Mechanisms of pluripotent cell state transitions

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Daniela Mayer

aus Österreich

Basel, 2020

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Prof. Dr. Susan Gasser

Dr. Jörg Betschinger

Dr. Christa Bücker

Basel, den 19.11.2019

Prof. Dr. Martin Spiess

Dekan

TABLE OF CONTENTS

1. SUMMARY

Cell state transitions enable the differentiation of stem and progenitor cells into more mature and specialized cell types and are, thus, fundamental to the formation of multicellular organisms. Developmental progression is largely a unidirectional process. However, expression of reprogramming factors is sufficient to de-differentiate mature somatic cells, suggesting that cellular plasticity persists even in terminally differentiated cell types. Multiple signaling pathways, epigenetic regulators, metabolic sensing cascades and transcription factors (TFs) contribute to differentiation and de-differentiation. However, if reprogramming requires the reversion of naturally occurring developmental mechanisms remains unknown.

A suitable model system to study cell state transitions *in vitro* are lineage-related mouse embryonic stem cells (ESCs) and epiblast stem cells (EpiSCs) which are derivatives of the pre-implantation blastocyst and the post-implantation epiblast, respectively. Interconvertibility of ESCs and EpiSCs provides an experimental model to explore to which extent lineage progression and reprogramming overlap mechanistically.

In a collaborative project, I contributed to the characterization of a novel ESC differentiation pathway: in a genome-wide clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated 9 (Cas9) screen we identified multiple components of a conserved amino acid signaling pathway as crucial drivers of ESC progression. Mechanistically, the lysosome activity, the Ragulator protein complex, and the tumor-suppressor Folliculin (Flcn) enable the Rag GTPases C and D to bind and seclude the TF Tfe3 in the cytoplasm. Ectopic nuclear Tfe3 represses specific developmental and activates metabolic transcriptional programs which are associated with *in vivo* development. In collaboration with geneticists, we identified point mutations in a Tfe3 domain required for cytoplasmic inactivation as a potential cause of a human developmental disorder. This work reveals an instructive and biomedically relevant role for metabolic signaling in licensing embryonic cell fate transitions.

In my main PhD project, we aimed to identify cell state transition regulators which both are required for exit from the ESC state and inhibit acquisition of the induced pluripotent cell (iPSC) identity upon reprogramming of EpiSCs. We therefore performed a large-scale loss-of-function reprogramming screen in sensitized EpiSCs. Comparison with ESC differentiation screens revealed the constitutively expressed TF Zfp281 as a unique bidirectional regulator of cell state interconversion. We identified the histone methyltransferase Ehmt1 and the zinc finger TF Zic2 as differentiation-specific protein interaction partners of Zfp281 and showed that subtle chromatin binding changes of Zfp281 during ESC progression translate into activation of Ehmt1 and stabilization of Zic2 on promoters and enhancers. Genetic gain- and loss-of-function experiments confirmed a critical role of Ehmt1 and Zic2 downstream of Zfp281 both in driving exit from the ESC state, and in restricting reprogramming of EpiSCs. This study reveals that the cell type-invariant chromatin association of Zfp281 provides an interaction platform for remodeling the cis-regulatory network underlying cellular plasticity.

2. INTRODUCTION

2.1 Pluripotency in the embryo and in culture

Pluripotency describes the capacity of a cell to give rise to all somatic lineages and the germline. Pluripotency is a transient cellular feature during embryonic development and exists only for a few days *in vivo*. In the past 30 years, culture conditions to capture this property *in vitro* have been developed revealing that distinct pluripotent cell states can be maintained which recapitulate the establishment and progression of pluripotency in the embryo. As a model system, I used murine pluripotent stem cells and hence this introduction will focus on mouse pluripotency in the embryo and in culture.

2.1.1 Pluripotent cell states

Embryogenesis is the process by which a single cell, the zygote, gives rise to a fully developed organism, which is comprised of many different and highly specialized cell types. After oocyte fertilization by a mature sperm, the zygote undergoes several rounds of cleavage division ultimately giving rise to a spherical blastocyst which occurs at embryonic day (E) 3.5 (Figure **1**). At this developmental stage the embryo is structured into two lineages: the inner cell mass (ICM) and the trophectoderm (TE), which surrounds the ICM and later contributes in large to the placenta. Pluripotency emerges in the late pre-implantation blastocyst at E4.0-4.5 when the ICM segregates into the primitive endoderm (PrE), which will contribute to the yolk sac, and the pluripotent epiblast (Figure **1**). The latter is comprised of roughly 10-20 cells each harboring the potential to generate all cell types of the embryo proper: ectoderm, endoderm, mesoderm and the germline (Rossant & Tam, 2009). Epiblast cells at this developmental stages can be isolated and cultured *ex vivo* as ESCs (Evans & Kaufman, 1981; Martin, 1981) (Figure **1**). *In vitro*, ESCs self-renew to regenerate clones of themselves and can be propagated in theory indefinitely while retaining pluripotency: to differentiate into all germ layers and germ cells *in vitro*, and to generate whole animals when injected back into the pre-implantation epiblast (Bradley et al, 1984).

Upon embryo implantation at around E5.0-5.5, the epiblast transforms into a cup-shaped epithelium, the egg cylinder epiblast, and enters gastrulation at E6.5, including formation of the primitive streak (PS) (Rossant & Tam, 2009) (Figure **1**). Although morphologically different, the post-implantation egg cylinder epiblast cells, similar to the pre-implantation epiblast, can develop into all lineages of the embryo including germ cells (Lawson et al, 1991). However, at this embryonic stage ESCs can no longer be derived but instead EpiSCs can be isolated using distinct culture conditions (Brons et al, 2007; Tesar et al, 2007). EpiSCs retain certain features of the post-implantation epiblast such as the ability to differentiate into multiple somatic lineages *in vitro* and, upon transplantation into post- but not pre-implantation epiblasts, to contribute to embryonic development (Brons et al, 2007; Huang et al, 2012; Tesar et al, 2007). In contrast to ESCs, EpiSCs are not capable of inducing primordial germ cell formation, suggesting a restricted developmental potential (Hayashi et al, 2011).

The different pluripotent states of ESCs and EpiSCs corresponding to the E4.5 pre- and E5.5 post-implantation epiblast have been therefore termed “naïve” and “primed”, respectively (Nichols & Smith, 2009) (Figure **1**).

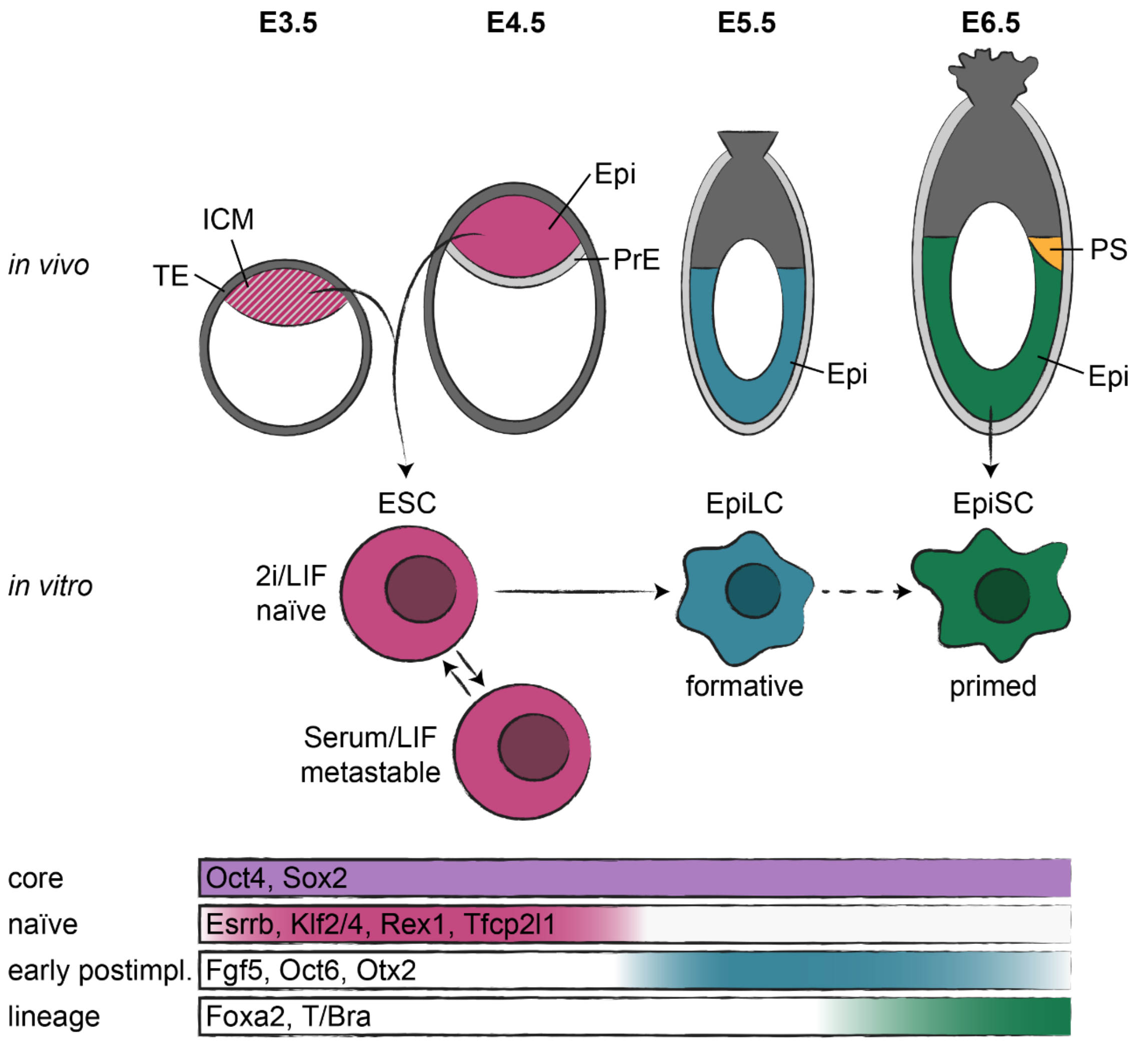


Figure 1: Pluripotency *in vivo* and *in vitro*.
A schematic overview showing the relationship between *in vivo* and *in vitro* pluripotent state progression. Pluripotent cells emerge in the inner cell mass (ICM) of the E3.5 blastocyst. Striped ICM indicates priming of epiblast (Epi) and primitive endoderm (PrE) identities which segregate into two distinct lineages in the E4.5 blastocyst. Upon implantation at around E5.0-5.5, the epiblast transforms into a cup-shaped epithelium and further initiates gastrulation, including formation of the primitive streak (PS), at E6.5. Conventionally, ESCs were cultured in Serum/LIF resulting in a heterogeneous cell population in a metastable state (pink-cyan gradient) which show partial transcriptional similarities to the post-implantation epiblast. ESCs cultured in chemically defined conditions, 2i/LIF (pink), maintain a more naïve pluripotent state and resemble most closely the E4.5 epiblast. Differentiation of naïve ESCs in defined conditions (Fgf2/ActivinA/knockout serum replacement, FAK) results in the transient appearance of epiblast like cells (EpiLCs) that most closely resemble the peri-implantation epiblast at E5.5, termed as formative pluripotency (cyan). In contrast, primed epiblast stem cells (EpiSCs, green) can be isolated from the E5.5-E8.0 post-implantation epiblast and resemble transcriptionally the anterior PS at E7.0. EpiSCs can be also derived from ESCs *in vitro* by continuous exposure to FA via an EpiLC state. Bottom graph depicts expression pattern of core pluripotency (purple), naïve pluripotency (pink), early post-implantation (cyan) and lineage (green) markers that are characteristic of distinct *in vivo* and *in vitro* pluripotent cell states. Top: extraembryonic lineages are depicted in gray; dark gray lineages are trophectoderm (TE)-derived and light gray lineages PrE-derived.

2.1.2 The naïve pluripotency gene regulatory network

TFs control gene expression by binding to specific deoxyribonucleic acid (DNA) sequences and recruiting various co-factors. ESC pluripotency is maintained by a highly interconnected gene regulatory network (GRN) of TFs (Martello & Smith, 2014). At its core are the POU domain TF Oct4 (Pou5f1) and the SRY box TF Sox2. Both are indispensable for self-renewal of ESCs and epiblast development *in vivo* (Avilion et al, 2003; Masui et al, 2007; Nichols et al, 1998; Niwa et al, 2000). Oct4 and Sox2 physically interact (Ambrosetti et al, 1997; Pardo et al, 2010; van den Berg et al, 2010), bind together on DNA at Oct/Sox elements (Ambrosetti et al, 2000; Chen et al, 2008), and positively regulate each other's transcription (Chew et al, 2005). Surprisingly, the expression of Oct4 and Sox2 is not restricted to naïve pluripotent cells, but is retained in the post-implantation epiblast and in EpiSCs (Avilion et al, 2003; Brons et al, 2007; Rosner et al, 1990; Scholer et al, 1990; Tesar et al, 2007), where they have subsequent roles in further lineage specification (Mulas et al, 2018; Thomson et al, 2011). In contrast, a set of naïve TFs are specifically expressed in ESCs and the pre-implantation epiblast and are rapidly downregulated in EpiSCs and upon implantation. These include Esrrb, Klf2, Klf4, Nanog and Tfcp2l1, which have all been demonstrate to functionally contribute to ESC self-renewal (Chambers et al, 2003; Festuccia et al, 2012; Hall et al, 2009; Martello et al, 2013; Martello et al, 2012; Niwa et al, 2009; Ye et al, 2013). In contrast, the TF Rex1 (Zfp42) follows similar expression kinetics, but it is not required for naïve pluripotency *in vitro* or *in vivo* (Masui et al, 2008; Pelton et al, 2002). Of note, Nanog although being transcriptionally downregulated upon implantation is re-activated in EpiSCs and the post-implantation epiblast, suggesting additional regulatory mechanisms (Brons et al, 2007; Hart et al, 2004; Tesar et al, 2007). The naïve pluripotent TFs are interconnected with one other and with core factors Oct4 and Sox2 (Chen et al, 2008; Marson et al, 2008) forming a self-reinforcing GRN to stabilize the ESC state (Dunn et al, 2014; Kim et al, 2008; Loh et al, 2006).

2.1.3 Metastable and ground state ESCs

Conventionally, ESCs were derived by cultivating blastocysts in medium supplemented with fetal calf serum on a feeder layer containing inactivated fibroblasts (Evans & Kaufman, 1981; Martin, 1981). How this culture condition enables faithful derivation of ESCs remained initially unknown. The crucial component of fetal calf serum is bone morphogenetic protein (BMP) 4 which activates Inhibitor of Differentiation (Id) genes via SMAD signaling pathways (Ying et al, 2003) (Figure **2**). Addition of leukemia inhibitory factor (LIF) in ESC cultures can substitute for the feeder layer when coating culture dishes with gelatin (Smith et al, 1988; Williams et al, 1988). Mechanistically, LIF acts primarily via Janus kinase (JAK) which mediates phosphorylation of signal transducer and activator of transcription (STAT) 3 and subsequently stimulates the expression of naïve pluripotency genes Klf2, Klf4 and Tfcp2l1 (Hall et al, 2009; Martello et al, 2013; Niwa et al, 2009; Ye et al, 2013) (Figure **2**). This led to the establishment of widely used feeder-free Serum/LIF (S/L) culture condition for ESCs. However, this condition results in a heterogeneous cell population in regards to expression of naïve pluripotency genes such as Rex1, Nanog, Esrrb and Klf4 (Chambers et al, 2007; Hayashi et al, 2008; Toyooka et al, 2008; van den Berg et al, 2008). S/L ESCs therefore exist in at least two distinct subpopulations corresponding to a naïve state, which recapitulates features of the pre-implantation epiblast, and a more advanced state, which is marked by expression of early post-implantation genes and reduced self-renewal (Kolodziejczyk et al, 2015). However, cells can fluctuate between these two states in culture suggesting a dynamic equilibrium *in vitro*. As a result, S/L ESCs are termed “metastable” (Figure **1**).

The question therefore arose whether it would be possible to develop culture conditions capturing a more homogenous ESC population. S/L ESCs produce FGF4 which activates extracellular signal-regulated kinase (ERK)/mitogen-activated protein kinase (MAPK) pathway in an autocrine manner and in turn promotes differentiation (Kunath et al, 2007; Stavridis et al, 2007). Paradoxically, Oct4 and Sox2 stimulate expression of Fgf4, suggesting that the pluripotent GRN primes its own dismantling (Ambrosetti et al, 1997). LIF and BMP4 signaling do not inhibit the Fgf4 activated pathway resulting in a constant battleground of pro- and anti-differentiation signals most likely causing the metastable cell state. In fact, knockout (KO) of Fgf4 reduces heterogeneity in ESCs and severely impairs differentiation (Kunath et al, 2007). Moreover, depletion of Erk2 results in a comparable phenotype implying the requirement of FGF/ERK signaling for exit from self-renewal (Stavridis et al, 2007). Mechanistically, Erk2 phosphorylates the naïve pluripotency factor KLF2, leading to its degradation and therefore destabilization of the pluripotency network (Yeo et al, 2014). Additionally, Erk1/2 also reinforce developmental programs by promoting expression of differentiation-associated genes (Tee et al, 2014). Indeed, chemical inhibition of mitogen-activated protein kinase kinases (MEK) 1 and 2 by PD0325901 (PD03), which blocks phosphorylation and therefore activation of Erk1/2, in appropriate culture conditions is sufficient to maintain ESCs in an undifferentiated state (Ying et al, 2008) (Figure **2**). However, PD03 alone is not able to substitute LIF in culture condition. Earlier studies showed that suppression of glycogen synthase kinase (GSK) 3 enhances ESC self-renewal (Ogawa et al, 2006; Sato et al, 2004). Inhibition of GSK3 mimics canonical WNT stimulation and thereby stabilizes cytoplasmic β-catenin. In turn, β-catenin translocates to the nucleus where it interacts with the transcriptional repressor TCF7L1 and therefore titrates it away from binding to and repressing key pluripotency genes (Martello et al, 2012; Wray et al, 2011; Yi et al, 2011). The combination of the GSK3 inhibitor Chir99021 (Chiron) and PD03, the so called 2i culture system, is sufficient to maintain ESCs in a pluripotent state without the requirement of Serum and LIF (Ying et al, 2008) (Figure **2**). Typically 2i cultures are carried out in feeder- and Serum-free medium such as N2B27 with the optional addition of LIF, which further stabilizes the pluripotent network (Martello et al, 2013; Ye et al, 2013).

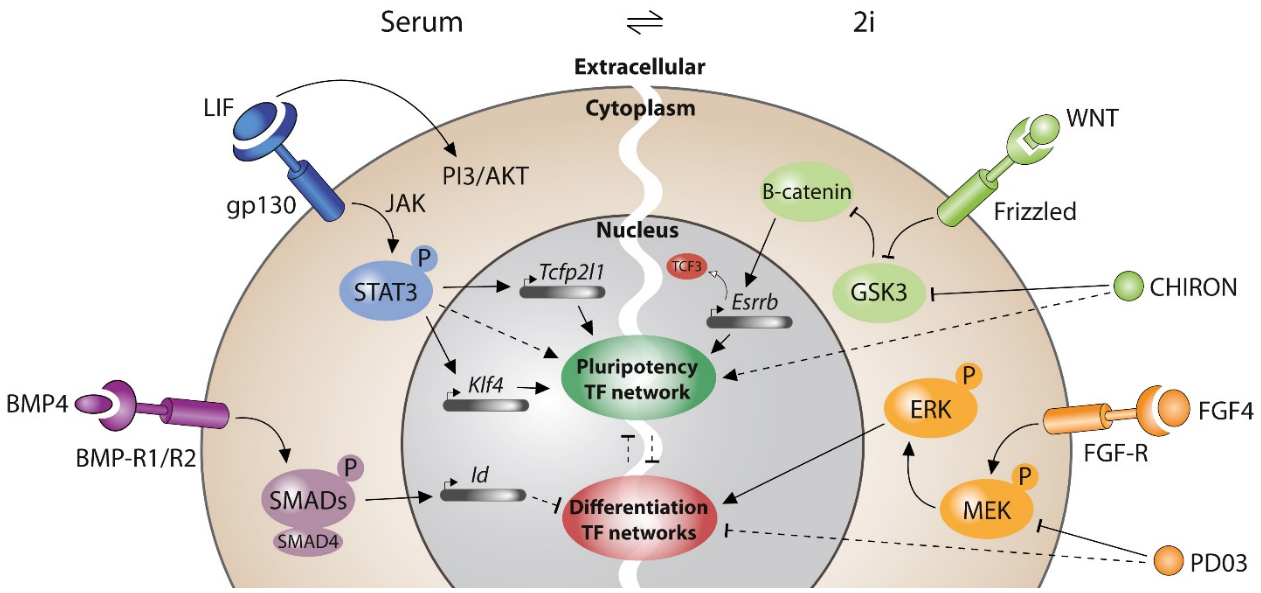


Figure 2: Signaling cascades promoting pluripotency and inhibiting differentiation.
Schematic overview of various signaling pathways that affect ESC self-renewal in metastable S/L (left side) and ground state 2i (right side) culture conditions. Clockwise: BMP4 is the crucial component in Serum and stimulates SMADs to activate Id genes. LIF acts primarily via JAK/STAT3 and activates expression of naïve pluripotency genes Tcfp2l1 and Klf4. Inhibition of GSK3 by Chiron mimics canonical WNT signaling which results in stabilization of β-catenin, subsequent inhibition of TCF7L1 (TCF3) and induction of Esrrb. Autocrine Fgf signaling activates the MAPK pathway which triggers differentiation. PD03 inhibits MEK and therefore maintains ESCs in an undifferentiated state. Figure adapted from Hackett & Surani, 2014.

Although S/L and 2i conditions are both sufficient to functionally capture pluripotency, ESCs cultured in these regimes behave very distinctly and differentially express roughly 3500 genes (Marks et al, 2012). This is partially explained by increased homogeneity in the chemically defined 2i condition, such as the uniform expression of various naïve pluripotency genes, e.g. Rex1, Nanog, Esrrb and Klf4, and the complete silencing of differentiation-associated markers (Figure **1**). However, the core pluripotency factors Oct4 and Sox2 are expressed at similar levels in both culture regimes (Marks et al, 2012). Besides transcriptional differences, a wide range of epigenetic features discriminates naïve and metastable pluripotency. DNA methylation at CpG dinucleotides is a repressive epigenetic mark usually associated with gene silencing. *In vivo*, the ICM is characterized by a hypomethylated genome, which is recapitulated in ESCs in 2i. Upon embryo implantation, DNA methylation rapidly increases, which is also observed upon converting ESCs from chemical defined to S/L conditions (Ficz et al, 2013; Habibi et al, 2013; Leitch et al, 2013; Monk et al, 1991; Monk et al, 1987). Similar to DNA methylation, the repressive trimethylation of lysine 27 of histone H3 (H3K27me3) is prominent in ESCs cultured in S/L, but not in 2i conditions (Marks et al, 2012). Consequentially, there are not many bivalent promoters, which are characterized by the presence of both repressive H3K27me3 and active trimethylation at lysine 4 of histone H3 (H3K4me3) histone marks (Azuara et al, 2006; Bernstein et al, 2006; Mikkelsen et al, 2007) and associated with rapid gene activation upon developmental progression, in ESCs cultured in chemically defined conditions (Marks et al, 2012). The difference between 2i and S/L cultured ESCs is even further exemplified by the function of *Prdm14*. While this TF is required to maintain pluripotency in conventional conditions, it is dispensable for self-renewal in 2i ESCs (Grabole et al, 2013; Yamaji et al, 2013). Thus, mechanisms controlling pluripotency can be context- and culture-dependent. Nevertheless, S/L ESCs converge on a naïve-specific transcriptome and hypomethylated genome when switched to 2i/LIF, suggesting high cellular plasticity (Ficz et al, 2013; Habibi et al, 2013; Marks et al, 2012). Taken together, two distinct culture methods for ESCs have been developed whereby 2i resembles the pre-implantation epiblast at E4.5 and S/L ESCs the post-implantation epiblast at E5.5 (Boroviak et al, 2014). Given the overall increase in homogeneity and closer resemblance to an early epiblast *in vivo*, ESCs cultured in 2i are therefore termed “ground state” pluripotent.

2.1.4 EpiSCs and primed pluripotency

EpiSCs can be obtained from a wide range of post-implantation embryonic stages (E5.5-E8.0) (Kojima et al, 2014; Osorno et al, 2012) in the presence of Fgf2 and Activin A (FA) on fibronectin coated culture dishes (Brons et al, 2007; Tesar et al, 2007). When injected into the post-implantation epiblast, EpiSCs most efficiently integrate into the PS of the developing embryo and their *in vitro* transcriptome resembles most closely the anterior PS of the E7.0 gastrula (Kojima et al, 2014). Similarly to ESCs, EpiSCs express the core pluripotency factors Oct4 and Sox2 but not naïve markers such as Rex1, Esrrb or Klf4 (Brons et al, 2007; Tesar et al, 2007). Notably, Nanog is detectable in EpiSCs, correlating with its re-activation in the posterior epiblast at around E6.0 (Hart et al, 2004). In contrast, early post-implantation markers, e.g. Fgf5, Otx2 and Oct6 (Pou3f1), are expressed and lineage markers such as T/Bra and Foxa2 are upregulated but only in a subset of cells (Tsakiridis et al, 2014) (Figure **1**). Additionally, EpiSCs display significant DNA methylation at promoters of pluripotency genes (Veillard et al, 2014). EpiSCs can also be derived from differentiating ESCs in the continuous presence of FA, however stable cultures are only obtained after several passages (Guo et al, 2009). Interestingly, upon neural differentiation, ESCs pass through an epiblast-like stage, from which EpiSCs can be derived, before acquiring neuronal identity (Zhang et al, 2010). This suggests that differentiation *in vitro* recapitulates developmental progression *in vivo* from pre- to post-implantation pluripotency before lineage specification.

2.1.5 Formative pluripotency

Naïve ESCs and primed EpiSCs resemble the E4.5 pre-implantation epiblast and the PS of E7.0 post-implantation epiblasts, respectively. Can one capture a cell state representing the peri-implantation epiblast? Single cell transcriptome analysis of early post-implantation epiblasts at E5.5 revealed a gene expression pattern distinct from the naïve pre-implantation and the primed gastrula stage, termed as formative pluripotency, which describes a progression phase of pluripotency (Mohammed et al, 2017; Smith, 2017). *In vitro*, ESCs exit the naïve pluripotent state before engaging into lineage specification (Kalkan et al, 2017; Mulas et al, 2017). In fact, a transient cell population with transcriptional similarities to the E5.5 epiblast has been identified upon differentiation of ESCs. These epiblast-like cells (EpiLCs) are generated from 2i/LIF ESCs by culturing in EpiSCs condition, FA, supplemented with knockout serum replacement (FAK) for 48 hours (h) (Hayashi et al, 2011) (Figure **1**). Activin seems to be dispensable for the transition to EpiLCs, suggesting that the predominant driver of EpiLC differentiation is Fgf signaling (Buecker et al, 2014). EpiLCs, similarly to the post-implantation epiblast, do not express naïve pluripotency markers, e.g. Rex1, Esrrb or Klf4, but instead early post-implantation genes such as Fgf5, Otx2 and Oct6, while expressing

Oct4 and Sox2 similarly to naïve ESCs. In contrast to primed EpiSCs, EpiLCs do not transcribe lineage markers, such as T/Bra or Foxa2. Moreover EpiLCs, unlike primed EpiSCs, can be differentiated into primordial germ cells *in vitro*, recapitulating potency of the early post-implantation epiblast *in vivo* (Hayashi et al, 2011). These findings suggest the existence of a continuum of pluripotent states between naïve and primed stem cells.

2.2 Exit from and acquisition of pluripotency

Historically, the Waddington landscape describes a model for normal development as a marble rolling downhill to its final differentiated state in which it preferentially resides (Waddington, 1957) (Figure **3**). While suggestive of continuous lineage restriction and irreversible commitment, TF overexpression experiments *in vitro* and *in vivo* have significantly challenged this view and revealed a profound degree of developmental cell state plasticity even in terminally differentiated cell types. Interconvertibility between distinct cell states offers the potential to study the dynamics of cell state transitions in a controllable system.

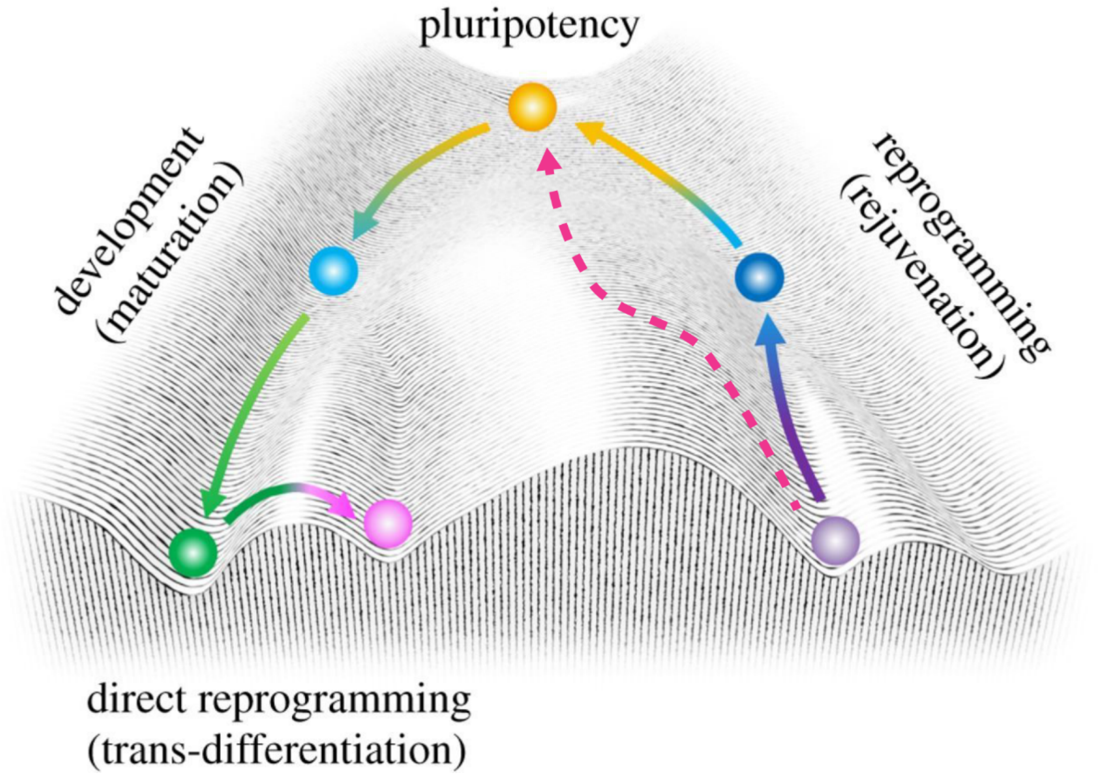


Figure 3: Cell state changes on Waddington's landscape.
Waddington's landscape model describes development as a strictly unidirectional process represented as a marble rolling downhill from a pluripotent (yellow, top) to a differentiated cell state (green, bottom) via a progenitor state (light blue, middle). Ectopic expression of tissue-specific TFs can convert the fate of a lineage-committed cell (green, bottom) to another lineage (pink, bottom), a process known as trans-differentiation. Transfer of a somatic nuclei into enucleated oocytes, cell fusion of a somatic with a pluripotent cell and ectopic expression of reprogramming factors is sufficient to erase the current cell state (purple, bottom) and to reprogram (rejuvenate) into a pluripotent state (yellow, top), which can occur through distinct trajectories (pink arrow). Figure adapted from Ohnuki & Takahashi, 2015.

2.2.1 ESC differentiation

Differentiation of ESCs and progression of the pre-implantation epiblast requires the resolution of the naïve pluripotent state in order to allow establishment of post-implantation pluripotency and eventually lineage specification and gastrulation. Various two- and three-dimensional systems have been established to differentiate ESCs *in vitro*. Embryoid bodies (EBs) are cultured as three-dimensional, round spheroids and are able to give rise to all three germ layers (Evans & Kaufman, 1981; Martin, 1981). However, EB cultures are heterogeneous and pose challenges to directed differentiation, while monolayer cultures are more controllable however usually recapitulate only one specific lineage (Keller, 2005). Conventional S/L ESCs are heterogeneous in terms of expression of pluripotency and lineage-associated markers. For a long time, it was believed that this heterogeneity is required for random exploration of available cell states and further lineage commitment (Moris et al, 2016). However, naïve ESCs are homogenous, suggesting that heterogeneity is not a property of stem cell pluripotency but rather dependent on the applied culture condition (Marks et al, 2012). If lineage specification requires priming and progression through a heterogeneous cell state remains, however, to be determined.

