19 built-in genome by HISAT2 after assessing their quality by FastQC. DESeq2 package was used to find differentially expressed genes between samples. Genes were considered to be differentially regulated based on $\log_2$ fold change > 0.5 and $P < 0.05$ for Chapter 3 and adj-p-value<0.05 for Chapter 4. Rank-ordered gene lists were used for gene-set enrichment analysis (Subramanian et al., 2005).

2.20. ChIP-Sequencing

5×10^6 dH1f cells were used per chromatin immunoprecipitation experiment. ChIP-seq assays were performed in triplicate using ChIP-IT High-sensitivity kit (Active Motif, catalog no. 53040) following manufacturer's instructions at the same time points as RNA-seq experiments. For quantitative ChIP-Seq experiments, SF9 cells were spiked into the fibroblast pool at 1:5 ratio. The cells were then cross-linked with 1% formaldehyde for 15 minutes and chromatin was fragmented to 200–300 base pairs by sonication (4X10 cycles, 30 second on, 30 seconds off, highest amplitude) using a Biorupter Pico (Diagenode). 10% of IP reaction was used as input and each lysate was immunoprecipitated with 5 µg of primary antibody: anti-H3K27Ac (Active Motif, 39133), anti-H3K18Ac (Abcam, ab1191), anti-H3K4me3 (Merck Millipore, 07-473), anti-H3K4me1 (Diagenode, C15410194). Purified DNA was used for library preparation using a NEBNext Ultra DNA sample preparation kit (NEB) according to the manufacturer's recommendations. The samples were multiplexed, quantified on Tapestation (Agilent), and sequenced on a NextSeq 500 (Illumina) platform (paired-end, 2×41 bp). Sequencing depth was in excess of 20 million reads per sample, suggesting sufficient coverage.

ChIP-seq, ATAC-Seq and RNA-Seq analysis related to Chapter 3 of this thesis was carried out by Dr. Adam Cribbs from Botnar Research Centre, University of Oxford. A computational pipeline was written calling scripts from the CGAT toolkit to analyze the next-generation sequencing data (Sims et al., 2014). Bowtie software v.0.12.5 was used to align the reads to the human hg19 reference genome (Langmead et al., 2009). Reads were only considered that were uniquely aligned to the genome with up to two mismatches. For quantitative ChIP, the number of reads mapping to SF9 cells was determined for each sample. Bedtools v.2.2.24 was used to generate bedgraph files from the mapped BAM files using the scaling factor derived from the SF9 read count. Averaged tracks for each condition were produced representing the mean of the scaled values for biological replicates. Homer tag directories were then produced using the raw read function, and coverage plots around the transcription start sites were plotted in R. MACS software (v.1.4.2) was used to identify enrichment of intervals of H3K27Ac following ChIP-seq and regions of open chromatin following ATAC-seq. Sequencing of the whole cell extract was performed to determine the background model when analyzing ChIP-seq. For ATAC-seq, peak callers were run without a background. For visualization as a

University of California Santa Cruz Genome Browser track, the bam files generated from Bowtie were converted to BigWig files.

2.21. ATAC-Seq

ATAC-seq was performed using 65000 cells for each triplicates per sample for the transposition reaction. Cells were trypsinized and centrifuged at 500g, 4 °C for 5 minutes. Cell pellet was gently resuspended in 50 µL ice-cold lysis buffer (10 mM Tris-HCl (pH: 7.4), 10 mM NaCl, 3 mM $MgCl_2$, 0.1% IGEPAL CA-630) and immediately centrifuged at 500g, 4 °C for 15 minutes. Cell pellet was gently resuspended in 50 µL reaction mixture (25 µL 2X TD buffer (homemade), 2.5 µL Tn5 transposase (gift of Dr. Udo Oppermann, Botnar Research Centre, University of Oxford), 22.5 µL nuclease-free water) on ice and this mixture was incubated at 37 °C for 30 minutes. Immediately after incubation, genomic DNA was isolated using Qiagen MinElute Kit by following manufacturer's guide. PCR amplification was performed using the following protocol: 3 min at 72 °C, 30 s at 98 °C and 11 cycles of 10 s at 98 °C, 30 s at 63 °C, and 3 min at 72 °C. The samples were then purified using the GeneJET PCR purification kit and eluted with 20 µl of TE buffer. Samples were then validated on a Tapestation (Agilent) to determine library size and quantification before paired-end (2 × 41 bp) sequencing on a NextSeq 500 (Illumina) platform to aim 20-25 million reads per sample.

Sequence reads were quality controlled with FastQC and mapped to the most recent human reference genome (hg38) using bowtie (Langmead et al., 2009), allowing a maximum of 2 mismatches and only uniquely mapped reads. Reads with mapping quality score below 30 and blacklisted regions were filtered. DeepTools (Ramírez et al., 2016), ComputeMatrix and plotProfile commands were used to generate aggregate ATAC plots. For this, BigWig files were generated using deepTools bamCoverage command, eliminating duplicates and normalizing by sequencing depth and effective genome size. Intersecting regions among fibroblast ChIP peaks for H3K27ac, H3K4me1 and H3K4me3 were identified using previously published fibroblast ChIP-seq data from Chapter 3 (Ebrahimi et al., 2019), calling the peaks using macs2 (Zhang et al., 2008) with the parameters --keep-dup 1 and –broad, and finding the intersecting regions using bedtools intersect (Quinlan and Hall, 2010).

2.22. Statistical Analysis

Statistical analysis was performed using GraphPad Prism 8 for t-tests or R 4.0.2 for Wilcoxon-signed rank test. The test details, number of biological replicates and exact p-values are provided in related figure legends.

2.23. Data availability

Chapter 3 related RNA-sequencing, ChIP-sequencing and ATAC-sequencing data are deposited to the NCBI GEO database with the accession number GSE118220. Chapter 4 related RNA-sequencing and ATAC-sequencing data are deposited to the NCBI GEO database with the accession number GSE161640.

Chapter 3. INVESTIGATING THE ROLE OF CREBBP/EP300 IN REPROGRAMMING

3.1.Initial stage of reprogramming is most sensitive to CREBBP/EP300 bromodomain inhibition.

To gain insight on the mechanism by which CREBBP/EP300 bromodomain inhibition increases human somatic cell reprogramming to pluripotency, I conducted a time-course treatment experiment with I-CBP112 for 1 and 2 week time windows (Figure 3.1.a and Figure 3.1.b). We observed that CREBBP/EP300 bromodomain inhibition increases somatic cell reprogramming efficiency when applied in the first week of reprogramming (Figure 3.1.c and Figure 3.1.d). This result shows that inhibiting CREBBP/EP300 bromodomain acts at the initiation stage to enhance reprogramming when somatic cell identity should properly silenced for successful iPSC generation.

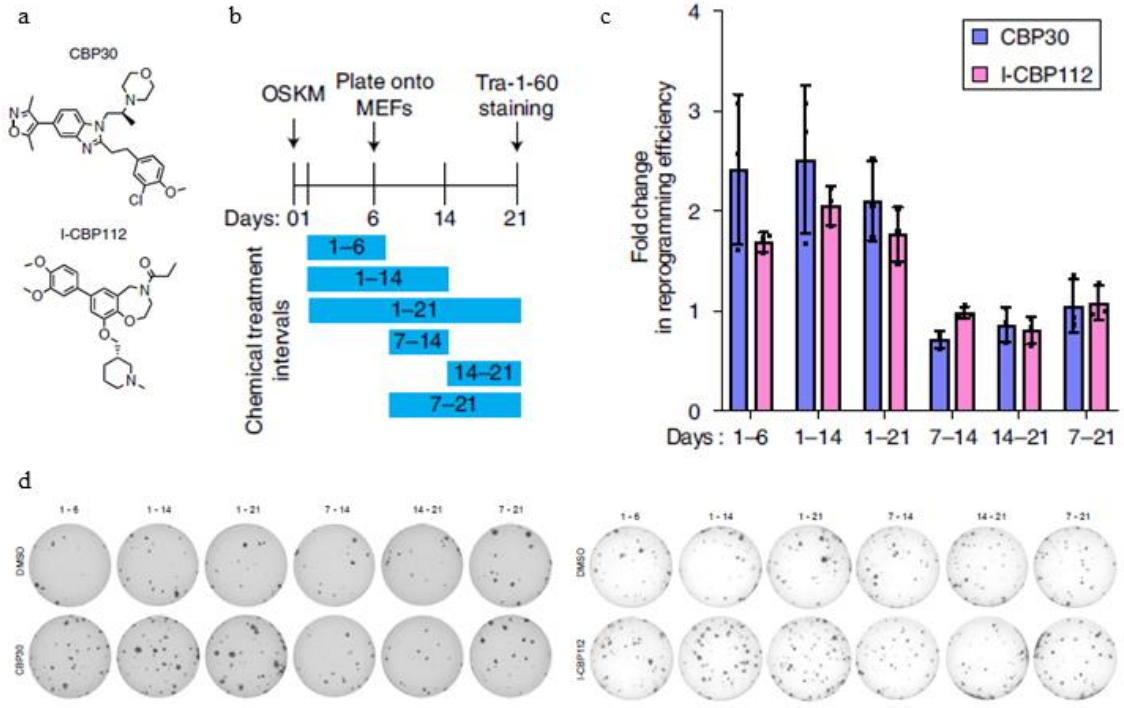


Figure 3.1 Initial week of reprogramming is most sensitive to CREBBP/EP300 bromodomain inhibition. (a) Chemical structures of the CREBBP/EP300 bromodomain inhibitors, CBP30 (top) (Hay et al., 2014) and I-CBP112 (bottom) (Picaud et al., 2015). (b) Schematic depicting the time intervals for treatment of CBP30 or I-CBP112 during reprogramming. (c) Fold change in the number of TRA-1-60 positive colonies with CBP30 or I-CBP112 treatments for the indicated time intervals normalized to DMSO-

treated wells. Bar graphs show the mean and error bars represent standard deviation. n=3, independent biological replicates. Dr. Ayyub Ebrahimi and Gülben Gürhan Sevinç conducted CBP30-related experiments. (d) Representative well images of TRA-1-60 stained wells from Figure 3.1c.

3.2. CREBBP/EP300 bromodomain inhibition increases reprogramming efficiency induced by fewer exogenous factors.

As DOT1L inhibition enables iPSC generation by overexpressing OCT4 and SOX2 (OS) only, I asked whether CREBBP/EP300 bromodomain inhibition could achieve this, too (Onder et al., 2012). I observed an increase in OS-mediated reprogramming efficiency by either I-CBP112, DOT1L inhibitor or their combination (Figure 3.2.a). A small molecule cocktail (IVFZ) consisting of DOT1L inhibitor (I), valproic acid (V), forskolin (F) and DZNep (Z) has been shown to increase OS-mediated reprogramming substantially by our group members. Adding I-CBP112 to this small molecule cocktail further enhanced OS-mediated reprogramming (Figure 3.2.b). These results suggest that CREBBP/EP300 bromodomain inhibitors improves reprogramming efficiency induced by OS overexpression and could potentially contribute to generating human chemically induced pluripotent stem cells.

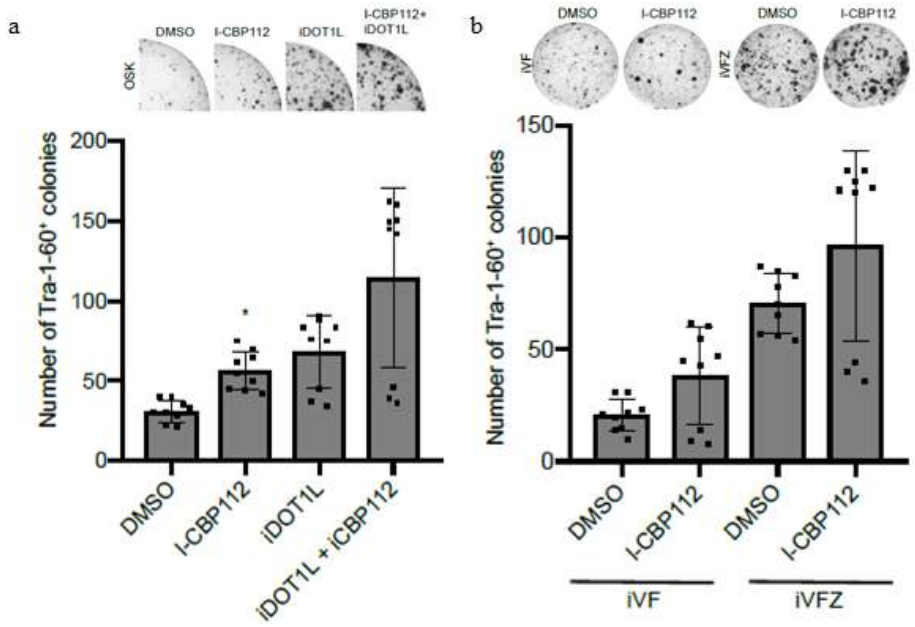


Figure 3.2 CREBBP/EP300 bromodomain inhibition increases the reprogramming efficiency induced by fewer exogenous factors. (a) Number of TRA-1-60 positive colonies generated by OSK-infected fibroblasts treated with DMSO, I-CBP112, iDOT1L and combination of I-CBP112 and iDOT1L. Bar graph shows the mean and error bars

represent s.e.m. Representative well images are above the graph. p-values were determined by a two-tailed Student's t-test; * denotes $p<0.05$. n=3, biologically independent experiments each with three technical replicates. p-values were 0.033, 0.067, 0.087 from left to right. (b) Number of TRA-1-60 positive colonies generated by OS-infected fibroblasts treated with iDOT1L (i), Valproic Acid (V), Forskolin (F) or DZNep (Z) in combination with I-CBP112. Representative well images are above the graph. n=3, biologically independent experiments each with three technical replicates.

3.3.Acetyltransferase activity of CREBBP/EP300 is necessary for iPSC generation.

To investigate the changes in histone acetylation upon CREBBP/EP300 inhibitions, I conducted a time-course treatment experiment with bromodomain inhibitors CBP30 and I-CBP112 and a potent acetyltransferase inhibitor A485 (Lasko et al., 2017). Although the global H3K27ac and H3K18ac levels fluctuate around 12 and 24 hour treatment with A485 at 1 μM concentration, overall inhibiting acetyltransferase activity of CREBBP/EP300 decreased global histone acetylation levels (Figure 3.3). However, inhibiting bromodomains of CREBBP/EP300 had minimal effect on global histone acetylation levels, only 5 day treatment with CBP30 and I-CBP112 moderately decreased the H3K27ac levels (Figure 3.3). This time-course treatment experiment showed that acetyltransferase inhibitor efficiently decreases histone targets of CREBBP/EP300 and bromodomain inhibitors has mild effect which takes 5 days to decrease global histone levels subtly.

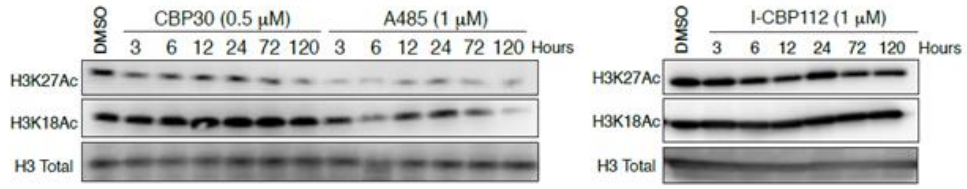


Figure 3.3 CREBBP/EP300 acetyltransferase inhibition decreases histone acetylation levels. Western blots for histone H3 lysine 27 acetylation (H3K27ac) and H3 lysine 18 acetylation (H3K18ac) in cells treated CBP30, A485 (left) or I-CBP112 (right) for the indicated numbers of hours. Total histone H3 was used as a loading control.

To investigate the catalytic function of CREBBP/EP300 in reprogramming, I utilized A485, a recently described potent acetyltransferase inhibitor of these factors. First, I assayed for inhibition of the HATs by treating fibroblasts with increasing concentrations of A485 and examining global H3K27ac levels by western blot. Starting

with 1 µM concentration, acetyltransferase activity of CREBBP/EP300 was significantly inhibited by A485 whereas the inactive control compound (A486) treatment did not change the acetylation levels (Figure 3.4.a).

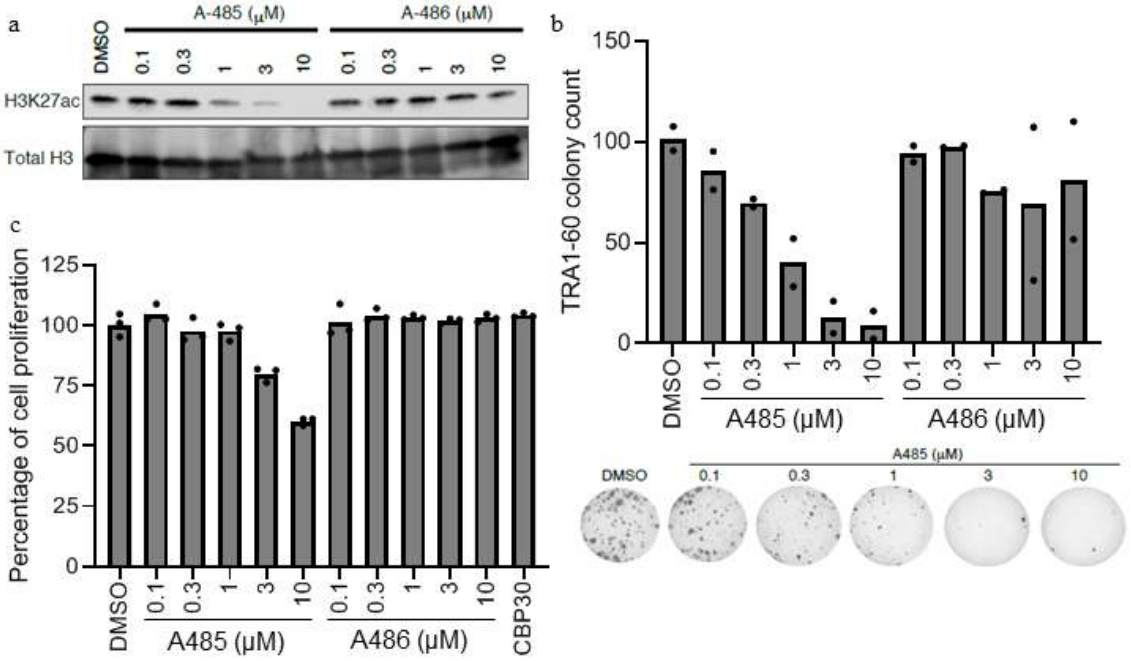


Figure 3.4 Acetyltranferase activity of CREBBP/EP300 is necessary of iPSC generation. (a) Western blots for histone H3 lysine 27 acetylation (H3K27ac) in cells treated with indicated concentrations of A485 and A486 for 5 days. Total histone H3 was used as a loading control. (b) Number of TRA-1-60 positive colonies generated upon treatment with A485 and A486 at indicated concentrations for 5 days. Bar graphs show mean. Representative well images are shown below the graph. n=2, biologically independent experiments each with three technical replicates. (c) Proliferation rate of cells treated with A485 and A486 at the indicated concentrations and CBP30 relative to DMSO control after 5 day treatment. Bar graphs show the mean. Experiments were performed with Gülben Gürhan Sevinç.

To assess the role of acetyltransferase activity of CREBBP/EP300 in reprogramming, we treated cells with increasing concentrations of A485 and A486 for the first week of reprogramming. While control A486 had no effect, A485 decreased repogramming efficiency in a dose-dependent manner (Figure 3.4.b). We, also, observed a 20-40% decrease in cell proliferation rate of fibroblasts starting at 3 µM A485, whereas, A486 or CBP30 treatment had no effect (Figure 3.4.c). At the lower 1 µM dose, A485 did not change cell proliferation rate, but decreased both histone acetylation levels and

reprogramming efficiency more than half compared to control treatment. These results suggest that acetyltransferase activity of CREBBP/EP300 is necessary for iPSC generation and somatic cell proliferation.

To examine if acetylation, itself, is necessary for successful iPSC generation, we knocked down enzymes responsible for producing cytosolic acetyl-CoA which is used by acetyltransferases as the acetyl donor. At least 50% knockdown was achieved for each of three shRNAs targeting ACLY or ACSS2 enzymes as assessed by RT-qPCR (Figure 3.5.a, Figure 3.5.b.). Suppression of ACSS2 proteins levels were confirmed by immunoblotting (Figure 3.5.c). Knockdown of either ACLY or ACSS2 decreased reprogramming efficiency to approximately 50% compared to control shRNA expression (Figure 3.5.d). These results suggest that acetyl-CoA producing enzymes are necessary for iPSC generation.

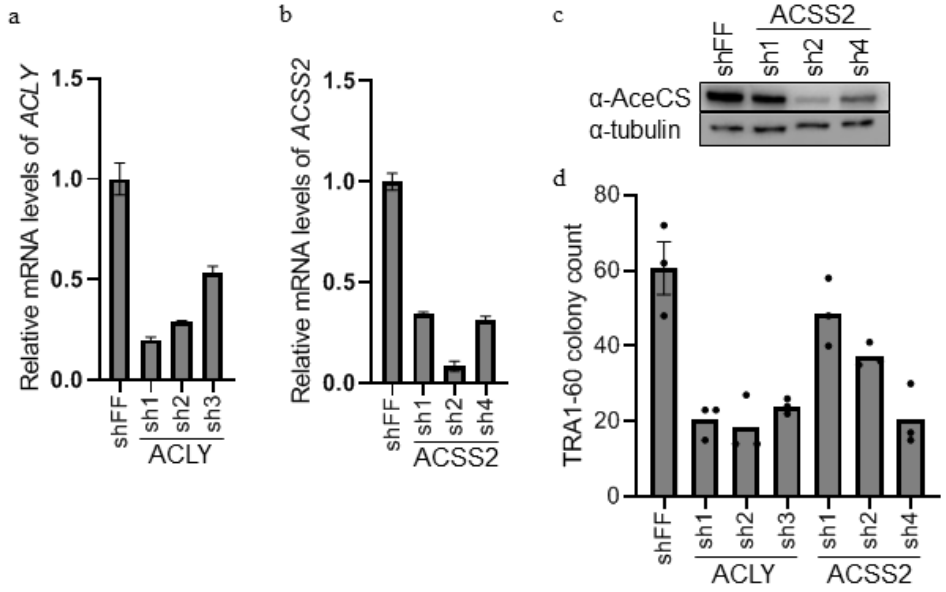


Figure 3.5 Knocking down enzymes responsible for acetyl-CoA production impairs reprogramming. (a) Fold change in expression levels of ACLY by indicated shRNA expressions compared to shFF control expression. (b) Fold change in expression levels of ACSS2 by indicated shRNA expressions compared to shFF control expression. (c) Western blots for ACSS2 from cells expressing indicated shRNAs. Tubulin serves as a loading control. (d) TRA-1-60 positive colonies generated from fibroblasts expressing indicated shRNAs. Bar graph shows the mean of 3 technical replicates. These experiments were performed together with undergraduate students Zeynep Ece Yıldız and Öznur Beyhun

3.4.Inhibiting acetyltransferase activity of CREBBP/EP300 has more effect on transcriptome than bromodomain inhibition.

To understand the transcriptional effects of inhibiting CREBBP/EP300 bromodomain-mediated interactions as well as their acetyltransferase function, I performed RNA-seq of fibroblasts and reprogramming cells at 6th day of OSKM overexpression treated with CREBBP/EP300 inhibitors for 5 days. To understand if there is any problem associated with sequencing and analysis, I performed GSEA on pre-ranked gene list comparing OSKM-induced fibroblasts to uninduced fibroblasts. As expectedly, EMT gene set was negatively and embryonic stem cell related gene set was positively enriched by OSKM induction (Figure 3.6.a).

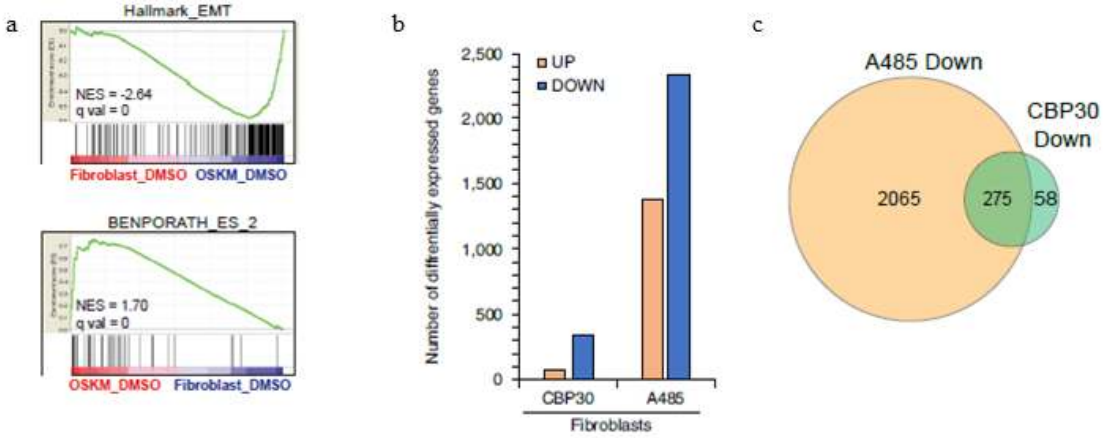


Figure 3.6 Inhibiting acetyltransferase activity of CREBBP/EP300 has more effect on transcriptome than bromodomain inhibition. (a) GSEA analysis of control DMSO-treated fibroblast upon OSKM expression on day 6 with respect to the indicated gene sets. (b) Number of differentially expressed genes upon treatments of fibroblasts with CBP30 or A485 for 5 days compared to DMSO control treatment. (c) Venn-diagram shows the number of common downregulated genes upon CBP30 and A485 treatment of uninfected fibroblasts. RNA librariers were generated and sequenced with the help from Dr. Martin Philpott and RNA-Seq analysis was performed by Dr. Adam Cribbs.

To understand the different effects of acetyltransferase and bromodomain of CREBBP/EP300 on reprogramming phenotype, number of differentially expressed genes were compared. This comparison showed a wider impact of inhibiting acetyltransferase function compared to inhibiting bromodomain of CREBBP/EP300 (Figure 3.6.b). More than 80% of differentially downregulated genes by CREBBP/EP300 bromodomain inhibition were downregulated by acetyltransferase inhibition (Figure 3.6.c). These

results suggest that CREBBP/EP300-induced acetylation has wider impact on transcriptome and CREBBP/EP300 mediated acetyllysine interactions are largely a subset of acetyltransferase activity.

3.5.CREBBP/EP300 bromodomain inhibition downregulates somatic cell-specific gene expression.

To understand the transcriptional changes upon CREBBP/EP300 bromodomain inhibition, I performed GSEA for the pre-ranked gene list comparing CBP30 treatment to DMSO treatment. This analysis showed that EMT-related gene set was negatively enriched by CREBBP/EP300 bromodomain inhibition (Figure 3.7.a). Furthermore, genes downregulated in both fibroblasts treated with CBP30 and reprogramming cells treated by either CBP30 or I-CBP112 were enriched in EMT, tissue development and extracellular structure organization related gene sets (Figure 3.7.b and Figure 3.7.c). These gene ontology and GSEA results show that CREBBP/EP300 bromodomain maintains fibroblast identity related transcriptional program as this cell type is mesenchymal.

To specifically address if CREBBP/EP300 bromodomain-mediated acetyllysine interactions are important in maintaining somatic cell-specific transcriptional program, we generated a fibroblast-related gene set from genes based on genes that were highly expressed ($|logFC|>3$ and $p<0.05$; n=286 genes) in fibroblasts compared to pluripotent stem cells (Figure 3.8.a). As expectedly, fibroblast-related genes were highly enriched in EMT and extracellular matrix-related gene sets (Figure 3.8.b). I performed GSEA for fibroblast-related gene set on pre-ranked gene lists generated by comparing CBP30, I-CBP112 and A485 treatments to DMSO control treatment. I observed a significant and negative enrichment of this gene set upon CREBBP/EP300 inhibitions with bromodomain inhibition being more potent than acetyltransferase inhibition (Figure 3.9.a and Figure 3.9.b). Similar to fibroblast-related gene set, I also generated a pluripotency-related gene set from highly expressed genes in pluripotent stem cells compared to fibroblasts (n=246 genes). Upon GSEA analysis, I observed that highly expressed genes in pluripotent cells were significantly and positively enriched by CREBBP/EP300 bromodomain inhibition. (Figure 3.9.c). On the other hand, this gene set was significantly and negatively enriched with A485 treatment. Average expression values for all pluripotency-related genes indicates that bromodomain inhibition upregulates and acetyltransferase inhibition downregulates this set of genes (Figure 3.9.d). Overall, these results showed that both CREBBP/EP300 inhibitions during reprogramming facilitates

the suppression of somatic cell-specific transcription while bromodomain inhibition favors iPSC generation through upregulating pluripotency-specific genes.

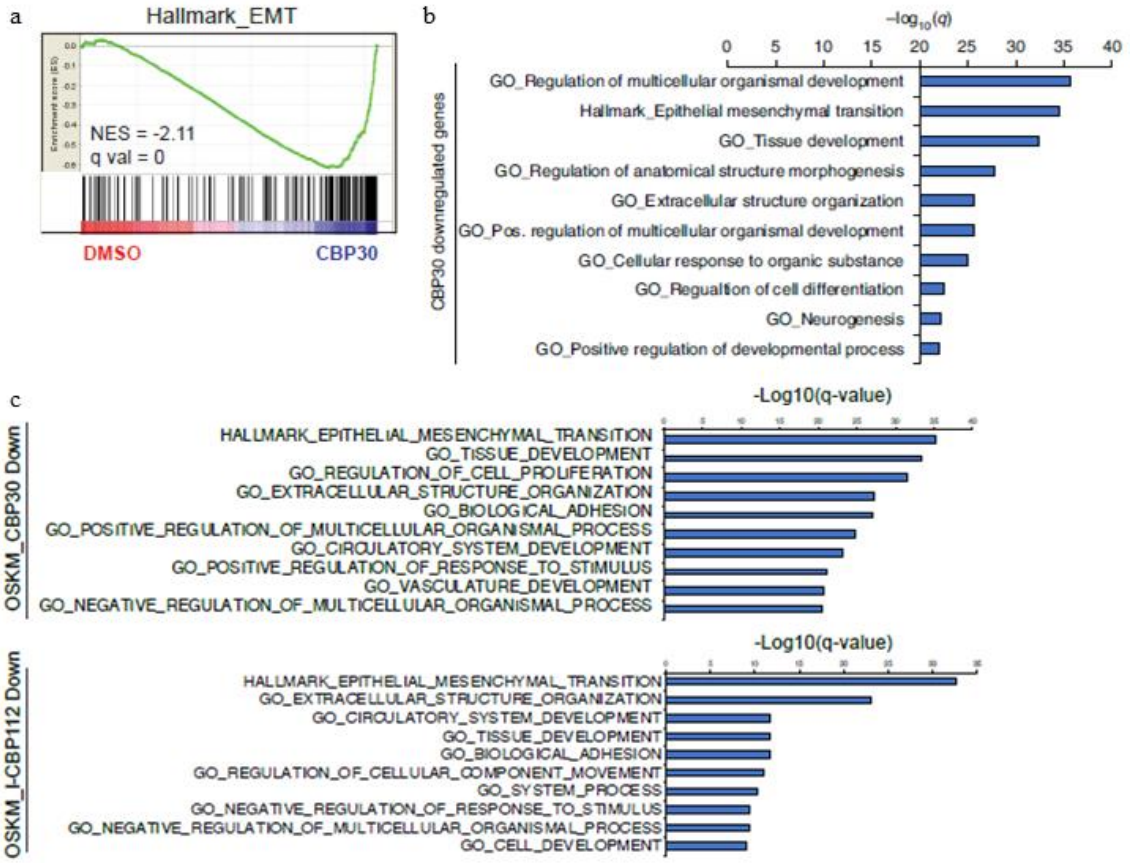


Figure 3.7 CREBBP/EP300 bromodomain inhibition downregulates genes related to EMT. (a) GSEA analysis of CBP30-induced changes in uninfected fibroblasts with respect to the Hallmark_EMT gene set. p-values for the significance of enrichment scores was calculated by permuting the gene sets, 1000 permutations, and false discovery rate (q-val) was used as an adjustment for multiple hypothesis testing. (b) Top ten gene ontology (GO) and hallmark gene sets enriched in downregulated genes upon CBP30 treatment. p-values were calculated by hypergeometric distribution. Number of genes (n) in comparison were 316. (c) Top ten GO and hallmark gene sets associated with downregulated genes on day 6 of reprogramming with CBP30 (top) or I-CBP112 (bottom) treatments. There were 427 and 118 genes respectively in each group. p-values were calculated by hypergeometric distribution.

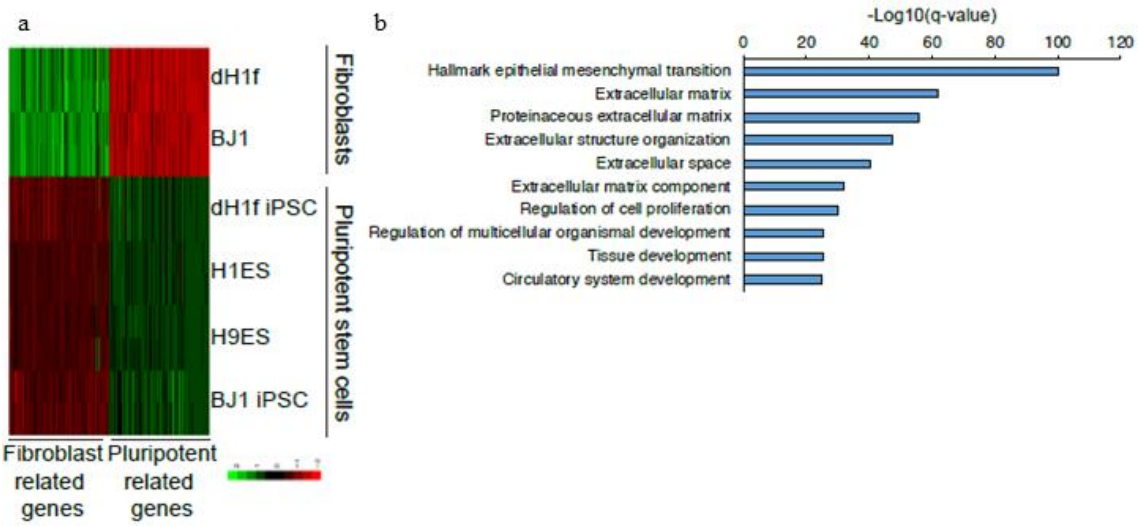


Figure 3.8 Fibroblast- and pluripotency-related gene sets are generated from genes highly expressed in respective cell types. (a) Heatmap for highly differentially expressed genes between fibroblasts and pluripotent stem cell lines. (b) Top ten GO and Hallmark gene sets enriched in fibroblast-related gene set. Number of genes (n) in comparison were 286. p-values were calculated by hypergeometric distribution. Microarray analysis was performed by Tunç Morova.

I, then, asked whether downregulation of somatic cell-specific transcription upon CREBBP/EP300 bromodomain inhibition depends on OSKM expression, or whether bromodomain inhibition, itself, can weaken somatic cell identity and thus facilitate reprogramming. To address this question, I performed GSEA for fibroblast-specific gene set on pre-ranked gene list generated by comparing CBP30 to DMSO treatment in OSKM naïve (non-OSKM) fibroblasts. I observed a significant and negative enrichment of this gene set. Furthermore, average expression of fibroblast-related genes including important developmental regulators and mesenchyme-specific genes such as *LUM*, *POSTN*, *PRRX1*, *GREM1* and *DKK1* were downregulated upon CREBBP/EP300 bromodomain inhibition (Figure 3.10). To understand if CREBBP/EP300 bromodomain inhibition is sufficient to upregulate pluripotency-related gene expression, GSEA with pluripotency-specific gene set was performed. However, the CREBBP/EP300 bromodomain inhibition without OSKM is not enough to activate pluripotency-circuit (Figure 3.10.a). Together, these results showed that CREBBP/EP300 bromodomain inhibition downregulates somatic cell-specific genes independent of OSKM and facilitate upregulation pluripotency-specific genes only in the presence of OSKM compared to relative control treatment.

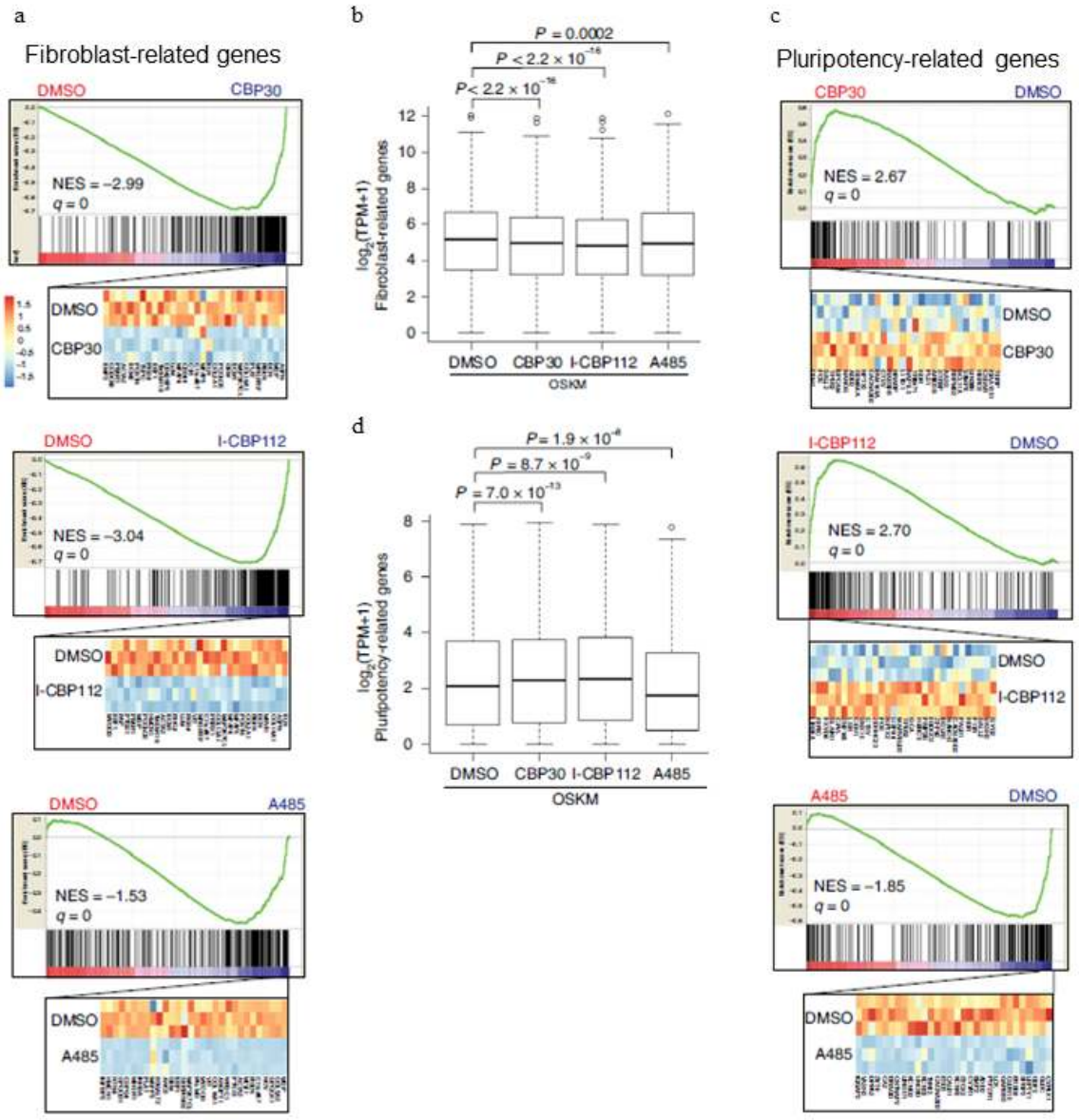


Figure 3.9 Pluripotency-related genes are expressed more whereas fibroblast-related gene expression decreases upon CREBBP/EP300 bromodomain inhibition during reprogramming. (a) GSEA analysis of gene expression changes induced by CBP30, I-CBP112 or A485 treatment 6 days after OSKM expression with respect to fibroblast-related gene set. p-values for the significance of enrichment scores were calculated by permuting the gene sets (1,000 permutations) and false discovery rate (q-val) was used as an adjustment for multiple hypothesis testing. Heatmaps show the relative expression of the top 30 genes gene set. NES, normalized enrichment score. (b) Average expression levels of the 286 genes in the fibroblast-related gene set in OSKM-transduced fibroblasts upon treatment with the indicated compounds on day 6 of reprogramming. Whiskers indicate 95% confidence interval. p-values were calculated by a two-sided Wilcoxon signed-rank test. (c) GSEA analysis of gene expression changes

induced by CBP30, I-CBP112 or A485 treatment 6 days after OSKM expression with respect to pluripotency-related gene set. p-values for the significance of enrichment scores were calculated by permuting the gene sets (1,000 permutations) and false discovery rate (q-val) was used as an adjustment for multiple hypothesis testing. Heatmaps show the relative expression of the top 30 genes gene set. NES, normalized enrichment score. (d) Average expression levels of the 246 genes in the pluripotency-related gene set in OSKM-transduced fibroblasts upon treatment with the indicated compounds on day 6 of reprogramming. Whiskers indicate 95% confidence interval. p-values were calculated by a two-sided Wilcoxon signed-rank test.

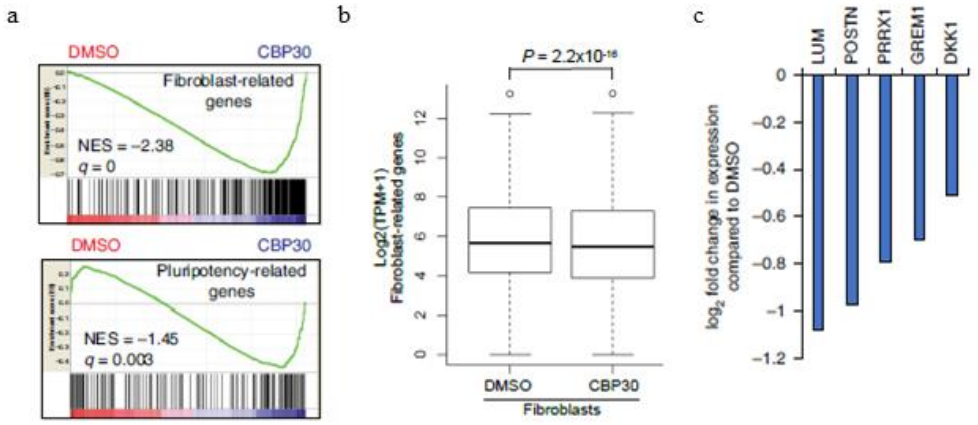


Figure 3.10 Pluripotency-related genes are expressed more whereas fibroblast-related gene expression decreases upon CREBBP/EP300 bromodomain inhibition during reprogramming. (a) GSEA analysis of CBP30-induced changes in uninfected fibroblasts with respect to fibroblast-related gene set (top) and pluripotency-related gene set (bottom). RNA-sequencing was carried out on three biologically independent samples. p-values for the significance of enrichment scores were calculated by permuting the gene sets (1,000 permutations) and false discovery rate (q-val) was used as an adjustment for multiple hypothesis testing. (b) Average expression levels of the 286 genes in the fibroblast-related gene set in DMSO and CBP30 treated cells. Whiskers indicate 95% confidence interval. p-values were calculated by a two-sided Wilcoxon signed-rank test. (c) $\log_2$ fold change in the expression levels of the indicated genes by CBP30 treatment compared with DMSO treatment.

3.6.Genome-wide chromatin changes upon CREBBP/EP300 bromodomain inhibition.

To characterize genome-wide changes in histone acetylation levels around active

promoters, putative enhancers and downstream transcriptional targets of CREBBP/EP300, I conducted quantitative chromatin immunoprecipitation assays followed by sequencing (ChIP-Seq) for H3K18ac, H3K27ac, H3K4me1 and H3K4me3. In addition, assay for transposase accessible chromatin followed by sequencing (ATAC-Seq) was performed upon 5 day CBP30 treatment in fibroblasts to assess changes in genome wide chromatin accessibility. Aggregate analysis of H3K27ac peaks around transcription start sites (TSS) demonstrated a decrease of this histone mark around TSS upon CREBBP/EP300 bromodomain. H3K18ac enrichment across the same regions were lower in magnitude and was slightly elevated upon inhibitor treatment (Figure 3.11.a and Figure 3.11.b). I next asked whether H3K27ac was reduced at accessible chromatin upon inhibitor treatment. I observed a decrease in H3K27ac levels around accessible chromatin upon CREBBP/EP300 bromodomain inhibition (Figure 3.11.c). The decrease in H3K27ac levels was observed around putative primed/active enhancers marked by H3K4me1, too (Figure 3.11.d). These results suggest that CREBBP/EP300 bromodomain inhibition decreases H3K27ac levels around regulatory elements.

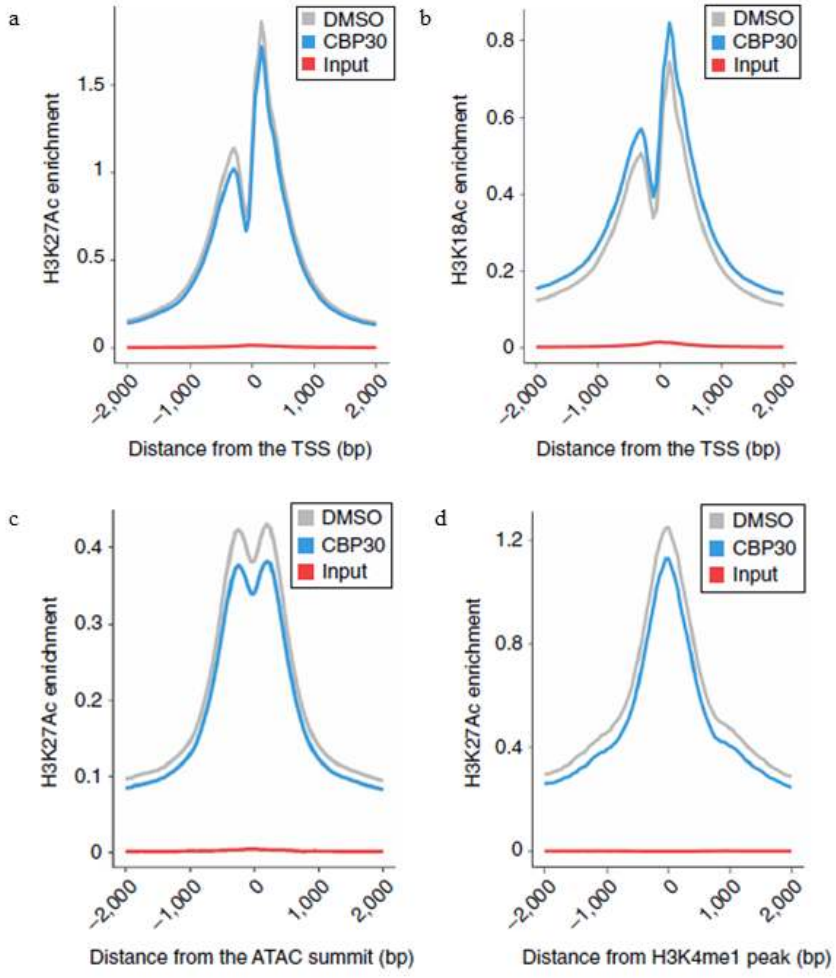


Figure 3.11 Genome-wide chromatin changes upon CREBBP/EP300 bromodomain

inhibition. (a) Aggregate H3K27ac profile plots showing enrichment centered at the transcription start sites (TSS) of CBP30-downregulated genes by DMSO and CBP30 treatments for 5 days. (b) Aggregate H3K18ac profile plots showing enrichment centered at the TSS of CBP30-downregulated genes by DMSO and CBP30 treatments for 5 days. (c) Aggregate plots of H3K27ac coverage centered at the consensus ATAC-seq peaks by DMSO and CBP30 treatments for 5 days. (d) Aggregate plots of H3K27ac coverage centered at H3K4me1-marked regions by DMSO and CBP30 treatments for 5 days. DNA libraries were generated and sequenced with Dr. Martin Philpott. ATAC and ChIP-Seq analysis was performed by Dr. Adam Cribbs.

Effect of CREBBP/EP300 bromodomain inhibition on regulatory elements of downstream transcriptional targets was investigated. RNA, accessible chromatin and H3K27ac peaks were decreased around promoter regions of mesenchyme-specific, CREBBP/EP300 bromodomain targets genes such as *LUM* and *POSTN* (Figure 3.12). Similarly, accessible chromatin, H3K27ac and H3K4me1 levels were decreased around previously defined fibroblast-enhancers (Figure 3.12). These results show that CREBBP/EP300 bromodomain maintains somatic cell-specific transcriptional program through regulating accessible chromatin, active histone marks such as H3K27ac and H3K4me1 around both enhancers and promoters of related genes.

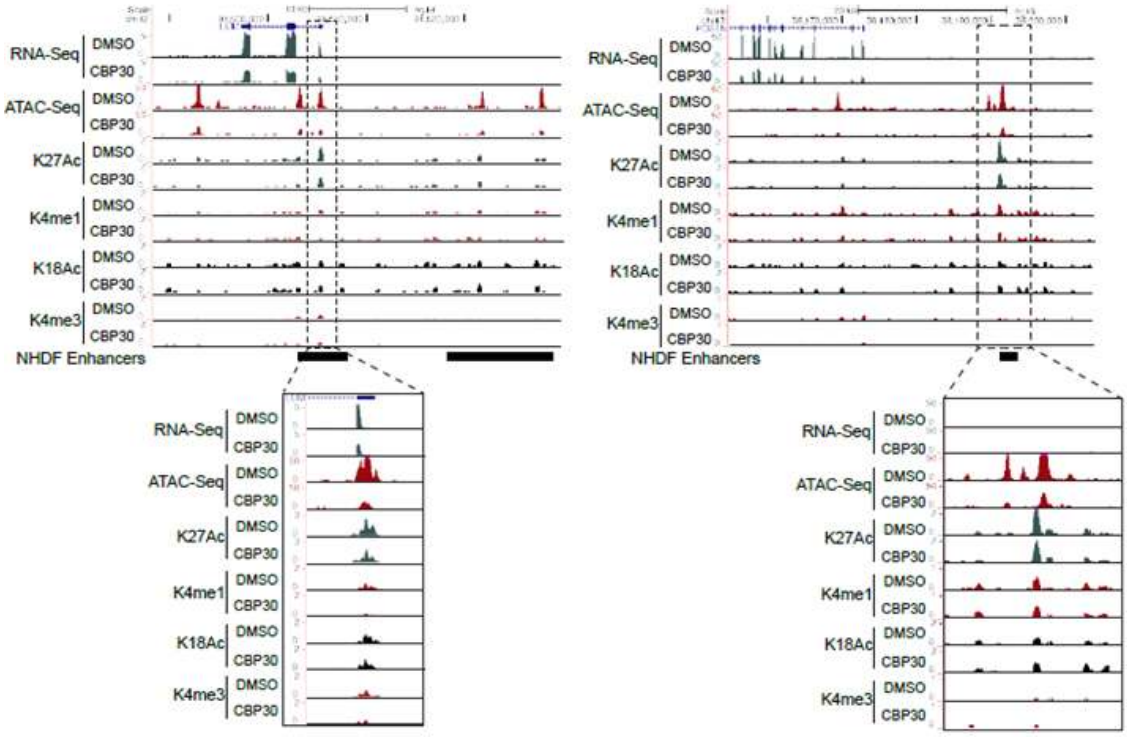


Figure 3.12 Examples of genome-wide sequencing data on a set of CREBBP/EP300

putative target genes. Aggregate ChIP-seq, ATAC-seq, RNA-Seq profile changes associated with LUM (left) and POSTN (right) in fibroblasts following treatment with DMSO or CBP30. The zoomed in section below shows region around the TSS and enhancer in NHDFs. The UCSC browser view of the ChIP-seq intensity from one of three independent experiments is shown.

3.7.Continuous expression of PRRX1 blocks somatic cell reprogramming

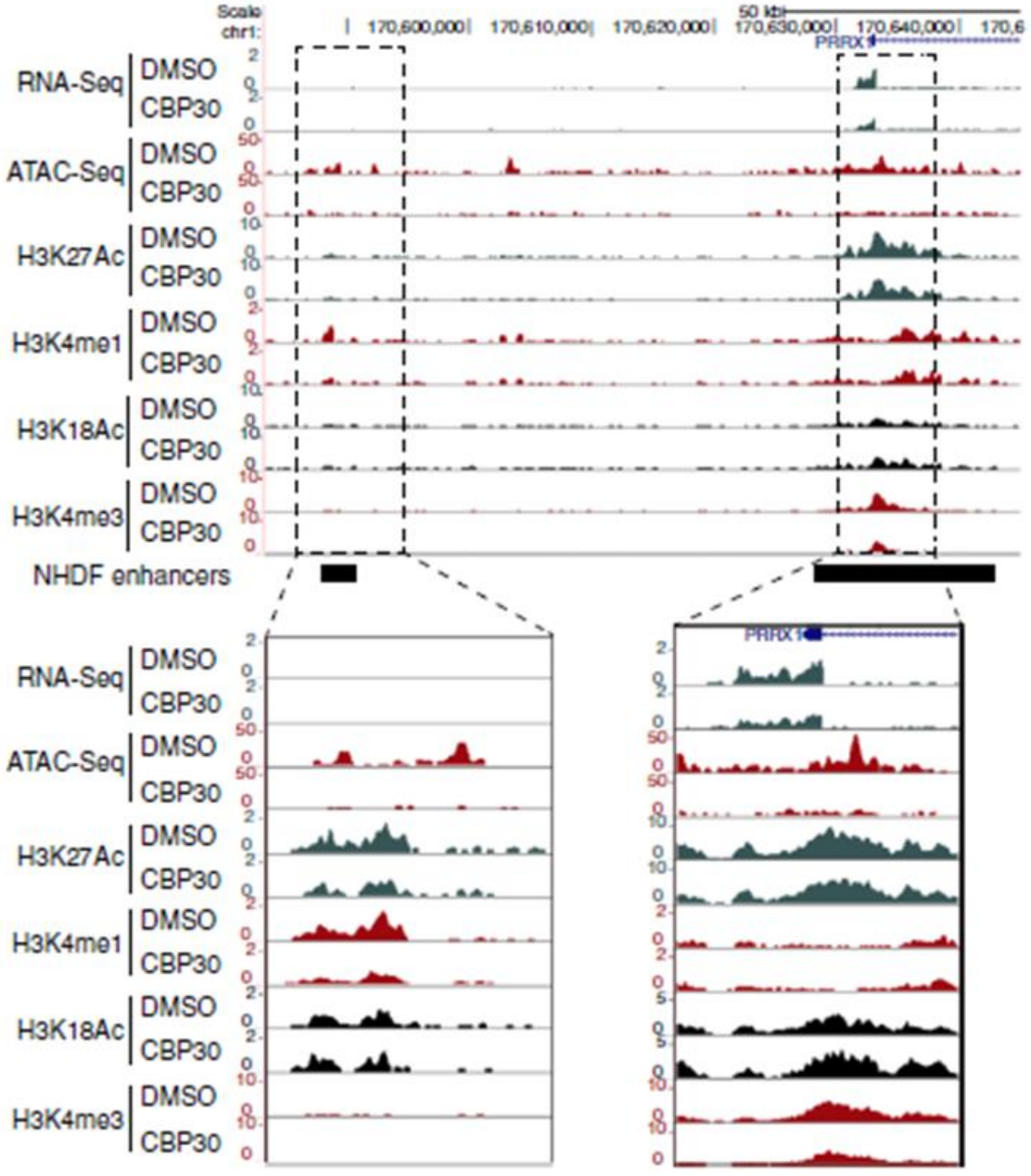


Figure 3.13 Aggregate ChIP-seq, ATAC-seq, RNA-Seq profile changes associated with PRRX1 in fibroblasts following treatment with DMSO or CBP30. The zoomed in section below shows region around the TSS and enhancer in NHDFs. The UCSC browser view of the ChIP-seq intensity from one of three independent experiments is

shown. ATAC-seq tracks is from one of two independent samples. The locations of NHDF (normal adult human diploid fibroblasts) enhancers and TSS are zoomed in.

To further understand the mechanistic basis for barrier functions of CREBBP/EP300 bromodomain mediated interactions in somatic cell reprogramming, I focused on downregulated transcriptional targets upon CBP30 treatment. PRRX1, a master transcription factor of fibroblasts, was downregulated by CBP30 treatment. Both promoter and enhancer regions of this gene were direct targets of CREBBP/EP300 bromodomain as assessed by decreasing levels of accessible chromatin, H3K27ac and H3K4me1 marks upon CBP30 treatment (Figure 3.13).

To assess the role of PRRX1 in somatic cell reprogramming, I performed shRNA-mediated loss-of-function experiments targeting PRRX1. The knockdown efficiencies by 6 different shRNAs were between 50 and 80% except for PRRX1 sh2 which was eliminated from further experiments (Figure 3.14.a). I observed that the remaining 4 out 5 PRRX1 shRNAs resulted increased reprogramming efficiency (Figure 3.14.b). Interestingly, except for PRRX1 sh7 whose knockdown efficiency was around 50%, I observed a decline in the effect of CBP30 on reprogramming efficiency upon PRRX1 knockdown. This set of loss-of-function experiments showed that one of the targets of CREBBP/EP300, PRRX1, itself is likely to be a barrier to reprogramming.

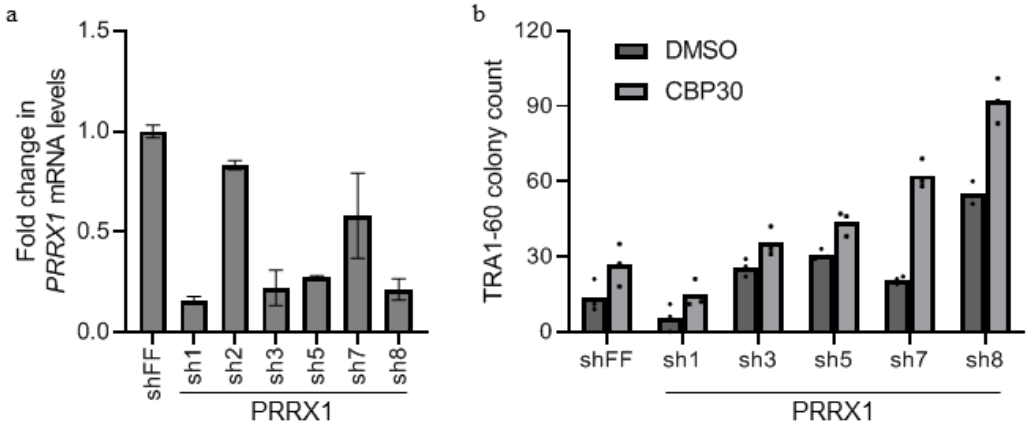


Figure 3.14 Knocking down *PRRX1* enhances somatic cell reprogramming. (a) Fold change in expression levels of *PRRX1* by indicated shRNA expressions compared to shFF control expression. (a) Number of TRA-1-60 positive colonies upon indicated shRNA expressions by CBP30 or DMSO treatments. Bar graphs show the average of three technical replicates.

Next, I hypothesized that CBP30-mediated downregulation of *PRRX1* is necessary for successful reprogramming. To address this notion, I overexpressed *PRRX1* in fibroblasts and reprogram them to iPSCs (Figure 3.15.a). I observed a significant decrease in reprogramming efficiency at both early and late time points of the process even in the presence of CREBBP/EP300 bromodomain inhibitors (Figure 3.15.b, Figure 3.15.c and Figure 3.15.d). To understand the barrier nature of continuous PRRX1 expression in reprogramming, I checked the expression levels of epithel and pluripotency marker genes. I observed that these markers failed to get induced during reprogramming when PRRX1 was overexpressed (Figure 3.15.e). These results indicated that continuous expression of a CREBBP/EP300 bromodomain target gene, PRRX1, prevents activation of epithelial and pluripotency genes and impairs reprogramming. Taken together with loss of function experiments, I concluded that CREBBP/EP300 acts to suppress somatic cell reprogramming, in part, by regulating the expression of the mesenchymal master regulator PRRX1.

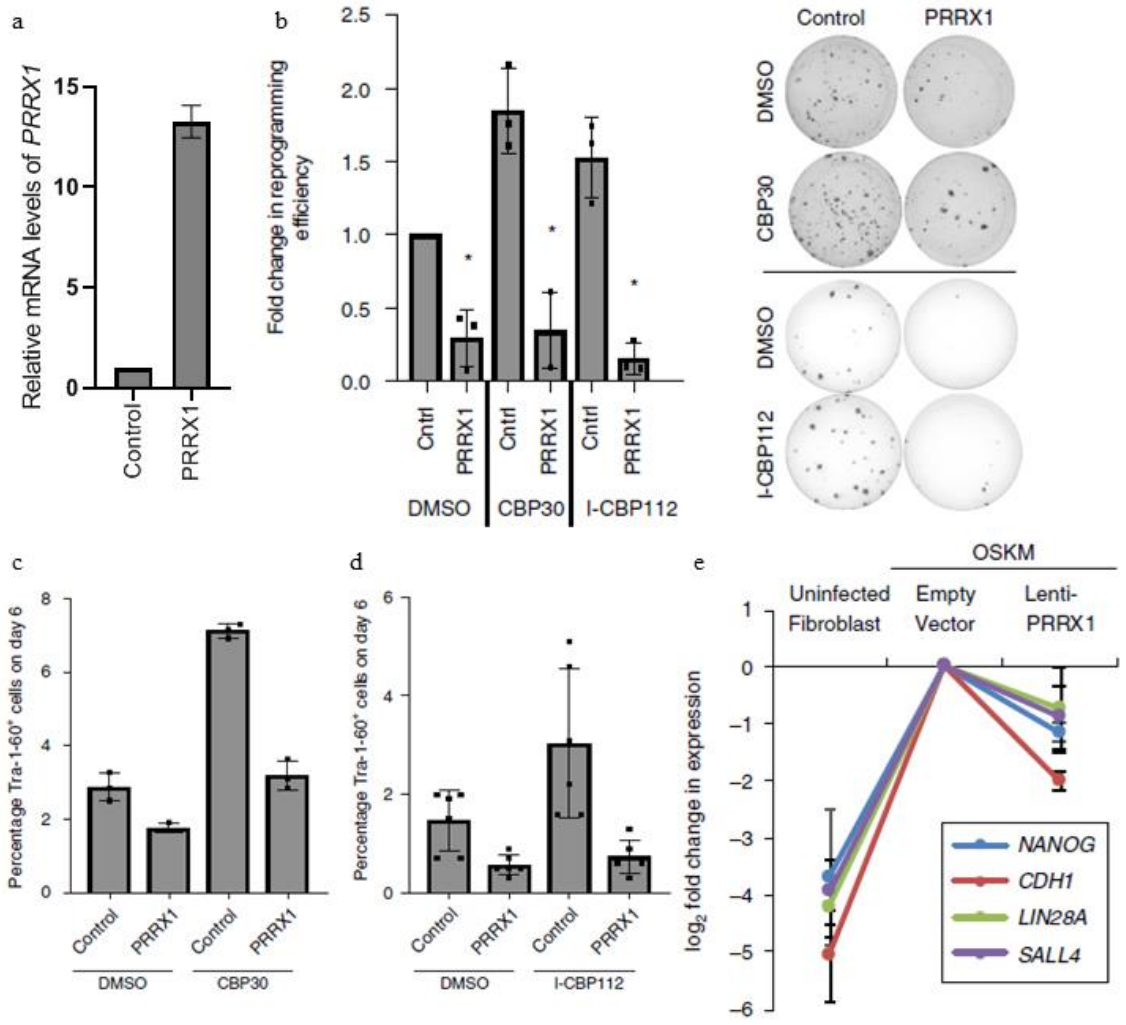


Figure 3.15 Continuous expression of *PRRX1* blocks somatic cell reprogramming.

(a) Fold change in expression levels of *PRRX1* upon overexpression. (b) Fold change in the reprogramming efficiency of *PRRX1*-expressing fibroblasts compared with controls with or without CBP30 and I-CBP112. Bar graphs show the mean and error bars represent standard deviation. Representative well images are on right side of graph. p-values were determined by a two-tailed Student's t-test; * denotes $p<0.05$ compared with control DMSO. n=3, independent biological experiments. p-values were 0.0238, 0.0481, 0.005 from left to right. (c) Percentage TRA-1-60 positive cells as analyzed by flow cytometry on day 6 of reprogramming of control and PRRX1-expressing fibroblasts treated with DMSO or CBP30. Bar graphs show the mean and error bars represent standard deviation. n=3, independent biological experiments. (d) Percentage TRA-1-60 positive cells as analyzed by flow cytometry on day 6 of reprogramming of control and PRRX1-expressing fibroblasts treated with DMSO or I-CBP112. Bar graphs show the mean and error bars represent standard deviation. n = 2, independent biological experiments each with three technical replicates. (e) log2 fold change in relative mRNA levels for the indicated genes on day 6 of reprogramming. Gene expression values were normalized to those observed in empty vector expressing fibroblasts. Error bars represent standard deviation. n=3, biologically independent experiments. These experiments were conducted with Gülben Gürhan Sevinç.

Chapter 4. INVESTIGATING THE ROLE OF BRD9 IN REPROGRAMMING

4.1.SWI/SNF focused CRISPR screen identified ncBAF complex as a barrier to human somatic cell reprogramming to pluripotency.

To interrogate the role of SWI/SNF complexes in human somatic cell reprogramming, I performed a focused CRISPR-Cas9 mediated knockout screen. First, I assessed the effect of individual knockout on human fibroblast cell proliferation. I observed that ACTB and ACTL6A knockouts decrease cell proliferation rates by 30-40% whereas most other knockouts have no impact on cell proliferation (Figure 4.1.a). One gRNA for each of BRD9, BCL7A, BCL7B and SS18L1 decreased cell proliferation rate by about 20% (Figure 4.1.a). These cell proliferation results show that ACTB and ACTL6A are necessary to sustain fibroblast growth and further data are needed to conclude a role for BRD9, BCL7A/B and SS18L1.

To reveal contribution of each SWI/SNF-related gene to human somatic cell reprogramming, sgRNA-expressing fibroblasts were subjected to reprogramming. Compared to non-targeting control, a 50% decrease in the efficiency of reprogramming achieved by both of two sgRNAs per gene was set as a threshold to determine if the target gene is necessary for pluripotency acquisition. Using this criteria, common SWI/SNF subunits ACTB, ACTL6A, BCL7B, SMARCA4, SMARCC1 and BAF-specific subunit DPF2 were identified as necessary factors to induce pluripotency (Figure 4.1.b). Conversely, if a 50% increase in the efficiency of reprogramming was achieved by both of two sgRNAs targeting the same gene, then that gene was considered to be a barrier to pluripotency acquisition. ncBAF-specific subunits BRD9 and GLTSCR1 met this criteria (Figure 4.1.b) and their role in somatic cell reprogramming was further investigated with additional sgRNAs. In conclusion, this SWI/SNF-focused knockout screen identified a select set of common subunits including SMARCA4 and SMARCC1 as necessary factors and ncBAF-specific subunits BRD9 and GLTSCR1 as barriers to human pluripotency induction.

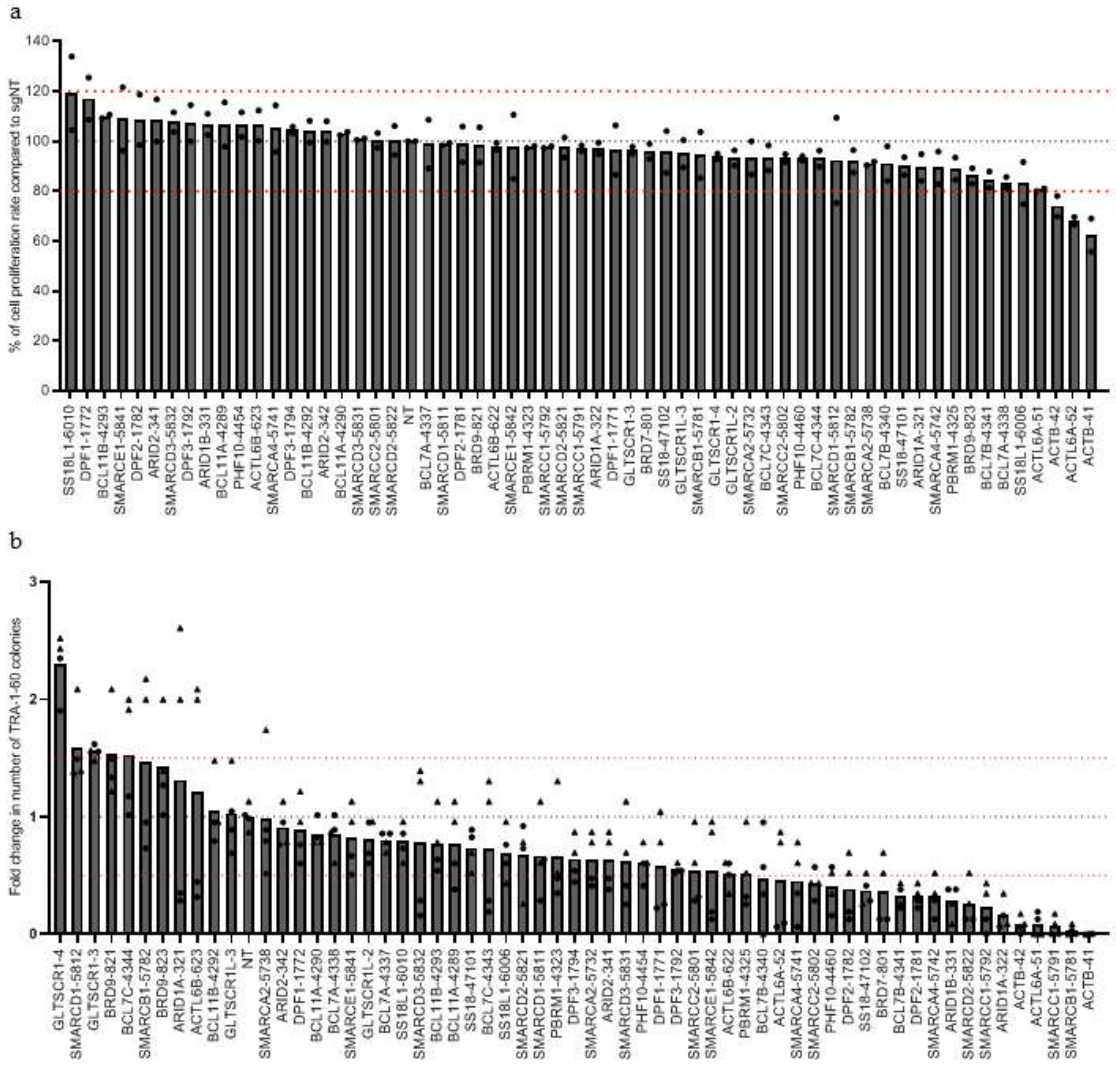


Figure 4.1 SWI/SNF focused CRISPR-Cas9 mediated knockout screen identified ncBAF complex-specific members as barriers to reprogramming. (a) Percentages of cell proliferation rates after 2 day culture for indicated gRNA expressions compared to control gRNA (sgNT1) expression. Black dashed line indicates 100% rate, red dashed lines indicate 20% change in cell proliferation rate. n=2, biological replicates in technical triplicates. (b) Fold change in TRA-1-60 positive colony count for indicated sgRNA expressions compared to control sgRNA expression. Black dashed line indicates 1-fold, red dashed lines indicate 0.5-fold change in reprogramming efficiency. n=2, biological replicates with technical duplicates. Majority of sgRNAs used in the screen were cloned by Hazal Can.