To monitor dynamics of ESC progression, various cell lines with fluorescently labeled proteins have been established. A particularly powerful system is the knock-in of a destabilized green fluorescent protein (GFP) with a half-life of 2h driven by the endogenous Rex1 promoter (RGd2) (Wray et al, 2010). Rex1 is homogeneously expressed in naïve 2i ESCs and the pre-implantation epiblast and, although Rex1 is not required for developmental progression, its expression is rapidly downregulated upon blastocyst implantation and during ESC differentiation (Boroviak et al, 2014; Kalkan et al, 2017; Masui et al, 2008; Pelton et al, 2002; Wray et al, 2010). Notably, silencing of Rex1 is heterogeneous *in vitro* (Wray et al, 2011), commencing after 24h of release from 2i in a subpopulation, but continuing for another 24h until Rex1-GFP expression is extinguished in all cells (Wray et al, 2011). Release of ESCs from 2i/LIF prolongs this process for an additional 12h (Dunn et al, 2014). Interestingly, downregulation of other naïve pluripotency factors, such as Nanog, Klf2 and Tfcp2l1, can be observed as early as 4h after removal of 2i (Leeb et al, 2014). Nevertheless, loss of Rex1 expression strongly correlates with loss of self-renewal indicating that downregulation of Rex1 marks irreversible cell fate commitment (Kalkan et al, 2017). Therefore, the RGd2 reporter enables near real-time tracking of exit from naïve pluripotency and purification of discrete and functionally defined cell populations by flow cytometry.

2.2.2 Somatic cell reprogramming

Reprogramming describes the process of erasing an existing cellular identity and reversion to a less differentiated cell state. However, reprogramming events are extremely rare *in vivo* and occur for example during tissue injury to initiate repair (Jessen et al, 2015). The idea of a strict unidirectionality during development as postulated by Waddington's landscape model (Waddington, 1957) (Figure **3**) was further challenged by somatic nuclear transfer experiments in frogs (Gurdon, 1962; Gurdon et al, 1958) and cell fusion experiments (Tada et al, 2001; Takagi et al, 1983). These findings suggested the existence of cell-intrinsic factors that can erase differentiated cellular programs. The first mechanistic insight into cell fate conversion came from studies showing that ectopic expression of myoblast determination protein (MYOD) is sufficient to trans-differentiate mouse fibroblasts into myoblasts (Davis et al, 1987) (Figure **3**). However, the breakthrough discovery was made by Shinya Yamanaka and colleagues who showed that somatic cells can be reprogrammed into iPSCs by the expression of four TFs, OCT4, SOX2, KLF4 and MYC (OSKM; Yamanaka factors) (Takahashi & Yamanaka, 2006). Mouse iPSCs are almost identical to ESCs in terms of morphology, gene expression and potency upon blastocyst injection (Okita et al, 2007; Takahashi & Yamanaka, 2006; Wernig et al, 2007). This discovery put forward the concept of TFs as determinants of mammalian cell identity. However, fibroblasts give rise to iPSCs at an efficiency of less than 0.1% (Wernig et al, 2007), therefore efforts have been made to identify reprogramming enhancers. Surprisingly, all four Yamanaka factors can be replaced by other, mainly related, factors (Heng et al, 2010; Jiang et al, 2008; Nakagawa et al, 2008) and expression of naïve pluripotency TFs can further enhance reprogramming efficiency (Takahashi & Yamanaka, 2016). Moreover, repressive epigenetic regulators have been identified as roadblocks for somatic reprogramming by restricting DNA binding of reprogramming TFs (Ebrahimi, 2015). These findings therefore suggest that formation of iPSCs can be achieved via distinct trajectories (Takahashi & Yamanaka, 2015) (Figure **3**). Reprogramming of fibroblasts is a multistep process with an early stochastic and late deterministic phase, which are characterized by specific molecular and transcriptional events. Especially the late phase of reprogramming recapitulates the reversion of certain developmental programs, e.g. a mesenchymal-to-epithelial transition is required for iPSC formation (Li et al, 2010; Samavarchi-Tehrani et al, 2010), while the opposite epithelial-to-mesenchymal transition is crucial for gastrulation (Acloque et al, 2009). Although it is debated whether a common developmental intermediate exists (Raab et al, 2017), these observations suggest that differentiation and reprogramming may employ common mechanisms in inverse directions.

2.2.3 Reprogramming of EpiSCs

As outlined above, conventional reprogramming of somatic cells into iPSCs requires overexpression of 4 TFs, proceeds over many days and is associated with cell death, therefore hampering mechanistical dissection of this process. EpiSCs are more closely related to iPSCs than somatic cells, suggesting that trajectories during EpiSC reprogramming may be less complicated and more tractable. Although EpiSCs, when transferred in either conventional S/L or chemically defined 2i/LIF culture condition, do not acquire the naïve pluripotent state but in fact differentiate and eventually die (Brons et al, 2007; Guo et al, 2009), overexpression of only one TF is sufficient for formation of EpiSC-derived iPSCs (Epi-iPSCs) (Guo et al, 2009). Specifically, activation of Klf2, Klf4, Nanog, Esrrb, Tfcp2l1 or Gbx3 (Festuccia et al, 2012; Guo et al, 2009; Hall et al, 2009; Martello et al, 2013; Silva et al, 2009; Tai & Ying, 2013; Ye et al, 2013) or hyperactivation of Stat3 (Yang et al, 2010) in EpiSCs generates Epi-iPSCs in the presence of 2i or 2i/LIF. Interestingly, different EpiSC reprogramming regimes proceed along distinct transcriptional trajectories before converging on the same naïve endpoint, suggesting a remarkable flexibility for the acquisition of cell identity from the same starting population (Stuart et al, 2019). Thus, interconvertibility of developmentally related naïve ESCs and primed EpiSCs provides an attractive system to identify and characterize regulators of pluripotent cell state transitions.

2.3 Regulators of cell state transitions

Cell fate determination is a highly dynamic and coordinated process that requires dissolution of the starting cell identity and acquisition of a new cell fate. Previous studies of differentiation, trans-differentiation and reprogramming showed a particular importance for TFs in these processes. However, many other mechanisms acting on multiple levels such as epigenetic regulation, nuclear transport, metabolic sensing and protein stability are also required for cell state transitions.

The introduction of ribonucleic acid interference (RNAi) and later CRISPR/Cas9 offered the possibility to perform genome-wide disruption of gene function. RNAi is a highly conserved biological process in which binding of small RNA molecules to complementary messenger RNA (mRNA) results in degradation of target mRNAs and thus reduction of protein levels (Fire et al, 1998). CRISPR/Cas9 is a prokaryotic defense mechanism that provides resistance to foreign genetic elements by recognizing and cleaving foreign DNA (Wiedenheft et al, 2012). This system was quickly adapted for mammalian cells in which a single guide RNA (sgRNA) guides the Cas9 nuclease to complementary DNA to create a site-specific double-strand break which can result in error-prone repair and thus frameshift mutations (Cong et al, 2013; Jinek et al, 2012; Mali et al, 2013). These tools have been extensively exploited to identify regulators of ESC differentiation by loss-of-function screens (Betschinger et al, 2013; Hackett et al, 2018; Leeb et al, 2014; Li et al, 2018; MacDougall et al, 2019b; Yang et al, 2012). Overall, more than 600 protein-coding genes have been implicated in exit from naïve pluripotency however only few were studied in detail. Although similar loss-of-function screens have been performed in somatic reprogramming (Qin et al, 2014; Yang et al, 2014a), knowledge about regulators of EpiSC de-differentiation is still limited. Here I will describe regulators of exit from and acquisition of naïve pluripotency that are relevant for the thesis.

2.3.1 Guardians of ESC differentiation

Tcf7l1 is a transcriptional repressor and the main downstream effector of Wnt/GSK3 signaling in ESCs. Withdrawal of Chiron from the 2i culture medium leads to the de-repression of TCF7L1 and subsequently repression of its direct targets Klf2, Nanog and Esrrb (Martello et al, 2012; Pereira et al, 2006). Among these, repression of Esrrb seems to be most critical for exit from naïve pluripotency. Thus, depletion of Tcf7l1 inhibits ESC differentiation (Guo et al, 2011; Pereira et al, 2006) and displays gastrulation defects *in vivo* (Merrill et al, 2004), while overexpression of Esrrb is sufficient to inhibit ESC progression (Martello et al, 2012). The important role of Tcf7l1 is highlighted by the fact that it scores as a top hit in most large-scale ESC differentiation screens (Betschinger et al, 2013; Hackett et al, 2018; Leeb et al, 2014; Li et al, 2018; MacDougall et al, 2019b; Yang et al, 2012).

Similar to Tcf7l1, the tumor-suppressor Flcn is also required upstream for repression of Esrrb during exit from naïve pluripotency (Betschinger et al, 2013). Flcn together with its interaction partners Fnip1 and Fnip2 drives ESC progression by cytoplasmic retention of the basic helix-loop-helix (bHLH) TF Tfe3. Thus, differentiation is strongly impaired by either depletion of Flcn or ectopic expression of nuclear Tfe3. In naïve ESCs, Tfe3 is localized in both the nucleus and cytoplasm but upon 2i withdrawal translocates into the cytoplasm. Nuclear exclusion of Tfe3 is also observed in the post-implantation epiblast, suggesting that the same mechanism operates during development *in vivo* (Betschinger et al, 2013). The Flcn-Fnip1/2 complex and the microphthalmia-associated transcription factor (MiTF)/Tfe family member Tfeb are known components of the somatic amino acid signaling pathway (Saxton & Sabatini, 2017). Mechanistically, Flcn activates mechanistic target of rapamycin complex (mTORC) 1 in response to amino acids, which in turn phosphorylates and inactivates Tfeb via cytoplasmic sequestration (Petit et al, 2013; Roczniak-Ferguson et al, 2012). Paradoxically, depletion of Flcn or tuberous sclerosis complex (Tsc) 2, an mTORC1 repressor, in naïve ESCs causes nuclear retention of Tfe3 and impairs exit of the ESC state (Betschinger et al, 2013). How the subcellular localization of Tfe3 is regulated and to which extent the amino acid signaling machinery is involved therefore remains to be explored.

The zinc finger TF Zfp281 was identified in several screens as a driver of ESC differentiation (Betschinger et al, 2013; Hackett et al, 2018; Li et al, 2018). This is in line with abnormal embryonic development of Zfp281-null mice, specifically defective activation of Nodal signaling and induction of genes associated with the primed pluripotent state, leading to embryonic lethality around E8.5 (Huang et al, 2017). Initial studies showed that Zfp281 recruits the NuRD complex, a multi-subunit complex which deacetylates H3K27 and thereby facilitates H3K27me3 and gene repression, to the Nanog promoter in metastable S/L ESCs to mediate Nanog repression (Fidalgo et al, 2012; Fidalgo et al, 2011). Thus, depletion of Zfp281 causes increased expression of Nanog and impairs EB differentiation. More recently, it was proposed that Zfp281 expression is upregulated upon differentiation of S/L cultured ESCs, which in turn destabilizes metastable pluripotency by direct protein interaction with the methylcytosine hydroxylase Tet1 and transcriptional repression of Tet2 (Fidalgo et al, 2016). These studies suggest various mechanisms for Zfp281 regulated cell state transition *in vitro* and *in vivo*. However, if these mechanisms also regulate exit from naïve pluripotency downstream of Zfp281 is unclear.

The homeobox TF Otx2 and the zinc finger TF Zic2 were identified as regulators of ESC differentiation in a genome-wide loss of function screen (Li et al, 2018). Homozygous mutant mice of Otx2 and Zic2 are embryonic lethal around E9.5 and E13.5 with severe defects in rostral head and neural crest cell formation, respectively, demonstrating an essential role in embryogenesis (Acampora et al, 1995; Elms et al, 2003; Matsuo et al, 1995). Zic2 together with the NuRD complex occupies enhancers of bivalent genes in metastable S/L ESCs to maintain H3K27me3 and depletion of Zic2 results in impaired upregulation of primed pluripotency markers upon EB differentiation (Luo et al, 2015). However, its precise role in exit from naïve pluripotency is unknown. Depletion of Otx2 in metastable S/L ESCs results in increased self-renewal and homogenous expression of the TF Nanog (Acampora et al, 2013).

In contrast, Otx2 is lowly expressed in 2i/LIF condition and its loss does not influence the naïve pluripotent state (Buecker et al, 2014). However, Otx2 interacts with Oct4 and is crucial to target Oct4 to previously inaccessible enhancer sites during ESC differentiation (Buecker et al, 2014; Yang et al, 2014b). Differentiating Otx2 KO ESCs therefore fail to repress naïve and induce primed marker genes. Ectopic expression of tagged Otx2 and Zic2 transgenes in EpiSCs revealed co-localization of these two TFs at active genomic regions in primed pluripotent cells (Matsuda et al, 2017), suggesting that Otx2 and Zic2 functionally act together in the exit of pluripotency.

While various TFs have important roles in safeguarding ESC progression, epigenetic modifications are also key to control precise gene expression during this process (Festuccia et al, 2017). A well-studied example is the heterodimeric histone methylation complex Ehmt1 (GLP)/Ehmt2 (G9a) which mediates repressive mono- and dimethylation at histone H3 lysine 9 (H3K9me1 and H3K9me2) (Shinkai & Tachibana, 2011). Both complex members were identified as regulators of exit from naïve pluripotency in a genome-wide loss of function screen (Li et al, 2018), which is consistent with increased H3K9me2 and subsequent silencing of associated developmental genes observed during embryo implantation and ESC differentiation (Zylicz et al, 2015). In fact, disruption of either Ehmt1 or Ehmt2 in mice results in widespread loss of H3K9me1 and H3K9me2, growth retardation and embryonic lethality around E9.5 (Tachibana et al, 2002; Tachibana et al, 2005). Co-depletion of Ehmt1 and Ehmt2 does not further reduce H3K9me1 and H3K9me2 levels, suggesting non-redundant functions (Tachibana et al, 2005). However, the exact mechanism that stimulates deposition of H3K9me2 during development remains unclear.

2.3.2 Roadblocks of EpiSC reprogramming

While roadblocks of EpiSC reprogramming have not been systematically identified, several genes inhibiting formation of Epi-iPSCs have been reported. These include the TFs Otx2 and

Zfp281 which also function in exit from naïve pluripotency. Ablation of Otx2 in primed EpiSCs leads to acquisition of an ESC-like morphology as well as upregulation of naïve and downregulation of primed pluripotency genes, even without reprogramming stimuli. However this dramatic phenotype is only apparent after several passages in EpiSC promoting conditions, thus suggesting a cellular adaption upon loss of Otx2 (Acampora et al, 2013). In contrast, depletion of Zfp281 in EpiSCs in the presence of FA results in destabilization of this cell state with proliferation defects and, ultimately, cell death (Fidalgo et al, 2016). However, upon shifting to naïve 2i/LIF culture conditions, Zfp281 depletion results in formation of Epi-iPSCs, potentially via transcriptional upregulation of Tet2. Thus, these examples show that exit from and acquisition of naïve pluripotency might be regulated via similar mechanisms.

3. AIM OF THIS THESIS

As outlined above, cell identities are specified by so called “master” or “lineage-determining” TFs that are sufficient to even convert terminally differentiated cells into the naïve pluripotent state. However, how such TFs are regulated in the context of developmental cell state transitions that orchestrate cell type specification in a spatially and temporally controlled manner, is unclear. The capture of developmentally related ESCs and EpiSCs provides a controllable *in vitro* setup to systematically annotate mechanisms of cellular plasticity.

Within my PhD project I set out to address the following questions:

- How is ESC identity silenced during developmental progression?
- Is reprogramming into iPSCs a reversion of “natural” differentiation?

4. RESULTS

4.1 Lysosomal signaling licenses embryonic stem cell differentiation via inactivation of Tfe3

Lysosomal signaling licenses embryonic stem cell differentiation via inactivation of Tfe3

Villegas F*, Lehalle D*, Mayer D*, Rittirsch M, Stadler MB, Zinner M, Olivieri D, Vabres P, Duplomb-Jego L, De Bont E, Duffourd Y, Duijkers F, Avila M, Genevieve D, Houcinat N, Jouan T, Kuentz P, Lichtenbelt KD, Thauvin-Robinet C, St-Onge J, Thevenon J, van Gassen KLI, van Haelst M, van Koningsbruggen S, Hess D, Smallwood SA, Riviere JB, Faivre L, Betschinger J

Cell Stem Cell. 2019 Feb; 24(2):257-270.e8 * Equal contribution

Abstract

Self-renewal and differentiation of pluripotent murine embryonic stem cells (ESCs) is regulated by extrinsic signaling pathways. It is less clear whether cellular metabolism instructs developmental progression. In an unbiased genome-wide CRISPR/Cas9 screen, we identified components of a conserved aminoacid-sensing pathway as critical drivers of ESC differentiation. Functional analysis revealed that lysosome activity, the Ragulator protein complex, and the tumor-suppressor protein Folliculin enable the Rag GTPases C and D to bind and seclude the bHLH transcription factor Tfe3 in the cytoplasm. In contrast, ectopic nuclear Tfe3 represses specific developmental and metabolic transcriptional programs that are associated with peri-implantation development. We show differentiation-specific and non-canonical regulation of Rag GTPase in ESCs and, importantly, identify point mutations in a Tfe3 domain required for cytoplasmic inactivation as potentially causal for a human developmental disorder. Our work reveals an instructive and biomedically relevant role of metabolic signaling in licensing embryonic cell fate transitions.

The entire article can be found in the appendix section.

4.1.1 Contributions

In this study, a master's student, Marietta Zinner, designed and performed the genome-wide CRISPR/Cas9 screen under my supervision. Additionally, I performed the computational analysis of RNA- and chromatin immunoprecipitation coupled to deep sequencing (RNA-seq and ChIP-seq) data in collaboration with Michael Stadler and assisted in writing the manuscript with Jörg Betschinger.

4.2 Zfp281 orchestrates interconversion of pluripotent states by engaging Ehmt1 and Zic2

Zfp281 orchestrates interconversion of pluripotent states by engaging Ehmt1 and Zic2

Mayer D, Stadler MB, Rittirsch M, Hess D, Lukonin I, Winzi M, Smith A, Buchholz F, Betschinger J

EMBO Journal. 2019 Nov; 29:e102591

Abstract

Developmental cell fate specification is a unidirectional process that can be reverted in response to injury or experimental reprogramming. Whether differentiation and de-differentiation trajectories intersect mechanistically is unclear. Here, we performed comparative screening in lineage-related mouse naïve embryonic stem cells (ESCs) and primed epiblast stem cells (EpiSCs), and identified the constitutively expressed zinc finger transcription factor (TF) Zfp281 as a bi-directional regulator of cell state interconversion. We showed that subtle chromatin binding changes in differentiated cells translate into activation of the histone H3 lysine 9 (H3K9) methyltransferase Ehmt1 and stabilization of the zinc finger TF Zic2 at enhancers and promoters. Genetic gain- and loss-of-function experiments confirmed a critical role of Ehmt1 and Zic2 downstream of Zfp281 both in driving exit from the ESC state, and in restricting reprogramming of EpiSCs. Our study reveals that cell type-invariant chromatin association of Zfp281 provides an interaction platform for remodeling the cis-regulatory network underlying cellular plasticity.

The entire article can be found in the appendix section.

4.2.1 Contributions

This work comprises my main PhD project and hence I was developing and performing most of the experiments. The large-scale EpiSC reprogramming screen was performed in the lab of Frank Buchholz at the technical University in Dresden by Maria Winzi and Jörg Betschinger. Melanie Rittirsch assisted in multiple experiments, Daniel Hess performed the mass spectrometry analysis and Ilya Lukonin the spheroid image analysis. I performed the computational analysis in collaboration with Michael Stadler and assisted in writing the manuscript with Jörg Betschinger.

4.3 Additional Results I: Zfp281 genetically interacts with Otx2

4.3.1 Screen of selected TFs for genetic interaction with Zfp281

We showed that transcription and DNA occupancy of Zfp281 remains largely unchanged during ESC differentiation (Mayer et al, 2019; Figure **2F**, **3E**). We therefore hypothesized that cell state specific protein interaction partners of Zfp281 might explain differentiation defects observed in the absence of Zfp281. This experimental strategy led to the identification of Ehmt1 and Zic2 as functional downstream effectors of Zfp281 specifically during ESC progression (Mayer et al, 2019; Figure **4A**, **C**).

In parallel to the above mentioned approach, we screened for genetic interactions of Zfp281 with candidate lineage-determining TFs with the following characteristics: (1) scored as a hit in at least one of the published knockdown ESC differentiation screens (Betschinger et al, 2013; Leeb et al, 2014; Yang et al, 2012), (2) mRNA is expressed in ESCs (RPKM>1) and (3) mRNA is upregulated 32h after 2i withdrawal (> 2 fold). To assess for genetic interactions of the 11 TFs passing these three criteria with Zfp281, we depleted each in wildtype (*WT*) and *Zfp281* KO RGd2 ESCs, which express a destabilized GFP downstream of the *Rex1* promoter (Wray et al, 2011), with small interfering RNAs (siRNAs) and differentiated them for 48h towards EpiLCs. While depletion of individual TFs in *WT* cells showed only little to no effect compared to control siRNA (neg), in *Zfp281* KO cells knockdown of one candidate, the homeobox TF Otx2, strongly enhanced retention of self-renewal from 8 to 20% (Figure **4A**). Otx2 has been identified as an important regulator for induction and maintenance of the EpiSC state (Acampora et al, 2013), likely via redirecting Oct4 to previously inaccessible enhancer sites during ESC differentiation (Buecker et al, 2014; Yang et al, 2014b). To test whether this genetic interaction is also functionally important during reprogramming, we used 796.4 EpiSCs which express GFP and Puromycin N-acetyl-transferase under the control of the regulatory sequence of the *Oct4* gene (Guo et al, 2011; Guo et al, 2009), and a constitutively expressed Stat3 activating receptor (GY118F) responsive to granulocyte colony stimulating factor (Gcsf)

(Yang et al, 2010). Stimulation of 796.4 EpiSCs with 2i and Gcsf for 4 days (d), followed by selection with Puromycin resulted in little to no self-renewing Epi-iPSCs which was strongly increased upon depletion of Zfp281 (Mayer et al, 2019; Figure **1D**). To ensure unambiguous detection of genetic interactions with Zfp281, we reduced treatment duration with Gcsf to 2d. Similar to what we observed during ESC differentiation, depletion of selected TFs did not enhance EpiSC reprogramming on their own. However, knockdown of Foxp1, Hand1, Jun and Otx2 significantly increased the Zfp281-dependent phenotype (Figure **4B**, **C**). Although knockdown of Foxp1 showed the strongest increase in formation of Epi-iPSC colonies when co-depleted with Zfp281, it did not enhance the *Zfp281* KO phenotype in ESC differentiation. In fact, Otx2 was the only gene which augmented Zfp281 function in both ESC differentiation and reprogramming of EpiSCs, suggesting a genetic interaction of these two TFs during pluripotent cell state transitions.

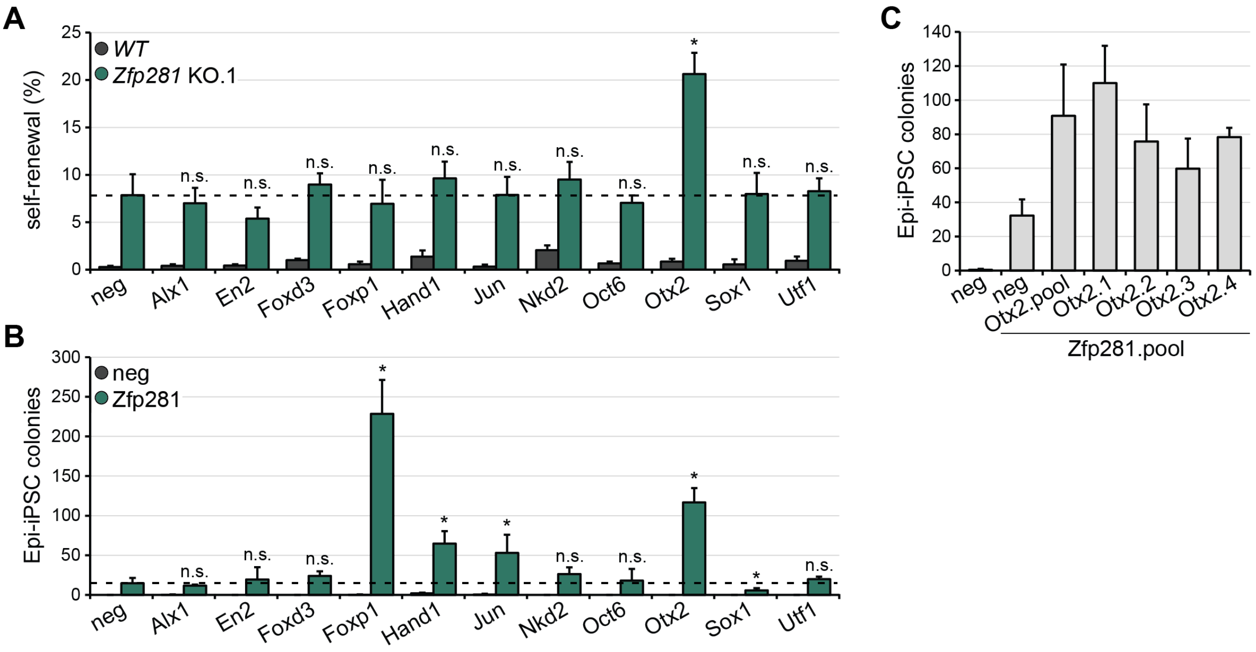


Figure 4: Screen for selected lineage-determining TFs downstream of Zfp281.
(**A**) Self-renewal of *WT* (grey) and *Zfp281* KO (green) cells transfected with indicated siRNAs after 48h of 2i withdrawal with addition of FAK. Dashed line marks self-renewal of *Zfp281* cells transfected with negative (neg) siRNA. Significance was determined using a Wilcoxon Mann-Whitney rank sum test compared to transfection with neg siRNA in *Zfp281* cells. (*) <0.05; not significant (n.s.). Average and standard deviation (SD) of 2 biological replicates.
(**B**) Number of Epi-iPSC colonies derived from 796.4 EpiSCs transfected with indicated siRNAs in combination with neg (grey) and Zfp281 (green) siRNAs, stimulated with Gcsf and 2i for 2d, and selected with Puromycin. Dashed line marks number of Epi-iPSC colonies transfected with Zfp281 siRNA. Significance was determined using a Wilcoxon Mann-Whitney rank sum test compared to transfection with Zfp281 siRNA. (*) <0.05; not significant (n.s.). Average and SD of 2 biological replicates.
(**C**) Deconvolution of Otx2 siRNA pool: Epi-iPSC colonies derived from 796.4 EpiSCs transfected with Otx2 siRNAs (individual siRNAs or pool) together with Zfp281 siRNA pool, stimulated for 2d with Gcsf and 2i, and selected with Puromycin. Average and SD of 2 biological replicates.

4.3.2 Otx2 genetically but not physically interacts with Zfp281 in pluripotent cell state transition

Deletion of Zfp281 causes deficient ESC differentiation and subsequent deregulation of several naïve pluripotency and early lineage markers during this process (Mayer et al, 2019; Figure **2A**, **EV3A**). Transcriptional upregulation of Otx2 upon exit from naïve pluripotency is required and sufficient for ESC progression (Buecker et al, 2014). To investigate mRNA levels of Otx2 in the absence of Zfp281, we made use of RNA-seq data from *WT* and *Zfp281* KO ESC differentiation (Mayer et al, 2019; Table **EV2**). Otx2 mRNA is more than 5-fold induced after 32h of 2i withdrawal in *WT* cells and is almost unaffected in *Zfp281* deficient cells (Figure **5A**). To further characterize the interaction of Zfp281 with Otx2, we generated *Otx2* individual and *Zfp281/Otx2* compound KO RGd2 ESCs (Figure **5B**). GFP is uniformly expressed in *WT* naïve ESCs (Rex1-GFP^{high}) and progressively downregulated during differentiation (Rex1-GFP^{low}). Reporter expression of two independent *Otx2* and *Zfp281/Otx2* KO clones was similar to parental *WT* cells and one untargeted sibling clone (*Otx2* WT) in 2i (Figure **5C**). 24h after 2i withdrawal, 90% of *Otx2* KO cells were GFP^{high} compared to 56% in *WT* cells. However, this phenotype only persisted moderately as at 72h after release from 2i and at 48h of EpiLC differentiation, less than 1% and 3% of *WT* cells were GFP^{high}, while 2% and 9% of *Otx2* KO cells maintained high GFP expression, respectively. Correspondingly, *Otx2* KO cells did not show any retention of self-renewal at 72h of differentiation and at 48h of EpiLC differentiation only 3% of *Otx2* deficient cells compared to 1% of *WT* cells could still self-renew (Figure **5D**). This therefore indicates that Otx2 depletion on its own induces predominantly a transient ESC differentiation phenotype. In contrast, a larger proportion of *Zfp281/Otx2* KO cells were Rex1-GFP^{high} after 72h of 2i withdrawal (43%) and EpiLC differentiation (49%) compared to *Zfp281* KO cells (14% and 31%, respectively) (Figure **5C**). Inability to repress Rex-GFP was also reflected by retention of self-renewal in 17% and 34% of *Zfp281/Otx2* KO cells compared to 10% and 15% of *Zfp281* KO cells after 72h of 2i withdrawal and 48h EpiLC differentiation, respectively (Figure **5D**). We therefore conclude that although individual *Otx2* KO cells only have a transient phenotype in ESC progression, co-depletion with Zfp281 strongly enhances the single mutant phenotypes, indicating a synergistic rather than additive interaction.

The genetic interaction of Zfp281 and Otx2 was not restricted to ESC differentiation but was also evident in reprogramming of 796.4 EpiSCs (Figure **4B**). To exclude a cell-line and/or Stat3-dependent effect, we used an independent O4GIP^{GY118F} EpiSC line as well as O4GIP EpiSCs conditionally expressing Esrrb or Klf4 via addition of Doxycycline (Dox) to induce reprogramming (Mayer et al, 2019). Surprisingly, co-depletion of Otx2 with Zfp281 in these

EpiSC lines did not increase the number of derived Epi-iPSCs when compared to Zfp281 knockdown (Figure **5E**, **F**). These results suggest that the synergy of Otx2 and Zfp281 only augmented Zfp281 in restricting reprogramming is cell line- and/or context-dependent.

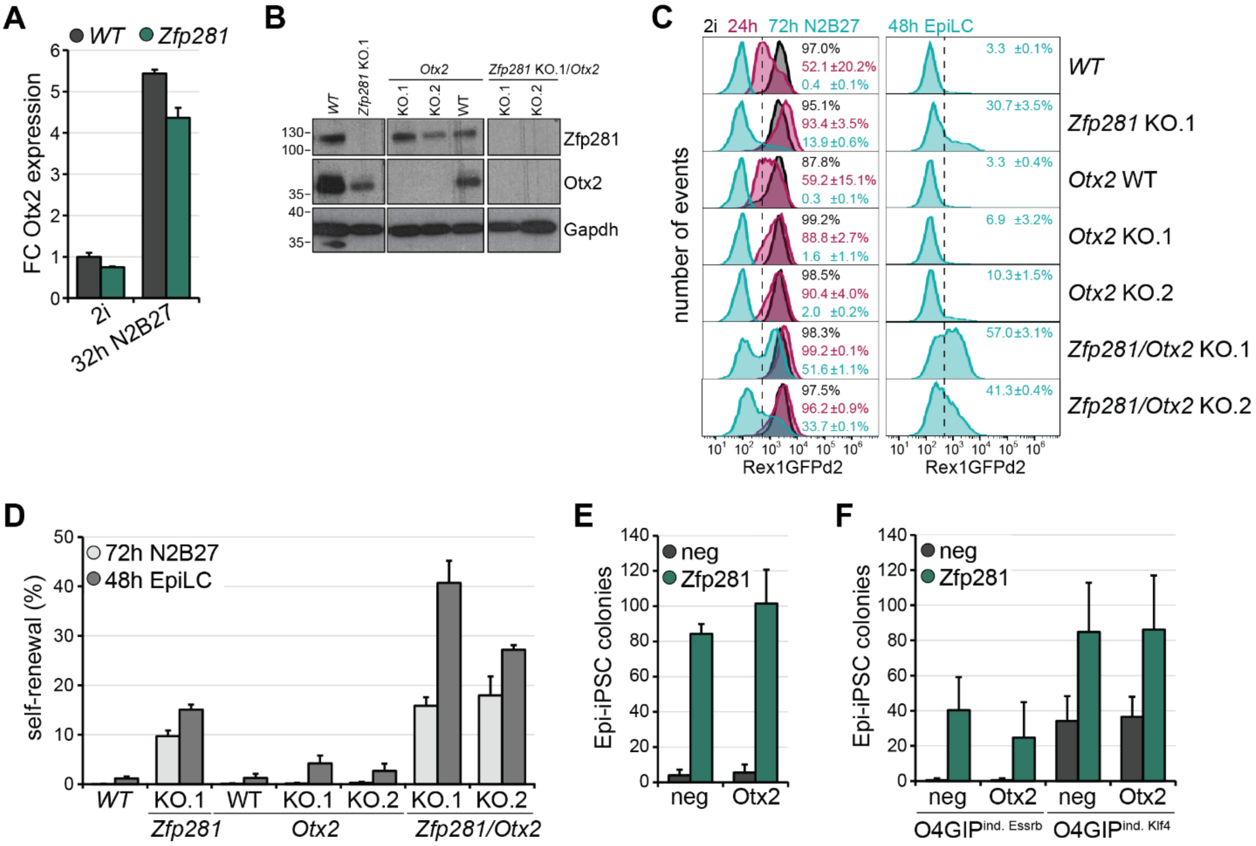


Figure 5: Otx2 genetically interacts with Zfp281 in pluripotent cell state transition.
(**A**) Otx2 mRNA fold change (FC) in *WT* and *Zfp281* naïve ESCs and 32h after 2i withdrawal. mRNA levels were extracted from RNA-seq datasets and normalized to *WT* ESCs (Mayer et al, 2019).
(**B**) Absence of proteins in indicated KO cells.
(**C**) Representative flow cytometry profiles of RGd2 cells of indicated genotypes in 2i, and after 24h and 72h of 2i withdrawal (left) and 48h of 2i withdrawal with addition of FAK (right). Dashed line indicates cut-off for Rex1-GFP^{high} and Rex1-GFP^{low} cells. Numbers are the average and SD of Rex1-GFP^{high} cells in 2 biological replicates.
(**D**) Self-renewal of cells with specified genotype after differentiation in indicated conditions. Average and SD of 2 biological replicates.
(**E**, **F**) Number of Epi-iPSC colonies derived from O4GIP^{GY118F} (**E**) and O4GIP EpiSCs containing Dox-inducible Esrrb or Klf4 transgenes (**F**) after transfection with indicated siRNA combinations, stimulated for 2d with Gcsf and 2i (**E**) or with Dox and 2i for 2d (**F**), and selection with Puromycin. Average and SD of 3 (O4GIP$^{ind.\ Klf4}$) or 2 (O4GIP^{GY118F} and O4GIP$^{ind.\ Esrrb}$) biological replicates.

Although we obtained non-conclusive results for a genetic interaction of Otx2 and Zfp281 during reprogramming of EpiSCs, increased resistance to ESC differentiation in compound KO cells was robust. To test whether this is due to co-binding of these two TFs to chromatin, we performed ChIP-seq of Otx2 in naïve ESCs (WT^{2i}) as well as 32h after 2i withdrawal (WT^{32h}). As described before (Buecker et al, 2014; Yang et al, 2014b), Otx2 binds dynamically to chromatin during ESC differentiation resulting in a weak correlation between these two cell states (Figure **6A**). However, overlap of Otx2 ChIP peaks with Zfp281 ChIP peaks (Mayer et al, 2019) was very sparse, and observed at only 15% of all Otx2 and 11% of all Zfp281 peaks respectively (Figure **6B**). Additionally, Zfp281 immunoprecipitation coupled to semi-quantitative mass spectrometry (IP-MS) in nuclear extracts of WT^{2i} and WT^{40h} cells did not reveal a direct protein interaction of Zfp281 with Otx2 (Mayer et al, 2019; Table **EV3**). Together, these data suggest that although not physically interacting, Zfp281 and Otx2 seem to synergistically regulate ESC progression *in vitro*, potentially by acting in parallel pathways.

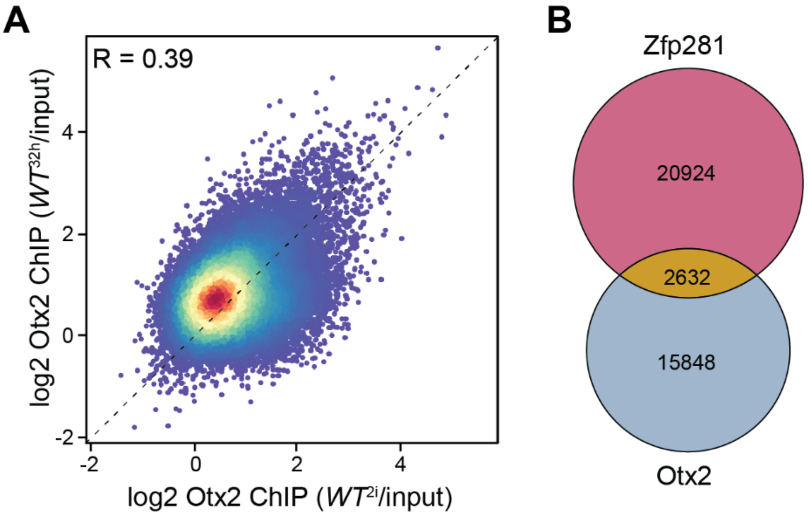


Figure 6: Comparison of DNA occupancy of Zfp281 and Otx2.
(**A**) Scatter plot comparing Otx2 log2 ChIP enrichment relative to matched inputs in WT^{2i} and WT^{32h} cells. (R) Pearson's correlation coefficient.
(**B**) Overlap of Zfp281 and Otx2 ChIP peaks.

4.4 Additional Results II: Similarities and differences between Zfp281 and Tfe3 regulated cell differentiation

4.4.1 No mechanistic overlap of Zfp281 and Tfe3 in pluripotent cell state transitions

Transitions of pluripotent cell states are controlled on different molecular levels including signaling pathways and metabolic sensing which ultimately lead to transcriptional changes ensuring stem cell progression. We have previously shown that the Ragulator protein complex, Tsc1/2 and Flcn together with Fnip1/2 enable the Rag GTPases C and D to bind and retain the bHLH TF Tfe3 in the cytoplasm, therefore enabling exit from the naïve pluripotent state (Villegas et al, 2019). In contrast, Zfp281 orchestrates pluripotent cell fate transition via recruitment and activation of the TF Zic2 and the methyltransferase Ehmt1 to chromatin albeit Zfp281 itself is relatively static during this process (Mayer et al, 2019). In both studies we showed that depletion of either Flcn, Lamtor1 or Zfp281 resulted in long-term maintenance of self-renewal in differentiation-promoting conditions (Villegas et al, 2019; Figure **1F**) (Mayer et al, 2019; Figure **2A, EV2B**), suggesting key roles in ESC progression. However, whether the Flcn-Lamtor-Tfe3 and Zfp281-Ehmt1-Zic2 axes converge on shared downstream targets remained to be determined.

In contrast to Zfp281, none of the Tfe3 upstream regulators scored positively in our EpiSC reprogramming screen (Mayer et al, 2019; Table **EV1**), nor did ectopic nuclear Tfe3 increased EpiSC de-differentiation efficiency (Betschinger et al, 2013). To validate these observations, we transfected three independent GY118F expressing EpiSC lines with siRNAs against Flcn, Fnip1/2, Lamtor3 and Tsc2 and induced reprogramming. As a negative control we included siRNAs targeting Tcf7l1, a differentiation-specific regulator (Pereira et al, 2006) with no described function in reprogramming of EpiSCs (Mayer et al, 2019; Figure **1D**), and as a positive control Zfp281. In all cell lines, depletion of none of the Tfe3 regulators increased the number of Epi-iPSCs compared to control transfection (Figure 7**A**), demonstrating distinct functions of the Flcn-Lamtor-Tfe3 and Zfp281-Ehmt1-Zic2 axes in stabilizing EpiSCs against reprogramming.

To test whether Zfp281 function in ESC differentiation depends on Tfe3, we made use of an independent differentiation assay using O4GIP ESCs (Betschinger et al, 2013). Knockdown of Tcf7l1, Zfp281 and Tfe3 upstream regulators caused resistance to differentiation (Figure **7B**). Co-depletion of Tfe3 but not related MiTF/Tfe family member Tfeb reverted the differentiation phenotype of Flcn, Fnip1/2, Lamtor3 and Tsc2 but not Tcf7l1 or Zfp281, suggesting that Zfp281 acts independently of Tfe3 in ESC progression (Figure **7B**). In line with these results, knockdown of Flcn but not Zfp281 or Tcf7l1 induced nuclear Tfe3 accumulation in naïve ESCs (Figure **7C**). Taken together these results suggest that Zfp281 acts independently of the Flcn-Lamtor-Tfe3 axis in pluripotent cell state transitions.

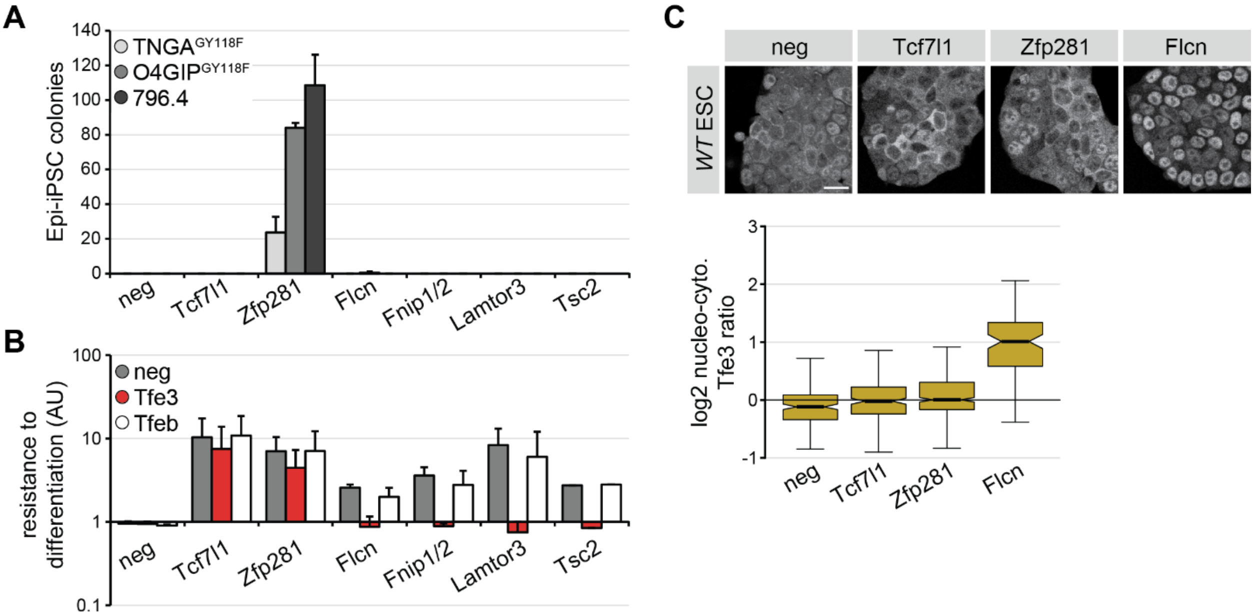


Figure 7: Zfp281 and Tfe3 regulate pluripotent cell state transitions independently from each other.
(**A**) Number of Epi-iPSC colonies derived from specified EpiSC lines after transfection with indicated siRNAs, stimulated with Gcsf and 2i for 4d, and selection with Puromycin. Average and SD of 3 (TNGA^{GY118F}) and 2 (OGIP^{GY118F} and 796.4) technical replicates.
(**B**) Resistance to differentiation of O4GIP ESCs transfected with indicated siRNAs in combination with neg (grey), Tfe3 (red) or Tfeb (white) siRNAs after 72h of 2i withdrawal. Results were normalized to neg siRNA and are shown on a log10-scaled axis. Average and SD of 2 biological replicates.
(**C**) Tfe3 localization (top) and log2 nucleo-cytoplasmic Tfe3 ratios (bottom) in O4GIP ESCs transfected with indicated siRNAs. Scale bar is 20µm.

4.4.2 Genome-wide comparison of impaired differentiation in *Zfp281* and *Flcn* deficient cells

Although functionally independent, depletion of Zfp281 or activation of Tfe3 induce pronounced resistance to differentiation. To investigate whether these TFs control a similar set of target genes we compared their genome-wide DNA occupancy in naïve ESCs (Betschinger et al, 2013; Mayer et al, 2019) and observed very little overlap on chromatin, specifically at 9% of Tfe3 and 5% of Zfp281 peaks, respectively (Figure **8A**). Consistently, we were not able to detect a direct protein interaction between these TFs in IP-MS experiments (Mayer et al, 2019; Table **EV3**) (Villegas et al, 2019; Table **S1**). Zfp281 (67%) and Tfe3 (84%) bind preferentially distal to transcription start sites (TSSs) (Mayer et al, 2019; Figure **EV3G**) (Villegas et al, 2019; Figure **S2D, E**) however, the 1290 Zfp281/Tfe3 co-bound peaks are almost equally distributed between proximal (49%) and distal genomic regions (51%) (Figure **8B**). We showed that promoter-bound Tfe3 associates with coordinated lysosomal expression and regulation (CLEAR) target genes (Sardiello et al, 2009) (Villegas et al, 2019; Figure **2D, S2F**). Genes mapped by proximity to ChIP binding sites revealed that proximal but not distal Zfp281/Tfe3 co-bound peaks enrich more than expected by chance at CLEAR target genes however to a lesser extent than proximal Tfe3-only bound peaks (Figure **8C**). In contrast, proximal Zfp281-peaks are depleted for this gene set. In summary, overlap of Zfp281 and Tfe3 on chromatin is very sparse, but proximal Zfp281/Tfe3 co-bound genomic regions are enriched to a certain extent at CLEAR target genes.

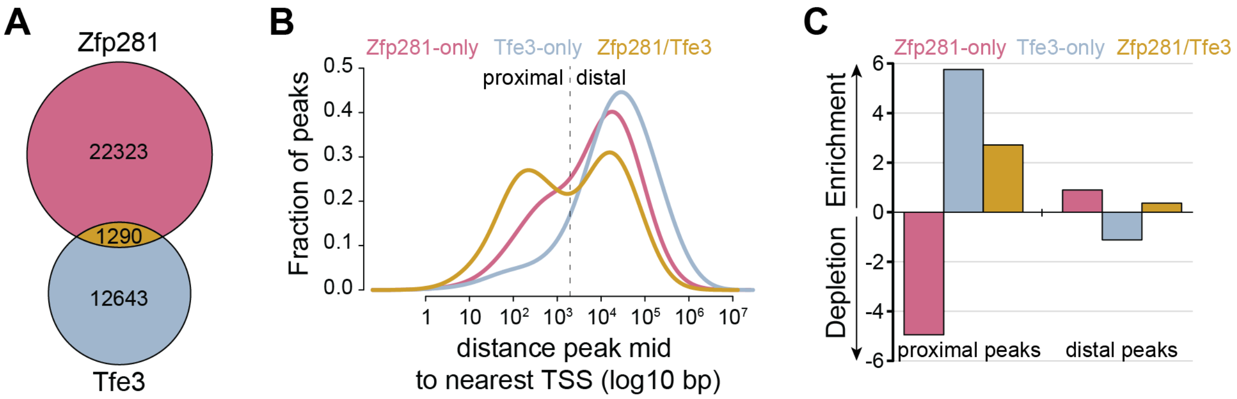


Figure 8: Comparison of DNA occupancy of Zfp281 and Tfe3.
(**A**) Overlap of Zfp281 and Tfe3 ChIP peaks. Zfp281-only (pink), Tfe3-only (blue) and Zfp281/Tfe3 co-bound peaks (yellow).
(**B**) Density plot showing log10 distance in base pairs (bp) of Zfp281-only (pink), Zic2-only (blue) and Zfp281/Zic2 co-bound peaks (yellow) to nearest transcription start site (TSS). Dashed line (2000 bp) indicates cut-off for proximal and distal peaks.
(**C**) Enrichment and depletion (standardized residuals) of CLEAR genes (Sardiello et al, 2009) in indicated proximal and distal peak selections.

Despite strong experimental and computational evidence that Zfp281 and Tfe3 independently regulate ESC differentiation, we considered the possibility that they control a generic transcriptional cell state progression program. We therefore, first, compared genes differentially expressed between *Zfp281* KO and *WT* (*Zfp281*^{2i}), and *Flcn* KO and *WT* (*Flcn*^{2i}) ESCs (Figure **9A**). k-means clustering of this combined gene set (total: 1704 genes) revealed shared up- and down-regulated transcripts (clusters 1 and 2), as well as Zfp281- and Flcn-specific target genes (clusters 3 and 4, and clusters 5 and 6, respectively). Consistent with previous analysis (Mayer et al, 2019; Figure **3B, C**), Zfp281 target genes in ESCs remain deregulated to a similar extent during *Zfp281* KO differentiation (*Zfp281*^{diff}). These genes are mostly unaffected after 32h of *WT* differentiation (*WT*^{32h}), however are differentially expressed in EpiSCs, notably with inverse directionality to *Zfp281* deficient cells (Figure **9A**). Surprisingly, Flcn target genes in ESCs (clusters 1, 2, 5 and 6) behave similarly, raising the possibility that this might be a common mechanism for cell state transition regulators to prime forthcoming developmental progression already at an early developmental stage. While shared and Zfp281-specific deregulated transcripts (clusters 1-4) are mainly enriched for generic developmental terms using gene ontology (GO) analysis, Flcn-specific up-regulated genes (cluster 5) enrich strongly for lysosomal terms (Figure **9B**). To assess whether co-binding of Zfp281 and Tfe3 on chromatin could account for common transcriptional targets, Zfp281/Tfe3 co-bound, Zfp281- and Tfe3-only peaks were mapped by proximity to genes in clusters 1-6 (Figure **9C**). Zfp281-only peaks were enriched for genes in clusters 2 and 4 which are transcriptionally down-regulated either common for *Zfp281* and *Tfe3* or specifically for *Zfp281* deficient cells, respectively, suggesting that Zfp281 might act predominantly as a transcriptional activator. In contrast, Tfe3-only peaks map preferentially to genes in cluster 5 which are up-regulated specifically in *Flcn* KO cells. Since these genes associate with lysosomal GO terms, CLEAR target genes were also strongly enriched in this gene cluster (Figure **9C**). Surprisingly, Zfp281/Tfe3 co-bound peaks associate equally with genes in clusters 4 and 5, which are mainly deregulated in either *Zfp281* or *Flcn* deficient cells, respectively, but not in the shared target gene clusters 1 and 2. Zfp281 and Flcn therefore control a common set of target genes in naïve ESCs, but independent of co-localizing on chromatin, suggesting discrete regulation.

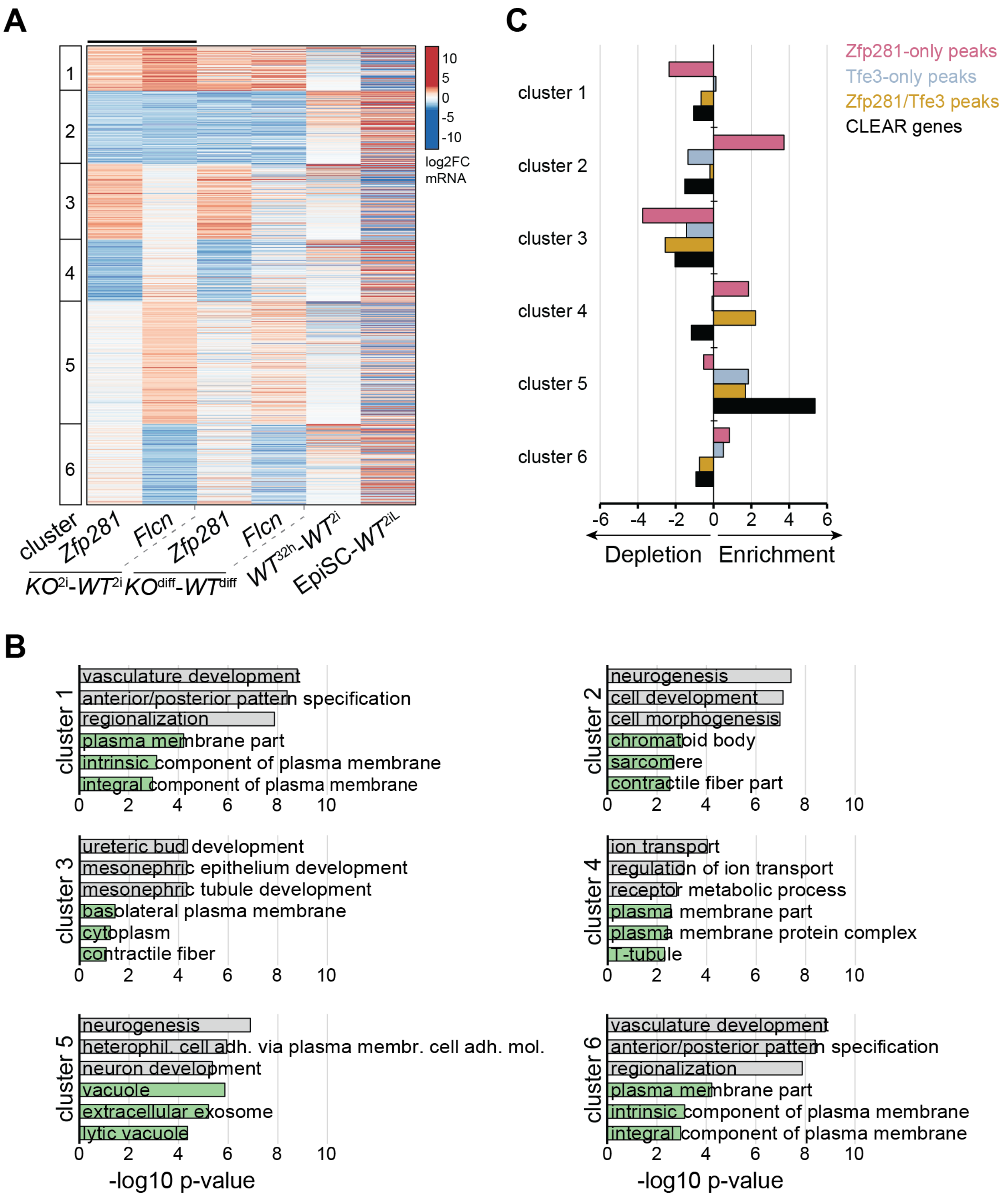


Figure 9: Transcriptional comparison of *Zfp281* and *Flcn* KO ESCs.
(**A**) Heatmap of mRNA log2 fold changes (log2FC) of *Zfp281* and *Flcn* KO ESCs and differentiated cells relative to respective *WT* controls, and *WT*^{32h} relative to *WT*^{2i} and EpiSCs relative to *WT*^{2iL} (Buecker et al, 2014; Factor et al, 2014) for genes significantly deregulated in *Zfp281*^{2i} and *Flcn*^{2i} (total: 1704 genes). *Zfp281*^{2i} and *Flcn*^{2i} samples were used for k-means clustering (black line on top of heatmap).
(**B**) Top 3 enriched GO terms for biological processes (grey) and cellular components (green) in heatmap clusters 1-6.
(**C**) Enrichment and depletion (standardized residuals) of genes associated with specified peak selections and CLEAR genes (Sardiello et al, 2009) in heatmap clusters 1-6.

To compare the function of Zfp281 and Tfe3 during early lineage progression, we used a similar approach however considering this time only genes which are significantly deregulated during *WT* differentiation (total: 1450 genes) (Figure **10A**). k-means clustering of genes limited to samples specific for *Zfp281* and *Flcn* KO differentiation (*Zfp281*$^{\text{diff}}$ and *Flcn*$^{\text{diff}}$, respectively) revealed 7 gene clusters (clusters 7-13) (Figure **10A**). Cluster 7 contains genes which are both up- and down-regulated in *WT*$^{\text{32h}}$ however remain unchanged in *Zfp281*$^{\text{diff}}$ and *Flcn*$^{\text{diff}}$. Clusters 8 and 9 are deregulated similarly between *Zfp281*$^{\text{diff}}$ and *Flcn*$^{\text{diff}}$, while clusters 10 and 11, and clusters 12 and 13 are specific for either *Zfp281*$^{\text{diff}}$ or *Flcn*$^{\text{diff}}$, respectively. In all cases these clusters are comprised of genes both up- and down-regulated in *WT*$^{\text{32h}}$, which show consistent regulation in EpiSCs (Figure **10A**). In line with previous analysis (Mayer et al, 2019; Figure **3B, C**), differentiation specific deregulated transcripts of *Zfp281* (clusters 8-11) showed minor mis-expression in *Zfp281*$^{\text{2i}}$ that became predominant only during differentiation. A similar pattern is observed in *Flcn* KO cells in clusters 8 and 9, and 12 and 13 (Figure **10A**), suggesting differentiation-specific gene regulation. Most gene clusters in differentiation (clusters 7, 9, 11, 12 and 13) are enriched for developmental associated GO terms, however clusters 8 and 12 are also enriched for vacuolar terms, which includes the lysosome (Figure **10B**). Proximal Zfp281/Tfe3 co-bound peaks are enriched at CLEAR target genes (Figure **8C**), suggesting that co-binding of these two TFs could explain shared transcriptional targets in cluster 8. Target genes mapped by proximity to Tfe3-only ChIP peaks revealed strong enrichment for clusters 8 and 12 genes, and weaker enrichment for genes belonging to cluster 10. Zfp281/Tfe3 co-bound peaks as well as CLEAR target genes associate predominantly with cluster 12 and CLEAR target genes to a lesser extent with cluster 8 (Figure **10C**). In contrast, Zfp281-only peaks were not more enriched at any gene clusters as expected by chance. This analysis therefore suggests that Zfp281 and Flcn share a set of common target genes during ESC differentiation, that includes lysosomal targets in cluster 8 which are strongly associated with Tfe3 binding to chromatin. However, Zfp281 and Flcn also regulate independent transcriptional targets which in the case of Flcn contain additional lysosomal genes (cluster 12) and are likely direct targets of Tfe3.

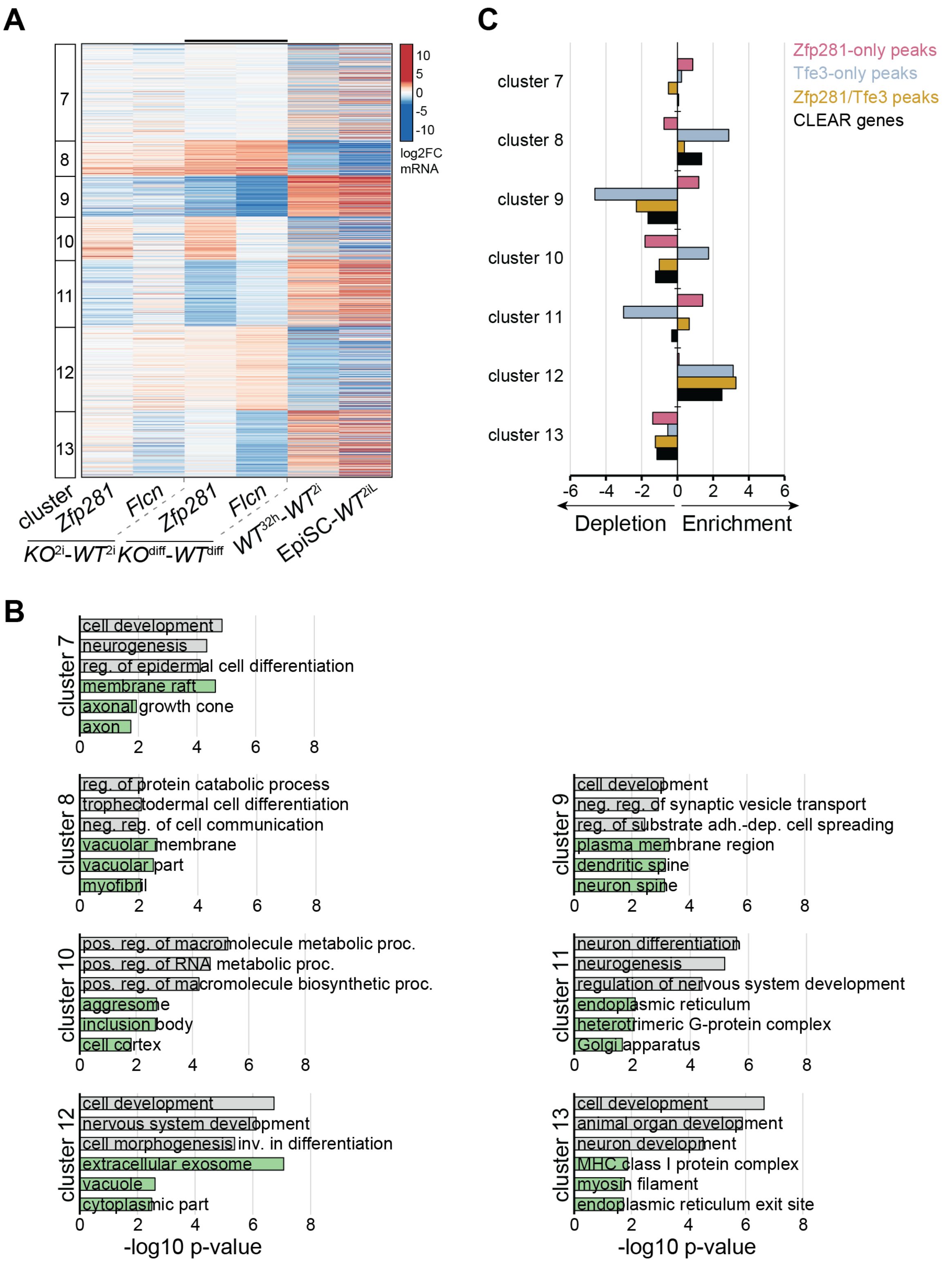


Figure 10: Transcriptional comparison of *Zfp281* and *Flcn* KO differentiation.
(**A**) Heatmap of mRNA log2FC of *Zfp281* and *Flcn* KO ESCs and differentiated cells relative to respective *WT* controls, and WT^{32h} relative to WT^{2i} and EpiSCs relative to WT^{2iL} (Buecker et al, 2014; Factor et al, 2014) for genes significantly deregulated in WT^{32h} relative to WT^{2i} (total: 1450 genes). $Zfp281^{diff}$ and $Flcn^{diff}$ samples were used for k-means clustering (black line on top of heatmap).
(**B**) Top 3 enriched GO terms for biological processes (grey) and cellular components (green) in heatmap clusters 7-13.
(**C**) Enrichment and depletion (standardized residuals) of genes associated with specified peak selections and CLEAR genes (Sardiello et al, 2009) in heatmap clusters 7-13.

4.5 Material and methods for additional results

Mouse ESCs and EpiSCs

Female RGd2 cells contain a Rex1:GFPd2-IRES-Blasticidin (Wray et al, 2011), male O4GIP ESCs contain a GFP-IRES-Puromycin transgene under control of an *Oct4* regulatory element (Betschinger et al, 2013) and *Zfp281* KO.1 cells are in female RGd2 background (Mayer et al, 2019).

O4GIP$^{\text{GY118F}}$ (Mayer et al, 2019) and 796.4 (Yang et al, 2010) EpiSCs contain a GFP-IRES-Puromycin transgene under control of an *Oct4* regulatory element (Guo et al, 2009) and express constitutively Stat3 activating receptor (GY118F) responsive to Gcsf (Yang et al, 2010). TNGA EpiSCs contain a GFP-IRES-Puromycin transgene under control of regulatory sequences of endogenous *Nanog* (Chambers et al, 2007). O4GIP$^{\text{ind. Esrrb}}$ and O4GIP$^{\text{ind. Klf4}}$ EpiSCs contain a conditionally expressed Esrrb or Klf4 transgene through addition of Dox (Mayer et al, 2019).

Cell Culture

Culturing of ESCs and EpiSCs, ESC differentiation, EpiSC reprogramming and siRNA transfection were performed as described (Mayer et al, 2019; Villegas et al, 2019). Following siRNAs (Qiagen) were used:

gene/control	catalog number
Alx1	GS216285
En2	GS13799
Flcn	GS216805
Fnip1	GS216742
Fnip2	GS329679
Foxd3	GS15221
Foxp1	GS108655
Hand1	GS15110
Jun	GS16476
Lamtor3	GS56692
negative control	1027281

gene/control	catalog number
Nkd2	GS72293
Oct6	GS18991
Otx2	GS18424
Sox1	GS20664
Tcf7l1	GS21415
Tfe3	GS209446
Tfeb	GS21425
Tsc2	GS22084
Utf1	GS22286
Zfp281	GS226442

Genome editing

CRISPR/Cas9 genome editing in RGd2 *WT* and *Zfp281* KO.1 ESCs was performed as described (Mayer et al, 2019). For generation of Otx2 deficient cell lines, two independent KO clones (specified in Figure **5B**) and, in the case of *Otx2* targeting in *WT* cells, one untargeted wildtype sibling clone were kept for further analysis. Following sgRNA sequences for deleting Exon 2 of *Otx2* were used:

sgRNA 1 (PAM) - targeting Otx2 Intron 1	**sgRNA 2** (PAM) - targeting Otx2 Intron 2
ATTTGTAGCTGGCGTCGGGGGG	GAACCAGGGTCTTTTAGCCTGG

Molecular biology

pPB-CAG-GY118F-pgk-hph (Yang et al, 2010) was stably integrated into TNGA EpiSCs after co-transfection with pBASE (Betschinger et al, 2013) and selection in the presence of 150µg/ml HygromycinB (Thermo Fisher).