4.2.Genetic suppression of BRD9 but not BRD7 enhances human somatic cell reprogramming to pluripotency.

To validate the results of the SWI/SNF focused CRISPR-Cas9 knockout screen, I

re-performed infection of fibroblasts with lentiviruses containing Cas9 and three independent sgRNAs for BRD9 and its closely related paralog BRD7. BRD9 protein levels were suppressed with BRD9-targeting gRNAs whereas BRD7 levels remained unaffected (Figure 4.2.a). Similarly, BRD7 protein levels decreased upon CRISPR-Cas9 targeting BRD7 (Figure 4.2.b). When these knockout fibroblasts were reprogrammed to iPSCs by overexpressing OSKM, I observed that knock-out of BRD9, but not BRD7, increases reprogramming efficiency (Figure 4.2.c). These results validated my SWI/SNF focused CRISPR-Cas9 screen and showed that ncBAF-specific subunit BRD9, but not PBAF-specific subunit BRD7, is a barrier to somatic cell reprogramming.

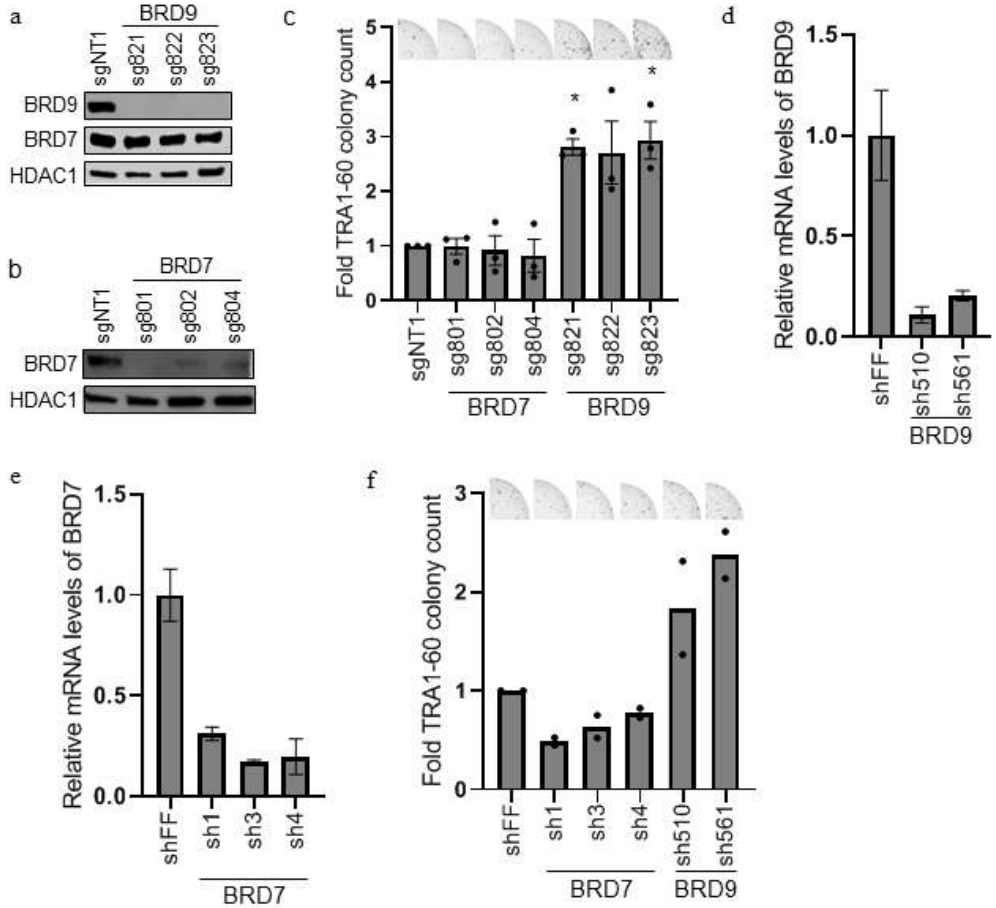


Figure 4.2 Genetic suppression of BRD9 but not BRD7 enhances human somatic cell reprogramming to pluripotency. (a) Western blots for BRD9 and BRD7 in fibroblasts expressing control (sgNT1) and BRD9-targeting gRNAs. HDAC1 serves as loading control. (b) Western blots for BRD7 in fibroblasts expressing control (sgNT1) and BRD7-targeting gRNAs. HDAC1 serves as loading control. (c) Fold change in the number of TRA-1-60-positive colonies for indicated sgRNA-expressing cells compared to control sgNT1-expressing cells. Representative well images are shown above the graph. Bar graphs show the mean and error bars represent standard error of mean. n=3, biological

replicates. p-values are calculated by one sample t-test for mu=1. * denotes p<0.05, exact p-values from left to right are: 0.9593, 0.7910, 0.6173, 0.0065, 0.0968, 0.0302. (d) RT-qPCR results for mRNA levels of BRD9 normalized to ACTB mRNA levels for indicated cells. (e) RT-qPCR results for mRNA levels of BRD7 normalized to ACTB mRNA levels for indicated cells. (f) Fold change in TRA-1-60-positive colony count normalized to shFF control. Representative well images are above the graph. n=2, two independent biological replicates with three technical replicates each. These experiments were conducted with Gülben Gürhan Sevinç.

To confirm our CRISPR-Cas9 mediated knockout of BRD9 and BRD7 results with an independent genetic approach, I utilized shRNA mediated knockdown as an additional loss-of-function assay. Upon respective shRNA expressions both BRD9 and BRD7 mRNA levels decreased below 30% compared to control shFF expression (Figure 4.2.d and Figure 4.2.e). Similar to CRISPR-Cas9 mediated knockout results, I observed that BRD9 knockdown but not BRD7 knockdown increases reprogramming efficiency (Figure 4.2.f). These results demonstrate that BRD9, but not BRD7, is a barrier to somatic cell reprogramming.

4.3.ncBAF is the specific barrier to reprograming among other SWI/SNF complexes.

To validate the results of the SWI/SNF focused CRISPR-Cas9 knockout screen, I re-performed infection of fibroblasts with lentiviruses containing Cas9 and additional independent sgRNAs to target *GLTSCR1* and its mutually exclusive subunit *GLTSCR1L*. I observed that sgRNAs targeting *GLTSCR1* increased reprogramming efficiency significantly, whereas, targeting *GLTSCR1L* did not have any effect (Figure 4.3.a). To eliminate the possibility that CRISPR-Cas9 did not target *GLTSCR1L* locus effectively, I performed T7 endonuclease assay to detect insertion-deletions (indel) at population level. At least 2 out of 3 sgRNAs targeting *GLTSCR1L* resulted in indel formation at expected genomic sites (Figure 4.3.b). These results define BRD9-GLTSCR1 containing ncBAF complex as a barrier to somatic cell reprogramming.

To examine if targeting core subunits of SWI/SNF complexes is detrimental to reprogramming, I utilized PROTAC degraders of catalytic subunits of all SWI/SNF complexes, SMARCA2/4, and observed a decrease in cell proliferation levels at the high concentrations (Figure 4.3.c). When fibroblasts were reprogrammed to pluripotency by overexpressing OSKM in the presence of these degraders, I observed a decrease in

reprogramming efficiency which may stem from the decrease in cell proliferation rates (Figure 4.3.d). However, at the lowest concentration of DL-2-4 and S31 that has the minimal effect on cell proliferation, the reprogramming efficiency decreased dramatically. These results suggest that ncBAF is the specific barrier to reprogramming and other SWI/SNF complexes may be essential for successful reprogramming.

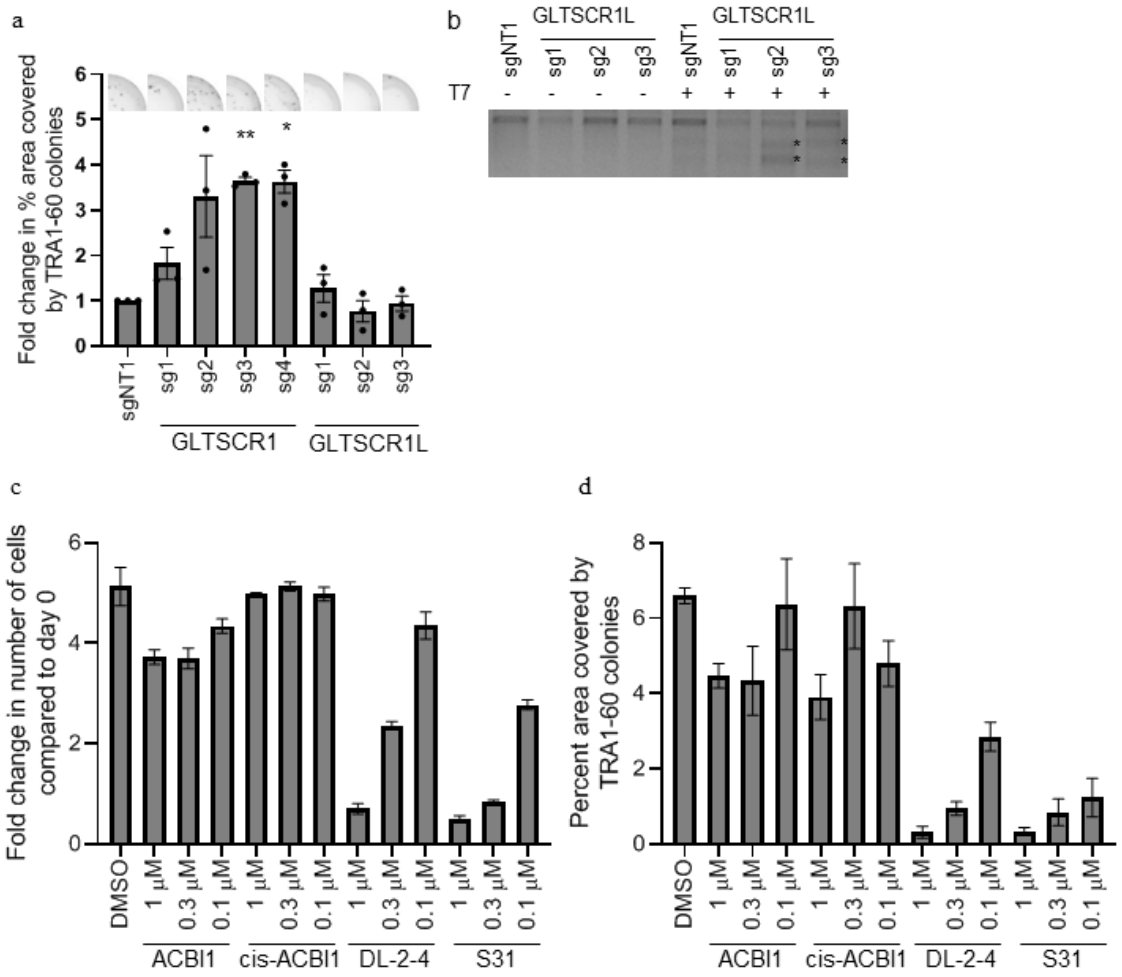


Figure 4.3 ncBAF is the specific barrier to reprograming among other SWI/SNF complexes. (a) Fold change in percent area covered by TRA-1-60-positive colonies generated from fibroblasts expressing *GLTSCR1* or *GLTSCR1L* sgRNAs compared to control sgNT1. Representative well images are above relative bars. Bar graphs show the mean and error bars represent standard error of mean. n=3, three biological replicates. p-values are calculated by one sample t-test for mu=1. * denotes $p<0.05$, ** denotes $p<0.005$, exact p-values from left to right are: 0.1442, 0.1248, 0.0009, 0.0092, 0.4638, 0.4252, 0.7585. (b) PCR products from genomic DNAs of indicated cells with and without T7 endonuclease to detect Cas9-mediated in-del formation. (c) Fold change in proliferation rates of fibroblasts treated with indicated SMARCA2/4 degraders and DMSO control compared to day 0. (d) Percent area covered by TRA-1-60-positive

colonies generated from fibroblasts treated with indicated SMARCA2/4 degraders. Bar graphs show the mean and error bars represent standard error of mean. n=1, three technical replicates.

4.4.BRD9 degradation and bromodomain inhibition increases reprogramming efficiency.

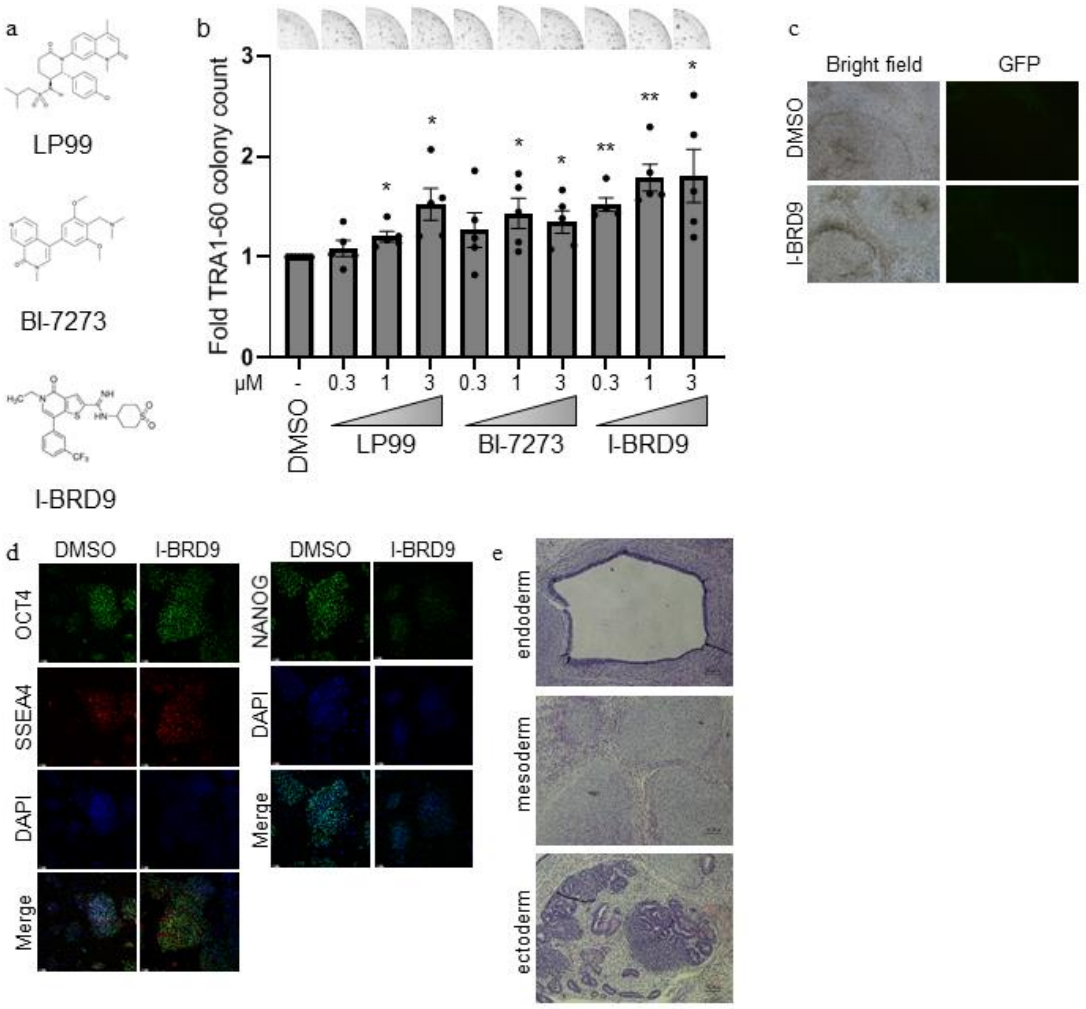


Figure 4.4 BRD9 bromodomain inhibition increases reprogramming efficiency. (a) Chemical structures of LP99, BI-7273 and I-BRD9. (b) Fold change in the number of TRA-1-60-positive colonies with the indicated compound treatments compared to DMSO control. Representative well images are shown above. Bar graphs show the mean and error bars represent standard error of mean. n=5, biological replicates for each treatment with 3 technical replicates. p-values are calculated by one sample t-test for mu=1. * denotes $p<0.05$, ** denotes $p<0.005$, exact p-values from left to right are: 0.3555, 0.0175, 0.0294, 0.1954, 0.0443, 0.0364, 0.0015, 0.0042, 0.0387. (c) Phase contrast and GFP fluorescence images of colonies derived by DMSO (top) and I-BRD9 (bottom) treatments

showing typical iPSC morphology and silencing of retroviral GFP transgene. (d) OCT4, SSEA4 (left) and NANOG (right) immunofluorescence of iPSCs derived by DMSO and I-BRD9 treatments. Hoechst 33324 was used to stain the nuclei. (e) Hematoxylin and eosin stained sections of teratomas of iPSCs derived from I-BRD9 treated-fibroblasts indicating tissues derived endoderm, ectoderm and mesoderm lineages. These experiments were conducted with Gülben Gürhan Sevinç, immunofluorescence experiments were conducted by Simge Kelekçi and teratoma sections were stained by Prof. Arzu Ruacan.

To further characterize the role of BRD9 in reprogramming and to validate the small molecule screen findings, I took advantage of small molecules, LP99, BI-7273 and I-BRD9 that inhibit the BRD9 bromodomain (Figure 4.4.a). Treatment of fibroblast with increasing concentrations of these bromodomain inhibitors, resulted in a maximum of 2-fold increase in reprogramming efficiency (Figure 4.4.b). iPSCs generated from I-BRD9 treated cells were expanded and characterized with respect to their pluripotency status. I-BRD9 treatment derived iPSCs could silence retroviral genes (Figure 4.4.c), express pluripotency markers such as OCT4, SSEA4 and NANOG (Figure 4.4.d) and contribute to cell types derived from three germ layers upon teratoma formation assay (Figure 4.4.e). These result indicate that BRD9 bromodomain inhibition during reprogramming does not impair the pluripotency of resulting iPSCs.

The original chemical screen was carried out in the presence of a DOT1L inhibitor, EPZ004777, suggesting that BRD9 inhibition can act additively with DOT1L inhibition in increasing reprogramming efficiency. To confirm this, I treated reprogramming cells with DOT1L and I-BRD9. I observed that BRD9 bromodomain and DOT1L inhibition have additive effect on reprogramming and together they increase reprogramming efficiency 5-fold (Figure 4.5.a). To further expand our repertoire to assess reprogramming efficiency from TRA-1-60 cell surface staining to reporter systems, I utilized differentiated fibroblasts derived from ESRG- or OCT4-GFP reporter pluripotency cell lines. ESRG-GFP positive colony count increased upon BRD9 bromodomain or DOT1L inhibitions, however, the combinatorial effect was somewhat lower than additive (Figure 4.5.b and Figure 4.5.c). Although the reprogrammed OCT4-GFP positive colonies were less compared to ESRG-GFP reporter line, I observed similar additive effect of BRD9 bromodomain inhibition with DOT1L inhibition (Figure 4.5.d and Figure 4.5.e). These results show that BRD9 bromodomain inhibition has additive

effect on reprogramming with DOT1L inhibition and ESRG- and OCT4-GFP pluripotent stem cell derived fibroblasts can be used to quantify reprogramming efficiency.

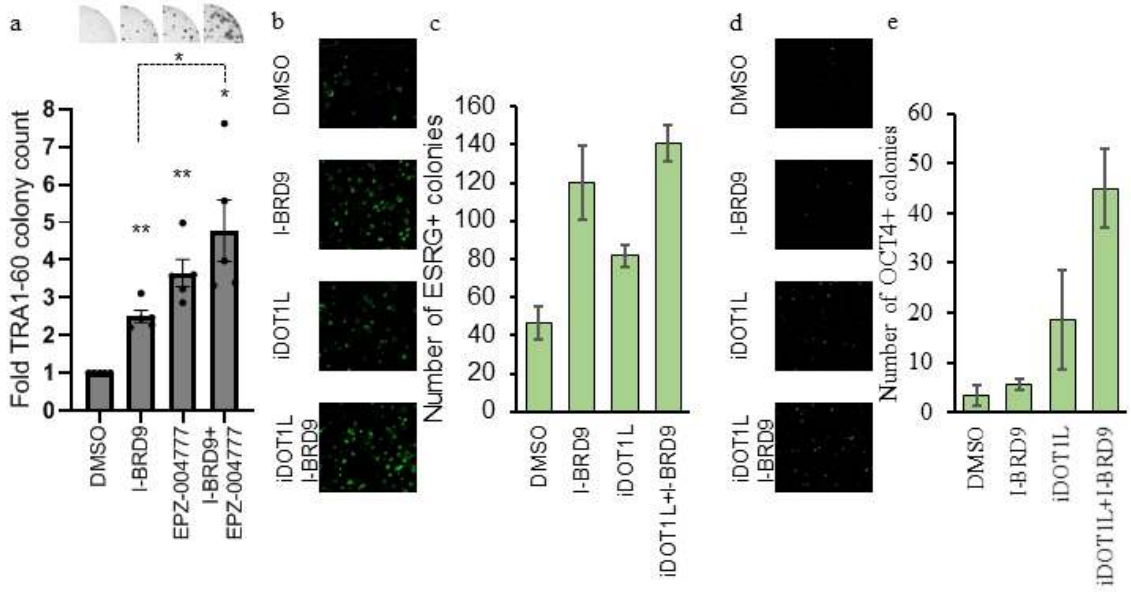


Figure 4.5 BRD9 bromodomain and DOT1L inhibition has additive effect on reprogramming assessed by different markers. (a) Fold change in the number of TRA-1-60-positive colonies upon I-BRD9, EPZ00477 (DOT1L inhibitor) and their combination treatment to DMSO. Representative well images are above the graph. Bar graphs show the mean and error bars represent standard error of mean. n=5, five independent biological replicates. p-values for DMSO comparisons are calculated by one sample t-test for mu=1. * denotes p<0.05, ** denotes p<0.005, exact p-values from left to right are: 0.0007, 0.0018, 0.0099. p-value for comparison of combination with I-BRD9 treatments is calculated by two-sample t-test. * denotes p<0.05, exact p-value is 0.0485. (b) GFP fluorescence images of ESRG-GFP positive colonies generated from indicated compound treatments. (c) Number of ESRG-GFP-positive colonies generated from fibroblasts treated with indicated compounds. Bar graphs show the mean and error bars represent standard error of mean. n=1, three technical replicates. (d) GFP fluorescence images of OCT4-GFP positive colonies generated from indicated compound treatments. (e) Number of OCT4-GFP-positive colonies generated from fibroblasts treated with indicated compounds. Bar graphs show the mean and error bars represent standard error of mean. n=1, three technical replicates.

I utilized PROTAC degrader of BRD9, dBRD9, to expand our toolkit to study BRD9's role in reprogramming (Figure 4.6.a). To evaluate the stability and minimum

time needed for degradation, fibroblasts were treated with 300 nM dBRD9 and lysates were collected after 6, 12, 24, 48 and 72 hours. We observed that dBRD9 degrades BRD9 at the 6 hour time-point and this effect was stable for at least 3 days (Figure 4.6.b). When reprogramming cells were treated with dBRD9, the number of iPSC colonies significantly increased up to 2-fold (Figure 4.6.c). To evaluate if the phenotype we observed with BRD9 inhibition and degradation on reprogramming is independent of fibroblast type and delivery method of OSKM, I reprogrammed adult human dermal fibroblasts (HDF) with episomal vectors in the presence of I-BRD9 and dBRD9. Both treatments increased adult fibroblast reprogramming efficiency with episomal vectors (Figure 4.6.d). Together, these results show that BRD9 loss increases reprogramming efficiency independent of somatic cell type and OSKM delivery method.

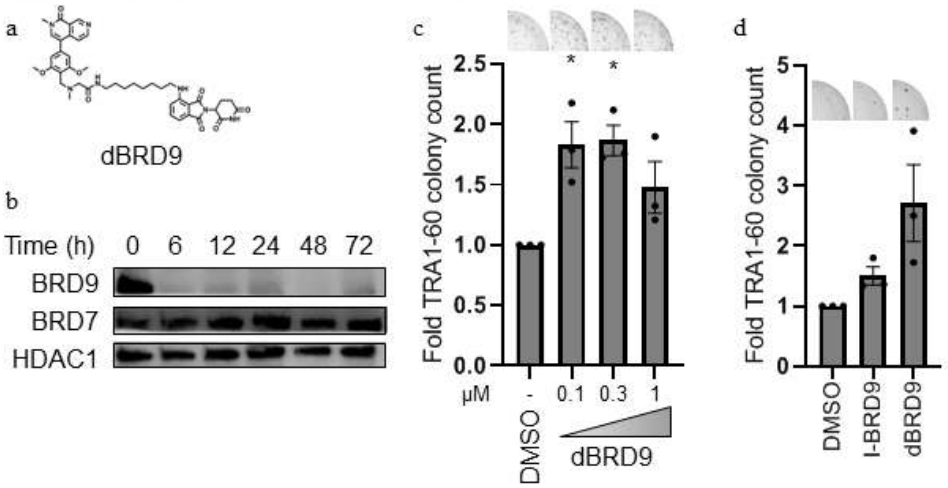


Figure 4.6 dBRD9 enhances cellular reprogramming. (a) Chemical structure of dBRD9. (b) Western blots for BRD9 and BRD7 at indicated time points after treatment with 0.3 μM dBRD9. HDAC1 serves as loading control. (c) Fold change in the number of TRA-1-60-positive colonies with increasing concentrations of dBRD9 compared to DMSO. Representative well images are above the graph. Bar graphs show the mean and error bars represent standard error of mean. n=3, biological replicates. p-values are calculated by one sample t-test for mu=1. * denotes $p<0.05$, exact p-values from left to right are: 0.0488, 0.0204, 0.1550. (d) Fold change in the number of TRA-1-60-positive colonies generated from HDF-A cells with episomal delivery of reprogramming factors. Representative well images are above the bars. n=3, three biological replicates. p-values are calculated by one sample t-test for mu=1. Exact p-values from left to right are: 0.0784, 0.1155. These experiments were conducted with Gülben Gürhan Sevinç.

4.5. BRD9 inhibition and degradation enable iPSC generation without KLF4 and c-MYC.

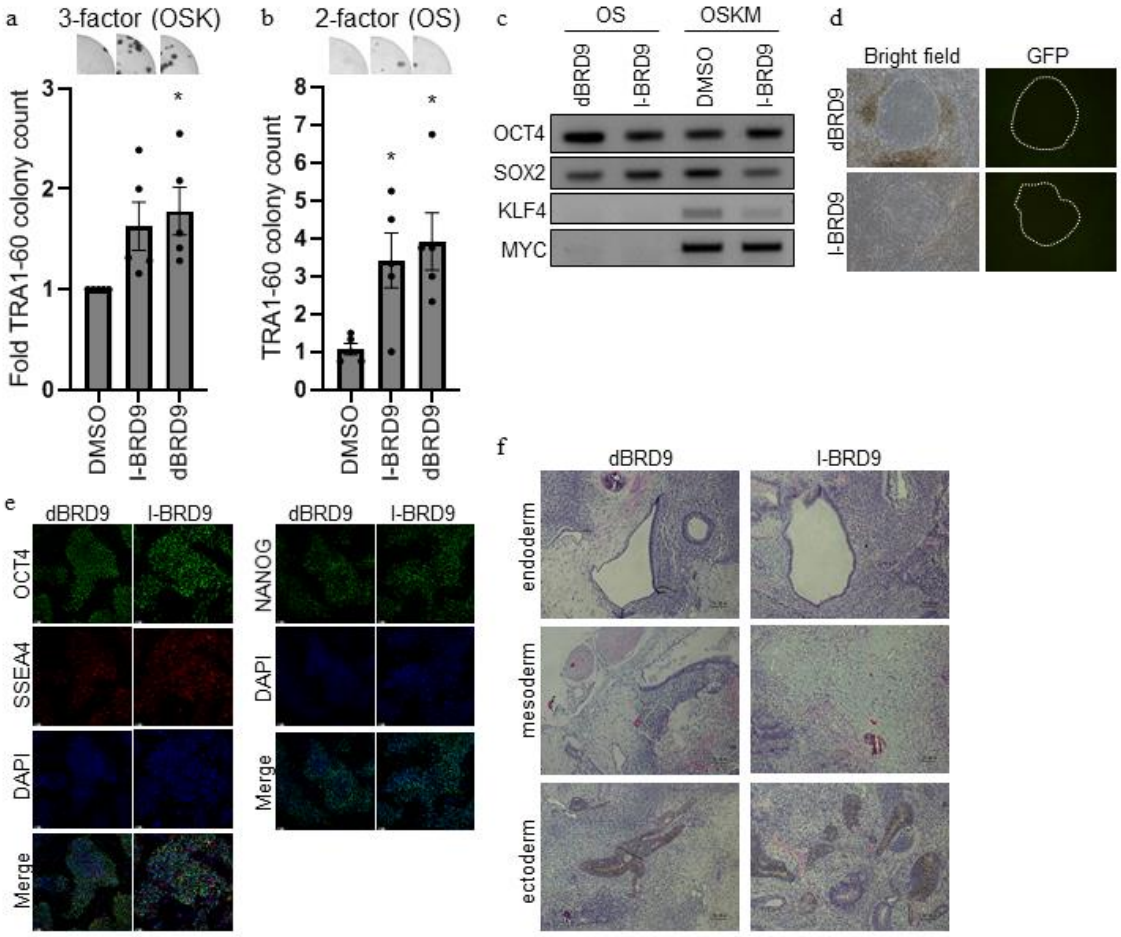


Figure 4.7 BRD9 inhibition and degradation enable iPSC generation without KLF4 and c-MYC. (a) Fold change in the number of TRA-1-60-positive colonies generated by OCT4, SOX2 and KLF4 (OSK) in the presence of DMSO, I-BRD9 or dBRD9. Representative well images are above the graph. Bar graphs show the mean and error bars represent standard error of mean. n=5, five biological replicates. p-values are calculated by one sample t-test for mu=1. Exact p-values from left to right are: 0.0583, 0.0297. (b) Number of TRA-1-60 positive colonies generated by OCT4 and SOX2 (OS) in the presence of DMSO, I-BRD9 or dBRD9. Representative well images are above the graph. n=5, biological replicates with at least 3 technical replicates. p-values are calculated by two sample t-test. * denotes $p<0.05$, exact p-values from left to right are: 0.0302, 0.0182. (c) Agarose gel images of PCR products for exogenous OCT4, SOX2, KLF4 and c-MYC transgenes in iPSCs generated from fibroblasts with indicated induction and treatments. (d) Phase contrast and GFP fluorescence images of colonies derived by OS induction and dBRD9 (upper) or I-BRD9 (lower) treatments showing typical iPSC morphology and

silencing of retroviral GFP transgene. (e) OCT4, SSEA4 (left) and NANOG (right) immunofluorescence of iPSCs derived from OS expressing fibroblasts treated with indicated treatments. Hoechst 33324 was used to stain the nuclei. (f) Hematoxylin and eosin stained sections of teratomas of iPSCs derived from OS-expressing fibroblasts treated with dBRD9 or I-BRD9 show tissues from endoderm, ectoderm and mesoderm lineages. These experiments were conducted with Gülben Gürhan Sevinç, immunofluorescence experiments were conducted by Simge Kelekçi and teratoma sections were stained by Prof. Arzu Ruacan.

To investigate if chemical perturbation of BRD9 can improve reprogramming by fewer exogenous factors. I reprogrammed fibroblasts without MYC in the presence of I-BRD9 and dBRD9. Both small molecules increased 3-factor (OSK) reprogramming efficiency which enables the exclusion of an oncogene, MYC, from reprogramming cocktail (Figure 4.7.a). To further decrease the number of transgenes necessary to reprogram fibroblast to iPSCs, I excluded KLF4 from the cocktail. In control DMSO treated wells, only one iPSC colony emerged whereas this number could be increased up to four in the presence of small molecules targeting BRD9 (Figure 4.7.b). Colonies generated from OCT4 and SOX2 overexpression and small molecule treatments were expanded for further pluripotency characterization experiments. Genomic DNA PCR with transgene-specific set of primers indicated that the OS-derived iPSCs did not contain KLF4 and MYC integrations (Figure 4.7.c). These iPSC colonies exhibited canonical features of pluripotent stem cells such as retroviral gene silencing, OCT4, SSEA4 and NANOG expression and contribution to tissues from three germ layers *in vivo* by teratoma formation assay (Figure 4.7.d, Figure 4.7.e and Figure 4.7.f). These results showed that BRD9 inhibition can enable iPSC generation with only 2-factors.

4.6.BRD9 is dispensable for human pluripotency induction and maintenance.

A previous study (Gatchalian et al., 2018) reported that BRD9 bromodomain safeguards mouse naïve pluripotency character by maintaining the expression of regulators of naïve pluripotency such as *Nanog*, *Klf4*, *Prdm14*, *Tet2*, *Fgf4* and *Tcl1*. As the short term inhibition or degradation of BRD9 during reprogramming did not interfere with the pluripotency character of resulting human iPSCs, I asked whether BRD9 expression is necessary for human pluripotency induction and maintenance. To address this question, I picked and expanded individual iPSC colonies generated by fibroblasts expressing sgRNAs targeting BRD9. Immunoblotting showed that majority of these

iPSCs (9 out of 14) did not express BRD9 protein and a minority (4 out of 14) expressed reduced amount of BRD9 (Figure 4.8.a). Four BRD9 knockout iPSC lines were characterized further for their pluripotency character and they did not show any difference in pluripotency marker expressions compared to control iPSCs (Figure 4.8.b). Similarly, treatment of OCT4-GFP reporter iPSCs with I-BRD9 and dBRD9 for 48 hours did not decrease cell proliferation levels and OCT4 levels compared to DMSO control (Figure 4.8.c, Figure 4.8.d and Figure 4.8.e). These results suggest that, in disagreement with previous findings regarding its role in mouse naïve pluripotency, BRD9 does not play a critical role in safeguarding human pluripotency.