Immunostaining

Immunostaining was performed as described (Villegas et al, 2019) using anti-Tfe3 (Sigma, HPA023881, 1:1000) antibody. DNA was visualized with Hoechst33342 (Life Technologies) and quantification of nucleo-cytoplasmic Tfe3 ratios was performed as described (Betschinger et al, 2013) for Flcn (97 cells), neg (198 cells), Tcf7l1 (208 cells) and Zfp281 (376 cells) siRNA transfected O4GIP ESCs.

Protein methods

Western blotting was performed as described (Mayer et al, 2019) using primary antibodies for anti-GAPDH (Sigma, G8795, 1:5000), anti-Otx2 (Abcam, ab21990, 1:500) and anti-Zfp281 (Bethyl Laboratories, A303-118A, 1:500).

ChIP and library preparation

ChIP for Otx2 (Abcam, ab21990, 2µg) and subsequent library preparation was performed as described (Mayer et al, 2019).

RNA-seq analysis

Published RNA-seq from *Zfp281* KO and corresponding *WT* ESCs in 2i and 32h after 2i withdrawal (Mayer et al, 2019), *Flcn* KO and corresponding *WT* ESCs in 2i and 34h after 2i withdrawal (Villegas et al, 2019), ESCs cultured in 2i/LIF (Buecker et al, 2014) and EpiSCs (Factor et al, 2014) were used. RNA-seq reads were aligned and quantified as described (Mayer et al, 2019; Villegas et al, 2019).

Transcripts were considered as expressed with at least 3 counts per million in at least two biological samples from *Zfp281* KO and corresponding *WT* cells (total: 13071 genes). Significantly deregulated genes were identified using edgeR (Robinson & Oshlack, 2010) (package version 3.24.0) and fitted to following generalized linear models:

~ *time* + *genotype* + *time:genotype*

This model was used separately for *Zfp281* and *Flcn* KO datasets including their corresponding *WT* samples. *time* refers to genes with a significant coefficient between WT^{2i} and WT^{32h} or WT^{34h} cells, genes with a significant *genotype* coefficient are differentially expressed between $Zfp281^{2i}$ or $Flcn^{2i}$ and WT^{2i} cells, and genes with a significant interaction term *time:genotype* are deregulated in *Zfp281* or *Flcn* KO cells specifically during differentiation.

Raw P values were corrected for multiple testing by calculating false discovery rates (FDR). Significant genes were identified as genes with an absolute log2FC greater than 1.0 and an FDR of less than 0.01.

Log2FC values obtained from edgeR were used for heatmap visualization. For ESC heatmap in Figure **9A**, only significantly deregulated genes in either $Zfp281^{2i}$ (781 genes) or $Flcn^{2i}$ (1167 genes) were considered (total: 1704 genes). For differentiation heatmap in Figure **10A**, only significantly deregulated genes between WT^{2i} and WT^{32h} were considered (total: 1450 genes). Heatmaps were generated using aheatmap function from the Bioconductor package NMF (Gaujoux & Seoighe, 2010) (package version 0.21.0). Analyses of enriched gene sets (Figure **9B**, **10B**) were performed using DAVID (Huang da et al, 2009) for GO terms of biological processes and cellular components. Differentially expressed genes upon transient TFEB overexpression in HeLa cells (CLEAR genes) (Sardiello et al, 2009) were converted to mouse homologs using HomoloGene (Geer et al, 2010) allowing only 1:1 mapping of genes.

ChIP-seq analysis

Published Zfp281 ChIP-seq data (Mayer et al, 2019) and Tfe3 ChIP-seq data (Betschinger et al, 2013) as well as one replicate each for Otx2 ChIP-seq in *WT*^{2i} and *WT*^{32h} were aligned, quantified and normalized as described (Mayer et al, 2019).

Zfp281, Otx2 and Tfe3 ChIP peaks were called using two replicates each of Zfp281 ChIP-seq reads in *WT*^{2i} and *WT*^{32h}, one replicate each of Otx2 ChIP-seq reads in *WT*^{2i} and *WT*^{32h} and one replicate each of Tfe3 ctrl and shFlcn ChIP-seq in *WT*^{2i}, respectively, using Macs2 (Zhang et al, 2008) (version 2.1.1.20160309) with default parameters. Peaks that were at least 2-fold enriched (IP over respective input for Zfp281 and Otx2 or IP over IgG for Tfe3 ChIP) in at least one of the four replicates for Zfp281 or in at least one of the two replicates for Otx2 or Tfe3 were considered (Zfp281: 23757 peaks, Otx2: 18207 peaks, Tfe3: 13513 peaks). To quantitatively compare Zfp281 and Otx2 chromatin binding, both peak sets were combined and overlapping peak regions were merged using the function reduce from Bioconductor package GenomicRanges (Lawrence et al, 2013) (package version 1.34.0). Combined peaks were classified into single- or co-bound as follows: Peaks that were enriched at least 2-fold (IP over respective input) in at least one of the four Zfp281 ChIP samples and in at least one of the two Otx2 ChIP samples were considered as co-bound (total: 2632 peaks), while if detected only in Zfp281 ChIP or only in Otx2 ChIP samples were considered as Zfp281-only (total: 20924 peaks) or Otx2-only (total: 15848 peaks), respectively. The same approach was used to determine overlapping chromatin binding of Zfp281 and Tfe3, revealing 1290 co-bound, 22323 Zfp281-only and 12643 Tfe3-only bound peaks.

Zfp281-only, Tfe3-only and Zfp281/Tfe3 co-bound peaks were assigned to genes by calculating the distance of the peak mid to nearest TSS using a set of non-redundant TSSs with a single start site randomly selected for each gene. Peaks were classified as proximal if the distance to the nearest TSS was less than 2000bp (Zfp281-only: 7052 peaks, Tfe3-only: 1645 peaks, Zfp281/Tfe3 co-bound: 633 peaks) and as distal otherwise (Zfp281-only: 15270 peaks, Tfe3-only: 10997 peaks, Zfp281/Tfe3 co-bound: 657 peaks).

Enrichment or depletion of CLEAR target genes (Sardiello et al, 2009) in proximal or distal Zfp281-only, Tfe3-only and Zfp281/Tfe3 co-bound peaks (Figure **8C**) and CLEAR target genes, Zfp281-only, Tfe3-only and Zfp281/Tfe3 co-bound peaks in gene clusters (Figure **9C**, **10C**) were quantified as standardized residuals by using the chisq.test function in R.

5. DISCUSSION

Developmental progression is an intrinsic and hard-wired cellular program which relies on stem and progenitor cells being able to sense external cues and accordingly activate differentiation programs. We have shown that metabolic signaling at the lysosomal membrane is essential for ESC progression by inactivating the TF Tfe3, which allows activation of differentiation and repression of pluripotency-associated transcriptional programs. Identification of TFE3 gain-of-function mutations in patients with a novel developmental disorders suggests that the ability of Tfe3 to repress lineage progression is operational throughout development (Villegas et al, 2019). Reversal of cellular specification can be achieved *in vivo* upon injury or *in vitro* via TF overexpression. We identified the bidirectionally acting TF Zfp281 to be required for the exit from and to inhibit the acquisition of naïve pluripotency, suggesting reversibility of developmental mechanisms. In contrast to classic lineage-instructive TFs, Zfp281 chromatin accessibility remains largely unchanged during this process and instead mediates the activation and recruitment of the histone methyltransferase Ehmt1 and the TF Zic2, respectively (Mayer et al, 2019). Detailed discussions for both studies can be found in the manuscripts in the appendix section.

In this discussion, I will focus first on additional results regarding a synergistic genetic interaction of Zfp281 with the homeobox TF Otx2 during pluripotent cell state transitions and how this could enhance our current understanding of the GRN acting during this process. In the second part, I will discuss the similarities and differences in the mode-of-action for the stem cell progression regulators Tfe3 and Zfp281. Finally, I will bring into perspective the role of TFs in regulating cellular plasticity to enable pluripotent cell state transition.

5.1 TF interplay in pluripotent cell state transitions

5.1.1 Otx2 synergistically interacts with Zfp281 in ESC differentiation and partially in reprogramming of EpiSCs

Given that the chromatin occupancy of Zfp281 remains largely static during ESC differentiation, we hypothesized that Zfp281 may interact with lineage-determining TFs which are usually characterized by upon activation, transcriptionally or via protein modification, instructing lineage progression. We therefore selected based on transcriptional activation and putative function in ESC differentiation (Betschinger et al, 2013; Leeb et al, 2014; Yang et al, 2012), 11 candidate lineage-determining TFs. Surprisingly, depletion of only one of those, the homeobox TF Otx2, strongly enhanced phenotypes induced by loss of Zfp281 in both ESC differentiation and EpiSC de-differentiation, but had little impact on its own. Otx2 has been identified as a regulator of EpiSC induction and maintenance (Acampora et al, 2013) by recruiting Oct4 to new enhancer sites (Buecker et al, 2014; Yang et al, 2014b). In fact, transcriptional induction of Otx2 during ESC differentiation is mediated by Oct4 itself (Williams et al, 2016; Yang et al, 2014b), suggesting a positive feedforward loop for robust cell fate transition. In the meantime, several other genome-wide loss-of-function screens for ESC progression have been performed (Hackett et al, 2018; Li et al, 2018; MacDougall et al, 2019a; Villegas et al, 2019). Including genes identified in these screens, such as Zfp217 or Zic3, into our candidate list of Zfp281 interactors may identify additional factors synergizing with Zfp281 in ESC differentiation and EpiSC reprogramming.

Loss of Otx2 in EpiSCs does not only lead to a destabilization of this cell state, manifested by derepression of Sox1 and neural differentiation, but also acquisition of naïve pluripotency features, such as an ESC-like morphology and upregulation of Nanog (Acampora et al, 2013), highlighting a putative function in reprogramming of EpiSCs. In our large-scale EpiSC de-differentiation screen, Otx2 scored just below the significant threshold (average Z-score: 1.98) and thus did not classify as a primary screen hit (Mayer et al, 2019; Table **EV1**).

However, depletion of Otx2 alone did not enhance reprogramming of EpiSCs (Figure **4B**, **5D**, **E**) which is likely due to the less potent reprogramming setup (2 instead of 4d of Gcsf treatment) that we employed in this experiments to detect alterations of the strong Zfp281 phenotype. Moreover, acquisition of pluripotency features in EpiSCs upon loss of Otx2 is only achieved after several passages and increases over time (Acampora et al, 2013). In contrast, our reprogramming setup involves acute siRNA knockdown. More specific experiments will be required to elucidate if and how Otx2 regulates reprogramming of EpiSCs. In contrast to validations in two independent ESC differentiation regimes, synergistic interaction of Zfp281 and Otx2 only persisted in one out of four tested EpiSC lines/reprogramming setups (Figure **4B**, **5D**, **E**). Whether this phenotypic discrepancy is due to the experimental setting, which might not allow us to see an enhancement of Zfp281 phenotype, or a cell-line dependent effect remains to be determined. Supporting the latter, depletion of Foxp1, which showed the strongest increase in Epi-iPSC colonies upon co-depletion with Zfp281 in 796.4 EpiSCs (Figure **4B**), in independent O4GIP^{GY118F} EpiSCs under the same condition used in Figure **5E** did indeed enhance Zfp281 reprogramming phenotype (data not shown).

5.1.2 Potential mechanisms of Zfp281 and Otx2 interaction

Removal of Otx2 on its own induces only minor defects in ESC differentiation, while co-depletion with Zfp281 showed a strong increase in retention of self-renewal, particularly in EpiLC differentiation (Figure **4A**, **5C**). This is consistent with the relatively mild transcriptional changes seen in differentiating *Otx2* KO cells (Buecker et al, 2014), suggesting a synergistic rather than additive interaction between Zfp281 and Otx2. This can be either achieved by co-regulating a common set of target genes, which are especially sensitive to the absence of both proteins, or by acting in independent pathways. Supporting the latter hypothesis, we do not detect a direct protein interaction of Zfp281 with Otx2 in WT^{2i} or WT^{40h} (Mayer et al, 2019; Table **EV3**) or in a published Zfp281 protein interactome in metastable S/L ESCs (Fidalgo et al, 2016). Moreover, ChIP-seq of these two TFs revealed only 15% overlap in chromatin localization (Figure **6B**). Fgf5, a bona fide early differentiation marker, is a well described downstream target gene of Otx2 with 5 binding sites in close proximity to its TSS (Buecker et al, 2014). While differentiation-induced transcriptional upregulation of Fgf5 is strongly perturbed in the absence of Otx2 (Acampora et al, 2013; Buecker et al, 2014), Fgf5 expression remains almost unaffected during *Zfp281* KO cell differentiation (Mayer et al, 2019; Figure **EV3A,** Table **EV2**). However, to validate a synergistic interaction of Zfp281 with Otx2, a comparison of transcriptional changes in individual and compound mutant cells would be required. Interestingly, for *Ehmt1/Zic2* compound KO cells, which also show increased resistance to differentiation compared to the individual mutants, transcriptional analysis revealed rather a reinforcement of already deregulated genes rather than two independent genetic programs (Mayer et al, 2019; Figure **6D**, **EV4E**).

5.1.3 Zfp281 and Otx2 as part of a larger GRN

Otx2 is required to redirect Oct4 to new enhancer sites during ESC differentiation (Buecker et al, 2014; Yang et al, 2014b). In cell state specific IP-MS for Oct4 not only Otx2 was identified to preferentially interact with Oct4 in the EpiLC state but surprisingly also Zfp281 and Zic2/3 (Buecker et al, 2014). Indeed, we also recovered a direct protein interaction between Zfp281 with Oct4 however without a preference for naïve ESCs or differentiated cells (Mayer et al, 2019; Table **EV3**). Although we only observed little co-binding of Zfp281 and Otx2 on chromatin, it is appealing to speculate that this could be achieved via Oct4 as a common protein interactor. Careful comparison of ChIP-seq data of these three TFs will allow to address this question. It has been shown that overexpression of Otx2 in naïve 2i/LIF condition is sufficient to induce differentiation and redirect Oct4 chromatin binding (Buecker et al, 2014; Yang et al, 2014b). However, absence of a strong differentiation phenotype in *Otx2* mutant cells suggests that other TFs can compensate for loss of Otx2. In light of these results, it would be interesting to investigate whether overexpression of Zfp281, which similarly to Otx2 induces

ESC differentiation (Mayer et al, 2019; Figure **2D**), is also sufficient to redirect Oct4 chromatin binding.

Where can we place Zic2 into this ESC differentiation specific GRN? Zic2 directly interacts with Zfp281 and Oct4 preferentially during differentiation (Buecker et al, 2014; Mayer et al, 2019). We showed that Zic2 is partially required for Zfp281-induced differentiation and that Zfp281 recruits and stabilizes Zic2 on chromatin during ESC progression. Similar to Otx2 and Zfp281, overexpression of Zic2 induces ESC differentiation even under highly restrictive naïve culture conditions (Mayer et al, 2019; Figure **EV4D**), however loss of Zic2 only had minor effects on gene expression during differentiation (Mayer et al, 2019; Figure **6**). Furthermore, *de novo* DNA motif analysis of top EpiLC specific Otx2 peaks revealed the Zic binding motif as the second most common one following only the Otx2 motif itself (Buecker et al, 2014). Although there is currently no IP-MS data in pluripotent cells available to validate a direct interaction of Otx2 and Zic2, one can speculate about a common mechanism of these two TFs. Indeed, transient expression of tagged Otx2, Zic2 and Oct4 in EpiSCs followed by ChIP-seq revealed preferential co-binding of Otx2 to chromatin with Zic2 rather than Oct4 in this developmentally more advanced cell state (Matsuda et al, 2017). Similarly to Otx2, Zic2 also binds all five enhancer sites in close proximity to the TSS of Fgf5 in EpiSCs (Matsuda et al, 2017) and expression of Fgf5 is drastically reduced upon Zic2 KO differentiation (Mayer et al, 2019; Table **EV2**). Although the transient transfection of tagged transgenes by Matsuda et al. may reveal chromatin binding features that are distinct to the endogenous TFs, these data suggest further dynamic changes in the GRN responsible for EpiSC progression. It would be interesting to compare ChIP-seq binding sites of Otx2 and Zic2 during ESC differentiation or assay potential changes in Oct4 binding upon Zic2 overexpression.

Taken together, these observations suggest a potential link between Zfp281, Otx2, Oct4 and Zic2 during ESC differentiation. However, more data, e.g. transcriptional consequences and effects on Oct4 chromatin binding upon combinatorial depletion of Otx2 and Zic2, are required to elucidate the dynamics of the GRN responsible for developmental progression. While Oct4, Otx2 and Zic2 dynamically change their DNA occupancy during ESC differentiation, this is not the case for Zfp281 whose binding remains relatively static, suggesting that Zfp281 might act as a molecular platform on DNA for other TFs and epigenetic regulators to enable fast and efficient remodeling of the chromatin landscape during development.

5.2 Transcriptional programs regulating ESC progression

We have identified and characterized the two TFs Tfe3 and Zfp281 as having key roles in transcriptional remodeling during ESC differentiation. While Flcn depletion results in a gain-of-function phenotype induced by nuclear Tfe3 accumulation, Zfp281 loss-of-function impairs exit from naïve pluripotency. Notably, Zfp281 regulates ESC progression independently from Tfe3 (Figure **7B**, **C**) and we did not detect peptides corresponding to the respective other TF in Zfp281 or Tfe3 IP-MS (Mayer et al, 2019; Table **EV3**) (Villegas et al, 2019; Table **S1**), suggesting no direct physical interaction. Absence of a protein interaction is also supported by low overlap of these two TFs on chromatin. Interestingly, roughly 50% of these Zfp281/Tfe3 co-bound peaks reside proximal to TSSs and show partial enrichment for CLEAR target genes (Figure **8B**, **C**). This is quite surprising since in general both TFs bind preferentially to distal genomic regions (Zfp281: 32% proximal and 68% distal, Tfe3: 15% proximal and 85% distal). Additionally, the binding motifs of Zfp281, CCCCTCCCCC (Mayer et al, 2019; Figure **EV3C**) and of Tfe3, TCACGTGAT (Betschinger et al, 2013), do not share any sequence similarities, therefore raising the possibility that co-binding at promoters could be rather by chance than dependent on each other. Together, these data strongly suggest that Zfp281 and the Flcn-Lamtor-Tfe3 axis impact ESC progression via two independent pathways, that may, however, converge on the same target genes. Indeed, in both naïve Zfp281 and Flcn mutant ESCs and early differentiation conditions, we observe a commonly deregulated gene set but also an equally large fraction of pathway-specific targets (Figure **9A**, **10A**), suggesting overlapping and independent functions.

In naïve ESCs, we showed that loss of Zfp281 results in strong transcriptional deregulation which persists throughout differentiation. Interestingly, these genes remain largely unchanged during early *WT* differentiation but show inverse deregulation in developmentally more advanced EpiSCs (Mayer et al, 2019; Figure **3B**). Surprisingly, we observed a similar behavior for the ESC-specific target genes of Tfe3 (Figure **9A**), suggesting that this might be a common mechanism of cell fate regulators to stabilize advanced developmental progression already at an earlier developmental stage. It would be interesting to investigate whether this hypothesis also applies to other differentiation-promoting pathways by re-analyzing published RNA-seq datasets for e.g. Foxd3 (Krishnakumar et al, 2016; Respuela et al, 2016), Otx2 (Buecker et al, 2014) or Tcf7l1 (Yi et al, 2011) deficient ESCs.

While in ESCs CLEAR target genes as well as GO terms for lysosomal/vacuolar function are enriched specifically in the Zfp281-independent/Flcn-dependent cluster 5, two separate clusters, 8 and 12, which are Zfp281-dependent and Zfp281-independent/Flcn-dependent, respectively, are enriched during differentiation (Figure **10C**). Zfp281 may therefore control cluster 8 lysosomal gene transcription during ESC progression. Absence of Flcn in ESCs does not only cause upregulation of lysosomal genes but also increased lysosomal proteolytic activity (Villegas et al, 2019; Figure **2A**, **S3C**). It is therefore appealing to speculate that loss of Zfp281 might cause a similar phenotype. In contrast to Tfe3, Zfp281 chromatin binding is not enriched at cluster 8 genes, indicating that this effect is consequential to deficient exit from pluripotency rather than directly mediated by Zfp281. Whether activation of lysosomal genes is a common and general consequence of impaired ESC differentiation remains to be determined.

Besides lysosomal genes, we also found several naïve pluripotency factors in cluster 8, e.g. Esrrb, Tbx3, Tfcp2l1 and Rex1, which are not efficiently repressed during differentiation in *Zfp281* and *Flcn* mutant cells. With the exception of Tbx3, Zfp281 and Tfe3 bind to the TSS of these genes. Cluster 9 genes, which show inefficient transcriptional upregulation in *Zfp281* and *Flcn* KO differentiation, are comprised of several differentiation-associated genes e.g. Oct6, Dnmt3b, Sox11 and Zic1. While Zfp281 binds close to the TSS of all these genes, Tfe3 binding is completely absent. This is consistent with depleted Tfe3 chromatin association at cluster 9 genes and enrichment for Zfp281 binding (Figure **10C**). Although we find several prominent genes being commonly deregulated between *Zfp281* and *Flcn* KO cells, it is rather unlikely that only a handful of genes cause such a strong differentiation defect. More likely an interplay between a larger set of factors, maybe only showing moderate deregulation, will regulate such a complex process. However, more transcriptional analysis comparing different lineage deficient mutants is required to draw stronger conclusion regarding a relevant transcriptional program for ESC progression.

While depletion of Zfp281 causes not only a deficiency in the exit from but also enhances the acquisition of naïve pluripotency (Mayer et al, 2019), neither ectopic nuclear Tfe3 (Betschinger et al, 2013) nor depletion of Flcn, Fnip1/2, Lamtor3 or Tsc2 resulted in increased numbers of Epi-iPSC colonies (Figure **7A**). In line with these results, none of the components of the Flcn-Lamtor-Tfe3 axis scored positively in our EpiSC reprogramming screen (Mayer et al, 2019; Table **EV1**). What discriminates Zfp281 from Tfe3? Nuclear induction of Tfe3 after exit from naïve pluripotency impaired proper acquisition of neuronal cell fates (Villegas et al, 2019; Figure **S7G**) which is consistent with gain-of-function mutations in human TFE3 causing a pleiotropic disorder with severe developmental defects (Villegas et al, 2019; Figure **7D, E, F**), suggesting that Tfe3 function is not limited to early development. Ectopic induction of nuclear Tfe3 enables binding of this TF to DNA, however whether Tfe3 has the capacity to bind to inaccessible chromatin in EpiSCs e.g. at pluripotency genes Esrrb, Tfcp2l1 and Rex1, remains unknown. It has been shown that Zfp281 depletion in EpiSCs results in cell cycle arrest and apoptosis (Fidalgo et al, 2016), suggesting a destabilization of the EpiSC state, thus making cells more vulnerable to external cues such as reprogramming stimuli. Depletion of Zfp281 will not only cause loss of associated binding partners, e.g. Zic2, but potentially also changes the cell-state specific chromatin environment.

Taken together, thorough computational comparison of two important pluripotent cell fate regulators, Zfp281 and Flcn, allowed us to speculate about general mechanisms of ESC progression, e.g. inverse transcriptional deregulation of late developmental genes already in ESCs and the lysosome as a potential central organelle in the exit from naïve pluripotency. Obviously, more data are required to validate these hypotheses and to further investigate their functional robustness and importance.

5.3 Conclusion and future perspective

Within this thesis, I discovered a bidirectional function for the TF Zfp281 in the exit from and acquisition of the naïve pluripotent state via recruitment and activation of the TF Zic2 and the histone methyltransferase Ehmt1. Moreover, I identified a synergistic genetic interaction of Zfp281 with the TF Otx2 most likely via Oct4. Although Zfp281 acts independent of the Flcn-Lamtor-Tfe3 axis during ESC differentiation, similarities in transcriptional deregulation indicate a general role for the lysosome as a check-point for developmental progression.

This combined study highlights the importance of TFs as key modulators of cell state transitions and underlines their diverse mechanisms that ensure robust and precise cellular programs. We can therefore classify them in: (a) lineage-instructive TFs characterized by enhanced activity or expression during cell state conversion e.g. Otx2, (b) lineage-repressive TFs which have to be inactivated to allow lineage progression e.g. Tfe3 and (c) lineage-permissive TFs that serve as stable molecular platforms on chromatin without changing their own activity e.g. Zfp281.

We know that there is a highly interconnected pluripotency GRN stabilizing the ESC state (Boyer et al, 2005; Kim et al, 2008; Loh et al, 2006), which needs to be dissolved and remodeled to allow developmental progression. While some components of this network are functionally required in both pluripotent and differentiated cells, e.g. Oct4 and Sox2, others only act cell state specifically, e.g. Tfe3 and Otx2. To further dissect and also better understand individual contributions to this network, combinatorial perturbations of different components have to be analyzed and computationally modelled (Dunn et al, 2014). Moreover, technical advances in single cell RNA-seq have enabled first insights in co-expression and co-regulation of key network genes in ESCs (Kolodziejczyk et al, 2015). Importantly, this GRN is not only a static construct but changes dynamically upon developmental progression. First attempts have been made to apply single cell analysis during ESC differentiation proposing a stochastic process of cell state transitions (Stumpf et al, 2017).

Can we ultimately use the knowledge of the GRN during developmental progression to enhance our understanding of cellular reprogramming? While ESC differentiation is a highly efficient process, de-differentiation is usually characterized by low reprogramming efficiency and accompanied by high cell death. Zfp281, as a novel bidirectional regulator of these processes, could serve as a prime examples to study the relationship of network dynamics in differentiation and de-differentiation.

6. REFERENCES

Acampora D, Di Giovannantonio LG, Simeone A (2013) Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition. *Development* **140:** 43-55

Acampora D, Mazan S, Lallemand Y, Avantaggiato V, Maury M, Simeone A, Brulet P (1995) Forebrain and midbrain regions are deleted in Otx2-/- mutants due to a defective anterior neuroectoderm specification during gastrulation. *Development* **121:** 3279-3290

Acloque H, Adams MS, Fishwick K, Bronner-Fraser M, Nieto MA (2009) Epithelial-mesenchymal transitions: the importance of changing cell state in development and disease. *The Journal of clinical investigation* **119:** 1438-1449

Ambrosetti DC, Basilico C, Dailey L (1997) Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites. *Molecular and cellular biology* **17:** 6321-6329

Ambrosetti DC, Scholer HR, Dailey L, Basilico C (2000) Modulation of the activity of multiple transcriptional activation domains by the DNA binding domains mediates the synergistic action of Sox2 and Oct-3 on the fibroblast growth factor-4 enhancer. *The Journal of biological chemistry* **275:** 23387-23397

Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R (2003) Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes & development* **17:** 126-140

Azuara V, Perry P, Sauer S, Spivakov M, Jorgensen HF, John RM, Gouti M, Casanova M, Warnes G, Merkenschlager M, Fisher AG (2006) Chromatin signatures of pluripotent cell lines. *Nature cell biology* **8:** 532-538

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125:** 315-326

Betschinger J, Nichols J, Dietmann S, Corrin PD, Paddison PJ, Smith A (2013) Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. *Cell* **153:** 335-347

Boroviak T, Loos R, Bertone P, Smith A, Nichols J (2014) The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nature cell biology* **16:** 516-528

Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, Gifford DK, Melton DA, Jaenisch R, Young RA (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122:** 947-956

Bradley A, Evans M, Kaufman MH, Robertson E (1984) Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* **309:** 255-256

Brons IG, Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA, Vallier L (2007) Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448:** 191-195

Buecker C, Srinivasan R, Wu Z, Calo E, Acampora D, Faial T, Simeone A, Tan M, Swigut T, Wysocka J (2014) Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell stem cell* **14:** 838-853

Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A (2003) Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113:** 643-655

Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B, Robertson M, Vrana J, Jones K, Grotewold L, Smith A (2007) Nanog safeguards pluripotency and mediates germline development. *Nature* **450:** 1230-1234

Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, Loh YH, Yeo HC, Yeo ZX, Narang V, Govindarajan KR, Leong B, Shahab A, Ruan Y, Bourque G, Sung WK, Clarke ND, Wei CL, Ng HH (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133:** 1106-1117

Chew JL, Loh YH, Zhang W, Chen X, Tam WL, Yeap LS, Li P, Ang YS, Lim B, Robson P, Ng HH (2005) Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Molecular and cellular biology* **25:** 6031-6046

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* **339:** 819-823

Davis RL, Weintraub H, Lassar AB (1987) Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51:** 987-1000

Dunn SJ, Martello G, Yordanov B, Emmott S, Smith AG (2014) Defining an essential transcription factor program for naive pluripotency. *Science* **344:** 1156-1160

Ebrahimi B (2015) Reprogramming barriers and enhancers: strategies to enhance the efficiency and kinetics of induced pluripotency. *Cell regeneration* **4:** 10

Elms P, Siggers P, Napper D, Greenfield A, Arkell R (2003) Zic2 is required for neural crest formation and hindbrain patterning during mouse development. *Developmental biology* **264:** 391-406

Evans MJ, Kaufman MH (1981) Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292:** 154-156

Factor DC, Corradin O, Zentner GE, Saiakhova A, Song L, Chenoweth JG, McKay RD, Crawford GE, Scacheri PC, Tesar PJ (2014) Epigenomic comparison reveals activation of "seed" enhancers during transition from naive to primed pluripotency. *Cell stem cell* **14:** 854-863

Festuccia N, Gonzalez I, Navarro P (2017) The Epigenetic Paradox of Pluripotent ES Cells. *Journal of molecular biology* **429:** 1476-1503

Festuccia N, Osorno R, Halbritter F, Karwacki-Neisius V, Navarro P, Colby D, Wong F, Yates A, Tomlinson SR, Chambers I (2012) Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. *Cell stem cell* **11:** 477-490

Ficz G, Hore TA, Santos F, Lee HJ, Dean W, Arand J, Krueger F, Oxley D, Paul YL, Walter J, Cook SJ, Andrews S, Branco MR, Reik W (2013) FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. *Cell stem cell* **13:** 351-359

Fidalgo M, Faiola F, Pereira CF, Ding J, Saunders A, Gingold J, Schaniel C, Lemischka IR, Silva JC, Wang J (2012) Zfp281 mediates Nanog autorepression through recruitment of the NuRD complex and inhibits somatic cell reprogramming. *Proceedings of the National Academy of Sciences of the United States of America* **109:** 16202-16207

Fidalgo M, Huang X, Guallar D, Sanchez-Priego C, Valdes VJ, Saunders A, Ding J, Wu WS, Clavel C, Wang J (2016) Zfp281 Coordinates Opposing Functions of Tet1 and Tet2 in Pluripotent States. *Cell stem cell* **19:** 355-369

Fidalgo M, Shekar PC, Ang YS, Fujiwara Y, Orkin SH, Wang J (2011) Zfp281 functions as a transcriptional repressor for pluripotency of mouse embryonic stem cells. *Stem cells* **29:** 1705-1716

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391:** 806-811

Gaujoux R, Seoighe C (2010) A flexible R package for nonnegative matrix factorization. *BMC bioinformatics* **11:** 367

Geer LY, Marchler-Bauer A, Geer RC, Han L, He J, He S, Liu C, Shi W, Bryant SH (2010) The NCBI BioSystems database. *Nucleic acids research* **38:** D492-496

Grabole N, Tischler J, Hackett JA, Kim S, Tang F, Leitch HG, Magnusdottir E, Surani MA (2013) Prdm14 promotes germline fate and naive pluripotency by repressing FGF signalling and DNA methylation. *EMBO reports* **14:** 629-637

Guo G, Huang Y, Humphreys P, Wang X, Smith A (2011) A PiggyBac-based recessive screening method to identify pluripotency regulators. *PloS one* **6:** e18189

Guo G, Yang J, Nichols J, Hall JS, Eyres I, Mansfield W, Smith A (2009) Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* **136:** 1063-1069

Gurdon JB (1962) The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. *Journal of embryology and experimental morphology* **10:** 622-640

Gurdon JB, Elsdale TR, Fischberg M (1958) Sexually mature individuals of Xenopus laevis from the transplantation of single somatic nuclei. *Nature* **182:** 64-65

Habibi E, Brinkman AB, Arand J, Kroeze LI, Kerstens HH, Matarese F, Lepikhov K, Gut M, Brun-Heath I, Hubner NC, Benedetti R, Altucci L, Jansen JH, Walter J, Gut IG, Marks H, Stunnenberg HG (2013) Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. *Cell stem cell* **13:** 360-369

Hackett JA, Huang Y, Gunesdogan U, Gretarsson KA, Kobayashi T, Surani MA (2018) Tracing the transitions from pluripotency to germ cell fate with CRISPR screening. *Nature communications* **9:** 4292

Hackett JA, Surani MA (2014) Regulatory principles of pluripotency: from the ground state up. *Cell stem cell* **15:** 416-430

Hall J, Guo G, Wray J, Eyres I, Nichols J, Grotewold L, Morfopoulou S, Humphreys P, Mansfield W, Walker R, Tomlinson S, Smith A (2009) Oct4 and LIF/Stat3 additively induce Kruppel factors to sustain embryonic stem cell self-renewal. *Cell stem cell* **5:** 597-609

Hart AH, Hartley L, Ibrahim M, Robb L (2004) Identification, cloning and expression analysis of the pluripotency promoting Nanog genes in mouse and human. *Developmental dynamics : an official publication of the American Association of Anatomists* **230:** 187-198

Hayashi K, de Sousa Lopes SMC, Tang F, Lao K, Surani MA (2008) Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell stem cell* **3:** 391-401

Hayashi K, Ohta H, Kurimoto K, Aramaki S, Saitou M (2011) Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146:** 519-532

Heng JC, Feng B, Han J, Jiang J, Kraus P, Ng JH, Orlov YL, Huss M, Yang L, Lufkin T, Lim B, Ng HH (2010) The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. *Cell stem cell* **6:** 167-174

Huang da W, Sherman BT, Lempicki RA (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic acids research* **37:** 1-13

Huang X, Balmer S, Yang F, Fidalgo M, Li D, Guallar D, Hadjantonakis AK, Wang J (2017) Zfp281 is essential for mouse epiblast maturation through transcriptional and epigenetic control of Nodal signaling. *eLife* **6**

Huang Y, Osorno R, Tsakiridis A, Wilson V (2012) In Vivo differentiation potential of epiblast stem cells revealed by chimeric embryo formation. *Cell reports* **2:** 1571-1578

Jessen KR, Mirsky R, Arthur-Farraj P (2015) The Role of Cell Plasticity in Tissue Repair: Adaptive Cellular Reprogramming. *Developmental cell* **34:** 613-620

Jiang J, Chan YS, Loh YH, Cai J, Tong GQ, Lim CA, Robson P, Zhong S, Ng HH (2008) A core Klf circuitry regulates self-renewal of embryonic stem cells. *Nature cell biology* **10:** 353-360

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337:** 816-821

Kalkan T, Olova N, Roode M, Mulas C, Lee HJ, Nett I, Marks H, Walker R, Stunnenberg HG, Lilley KS, Nichols J, Reik W, Bertone P, Smith A (2017) Tracking the embryonic stem cell transition from ground state pluripotency. *Development* **144:** 1221-1234

Keller G (2005) Embryonic stem cell differentiation: emergence of a new era in biology and medicine. *Genes & development* **19:** 1129-1155

Kim J, Chu J, Shen X, Wang J, Orkin SH (2008) An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* **132:** 1049-1061

Kojima Y, Kaufman-Francis K, Studdert JB, Steiner KA, Power MD, Loebel DA, Jones V, Hor A, de Alencastro G, Logan GJ, Teber ET, Tam OH, Stutz MD, Alexander IE, Pickett HA, Tam PP (2014) The transcriptional and functional properties of mouse epiblast stem cells resemble the anterior primitive streak. *Cell stem cell* **14:** 107-120

Kolodziejczyk AA, Kim JK, Tsang JC, Ilicic T, Henriksson J, Natarajan KN, Tuck AC, Gao X, Buhler M, Liu P, Marioni JC, Teichmann SA (2015) Single Cell RNA-Sequencing of Pluripotent States Unlocks Modular Transcriptional Variation. *Cell stem cell* **17:** 471-485

Krishnakumar R, Chen AF, Pantovich MG, Danial M, Parchem RJ, Labosky PA, Blelloch R (2016) FOXD3 Regulates Pluripotent Stem Cell Potential by Simultaneously Initiating and Repressing Enhancer Activity. *Cell stem cell* **18:** 104-117

Kunath T, Saba-El-Leil MK, Almousailleakh M, Wray J, Meloche S, Smith A (2007) FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development* **134:** 2895-2902

Lawrence M, Huber W, Pages H, Aboyoun P, Carlson M, Gentleman R, Morgan MT, Carey VJ (2013) Software for computing and annotating genomic ranges. *PLoS computational biology* **9:** e1003118

Lawson KA, Meneses JJ, Pedersen RA (1991) Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development* **113:** 891-911

Leeb M, Dietmann S, Paramor M, Niwa H, Smith A (2014) Genetic exploration of the exit from self-renewal using haploid embryonic stem cells. *Cell stem cell* **14:** 385-393

Leitch HG, McEwen KR, Turp A, Encheva V, Carroll T, Grabole N, Mansfield W, Nashun B, Knezovich JG, Smith A, Surani MA, Hajkova P (2013) Naive pluripotency is associated with global DNA hypomethylation. *Nature structural & molecular biology* **20:** 311-316

Li M, Yu JSL, Tilgner K, Ong SH, Koike-Yusa H, Yusa K (2018) Genome-wide CRISPR-KO Screen Uncovers mTORC1-Mediated Gsk3 Regulation in Naive Pluripotency Maintenance and Dissolution. *Cell reports* **24:** 489-502

Li R, Liang J, Ni S, Zhou T, Qing X, Li H, He W, Chen J, Li F, Zhuang Q, Qin B, Xu J, Li W, Yang J, Gan Y, Qin D, Feng S, Song H, Yang D, Zhang B, Zeng L, Lai L, Esteban MA, Pei D (2010) A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell stem cell* **7:** 51-63

Loh YH, Wu Q, Chew JL, Vega VB, Zhang W, Chen X, Bourque G, George J, Leong B, Liu J, Wong KY, Sung KW, Lee CW, Zhao XD, Chiu KP, Lipovich L, Kuznetsov VA, Robson P, Stanton LW, Wei CL, Ruan Y, Lim B, Ng HH (2006) The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nature genetics* **38:** 431-440

Luo Z, Gao X, Lin C, Smith ER, Marshall SA, Swanson SK, Florens L, Washburn MP, Shilatifard A (2015) Zic2 is an enhancer-binding factor required for embryonic stem cell specification. *Molecular cell* **57:** 685-694

MacDougall MS, Clarke R, Merrill BJ (2019a) Intracellular Ca(2+) Homeostasis and Nuclear Export Mediate Exit from Naive Pluripotency. *Cell stem cell*

MacDougall MS, Clarke R, Merrill BJ (2019b) Intracellular Ca(2+) Homeostasis and Nuclear Export Mediate Exit from Naive Pluripotency. *Cell stem cell* **25:** 210-224 e216

Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM (2013) RNA-guided human genome engineering via Cas9. *Science* **339:** 823-826

Marks H, Kalkan T, Menafra R, Denissov S, Jones K, Hofemeister H, Nichols J, Kranz A, Stewart AF, Smith A, Stunnenberg HG (2012) The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* **149:** 590-604

Marson A, Levine SS, Cole MF, Frampton GM, Brambrink T, Johnstone S, Guenther MG, Johnston WK, Wernig M, Newman J, Calabrese JM, Dennis LM, Volkert TL, Gupta S, Love J, Hannett N, Sharp PA, Bartel DP, Jaenisch R, Young RA (2008) Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* **134:** 521-533

Martello G, Bertone P, Smith A (2013) Identification of the missing pluripotency mediator downstream of leukaemia inhibitory factor. *The EMBO journal* **32:** 2561-2574

Martello G, Smith A (2014) The nature of embryonic stem cells. *Annual review of cell and developmental biology* **30:** 647-675

Martello G, Sugimoto T, Diamanti E, Joshi A, Hannah R, Ohtsuka S, Gottgens B, Niwa H, Smith A (2012) Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. *Cell stem cell* **11:** 491-504

Martin GR (1981) Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America* **78:** 7634-7638

Masui S, Nakatake Y, Toyooka Y, Shimosato D, Yagi R, Takahashi K, Okochi H, Okuda A, Matoba R, Sharov AA, Ko MS, Niwa H (2007) Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nature cell biology* **9:** 625-635

Masui S, Ohtsuka S, Yagi R, Takahashi K, Ko MS, Niwa H (2008) Rex1/Zfp42 is dispensable for pluripotency in mouse ES cells. *BMC developmental biology* **8:** 45

Matsuda K, Mikami T, Oki S, Iida H, Andrabi M, Boss JM, Yamaguchi K, Shigenobu S, Kondoh H (2017) ChIP-seq analysis of genomic binding regions of five major transcription factors highlights a central role for ZIC2 in the mouse epiblast stem cell gene regulatory network. *Development* **144:** 1948-1958

Matsuo I, Kuratani S, Kimura C, Takeda N, Aizawa S (1995) Mouse Otx2 functions in the formation and patterning of rostral head. *Genes & development* **9:** 2646-2658

Mayer D, Stadler MB, Rittirsch M, Hess D, Lukonin I, Winzi M, Smith A, Buchholz F, Betschinger J (2019) Zfp281 orchestrates interconversion of pluripotent states by engaging Ehmt1 and Zic2. *The EMBO journal*: e102591

Merrill BJ, Pasolli HA, Polak L, Rendl M, Garcia-Garcia MJ, Anderson KV, Fuchs E (2004) Tcf3: a transcriptional regulator of axis induction in the early embryo. *Development* **131:** 263-274

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES, Bernstein BE (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448:** 553-560

Mohammed H, Hernando-Herraez I, Savino A, Scialdone A, Macaulay I, Mulas C, Chandra T, Voet T, Dean W, Nichols J, Marioni JC, Reik W (2017) Single-Cell Landscape of Transcriptional Heterogeneity and Cell Fate Decisions during Mouse Early Gastrulation. *Cell reports* **20:** 1215-1228

Monk M, Adams RL, Rinaldi A (1991) Decrease in DNA methylase activity during preimplantation development in the mouse. *Development* **112:** 189-192

Monk M, Boubelik M, Lehnert S (1987) Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **99:** 371-382

Moris N, Pina C, Arias AM (2016) Transition states and cell fate decisions in epigenetic landscapes. *Nature reviews Genetics* **17:** 693-703

Mulas C, Chia G, Jones KA, Hodgson AC, Stirparo GG, Nichols J (2018) Oct4 regulates the embryonic axis and coordinates exit from pluripotency and germ layer specification in the mouse embryo. *Development* **145**

Mulas C, Kalkan T, Smith A (2017) NODAL Secures Pluripotency upon Embryonic Stem Cell Progression from the Ground State. *Stem cell reports* **9:** 77-91

Nakagawa M, Koyanagi M, Tanabe K, Takahashi K, Ichisaka T, Aoi T, Okita K, Mochiduki Y, Takizawa N, Yamanaka S (2008) Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nature biotechnology* **26:** 101-106

Nichols J, Smith A (2009) Naive and primed pluripotent states. *Cell stem cell* **4:** 487-492

Nichols J, Zevnik B, Anastassiadis K, Niwa H, Klewe-Nebenius D, Chambers I, Scholer H, Smith A (1998) Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95:** 379-391

Niwa H, Miyazaki J, Smith AG (2000) Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature genetics* **24:** 372-376

Niwa H, Ogawa K, Shimosato D, Adachi K (2009) A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature* **460:** 118-122

Ogawa K, Nishinakamura R, Iwamatsu Y, Shimosato D, Niwa H (2006) Synergistic action of Wnt and LIF in maintaining pluripotency of mouse ES cells. *Biochemical and biophysical research communications* **343:** 159-166

Ohnuki M, Takahashi K (2015) Present and future challenges of induced pluripotent stem cells. *Philosophical transactions of the Royal Society of London Series B, Biological sciences* **370:** 20140367

Okita K, Ichisaka T, Yamanaka S (2007) Generation of germline-competent induced pluripotent stem cells. *Nature* **448:** 313-317

Osorno R, Tsakiridis A, Wong F, Cambray N, Economou C, Wilkie R, Blin G, Scotting PJ, Chambers I, Wilson V (2012) The developmental dismantling of pluripotency is reversed by ectopic Oct4 expression. *Development* **139:** 2288-2298

Pardo M, Lang B, Yu L, Prosser H, Bradley A, Babu MM, Choudhary J (2010) An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. *Cell stem cell* **6:** 382-395

Pelton TA, Sharma S, Schulz TC, Rathjen J, Rathjen PD (2002) Transient pluripotent cell populations during primitive ectoderm formation: correlation of in vivo and in vitro pluripotent cell development. *Journal of cell science* **115:** 329-339

Pereira L, Yi F, Merrill BJ (2006) Repression of Nanog gene transcription by Tcf3 limits embryonic stem cell self-renewal. *Molecular and cellular biology* **26:** 7479-7491

Petit CS, Roczniak-Ferguson A, Ferguson SM (2013) Recruitment of folliculin to lysosomes supports the amino acid-dependent activation of Rag GTPases. *The Journal of cell biology* **202:** 1107-1122

Qin H, Diaz A, Blouin L, Lebbink RJ, Patena W, Tanbun P, LeProust EM, McManus MT, Song JS, Ramalho-Santos M (2014) Systematic identification of barriers to human iPSC generation. *Cell* **158:** 449-461

Raab S, Klingenstein M, Moller A, Illing A, Tosic J, Breunig M, Kuales G, Linta L, Seufferlein T, Arnold SJ, Kleger A, Liebau S (2017) Reprogramming to pluripotency does not require transition through a primitive streak-like state. *Scientific reports* **7:** 16543

Respuela P, Nikolic M, Tan M, Frommolt P, Zhao Y, Wysocka J, Rada-Iglesias A (2016) Foxd3 Promotes Exit from Naive Pluripotency through Enhancer Decommissioning and Inhibits Germline Specification. *Cell stem cell* **18:** 118-133

Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome biology* **11:** R25

Roczniak-Ferguson A, Petit CS, Froehlich F, Qian S, Ky J, Angarola B, Walther TC, Ferguson SM (2012) The transcription factor TFEB links mTORC1 signaling to transcriptional control of lysosome homeostasis. *Science signaling* **5:** ra42

Rosner MH, Vigano MA, Ozato K, Timmons PM, Poirier F, Rigby PW, Staudt LM (1990) A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. *Nature* **345:** 686-692

Rossant J, Tam PP (2009) Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. *Development* **136:** 701-713

Samavarchi-Tehrani P, Golipour A, David L, Sung HK, Beyer TA, Datti A, Woltjen K, Nagy A, Wrana JL (2010) Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell stem cell* **7:** 64-77

Sardiello M, Palmieri M, di Ronza A, Medina DL, Valenza M, Gennarino VA, Di Malta C, Donaudy F, Embrione V, Polishchuk RS, Banfi S, Parenti G, Cattaneo E, Ballabio A (2009) A gene network regulating lysosomal biogenesis and function. *Science* **325:** 473-477

Sato N, Meijer L, Skaltsounis L, Greengard P, Brivanlou AH (2004) Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nature medicine* **10:** 55-63

Saxton RA, Sabatini DM (2017) mTOR Signaling in Growth, Metabolism, and Disease. *Cell* **168:** 960-976

Scholer HR, Dressler GR, Balling R, Rohdewohld H, Gruss P (1990) Oct-4: a germline-specific transcription factor mapping to the mouse t-complex. *The EMBO journal* **9:** 2185-2195

Shinkai Y, Tachibana M (2011) H3K9 methyltransferase G9a and the related molecule GLP. *Genes & development* **25:** 781-788

Silva J, Nichols J, Theunissen TW, Guo G, van Oosten AL, Barrandon O, Wray J, Yamanaka S, Chambers I, Smith A (2009) Nanog is the gateway to the pluripotent ground state. *Cell* **138:** 722-737

Smith A (2017) Formative pluripotency: the executive phase in a developmental continuum. *Development* **144:** 365-373

Smith AG, Heath JK, Donaldson DD, Wong GG, Moreau J, Stahl M, Rogers D (1988) Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* **336:** 688-690

Stavridis MP, Lunn JS, Collins BJ, Storey KG (2007) A discrete period of FGF-induced Erk1/2 signalling is required for vertebrate neural specification. *Development* **134:** 2889-2894

Stuart HT, Stirparo GG, Lohoff T, Bates LE, Kinoshita M, Lim CY, Sousa EJ, Maskalenka K, Radzisheuskaya A, Malcolm AA, Alves MRP, Lloyd RL, Nestorowa S, Humphreys P, Mansfield W, Reik W, Bertone P, Nichols J, Gottgens B, Silva JCR (2019) Distinct Molecular Trajectories Converge to Induce Naive Pluripotency. *Cell stem cell* **25:** 388-406 e388

Stumpf PS, Smith RCG, Lenz M, Schuppert A, Muller FJ, Babtie A, Chan TE, Stumpf MPH, Please CP, Howison SD, Arai F, MacArthur BD (2017) Stem Cell Differentiation as a Non-Markov Stochastic Process. *Cell systems* **5:** 268-282 e267

Tachibana M, Sugimoto K, Nozaki M, Ueda J, Ohta T, Ohki M, Fukuda M, Takeda N, Niida H, Kato H, Shinkai Y (2002) G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes & development* **16:** 1779-1791

Tachibana M, Ueda J, Fukuda M, Takeda N, Ohta T, Iwanari H, Sakihama T, Kodama T, Hamakubo T, Shinkai Y (2005) Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes & development* **19:** 815-826

Tada M, Takahama Y, Abe K, Nakatsuji N, Tada T (2001) Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. *Current biology : CB* **11:** 1553-1558

Tai CI, Ying QL (2013) Gbx2, a LIF/Stat3 target, promotes reprogramming to and retention of the pluripotent ground state. *Journal of cell science* **126:** 1093-1098

Takagi N, Yoshida MA, Sugawara O, Sasaki M (1983) Reversal of X-inactivation in female mouse somatic cells hybridized with murine teratocarcinoma stem cells in vitro. *Cell* **34:** 1053-1062

Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126:** 663-676

Takahashi K, Yamanaka S (2015) A developmental framework for induced pluripotency. *Development* **142:** 3274-3285

Takahashi K, Yamanaka S (2016) A decade of transcription factor-mediated reprogramming to pluripotency. *Nature reviews Molecular cell biology* **17:** 183-193

Tee WW, Shen SS, Oksuz O, Narendra V, Reinberg D (2014) Erk1/2 activity promotes chromatin features and RNAPII phosphorylation at developmental promoters in mouse ESCs. *Cell* **156:** 678-690

Tesar PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, Gardner RL, McKay RD (2007) New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448:** 196-199

Thomson M, Liu SJ, Zou LN, Smith Z, Meissner A, Ramanathan S (2011) Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell* **145:** 875-889

Toyooka Y, Shimosato D, Murakami K, Takahashi K, Niwa H (2008) Identification and characterization of subpopulations in undifferentiated ES cell culture. *Development* **135:** 909-918

Tsakiridis A, Huang Y, Blin G, Skylaki S, Wymeersch F, Osorno R, Economou C, Karagianni E, Zhao S, Lowell S, Wilson V (2014) Distinct Wnt-driven primitive streak-like populations reflect in vivo lineage precursors. *Development* **141:** 1209-1221

van den Berg DL, Snoek T, Mullin NP, Yates A, Bezstarosti K, Demmers J, Chambers I, Poot RA (2010) An Oct4-centered protein interaction network in embryonic stem cells. *Cell stem cell* **6:** 369-381

van den Berg DL, Zhang W, Yates A, Engelen E, Takacs K, Bezstarosti K, Demmers J, Chambers I, Poot RA (2008) Estrogen-related receptor beta interacts with Oct4 to positively regulate Nanog gene expression. *Molecular and cellular biology* **28:** 5986-5995

Veillard AC, Marks H, Bernardo AS, Jouneau L, Laloe D, Boulanger L, Kaan A, Brochard V, Tosolini M, Pedersen R, Stunnenberg H, Jouneau A (2014) Stable methylation at promoters distinguishes epiblast stem cells from embryonic stem cells and the in vivo epiblasts. *Stem cells and development* **23:** 2014-2029

Villegas F, Lehalle D, Mayer D, Rittirsch M, Stadler MB, Zinner M, Olivieri D, Vabres P, Duplomb-Jego L, De Bont E, Duffourd Y, Duijkers F, Avila M, Genevieve D, Houcinat N, Jouan T, Kuentz P, Lichtenbelt KD, Thauvin-Robinet C, St-Onge J, Thevenon J, van Gassen KLI, van Haelst M, van Koningsbruggen S, Hess D, Smallwood SA, Riviere JB, Faivre L, Betschinger J (2019) Lysosomal Signaling Licenses Embryonic Stem Cell Differentiation via Inactivation of Tfe3. *Cell stem cell* **24:** 257-270 e258

Waddington CH (1957) *The Strategy of Genes. A Discussion of Some Aspects of Theoretical Biology*: Allen & Unwin.

Wernig M, Meissner A, Foreman R, Brambrink T, Ku M, Hochedlinger K, Bernstein BE, Jaenisch R (2007) In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* **448:** 318-324

Wiedenheft B, Sternberg SH, Doudna JA (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **482:** 331-338

Williams EO, Taylor AK, Bell EL, Lim R, Kim DM, Guarente L (2016) Sirtuin 1 Promotes Deacetylation of Oct4 and Maintenance of Naive Pluripotency. *Cell reports* **17:** 809-820

Williams RL, Hilton DJ, Pease S, Willson TA, Stewart CL, Gearing DP, Wagner EF, Metcalf D, Nicola NA, Gough NM (1988) Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* **336:** 684-687

Wray J, Kalkan T, Gomez-Lopez S, Eckardt D, Cook A, Kemler R, Smith A (2011) Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nature cell biology* **13:** 838-845

Wray J, Kalkan T, Smith AG (2010) The ground state of pluripotency. *Biochemical Society transactions* **38:** 1027-1032

Yamaji M, Ueda J, Hayashi K, Ohta H, Yabuta Y, Kurimoto K, Nakato R, Yamada Y, Shirahige K, Saitou M (2013) PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. *Cell stem cell* **12:** 368-382

Yang CS, Chang KY, Rana TM (2014a) Genome-wide functional analysis reveals factors needed at the transition steps of induced reprogramming. *Cell reports* **8:** 327-337

Yang J, van Oosten AL, Theunissen TW, Guo G, Silva JC, Smith A (2010) Stat3 activation is limiting for reprogramming to ground state pluripotency. *Cell stem cell* **7:** 319-328

Yang SH, Kalkan T, Morissroe C, Marks H, Stunnenberg H, Smith A, Sharrocks AD (2014b) Otx2 and Oct4 drive early enhancer activation during embryonic stem cell transition from naive pluripotency. *Cell reports* **7:** 1968-1981

Yang SH, Kalkan T, Morrisroe C, Smith A, Sharrocks AD (2012) A genome-wide RNAi screen reveals MAP kinase phosphatases as key ERK pathway regulators during embryonic stem cell differentiation. *PLoS genetics* **8:** e1003112

Ye S, Li P, Tong C, Ying QL (2013) Embryonic stem cell self-renewal pathways converge on the transcription factor Tfcp2l1. *The EMBO journal* **32:** 2548-2560

Yeo JC, Jiang J, Tan ZY, Yim GR, Ng JH, Goke J, Kraus P, Liang H, Gonzales KA, Chong HC, Tan CP, Lim YS, Tan NS, Lufkin T, Ng HH (2014) Klf2 is an essential factor that sustains ground state pluripotency. *Cell stem cell* **14:** 864-872

Yi F, Pereira L, Hoffman JA, Shy BR, Yuen CM, Liu DR, Merrill BJ (2011) Opposing effects of Tcf3 and Tcf1 control Wnt stimulation of embryonic stem cell self-renewal. *Nature cell biology* **13:** 762-770

Ying QL, Nichols J, Chambers I, Smith A (2003) BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* **115:** 281-292

Ying QL, Wray J, Nichols J, Batlle-Morera L, Doble B, Woodgett J, Cohen P, Smith A (2008) The ground state of embryonic stem cell self-renewal. *Nature* **453:** 519-523

Zhang K, Li L, Huang C, Shen C, Tan F, Xia C, Liu P, Rossant J, Jing N (2010) Distinct functions of BMP4 during different stages of mouse ES cell neural commitment. *Development* **137:** 2095-2105

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS (2008) Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9:** R137

Zylicz JJ, Dietmann S, Gunesdogan U, Hackett JA, Cougot D, Lee C, Surani MA (2015) Chromatin dynamics and the role of G9a in gene regulation and enhancer silencing during early mouse development. *eLife* **4**

7. ACKNOWLEDGEMENTS

First and foremost, I would like to thank Jörg for hiring me as his first PhD student in the lab, for showing me how to do a Western blot, for inspiring discussion about my own project but also science in general, for challenging me constantly and for his endless support throughout the years inside and outside the lab. I would like to further thank my PhD committee member Susan, Michael and Christa for valuable feedback and support.

I would also like to thank all former and current members of the Betschinger lab for always finding our way out from the room escapes and great input for the project. I want to especially thank Mel for her experimental support during my project, I would have not managed to genotype all my clones without her. I am also very grateful to Jeff's, Dirk's, Antoine's, Prisca's, Luca's and Charisos' lab for various combination of joint lab meetings and feedback for my project. A special thanks goes to Antoine's group, Paul, Leslie, Philip, Matyas, Fabio, Marietta and Ilya for sharing reagents and protocols, troubleshooting my ChIP, and imaging the small wheels and raspberries. This work would have not been possible without the amazing support from the FMI facilities: Hubertus for endless cell sorts and fun conversations about music, Laurent and Steve for explaining me how to use the microscope at least five times during my PhD, the whole sequencing facility for preparing more than hundred libraries for me, Daniel who literally saved my project by performing a novel protein digestion method followed by mass spec analysis and of course Michael for teaching me everything I know about bioinformatics, for spending hours over hours writing scripts and discussing my projects, and for being extremely supportive over all the years.

Of course a huge thanks goes to “the office” for creating a really fun work environment and for tolerating my grumpy mood days. Zhan was an amazing help with my millions of “short” bioinformatics questions and for sharing lots of scripts, Verena for being a great help with color-related questions (palevioletred3) and my postdoc applications, Marco for being the best coffee buddy, and of course Franka for being an incredible inspiration and mentor! I also want to thank everyone else whom I have met during my time in Basel. I have made so many great friends over the last 6 years. A very special thank goes to Conny, Mel and Tina for becoming incredibly important in my life, all the fun we had together and for being extremely supportive.

Zum Schluss möchte ich mich noch bei meinen liebsten drei Mädels für eure tolle Unterstützung in den letzten Jahren bedanken, dass ihr euch so oft auf den Weg nach Basel gemacht habt und für die wunderbare Zeit, die wir gemeinsam verbracht haben! Ein großer Dank gilt auch meiner Familie, für die vielen Telefonate und die aufmunternden Worte in den schwierigen Zeiten. Diese Arbeit wäre nicht möglich gewesen ohne dich, Philipp. Danke dass du immer für mich da bist, nach Basel gezogen bist um bei mir zu sein und mir auch ein Leben außerhalb der Wissenschaft zeigst! Ich freue mich schon auf unsere nächsten Abenteuer!

8. APPENDICES

8.1 Abbreviations

bHLH	basic helix-loop-helix
Bmp	bone morphogenetic protein
bp	base pairs
Cas9	clustered regularly interspaced short palindromic repeats-associated 9
ChIP	chromatin immunoprecipitation
Chiron	Chir99021
CLEAR	coordinated lysosomal expression and regulation
CRISPR	clustered regularly interspaced short palindromic repeats
d	days
DNA	deoxyribonucleic acid
dox	doxycycline
E	embryonic day
EB	embryoid body
Epi	epiblast
Epi-iPSC	epiblast stem cell derived induced pluripotent stem cell
EpiLC	epiblast-like cell
EpiSC	epiblast stem cell
Erk	extracellular signal-regulated kinase
ESC	embryonic stem cell
FA	bFGF and ActivinA
FAK	bFGF, ActivinA and knockout serum replacement
FC	fold change
Flcn	Folliculin
Gcsf	granulocyte colony stimulating factor

GFP	green fluorescent protein
GO	gene ontology
GRN	gene regulatory network
Gsk	glycogen synthase kinase
h	hours
H3K4me3	trimethylation at histone H3 lysine 4
H3K9me1	monomethylation at histone H3 lysine 9
H3K9me2	dimethylation at histone H3 lysine 9
H3K27me3	trimethylation at histone H3 lysine 27
ICM	inner cell mass
Id	Inhibitor of Differentiation
IP-MS	immunoprecipitation coupled to semi-quantitative mass spectrometry
iPSC	induced pluripotent stem cell
Jak	Janus kinase
KO	knockout
LIF	leukemia inhibitory factor
Mapk	mitogen-activated protein kinase
Mek	mitogen-activated protein kinase kinases
MiTF	microphthalmia-associated transcription factor
mTORC	mechanistic target of rapamycin complex
mRNA	messenger ribonucleic acid
Myod	myoblast determination protein
neg	negative
n.s.	not significant
OSKM	Oct3/4, Sox2, Klf4 and Myc
PD03	PD0325901
PrE	primitive endoderm
PS	primitive streak

R	Pearson’s correlation
RGd2	Rex1:GFPd2-IRES-Blasticidin
RNA	ribonucleic acid
RNAi	ribonucleic acid interference
SD	standard deviation
seq	sequencing
sgRNA	single guide ribonucleic acid
siRNA	small interfering ribonucleic acid
S/L	Serum/Lif
Stat	signal transducer and activator of transcription
TE	trophectoderm
TF	transcriptiona factor
Tsc	tuberous sclerosis complex
TSS	transcription start site
WT	wildtype

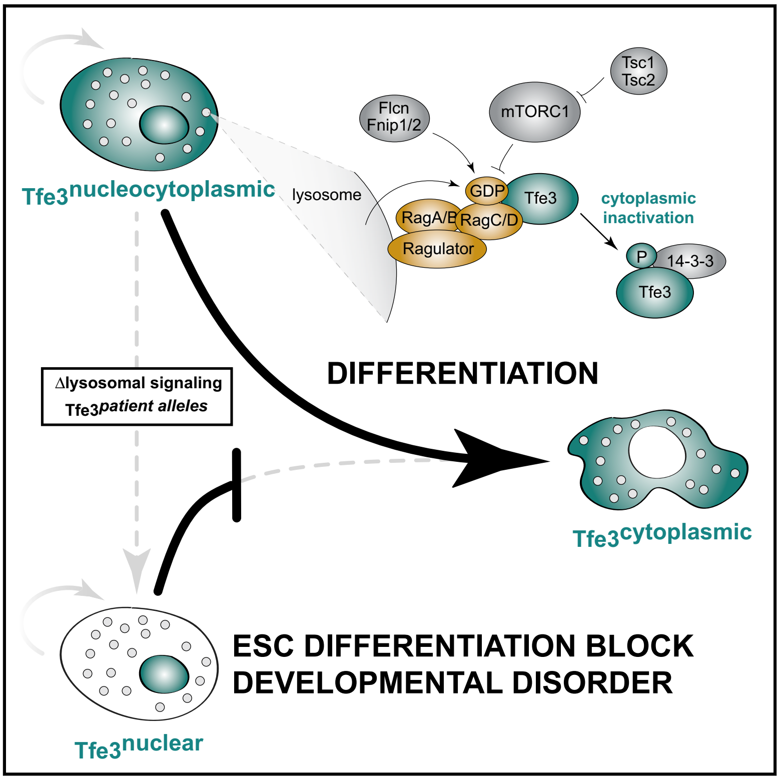
8.2 Manuscripts

Article

Cell Stem Cell

Lysosomal Signaling Licenses Embryonic Stem Cell Differentiation via Inactivation of Tfe3

Graphical Abstract



Highlights

- Genome-wide CRISPR/Cas9 screen for differentiation resistance in mouse ESCs
- Lysosomal Rag GTPase signaling inactivates Tfe3 to license exit from self-renewal
- Rag GTPase regulation in steady-state cells and starvation is distinct
- Tfe3 inactivation mutations found in a human mosaic developmental disorder

Authors

Florian Villegas, Daphné Lehalle,
Daniela Mayer, ...,
Jean-Baptiste Rivière,
Laurence Faivre, Joerg Betschinger

Correspondence

joerg.betschinger@fmi.ch

In Brief

Villegas et al. identify mouse embryonic stem cell differentiation drivers in a genome-wide CRISPR/Cas9 screen. The majority of these are part of a lysosomal signaling pathway that licenses differentiation by inactivating the transcription factor Tfe3. The authors discover lysosomal-signaling-insensitive Tfe3 mutations as potentially causal for a human developmental disorder.