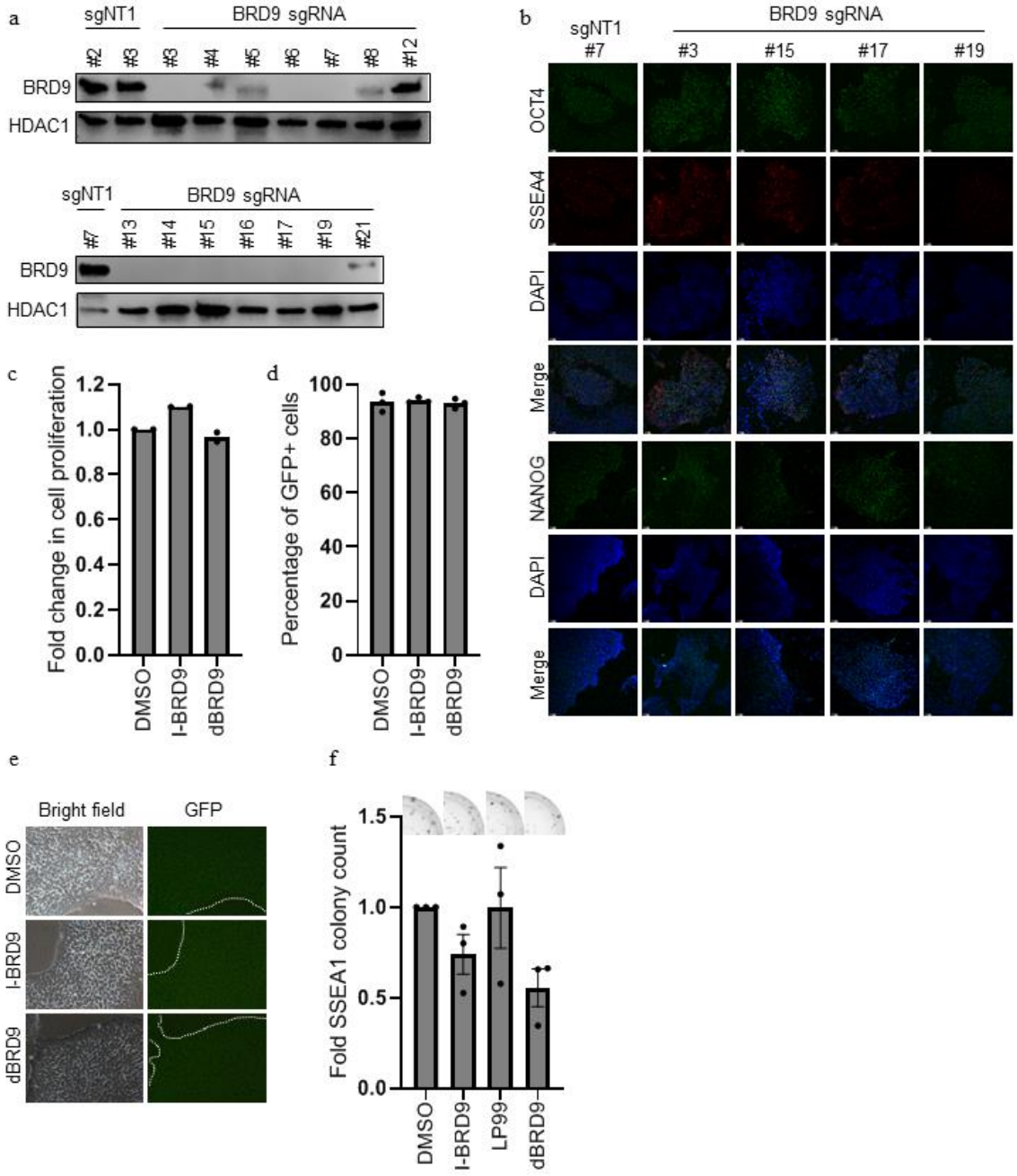


Figure 4.8 BRD9 is dispensable for human pluripotency induction and maintenance.

(a) Western blot for BRD9 from iPSC clones generated from either sgNT1 or BRD9 sg823-expressing fibroblasts. HDAC1 serves as loading control. (b) Immunofluorescence for OCT4, SSEA4 (top) and NANOG (bottom) in iPSCs generated from control sgNT1 and BRD9 sg823-expressing cells. Hoechst 33324 was used to stain the nuclei. Scale bars denote 100 μm. (c) Cell proliferation analysis upon 2 day treatment of OCT4-GFP reporter iPSCs with indicated small molecules. n=2, two biological replicates with technical triplicates. (d) Percentage of GFP-positive cells in OCT4-GFP reporter iPSCs upon indicated treatments. n=3, biological replicates. p-values are calculated by unpaired two-sided two sample t-test. Exact p-values from left to right are: 0.7636 and 0.8739. (e) Phase contrast and GFP fluorescence images of OCT4-GFP reporter iPSCs after 2 day treatment. (f) Fold change in the number of SSEA1-positive colonies generated by MEFs upon OSKM induction with indicated treatments. Representative well images are above the graph. Bar graphs show the mean and error bars represent standard error of mean. n=3, three biological replicates. p-values are calculated by one sample t-test for mu=1. Exact p-values from left to right are: 0.1415, 0.9861 and 0.05090. These experiments were conducted with Gülben Gürhan Sevinç and immunofluorescence experiments were conducted by Simge Kelekçi.

This observation led us to hypothesize that BRD9 has a species-specific role in regulating cell identity. To address this, I reprogrammed mouse embryonic fibroblasts (MEF) in the presence of BRD9 degrader and bromodomain inhibitors. BRD9 perturbations did not increase murine reprogramming to pluripotency (Figure 4.8.f). These results suggest that BRD9 has a species-specific role in reprogramming context and it is not essential for human pluripotency maintenance unlike its role in murine naïve pluripotency.

4.7.BRD9 inhibition does not interfere with pluripotency exit but decreases the mesendoderm differentiation capacity.

Next, I asked whether BRD9 expression is necessary for human pluripotent stem cells to differentiate into three germ layers which is an essential feature for this cell type. To answer this question, BRD9 knockout iPSCs were injected into immuno-compromised mice and teratomas were removed from sacrificed animals. Teratoma sections were stained with hemotoxylin and eosin. Upon investigating these sections, I observed that tissues from mesoderm lineage were restricted in BRD9 knockout iPSCs and in clone #19, no tissue from mesoderm lineage was detected (Figure 4.9.a). These teratoma assay

results suggested that BRD9 has a role in early cell fate decisions of human pluripotent stem cells towards mesoderm lineage specification.

To address if BRD9 has a role in mesoderm lineage specification *in vitro*, several mesoderm inducing methods from pluripotent stem cells were compared. Although embryoid body formation assay does not differentiate cells into a specific lineage, WNT3A-ActivinA or WNT-agonist CHIR99021 treatments were shown to be specific to induce mesendodermal differentiation from human pluripotent stem cells (Lam et al., 2014). I observed by RT-qPCR for a panel of pluripotency and mesendoderm lineage-specific marker expressions such as BRACHURY and EOMES that 2 day exposure of iPSCs to CHIR99021 was superior to other methods to induce mesoderm lineage gene expression (Figure 4.9.b).

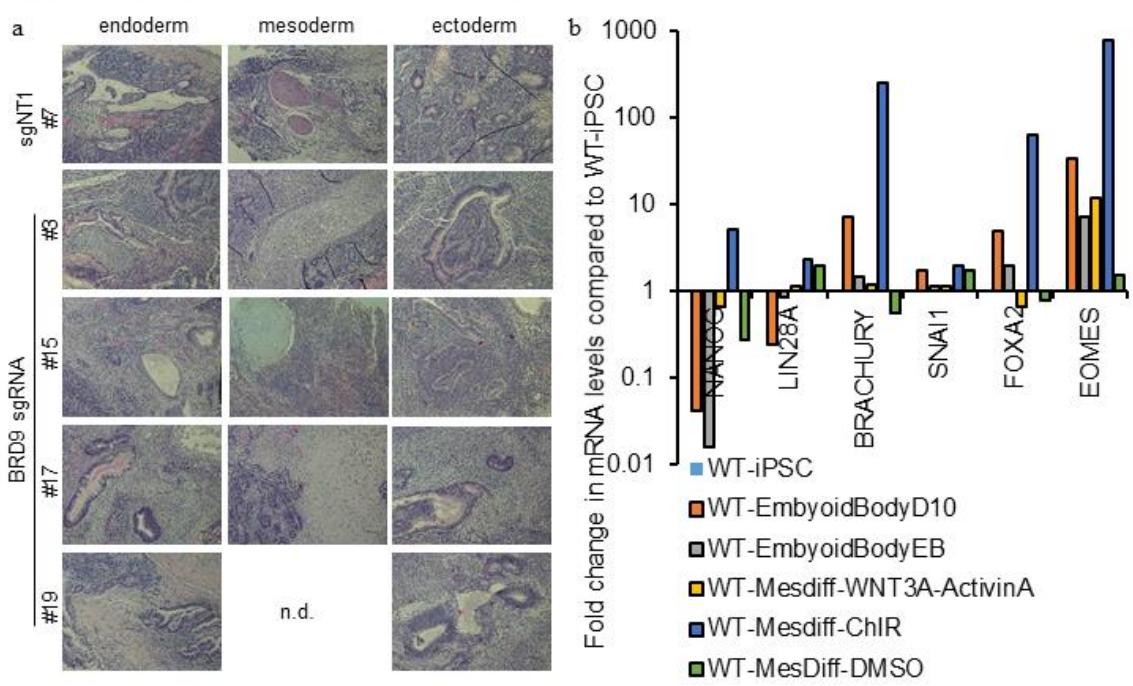


Figure 4.9 BRD9 knockout limits mesoderm differentiation capacity of human pluripotent stem cells *in vivo*. (a) Hematoxylin and eosin-stained sections of teratomas of iPSCs derived from sgNT1- or BRD9 sg823-expressing fibroblasts show tissues from endoderm, ectoderm and mesoderm lineages. n.d.: not detected. (b) Comparison of differentiation protocols by quantifying change in pluripotency and mesendodermal marker gene expressions compared to iPSC. WT: sgNT1 #7, EmbryoidBodyD10: embryoid body formation protocol with D10 media, EmbryoidBodyEB: embryoid body formation protocol with hES media without FGF. MesDiff: when 50% confluent, iPSCs are changed to D10 media supplemented with indicated treatments. Teratoma sections were stained by Prof. Arzu Ruacan.

To understand if BRD9 expression is necessary to differentiate pluripotent stem cells into mesoderm, BRD9 knockout iPSCs were treated with CHIR99021 for 2 days. BRD9 knockout iPSC lines #15 and #19 expressed mesendoderm markers such as BRACHURY and EOMES less compared to control iPSCs (Figure 4.10.a). Pluripotency genes were as efficiently repressed in BRD9 knockout cell as control iPSCs during the course of differentiation. These preliminary results indicate BRD9 is not necessary pluripotency exit.

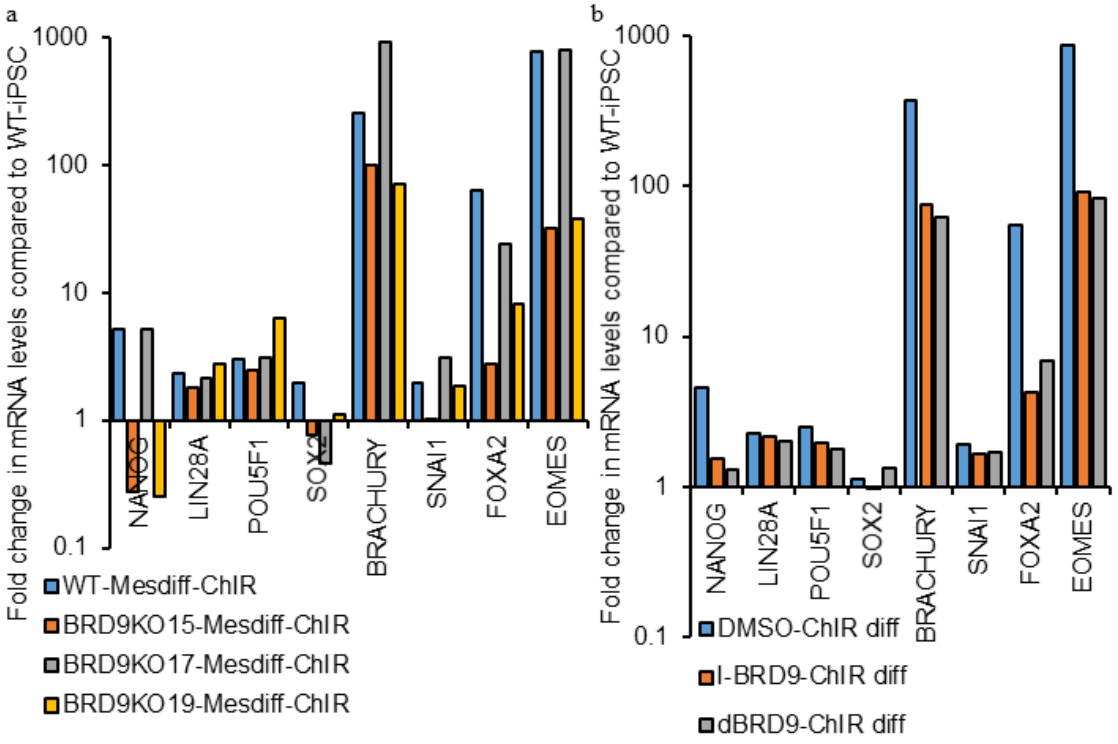


Figure 4.10 BRD9 is required for mesendoderm differentiation. (a) Indicated gene expressions of CHIR-induced mesendoderm differentiating BRD9 knockout iPSCs compared to undifferentiated sgNT1 #7 iPSC. (b) Indicated gene expressions of CHIR-induced mesendoderm differentiating sgNT1 #7 iPSCs treated with BRD9 inhibitors compared to undifferentiated sgNT1 #7 iPSC. These experiments were conducted with Gülben Gürhan Sevinç.

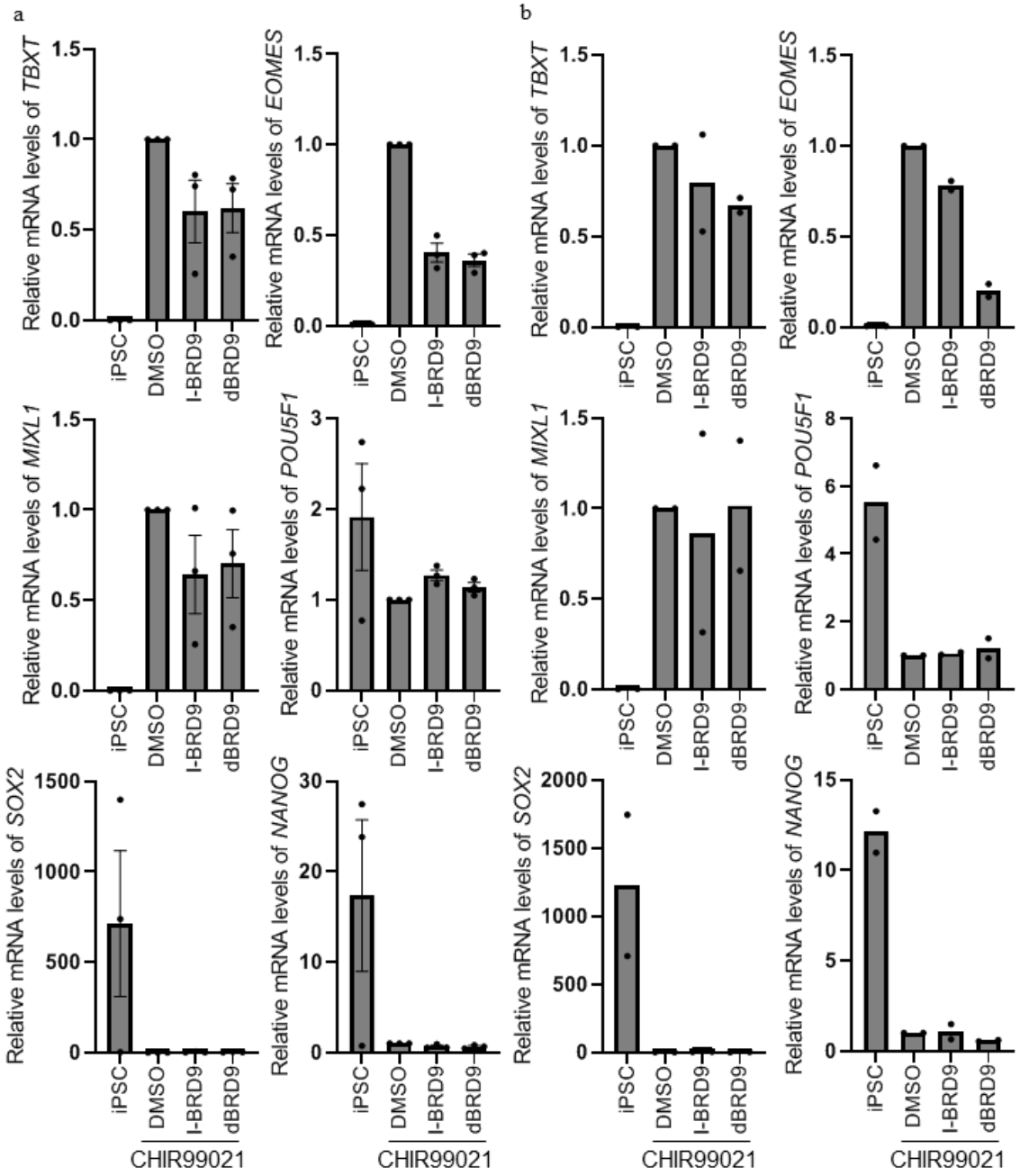


Figure 4.11 BRD9 inhibition does not interfere with pluripotency exit but decreases the mesendoderm differentiation capacity. (a) Fold change in expression levels of indicated genes upon compound treatments compared to DMSO control for a wild type iPSC line generated from adult fibroblasts by episomal reprogramming. (b) Fold change in expression levels of indicated genes upon compound treatments compared to DMSO control for an additional wild type iPSC line generated from adult fibroblasts by episomal reprogramming. These experiments were conducted with Gülben Gürhan Sevinç.

Although two out of three BRD9 knockout iPSC lines expressed reduced mesendoderm markers, BRD9 knockout iPSC line #17 expressed the similar levels of

such genes compared to control iPSCs. To understand if this is a clonal effect, I treated control iPSCs with I-BRD9 or dBRD9 during CHIR99021-induced mesendoderm differentiation. I observed that small molecules targeting BRD9 decreases the expression levels of mesendoderm lineage-specific markers such as BRACHURY and EOMES compared to DMSO control (Figure 4.10.b). This chemical perturbation based method complements the genetic data and shows that BRD9 is necessary for mesoderm-lineage specification.

To extend our observation on BRD9's role in mesoderm differentiation, I utilized two fully characterized iPSC lines generated in our laboratory with non-integrating episomal vectors. iPSCs used in Figure 4.9 and 4.10 were generated with lentiviral delivery of OSKM which may not be silenced efficiently to during differentiation and confound the results. Therefore transgene-free iPSCs derived by episomal plasmid delivery were utilized. I observed that small molecule treatments targeting BRD9 did not interfere with pluripotency exit of transgene-free iPSCs but downregulated mesendoderm lineage markers such as BRACHURY, EOMES and MIXL1 (Figure 4.11.a and Figure 4.11.b). These results confirmed the previous genetic and chemical-based results from iPSCs generated from lentiviral delivery of OSKM that BRD9 is a positive regulator of mesendodermal lineage specification.

4.8.Initial week of reprogramming is most sensitive to BRD9 inhibition.

To gain insight into the mechanism by which BRD9 perturbations increase reprogramming efficiency, I treated cells with I-BRD9 or dBRD9 at different time windows of reprogramming (Figure 4.12.a). I observed that treating cells with both I-BRD9 and dBRD9 only during the initial week of reprogramming significantly increase the number of iPSCs generated (Figure 4.12.b and Figure 4.12.c). To show if this effect stems from more reprogrammed cells at the initial week of reprogramming, I used flow-cytometry to quantify TRA-1-60 stained cells on day 6. Flow cytometry results showed that BRD9 bromodomain inhibition significantly increased the number of TRA-1-60 positive cells at this early time point of reprogramming (Figure 4.12.d and Figure 4.12.e). Since the initial phase of reprogramming is characterized by downregulation of the somatic-specific genes, these results suggest that BRD9 inhibition may increase reprogramming through regulation of fibroblast specific genes.

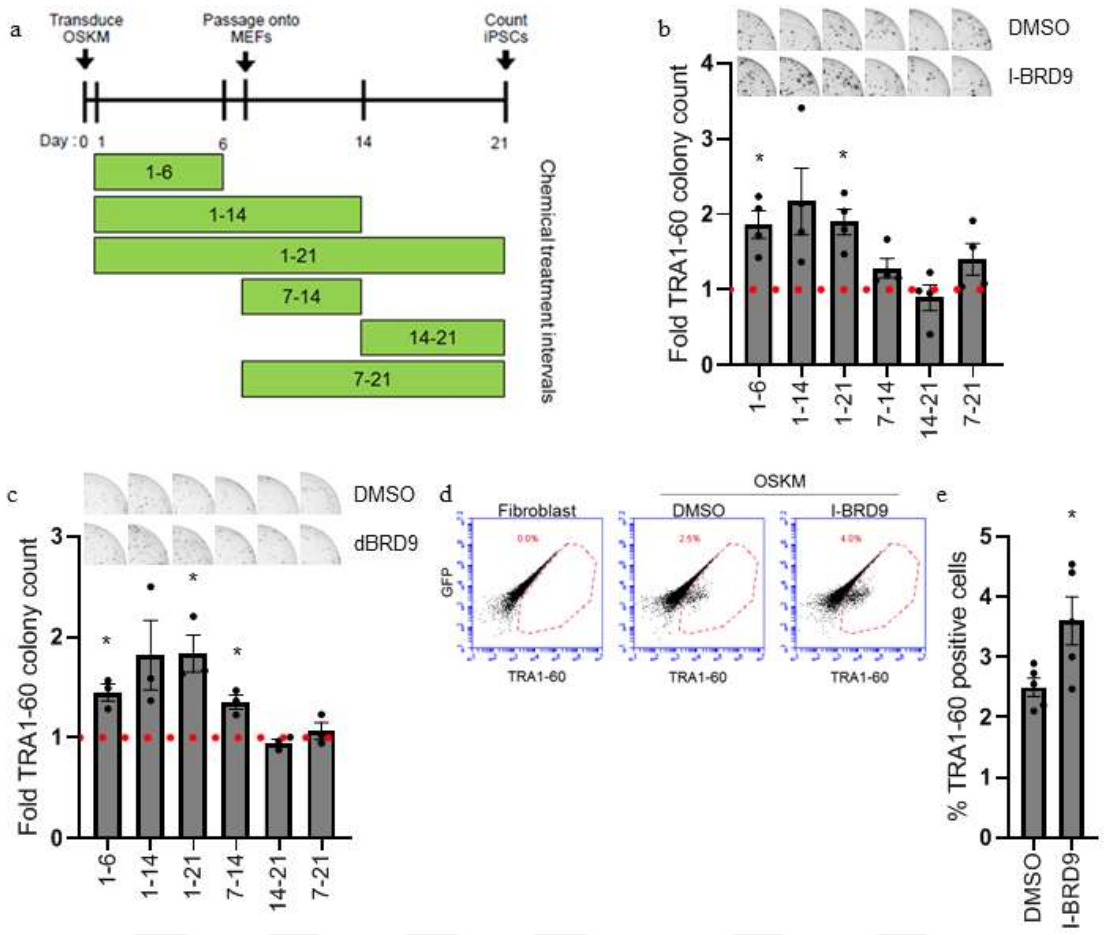


Figure 4.12 Initial week of reprogramming is most sensitive to BRD9 inhibition. (a) Schematic depicting the time intervals for treatment of compounds during reprogramming. (b) Fold change in the number of TRA-1-60-positive colonies upon I-BRD9 treatment compared to DMSO between indicated days of reprogramming. Representative well images are above the graph. Bar graphs show the mean and error bars represent standard error of mean. n=4, independent biological replicates. p-values were calculated by one sample t-test for mu=1. * denotes $p<0.05$, exact p-values from left to right are: 0.0178, 0.0771, 0.0139, 0.1116, 0.5826, 0.1492. (c) Fold change in the number of TRA-1-60-positive colonies upon dBRD9 treatment compared to DMSO between indicated days of reprogramming. Representative well images are above the graph. Bar graphs show the mean and error bars represent standard error of mean. n=3, independent biological replicates. p-values are calculated by one sample t-test for mu=1. * denotes $p<0.05$, exact p-values from left to right are: 0.0332, 0.1417, 0.0456, 0.0376, 0.3099, 0.5208. (d) Representative flow cytometry gates for indicated treatments. (e) Percentage of TRA-1-60-positive cells on day 6 of reprogramming with DMSO and I-BRD9 treatment. Bar graphs show the mean and error bars represent standard error of mean. n=5, five biological replicates. p-values are calculated by two sample t-test. * denotes

p<0.05, exact p-value: 0.0474. These experiments were conducted with Gülben Gürhan Sevinç.

4.9.Chemical perturbations of BRD9 have common transcriptional targets with genetic suppression.

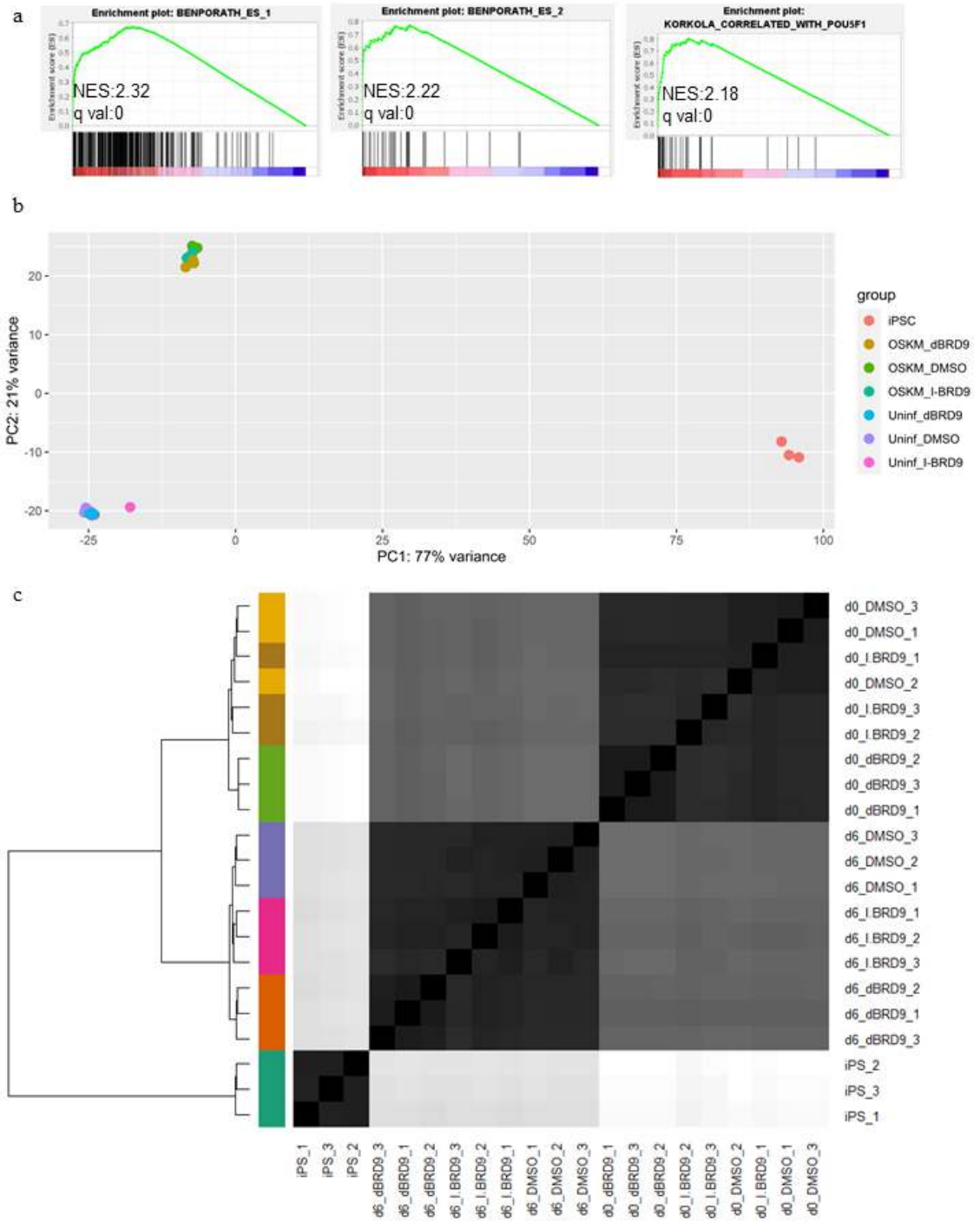


Figure 4.13 Replicates of RNA-Seq samples cluster together. (a) Gene set enrichment analysis (GSEA) of transcriptome data comparing iPSCs to fibroblasts for the gene sets

related to pluripotency. NES: normalized enrichment score, q val: False discovery rate (FDR) q-value. (b) Principle component analysis of transcriptome data with indicated cells. (c) Sample distance matrix and hierarchical clustering of indicated cells. RNA libraries were generated and sequenced by Dr. Martin Philpott.

To address the hypothesis that BRD9 maintains the expression of somatic cell-specific gene expression and its inhibition, thus, enhances reprogramming, I performed RNA-Seq for fibroblasts and reprogramming cells (at day 6) that were treated with compounds targeting BRD9 for 5 days or sgRNA targeting BRD9 along with relative controls and iPSCs. First, I compared the gene expression profiles of iPSCs and fibroblasts as controls. When I performed gene set enrichment analysis for this comparison, I observed that gene sets related to pluripotency were positively enriched in iPSCs relative to fibroblasts (Figure 4.13.a). Next, to check if relatively similar cells could cluster together, I performed principle component anaylsis (PCA). PCA results showed that fibroblasts, reprogramming cells and iPSCs clustered apart from each other but biological replicates clustered closer (Figure 4.13.b). Using a distance map, I observed that biological replicates clustered together (Figure 4.13.c). These initial quality control indicated that experimental and analytic parts of RNA-Seq were performed properly.

To understand the effect of different BRD9 perturbations on the transcriptome, I examined the number of differentially expressed genes for each conditions relative to respective controls. I observed that knocking out BRD9 have a greater effect on the transcriptome than all other treatments, and BRD9 degradation has more effect than BRD9 bromodomain inhibition (Figure 4.14.a). To address whether BRD9 KO affects the same set of common genes affected by BRD9 inhibitors, I first generated a gene list of BRD9 targets based on BRD9 knockout data. I called differentially upregulated genes BRD9 KO as BRD9-negative targets and differentially downregulated genes by BRD9 KO as BRD9-positive targets. Upon treatment of fibroblasts with small molecules targeting BRD9 I observed a positive enrichment for BRD9-negative targets wit (Figure 4.14.b) and a negative enrichment for BRD9-positive targets (Figure 4.14.c). To reveal common set of genes regulated by genetic suppression and chemical inhibition of BRD9, I generated venn-diagrams showing the number of overlapping differentially regulated genes between these two approaches. I observed that majority of genes (61 out of 92) regulated commonly by chemical inhibition of BRD9 were, also, significantly regulated by genetic suppression of BRD9 (Figure 4.14.d and Figure 4.14.e). To reveal the impact

of various approaches to inhibit BRD9 on human fibroblast transcriptome, I generated heatmaps of all differentially regulated genes by at least one BRD9 perturbation for all samples. Beyond showing highly reproducibility between biological replicates, this analysis showed that BRD9 knockout has wider effect on the transcriptome than chemical inhibitions (Figure 4.15). Additionally, BRD9 degradation has more impact on transcriptome than bromodomain inhibition. These results show that BRD9 perturbations affect the similar set of genes to varying degrees which may explain their different effect size on reprogramming.

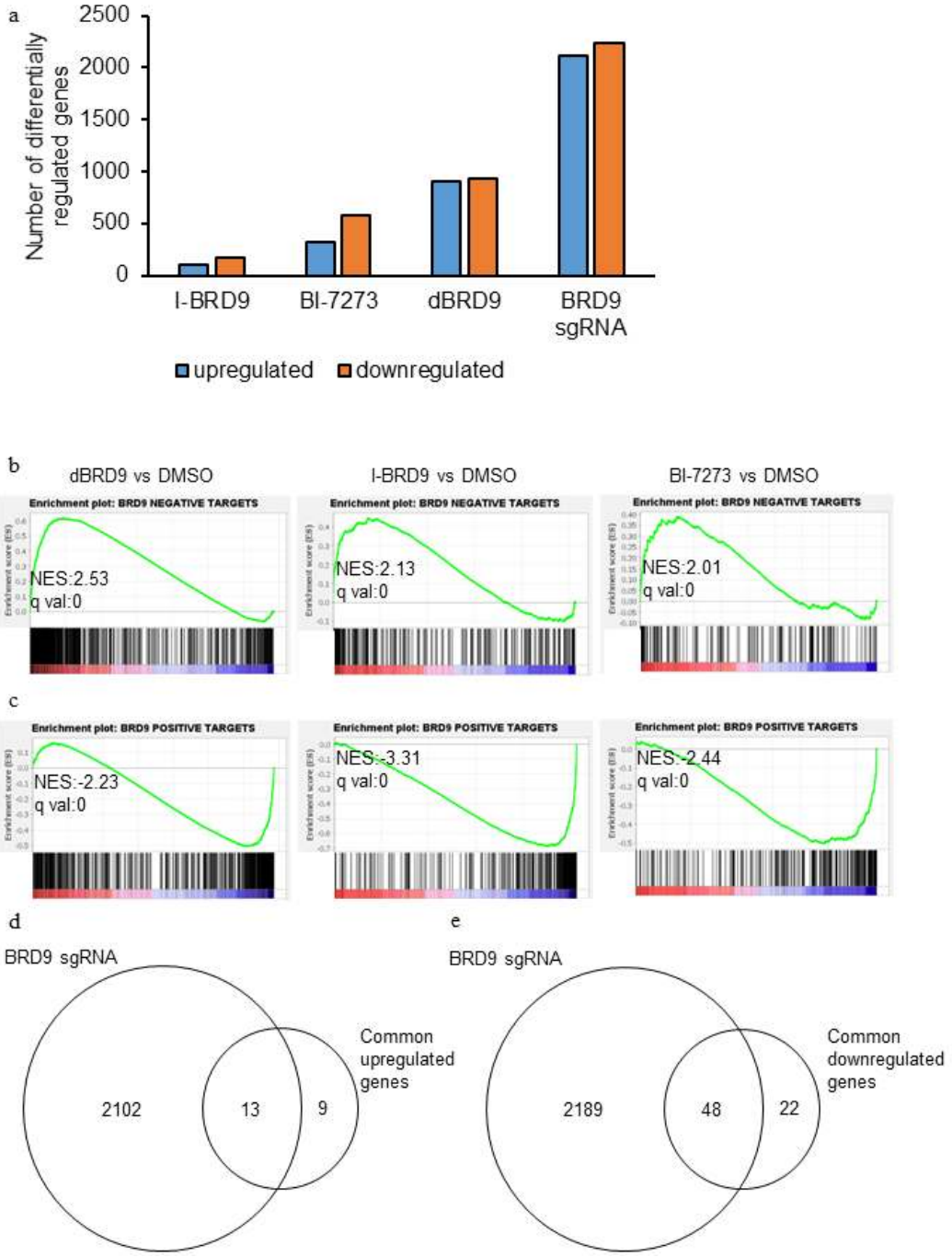


Figure 4.14 Chemical perturbations of BRD9 have common transcriptional targets with genetic suppression. (a) Number of differentially regulated genes by indicated compound treatments compared to DMSO control and by BRD9 sgRNA expression compared to control sgRNA expression. (b) Gene set enrichment analysis (GSEA) of transcriptome data with indicated treatments for BRD9 negative targets gene set generated by differentially upregulated genes upon BRD9 sgRNA expression. NES: normalized enrichment score, q val: False discovery rate (FDR) q-value. (c) Gene set enrichment analysis (GSEA) of transcriptome data with indicated treatments for BRD9 positive targets gene set generated by differentially downregulated genes upon BRD9 sgRNA expression. NES: normalized enrichment score, q val: False discovery rate (FDR) q-value. (d) Venn-diagram showing the number of differentially upregulated genes upon BRD9 knockout and common upregulated genes by dBRD9, BI-7273 and I-BRD9 treatments. (e) Venn-diagram showing the number of differentially downregulated genes upon BRD9 knockout and common downregulated genes by dBRD9, BI-7273 and I-BRD9 treatments.

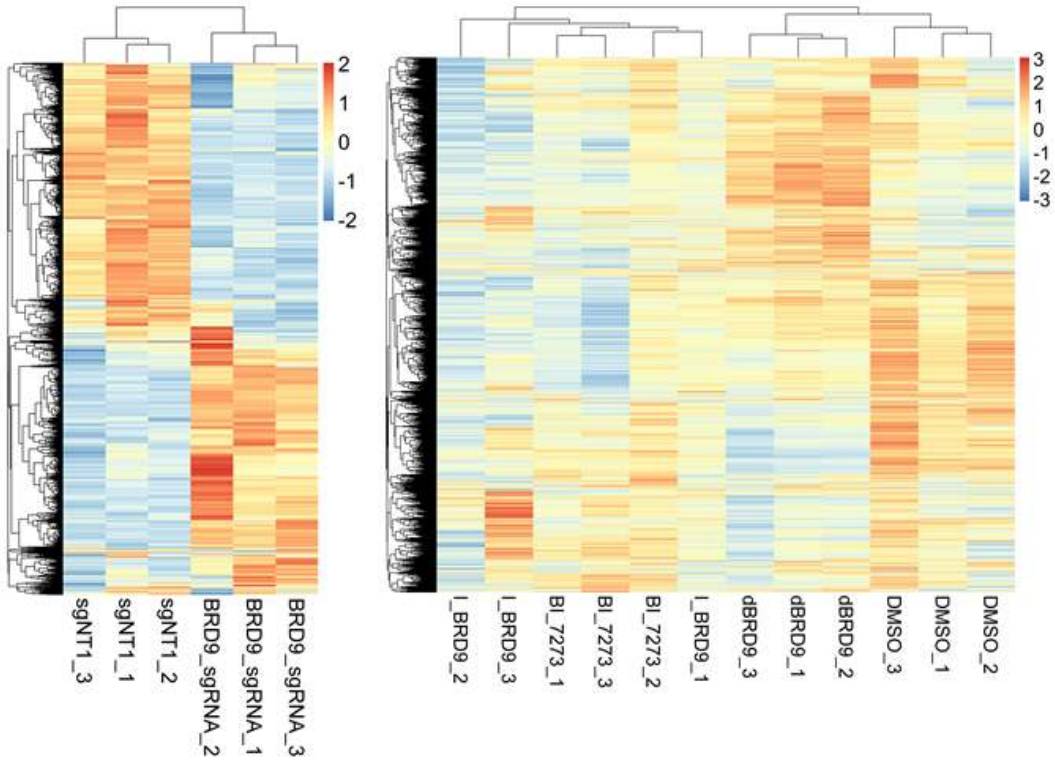


Figure 4.15 Genetic suppression of BRD9 has wider and more impact on transcriptome than chemical perturbation of BRD9. Heatmap shows relative

expression levels of differentially regulated genes in fibroblasts by at least one BRD9 perturbation across all samples.

4.10. BRD9 inhibition does not increase reprogramming efficiency due to increasing OSKM factor expression or suppressing cell cycle checkpoints.

To answer if the effect of BRD9 on reprogramming results from more efficient expression of OSKM factors, I looked at the expression levels of OSKM across BRD9 inhibitions at 6th day of reprogramming. I observed that chemical inhibitions of BRD9 do not increase OSKM expression levels, however, BRD9 knockout non-significantly upregulates them (Figure 4.16). This observation eliminates the possibility of more efficient OSKM expression upon BRD9 inhibition and suggests different mechanism(s) for BRD9 to act as a barrier to reprogramming.

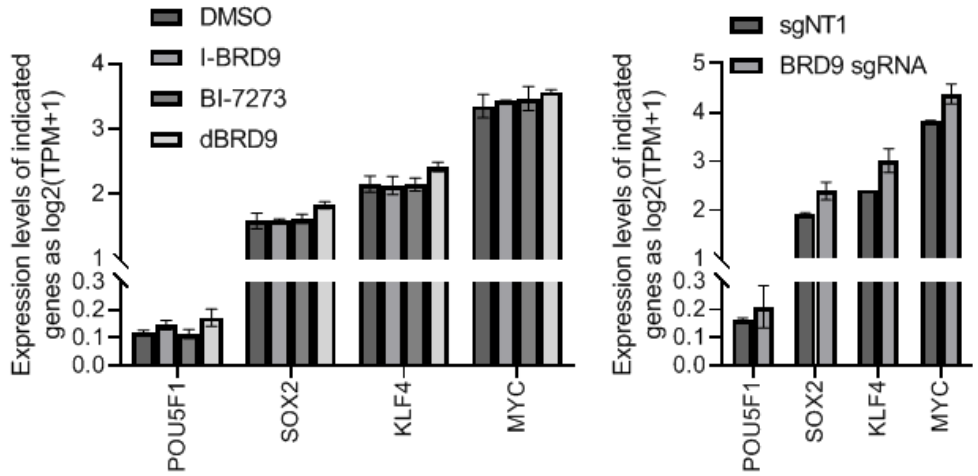


Figure 4.16 BRD9 inhibition does not increase expression levels of reprogramming factors. Average expression levels of OSKM genes as log2(TPM+1) across fibroblasts induced by OSKM for indicated treatments.

As the cell cycle checkpoints are barriers to somatic cell reprogramming, I hypothesized that BRD9 inhibition enhances reprogramming by downregulating such genes. To test this, I checked the change in expression levels of genes coding for cell cycle checkpoints. CDKN2B was commonly downregulated by all BRD9 perturbations whereas CDKN1A was downregulated by only knocking out BRD9 (Figure 4.17.a). To test the notion that inhibiting cell cycle checkpoints increases reprogramming efficiency in our fibroblast model, I used CRISPR-Cas9 mediated knockout of CDKN1A and CDKN2A. Indeed, knocking out either of them increased reprogramming efficiency

while CDKN2A knockout seems to be much more potent (Figure 4.17.b). These results show that cell cycle checkpoints are barriers to reprogramming and BRD9 inhibition downregulates one of them.

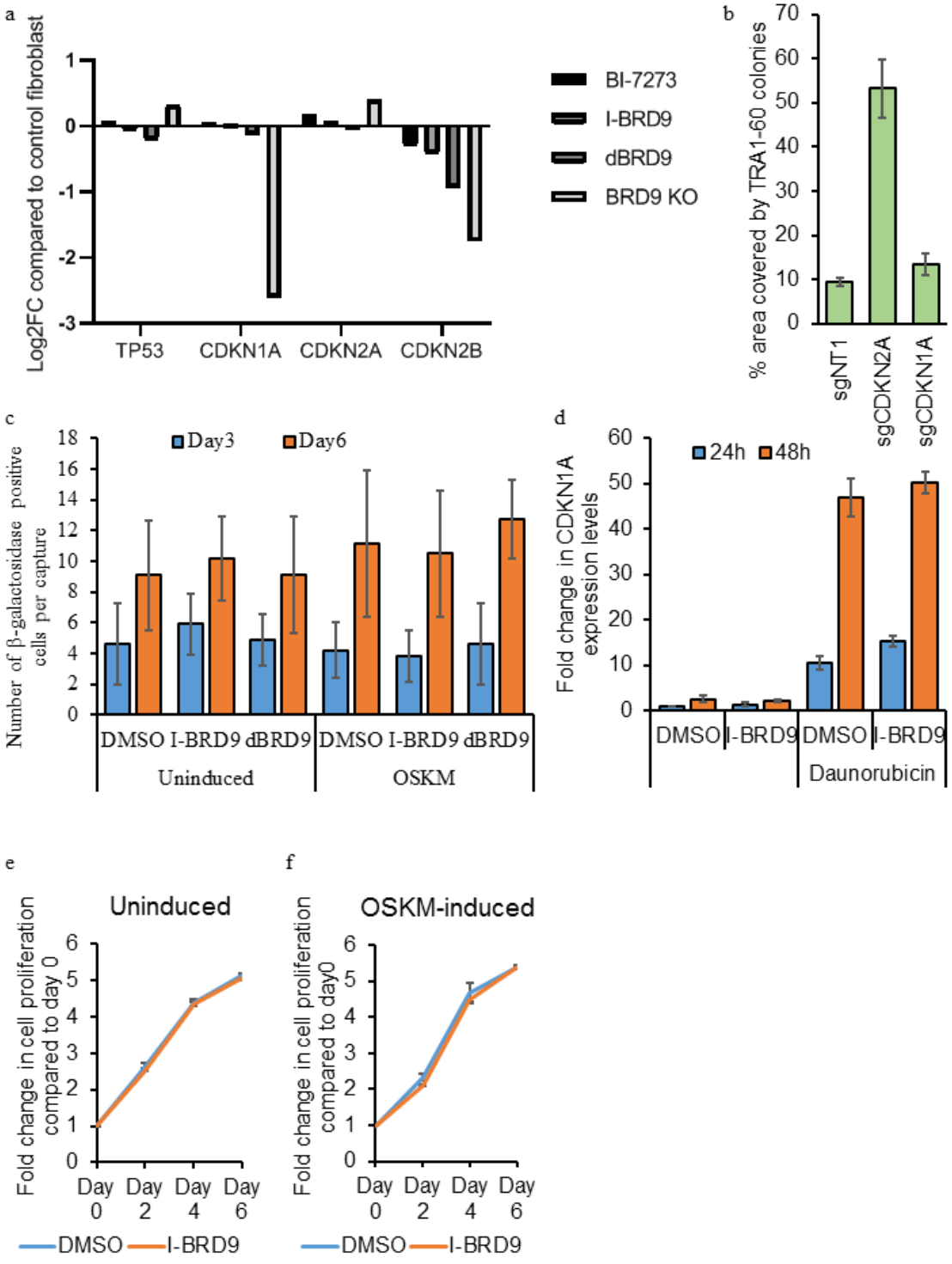


Figure 4.17 BRD9 inhibition does not increase reprogramming efficiency due to suppressing cell cycle checkpoints. (a) Log2 fold change in expression levels of indicated cell cycle checkpoints upon BRD9 inhibition compared to control cells. (b) Percentage of area covered by TRA-1-60 positive colonies upon CDKN1A and CDKN2A

sgRNA expression. (c) Number of β-gal positive cells per capture for indicated compound treatments. (d) Fold change in p21 expression levels for indicated compound treatments compared to DMSO control. (e) Fold change in cell proliferation rates of fibroblasts treated with DMSO or I-BRD9. (f) Fold change in cell proliferation rates of reprogramming fibroblasts treated with DMSO or I-BRD9.

To address the functional link between BRD9 inhibition, cell cycle checkpoints and increased reprogramming efficiency, I, first, asked if fibroblasts become less susceptible to senescence induction upon BRD9 inhibition. To answer this, I treated uninduced and OSKM-induced fibroblasts with DMSO, I-BRD9 and dBRD9 for 3 and 6 days and perform β-galactosidase assay. I observed that BRD9 inhibition does not have an effect on the number of senescent cells compared to controls (Figure 4.17.c). Similarly, when fibroblasts were treated with daunorubicin for 1 and 2 days, CDKN2A mRNA levels were upregulated to 15- and 50-fold, respectively (Figure 4.17.d). In this assay, BRD9 bromodomain inhibition did not delay or lessen the upregulation level of CDKN1A. To further eliminate the possibility that BRD9 inhibition causes fibroblasts to proliferate more than controls to result in more iPSC colonies, I measured the proliferation rates of uninduced and OSKM-induced fibroblasts treated with BRD9 inhibitors using an ATP-based cell viability assay. I observed that BRD9 bromodomain inhibition did not result in a proliferation advantage compared to controls (Figure 4.17.e and Figure 4.17.f). Together, these results show that BRD9 inhibition does not increase reprogramming efficiency through delaying senescence or providing proliferation advantage.

4.11. BRD9 maintains the expression of somatic-specific genes.

To understand the common transcriptional targets of different small molecules inhibiting BRD9, I checked the number of upregulated genes by all treatments. A small number of genes overlapped and these 22 common upregulated genes were not highly enriched in any specific gene set (Figure 4.18.a). On the other hand, 70 genes were commonly downregulated by BRD9 inhibition and these genes were highly enriched in EMT and extracellular matrix-related gene sets (Figure 4.18.b). To understand the significance of EMT-enrichment from gene ontology analysis, I performed GSEA on pre-ranked gene lists comparing I-BRD9 or dBRD9 to DMSO control for all gene sets available. EMT-related gene sets were among the most highly negatively enriched gene sets (Figure 4.18.c). These results show that inhibiting BRD9 in fibroblasts downregulates

EMT-related genes.

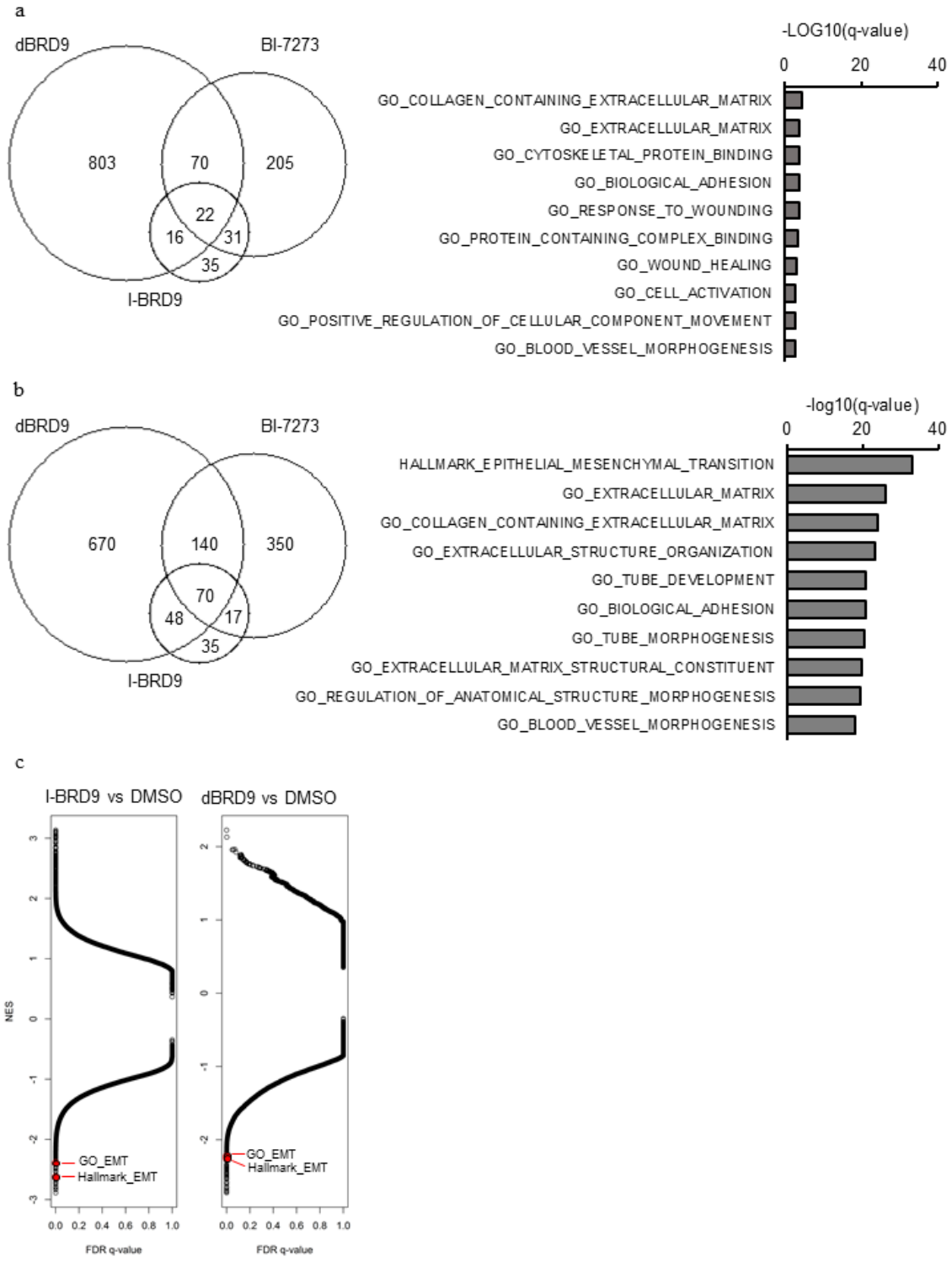


Figure 4.18 Genes related to epithelial-to-mesenchymal transition are downregulated by BRD9 perturbation. (a) Venn-diagram showing the number of differentially upregulated genes upon dBRD9, BI-7273 and I-BRD9 treatments (left) and top ten GO and Hallmark gene sets enriched in common upregulated genes upon dBRD9, BI-7273 and I-BRD9 treatments (right). Number of genes (n) in comparison were 22. p-

values were calculated by hypergeometric distribution. (b) Venn-diagram showing the number of differentially downregulated genes upon dBRD9, BI-7273 and I-BRD9 treatments (left) and top ten GO and Hallmark gene sets enriched in common downregulated genes upon dBRD9, BI-7273 and I-BRD9 treatments (right). Number of genes (n) in comparison were 70. p-values were calculated by hypergeometric distribution. (c) Gene set enrichment analysis on pre-ranked gene lists according to log2FC value for comparisons of fibroblasts treated with I-BRD9 and DMSO (left) and dBRD9 and DMSO (right) for all gene sets available at The Molecular Signatures Database (MSigDB). Red circles indicate Hallmark_EMT and GO_EMT gene sets.

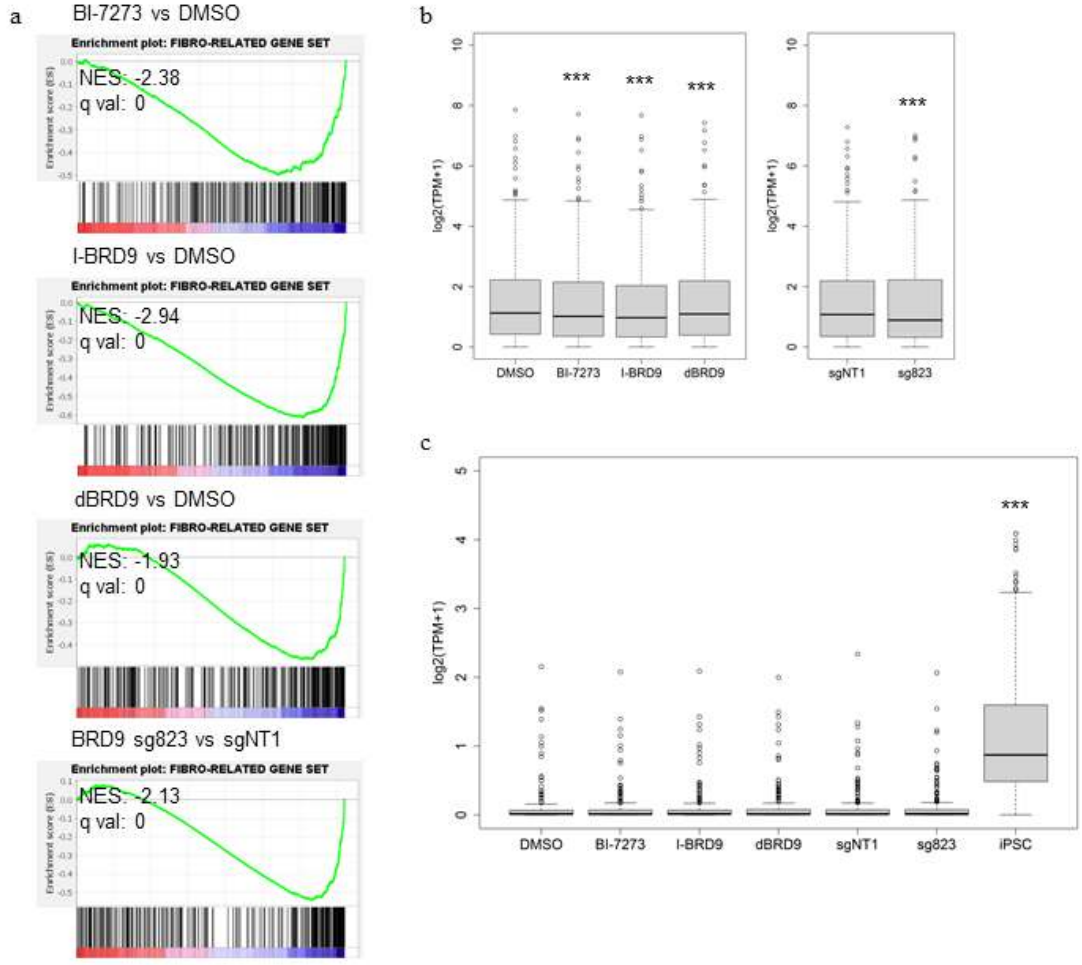


Figure 4.19 BRD9 maintains the expression of somatic-specific genes. (a) Gene set enrichment analysis (GSEA) of transcriptome data with indicated treatments for the fibroblast-related gene set. NES: normalized enrichment score, q val: False discovery rate (FDR) q-value. (b) Average expression levels of the 307 genes in fibroblast-related gene set across indicated fibroblasts. Whiskers indicate 95% confidence interval. p-values were calculated by Wilcoxon signed-rank test. *** denotes p<0.005, exact p-values from left to right are: 2.2e-16, 2.2e-16, 3.8e-10 and 5.057e-05. (c) Average expression levels of the

246 genes in pluripotency-related gene set across indicated cells. Whiskers indicate 95% confidence interval. p-values were calculated by one-sided Wilcoxon signed-rank test. *** denotes p<0.005, exact p-value for iPSC and DMSO comparison is 2.2e-16 and other comparisons are not significant.

To specifically address if BRD9 has a role in maintaining the expression of fibroblast-related genes, I performed GSEA on pre-ranked gene lists from comparisons of BRD9 perturbations to relative controls for the previously generated fibroblast-related gene set. I observed a strong and significant negative enrichment of this gene set by all BRD9 inhibitions (Figure 4.19.a). Further, the mean expression of fibroblast related genes decreased upon all BRD9 inhibitions (Figure 4.19.b). However, inhibiting BRD9 in fibroblasts did not activate pluripotency-related genes on its own (Figure 4.19.c). To specifically observe the effect size of BRD9 inhibitions on weakening fibroblast cell identity, I intersected the downregulated genes upon BRD9 inhibition with fibroblast-specific gene set. I observed that 131 and 117 out of 307 fibroblast-specific genes were, already, downregulated in fibroblasts by genetic and chemical inhibition of BRD9, respectively (Figure 4.20). These results support the notion that BRD9 maintains the fibroblast-specific gene expression and its inhibition erases somatic cell identity in a more efficient way during reprogramming.

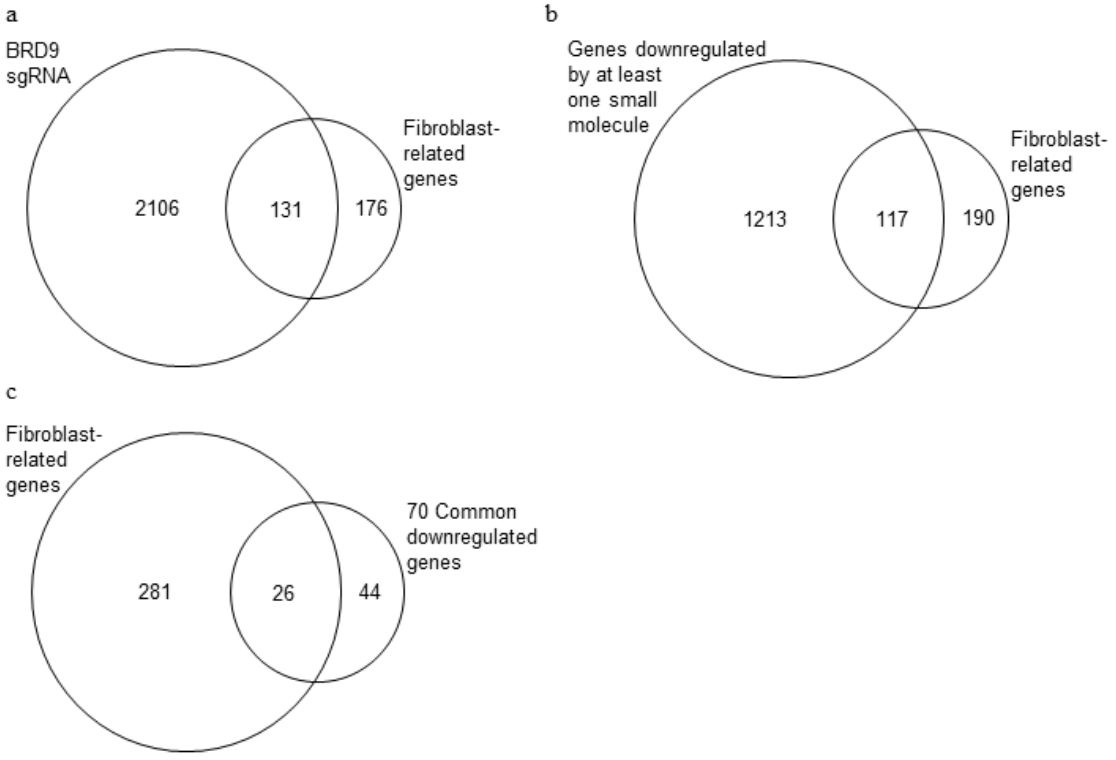


Figure 4.20 BRD9 inhibition downregulates genes related to fibroblast identity. (a) Venn-diagram showing the number of differentially downregulated genes upon BRD9 knockout and fibroblast-specific gene set. (b) Venn-diagram showing the number of differentially downregulated genes by at least one chemical perturbation of BRD9 and fibroblast-specific gene set. (c) Venn-diagram showing the number of common differentially downregulated genes by chemical perturbation of BRD9 and fibroblast-specific gene set.

4.12. BRD9 inhibition results in less accessible chromatin around putative enhancers.

To understand the mechanism of downregulation of genes related to somatic cell identity by BRD9 inhibition at the chromatin level, I performed an assay for transposase accessible chromatin followed by next generation sequencing (ATAC-Seq) for fibroblasts treated with I-BRD9, dBRD9 and DMSO control. Interestingly, when all active promoters marked by H3K4me3 and H3K27ac around transcription start sites (TSS) were taken into consideration, neither of BRD9 inhibitions resulted in less chromatin accessibility (Figure 4.21.a). However, when putative active enhancers marked by H3K4me1 and H3K27ac were examined, both I-BRD9 and dBRD9 decreased the accessible chromatin signal compared to DMSO control (Figure 4.21.b). Importantly, these putative fibroblast enhancers were also subject to decrease in chromatin accessibility upon OSKM induction (Figure 4.21.c). Overall, these results confirmed previous observations that BRD9 and ncBAF complex act on enhancers to regulate gene expression and these factors are important in maintaining fibroblast-specific gene expressions through enhancers.

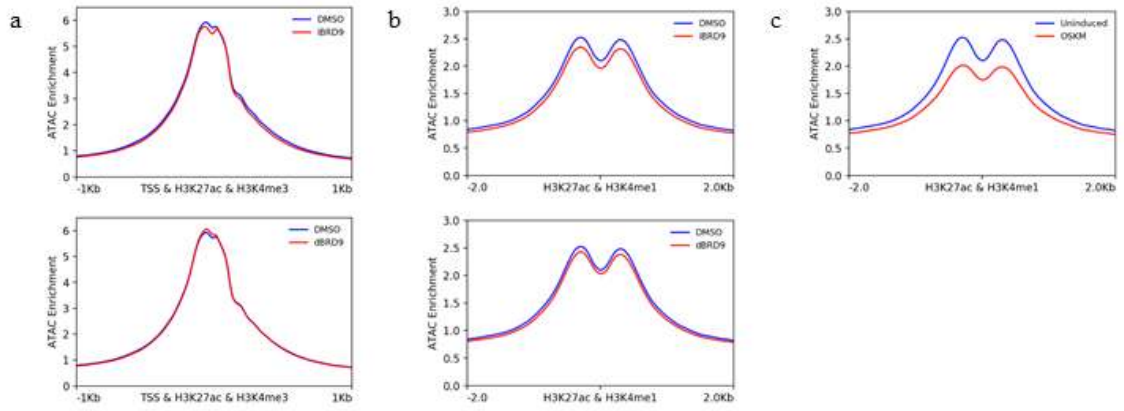


Figure 4.21 BRD9 inhibition results in less accessible chromatin around putative enhancers. (a) Aggregate ATAC-seq plots from fibroblasts treated with I-BRD9 (top), dBRD9 (bottom) and DMSO on the +-1kb of transcription start site, H3K27ac and

H3K4me3 summit. n=3, three biological replicates. (b) Aggregate ATAC-seq plots from fibroblasts treated with I-BRD9 (left), dBRD9 (right) and DMSO around +/-2kb of H3K27ac and H3K4me1 summit. n=3, three biological replicates. (c) Aggregate ATAC-seq plot from uninduced and OSKM-induced fibroblasts at day 6 of reprogramming around +/-2kb of H3K27ac and H3K4me1 summit. n=3, three biological replicates. DNA libraries were generated and sequenced by Dr. Martin Philpott and Dr. Derya Cavga analyzed ATAC-Seq data.

4.13. BRD9 and CREBBP/EP300 regulate the expression of common set of genes and their combined inhibition has no additive effect on reprogramming efficiency.

As I showed in Figure 4.5.a that DOT1L and BRD9 inhibitions have additive effects on reprogramming, I asked whether these chromatin factors regulate different set of genes or their combined effect has additive impact on regulating common set of genes. To answer this, I focused on the number of common differentially regulated genes by DOT1L and BRD9 inhibitions. I observed that only a few genes regulated commonly with these factors (Figure 4.22). This analysis suggest that BRD9 and DOT1L regulate mostly different set of genes and, thus, their combined inhibition has an additive impact on reprogramming phenotype.

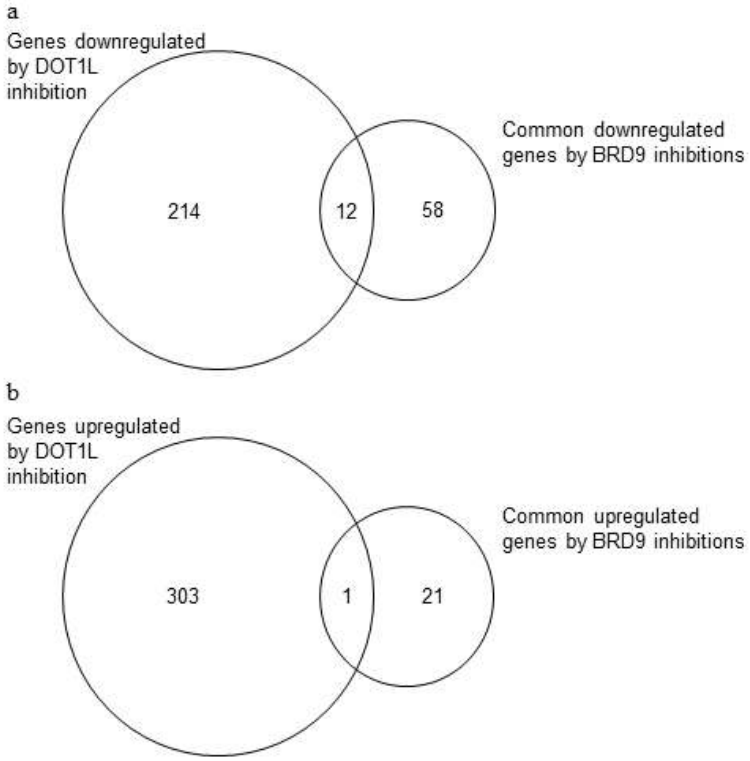


Figure 4.22 DOT1L and BRD9 do not have common targets in fibroblasts. (a) Venn-diagram showing the number of differentially downregulated genes upon DOT1L inhibition (GSE29253) and common differentially downregulated genes by chemical perturbation of BRD9. (b) Venn-diagram showing the number of differentially upregulated genes upon DOT1L inhibition (GSE29253) and common differentially upregulated genes by chemical perturbation of BRD9.

As there were common features such as acting early in reprogramming, maintaining somatic cell identity-related gene expression program through enhancers between BRD9 and CREBBP/EP300-acetyl lysine mediated interactions, I asked whether both factors affect the reprogramming process in a parallel but non-redundant pathways. To answer this, I combined CREBBP/EP300 bromodomain inhibitor, CBP30, with either I-BRD9 or dBRD9. I observed that combination of CBP30 with either of compounds targeting BRD9 did not result in further increase in reprogramming efficiency (Figure 4.23.a and Figure 4.23.b). To check if BRD9 inhibition affects the common set of genes that CREBBP/EP300 bromodomain inhibition does, I generated "CBP/EP300 Targets" gene set composed of genes downregulated upon CBP30 treatment in fibroblasts. When I performed GSEA with this gene set, I observed that CBP/EP300 Targets were significantly negatively enriched in all BRD9 perturbations (Figure 4.23.c). These results suggest that BRD9 and CREBBP/EP300 maintain the somatic cell identity-related transcriptome in parallel and non-redundant ways and their combined inhibitions have no additive effect on reprogramming.

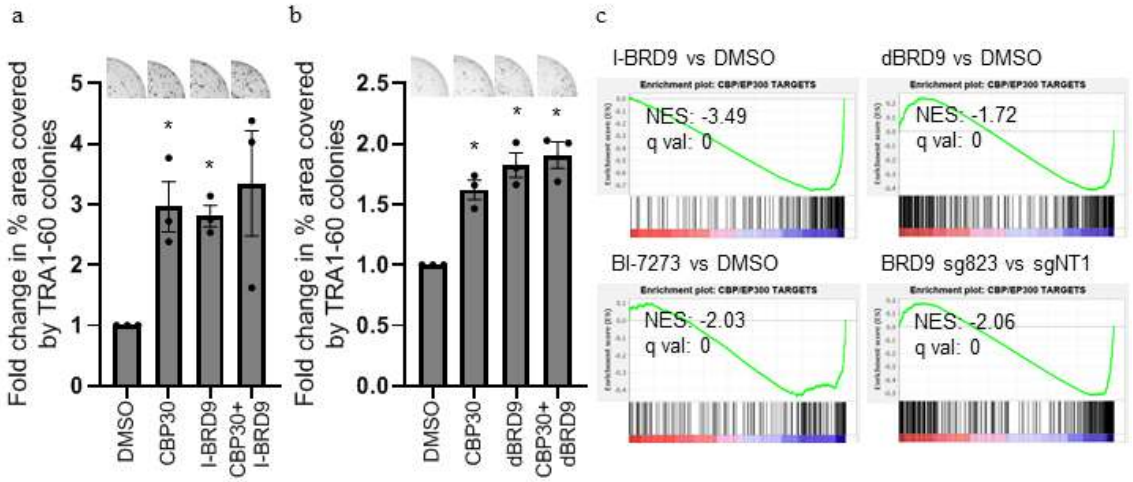


Figure 4.23 BRD9 and CREBBP/EP300 bromodomain have non-redundant and overlapping roles in reprogramming. (a) Fold change in the percent area covered by

TRA-1-60-positive colonies by indicated treatments compared to DMSO. Representative well images are above the graph. n=3, three independent biological replicates. p-values are calculated by one sample t-test for mu=1. * denotes p<0.05, exact p-values from left to right are: 0.0422, 0.0095, 0.1137. (b) Fold change in the percent area covered by TRA-1-60-positive colonies by indicated treatments compared to DMSO. Representative well images are above the graph. n=3, three independent biological replicates. p-values are calculated by one sample t-test for mu=1. * denotes p<0.05, exact p-values from left to right are: 0.0166, 0.0148, 0.0142. (c) Gene set enrichment analysis (GSEA) of transcriptome data with indicated treatments for the CBP/EP300 target gene set. NES: normalized enrichment score, q val: False discovery rate (FDR) q-value.