Villegas et al., 2019, Cell Stem Cell *24*, 257–270
February 7, 2019 © 2018 Elsevier Inc.
https://doi.org/10.1016/j.stem.2018.11.021

Cell Stem Cell
Article

Lysosomal Signaling Licenses Embryonic Stem Cell Differentiation via Inactivation of Tfe3

Florian Villegas,[1,2,13] Daphné Lehalle,[3,4,13] Daniela Mayer,[1,2,13] Melanie Rittirsch,[1] Michael B. Stadler,[1,5] Marietta Zinner,[1,2] Daniel Olivieri,[1] Pierre Vabres,[3,4,6] Laurence Duplomb-Jego,[4] Eveline S.J.M. De Bont,[7] Yannis Duffourd,[4] Floor Duijkers,[8] Magali Avila,[4] David Geneviève,[9] Nada Houcinat,[3,4] Thibaud Jouan,[4] Paul Kuentz,[3,4] Klaske D. Lichtenbelt,[10] Christel Thauvin-Robinet,[3,4] Judith St-Onge,[4,11] Julien Thevenon,[3,4] Koen L.I. van Gassen,[10] Mieke van Haelst,[8] Silvana van Koningsbruggen,[8] Daniel Hess,[1] Sebastien A. Smallwood,[1] Jean-Baptiste Rivière,[4,11,12] Laurence Faivre,[3,4] and Joerg Betschinger[1,14,*]

[1]Friedrich Miescher Institute for Biomedical Research, 4058 Basel, Switzerland
[2]Faculty of Sciences, University of Basel, 4003 Basel, Switzerland
[3]Fédération Hospitalo-Universitaire Médecine Translationnelle et Anomalies du Développement (TRANSLAD), Centre Hospitalier Universitaire Dijon et Université de Bourgogne, 21079 Dijon, France
[4]Equipe GAD, INSERM LNC UMR 1231, Faculté de Médecine, Université de Bourgogne Franche-Comté, Dijon, France
[5]Swiss Institute of Bioinformatics, 4058 Basel, Switzerland
[6]Département de Dermatologie, CHU Dijon, Dijon, France
[7]Department of Pediatric Oncology/Hematology, Beatrix Children's Hospital, University Medical Centre Groningen, Groningen, the Netherlands
[8]Department of Clinical Genetics, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands
[9]Department of Clinical Genetics, University Medical Centre Montpellier, Montpellier, France
[10]Department of Genetics, University Medical Center Utrecht (UMCU), Utrecht, the Netherlands
[11]Child Health and Human Development Program, Research Institute of the McGill University Health Centre, Montreal, QC H4A 3J1, Canada
[12]Department of Human Genetics, Faculty of Medicine, McGill University, Montreal, QC H3A 1B1, Canada
[13]These authors contributed equally
[14]Lead Contact
*Correspondence: joerg.betschinger@fmi.ch
https://doi.org/10.1016/j.stem.2018.11.021

SUMMARY

Self-renewal and differentiation of pluripotent murine embryonic stem cells (ESCs) is regulated by extrinsic signaling pathways. It is less clear whether cellular metabolism instructs developmental progression. In an unbiased genome-wide CRISPR/Cas9 screen, we identified components of a conserved amino-acid-sensing pathway as critical drivers of ESC differentiation. Functional analysis revealed that lysosome activity, the Ragulator protein complex, and the tumor-suppressor protein Folliculin enable the Rag GTPases C and D to bind and seclude the bHLH transcription factor Tfe3 in the cytoplasm. In contrast, ectopic nuclear Tfe3 represses specific developmental and metabolic transcriptional programs that are associated with peri-implantation development. We show differentiation-specific and non-canonical regulation of Rag GTPase in ESCs and, importantly, identify point mutations in a Tfe3 domain required for cytoplasmic inactivation as potentially causal for a human developmental disorder. Our work reveals an instructive and biomedically relevant role of metabolic signaling in licensing embryonic cell fate transitions.

INTRODUCTION

Pluripotency emerges in the mouse blastocyst at embryonic day (E) 3.75 and declines upon implantation. Pausing developmental progression of the pre-implantation epiblast by chemical inhibition of glycogen synthetase kinase 3 (GSK3) and fibroblast growth factor (FGF) signaling (2i) allows stabilization of naive pluripotency in self-renewing embryonic stem cells (ESCs) *in vitro* (Smith, 2017). Besides FGF4-elicited mitogen-activated protein kinase signaling and GSK3-dependent disinhibition of the transcriptional repressor Tcf7l1, multiple additional transcriptional and post-transcriptional mechanisms enforce loss of ESC identity (Betschinger, 2017). Among those are metabolic pathways that provide cofactors for chromatin-modifying enzymes and contribute to differentiation-associated epigenetic changes (Zhang et al., 2018). Whether these instruct or permissively facilitate extinction of the ESC state is unclear. Furthermore, perturbation of respective metabolites delays differentiation (Hwang et al., 2016; Moussaieff et al., 2015), suggesting a fine balancing function rather than an essential requirement. In contrast, depletion of the tumor suppressors *Folliculin* (*Flcn*) and *Tuberous sclerosis complex 1* or *2* (*Tsc1/2*) impairs exit from ESC self-renewal upstream of nucleocytoplasmic distribution of the basic-helix-loop-helix transcription factor (TF) Tfe3 (Betschinger et al., 2013). Flcn, Tsc, and Tfe3 also function in signaling of the mechanistic target of rapamycin (mTOR) complex 1 (mTORC1), indicating that metabolism controls developmental progression of ESCs.

The mTORC1 protein kinase complex coordinates cellular growth in response to environmental conditions (Saxton and Sabatini, 2017; Figure S1A). Besides enabling protein, lipid, and nucleotide synthesis, mTORC1 suppresses catabolic processes such as autophagy. This is done in part by phosphorylating members of the MiTF/Tfe family of TFs to which Tfe3 belongs (Raben and Puertollano, 2016). Phosphorylation induces binding to 14-3-3 proteins, leading to cytoplasmic sequestration and inactivation. When mTORC1 is inactive, these TFs translocate into the nucleus and induce both autophagy and lysosome target genes. This provides a transcriptional feedback mechanism during cellular starvation that enhances the production of biosynthetic precursors through protein degradation.

Flcn and Tsc proteins act in pathways that signal nutritional stimuli to mTORC1 (Saxton and Sabatini, 2017). Tsc1 and 2 are part of the heterotrimeric tuberous sclerosis protein complex, which is inhibited by extrinsic growth factors and acts as a negative regulator of the mTORC1 activator Rheb (Figure S1A). Flcn functions in a parallel signaling branch that conveys amino acid (aa) levels to Rag guanosine triphosphatases (GTPases). Rag GTPases form obligate heterodimers in which RagA or RagB pairs with RagC or RagD subunits and that are recruited to the lysosomal membrane by the pentameric Ragulator protein complex (Bar-Peled et al., 2012; Sancak et al., 2010). Binding of guanosine triphosphate (GTP) by the RagA/B subunit and of guanosine diphosphate (GDP) by the RagC/D subunit induces a conformation that interacts with and activates mTORC1 at lysosomes (Saxton and Sabatini, 2017). The nucleotide state of Rag GTPases is controlled by GTPase-activating proteins (GAPs) and guanine nucleotide exchange factors (GEFs) that act downstream of aa-sensing protein complexes; sensors of cytosolic aas converge on inhibition of the RagA/B GAP complex GATOR1, whereas Slc38a9 and the RagA/B GEF activity of Ragulator are downstream of intralysosomal aas (Saxton and Sabatini, 2017). Flcn, in contrast, interacts with Flcn-interacting proteins (Fnip) 1 and 2 and is a RagC/D GAP that facilitates activation of the Rag heterodimer and, consequentially, mTORC1 in response to aas (Meng and Ferguson, 2018; Petit et al., 2013; Tsun et al., 2013).

The function of Flcn and Tsc in ESC differentiation therefore poses a conundrum; Flcn is an activator and Tsc an inhibitor of mTORC1, but depletion of either protein impairs exit from self-renewal upstream of nuclear Tfe3 (Betschinger et al., 2013; Li et al., 2018). Canonical mTORC1 signaling is therefore unlikely to instruct differentiation. Consistent with this, chemical inhibition of mTORC1 in ESCs compromises proliferation and translational regulation but not extinction of naive pluripotency (Betschinger et al., 2013; Gangloff et al., 2004; Murakami et al., 2004; Sampath et al., 2008). It is therefore unclear how Flcn and Tsc drive ESC differentiation. Furthermore, it remains to be determined how Tfe3 prevents resolution of the pluripotent transcriptional circuitry and whether its anti-differentiation function is restricted to murine pluripotent stem cells.

Here we apply a genome-wide functional CRISPR/Cas9 screen to identify genes that are essential for exit from ESC self-renewal. We show that the majority of these, including *Flcn*, are part of a signaling pathway that activates RagC/D at the lysosome, which, in turn, induces cytoplasmic Tfe3 inactivation. In contrast to canonical aa sensing, this function of Rag GTPases is mTORC-independent. In fact, mTORC1 antagonizes RagC/D activation in ESCs, providing a mechanistic explanation for why Flcn and Tsc proteins act antonymously in mTORC1 regulation but synonymously in differentiation. We demonstrate that the catabolic activity of lysosomes controls differentiation through RagC/D and Tfe3, providing a direct mechanistic link between ESC catabolism and transcriptional control of developmental progression. Furthermore, we identify a Tfe3 protein region that is essential for lysosome-mediated inactivation. This part of the protein is frequently deleted in *TFE3* translocation-induced human renal cell carcinomas and mutated in a human mosaic developmental disorder, indicating a conserved role of lysosomal signaling in cell state transitions.

RESULTS

Ragulator Drives ESC Differentiation Upstream of Subcellular Tfe3 Localization

Release from 2i induces exit from ESC self-renewal and terminal differentiation (Kalkan et al., 2017). To identify the molecular mechanisms essential for this process, we screened for genes whose deletion would block differentiation. Toward this goal, we infected Cas9-expressing RGd2 ESCs with a genome-wide lentiviral guide RNA (gRNA) library (Koike-Yusa et al., 2014) and continuously passaged these cells in the absence of 2i. To enrich for differentiation-resistant mutant cells, we concomitantly selected for expression of the Rex1 (Zfp42) reporter in RGd2 cells, which marks self-renewing ESCs (Kalkan et al., 2017; Figure 1A). This screen was performed independently in two female ESC clones and identified 19 high-confidence target genes, including *Flcn* (Figure 1B; Table S1). We validated the differentiation resistance phenotype for 15 of those in a male RGd2 ESC line that was transiently transfected with Cas9 and individual gRNAs (Table S3). To determine genes that act upstream of subcellular Tfe3 distribution, we analyzed Tfe3 localization in these (Figure 1C). Similar to differentiated wild-type cells (Figure S1H), seven of the differentiation-resistant mutants showed cytoplasmic Tfe3 accumulation. These contained members of the pathways inhibited by 2i: the GSK3-signaling components APC and Tcf7l1 and the FGF receptor signal transducers Ptpn11 and Grb2. In contrast, Tfe3 was nuclear in the absence of Flcn and the Ragulator subunits Lamtor1, 2, 4, and 5 and nucleo-cytoplasmic in the absence of Tsc1 and 2. Consistent with a predominant role of Tfe3 distribution in exit from ESC self-renewal, we also identified two gRNAs targeting exon 4 of *Tfe3* that generated Tfe3 alleles with ectopic nuclear localization. Similar roles in differentiation and aa sensing suggest that Ragulator and Flcn function together in exit from self-renewal.

We first confirmed the role of Lamtor1, 2, and 3 in an independent assay (Betschinger et al., 2013). Their knockdown in Oct4GIP ESCs by small interfering RNA (siRNA) transfection induced nuclear enrichment of Tfe3 and impaired differentiation, similar to absence of Flcn and the redundantly acting Folliculin-interacting proteins Fnip1 and 2 (Figures S1B–S1D; Betschinger et al., 2013). To test for Tfe3 dependency, we co-depleted MiTF/Tfe family members expressed in ESCs (Figure S1E). Removal of Tfe3, but not Tfeb or MiTF, reverted differentiation defects caused by knockdown of Flcn and Lamtor1 and 2 but not Tcf7l1 (Figure 1D), demonstrating a specific requirement for Tfe3

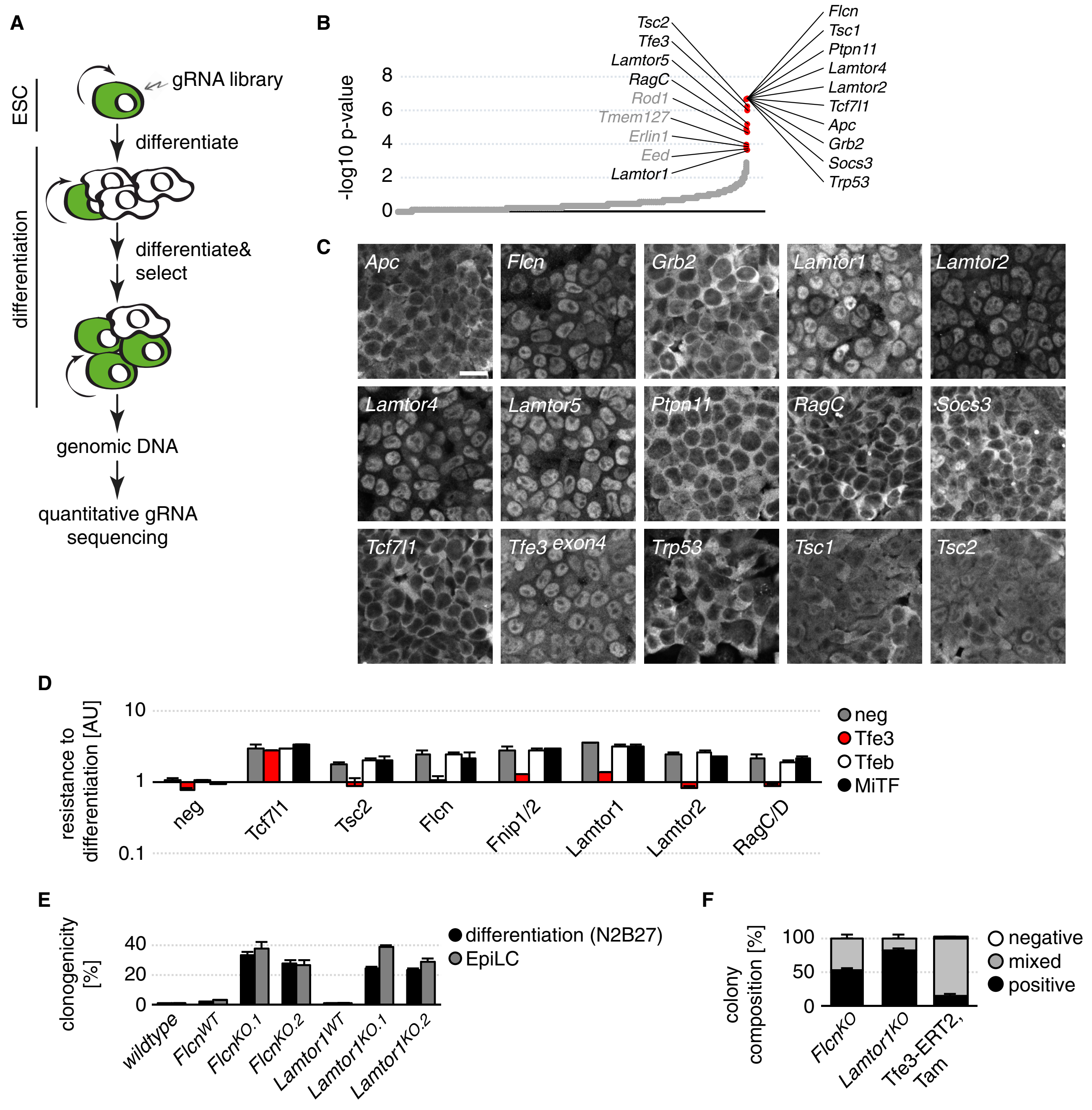


Figure 1. Identification of Differentiation Drivers in a Genome-wide CRISPR/Cas9 Screen

(A) Schematic outline of the screening procedure. Green indicates undifferentiated cells.

(B) Screen results ordered by statistical significance. Genes labeled in red were retested, of which genes annotated in black were validated.

(C) Tfe3 localization in RGd2 cells transiently transfected with Cas9 and gRNAs targeting the indicated genes after 3–5 passages in the absence of 2i. Note that control cells are lost as expected during the procedure because of terminal differentiation. Scale bar is 20 μm.

(D) Resistance to differentiation of O4GIP ESCs transfected with the indicated siRNA combinations after 3 days of 2i withdrawal. Results were normalized to negative (neg) siRNA cells and are shown on a log(10)-scaled axis. Average and standard deviation (SD) of 2 technical replicates are shown.

(E) Retention of self-renewal after 2 days of 2i withdrawal under the indicated conditions. Average and SD of 3 biological replicates are shown.

(F) Alkaline phosphatase composition of colonies derived from the indicated long-term differentiated cell lines subjected to clonal self-renewal in 2i. Average and SD of 2 biological replicates are shown.

downstream of these genes. For further analysis, we generated isogenic *Flcn* and *Lamtor1* knockout (KO) ESC clones (Figures S1F and S1G). These maintained nuclear Tfe3 localization (Figure S1H) and failed to exit self-renewal under two different differentiation regimes: removal of 2i and the strongly inductive epiblast-like cell (EpiLC) differentiation (Hayashi et al., 2011;

Figure 1E). Both phenotypes were rescued by Flcn and Lamtor1 transgene expression in the respective mutants (Figures 3B and 3C). Long-term culture of *Flcn* and *Lamtor1* KO ESCs in the absence of 2i gave rise to cell populations retaining self-renewal, which is similar to induction of ectopic nuclear Tfe3 in 4-hydroxytamoxifen (Tam)-treated Tfe3-ERT2 expressing cells (Figure 1F; Betschinger et al., 2013). In summary, Ragulator and Flcn function upstream of Tfe3 and are required for ESC differentiation.

The Flcn-Ragulator-Tfe3 Axis Regulates Lysosomal and Developmental Transcription

To gain insight into Flcn and Ragulator function in ESCs, we compared the transcriptomes of two independent *Flcn* and *Lamtor1* KO cell lines and their respective controls (Table S1). Transcriptional deregulation in mutant ESCs was strongly correlated (Pearson correlation coefficient [R] = 0.70) (Figure S2A), suggesting that both genes act in the same pathway. To explore the relation with Tfe3 activity, we also profiled Tfe3-ERT2-expressing ESCs after 3 hr and 34 hr of Tam induction. k-means clustering (Figures 2A and S2B) of deregulated transcripts identified groups of early (cluster 1), late, and indirect Tfe3 target genes (clusters 2–6) and Tfe3-independent alterations in established *Flcn* and *Lamtor1* KO cell lines (clusters 7 and 8). Comparison with published datasets of blastocyst development (Boroviak et al., 2015) showed, in particular upon 34 hr Tfe3 activation in Tfe3-ERT2 ESCs, a correlation with induction of diapause, an embryonic state of suspended development (Figure S2A). Although this correlation is not strong (R = 0.22), it is comparable with reported diapause—mimicking ESC regimes using chemical inhibitors (Bulut-Karslioglu et al., 2016; Scognamiglio et al., 2016).

To directly explore the role of Tfe3 in differentiation, we sequenced RNA of control, *Flcn* KO, and Tam-induced Tfe3-ERT2 cells after release from 2i (Table S1). Principal-component analysis (PCA) separated control ESCs and the E4.5 epiblast (Boroviak et al., 2015) from 34-hr differentiated control cells, EpiLCs (Buecker et al., 2014), and the E5.5 epiblast along PC2 (Figure 2B). In contrast, 34-hr differentiated *Flcn* KO and Tfe3-induced cells as well as their long-term differentiated progeny, including *Lamtor1* KO cells, clustered close to wild-type ESCs, suggesting transcriptome-wide impairment of peri-implantation programs. k-means clustering (Figure 2C) of differentiation-specific transcriptional changes identified insensitive genes (clusters 9 and 10), transiently induced Tfe3 targets (cluster 11) and two clusters of genes (clusters 12 and 13) whose dynamic behavior during differentiation was significantly (data not shown) dampened by genetic perturbations. Gene ontology (GO) analysis revealed enrichment of developmental and lysosomal regulators in clusters 12 and 13, respectively (Figure 2C; Table S1). Deregulation of these two clusters was already detectable in *Flcn* and *Lamtor1* KO ESCs and Tfe3-ERT2 ESCs induced for 34 hr. This is, however, not causal to aberrant transcription during differentiation because Tam was added to Tfe3-ERT2 cells only upon release from 2i. To test whether the induction of lysosomal regulators has a functional consequence, we exposed ESCs to DQ-BSA, a substrate that fluoresces after proteolysis in lysosomes (Figure S2C). This revealed an increase in lysosomal proteolytic activity in *Flcn* (+40%) and *Lamtor1* (+57%) KO and Tfe3-ERT2 Tam-treated (+61%) cells. The opposite was observed in the presence of Vacuolin-1 (−33%), a compound that compromises delivery and maturation of lysosomal enzymes (Sano et al., 2016). When Tfe3 was acutely induced, we also observed an expansion of the LysoTracker-labeled endolysosomal compartment. Tfe3 overactivation therefore increases the activity of lysosomes in *Flcn*- and *Lamtor1*-mutant ESCs.

Alterations of mRNA levels in clusters 12 and 13 were by and large undiscernible in Tfe3-ERT2 ESCs induced for 3 hr (Figure 2C), indicating that they are late or indirect Tfe3 targets. To discriminate between these possibilities, we analyzed genome-wide Tfe3 occupancy in ESCs (Betschinger et al., 2013). Tfe3 binds predominantly distal of transcriptional start sites (TSSs), but the overlap with enhancer features is scarce (2,387 putative enhancer within 11,459 total peaks) (Figures S2D and S2E). Target genes mapped by proximity to all distal binding sites (3,996 genes) or the 2,387 putative enhancers (1,440 genes) are predominantly enriched at Tfe3-insensitive genes in cluster 10 (Figure 2D). In contrast, the majority of Tfe3-bound proximal sites are associated with marks of active transcription and enriched at Tfe3-sensitive genes in clusters 11 and 13 (Figures 2D and S2E). Tfe3, thus, antagonizes differentiation-specific transcription by directly activating promoters of cluster 13 targets and indirectly repressing expression of cluster 12 genes. Promoter-driven induction of lysosomal regulators is reminiscent of MiTF/Tfe family TFs driving the CLEAR (coordinated lysosomal expression and regulation) gene network in somatic cells (Sardiello et al., 2009). CLEAR genes are indeed enriched in clusters 11 and 13 and also mirror the Tfe3 binding preference to promoters of genes in clusters 1–8 (Figures 2D and S2F). This suggests that upregulation of at least a subset of direct metabolic targets in ESCs is a cell-type-independent response to Tfe3 activation. Taken together these findings indicate that disinhibition of Tfe3 in the absence of Flcn and the Ragulator subunit Lamtor1 triggers a specific transcriptional cascade that simultaneously augments cellular catabolism and inhibits forthcoming developmental transcription.

Flcn-Dependent RagC/D Activation Controls Exit from ESC Self-Renewal

Ragulator and Flcn-Fnip are required for activation of RagA/B and RagC/D, respectively (Figure S1A). As has been described for somatic cells, RagC, Lamtor1 (Sancak et al., 2010), and a subfraction of endogenously tagged Flcn (Martina et al., 2014) colocalize with lysosome-associated membrane protein 1 (LAMP1) in ESCs (Figure 3A). Lysosomal RagC localization requires Lamtor1, 2, and 3 but not Flcn (Figure S3A), corroborating the previously reported Ragulator-dependent Rag GTPase recruitment (Sancak et al., 2010). To probe their role in exit from self-renewal, we depleted Rag GTPases in O4GIP ESCs by siRNA transfection (Figure S3B). The combinatorial knockdown of RagA/B or RagC/D caused nuclear Tfe3 accumulation and Tfe3-dependent differentiation impairment (Figures 1D and S1B–S1D). Rag GTPases therefore act redundantly and, similar to Ragulator and Flcn, drive developmental progression upstream of Tfe3. We note that our primary screen identified RagC (Figure 1B) and that individual knockdown of RagC induced resistance to differentiation, although less pronounced than together with RagD (Figure S1B). Although we did not observe

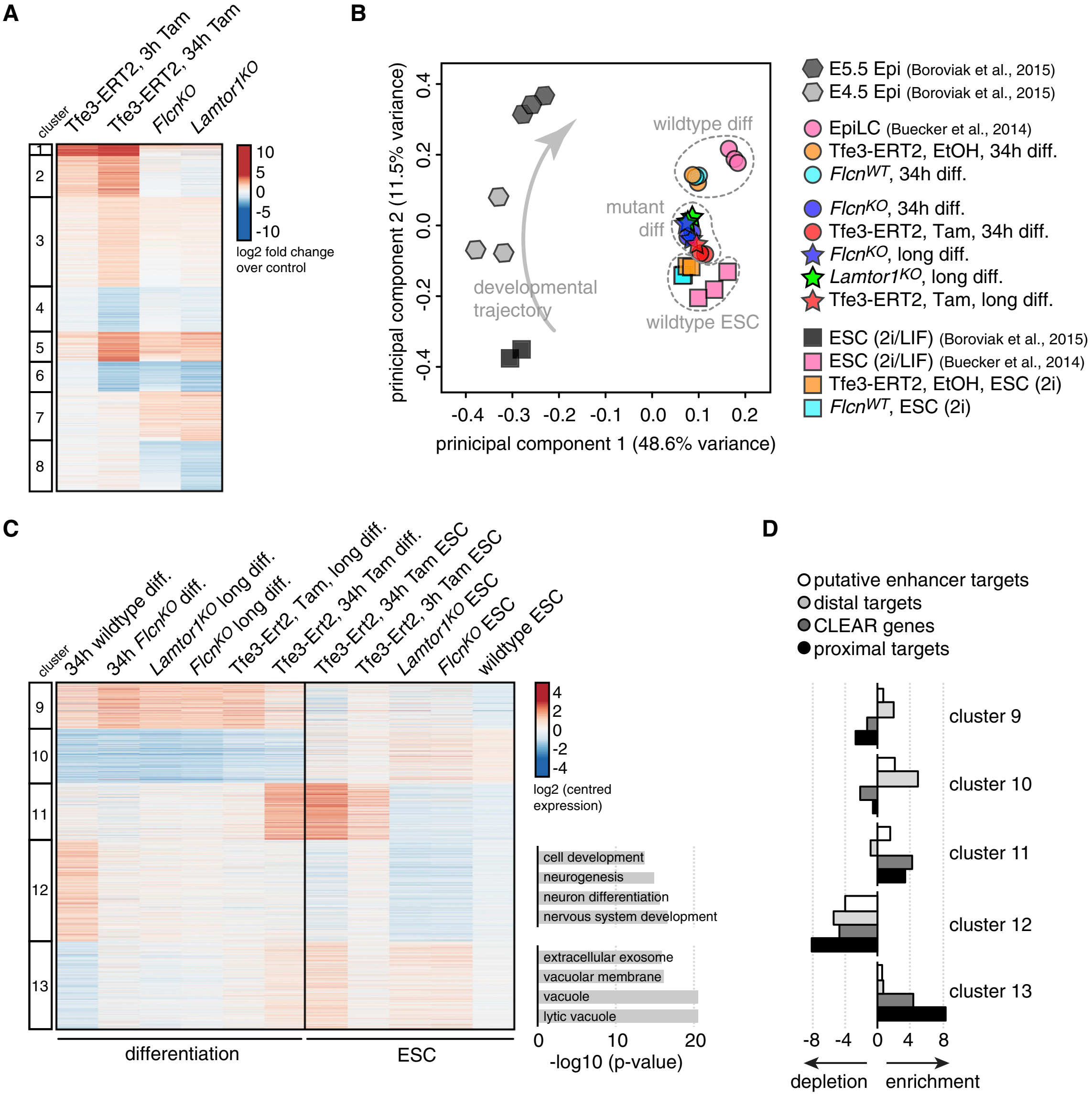


Figure 2. Flcn-Lamtor1-Tfe3 Control Differentiation-Specific Transcription

(A) k-means clustering of log2 fold changes in genes significantly deregulated in ESCs of the indicated genotypes.

(B) PCA analysis of the indicated samples. Note that PC1 distinguishes experimental variations between studies.

(C) k-means clustering of centered expression (log2 fold changes relative to gene average) in ESCs and during differentiation of the indicated genotypes. Genes dynamically regulated during wild-type ESC differentiation and specifically deregulated during differentiation of Tfe3-ERT2 Tam-treated and *Flcn* knockout (KO) cells were considered. The top five GO term enrichments of gene clusters 12 and 13 are shown. Samples (columns) were ordered by hierarchical clustering.

(D) Enrichment and depletion (standardized residuals) of genes associated with proximal and distal Tfe3 peaks and CLEAR genes (Sardiello et al., 2009) in the indicated clusters.

significant nuclear Tfe3 enrichment in cells without RagC (Figures 1C and S1D), co-depletion of Tfe3 restored exit from self-renewal (Figure S3C), suggesting subtle but functionally relevant Tfe3 mislocalization.

We next tested whether Flcn and Ragulator act upstream of Rag GTPases by using mutants preferentially binding GDP (RagB/C/D^{GDP}) or encoding for GTPase-inactive alleles (RagB/C/D^{GTP}) (Tsun et al., 2013). Overexpression of RagC/D^{GDP} reverted nuclear Tfe3 localization and differentiation impairment in *Flcn* KO cells (Figures S3D and S3E), whereas RagC/D^{GTP} or any nucleotide-loading RagB mutant did not, suggesting that Flcn acts as a GAP for RagC/D in ESCs. To determine the role of Ragulator in activating RagA/B, we exploited the notion that Lamtor1 anchors the Ragulator-Rag GTPase complex at the lysosome (Bar-Peled et al., 2012; Nada et al., 2009). We reasoned that cytoplasmic delocalization of Rag GTPases in the absence of Lamtor1 (Figure S3A) would allow us to assemble lysosomal signaling by targeting Rag GTPases ectopically to the lysosomal membrane. We therefore engineered hybrid proteins in which the N-terminal lysosomal targeting sequence of

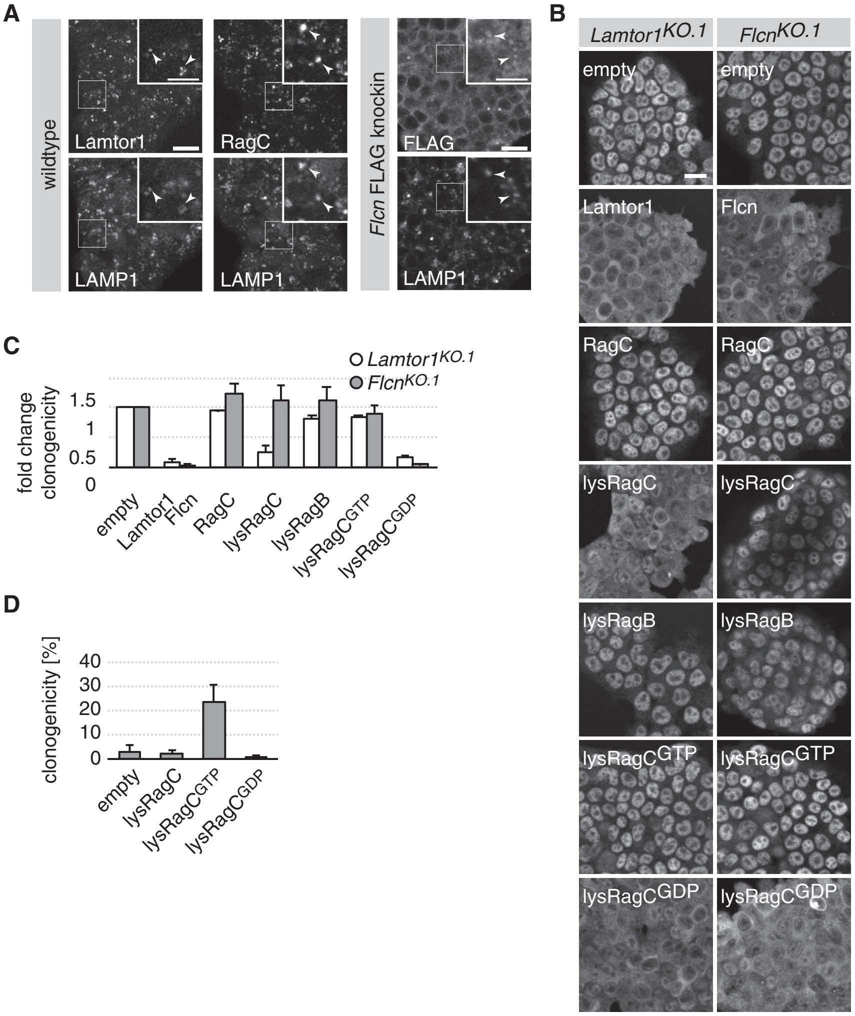


Figure 3. Lysosomal Rag GTPases Drive ESC Differentiation

(A) Colocalization of Lamtor1, RagC, and tagged Flcn with LAMP1 (arrowheads) in ESCs. Scale bars, 20 μm and 10 μm (inset).