4.14. Transcriptional targets of BRD9, MN1 and ZBTB38 are barriers to reprogramming.

To gain more insight on how BRD9 inhibition increases reprogramming efficiency, I focused on transcriptional targets of BRD9 that may act as barriers to reprogramming. To this end, transcriptional regulators were inspected closely as they orchestrate the expression of many genes and subtle changes in transcription of such regulators may have a significant downstream effect. I observed that two such regulators, MN1 and ZBTB38 were downregulated upon all BRD9 perturbations (Figure 4.24.a). MN1 regulates palate development and can act as co-factor for various transcription factors such as retinoic acid receptor/retinoic X receptor (RAR/RXR) (Mak et al., 2020; Meester-Smoor et al., 2008). However, MN1's role in reprogramming has not been studied to date. Overexpression of MN1 in fibroblasts significantly decreased their reprogramming efficiency (Figure 4.24.b and Figure 4.24.c). ZBTB38 is a potential master regulator in fibroblasts and its expression is controlled by a fibroblast specific super-enhancer (HniszBrian et al., 2013). Overexpressing this transcription factor decreased the number of iPSCs generated even in the presence of small molecules targeting BRD9 and reprogramming cells at an early time point (Figure 4.24.b, Figure 4.24.d and Figure 4.24.e). Expectedly, ZBTB38 overexpression downregulated pluripotency-related gene expression such as NANOG, LEFTY2 and LIN28A during reprogramming (Figure 4.24.f). Overall, these results suggest that BRD9 inhibition increases reprogramming efficiency at least partly by downregulating the expression of transcription factors that act as reprogramming barriers.

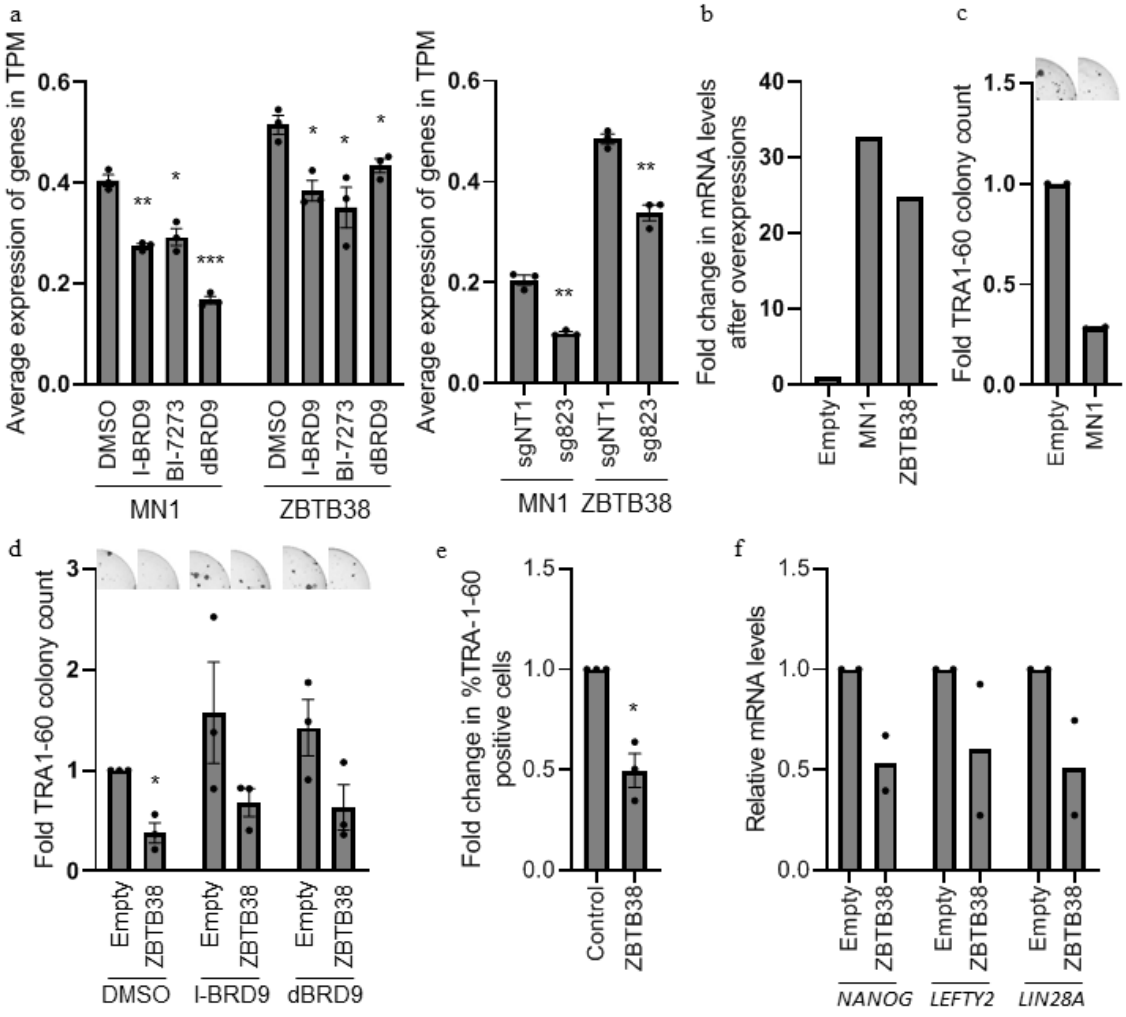


Figure 4.24 Transcriptional targets of BRD9, MN1 and ZBTB38 are barriers to reprogramming. (a) Average expression of *MN1* and *ZBTB38* in TPM across indicated fibroblasts. n=3, biological replicates. p-values were calculated by two sample t-test. * denotes p<0.05, ** denotes p<0.005, *** denotes p<0.0005. Exact p-values for small molecule treatments compared to DMSO control from left to right are: 0.0031, 0.0076 and 0.0002 for *MN1* and 0.0089, 0.0378 and 0.0284 for *ZBTB38*. Exact p-value of BRD9 gRNA expressing cells to control cells is 0.0045 for *MN1* and 0.0026 for *ZBTB38*. (b) Fold change in relative *MN1 and ZBTB38* mRNA levels upon their overexpression in fibroblasts compared empty vector controls (c) Fold change in the number of TRA-1-60-positive colonies upon *MN1* overexpression. Representative well images are above the graph. Bar graphs show the mean. n=2, independent biological replicates. (d) Fold change in the number of TRA-1-60-positive colonies upon *ZBTB38* overexpression and indicated treatments. Representative well images are above the graph. Bar graphs show the mean and error bars represent standard error of mean. n=3, independent biological replicates.

p-values are calculated by one sample t-test for mu=1. * denotes p<0.05, exact p-value is 0.0255. (e) Fold change in the percentage of TRA-1-60-positive cells on day 6 of reprogramming by comparing *ZBTB38* expressing cells to control. Bar graphs show the mean and error bars represent standard error of mean. n=3, three biological replicates. p-values were calculated by two sample t-test. * denotes p<0.05, exact p-value: 0.0272. (f) Fold change in relative mRNA levels for the *NANOG*, *LEFTY2* and *LIN28* genes on day 6 of reprogramming. Gene expression values were normalized to those observed in empty vector expressing fibroblasts. Bar graphs show the mean and dots indicate two biological replicates. These experiments were conducted with Gülben Gürhan Sevinç.

Chapter 5. DISCUSSION

5.1. Small molecule screen identified reprogramming boosting compounds that could replace KLF4 and MYC.

In this thesis, I investigated the effects of CREBBP/EP300 and BRD9-associated ncBAF complex on human somatic cell reprogramming to pluripotency. Various epigenetic probes enabled the study of these epigenetic factors in reprogramming. CREBBP/EP300 bromodomain inhibitor SGC-CBP30 increased the efficiency of human somatic cell reprogramming to iPSC 2.5-3-fold while I-CBP112 resulted in a relatively modest (2-fold) increase in the efficiency (Ebrahimi et al., 2019). Combining CREBBP/EP300 bromodomain inhibition with DOT1L inhibition boosted reprogramming efficiency up to 10-fold and enabled an efficient OS- and OSK-mediated reprogramming (Ebrahimi et al., 2019). However, inhibiting acetyltransferase function of these transcriptional co-activators impaired pluripotency acquisition. BRD9 bromodomain inhibitors; LP99, BI-7273 and I-BRD9 increased reprogramming efficiency 1.5-2-fold while dBRD9 enabled 2-fold increase in the efficiency. Importantly, both I-BRD9 and dBRD9 enabled OS- and OSK-mediated reprogramming to fully characterized iPSCs. Both BRD9 and CREBBP/EP300 bromodomain inhibitions increased the number of TRA-1-60 positive cells at the early stage of human somatic cell reprogramming (Ebrahimi et al., 2019; Sevinç et al., 2021). These results showed that these reprogramming boosting compounds could enable more efficient, rapid reprogramming and generate iPSCs by less transgenes.

Achieving chemically induced human pluripotency will facilitate the use of iPSCs clinical applications as the transgenes used in traditional reprogramming methods are oncogenic (Lin and Wu, 2015). Small molecule compound cocktails enabled efficient iPSC generation from MEFs without transgene expression (chemical reprogramming) (Hou et al., 2013; Zhao et al., 2015). However, these chemical cocktails failed to generate iPSCs without transgenes from human somatic cells, which suggest additional perturbations are needed for human cells. We achieved above 1% efficiency by OS-induced human pluripotency by treating cells with IVFC-Z (EPZ004777 (DOT1L inhibitor, iDOT1L), valproic acid (HDAC inhibitor), forskolin (adenylate cyclase activator) and CREBBP/EP300 bromodomain inhibitor (SGC-CBP30 or I-CBP112) at the first week of reprogramming and DZNep (S-adenosylhomocysteine hydrolase

inhibitor) at the second week of reprogramming). Interestingly, a recent study showed that in addition to DOT1L inhibitor (SGC0946), treating human somatic cells with SB431542 (TGF-β inhibitor) and RN-1 (LSD1 inhibitor) (S-SB-RN) enabled replacing OCT4 with most of POU families and a set of TFs along with retroviral SKM overexpression (Kim et al., 2021). It would be an interesting future direction to test if iPSC generation is possible by combining our small molecule cocktail with this in only SOX2-mediated human cellular reprogramming. Importantly, Kim et al. showed that overexpressing OCT4-substituting TFs did not necessarily enhanced OSKM-mediated reprogramming (Kim et al., 2021). Thus, it would be a promising approach to screen additional small molecules to replace SOX2 in OKM-induced reprogramming instead of focusing on previously identified reprogramming boosting compounds. If this screen reveals a potential compound, in combination with IVFC-Z and S-SB-RN, it could accomplish long sought chemically induced human pluripotency.

5.2. SWI/SNF focused CRISPR-Cas9 mediated knockout screen during human somatic cell reprogramming to pluripotency identified ncBAF complex as a barrier to the process.

By using CRISPR-Cas9 mediated knockout vectors, I screened all SWI/SNF subunits in fibroblast proliferation and reprogramming to pluripotency. Depleting subset of common subunits and BAF-specific subunit DPF2 impaired iPSC generation. The screen also revealed ACTL6A, SMARCA4, SMARCC1 and DPF2 as necessary factors of reprogramming. Previously, these proteins were shown to be subunits of esBAF complex and their depletion was shown to decrease mESC self-renewal (Alfert et al., 2019; Ho et al., 2009, 2011; Innis and Cabot, 2020). Specifically, esBAF co-operates with Oct4, Sox2 and Nanog to maintain the expression of pluripotency-specific genes (Ho et al., 2009, 2011). Knocking out subunits of this esBAF decreased reprogramming efficiency probably due to impaired activation of pluripotency circuit. Additionally, overexpressing esBAF-specific subunits, Brg1 and Smarcc1, has been shown to increase mouse somatic cell reprogramming to iPSC by facilitating Oct4 binding to its targets (Singhal et al., 2010). Importantly, this screen showed that knocking out ACTB and ACTL6A decreased reprogramming efficiency probably due to a concomitant decrease in fibroblast cell proliferation. On the contrary, a previous report showed that shRNA-mediated Actb knockdown increases mouse neural progenitor cell reprogramming to iPSCs. However, as the initial somatic cell type and the genetic approach to deplete ACTB is different with this study and my screen, it is difficult to compare. The most

plausible explanation is that retroviral shRNAs are silenced while cells acquire pluripotency while CRISPR-Cas9 mediated knockout is continuous during reprogramming. A transient ACTB depletion might increase the efficiency while continuous ACTB knockout is detrimental to pluripotency induction.

SWI/SNF focused knockout screen identified ncBAF-specific subunits, BRD9 and GLTSCR1 as barriers to human somatic cell reprogramming to pluripotency. Further validation experiments confirmed the screen with additional gRNAs and perturbation methods. Interestingly, this effect seems to be specific to BRD9 and GLTSCR1 as knocking out close paralogs BRD7 and GLTSCR1L, respectively, did not enhance the efficiency of somatic cell reprogramming. While BRD7 is the PBAF-specific subunit, GLTSCR1L is the specific subunit of ncBAF complex of which other specific subunits were shown to enhance reprogramming. There could be two possible explanations for this. The first is that CRISPR-Cas9 system failed to knockout GLTSCR1L at protein level or functional truncated versions of the protein were produced. We could not eliminate these possibilities due to lack of commercially available antibodies against the protein at the time of this thesis. The second explanation is that GLTSCR1 is the primary subunit of somatic ncBAF and could not be replaced by mutually exclusive subunit GLTSCR1L. To support the latter hypothesis, a previous study showed that GLTSCR1L, rather than GLTSCR1, binds primarily to SMARCA4 in mESC (Hota et al., 2019). Further biochemical and proteomics studies with available antibodies against GLTSCR1 and GLTSCR1L would be helpful to conclude which of these explanations is valid.

Due to limited representation caused by two sgRNAs per gene, possible off-targets caused ambiguous reprogramming outcome for certain genes. For example, two sgRNAs targeting SMARCB1 had opposing outcomes on reprogramming efficiency. Previously, SMARCB1 was shown to positively regulate mESC differentiation and it has a negative impact on mESC self-renewal which strengthen the possibility of its barrier role in reprogramming (You et al., 2013). The role of esBAF subunits ARID1A and SMARCD2 is inconclusive as only one sgRNA decreased the efficiency of human somatic cell reprogramming (Zhang et al., 2021a). On contrary, one sgRNA expression per each of esBAF subunits, BCL7C and SMARCD1, increased the reprogramming efficiency which was unexpected given their role in maintaining pluripotency (Zhang et al., 2021a). Further studies with additional sgRNAs or different depletion methods to conclude the role of these esBAF subunits in human somatic cell reprogramming are necessary.

5.3. The role of BRD9 in pluripotency and mesoderm differentiation.

We could generate iPSCs from BRD9 knockout fibroblasts and they could still show the canonical properties of pluripotent stem cells. This observations is interesting because previous study on naïve mESCs show that the cell type-specific transcription, proliferation, clonogenecity and self-renewal capacity decreased upon BRD9 bromodomain inhibition (Gatchalian et al., 2018), suggesting a role for BRD9 in maintaining pluripotency. However, we were able to propagate human BRD9 knockout iPSCs several passages without an observable difference compared to wild type cells. Additionally, 48 hour dBRD9 and I-BRD9 treatments did not decrease cell proliferation or OCT4-positive cell percentage compared to control treatment. Interestingly, continuous BRD9 inhibition/depletion during mouse somatic cell reprogramming to pluripotency did not increase the efficiency whereas human cellular reprogramming efficiency was increased upon BRD9 depletion (Janiszewski et al., 2019; Sevinç et al., 2021). However, the pluripotency state of iPSC achieved by mouse somatic cell reprogramming is more naïve-like, whereas human iPSCs are primed. Human iPSCs generated with conventional reprogramming as I employed in this thesis form flat colonies, contribute little to chimeric embryo when transplanted to other mammalian blastocysts and do not reactivate their inactive X chromosomes (Kilens et al., 2018). On the contrary, mouse iPSCs form dome-shaped colonies, contribute to chimeric embryo when transplanted into blastocyst and reactivate inactive X chromosome (Ghimire et al., 2018). The different roles of BRD9 on these two systems may stem from the difference between naïve and primed state of pluripotency or from the difference between species. For example, H3K9 methyltransferases EHMT1 and SETDB1 were shown to be barriers to reprogramming MEFs and pre-iPSCs into iPSCs (Chen et al., 2013; Soufi et al., 2012; Sridharan et al., 2013). Contrarily, both knockdowns decreased human somatic cell reprogramming efficiency (Onder et al., 2012). To conclude if the role of BRD9 in safeguarding cell identity is species- or pluripotent state-specific, further investigations on BRD9's effect on inducing and maintaining human naïve pluripotent stem cells are necessary.

BRD9 knockout impaired mesoderm differentiation of iPSCs revealed by investigating teratoma sections and marker gene expression upon CHIR-induced *in vitro* mesoderm differentiation. Molecular mechanisms governing mesoderm differentiation could be studied further by directed mammalian ESC differentiation into primitive streak, mesendoderm lineages *in vitro* and several factors were identified as regulators of this

process (Lam et al., 2014; Lei et al., 2019; Liu et al., 2017; Magli et al., 2019; Moris et al., 2020; Yilmaz et al., 2020; Zhang et al., 2019). Cells from mesoderm lineage were absent but ectoderm-lineage cells were present in embryoid bodies generated by spontaneous differentiation of Arid1a knockout mESCs in suspension and mesoderm and cardiomyocyte differentiations from hESCs were impaired upon ARID1A depletion (Gao et al., 2008; Lei et al., 2019). Deleting esBAF subunit Dpf2 impaired mesoderm lineage commitment which could be rescued by restoring Tbx3 expression (Zhang et al., 2019). These studies showed that esBAF subunits not only promote ESC self-renewal but also have a role in differentiation into embryonic lineages such as mesoderm.

A dual reporter mESC line for Foxa2 and Brachury expression was used to identify a role for epigenetic factors by shRNA mediated screen during mesendodermal lineage differentiation (Terzi Cizmecioglu et al., 2020). Interestingly, this screen did not identify Brd9 as a positive regulator of mesendoderm differentiation but focused on Arid4b whose suppression decreased mesendoderm lineage differentiation (Terzi Cizmecioglu et al., 2020). Similarly, a genome-wide CRISPR-Cas9 knockout screen performed during early lineage commitments of haploid hESCs did not identify BRD9 as a regulator of mesoderm differentiation. However, by using *TBXT* (Brachury, T), *EOMES* and *MIXL1* expressions to assess mesoderm differentiation induced by two-day treatment with a WNT pathway agonist, GSK3β inhibitor, CHIR99021, I showed that BRD9 positively regulates mesoderm differentiation *in vitro.* It is tempting to speculate that BRD9 loss impairs differentiation into mesoderm lineage from which fibroblasts, the initial somatic cell type used in reprogramming assays, were originated and, thus, factors safeguarding fibroblast identity during reprogramming could be used to predict their role in mesoderm differentiation. However, future experiments to assess the role of BRD9 in ectoderm or endoderm differentiation and the mechanism of impaired mesoderm differentiation upon BRD9 inhibition should be performed to reveal an antagonistic relationship between reprogramming and mesoderm differentiation as speculated.

5.4. CREBBP/EP300 bromodomains and BRD9-associated ncBAF complex maintain the cell type-specific gene expression.

In this thesis, I showed that inhibiting CREBBP/EP300 bromodomain and BRD9 restricts accessible chromatin at putative regulatory sites of somatic genes. This chromatin re-arrangement caused downregulation of somatic-specific genes in fibroblasts even without OSKM overexpression. As the somatic program was suppressed by inhibiting CREBBP/EP300 bromodomain or BRD9, questions arise such as what happens

to fibroblasts with continuous CREBBP/EP300 or BRD9 inhibition? Would they de-differentiate spontaneously, start dying or stop proliferating? For example, CAF-1 knockdown, another reprogramming barrier, decreased the cell proliferation rate around 30% at the tenth day without OSKM induction (Cheloufi et al., 2015). Additionally, five day treatment of BET inhibitor led fibroblasts to a more round/polygonal morphology rather than their conventional stretch and long morphology (Shao et al., 2016). However, the effect of inhibiting CREBBP/EP300 or BRD9 on somatic cell identity should be investigated by morphology change assessment and collagen assays to understand the level of safeguarding roles of these factors in cell identity. Additional long term cell proliferation, apoptosis and cell cycle profiling analysis would strengthen the argument for safeguarding effect of these chromatin factors on fibroblast cell identity. Yet, from the observations presented in this thesis, I speculate that although the weakening of somatic cell identity seems subtle and at the chromatin/transcriptomic levels, it is enough to enhance transcription factor-mediated cellular reprogramming.

CREBBP/EP300 were shown to safeguard cell identity and play roles in cell type conversions including differentiation and reprogramming. Knocking out CBP and p300 has been shown to be embryonic lethal and conditional *CBP/p300* knockout was used to assess these factors' role in development (Tanaka et al., 2000; Yao et al., 1998). Such studies revealed that these factors are necessary in hematopoiesis, expression of genes related to excitatory neuron identity, normal energy storage regulated by liver, CD8 positive T-cell development, Treg cell development, juxtaglomerular cell development in kidney and lens induction (Bedford et al., 2011; Gomez et al., 2009; Kasper et al., 2002; Lipinski et al., 2020; Liu et al., 2014; Tomofusa et al., 2009; Wolf et al., 2013). Using A-485, CBP/P300-mediated acetylation was shown to be essential for pancreatic β-cell identity and function in mice (Zhang et al., 2021b). Similarly, A485, a catalytic inhibitor of P300/CBP, downregulated the expression of *MITF*, a melanocyte-specific TF, and its downstream targets and induced senescence in human melanoma cell lines (Wang et al., 2018a). Super-enhancers characterized by p300 binding in T-helper cells were associated with non-coding RNA expressions and their associated genes were the most responsive ones to LPS induction and BET inhibition (Witte et al., 2015). CPI703, a CREBBP/EP300 bromodomain inhibitor, impaired differentiation of human naïve T cells into FOXP3 positive Treg cells and CPI703 treatment of Treg cells resulted in suppression of Treg-specific gene network (Ghosh et al., 2016). CBP30 and I-CBP112 increased TF-mediated transdifferentiation of MEFs into cardiomyocytes which shows that CREBBP/EP300

bromodomain is not a pluripotency-specific barrier (Lim et al., 2021). These findings are consistent with results presented in this thesis and point to a role of CREBBP/EP300 in safeguarding cell identity (Figure 5.1).

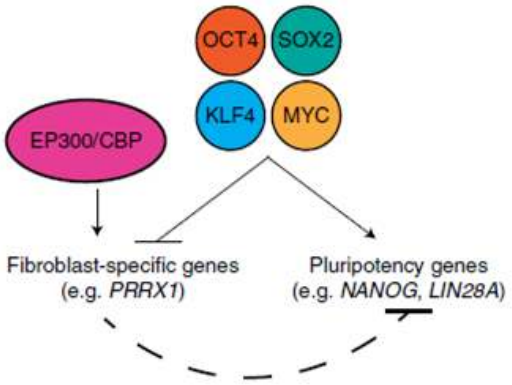


Figure 5.1 Model for CREBBP/EP300's role in maintaining somatic cell identity during reprogramming.

BRD9 has been shown to play a role in safeguarding cell identity, oncogenic transcription and various differentiation processes. Disruption of BRD9 splicing via mutant SF3B1 that includes a poison exon leading to BRD9 mRNA degradation in melanoma was shown to promote the growth of this cell type and correcting this splicing event increased BRD9 protein levels and suppressed the cell proliferation (Inoue et al., 2019). Similarly, BRD9 inhibition was shown to rescue the phenotype of decreased proliferation and tumor growth of merkel cell carcinoma achieved by LSD1 inhibition (Park et al., 2020). On the contrary, in acute myeloid leukemia (AML) cells, shRNA-mediated BRD9 knockdown and bromodomain inhibition by BI-7273 resulted in a decrease in cell proliferation and myeloid differentiation phenotype which could be rescued by overexpressing MYC (Del Gaudio et al., 2019; Hohmann et al., 2016). Additionally, BRD9 was shown to bind to putative enhancers to maintain a cell type-specific transcription network in AML cells (Del Gaudio et al., 2019). These studies did not investigate if cancer stem cells are the major drivers for decreased cell proliferation upon BRD9 inhibition. However, it is tempting to speculate that differentiated and highly proliferative cancer cells rather than cancer stem cells would be affected by BRD9 loss as this thesis found out a role for BRD9 in maintaining human somatic cells rather than pluripotent stem cells. It is tempting to speculate that BRD9 depletion mainly Chemical and genetic depletion of BRD9 impaired cell survival *in vitro* and tumor growth *in vivo*, in SS18-SSX fusion protein containing synovial sarcoma and SMARCB1 mutated

malignant rhabdoid tumor (Brien et al., 2018; Michel et al., 2018). These studies show that BRD9-containing ncBAF complex mediates cancer cell-specific oncogenic transcriptional programs in the absence/impairment of canonical BAF complexes. A recent study showed that BRD9 and ncBAF complex interact with androgen receptor to activate downstream genes whose expressions are necessary for prostate cancer cell survival *in vitro* and *in vivo* (Alpsoy et al., 2021). Brd9 was, also, shown to promote Foxp3 binding to its targets to regulate gene expression which is necessary to safeguard Treg cell identity *in vitro* and *in vivo* (Loo et al., 2020). The growing literature and the results presented in this thesis clearly show that BRD9 and associated ncBAF complex are important in safeguarding cell identity through interacting with cell type-specific TFs at regulatory sites to activate downstream genes (Figure 5.2).

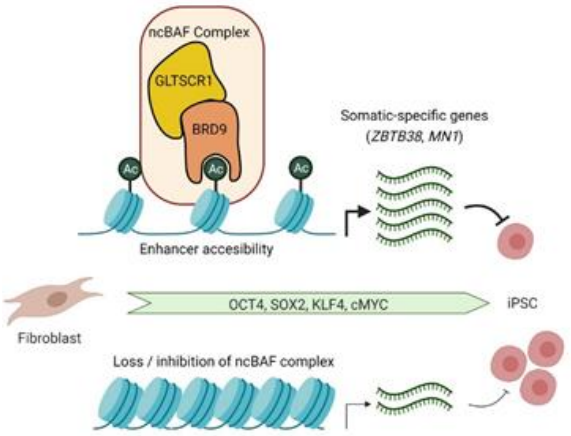


Figure 5.2 Model for BRD9-containing BAF complex's role in maintaining somatic cell identity during reprogramming.

CBP/EP300 bromodomain and BRD9 inhibitions do not increase reprogramming efficiency additively. This observation suggests that both BRD9 and CBP/EP300 bromodomains function in the same pathway to preserve somatic cell identity in a non-redundant manner. Interestingly, previous study showed that BRD9 promotes p300 binding to regulatory sites to regulate gene expression in hepatocellular carcinoma (Dou et al., 2020). Indeed, we show that CBP/EP300 bromodomain target genes are negatively enriched upon BRD9 inhibitions in fibroblasts. If BRD9 promotes p300 binding to fibroblast enhancers as in hepatocellular carcinoma, I would expect to observe depletion of p300 from these sites after BRD9 depletion. On the other hand, if p300 localizes BRD9 to putative fibroblast enhancers, I would expect a depletion of BRD9 from these sites upon CREBBP/EP300 inhibition. It would be motivating to study if such a mechanism

exist in human fibroblasts by such genomic and biochemical assays.

5.5. Role of somatic transcription factors in resisting pluripotency acquisition

Previous studies show barrier characteristics of somatic transcription factors such as TWIST1, SNAI1, ZEB1, PRRX1, RUNX1 and PRRX2 for reprogramming. Downregulating or overexpressing such factors would enhance or impair the process, respectively (Ebrahimi et al., 2019; Mellis et al., 2020; Onder et al., 2012; Yang et al., 2011). In this thesis, I showed that *PRRX1* knockdown increased human reprogramming efficiency, consistent with what was previously shown for mouse somatic cell reprogramming (Yang et al., 2011). In addition, I showed that overexpressing PRRX1, ZBTB38 and MN1 blocks human somatic cell reprogramming to pluripotency. These results showed that CREBBP/EP300 bromodomain and BRD9 sustain the expression of reprogramming barriers. However, the way I overexpressed these TFs was beyond the physiological levels and a less active or TET-inducible promoter could be used to overexpress these genes just around 2-3-folds. Also, the better way to interrogate a possible cause and effect in such system would be knocking down TF and combining this with small molecule treatment during reprogramming. If these two perturbations individually increase the reprogramming efficiency but their combination has no further effect on reprogramming, then it would be fair to conclude that CREBBP/EP300 and BRD9 are barriers to reprogramming through maintaining the expression of TFs *PRRX1*, *ZBTB38* and *MN1*. Nonetheless, results presented in this thesis showed that continuous expression of *PRRX1*, *ZBTB38* and *MN1* is detrimental to pluripotency induction and CREBBP/EP300 and BRD9 positively regulate the expression of these genes.

In conclusion, the results presented in this thesis revealed CREBBP/EP300 bromodomain and non-canonical BAF complex as chromatin-based barriers to human somatic cell reprogramming to pluripotency. These factors were shown to maintain the expression of somatic-specific genes and inhibitors utilized in this thesis, thereby, may increase the efficiency of other reprogramming approaches. Additionally, small molecule combinations and BRD9 inhibition that were demonstrated to substitute KLF4 and MYC transgene expression could be included in long-sought chemical cocktails to generate human iPSCs without OSKM.

Bibliography

Aarts, M., Georgilis, A., Beniazza, M., Beolchi, P., Banito, A., Carroll, T., Kulisic, M., Kaemena, D.F., Dharmalingam, G., Martin, N., et al. (2017). Coupling shRNA screens with single-cell RNA-seq identifies a dual role for mTOR in reprogramming-induced senescence. Genes Dev. *31*, 2085–2098.

Aasen, T., Raya, A., Barrero, M.J., Garreta, E., Consiglio, A., Gonzalez, F., Vassena, R., Bilić, J., Pekarik, V., Tiscornia, G., et al. (2008). Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. Nat. Biotechnol. *26*, 1276–1284.

Akbari, S., Sevinç, G.G., Ersoy, N., Basak, O., Kaplan, K., Sevinç, K., Ozel, E., Sengun, B., Enustun, E., Ozcimen, B., et al. (2019). Robust, Long-Term Culture of Endoderm-Derived Hepatic Organoids for Disease Modeling. Stem Cell Reports *13*, 627–641.

Alfert, A., Moreno, N., and Kerl, K. (2019). The BAF complex in development and disease. Epigenetics Chromatin *12*, 19.

Alici-Garipcan, A., Özçimen, B., Süder, I., Ülker, V., Önder, T.T., and Özören, N. (2020). NLRP7 plays a functional role in regulating BMP4 signaling during differentiation of patient-derived trophoblasts. Cell Death Dis. *11*, 658.

Allshire, R.C., and Madhani, H.D. (2018). Ten principles of heterochromatin formation and function. Nat. Rev. Mol. Cell Biol. *19*, 229–244.

Alpsoy, A., and Dykhuizen, E.C. (2018). Glioma tumor suppressor candidate region gene 1 (GLTSCR1) and its paralog GLTSCR1-like form SWI/SNF chromatin remodeling subcomplexes. J. Biol. Chem. *293*, 3892–3903.

Alpsoy, A., Utturkar, S.M., Carter, B.C., Dhiman, A., Torregrosa-Allen, S.E., Currie, M.P., Elzey, B.D., and Dykhuizen, E.C. (2021). BRD9 Is a Critical Regulator of Androgen Receptor Signaling and Prostate Cancer Progression. Cancer Res. *81*, 820–833.

Alver, B.H., Kim, K.H., Lu, P., Wang, X., Manchester, H.E., Wang, W., Haswell, J.R., Park, P.J., and Roberts, C.W.M. (2017). The SWI/SNF chromatin remodelling complex is required for maintenance of lineage specific enhancers. Nat. Commun. *8*, 1–10.

Aoi, T., Yae, K., Nakagawa, M., Ichisaka, T., Okita, K., Takahashi, K., Chiba, T., and Yamanaka, S. (2008). Generation of pluripotent stem cells from adult mouse liver and stomach cells. Science *321*, 699–702.

Arabacı, D.H., Terzioğlu, G., Bayırbaşı, B., and Önder, T.T. (2020). Going up the hill: chromatin-based barriers to epigenetic reprogramming. FEBS J. *n/a*.

Attar, N., and Kurdistani, S.K. (2017). Exploitation of EP300 and CREBBP Lysine Acetyltransferases by Cancer. Cold Spring Harb. Perspect. Med. *7*, a026534.

Bae, S., and Lesch, B.J. (2020). H3K4me1 Distribution Predicts Transcription State and Poising at Promoters. Front. Cell Dev. Biol. *8*, 289.

Baker, C.L., and Pera, M.F. (2018). Capturing Totipotent Stem Cells. Cell Stem Cell *22*, 25–34.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res. *21*, 381–395.

Beagan, J.A., and Phillips-Cremins, J.E. (2020). On the existence and functionality of topologically associating domains. Nat. Genet. *52*, 8–16.

Bedford, D.C., and Brindle, P.K. (2012). Is histone acetylation the most important physiological function for CBP and p300? Aging (Albany. NY). *4*, 247–255.

Bedford, D.C., Kasper, L.H., Wang, R., Chang, Y., Green, D.R., and Brindle, P.K. (2011). Disrupting the CH1 Domain Structure in the Acetyltransferases CBP and p300 Results in Lean Mice with Increased Metabolic Control. Cell Metab. *14*, 219–230.

Blum, R. (2015). Stepping inside the realm of epigenetic modifiers. Biomol. Concepts *6*, 119–136.

Borkent, M., Bennett, B.D., Lackford, B., Bar-Nur, O., Brumbaugh, J., Wang, L., Du, Y., Fargo, D.C., Apostolou, E., Cheloufi, S., et al. (2016). A Serial shRNA Screen for Roadblocks to Reprogramming Identifies the Protein Modifier SUMO2. Stem Cell Reports *6*, 704–716.

Bredenkamp, N., Yang, J., Clarke, J., Stirparo, G.G., von Meyenn, F., Dietmann, S., Baker, D., Drummond, R., Ren, Y., Li, D., et al. (2019). Wnt Inhibition Facilitates RNA-Mediated Reprogramming of Human Somatic Cells to Naive Pluripotency. Stem Cell Reports *13*, 1083–1098.

Brien, G.L., Remillard, D., Shi, J., Hemming, M.L., Chabon, J., Wynne, K., Dillon, E.T., Cagney, G., Van Mierlo, G., Baltissen, M.P., et al. (2018). Targeted degradation of BRD9 reverses oncogenic gene expression in synovial sarcoma. Elife *7*, e41305.

Bright, J., Hussain, S., Dang, V., Wright, S., Cooper, B., Byun, T., Ramos, C., Singh, A., Parry, G., Stagliano, N., et al. (2015). Human secreted tau increases amyloid-beta production. Neurobiol. Aging *36*, 693–709.

Brumbaugh, J., Di Stefano, B., and Hochedlinger, K. (2019). Reprogramming:

identifying the mechanisms that safeguard cell identity. Development *146*, dev182170.

Buganim, Y., Faddah, D.A., and Jaenisch, R. (2013). Mechanisms and models of somatic cell reprogramming. Nat. Rev. Genet. *14*, 427–439.

Byrne, J.A., Simonsson, S., Western, P.S., and Gurdon, J.B. (2003). Nuclei of adult mammalian somatic cells are directly reprogrammed to oct-4 stem cell gene expression by amphibian oocytes. Curr. Biol. *13*, 1206–1213.

Cacchiarelli, D., Trapnell, C., Ziller, M.J., Soumillon, M., Cesana, M., Karnik, R., Donaghey, J., Smith, Z.D., Ratanasirintrawoot, S., Zhang, X., et al. (2015). Integrative Analyses of Human Reprogramming Reveal Dynamic Nature of Induced Pluripotency. Cell *162*, 412–424.