(B and C) Tfe3 localization (B) and retention of self-renewal relative to empty vector-transfected cells after 3 days of 2i withdrawal (C) in cells of the indicated genotypes expressing the indicated Rag transgenes. Average and SD of 2 biological replicates are shown. Scale bar, 20 μm.

(D) Retention of self-renewal after 3 days of 2i withdrawal in wild-type cells expressing indicated RagC transgenes. Average and SD of 2 biological replicates.

Lamtor1 (Nada et al., 2009) was fused to RagB and C (lysRagB/C) and that colocalized with LAMP1 in *Lamtor1* KO ESCs (Figures S3F–S3H). Importantly, lysRagC but not lysRagB reverted nuclear Tfe3 localization and differentiation resistance (Figures 3B and 3C) in *Lamtor1* KO ESCs. Although we cannot exclude that the lysRagB hybrid protein is non-functional, this suggests that lysosomal targeting of RagC is sufficient for exit from self-renewal in the absence of Ragulator. Consistent with the notion that Flcn acts upstream of the RagC/D nucleotide state and independent of lysosomal recruitment (Figure S3A), lysRagC^{GDP} rescued *Flcn* KO phenotypes whereas lysRagC did not (Figures 3B and 3C). Conversely, lysRagC^{GTP} dominantly induced nuclear Tfe3 localization and Tfe3-dependent differentiation impairment in wild-type cells (Figures 3D, S3I, and S3J). These results suggest that the principal role of Ragulator in exit from ESC self-renewal is lysosomal recruitment of RagC/D.

Non-canonical RagGTPase Regulation and Signaling in ESCs

The RagA/B nucleotide state depends on the GATOR1 GAP complex, which is antagonized by GATOR2 (Bar-Peled et al., 2013; Figure S1A). To probe their role in differentiation, we generated *Depdc5* (a GATOR1 subunit) and *Wdr24* (a GATOR2 subunit) KO ESCs (Figure S4A). Neither featured nuclear Tfe3 enrichment or resistance to differentiation (Figures 4A and 4B), suggesting that RagA/B regulation is dispensable. This is surprising because activation of the Rag heterodimer by aas in somatic cells critically depends on the RagA/B nucleotide state (Saxton and Sabatini, 2017). We therefore tested whether ESCs are at all able to transduce changes in aa levels to mTORC1 by monitoring phosphorylation of the mTORC1 substrates S6 kinase (S6K) and 4EBP1. Aa deprivation in ESCs caused Depdc5-dependent mTORC1 inactivation, whereas readdition led to its reactivation, which was blunted in *Wdr24* and severely impaired in *Lamtor1* KO ESCs (Figure 4C). These observations are consistent with canonical aa sensing in ESCs. In contrast, mTORC1 regulation was unperturbed in *Flcn* KO ESCs. Although unexpected, Flcn- and Fnip1/2-independent mTORC1 activation in response to aas has been described before in other cell types (Nagashima et al., 2017; Wada et al., 2016). We conclude that GATOR-dependent regulation of RagA/B mediates mTORC1 activation by aas but not differentiation of non-starved ESCs. Consistent with decoupling of RagGTPase regulation and mTORC1 in steady-state ESCs, phosphorylation of mTORC1 substrates was not significantly altered in *Flcn* and *Lamtor1* KO cells (Figure S4B), and transcriptional deregulation in *Flcn* and *Lamtor1* KO ESCs does not correlate with mTOR inhibition (Bulut-Karslioglu et al., 2016; Figure S2A).

This may explain why exit from self-renewal is unperturbed in the presence of the allosteric mTORC1 inhibitor rapamycin (Rapa) (Betschinger et al., 2013). To show that this is not because of Rapa-insensitive mTORC1 activity (Saxton and Sabatini, 2017), we treated ESCs with the catalytic mTOR inhibitor AZD8055 and knocked down the mTORC1 subunit Raptor by siRNA transfection. Similar to Rapa treatment, this did not induce nuclear Tfe3 translocation or retention of self-renewal (Figures S4C–S4E), demonstrating that Rag GTPases drive differentiation independent of activating mTORC1. How, then, does mTORC1 hyperactivation upon depletion of Tsc1 and

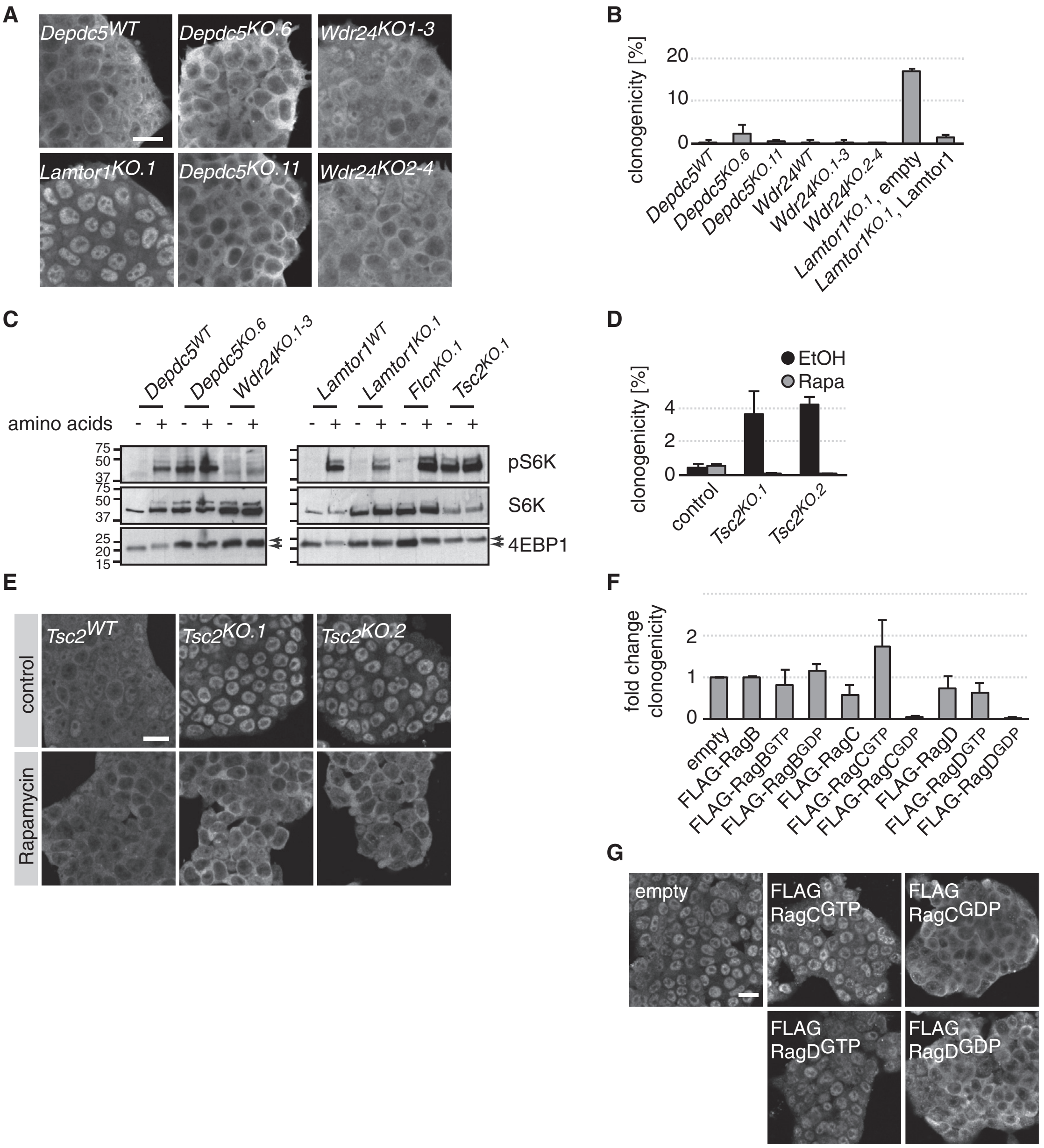


Figure 4. Non-canonical Rag GTPase Regulation in Steady-State ESCs

(A and B) Tfe3 localization (A) and retention of self-renewal after 3 days of 2i withdrawal (B) of the indicated genotypes. Average and SD of 2 biological replicates are shown. Scale bar, 20 μm.

(C) mTORC1 activity in ESCs in response to amino acids. Arrows mark fast-migrating (unphosphorylated) and slow-migrating (phosphorylated) 4EBP1.

(D) Retention of self-renewal after 3 days of 2i withdrawal in the indicated cell lines in the absence or presence of 20 nM rapamycin. Average and SD of 2 biological replicates are shown.

(E and G) Tfe3 subcellular localization in the indicated ESCs with or without 20 nM rapamycin (E) and *Tsc2* $^{KO.1}$ ESCs expressing the indicated Rag constructs (G). Scale bar, 20 μm.

(F) Retention of self-renewal relative to empty vector-transfected cells in *Tsc2*$^{KO.1}$ cells expressing the indicated transgenes after 3 days of 2i withdrawal. Average and SD of 2 biological replicates are shown.

Tsc2 impair ESC differentiation upstream of nuclear Tfe3 (Betschinger et al., 2013; Figures 1B–1D, S1D, S4C, and S4D)? We first compared transcriptional changes in *Flcn* and *Lamtor1* KO ESCs with recently published alterations in *Tsc2* KO ESCs (Li et al., 2018). This revealed a similar, although weaker, deregulation of Tfe3 target genes, particularly in clusters 5 and 6 (Figure S2B). The Tsc2 and Flcn-Ragulator pathways thus overlap significantly at the level of Tfe3 regulation. We reasoned that Rag GTPases are a potential intersection point of both pathways. To functionally test this, we generated *Tsc2* KO ESCs

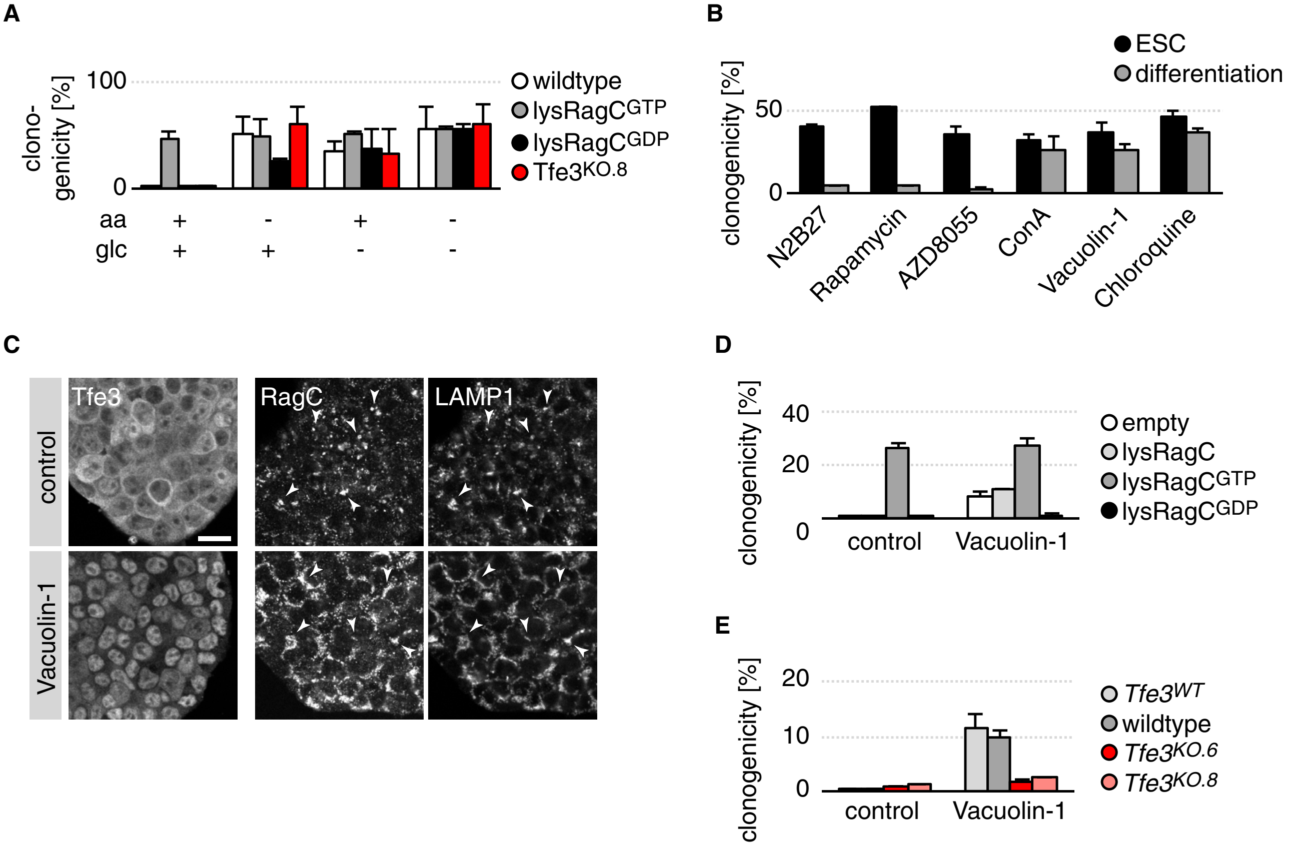


Figure 5. Lysosomes License ESC Differentiation

(A) Retention of self-renewal after 2 days of 2i withdrawal in the absence of essential amino acids (aa) and/or glucose (glc) in the indicated genotypes. Average and SD of 2 biological replicates are shown.

(B) Self-renewal of wild-type cells in 2i after 24 hr treatment with 40 nM rapamycin, 100 nM AZD8055, 20 nM Con A, 100 nM Vacuolin-1, or 10 μM chloroquine under 2i or differentiation conditions. Average and SD of two technical replicates are shown.

(C) Tfe3, RagC, and LAMP1 localization in wild-type ESCs treated for 24 hr with 100 nM Vacuolin-1. Arrowheads indicate colocalization of RagC and LAMP1. Scale bar, 20 μm.

(D and E) Retention of self-renewal in wild-type cells expressing the indicated Rag transgenes (D) and in the indicated cell lines (E) differentiated for 3 days in the absence or presence of 100 nM Vacuolin-1. Average and SD of 2 technical replicates are shown.

(Figures S4A and S4F). These cells featured hyperactive mTORC1, nuclear Tfe3, and impaired differentiation, which was reversed by addition of Rapa (Figures 4D, 4E, and S4G). Remarkably, expression of RagC^{GDP} or RagD^{GDP} specifically rescued these phenotypes (Figures 4F and 4G), which is similar to *Flcn* KO cells (Figures S3D and S3E). This suggests that Tsc-dependent mTORC1 inhibition and Flcn-Ragulator function converge on lysosomal RagC/D to sequester Tfe3 in the cytosol and promote exit from self-renewal.

Lysosomal Catabolism Drives ESC Differentiation

Our results indicate that Flcn specifies differentiation-specific RagGTPase signaling. To characterize upstream signals, we tested the role of aas and glucose in this pathway. Removal of essential aas or glucose from the ESC culture medium induced nuclear Tfe3 localization and impaired exit from self-renewal (Figures 5A and S5A). This was accompanied by inhibition of proliferation (data not shown), which was not observed in *Flcn* or *Lamtor1* KO ESCs. Consistent with a Rag GTPase-independent response, resistance to differentiation was not reverted by lysRagC^{GDP} expression or in *Tfe3* KO ESCs (Figures 5A and S5B-S5D). Although these observations do not rule out nutritional control of RagC/D activation, they likely reflect a multitude of stress-related responses in starving ESCs.

Lysosomes degrade proteins, lipids, carbohydrates, and organelles to generate aas that signal to Rag GTPases (Saxton and Sabatini, 2017). We therefore decided to specifically probe the role of lysosomal catabolism in ESC differentiation. Addition of the vacuolar H$^+$ ATPase (v-ATPase) inhibitors bafilomycin A (Baf A) and concanamycin A (Con A), the lysosomotropic compound chloroquine, or Vacuolin-1 impaired exit from self-renewal (Figures 5B and S5E). Because Baf A, Con A, and chloroquine also triggered dose-dependent cell lethality (data not shown), we focused our experiments on Vacuolin-1. Treatment with Vacuolin-1 induced nuclear Tfe3 translocation (Figure 5C), but expression of lysRagC^{GDP} or KO of *Tfe3* rescued Vacuolin-1-elicited resistance to differentiation (Figures 5D and 5E). Compound treatment altered lysosomal morphology, but RagC still

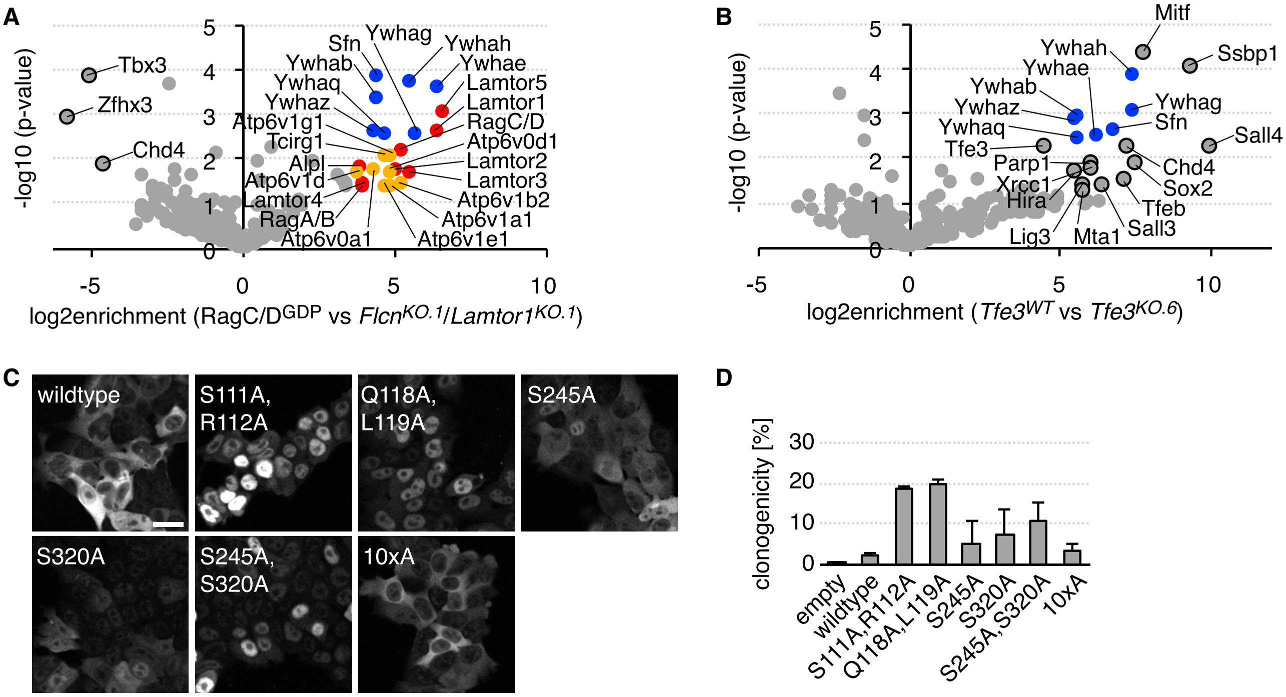


Figure 6. Association with Rag GTPases Drives Phosphorylation-Dependent Tfe3 Inactivation

(A and B) Tfe3 interactome comparison between RagC/D^{GDP}-expressing and *Flcn* and *Lamtor1* KO ESCs (A) and wild-type and *Tfe3* KO ESCs (B). Red marks the Ragulator-Rag GTPase complex, blue 14-3-3 protein family members, and yellow v-ATPase subunits. Quantification is based on 3 biological replicates.

(C and D) Tfe3 localization (C) and retention of self-renewal after 3 days of 2i withdrawal (D) in *Tfe3$^{KO.6}$* cells conditionally expressing the indicated Tfe3 alleles. Average and SD of 2 biological replicates are shown. Scale bar, 20 μm.

colocalized with LAMP1 (Figure 5C). In the presence of Vacuolin-1, the Ragulator-RagC complex, despite localizing to lysosomes, is thus not able to sequester Tfe3 in the cytoplasm. This suggests that lysosomal catabolism per se activates Flcn and RagC/D to drive Tfe3 inactivation and exit from ESC self-renewal.

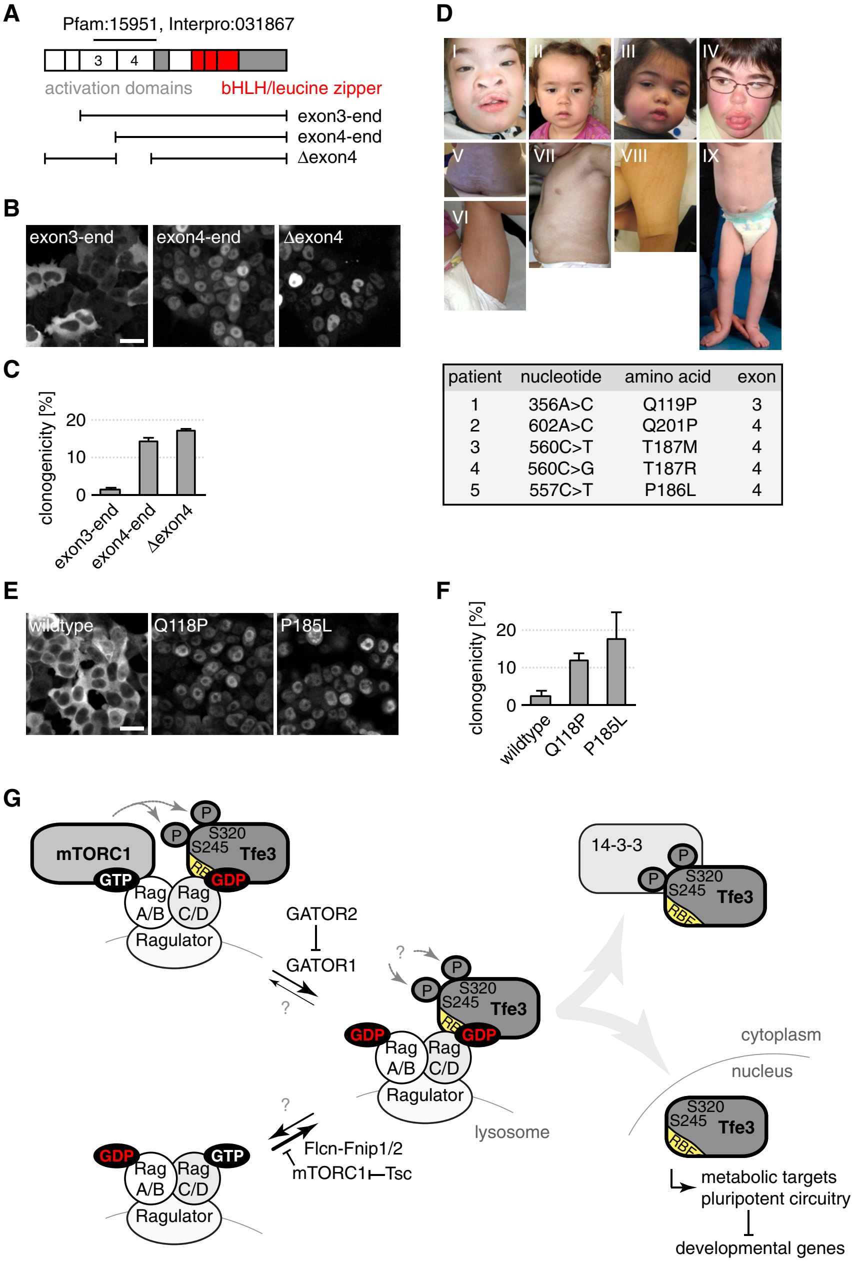
Binding to RagC/D Is Required for Phosphorylation and Inactivation of Tfe3

To explore how Rag GTPase signaling in ESCs regulates Tfe3, we determined the Tfe3 interactome using label-free quantitative mass spectrometry (Table S1). Tfe3 in RagC/D^{GDP}-overexpressing ESCs specifically interacted with the Ragulator-Rag GTPase protein complex, v-ATPase subunits, and members of the 14-3-3 protein family compared with nuclear Tfe3-specific binding partners in *Flcn* and *Lamtor1* KO cells (Figure 6A). Pull-down of Tfeb by the RagA/B^{GTP}-RagC/D^{GDP} heterodimer in human cell lines has been reported before (Martina and Puertollano, 2013), but Tfe3 failed to interact with any RagB nucleotide-loading mutant in ESCs (Figure S6A), demonstrating RagC/D^{GDP} specificity. We also noted that RagD^{GDP} co-immunoprecipitated more Tfe3 than RagC^{GDP} and that both induced Tfe3 protein stabilization. Association of Tfe3 with Rag GTPases was undetectable in wild-type ESCs (Figure 6B) and ESCs overexpressing wild-type RagC/D proteins (Figure S6A), suggesting transient and nucleotide-dependent binding. To test the necessity of this interaction in the context of the endogenous Rag nucleotide cycle, we expressed, in *Tfe3* KO ESCs, inducible Tfe3 alleles with mutations in residues homologous to the ones required for interaction of Tfeb with the active Rag heterodimer (Martina and Puertollano, 2013). In contrast to wild-type Tfe3, Tfe3(S111A, R112A) and Tfe3(Q118A, L119A) localized to the nucleus (Figure 6C) and impaired exit from self-renewal (Figure 6D), indicating that association of Tfe3 with endogenous RagC/D^{GDP} is required for cytoplasmic localization and ESC differentiation.

Phosphorylation by mTORC1 induces interaction of MiTF/Tfe TFs with 14-3-3 proteins in response to aas (Raben and Puertollano, 2016). To test whether association of cytoplasmic Tfe3 with 14-3-3 proteins in ESCs requires phosphorylation of the same sites, we generated Tfe3 alleles with Ala mutations in conserved mTORC1 substrate residues. This revealed that Tfe3(S245A) (Settembre et al., 2012) and Tfe3(S320A) (Roczniak-Ferguson et al., 2012) were partially and Tfe3(S245,S320A) exclusively nuclear compared with wild-type Tfe3 and Tfe3(10xA) (Peña-Llopis et al., 2011; Figure 6C). Single and double Tfe3 mutants also impaired ESC differentiation (Figure 6D). These findings suggest that mTORC1-independent S245 and S320 phosphorylation downstream of RagC/D^{GDP} is necessary for 14-3-3 protein-mediated cytoplasmic Tfe3 retention.

Point Mutations in Tfe3 Associated with a Human Developmental Disorder

Taken together, these findings demonstrate non-canonical Rag GTPase signaling in differentiation. This may not be unique to ESCs because *Tsc1/2* and *Flcn* loss-of-function mutations, hyperactive *mTOR* alleles, and deletion of *Tfe3* exons 1–3 by



patient	nucleotide	amino acid	exon
1	356A>C	Q119P	3
2	602A>C	Q201P	4
3	560C>T	T187M	4
4	560C>G	T187R	4
5	557C>T	P186L	4

Figure 7. Disease-Associated Tfe3 Alterations Produce Ectopic Nuclear Gain-of-Function Alleles

(A) Tfe3 domain and exon structure and Tfe3 deletion constructs used in (B) and (C). Exons 3 and 4 and the MiTF/Tfe TF family-specific homology stretch are indicated.

(B, C, E, and F) Tfe3 localization (B and E) and retention of self-renewal after 3 days of 2i withdrawal (C and F) in *Tfe3*$^{KO.6}$ cell lines conditionally expressing the indicated Tfe3 alleles. Average and SD of 2 biological replicates are shown. Scale bar, 20 μm.

(D) Clinical phenotype of patients and molecular details of corresponding Tfe3 mutations. I–IV: facial features of patients 1 (I), 2 (II), 3 (III), and 5 (IV). Note the coarse features, facial hypertrichosis, and full cheeks common to the individuals. V–VIII: cutaneous phenotype of patients 1 (V), 2 (VI and VII), and 3 (VIII). Note the whorls and streak hypopigmentation along Blaschko's line. IX: body asymmetry and umbilical hernia in patient 4.

(G) Model of lysosomal signaling in ESCs. The Rag heterodimer equilibrium under steady-state ESCs is shifted toward the di-GDP conformation. RagC/D^{GDP} interacts with the Rag binding fold (RBF) in Tfe3, catalyzing its phosphorylation and consequential cytoplasmic sequestration by 14-3-3 proteins. Unphosphorylated Tfe3 translocates into the nucleus to induce metabolic targets and stabilize the pluripotent transcriptional circuitry.

translocation to distinct partner genes are frequently detected in renal cell carcinoma (Grabiner et al., 2014; Kauffman et al., 2014). The molecular consequences of these Tfe3 alterations are poorly understood (Kauffman et al., 2014), but the results presented here suggest that they behave as nuclear gain-of-function alleles. We therefore conditionally expressed corresponding Tfe3 deletions in *Tfe3* KO ESCs (Figure 7A). Inclusion of exon 3 (exon 3-end) recapitulated cytoplasmic localization of wild-type Tfe3, whereas removal (exon 4-end) induced nuclear translocation and impaired differentiation (Figures 7B and 7C), indicating that the absence of exon 3 and not the presence of fusion partners may underly oncogenic activity of chromosomal Tfe3 rearrangements. Surprisingly, transcript sequencing of gain-of-function *Tfe3* alleles recovered in our screen (Figures 1B and 1C) revealed small in-frame-deletions around T180–P185 and loss of the entire exon 4 (Figure S7A). Nuclear localization of the Tfe3$^{\Delta exon4}$ protein and resistance to differentiation were validated in a Tfe3 exon 4-skipping mutant ESC clone generated by an independent gRNA (Figures S5B, S5D, and S7B) and in *Tfe3* KO cells harboring inducible Tfe3$^{\Delta exon4}$ transgenes (Figures 7A–7C). Both exon 3 and exon 4 are therefore required for Tfe3 inactivation. Alterations in exon 4, however, have not been associated with disease as of now.

In a program designed to identify genetic causalities of hypomelanosis of Ito (HI, OMIM: 300337), we identified four *de novo* germline mutations (female patient 1, Q119P; female patient 2, Q201P; female patient 3, T187M; female patient 5, P186L) and one *de novo* mosaic mutation (male patient 4, T187R) in *TFE3* (Figure 7D; Table S2; Data S1). HI and pigmentary mosaicism are unspecific terms encompassing a heterogeneous group of disorders characterized by hypopigmented whorls and streaks along Blaschko's lines and variable extracutaneous features affecting the musculoskeletal and nervous systems (Sybert, 1994). Phenotypes of patients with *TFE3* mutations included hypopigmentation on Blaschko's lines, severe intellectual disability (ID), coarse facial features, frontonasal dysplasia, obesity, epilepsy, and growth retardation (Table S2), suggesting a pleiotropic developmental disorder. All five *TFE3* variants affect conserved aas (Figure S7C), and the missense variants are *in silico* predicted to be pathogenic (Table S2). The Q119P mutation resides in exon 3 and affects an aa implicated in Rag GTPase interaction (Figures 6C and 6D; Martina and Puertollano, 2013), whereas the remaining four mutations (P186L, T187M, T187R, and Q201P) induce aa changes in exon 4. Nuclear localization and resistance to differentiation were validated in *Tfe3* KO ESCs expressing murine Tfe3 alleles (Q118P and P185L) (Figures 7E and 7F). Secondary structure analysis of Tfe3 using I-Tasser (Yang et al., 2015) predicts residues 110–215, encoded by exon 3 and exon 4, to form a domain of two stable α helices (Figure S7C). The Q-to-P mutations at positions 119 and 201 likely disrupt helix formation. The two helices are connected by a short loop and an induced α helix that contains the residues P186 and T187, which may be crucial for interaction with the Rag heterodimer. These observations suggest that exon 3 and exon 4 form a Rag binding fold whose structural integrity is indispensable for cytoplasmic Tfe3 inactivation.

Tfe3 Impairs Neuronal Progenitor Differentiation

Subcellular Tfe3 distribution is associated with developmental progression of human pluripotent stem cells (Gafni et al., 2013). It is therefore surprising that Tfe3 gain-of-function mutations do not induce lethality in humans. This is probably a consequence of mosaicism because *TFE3* is an X chromosome-linked gene. We indeed observed genetic mosaicism by postzygotic *TFE3* mutation in the only male patient, likely functional mosaicism by random X chromosome inactivation in three female patients (demonstrated for one patient), and skewed X inactivation in one patient without hypopigmentation (patient 5). Emergence of ectopic nuclear Tfe3 mutations in somatic but not pluripotent cells is, thus, likely causal for the disorder.