Calo, E., and Wysocka, J. (2013). Modification of Enhancer Chromatin: What, How, and Why? Mol. Cell *49*, 825–837.

Camp, J.G., Badsha, F., Florio, M., Kanton, S., Gerber, T., Wilsch-Bräuninger, M., Lewitus, E., Sykes, A., Hevers, W., Lancaster, M., et al. (2015). Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. Proc. Natl. Acad. Sci. U. S. A. *112*, 15672–15677.

Catarino, R.R., and Stark, A. (2018). Assessing sufficiency and necessity of enhancer activities for gene expression and the mechanisms of transcription activation. Genes Dev. *32*, 202–223.

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., et al. (2009). Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. Nat. Biotechnol. *27*, 1033–1037.

Chan, Y.-S., Göke, J., Ng, J.-H., Lu, X., Gonzales, K.A.U., Tan, C.-P., Tng, W.-Q., Hong, Z.-Z., Lim, Y.-S., and Ng, H.-H. (2013). Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. Cell Stem Cell *13*, 663–675.

Cheloufi, S., Elling, U., Hopfgartner, B., Jung, Y.L., Murn, J., Ninova, M., Hubmann, M., Badeaux, A.I., Euong Ang, C., Tenen, D., et al. (2015). The histone chaperone CAF-1 safeguards somatic cell identity. Nature *528*, 218–224.

Chen, Y., and Lai, D. (2015). Pluripotent states of human embryonic stem cells. Cell. Reprogram. *17*, 1–6.

Chen, J., Liu, H., Liu, J., Qi, J., Wei, B., Yang, J., Liang, H., Chen, Y., Chen, J., Wu, Y., et al. (2013). H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. Nat. Genet. *45*, 34–42.

Choi, J., Costa, M.L., Mermelstein, C.S., Chagas, C., Holtzer, S., and Holtzer, H. (1990). MyoD converts primary dermal fibroblasts, chondroblasts, smooth muscle, and retinal pigmented epithelial cells into striated mononucleated myoblasts and multinucleated myotubes. Proc. Natl. Acad. Sci. U. S. A. *87*, 7988–7992.

Chronis, C., Fiziev, P., Papp, B., Butz, S., Bonora, G., Sabri, S., Ernst, J., and Plath, K. (2017). Cooperative Binding of Transcription Factors Orchestrates Reprogramming. Cell *168*, 442-459.e20.

Cossec, J.-C., Theurillat, I., Chica, C., Búa Aguín, S., Gaume, X., Andrieux, A., Iturbide, A., Jouvion, G., Li, H., Bossis, G., et al. (2018). SUMO Safeguards Somatic and Pluripotent Cell Identities by Enforcing Distinct Chromatin States. Cell Stem Cell *23*, 742-757.e8.

Cowan, C.A., Klimanskaya, I., McMahon, J., Atienza, J., Witmyer, J., Zucker, J.P., Wang, S., Morton, C.C., McMahon, A.P., Powers, D., et al. (2004). Derivation of embryonic stem-cell lines from human blastocysts. N. Engl. J. Med. *350*, 1353–1356.

Cowan, C.A., Atienza, J., Melton, D.A., and Eggan, K. (2005). Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. Science *309*, 1369–1373.

Cyranoski, D. (2018). "Reprogrammed" stem cells approved to mend human hearts for the first time. Nature *557*, 619–620.

Dahéron, L., Opitz, S.L., Zaehres, H., Lensch, M.W., Andrews, P.W., Itskovitz-Eldor, J., and Daley, G.Q. (2004). LIF/STAT3 signaling fails to maintain self-renewal of human embryonic stem cells. Stem Cells *22*, 770–778.

Do, J.T., and Schöler, H.R. (2006). Cell-cell fusion as a means to establish pluripotency. Ernst Schering Res. Found. Workshop 35–45.

Dong, X., and Weng, Z. (2013). The correlation between histone modifications and gene expression. Epigenomics *5*, 113–116.

Dou, C., Sun, L., Wang, L., Cheng, J., Wu, W., Zhang, C., Xu, Q., Tu, K., and Liu, J. (2020). Bromodomain-containing protein 9 promotes the growth and metastasis of human hepatocellular carcinoma by activating the TUFT1/AKT pathway. Cell Death Dis. *11*, 730.

Duran Alonso, M.B., Lopez Hernandez, I., de la Fuente, M.A., Garcia-Sancho, J., Giraldez, F., and Schimmang, T. (2018). Transcription factor induced conversion of human fibroblasts towards the hair cell lineage. PLoS One *13*, e0200210.

Dye, B.R., Hill, D.R., Ferguson, M.A.H., Tsai, Y.-H., Nagy, M.S., Dyal, R., Wells, J.M., Mayhew, C.N., Nattiv, R., Klein, O.D., et al. (2015). In vitro generation of human

pluripotent stem cell derived lung organoids. Elife *4*.

Ebrahimi, A., Sevinç, K., Gürhan Sevinç, G., Cribbs, A.P., Philpott, M., Uyulur, F., Morova, T., Dunford, J.E., Göklemez, S., Arı, Ş., et al. (2019). Bromodomain inhibition of the coactivators CBP/EP300 facilitate cellular reprogramming. Nat. Chem. Biol. *15*, 519–528.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154–156.

Fauquier, L., Azzag, K., Parra, M.A.M., Quillien, A., Boulet, M., Diouf, S., Carnac, G., Waltzer, L., Gronemeyer, H., and Vandel, L. (2018). CBP and P300 regulate distinct gene networks required for human primary myoblast differentiation and muscle integrity. Sci. Rep. *8*, 12629.

Fidan, K., Kavaklıoğlu, G., Ebrahimi, A., Özlü, C., Ay, N.Z., Ruacan, A., Gül, A., and Önder, T.T. (2015). Generation of integration-free induced pluripotent stem cells from a patient with Familial Mediterranean Fever (FMF). Stem Cell Res. *15*, 694–696.

Firas, J., Liu, X., Lim, S.M., and Polo, J.M. (2015). Transcription factor-mediated reprogramming: epigenetics and therapeutic potential. Immunol. Cell Biol. *93*, 284–289.

FISCHBERG, M., GURDON, J.B., and ELSDALE, T.R. (1958). Nuclear Transplantation in Xenopus laevis. Nature *181*, 424.

Flynn, E.M., Huang, O.W., Poy, F., Oppikofer, M., Bellon, S.F., Tang, Y., and Cochran, A.G. (2015). A Subset of Human Bromodomains Recognizes Butyryllysine and Crotonyllysine Histone Peptide Modifications. Structure *23*, 1801–1814.

Fu, K., Chronis, C., Soufi, A., Bonora, G., Edwards, M., Smale, S.T., Zaret, K.S., Plath, K., and Pellegrini, M. (2018). Comparison of reprogramming factor targets reveals both species-specific and conserved mechanisms in early iPSC reprogramming. BMC Genomics *19*, 956.

Gabut, M., Samavarchi-Tehrani, P., Wang, X., Slobodeniuc, V., O'Hanlon, D., Sung, H.-K., Alvarez, M., Talukder, S., Pan, Q., Mazzoni, E.O., et al. (2011). An alternative splicing switch regulates embryonic stem cell pluripotency and reprogramming. Cell *147*, 132–146.

Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A., et al. (2013). Derivation of novel human ground state naive pluripotent stem cells. Nature *504*, 282–286.

Gao, X., Tate, P., Hu, P., Tjian, R., Skarnes, W.C., and Wang, Z. (2008). ES cell pluripotency and germ-layer formation require the SWI/SNF chromatin remodeling

component BAF250a. Proc. Natl. Acad. Sci. *105*, 6656 LP – 6661.

Garcez, P.P., Loiola, E.C., Madeiro da Costa, R., Higa, L.M., Trindade, P., Delvecchio, R., Nascimento, J.M., Brindeiro, R., Tanuri, A., and Rehen, S.K. (2016). Zika virus impairs growth in human neurospheres and brain organoids. Science *352*, 816–818.

Garcia-Carpizo, V., Ruiz-Llorente, S., Sarmentero, J., Graña-Castro, O., Pisano, D.G., and Barrero, M.J. (2018). CREBBP/EP300 bromodomains are critical to sustain the GATA1/MYC regulatory axis in proliferation. Epigenetics Chromatin *11*, 30.

Gatchalian, J., Malik, S., Ho, J., Lee, D.-S., Kelso, T.W.R., Shokhirev, M.N., Dixon, J.R., and Hargreaves, D.C. (2018). A non-canonical BRD9-containing BAF chromatin remodeling complex regulates naive pluripotency in mouse embryonic stem cells. Nat. Commun. *9*, 5139.

Del Gaudio, N., Di Costanzo, A., Liu, N.Q., Conte, L., Migliaccio, A., Vermeulen, M., Martens, J.H.A., Stunnenberg, H.G., Nebbioso, A., and Altucci, L. (2019). BRD9 binds cell type-specific chromatin regions regulating leukemic cell survival via STAT5 inhibition. Cell Death Dis. *10*, 338.

Ghimire, S., Van der Jeught, M., Neupane, J., Roost, M.S., Anckaert, J., Popovic, M., Van Nieuwerburgh, F., Mestdagh, P., Vandesompele, J., Deforce, D., et al. (2018). Comparative analysis of naive, primed and ground state pluripotency in mouse embryonic stem cells originating from the same genetic background. Sci. Rep. *8*, 5884.

Ghosh, S., Taylor, A., Chin, M., Huang, H.-R., Conery, A.R., Mertz, J.A., Salmeron, A., Dakle, P.J., Mele, D., Cote, A., et al. (2016). Regulatory T Cell Modulation by CBP/EP300 Bromodomain Inhibition. J. Biol. Chem. *291*, 13014–13027.

Golipour, A., David, L., Liu, Y., Jayakumaran, G., Hirsch, C.L., Trcka, D., and Wrana, J.L. (2012). A Late Transition in Somatic Cell Reprogramming Requires Regulators Distinct from the Pluripotency Network. Cell Stem Cell *11*, 769–782.

Gomez, R.A., Pentz, E.S., Jin, X., Cordaillat, M., and Sequeira Lopez, M.L.S. (2009). CBP and p300 are essential for renin cell identity and morphological integrity of the kidney. Am. J. Physiol. Circ. Physiol. *296*, H1255–H1262.

González, F., and Huangfu, D. (2016). Mechanisms underlying the formation of induced pluripotent stem cells. Wiley Interdiscip. Rev. Dev. Biol. *5*, 39–65.

GURDON, J.B. (1960). The developmental capacity of nuclei taken from differentiating endoderm cells of Xenopus laevis. J. Embryol. Exp. Morphol. *8*, 505–526.

GURDON, J.B. (1962a). The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. J. Embryol. Exp. Morphol. *10*, 622–640.

GURDON, J.B. (1962b). Adult frogs derived from the nuclei of single somatic cells. Dev. Biol. *4*, 256–273.

Haberle, V., and Stark, A. (2018). Eukaryotic core promoters and the functional basis of transcription initiation. Nat. Rev. Mol. Cell Biol. *19*, 621–637.

Hamdan, F.H., and Johnsen, S.A. (2018). Super enhancers - new analyses and perspectives on the low hanging fruit. Transcription *9*, 123–130.

Hanna, J., Carey, B.W., and Jaenisch, R. (2008). Reprogramming of somatic cell identity. Cold Spring Harb. Symp. Quant. Biol. *73*, 147–155.

Hanna, J., Cheng, A.W., Saha, K., Kim, J., Lengner, C.J., Soldner, F., Cassady, J.P., Muffat, J., Carey, B.W., and Jaenisch, R. (2010). Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. Proc. Natl. Acad. Sci. U. S. A. *107*, 9222–9227.

Hargreaves, D.C., and Crabtree, G.R. (2011). ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. Cell Res. *21*, 396–420.

Hay, D.A., Fedorov, O., Martin, S., Singleton, D.C., Tallant, C., Wells, C., Picaud, S., Philpott, M., Monteiro, O.P., Rogers, C.M., et al. (2014). Discovery and optimization of small-molecule ligands for the CBP/p300 bromodomains. J. Am. Chem. Soc. *136*, 9308–9319.

Heinz, S., Romanoski, C.E., Benner, C., and Glass, C.K. (2015). The selection and function of cell type-specific enhancers. Nat. Rev. Mol. Cell Biol. *16*, 144–154.

Hennig, A.K., Peng, G.-H., and Chen, S. (2013). Transcription Coactivators p300 and CBP Are Necessary for Photoreceptor-Specific Chromatin Organization and Gene Expression. PLoS One *8*, e69721.

Hernandez, C., Wang, Z., Ramazanov, B., Tang, Y., Mehta, S., Dambrot, C., Lee, Y.-W., Tessema, K., Kumar, I., Astudillo, M., et al. (2018). Dppa2/4 Facilitate Epigenetic Remodeling during Reprogramming to Pluripotency. Cell Stem Cell *23*, 396-411.e8.

Hirsch, C.L., Coban Akdemir, Z., Wang, L., Jayakumaran, G., Trcka, D., Weiss, A., Hernandez, J.J., Pan, Q., Han, H., Xu, X., et al. (2015). Myc and SAGA rewire an alternative splicing network during early somatic cell reprogramming. Genes Dev. *29*, 803–816.

HniszBrian, D., Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A., Young, R.A., and HniszBrian, D. (2013). Super-Enhancers in the Control of Cell Identity and Disease. Cell *155*, 934–947.

Ho, L., Ronan, J.L., Wu, J., Staahl, B.T., Chen, L., Kuo, A., Lessard, J., Nesvizhskii, A.I.,

Ranish, J., and Crabtree, G.R. (2009). An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. Proc. Natl. Acad. Sci. *106*, 5181 LP – 5186.

Ho, L., Miller, E.L., Ronan, J.L., Ho, W.Q., Jothi, R., and Crabtree, G.R. (2011). EsBAF facilitates pluripotency by conditioning the genome for LIF/STAT3 signalling and by regulating polycomb function. Nat. Cell Biol. *13*, 903–913.

Hohmann, A.F., Martin, L.J., Minder, J.L., Roe, J.-S., Shi, J., Steurer, S., Bader, G., McConnell, D., Pearson, M., Gerstberger, T., et al. (2016). Sensitivity and engineered resistance of myeloid leukemia cells to BRD9 inhibition. Nat. Chem. Biol. *12*, 672–679.

Hota, S.K., Johnson, J.R., Verschueren, E., Thomas, R., Blotnick, A.M., Zhu, Y., Sun, X., Pennacchio, L.A., Krogan, N.J., and Bruneau, B.G. (2019). Dynamic BAF chromatin remodeling complex subunit inclusion promotes temporally distinct gene expression programs in cardiogenesis. Development *146*.

Hou, P., Li, Y., Zhang, X., Liu, C., Guan, J., Li, H., Zhao, T., Ye, J., Yang, W., Liu, K., et al. (2013). Pluripotent Stem Cells Induced from Mouse Somatic Cells by Small-Molecule Compounds. Science (80-.). *341*, 651 LP – 654.

Hu, E., Tontonoz, P., and Spiegelman, B.M. (1995). Transdifferentiation of myoblasts by the adipogenic transcription factors PPAR gamma and C/EBP alpha. Proc. Natl. Acad. Sci. U. S. A. *92*, 9856–9860.

Huang, Y., Zhang, H., Wang, L., Tang, C., Qin, X., Wu, X., Pan, M., Tang, Y., Yang, Z., Babarinde, I.A., et al. (2020). JMJD3 acts in tandem with KLF4 to facilitate reprogramming to pluripotency. Nat. Commun. *11*, 5061.

Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., and Melton, D.A. (2008). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. Nat. Biotechnol. *26*, 795–797.

Huddy, T., Evans, R.M., Ross, B., Dai, Y., Pinto, A.F.M., Yoshihara, E., Fang, S., Liddle, C., Wei, Z., He, N., et al. (2018). Vitamin D Switches BAF Complexes to Protect β Cells. Cell *173*, 1135-1149.e15.

Huisinga, K.L., Brower-Toland, B., and Elgin, S.C.R. (2006). The contradictory definitions of heterochromatin: transcription and silencing. Chromosoma *115*, 110–122.

Hyun, K., Jeon, J., Park, K., and Kim, J. (2017). Writing, erasing and reading histone lysine methylations. Exp. Mol. Med. *49*, e324–e324.

Ichida, J.K., Blanchard, J., Lam, K., Son, E.Y., Chung, J.E., Egli, D., Loh, K.M., Carter, A.C., Di Giorgio, F.P., Koszka, K., et al. (2009). A small-molecule inhibitor of tgf-Beta

signaling replaces sox2 in reprogramming by inducing nanog. Cell Stem Cell *5*, 491–503.

Ieda, M. (2013). Direct reprogramming into desired cell types by defined factors. Keio J. Med. *62*, 74–82.

Innis, S.M., and Cabot, B. (2020). GBAF, a small BAF sub-complex with big implications: a systematic review. Epigenetics Chromatin *13*, 48.

Inoue, D., Chew, G.-L., Liu, B., Michel, B.C., Pangallo, J., D'Avino, A.R., Hitchman, T., North, K., Lee, S.C.-W., Bitner, L., et al. (2019). Spliceosomal disruption of the non-canonical BAF complex in cancer. Nature *574*, 432–436.

Iyer, N.G., Özdag, H., and Caldas, C. (2004). p300/CBP and cancer. Oncogene *23*, 4225–4231.

Janiszewski, A., Talon, I., Chappell, J., Collombet, S., Song, J., De Geest, N., To, S.K., Bervoets, G., Marin-Bejar, O., Provenzano, C., et al. (2019). Dynamic reversal of random X-Chromosome inactivation during iPSC reprogramming. Genome Res. *29*, 1659–1672.

Jedrusik, A. (2015). Making the first decision: lessons from the mouse. Reprod. Med. Biol. *14*, 135–150.

Jiang, Q., Huang, X., Hu, X., Shan, Z., Wu, Y., Wu, G., and Lei, L. (2020). Histone demethylase KDM6A promotes somatic cell reprogramming by epigenetically regulating the PTEN and IL-6 signal pathways. Stem Cells *38*, 960–972.

Jiang, Z., Tang, Y., Zhao, X., Zhang, M., Donovan, D.M., and Tian, X. (Cindy) (2015). Knockdown of Brm and Baf170, Components of Chromatin Remodeling Complex, Facilitates Reprogramming of Somatic Cells. Stem Cells Dev. *24*, 2328–2336.

Jin, F., Li, Y., Ren, B., and Natarajan, R. (2011). Enhancers: multi-dimensional signal integrators. Transcription *2*, 226–230.

Karlić, R., Chung, H.-R., Lasserre, J., Vlahoviček, K., and Vingron, M. (2010). Histone modification levels are predictive for gene expression. Proc. Natl. Acad. Sci. *107*, 2926 LP – 2931.

Kasper, L.H., Boussouar, F., Ney, P.A., Jackson, C.W., Rehg, J., van Deursen, J.M., and Brindle, P.K. (2002). A transcription-factor-binding surface of coactivator p300 is required for haematopoiesis. Nature *419*, 738–743.

Kilens, S., Meistermann, D., Moreno, D., Chariau, C., Gaignerie, A., Reignier, A., Lelièvre, Y., Casanova, M., Vallot, C., Nedellec, S., et al. (2018). Parallel derivation of isogenic human primed and naive induced pluripotent stem cells. Nat. Commun. *9*, 360.

Kim, E.J.Y., Anko, M.-L., Flensberg, C., Majewski, I.J., Geng, F.-S., Firas, J., Huang,

D.C.S., van Delft, M.F., and Heath, J.K. (2018). BAK/BAX-Mediated Apoptosis Is a Myc-Induced Roadblock to Reprogramming. Stem Cell Reports *10*, 331–338.

Kim, K.-P., Choi, J., Yoon, J., Bruder, J.M., Shin, B., Kim, J., Arauzo-Bravo, M.J., Han, D., Wu, G., Han, D.W., et al. (2021). Permissive epigenomes endow reprogramming competence to transcriptional regulators. Nat. Chem. Biol. *17*, 47–56.

Kim, S.-K., Jung, I., Lee, H., Kang, K., Kim, M., Jeong, K., Kwon, C.S., Han, Y.-M., Kim, Y.S., Kim, D., et al. (2012). Human histone H3K79 methyltransferase DOT1L protein [corrected] binds actively transcribing RNA polymerase II to regulate gene expression. J. Biol. Chem. *287*, 39698–39709.

Kimura, H. (2013). Histone modifications for human epigenome analysis. J. Hum. Genet. *58*, 439–445.

Klein, J.C., Agarwal, V., Inoue, F., Keith, A., Martin, B., Kircher, M., Ahituv, N., and Shendure, J. (2020). A systematic evaluation of the design and context dependencies of massively parallel reporter assays. Nat. Methods *17*, 1083–1091.

Klimanskaya, I., Chung, Y., Becker, S., Lu, S.-J., and Lanza, R. (2006). Human embryonic stem cell lines derived from single blastomeres. Nature *444*, 481–485.

Knaupp, A.S., Buckberry, S., Pflueger, J., Lim, S.M., Ford, E., Larcombe, M.R., Rossello, F.J., de Mendoza, A., Alaei, S., Firas, J., et al. (2017). Transient and Permanent Reconfiguration of Chromatin and Transcription Factor Occupancy Drive Reprogramming. Cell Stem Cell *21*, 834-845.e6.

Koche, R.P., Smith, Z.D., Adli, M., Gu, H., Ku, M., Gnirke, A., Bernstein, B.E., and Meissner, A. (2011). Reprogramming factor expression initiates widespread targeted chromatin remodeling. Cell Stem Cell *8*, 96–105.

Kolundzic, E., Ofenbauer, A., Bulut, S.I., Uyar, B., Baytek, G., Sommermeier, A., Seelk, S., He, M., Hirsekorn, A., Vucicevic, D., et al. (2018). FACT Sets a Barrier for Cell Fate Reprogramming in Caenorhabditis elegans and Human Cells. Dev. Cell *46*, 611-626.e12.

Kornberg, R.D. (1977). Structure of chromatin. Annu. Rev. Biochem. *46*, 931–954.

Kvon, E.Z. (2015). Using transgenic reporter assays to functionally characterize enhancers in animals. Genomics *106*, 185–192.

Lai, B., Lee, J.-E., Jang, Y., Wang, L., Peng, W., and Ge, K. (2017). MLL3/MLL4 are required for CBP/p300 binding on enhancers and super-enhancer formation in brown adipogenesis. Nucleic Acids Res. *45*, 6388–6403.

Lam, A.Q., Freedman, B.S., Morizane, R., Lerou, P.H., Valerius, M.T., and Bonventre, J. V (2014). Rapid and efficient differentiation of human pluripotent stem cells into

intermediate mesoderm that forms tubules expressing kidney proximal tubular markers. J. Am. Soc. Nephrol. *25*, 1211–1225.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Lasko, L.M., Jakob, C.G., Edalji, R.P., Qiu, W., Montgomery, D., Digiammarino, E.L., Hansen, T.M., Risi, R.M., Frey, R., Manaves, V., et al. (2017). Discovery of a selective catalytic p300/CBP inhibitor that targets lineage-specific tumours. Nature *550*, 128–132.

Lattanzi, L., Salvatori, G., Coletta, M., Sonnino, C., Cusella De Angelis, M.G., Gioglio, L., Murry, C.E., Kelly, R., Ferrari, G., Molinaro, M., et al. (1998). High efficiency myogenic conversion of human fibroblasts by adenoviral vector-mediated MyoD gene transfer. An alternative strategy for ex vivo gene therapy of primary myopathies. J. Clin. Invest. *101*, 2119–2128.

Lei, I., Tian, S., Chen, V., Zhao, Y., and Wang, Z. (2019). SWI/SNF Component BAF250a Coordinates OCT4 and WNT Signaling Pathway to Control Cardiac Lineage Differentiation. Front. Cell Dev. Biol. *7*, 358.

Leung, C.Y., and Zernicka-Goetz, M. (2015). Mapping the journey from totipotency to lineage specification in the mouse embryo. Curr. Opin. Genet. Dev. *34*, 71–76.

Li, Z., and Rana, T.M. (2012). A kinase inhibitor screen identifies small-molecule enhancers of reprogramming and iPS cell generation. Nat. Commun. *3*, 1085.

Li, D., Liu, J., Yang, X., Zhou, C., Guo, J., Wu, C., Qin, Y., Guo, L., He, J., Yu, S., et al. (2017). Chromatin Accessibility Dynamics during iPSC Reprogramming. Cell Stem Cell *21*, 819-833.e6.

Li, Y., Moretto-Zita, M., Soncin, F., Wakeland, A., Wolfe, L., Leon-Garcia, S., Pandian, R., Pizzo, D., Cui, L., Nazor, K., et al. (2013). BMP4-directed trophoblast differentiation of human embryonic stem cells is mediated through a ΔNp63+ cytotrophoblast stem cell state. Development *140*, 3965–3976.

Liao, M.-C., Muratore, C.R., Gierahn, T.M., Sullivan, S.E., Srikanth, P., De Jager, P.L., Love, J.C., and Young-Pearse, T.L. (2016). Single-Cell Detection of Secreted Aβ and sAPPα from Human IPSC-Derived Neurons and Astrocytes. J. Neurosci. *36*, 1730–1746.

Lim, C.K., Efthymios, M., Tan, W., Autio, M.I., Tiang, Z., Li, P.Y., and Foo, R.S.Y. (2021). Dimethyl sulfoxide (DMSO) enhances direct cardiac reprogramming by inhibiting the bromodomain of coactivators CBP/p300. J. Mol. Cell. Cardiol. *160*, 15–26.

Lin, T., and Wu, S. (2015). Reprogramming with Small Molecules instead of Exogenous Transcription Factors. Stem Cells Int. *2015*, 794632.

Lipinski, M., Muñoz-Viana, R., del Blanco, B., Marquez-Galera, A., Medrano-Relinque, J., Caramés, J.M., Szczepankiewicz, A.A., Fernandez-Albert, J., Navarrón, C.M., Olivares, R., et al. (2020). KAT3-dependent acetylation of cell type-specific genes maintains neuronal identity in the adult mouse brain. Nat. Commun. *11*, 2588.

Liu, Q., Jiang, C., Xu, J., Zhao, M.-T., Van Bortle, K., Cheng, X., Wang, G., Chang, H.Y., Wu, J.C., and Snyder, M.P. (2017). Genome-Wide Temporal Profiling of Transcriptome and Open Chromatin of Early Cardiomyocyte Differentiation Derived From hiPSCs and hESCs. Circ. Res. *121*, 376–391.

Liu, Y., Wang, L., Han, R., Beier, U.H., Akimova, T., Bhatti, T., Xiao, H., Cole, P.A., Brindle, P.K., and Hancock, W.W. (2014). Two histone/protein acetyltransferases, CBP and p300, are indispensable for Foxp3+ T-regulatory cell development and function. Mol. Cell. Biol. *34*, 3993–4007.

Lloyd, J.T., and Glass, K.C. (2018). Biological function and histone recognition of family IV bromodomain-containing proteins. J. Cell. Physiol. *233*, 1877–1886.

Loo, C.-S., Gatchalian, J., Liang, Y., Leblanc, M., Xie, M., Ho, J., Venkatraghavan, B., Hargreaves, D.C., and Zheng, Y. (2020). A Genome-wide CRISPR Screen Reveals a Role for the Non-canonical Nucleosome-Remodeling BAF Complex in Foxp3 Expression and Regulatory T Cell Function. Immunity *53*, 143-157.e8.

Lowry, W.E., Richter, L., Yachechko, R., Pyle, A.D., Tchieu, J., Sridharan, R., Clark, A.T., and Plath, K. (2008). Generation of human induced pluripotent stem cells from dermal fibroblasts. Proc. Natl. Acad. Sci. U. S. A. *105*, 2883–2888.

Lynch, C.J., Bernad, R., Calvo, I., Nóbrega-Pereira, S., Ruiz, S., Ibarz, N., Martinez-Val, A., Graña-Castro, O., Gómez-López, G., Andrés-León, E., et al. (2018). The RNA Polymerase II Factor RPAP1 Is Critical for Mediator-Driven Transcription and Cell Identity. Cell Rep. *22*, 396–410.

Lynch, C.J., Bernad, R., Calvo, I., and Serrano, M. (2020). Manipulating the Mediator complex to induce naïve pluripotency. Exp. Cell Res. *395*, 112215.

Lyssiotis, C.A., Foreman, R.K., Staerk, J., Garcia, M., Mathur, D., Markoulaki, S., Hanna, J., Lairson, L.L., Charette, B.D., Bouchez, L.C., et al. (2009). Reprogramming of murine fibroblasts to induced pluripotent stem cells with chemical complementation of Klf4. Proc. Natl. Acad. Sci. U. S. A. *106*, 8912–8917.

Magli, A., Baik, J., Mills, L.J., Kwak, I.-Y., Dillon, B.S., Mondragon Gonzalez, R.,

Stafford, D.A., Swanson, S.A., Stewart, R., Thomson, J.A., et al. (2019). Time-dependent Pax3-mediated chromatin remodeling and cooperation with Six4 and Tead2 specify the skeletal myogenic lineage in developing mesoderm. PLoS Biol. *17*, e3000153.

Maherali, N., and Hochedlinger, K. (2008). Induced pluripotency of mouse and human somatic cells. Cold Spring Harb. Symp. Quant. Biol. *73*, 157–162.

Mak, C.C.Y., Doherty, D., Lin, A.E., Vegas, N., Cho, M.T., Viot, G., Dimartino, C., Weisfeld-Adams, J.D., Lessel, D., Joss, S., et al. (2020). MN1 C-terminal truncation syndrome is a novel neurodevelopmental and craniofacial disorder with partial rhombencephalosynapsis. Brain *143*, 55–68.

Malik, N., and Rao, M.S. (2013). A review of the methods for human iPSC derivation. Methods Mol. Biol. *997*, 23–33.

Mandai, M., Watanabe, A., Kurimoto, Y., Hirami, Y., Morinaga, C., Daimon, T., Fujihara, M., Akimaru, H., Sakai, N., Shibata, Y., et al. (2017). Autologous Induced Stem-Cell-Derived Retinal Cells for Macular Degeneration. N. Engl. J. Med. *376*, 1038–1046.

Mansour, A.A., Gafni, O., Weinberger, L., Zviran, A., Ayyash, M., Rais, Y., Krupalnik, V., Zerbib, M., Amann-Zalcenstein, D., Maza, I., et al. (2012). The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. Nature *488*, 409–413.

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc. Natl. Acad. Sci. U. S. A. *78*, 7634–7638.

Martire, S., Nguyen, J., Sundaresan, A., and Banaszynski, L.A. (2020). Differential contribution of p300 and CBP to regulatory element acetylation in mESCs. BMC Mol. Cell Biol. *21*, 55.

Mashtalir, N., D'Avino, A.R., Michel, B.C., Luo, J., Pan, J., Otto, J.E., Zullow, H.J., McKenzie, Z.M., Kubiak, R.L., St Pierre, R., et al. (2018). Modular Organization and Assembly of SWI/SNF Family Chromatin Remodeling Complexes. Cell *175*, 1272–1288.

Masliah-Planchon, J., Bièche, I., Guinebretière, J.-M., Bourdeaut, F., and Delattre, O. (2015). SWI/SNF chromatin remodeling and human malignancies. Annu. Rev. Pathol. *10*, 145–171.

McNeish, J., Gardner, J.P., Wainger, B.J., Woolf, C.J., and Eggan, K. (2015). From Dish to Bedside: Lessons Learned While Translating Findings from a Stem Cell Model of Disease to a Clinical Trial. Cell Stem Cell *17*, 8–10.

Meester-Smoor, M.A., Janssen, M.J.F.W., Grosveld, G.C., de Klein, A., van IJcken,

W.F.J., Douben, H., and Zwarthoff, E.C. (2008). MN1 affects expression of genes involved in hematopoiesis and can enhance as well as inhibit RAR/RXR-induced gene expression. Carcinogenesis *29*, 2025–2034.

Mellis, I.A., Edelstein, H.I., Truitt, R., Beck, L.E., Symmons, O., Goyal, Y., Dunagin, M.C., Saldana, R.A.L., Shah, P.P., Yang, W., et al. (2020). Responsiveness to perturbations is a hallmark of transcription factors that maintain cell identity. BioRxiv 2020.06.11.147207.

Michel, B.C., D'Avino, A.R., Cassel, S.H., Mashtalir, N., McKenzie, Z.M., McBride, M.J., Valencia, A.M., Zhou, Q., Bocker, M., Soares, L.M.M., et al. (2018). A non-canonical SWI/SNF complex is a synthetic lethal target in cancers driven by BAF complex perturbation. Nat. Cell Biol. *20*, 1410–1420.

Michlits, G., Hubmann, M., Wu, S.H., Vainorius, G., Budusan, E., Zhuk, S., Burkard, T.R., Novatchkova, M., Aichinger, M., Lu, Y., et al. (2017). CRISPR-UMI: Single-cell lineage tracing of pooled CRISPR-Cas9 screens. Nat. Methods *14*, 1191–1197.

Miles, D.C., de Vries, N.A., Gisler, S., Lieftink, C., Akhtar, W., Gogola, E., Pawlitzky, I., Hulsman, D., Tanger, E., Koppens, M., et al. (2017). TRIM28 Is an Epigenetic Barrier to Induced Pluripotent Stem Cell Reprogramming. Stem Cells *35*, 147–157.

Miller, J.L., and Grant, P.A. (2013). The role of DNA methylation and histone modifications in transcriptional regulation in humans. Subcell. Biochem. *61*, 289–317.

Minkovsky, A., Patel, S., and Plath, K. (2012). Concise review: Pluripotency and the transcriptional inactivation of the female Mammalian X chromosome. Stem Cells *30*, 48–54.

Molè, M.A., Weberling, A., and Zernicka-Goetz, M. (2020). Comparative analysis of human and mouse development: From zygote to pre-gastrulation. Curr. Top. Dev. Biol. *136*, 113–138.

Moris, N., Anlas, K., van den Brink, S.C., Alemany, A., Schröder, J., Ghimire, S., Balayo, T., van Oudenaarden, A., and Martinez Arias, A. (2020). An in vitro model of early anteroposterior organization during human development. Nature *582*, 410–415.

Mossahebi-Mohammadi, M., Quan, M., Zhang, J.-S., and Li, X. (2020). FGF Signaling Pathway: A Key Regulator of Stem Cell Pluripotency . Front. Cell Dev. Biol. *8*, 79.

Musselman, C.A., Lalonde, M.-E., Côté, J., and Kutateladze, T.G. (2012). Perceiving the epigenetic landscape through histone readers. Nat. Struct. Mol. Biol. *19*, 1218–1227.

Naryshkin, N.A., Weetall, M., Dakka, A., Narasimhan, J., Zhao, X., Feng, Z., Ling, K.K.Y., Karp, G.M., Qi, H., Woll, M.G., et al. (2014). Motor neuron disease. SMN2

splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. Science *345*, 688–693.

Nashun, B., Hill, P.W.S., and Hajkova, P. (2015). Reprogramming of cell fate: epigenetic memory and the erasure of memories past. EMBO J. *34*, 1296–1308.

Nguyen, H.N., Byers, B., Cord, B., Shcheglovitov, A., Byrne, J., Gujar, P., Kee, K., Schüle, B., Dolmetsch, R.E., Langston, W., et al. (2011). LRRK2 mutant iPSC-derived DA neurons demonstrate increased susceptibility to oxidative stress. Cell Stem Cell *8*, 267–280.

Ning, B., Zhao, W., Qian, C., Liu, P., Li, Q., Li, W., and Wang, R.-F. (2017). USP26 functions as a negative regulator of cellular reprogramming by stabilising PRC1 complex components. Nat. Commun. *8*, 349.

Nishimura, K., Kato, T., Chen, C., Oinam, L., Shiomitsu, E., Ayakawa, D., Ohtaka, M., Fukuda, A., Nakanishi, M., and Hisatake, K. (2014). Manipulation of KLF4 expression generates iPSCs paused at successive stages of reprogramming. Stem Cell Reports *3*, 915–929.

Oh, Y., and Jang, J. (2019). Directed Differentiation of Pluripotent Stem Cells by Transcription Factors. Mol. Cells *42*, 200–209.

Ohtsuka, S., and Dalton, S. (2008). Molecular and biological properties of pluripotent embryonic stem cells. Gene Ther. *15*, 74–81.

Ohtsuka, S., Nakai-Futatsugi, Y., and Niwa, H. (2015). LIF signal in mouse embryonic stem cells. JAK-STAT *4*, e1086520–e1086520.

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., et al. (2012). Chromatin-modifying enzymes as modulators of reprogramming. Nature *483*, 598–602.

Panamarova, M., Cox, A., Wicher, K.B., Butler, R., Bulgakova, N., Jeon, S., Rosen, B., Seong, R.H., Skarnes, W., Crabtree, G., et al. (2016). The BAF chromatin remodelling complex is an epigenetic regulator of lineage specification in the early mouse embryo. Development *143*, 1271–1283.

Panigrahi, A., and O'Malley, B.W. (2021). Mechanisms of enhancer action: the known and the unknown. Genome Biol. *22*, 108.

Papp, B., and Plath, K. (2011). Reprogramming to pluripotency: stepwise resetting of the epigenetic landscape. Cell Res. *21*, 486–501.

Park, D.E., Cheng, J., McGrath, J.P., Lim, M.Y., Cushman, C., Swanson, S.K., Tillgren, M.L., Paulo, J.A., Gokhale, P.C., Florens, L., et al. (2020). Merkel cell polyomavirus

activates LSD1-mediated blockade of non-canonical BAF to regulate transformation and tumorigenesis. Nat. Cell Biol.

Park, I.-H., Lerou, P.H., Zhao, R., Huo, H., and Daley, G.Q. (2008a). Generation of human-induced pluripotent stem cells. Nat. Protoc. *3*, 1180–1186.

Park, I.-H., Zhao, R., West, J.A., Yabuuchi, A., Huo, H., Ince, T.A., Lerou, P.H., Lensch, M.W., and Daley, G.Q. (2008b). Reprogramming of human somatic cells to pluripotency with defined factors. Nature *451*, 141–146.

Pasque, V., Radzisheuskaya, A., Gillich, A., Halley-Stott, R.P., Panamarova, M., Zernicka-Goetz, M., Surani, M.A., and Silva, J.C.R. (2012). Histone variant macroH2A marks embryonic differentiation in vivo and acts as an epigenetic barrier to induced pluripotency. J. Cell Sci. *125*, 6094–6104.

Pei, D., Shu, X., Gassama-Diagne, A., and Thiery, J.P. (2019). Mesenchymal-epithelial transition in development and reprogramming. Nat. Cell Biol. *21*, 44–53.

Pennacchio, L.A., Bickmore, W., Dean, A., Nobrega, M.A., and Bejerano, G. (2013). Enhancers: five essential questions. Nat. Rev. Genet. *14*, 288–295.

Pfaff, N., Liebhaber, S., Möbus, S., Beh-Pajooh, A., Fiedler, J., Pfanne, A., Schambach, A., Thum, T., Cantz, T., and Moritz, T. (2017). Inhibition of miRNA-212/132 improves the reprogramming of fibroblasts into induced pluripotent stem cells by de-repressing important epigenetic remodelling factors. Stem Cell Res. *20*, 70–75.

Pfisterer, U., Kirkeby, A., Torper, O., Wood, J., Nelander, J., Dufour, A., Björklund, A., Lindvall, O., Jakobsson, J., and Parmar, M. (2011). Direct conversion of human fibroblasts to dopaminergic neurons. Proc. Natl. Acad. Sci. U. S. A. *108*, 10343–10348.

Picaud, S., Fedorov, O., Thanasopoulou, A., Leonards, K., Jones, K., Meier, J., Olzscha, H., Monteiro, O., Martin, S., Philpott, M., et al. (2015). Generation of a Selective Small Molecule Inhibitor of the CBP/p300 Bromodomain for Leukemia Therapy. Cancer Res. *75*, 5106–5119.

Płusa, B., and Piliszek, A. (2020). Common principles of early mammalian embryo self-organisation. Development *147*.

Qian, X., Kim, J.K., Tong, W., Villa-Diaz, L.G., and Krebsbach, P.H. (2016). DPPA5 Supports Pluripotency and Reprogramming by Regulating NANOG Turnover. Stem Cells *34*, 588–600.

Qin, H., Diaz, A., Blouin, L., Lebbink, R.J., Patena, W., Tanbun, P., LeProust, E.M., McManus, M.T., Song, J.S., and Ramalho-Santos, M. (2014). Systematic identification of barriers to human iPSC generation. Cell *158*, 449–461.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Rack, J.G.M., Lutter, T., Kjæreng Bjerga, G.E., Guder, C., Ehrhardt, C., Värv, S., Ziegler, M., and Aasland, R. (2014). The PHD finger of p300 Influences Its Ability to Acetylate Histone and Non-Histone Targets. J. Mol. Biol. *426*, 3960–3972.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44*, W160-5.

Rand, T.A., Sutou, K., Tanabe, K., Jeong, D., Nomura, M., Kitaoka, F., Tomoda, E., Narita, M., Nakamura, M., Nakamura, M., et al. (2018). MYC Releases Early Reprogrammed Human Cells from Proliferation Pause via Retinoblastoma Protein Inhibition. Cell Rep. *23*, 361–375.

Riedel, S.S., Haladyna, J.N., Bezzant, M., Stevens, B., Pollyea, D.A., Sinha, A.U., Armstrong, S.A., Wei, Q., Pollock, R.M., Daigle, S.R., et al. (2016). MLL1 and DOT1L cooperate with meningioma-1 to induce acute myeloid leukemia. J. Clin. Invest. *126*, 1438–1450.

Robinton, D.A., and Daley, G.Q. (2012). The promise of induced pluripotent stem cells in research and therapy. Nature *481*, 295–305.

Romito, A., and Cobellis, G. (2016). Pluripotent Stem Cells: Current Understanding and Future Directions. Stem Cells Int. *2016*, 9451492.

Samavarchi-Tehrani, P., Golipour, A., David, L., Sung, H., Beyer, T.A., Datti, A., Woltjen, K., Nagy, A., and Wrana, J.L. (2010). Functional Genomics Reveals a BMP-Driven Mesenchymal-to-Epithelial Transition in the Initiation of Somatic Cell Reprogramming. Cell Stem Cell *7*, 64–77.

Sevinç, K., Sevinç, G.G., Cavga, A.D., Philpott, M., Kelekçi, S., Can, H., Cribbs, A.P., Ayar, E.S., Arabacı, D.H., Dunford, J.E., et al. (2021). BRD9-containing non-canonical BAF complexes safeguard cell identity and prevent reprogramming. BioRxiv 2021.05.27.445940.

Shao, Z., Yao, C., Khodadadi-Jamayran, A., Xu, W., Townes, T.M., Crowley, M.R., and Hu, K. (2016). Reprogramming by De-bookmarking the Somatic Transcriptional Program through Targeting of BET Bromodomains. Cell Rep. *16*, 3138–3145.

Shi, Y., Desponts, C., Do, J.T., Hahm, H.S., Schöler, H.R., and Ding, S. (2008). Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. Cell Stem Cell *3*, 568–574.

Shi, Y., Inoue, H., Wu, J.C., and Yamanaka, S. (2017). Induced pluripotent stem cell technology: a decade of progress. Nat. Rev. Drug Discov. *16*, 115–130.

Shimizu, T., Ueda, J., Ho, J.C., Iwasaki, K., Poellinger, L., Harada, I., and Sawada, Y. (2012). Dual inhibition of Src and GSK3 maintains mouse embryonic stem cells, whose differentiation is mechanically regulated by Src signaling. Stem Cells *30*, 1394–1404.

Sims, D., Ilott, N.E., Sansom, S.N., Sudbery, I.M., Johnson, J.S., Fawcett, K.A., Berlanga-Taylor, A.J., Luna-Valero, S., Ponting, C.P., and Heger, A. (2014). CGAT: computational genomics analysis toolkit. Bioinformatics *30*, 1290–1291.

Singhal, N., Graumann, J., Wu, G., Araúzo-Bravo, M.J., Han, D.W., Greber, B., Gentile, L., Mann, M., and Schöler, H.R. (2010). Chromatin-Remodeling Components of the BAF Complex Facilitate Reprogramming. Cell *141*, 943–955.

Smith, K.P., Luong, M.X., and Stein, G.S. (2009). Pluripotency: Toward a gold standard for human ES and iPS cells. J. Cell. Physiol. *220*, 21–29.

Soldner, F., Stelzer, Y., Shivalila, C.S., Abraham, B.J., Latourelle, J.C., Barrasa, M.I., Goldmann, J., Myers, R.H., Young, R.A., and Jaenisch, R. (2016). Parkinson-associated risk variant in distal enhancer of α-synuclein modulates target gene expression. Nature *533*, 95–99.

Sommer, C.A., Stadtfeld, M., Murphy, G.J., Hochedlinger, K., Kotton, D.N., and Mostoslavsky, G. (2009). Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. Stem Cells *27*, 543–549.

Son, E.Y., Ichida, J.K., Wainger, B.J., Toma, J.S., Rafuse, V.F., Woolf, C.J., and Eggan, K. (2011). Conversion of mouse and human fibroblasts into functional spinal motor neurons. Cell Stem Cell *9*, 205–218.

Soufi, A., Donahue, G., and Zaret, K.S. (2012). Facilitators and Impediments of the Pluripotency Reprogramming Factors' Initial Engagement with the Genome. Cell *151*, 994–1004.

Spicuglia, S., and Vanhille, L. (2012). Chromatin signatures of active enhancers. Nucleus *3*, 126–131.

Sridharan, R., Gonzales-Cope, M., Chronis, C., Bonora, G., McKee, R., Huang, C., Patel, S., Lopez, D., Mishra, N., Pellegrini, M., et al. (2013). Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1γ in reprogramming to pluripotency. Nat. Cell Biol. *15*, 872–882.

Stadtfeld, M., and Hochedlinger, K. (2010). Induced pluripotency: history, mechanisms, and applications. Genes Dev. *24*, 2239–2263.

Staerk, J., Lyssiotis, C.A., Medeiro, L.A., Bollong, M., Foreman, R.K., Zhu, S., Garcia, M., Gao, Q., Bouchez, L.C., Lairson, L.L., et al. (2011). Pan-Src Family Kinase Inhibitors Replace Sox2 during the Direct Reprogramming of Somatic Cells. Angew. Chemie Int. Ed. *50*, 5734–5736.

Strelchenko, N., Verlinsky, O., Kukharenko, V., and Verlinsky, Y. (2004). Morula-derived human embryonic stem cells. Reprod. Biomed. Online *9*, 623–629.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. *102*, 15545 LP – 15550.

Sun, H., Liu, J., Zhang, J., Shen, W., Huang, H., Xu, C., Dai, H., Wu, J., and Shi, Y. (2007). Solution structure of BRD7 bromodomain and its interaction with acetylated peptides from histone H3 and H4. Biochem. Biophys. Res. Commun. *358*, 435–441.

Syed, K.M., Joseph, S., Mukherjee, A., Majumder, A., Teixeira, J.M., Dutta, D., and Pillai, M.R. (2016). Histone chaperone APLF regulates induction of pluripotency in murine fibroblasts. J. Cell Sci. *129*, 4576–4591.

Szabo, Q., Bantignies, F., and Cavalli, G. (2019). Principles of genome folding into topologically associating domains. Sci. Adv. *5*, eaaw1668.

Tada, M., Tada, T., Lefebvre, L., Barton, S.C., and Surani, M.A. (1997). Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells. EMBO J. *16*, 6510–6520.

Tada, M., Takahama, Y., Abe, K., Nakatsuji, N., and Tada, T. (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. Curr. Biol. *11*, 1553–1558.

Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell *126*, 663–676.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. Cell *131*, 861–872.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Sasaki, A., Yamamoto, M., Nakamura, M., Sutou, K., Osafune, K., and Yamanaka, S. (2014). Induction of pluripotency in human somatic cells via a transient state resembling primitive streak-like mesendoderm. Nat. Commun. *5*, 3678.

Takahashi, S., Kobayashi, S., and Hiratani, I. (2018). Epigenetic differences between

naïve and primed pluripotent stem cells. Cell. Mol. Life Sci. *75*, 1191–1203.

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W., et al. (2014). Resetting transcription factor control circuitry toward ground-state pluripotency in human. Cell *158*, 1254–1269.

Tamaru, H. (2010). Confining euchromatin/heterochromatin territory: jumonji crosses the line. Genes Dev. *24*, 1465–1478.

Tanaka, Y., Naruse, I., Hongo, T., Xu, M.-J., Nakahata, T., Maekawa, T., and Ishii, S. (2000). Extensive brain hemorrhage and embryonic lethality in a mouse null mutant of CREB-binding protein. Mech. Dev. *95*, 133–145.

Tao, J., Zhang, Y., Zuo, X., Hong, R., Li, H., Liu, X., Huang, W., Cao, Z., and Zhang, Y. (2017). DOT1L inhibitor improves early development of porcine somatic cell nuclear transfer embryos. PLoS One *12*, e0179436.

Terzi Cizmecioglu, N., Huang, J., Keskin, E.G., Wang, X., Esen, I., Chen, F., and Orkin, S.H. (2020). ARID4B is critical for mouse embryonic stem cell differentiation towards mesoderm and endoderm, linking epigenetics to pluripotency exit. J. Biol. Chem. *295*, 17738–17751.

Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan, Z.P., Maetzel, D., Ganz, K., Shi, L., et al. (2014). Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. Cell Stem Cell *15*, 471–487.

Thiery, J.P., Acloque, H., Huang, R.Y.J., and Nieto, M.A. (2009). Epithelial-mesenchymal transitions in development and disease. Cell *139*, 871–890.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic Stem Cell Lines Derived from Human Blastocysts. Science (80-.). *282*, 1145 LP – 1147.

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., et al. (2016). RNAi Reveals Phase-Specific Global Regulators of Human Somatic Cell Reprogramming. Cell Rep. *15*, 2597–2607.

Tomofusa, F., H., K.L., Fayçal, B., Trushar, J., Jan, van D., and K., B.P. (2009). Histone Acetyltransferase CBP Is Vital To Demarcate Conventional and Innate CD8+ T-Cell Development. Mol. Cell. Biol. *29*, 3894–3904.

Tran, K.A., Pietrzak, S.J., Zaidan, N.Z., Siahpirani, A.F., McCalla, S.G., Zhou, A.S., Iyer, G., Roy, S., and Sridharan, R. (2019). Defining Reprogramming Checkpoints from Single-Cell Analyses of Induced Pluripotency. Cell Rep. *27*, 1726-1741.e5.

Uğurlu-Çimen, D., Odluyurt, D., Sevinç, K., Özkan-Küçük, N.E., Özçimen, B., Demirtaş, D., Enüstün, E., Aztekin, C., Philpott, M., Oppermann, U., et al. (2021). AF10 (MLLT10) prevents somatic cell reprogramming through regulation of DOT1L-mediated H3K79 methylation. Epigenetics Chromatin *14*, 32.

Ursu, A., Illich, D.J., Takemoto, Y., Porfetye, A.T., Zhang, M., Brockmeyer, A., Janning, P., Watanabe, N., Osada, H., Vetter, I.R., et al. (2016). Epiblastin A Induces Reprogramming of Epiblast Stem Cells Into Embryonic Stem Cells by Inhibition of Casein Kinase 1. Cell Chem. Biol. *23*, 494–507.

Vallier, L., Alexander, M., and Pedersen, R.A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. J. Cell Sci. *118*, 4495–4509.

Velychko, S., Adachi, K., Kim, K.-P., Hou, Y., MacCarthy, C.M., Wu, G., and Schöler, H.R. (2019). Excluding Oct4 from Yamanaka Cocktail Unleashes the Developmental Potential of iPSCs. Cell Stem Cell *25*, 737-753.e4.

Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Südhof, T.C., and Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. Nature *463*, 1035–1041.

Voigt, P., Tee, W.-W., and Reinberg, D. (2013). A double take on bivalent promoters. Genes Dev. *27*, 1318–1338.

Waddington, C.H. (1956). Genetic Assimilation of the Bithorax Phenotype. Evolution (N. Y). *10*, 1–13.

Wagner, E.J., and Carpenter, P.B. (2012). Understanding the language of Lys36 methylation at histone H3. Nat. Rev. Mol. Cell Biol. *13*, 115–126.

Wang, F., Marshall, C.B., and Ikura, M. (2013). Transcriptional/epigenetic regulator CBP/p300 in tumorigenesis: structural and functional versatility in target recognition. Cell. Mol. Life Sci. *70*, 3989–4008.

Wang, J., Jia, S.T., and Jia, S. (2016). New Insights into the Regulation of Heterochromatin. Trends Genet. *32*, 284–294.

Wang, L., Su, Y., Huang, C., Yin, Y., Chu, A., Knupp, A., and Tang, Y. (2019a). NANOG and LIN28 dramatically improve human cell reprogramming by modulating LIN41 and canonical WNT activities. Biol. Open *8*.

Wang, M., Ling, W., Xiong, C., Xie, D., Chu, X., Li, Y., Qiu, X., Li, Y., and Xiao, X. (2019b). Potential Strategies for Cardiac Diseases: Lineage Reprogramming of Somatic Cells into Induced Cardiomyocytes. Cell. Reprogram. *21*, 63–77.

Wang, R., He, Y., Robinson, V., Yang, Z., Hessler, P., Lasko, L.M., Lu, X., Bhathena, A., Lai, A., Uziel, T., et al. (2018a). Targeting Lineage-specific MITF Pathway in Human Melanoma Cell Lines by A-485, the Selective Small-molecule Inhibitor of p300/CBP. Mol. Cancer Ther. *17*, 2543 LP – 2550.

Wang, T., Chen, K., Zeng, X., Yang, J., Wu, Y., Shi, X., Qin, B., Zeng, L., Esteban, M.A., Pan, G., et al. (2011). The histone demethylases Jhdm1a/1b enhance somatic cell reprogramming in a vitamin-C-dependent manner. Cell Stem Cell *9*, 575–587.

Wang, Y., Zhao, C., Hou, Z., Yang, Y., Bi, Y., Wang, H., Zhang, Y., and Gao, S. (2018b). Unique molecular events during reprogramming of human somatic cells to induced pluripotent stem cells (iPSCs) at naïve state. Elife *7*.

Wang, Y., Chen, S., Jiang, Q., Deng, J., Cheng, F., Lin, Y., Cheng, L., Ye, Y., Chen, X., Yao, Y., et al. (2020). TFAP2C facilitates somatic cell reprogramming by inhibiting c-Myc-dependent apoptosis and promoting mesenchymal-to-epithelial transition. Cell Death Dis. *11*, 482.

Wang, Z., Wang, P., Li, Y., Peng, H., Zhu, Y., Mohandas, N., and Liu, J. (2021). Interplay between cofactors and transcription factors in hematopoiesis and hematological malignancies. Signal Transduct. Target. Ther. *6*, 24.

Ware, C.B., Nelson, A.M., Mecham, B., Hesson, J., Zhou, W., Jonlin, E.C., Jimenez-Caliani, A.J., Deng, X., Cavanaugh, C., Cook, S., et al. (2014). Derivation of naive human embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. *111*, 4484–4489.

Wei, J., Antony, J., Meng, F., MacLean, P., Rhind, R., Laible, G., and Oback, B. (2017). KDM4B-mediated reduction of H3K9me3 and H3K36me3 levels improves somatic cell reprogramming into pluripotency. Sci. Rep. *7*, 7514.

Wei, Z., Gao, F., Kim, S., Yang, H., Lyu, J., An, W., Wang, K., and Lu, W. (2013). Klf4 Organizes Long-Range Chromosomal Interactions with the Oct4 Locus in Reprogramming and Pluripotency. Cell Stem Cell *13*, 36–47.

Weinberger, L., Ayyash, M., Novershtern, N., and Hanna, J.H. (2016). Dynamic stem cell states: naive to primed pluripotency in rodents and humans. Nat. Rev. Mol. Cell Biol. *17*, 155–169.

Wilmut, I., Schnieke, A.E., McWhir, J., Kind, A.J., and Campbell, K.H. (1997). Viable offspring derived from fetal and adult mammalian cells. Nature *385*, 810–813.

Witte, S., Bradley, A., Enright, A.J., and Muljo, S.A. (2015). High-density P300 enhancers control cell state transitions. BMC Genomics *16*, 903.

Wolf, L., Harrison, W., Huang, J., Xie, Q., Xiao, N., Sun, J., Kong, L., Lachke, S.A.,

Kuracha, M.R., Govindarajan, V., et al. (2013). Histone posttranslational modifications and cell fate determination: lens induction requires the lysine acetyltransferases CBP and p300. Nucleic Acids Res. *41*, 10199–10214.

Woo, H., Dam Ha, S., Lee, S.B., Buratowski, S., and Kim, T. (2017). Modulation of gene expression dynamics by co-transcriptional histone methylations. Exp. Mol. Med. *49*, e326–e326.

Wozniak, G.G., and Strahl, B.D. (2014). Hitting the "mark": interpreting lysine methylation in the context of active transcription. Biochim. Biophys. Acta *1839*, 1353–1361.

Yagi, T., Ito, D., Okada, Y., Akamatsu, W., Nihei, Y., Yoshizaki, T., Yamanaka, S., Okano, H., and Suzuki, N. (2011). Modeling familial Alzheimer's disease with induced pluripotent stem cells. Hum. Mol. Genet. *20*, 4530–4539.

Yang, C.-S., Lopez, C.G., and Rana, T.M. (2011). Discovery of nonsteroidal anti-inflammatory drug and anticancer drug enhancing reprogramming and induced pluripotent stem cell generation. Stem Cells *29*, 1528–1536.

Yang, C.S., Chang, K.Y., and Rana, T.M. (2014). Genome-wide Functional Analysis Reveals Factors Needed at the Transition Steps of Induced Reprogramming. Cell Rep. *8*, 327–337.

Yao, T.P., Oh, S.P., Fuchs, M., Zhou, N.D., Ch'ng, L.E., Newsome, D., Bronson, R.T., Li, E., Livingston, D.M., and Eckner, R. (1998). Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. Cell *93*, 361–372.

Yilmaz, A., Braverman-Gross, C., Bialer-Tsypin, A., Peretz, M., and Benvenisty, N. (2020). Mapping Gene Circuits Essential for Germ Layer Differentiation via Loss-of-Function Screens in Haploid Human Embryonic Stem Cells. Cell Stem Cell *27*, 679-691.e6.

You, J.S., De Carvalho, D.D., Dai, C., Liu, M., Pandiyan, K., Zhou, X.J., Liang, G., and Jones, P.A. (2013). SNF5 Is an Essential Executor of Epigenetic Regulation during Differentiation. PLOS Genet. *9*, e1003459.

Zhang, H., Wang, X., Li, J., Shi, R., and Ye, Y. (2021a). BAF Complex in Embryonic Stem Cells and Early Embryonic Development. Stem Cells Int. *2021*, 6668866.

Zhang, L., Sheng, C., Zhou, F., Zhu, K., Wang, S., Liu, Q., Yuan, M., Xu, Z., Liu, Y., Lu, J., et al. (2021b). CBP/p300 HAT maintains the gene network critical for β cell identity and functional maturity. Cell Death Dis. *12*, 476.

Zhang, T., Cooper, S., and Brockdorff, N. (2015). The interplay of histone modifications – writers that read. EMBO Rep. *16*, 1467–1481.

Zhang, W., Chronis, C., Chen, X., Zhang, H., Spalinskas, R., Pardo, M., Chen, L., Wu, G., Zhu, Z., Yu, Y., et al. (2019). The BAF and PRC2 Complex Subunits Dpf2 and Eed Antagonistically Converge on Tbx3 to Control ESC Differentiation. Cell Stem Cell *24*, 138-152.e8.

Zhang, X., Li, B., Li, W., Ma, L., Zheng, D., Li, L., Yang, W., Chu, M., Chen, W., Mailman, R.B., et al. (2014). Transcriptional repression by the BRG1-SWI/SNF complex affects the pluripotency of human embryonic stem cells. Stem Cell Reports *3*, 460–474.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based Analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

Zhao, Y., Zhao, T., Guan, J., Zhang, X., Fu, Y., Ye, J., Zhu, J., Meng, G., Ge, J., Yang, S., et al. (2015). A XEN-like State Bridges Somatic Cells to Pluripotency during Chemical Reprogramming. Cell *163*, 1678–1691.

Zhu, M., and Zernicka-Goetz, M. (2020). Principles of Self-Organization of the Mammalian Embryo. Cell *183*, 1467–1478.

Zhu, H., Swami, S., Yang, P., Shapiro, F., and Wu, J.Y. (2020). Direct Reprogramming of Mouse Fibroblasts into Functional Osteoblasts. J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. *35*, 698–713.

Zhu, S., Li, W., Zhou, H., Wei, W., Ambasudhan, R., Lin, T., Kim, J., Zhang, K., and Ding, S. (2010). Reprogramming of human primary somatic cells by OCT4 and chemical compounds. Cell Stem Cell *7*, 651–655.

Zviran, A., Mor, N., Rais, Y., Gingold, H., Peles, S., Chomsky, E., Viukov, S., Buenrostro, J.D., Scognamiglio, R., Weinberger, L., et al. (2019). Deterministic Somatic Cell Reprogramming Involves Continuous Transcriptional Changes Governed by Myc and Epigenetic-Driven Modules. Cell Stem Cell *24*, 328-341.e9.

Appendix A: BRD9 Chromatin Immunoprecipitation (ChIP) Trials

Several ChIP methods to map genome-wide distribution of BRD9 protein were performed on human fibroblasts and human iPSCs. 9×10^6 cells were used per chromatin immunoprecipitation experiments. ChIP assays were performed with ChIP-IT High-sensitivity kit (Active Motif, catalog no. 53040) with following modifications.

1. From the beginning of the protocol, cells were frozen at most once if only necessary.

2. Cells were fixated in complete cell fixation solution that contains 1% formaldehyde for 15 minutes at room temperature. Additionally, double crosslinking with disuccinimidyl glutarate (Cayman, cat no. 20646) at the final concentration of 1mM for 45 minutes before 1% formaldehyde fixation was tried, too.

3. During the cell lysis step, cells were homogenized using a tight fitting pestle for 40 strokes as 30 strokes did not release all nuclei.

4. Cells were resuspended in ChIP buffer at a ratio of 1×10^7 cells/mL and 250 µL (2.5 million cells) was carefully added at each 1.5 mL sonicator tube by avoiding any bubble formation.

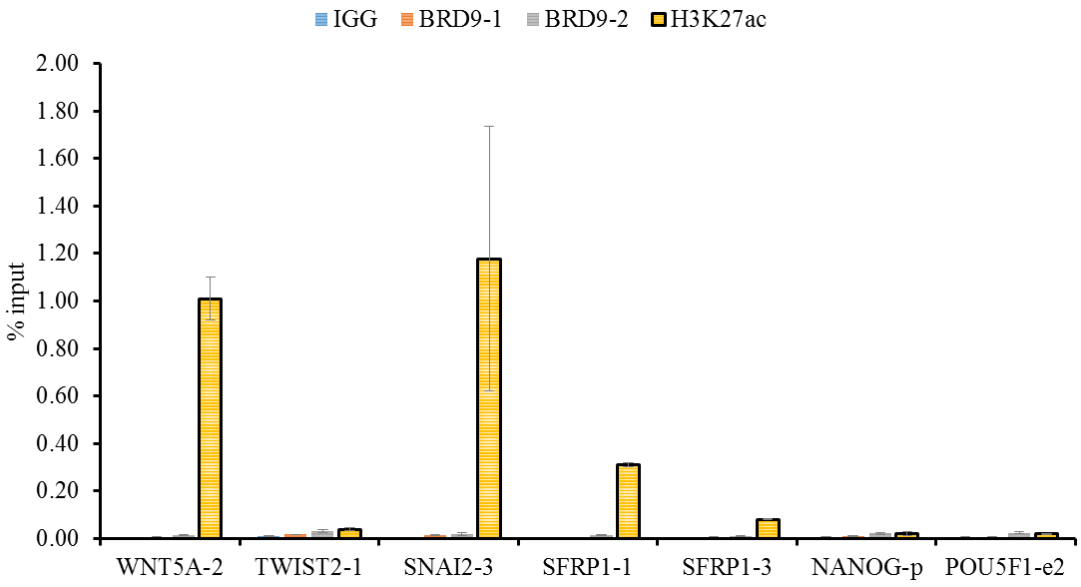
5. The fresh water at Biorupter Pico (Diagenode) sonicator was cooled down to 4^0C and all 6 slots available for 1.5 mL tubes were filled with tubes carrying the same volume of liquid even less slots were going to be used.

6. Chromatin was fragmented to 200–800 base pairs by sonication (5 × 10 cycles, 30 second on, 30 seconds off, highest amplitude) using Biorupter Pico (Diagenode) sonicator. After each cycle, suspension was vortexed and centrifuged briefly. This settings were optimized to yield the target chromatin fragmentation range but they depend on cell type, sonicator type, centrifuge tube and buffer.

7. 10% of ChIP reaction which is 9×10^5 cells were used as input. During input preparation and agarose gel visualization to ensure correct chromatin fragmentation, the rest of fragmented chromatin was kept on ice to avoid freeze-thaw cycle.

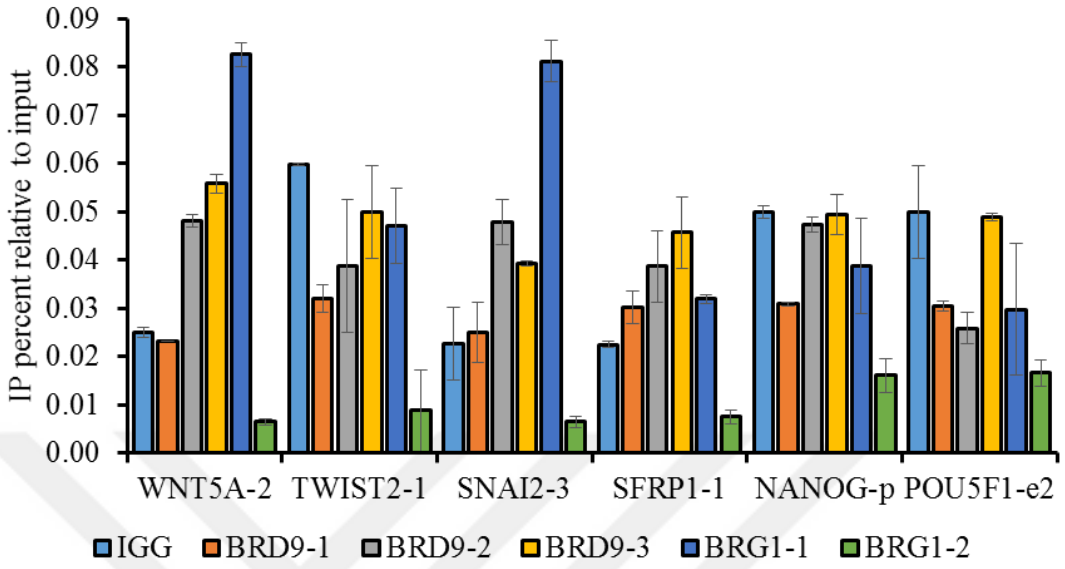
8. After ensuring chromatin fragmentation was as expected, each immunoprecipitation was performed for fragmented chromatin corresponding to 9×10^6 cells. 10 µg BRD9 antibodies (Active Motif, cat no. 61537 _ BRD9-1, Bethyl Laboratory, cat no. A303-781A _ BRD9-2 and abcam, cat no. ab137245 _ BRD9-3) and BRG1 antibodies (Santa Cruz, cat no. sc-17796 _ BRG1-1 and abcam cat no. ab110641 _ BRG1-2) were used separately. IGG (abcam, cat no. ab37415) and H3K27ac (Active Motif, cat no. 39133) were used as negative and positive control pull downs, respectively.

9. Although I observed around 50-fold enrichment for H3K27ac ChIP at positive regions versus negative regions, I failed to observe this enrichment for BRD9 and BRG1 ChIPs. That's why; single (Appendix A - Figure 1) or double cross-linking (Appendix A – Figure 2) human fibroblasts and antibodies listed at the previous step failed to pulldown BRD9 and BRG1 bound DNAs. Similarly, single cross-linking human iPSCs did not yield to enrichments after pull-down with BRD9 antibodies (Appendix A - Figure 3).

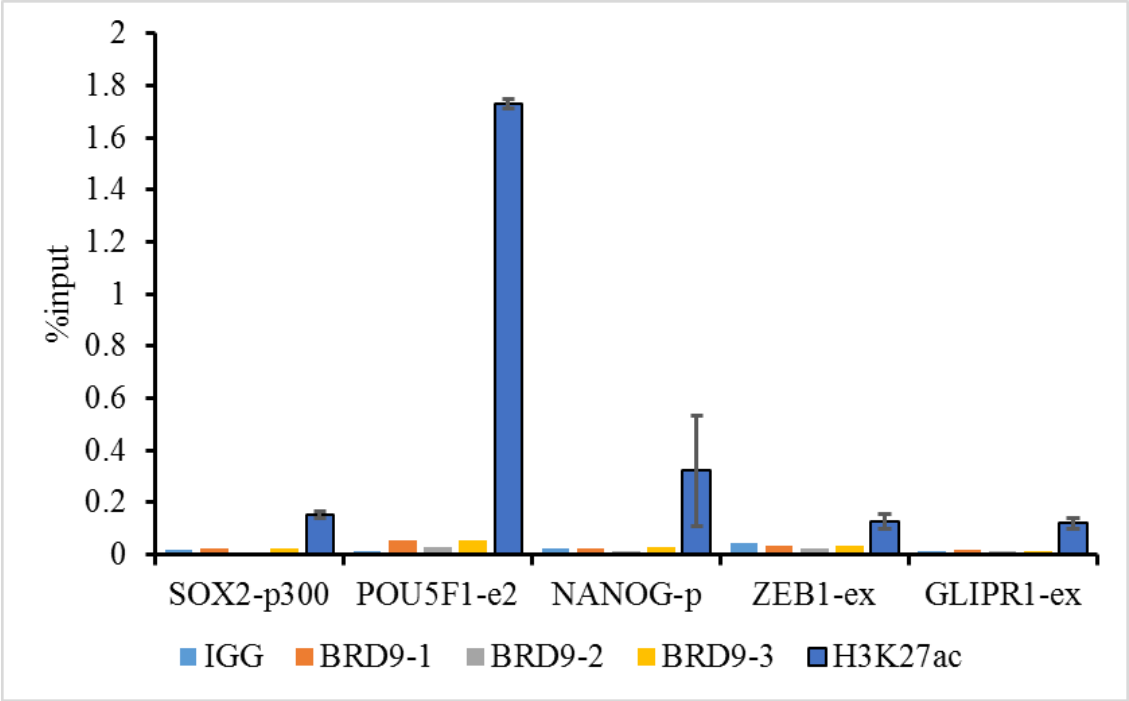


Appendix A – Figure 1. Percentage of input pulled down from single cross-linked human

fibroblasts with indicated antibodies.



Appendix A – Figure 2. Percentage of input pulled down from double cross-linked human fibroblasts with indicated antibodies.



Appendix A – Figure 3. Percentage of input pulled down from single cross-linked human iPSCs with indicated antibodies.