Although patient data did not allow us to infer when or where mutations emerged, defects in nervous system development would be consistent with some of the disease phenotypes. We therefore used the neural lineage as a proxy to test whether Tfe3 gain of function, similar to ESCs, induces differentiation defects in somatic cells. When released from self-renewal, ESCs recapitulate hallmarks of neural development *in vitro*, including expression of the neural progenitor markers Ascl1, Chrdl1, Lrp2, Hes5, Pax3, Sox1, and Zic1 by day 4, followed by a gradual increase in the neuronal differentiation markers Dcx and Tubb3 (Figure S7D; Abranches et al., 2009). Nuclear Tfe3 impaired exit from self-renewal and downregulation of the pluripotency TFs Esrrb, Nanog, and Oct4 only when induced on day 0 or day 1 but not day 2 of differentiation (Betschinger et al., 2013; Figures S7E and S7F), providing an experimental setup to study Tfe3 independent of its role in pluripotency. We induced Tfe3 daily from day 4 onward and observed a decrease in expression of Dcx and Tubb3 on day 10 compared with uninduced cells (Figure S7G). Notably, the magnitude of repression scaled with the timing of Tfe3 induction and was mirrored by an increase in the neural progenitor markers Lrp2, Chrdl1, and Zic1 but not Sox1, Ascl1, or Pax3. The lysosomal ESC Tfe3 targets Cd63, Ctsd, Hexa, and Uppl1 were upregulated to similar levels irrespective of the timing of Tfe3 induction, confirming that lysosomal gene transcription is a cell-type-invariant Tfe3 response. These findings suggest that nuclear Tfe3, similar to its function in blocking exit from the ESC state, impairs developmental progression of somatic progenitor cells.

DISCUSSION

Blastocyst development and ESC differentiation are associated with shifts in metabolic pathways, but the cause-and-effect relations are mostly unclear (Zhang et al., 2018). Here we show that lysosomal Rag GTPases instruct exit from ESC self-renewal. Rag GTPases have an evolutionarily conserved role in conveying aa signals to cellular metabolism through activation of mTORC1 (Saxton and Sabatini, 2017). Using genetic and biochemical approaches, we demonstrate a distinct mechanism in ESCs (Figure 7G): differentiation requires an alternate Rag heterodimer conformation that, instead of activating mTORC1, physically interacts with Tfe3 and induces its phosphorylation. Subsequent binding to 14-3-3 proteins sequesters Tfe3 in the cytoplasm, consequentially licensing downregulation of metabolic and upregulation of developmental transcription. Non-canonical regulation of Rag GTPases explains why Flcn and Tsc drive ESC differentiation but have opposite roles in mTORC1 activation. We identify mosaic *TFE3* gain-of-function mutations in a pleiotropic human developmental disorder that are located in a protein domain essential for this lysosome-dependent inactivation. Our findings reveal that metabolic signaling at the lysosomal membrane, in addition to regulating cellular homeostasis, controls embryonic cell fate transitions.

How are mTORC1-dependent and -independent Rag GTPase functions coordinated in ESCs? Our findings suggest two mutually non-exclusive possibilities. First, activation of mTORC1 by aas occurs within minutes and depends on prior aa starvation, whereas exit from the non-starved ESC state requires at least 24 hr (Kalkan et al., 2017). The difference in timing and environmental conditions may alter mTORC1 regulation; e.g., because of compensatory mechanisms or cellular stress responses. Depletion of the Ragulator subunit Lamtor1, which is crucial for Rag GTPase signaling in aa sensing and differentiation, supports this possibility. *Lamtor1* KO cells are deficient in activating mTORC1 upon aa refeeding but proficiently do so under non-starved steady-state conditions, demonstrating that mTORC1 regulation depends on the cellular context. Second, Rag GTPase GAPs have distinct functions in aa sensing and differentiation. We showed that the RagA/B GAP GATOR1 and its inhibitor GATOR2, but not the RagC/D GAP Flcn, control mTORC1 activation in response to aas. Conversely, Flcn, but not the GATORs, drives Tfe3 inactivation and exit from self-renewal. Ragulator is required for both, likely because it recruits Rag GTPases to the lysosomal membrane. This indicates that extrinsic regulation of the RagC/D but not RagA/B nucleotide state specifies mTORC1-independent signaling in non-starved ESCs.

How is Tfe3 inactivation instructed by the Rag heterodimer? Rag heterodimers cycle between stable mTORC1-activating RagA/B^{GTP}-RagC/D^{GDP} and inactive RagA/B^{GDP}-RagC/D^{GTP} conformations via a di-GDP state (Shen et al., 2017). We propose that these conformations co-exist at an equilibrium in steady-state ESCs that is shifted toward the di-GDP state, which catalyzes differentiation (Figure 7G). First, the di-GDP conformation is unable to interact with mTORC1 (Sancak et al., 2008), which is in line with the expandability of mTORC1 activation for exit from self-renewal. Second, the limited contribution of lysosomal Rag GTPases to mTORC1 activity in non-starved ESCs (Figure S4B) and the persistence of differentiation-competent Rag GTPases upon GATOR1 KO is consistent with low amounts of the RagA/B^{GTP}-RagC/D^{GDP} heterodimer under steady-state conditions. We showed that the RagC/D^{GDP} subunit specifies interaction and inactivation of Tfe3. This depends on a Rag GTPase binding fold encoded by exons 3 and 4 and phosphorylation of serine residues that induce binding to 14-3-3 proteins. The same domain and phosphorylation sites are required for mTORC1-dependent cytoplasmic sequestration of MiTF/Tfe TFs in somatic cells (Martina and Puertollano, 2013), suggesting that the RagA/B^{GDP} moiety in the di-GDP heterodimer recruits a yet to be identified kinase. Alternatively, inability to recruit mTORC1 may allow lysosome-associated kinases to phosphorylate Tfe3 in ESCs.

It is therefore surprising that hyperactive mTORC1 in *Tsc2* KO ESCs is an upstream inhibitor of RagC/D and antagonizes Flcn. mTORC1 in wild-type cells may have a similar activity because its chemical inhibition causes lysosomal Tfe3 recruitment (Figure S4E), indicating strengthening of the RagC/D-Tfe3 complex. mTORC1 has indeed been shown to induce phosphorylation (Piao et al., 2009) and reduce lysosomal localization (Martina et al., 2014; Meng and Ferguson, 2018) of Flcn, suggesting that hyperactive mTORC1 directly antagonizes Flcn. mTORC1 overactivation alone is, however, not sufficient to explain why *Tsc2* KO ESCs inefficiently exit self-renewal because absence of GATOR1 similarly induces mTORC1 but accelerates differentiation (Li et al., 2018) without perturbing the transcription of Tfe3 target genes (Figure S2B). The conflicting roles of these two mTORC1-inhibiting protein complexes may arise because they impose different quantitative and/or qualitative restrictions on mTORC1 substrate selection (Li et al., 2018). Alternatively, an additional, mTORC1 activation-independent function of Tsc2, such as its physical interaction with Rag heterodimers (Demetriades et al., 2014), may contribute to stabilization of RagC/D^{GDP}. We suspect that the consequences of mTORC1 hyperactivation are not cell type specific because *MTORC1* gain-of-function alleles, similar to *TFE3*, are associated with kidney cancer (Grabiner et al., 2014) and pigmentary mosaicism (Mirzaa et al., 2016), and patients with *TFE3* point mutations present classical features of hyperactive mTORC1 pathway-related disorders (Parrini et al., 2016) (epilepsy, frequently associated with cortical malformations) (Table S2).

We surmise that independent regulation of Tfe3 and mTORC1 in ESCs allows differential control of ESC proliferation (Gangloff et al., 2004; Murakami et al., 2004), exit from self-renewal, and translation (Sampath et al., 2008) in response to intra- and extracellular cues. Lysosome activity, Flcn-Fnip, and Tsc, which are all directly associated with cellular nutrition (Nagashima et al., 2017; Saxton and Sabatini, 2017), inactivate RagC/D and Tfe3 in ESCs. It is therefore conceivable that the quantitative and qualitative integration of multiple nutritional signals forms a metabolic differentiation checkpoint that anticipates energy requirements faced during developmental progression. This may not be unique to pluripotent stem cells because induction of Tfe3 in neural progenitor cells also delays differentiation while sustaining progenitor-specific transcription. Activation of lysosomal genes is a direct and generic catabolic function of Tfe3 (Martina et al., 2014) that could contribute to lysosomal storage disorder-related phenotypes in most individuals harboring *TFE3* mutations (coarse facial features, umbilical hernia, postnatal growth retardation, obesity, hepatomegaly, and hypoglycemia) (Table S2). Inhibition of differentiation, in contrast, is most likely caused by the concomitant repression of developmental genes, which, at least in ESCs, are predominantly indirect Tfe3 targets in cluster 12 (Figure 2C). Our genomic and proteomic analyses suggest two mechanisms for this cell-type-specific response to Tfe3 activation: Tfe3 may directly stabilize the ESC state by interacting with the pluripotent transcription factor network post-transcriptionally by physical association with, e.g., Sall4 and Sox2 (Figure 6B) and transcriptionally by binding to promoters of, e.g., the cluster 13 TFs Rex1 and Esrrb (Table S1; Betschinger et al., 2013). Mutually non-exclusive, direct metabolic targets may perpetuate ESC self-renewal by programing a metabolic state refractory to differentiation. Consistent with this possibility, MiTF/Tfe TFs have been reported to control cell type specification in brown adipose tissue (Wada et al., 2016) and the endoderm (Young et al., 2016) through metabolic target genes. How developmental defects in patients with somatic Tfe3 gain-of-function mutations relate to the protein's function in lysosomal biogenesis, cell-type-specific transcription, and exit from progenitor cell states remains to be determined. Clinically, the association of cutaneous pigmentary mosaicism, coarse facial features, and ID is recognizable, and broadening of the phenotypic spectrum by identification of more patients appears likely. In addition to regulating embryogenesis, mosaic *TFE3* mutations may thus be considered causative of syndromic ID.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
    - Mouse ESCs
    - Patients
- METHOD DETAILS
    - Cell Culture
    - Immunostaining
    - CRISPR/Cas9 screen
- GENOME EDITING
    - Molecular biology
    - Protein methods
- QUANTIFICATION AND STATISTICAL ANALYSIS
    - CRISPR/Cas9 screen quantification
    - Protein identification and quantification
    - Microarray data analysis
    - RNaseq and GROseq data analysis
    - ChIPseq and DHSseq data analysis
    - Whole exome sequencing (WES)
    - TFE3 sequencing
    - Complementary DNA analysis

  - ○ Determination of the X chromosome inactivation pattern at the HUMARA locus
  - ○ Details for quantification and statistical analysis in Figures
- ● DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures, three tables, and one data file and can be found with this article online at https://doi.org/10.1016/j.stem.2018.11.021.

ACKNOWLEDGMENTS

We would like to thank the patients and families who participated in this study. We also thank the University of Burgundy Centre de Calcul (CCuB) for technical support and management of the informatics platform; S. Dessus-Babus, K. Jacobeit, and T. Roloff (FMI) for processing sequencing samples; L. Gelman (FMI) for imaging assistance; H. Gut (FMI) for structural modelling; M. Flemr (FMI) for advice regarding genome editing; J. Chao (FMI), A. Smith (University of Cambridge), F. Stewart (University of Dresden), K. Yusa (Wellcome Trust Sanger Institute), and M. Leeb (MFPL) for providing reagents; and S. Gasser, P. Liberali, A. Peters, D. Schuebeler, N. Thomae (FMI), S. Stricker (Helmholtz Zentrum), and D. Teis (University of Innsbruck) for comments on the manuscript. This work was funded by EMBO and Marie Curie Actions (ALTF 1632-2014 to D.O.), the Programme Hospitalier de Recherche Clinique (PHRC) National (to P.V.), the Regional Council of Burgundy and Centre Hospitalo-Universitaire de Dijon PARI 2015 (to L.F.), and the Novartis Research Foundation (to J.B.).

AUTHOR CONTRIBUTIONS

F.V., M.R., M.A., L.D.J., J.S.-O., T.J., and J.B. performed experiments. D.M., M.S., Y.D., P.K., J.T., K.L.I.v.G., and S.V.K. performed bioinformatics analyses. M.Z. and D.O. executed the CRIPSR/Cas9 screen. D.H. performed mass spectrometry. P.V., E.S.J.M.D.B., F.D., D.G., N.H., K.D.L., C.T.R., and M.v.H. recruited and evaluated the patients. S.S. established quantitative gRNA sequencing. L.F., J.-B.R., and J.B. supervised the work. D.L. and J.B. wrote the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 29, 2017
Revised: September 21, 2018
Accepted: November 20, 2018
Published: February 7, 2019

REFERENCES

Abranches, E., Silva, M., Pradier, L., Schulz, H., Hummel, O., Henrique, D., and Bekman, E. (2009). Neural differentiation of embryonic stem cells in vitro: a road map to neurogenesis in the embryo. PLoS ONE *4*, e6286.

Allen, R.C., Zoghbi, H.Y., Moseley, A.B., Rosenblatt, H.M., and Belmont, J.W. (1992). Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. Am. J. Hum. Genet. *51*, 1229–1239.

Baba, M., Hong, S.-B., Sharma, N., Warren, M.B., Nickerson, M.L., Iwamatsu, A., Esposito, D., Gillette, W.K., Hopkins, R.F., 3rd, Hartley, J.L., et al. (2006). Folliculin encoded by the BHD gene interacts with a binding protein, FNIP1, and AMPK, and is involved in AMPK and mTOR signaling. Proc. Natl. Acad. Sci. USA *103*, 15552–15557.

Bar-Peled, L., Schweitzer, L.D., Zoncu, R., and Sabatini, D.M. (2012). Ragulator is a GEF for the rag GTPases that signal amino acid levels to mTORC1. Cell *150*, 1196–1208.

Bar-Peled, L., Chantranupong, L., Cherniack, A.D., Chen, W.W., Ottina, K.A., Grabiner, B.C., Spear, E.D., Carter, S.L., Meyerson, M., and Sabatini, D.M. (2013). A Tumor suppressor complex with GAP activity for the Rag GTPases that signal amino acid sufficiency to mTORC1. Science *340*, 1100–1106.

Betschinger, J. (2017). Charting Developmental Dissolution of Pluripotency. J. Mol. Biol. *429*, 1441–1458.

Betschinger, J., Nichols, J., Dietmann, S., Corrin, P.D., Paddison, P.J., and Smith, A. (2013). Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. Cell *153*, 335–347.

Boroviak, T., Loos, R., Lombard, P., Okahara, J., Behr, R., Sasaki, E., Nichols, J., Smith, A., and Bertone, P. (2015). Lineage-Specific Profiling Delineates the Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis. Dev. Cell *35*, 366–382.

Buecker, C., Srinivasan, R., Wu, Z., Calo, E., Acampora, D., Faial, T., Simeone, A., Tan, M., Swigut, T., and Wysocka, J. (2014). Reorganization of enhancer patterns in transition from naive to primed pluripotency. Cell Stem Cell *14*, 838–853.

Bulut-Karslioglu, A., Biechele, S., Jin, H., Macrae, T.A., Hejna, M., Gertsenstein, M., Song, J.S., and Ramalho-Santos, M. (2016). Inhibition of mTOR induces a paused pluripotent state. Nature *540*, 119–123.

Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N.V., Bogdanove, A.J., and Voytas, D.F. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res. *39*, e82–e82.

Demetriades, C., Doumpas, N., and Teleman, A.A. (2014). Regulation of TORC1 in response to amino acid starvation via lysosomal recruitment of TSC2. Cell *156*, 786–799.

Dorighi, K.M., Swigut, T., Henriques, T., Bhanu, N.V., Scruggs, B.S., Nady, N., Still, C.D., 2nd, Garcia, B.A., Adelman, K., and Wysocka, J. (2017). Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from Promoters Independently of H3K4 Monomethylation. Mol. Cell *66*, 568–576.e4.

Flemr, M., and Bühler, M. (2015). Single-Step Generation of Conditional Knockout Mouse Embryonic Stem Cells. Cell Rep. *12*, 709–716.

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.

Gaidatzis, D., Lerch, A., Hahne, F., and Stadler, M.B. (2015). QuasR: quantification and annotation of short reads in R. Bioinformatics *31*, 1130–1132.

Gangloff, Y.-G., Mueller, M., Dann, S.G., Svoboda, P., Sticker, M., Spetz, J.-F., Um, S.H., Brown, E.J., Cereghini, S., Thomas, G., and Kozma, S.C. (2004). Disruption of the mouse mTOR gene leads to early postimplantation lethality and prohibits embryonic stem cell development. Mol. Cell. Biol. *24*, 9508–9516.

Grabiner, B.C., Nardi, V., Birsoy, K., Possemato, R., Shen, K., Sinha, S., Jordan, A., Beck, A.H., and Sabatini, D.M. (2014). A diverse array of cancer-associated MTOR mutations are hyperactivating and can predict rapamycin sensitivity. Cancer Discov. *4*, 554–563.

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. Cell *146*, 519–532.

Hwang, I.-Y., Kwak, S., Lee, S., Kim, H., Lee, S.-E., Kim, J.-H., Kim, Y.A., Jeon, Y.K., Chung, D.H., Jin, X., et al. (2016). Psat1-Dependent Fluctuations in α-Ketoglutarate Affect the Timing of ESC Differentiation. Cell Metab. *24*, 494–501.

Kalkan, T., Olova, N., Roode, M., Mulas, C., Lee, H.J., Nett, I., Marks, H., Walker, R., Stunnenberg, H.G., Lilley, K.S., et al. (2017). Tracking the embryonic stem cell transition from ground state pluripotency. Development *144*, 1221–1234.

Kauffman, E.C., Ricketts, C.J., Rais-Bahrami, S., Yang, Y., Merino, M.J., Bottaro, D.P., Srinivasan, R., and Linehan, W.M. (2014). Molecular genetics and cellular features of TFE3 and TFEB fusion kidney cancers. Nat. Rev. Urol. *11*, 465–475.

Koike-Yusa, H., Li, Y., Tan, E.-P., Velasco-Herrera, Mdel.C., and Yusa, K. (2014). Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nat. Biotechnol. *32*, 267–273.

Kurimoto, K., Yabuta, Y., Hayashi, K., Ohta, H., Kiyonari, H., Mitani, T., Moritoki, Y., Kohri, K., Kimura, H., Yamamoto, T., et al. (2015). Quantitative Dynamics of Chromatin Remodeling during Germ Cell Specification from Mouse Embryonic Stem Cells. Cell Stem Cell *16*, 517–532.

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.

Li, W., Xu, H., Xiao, T., Cong, L., Love, M.I., Zhang, F., Irizarry, R.A., Liu, J.S., Brown, M., and Liu, X.S. (2014). MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome Biol. *15*, 554.

Li, M., Yu, J.S.L., Tilgner, K., Ong, S.H., Koike-Yusa, H., and Yusa, K. (2018). Genome-wide CRISPR-KO Screen Uncovers mTORC1-Mediated Gsk3 Regulation in Naive Pluripotency Maintenance and Dissolution. Cell Rep. *24*, 489–502.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823–826.

Martina, J.A., and Puertollano, R. (2013). Rag GTPases mediate amino acid-dependent recruitment of TFEB and MITF to lysosomes. J. Cell Biol. *200*, 475–491.

Martina, J.A., Diab, H.I., Lishu, L., Jeong-A, L., Patange, S., Raben, N., and Puertollano, R. (2014). The nutrient-responsive transcription factor TFE3 promotes autophagy, lysosomal biogenesis, and clearance of cellular debris. Sci. Signal. *7*, ra9.

Meng, J., and Ferguson, S.M. (2018). GATOR1-dependent recruitment of FLCN-FNIP to lysosomes coordinates Rag GTPase heterodimer nucleotide status in response to amino acids. J. Cell Biol. *217*, 2765.

Mirzaa, G.M., Campbell, C.D., Solovieff, N., Goold, C., Jansen, L.A., Menon, S., Timms, A.E., Conti, V., Biag, J.D., Adams, C., et al. (2016). Association of MTOR Mutations With Developmental Brain Disorders, Including Megalencephaly, Focal Cortical Dysplasia, and Pigmentary Mosaicism. JAMA Neurol. *73*, 836–845.

Moussaieff, A., Rouleau, M., Kitsberg, D., Cohen, M., Levy, G., Barasch, D., Nemirovski, A., Shen-Orr, S., Laevsky, I., Amit, M., et al. (2015). Glycolysis-mediated changes in acetyl-CoA and histone acetylation control the early differentiation of embryonic stem cells. Cell Metab. *21*, 392–402.

Murakami, M., Ichisaka, T., Maeda, M., Oshiro, N., Hara, K., Edenhofer, F., Kiyama, H., Yonezawa, K., and Yamanaka, S. (2004). mTOR is essential for growth and proliferation in early mouse embryos and embryonic stem cells. Mol. Cell. Biol. *24*, 6710–6718.

Nada, S., Hondo, A., Kasai, A., Koike, M., Saito, K., Uchiyama, Y., and Okada, M. (2009). The novel lipid raft adaptor p18 controls endosome dynamics by anchoring the MEK-ERK pathway to late endosomes. EMBO J. *28*, 477–489.

Nagashima, K., Fukushima, H., Shimizu, K., Yamada, A., Hidaka, M., Hasumi, H., Ikebe, T., Fukumoto, S., Okabe, K., and Inuzuka, H. (2017). Nutrient-induced FNIP degradation by SCFβ-TRCP regulates FLCN complex localization and promotes renal cancer progression. Oncotarget *8*, 9947–9960.

Ostapcuk, V., Mohn, F., Carl, S.H., Basters, A., Hess, D., Iesmantavicius, V., Lampersberger, L., Flemr, M., Pandey, A., Thomä, N.H., et al. (2018). Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes. Nature *557*, 739–743.

Parrini, E., Conti, V., Dobyns, W.B., and Guerrini, R. (2016). Genetic Basis of Brain Malformations. Mol. Syndromol. *7*, 220–233.

Peña-Llopis, S., Vega-Rubin-de-Celis, S., Schwartz, J.C., Wolff, N.C., Tran, T.A.T., Zou, L., Xie, X.-J., Corey, D.R., and Brugarolas, J. (2011). Regulation of TFEB and V-ATPases by mTORC1. EMBO J. *30*, 3242–3258.

Petit, C.S., Roczniak-Ferguson, A., and Ferguson, S.M. (2013). Recruitment of folliculin to lysosomes supports the amino acid-dependent activation of Rag GTPases. J. Cell Biol. *202*, 1107–1122.

Piao, X., Kobayashi, T., Wang, L., Shiono, M., Takagi, Y., Sun, G., Abe, M., Hagiwara, Y., Zhang, D., Okimoto, K., et al. (2009). Regulation of folliculin (the BHD gene product) phosphorylation by Tsc2-mTOR pathway. Biochem. Biophys. Res. Commun. *389*, 16–21.

Raben, N., and Puertollano, R. (2016). TFEB and TFE3: Linking Lysosomes to Cellular Adaptation to Stress. Annu. Rev. Cell Dev. Biol. *32*, 255–278.

Roczniak-Ferguson, A., Petit, C.S., Froehlich, F., Qian, S., Ky, J., Angarola, B., Walther, T.C., and Ferguson, S.M. (2012). The transcription factor TFEB links mTORC1 signaling to transcriptional control of lysosome homeostasis. Sci. Signal. *5*, ra42.

Sampath, P., Pritchard, D.K., Pabon, L., Reinecke, H., Schwartz, S.M., Morris, D.R., and Murry, C.E. (2008). A hierarchical network controls protein translation during murine embryonic stem cell self-renewal and differentiation. Cell Stem Cell *2*, 448–460.

Sancak, Y., Peterson, T.R., Shaul, Y.D., Lindquist, R.A., Thoreen, C.C., Bar-Peled, L., and Sabatini, D.M. (2008). The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1. Science *320*, 1496–1501.

Sancak, Y., Bar-Peled, L., Zoncu, R., Markhard, A.L., Nada, S., and Sabatini, D.M. (2010). Ragulator-Rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids. Cell *141*, 290–303.

Sano, O., Kazetani, K., Funata, M., Fukuda, Y., Matsui, J., and Iwata, H. (2016). Vacuolin-1 inhibits autophagy by impairing lysosomal maturation via PIKfyve inhibition. FEBS Lett. *590*, 1576–1585.

Sardiello, M., Palmieri, M., di Ronza, A., Medina, D.L., Valenza, M., Gennarino, V.A., Di Malta, C., Donaudy, F., Embrione, V., Polishchuk, R.S., et al. (2009). A gene network regulating lysosomal biogenesis and function. Science *325*, 473–477.

Saxton, R.A., and Sabatini, D.M. (2017). mTOR Signaling in Growth, Metabolism, and Disease. Cell *168*, 960–976.

Scognamiglio, R., Cabezas-Wallscheid, N., Thier, M.C., Altamura, S., Reyes, A., Prendergast, Á.M., Baumgärtner, D., Carnevalli, L.S., Atzberger, A., Haas, S., et al. (2016). Myc Depletion Induces a Pluripotent Dormant State Mimicking Diapause. Cell *164*, 668–680.

Settembre, C., Zoncu, R., Medina, D.L., Vetrini, F., Erdin, S., Erdin, S., Huynh, T., Ferron, M., Karsenty, G., Vellard, M.C., et al. (2012). A lysosome-to-nucleus signalling mechanism senses and regulates the lysosome via mTOR and TFEB. EMBO J. *31*, 1095–1108.

Shen, K., Choe, A., and Sabatini, D.M. (2017). Intersubunit Crosstalk in the Rag GTPase Heterodimer Enables mTORC1 to Respond Rapidly to Amino Acid Availability. Mol. Cell *68*, 552–565.e8.

Smith, A. (2017). Formative pluripotency: the executive phase in a developmental continuum. Development *144*, 365–373.

Sybert, V.P. (1994). Hypomelanosis of Ito: a description, not a diagnosis. J. Invest. Dermatol. *103* (5, Suppl), 141S–143S.

Tsun, Z.-Y., Bar-Peled, L., Chantranupong, L., Zoncu, R., Wang, T., Kim, C., Spooner, E., and Sabatini, D.M. (2013). The folliculin tumor suppressor is a GAP for the RagC/D GTPases that signal amino acid levels to mTORC1. Mol. Cell *52*, 495–505.

Wada, S., Neinast, M., Jang, C., Ibrahim, Y.H., Lee, G., Babu, A., Li, J., Hoshino, A., Rowe, G.C., Rhee, J., et al. (2016). The tumor suppressor FLCN mediates an alternate mTOR pathway to regulate browning of adipose tissue. Genes Dev. *30*, 2551–2564.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. Nat. Methods *12*, 7–8.

Young, N.P., Kamireddy, A., Van Nostrand, J.L., Eichner, L.J., Shokhirev, M.N., Dayn, Y., and Shaw, R.J. (2016). AMPK governs lineage specification through Tfeb-dependent regulation of lysosomes. Genes Dev. *30*, 535–552.

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.

Zhang, J., Zhao, J., Dahan, P., Lu, V., Zhang, C., Li, H., and Teitell, M.A. (2018). Metabolism in Pluripotent Stem Cells and Early Mammalian Development. Cell Metab. *27*, 332–338.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
FLAG	Sigma	Cat#F1804; RRID:AB_262044
LAMP1	DSHB	Cat#1D4B; RRID:AB_2134500
Lamtor1	Cell Signaling	Cat#8975S; RRID:AB_10860252
RagB	Cell Signaling	Cat#8150S; RRID:AB_11178806
RagC	Cell Signaling	Cat#9480S; RRID:AB_10614716
Tfe3	Sigma	Cat#HPA023881; RRID:AB_1857931
GAPDH	Sigma	Cat#G8795; RRID:AB_1078991
4EBP1	Cell Signaling	Cat#9452S; RRID:AB_331692
S6	Cell Signaling	Cat#2217; RRID:AB_331355
S6K1	Cell Signaling	Cat#9202S; RRID:AB_331676
pS6	Cell Signaling	Cat#2211; RRID:AB_331679
pS6K1	Cell Signaling	Cat#9205; RRID:AB_330944
Flcn	Baba et al., 2006	N/A
Tsc2	Cell Signaling	Cat#4308S; RRID:AB_10547134
Biological Samples		
Skin and blood samples from 26 affected patients and their parents.	Centre Hospitalier Universitaire Dijon et Université de Bourgogne, France	N/A
Chemicals, Peptides, and Recombinant Proteins		
CHIR99021	Steward lab, Dresden	N/A
PD0325901	Steward lab, Dresden	N/A
FGF	Smith lab, Cambridge	N/A
ActivinA	Smith lab, Cambridge	N/A
KO serum replacement	Thermo Fisher	Cat#10828028
Rapa	Cell Signaling	Cat# 9904
AZD8055	Santa Cruz	Cat#sc-364424
BafilomycinA	Merck	Cat#19-148
ConcanamycinA	Santa Cruz	Cat#sc-202111
Vacuolin-1	Santa Cruz	Cat#sc-216045
Chloroquine	Sigma	Cat#C6628
Puromycin	GIBCO	Cat#A1113803
Blasticidin	GIBCO	Cat#A1113903
HygromycinB	Thermo Fisher	Cat#10687010
G418	Thermo Fisher	Cat#G8168
Paraformaldehyde	Electron Microscopy Sciences	Cat#15710
Lipofectamin RNAiMAX	Thermo Fisher	Cat#13778075
Lipofectamin 2000	Thermo Fisher	Cat#11668019
Dynabeads Protein G for Immunoprecipitation	Thermo Fisher	Cat#10004D
Lysyl Endopeptidase	WAKO	Cat#129-02541
Sequencing Grade Modified Trypsin	Promega	Cat#V5111
Critical Commercial Assays		
Alkaline phosphatase kit	Sigma	Cat#86R-1KT
Golden Gate TALEN and TAL Effector Kit 2.0	Addgene	Cat#1000000024
Gateway BP Clonase II Enzyme mix	Thermo Fisher	Cat#11789020
Gateway LR Clonase II Enzyme mix	Thermo Fisher	Cat#11791020

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
TruSeq Stranded mRNA Library Prep	Illumina	Cat#20020594
SureSelect Human All Exon V5	Agilent	discontinued
SureSelect XT Human All Exon V6 kit	Agilent	discontinued
SeqCap EZ MedExome	Roche	Cat#07676581001
Nextera XT DNA Sample Preparation kit	Illumina	Cat#FC-131-1024
Deposited Data		
RNaseq	this paper; GEO:GSE108057	GEO: GSE108057
Experimental Models: Cell Lines		
Male RGd2 mouse embryonic stem cells	gift from T.Kalkan (Stem Cell Centre, Cambridge, UK); Kalkan et al., 2017	N/A
Male RGd2 mouse embryonic stem cells (mutant and transgenic lines)	this paper	N/A
Female RGd2 mouse embryonic stem cells	gift from M. Leeb (Max Perutz laboratories, Vienna)	N/A
Male O4GIP mouse embryonic stem cells	Betschinger et al., 2013	N/A
Male E14 mouse embryonic stem cells	gift from M. Leeb (Max Perutz laboratories, Vienna)	N/A
Male E14 mouse embryonic stem cells (endogenous 3xFLAG Flcn knockin)	this paper	N/A
HEK293	gift from B. Hemmings (FMI, Basel)	N/A
Oligonucleotides		
Primers used for qPCR, see Table S3	this paper	N/A
siRNAs, see Table S3	this paper	N/A
gRNA sequences, see Table S3	this paper	N/A
Flcn tagging template for genome editing: gtccaccgtccgaagccccacagctacagagtcacggagcGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGtgactccgagaactccttctggaaggtggtgtacagacca	this paper	N/A
UMI containing oligo nucleotide for CRISPR/Cas9 screen gRNA sequencing: 5′-TCCCTACACGACGCTCTTCCGATCTN(5-9)TCTTGTGGAAAGGACGAAACACC-3′	this paper	N/A
Forward primer for CRISPR/Cas9 screen gRNA amplification: 5′-TCCCTACACGACGCTCTTCCGATCT-3′	this paper	N/A
Reverse primer for CRISPR/Cas9 screen gRNA amplification: 5′-GTTCAGACGTGTGCTCTTCCGATCTCTTGTGTAGCGCCAAGTGCC-3′	this paper	N/A
Recombinant DNA		
Genome-wide mouse pooled gRNA library	Addgene; Koike-Yusa et al., 2014	Cat#50947
pPB-LR5.1-EF1a-hph2ACas	this paper	N/A
hCas9	Addgene; Mali et al., 2013	Cat#41815
U6-gRNA	Addgene; Mali et al., 2013	Cat#41824
SV40-ELD	Flemr and Bühler, 2015	N/A
SV40-KKR	Flemr and Bühler, 2015	N/A
pBase	Betschinger et al., 2013	N/A
pPB-CAG-dsRed-pgk-hph	this paper	N/A
pDONR221-Tfe3 (WT and mutant)	this paper	N/A
pDONR221-Flcn	this paper	N/A
pDONR221-Lamtor1	this paper	N/A
pDONR221-RagB/C/D (WT and mutant)	this paper	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pDONR221-lysRagB/C (WT and mutant)	this paper	N/A
pPB-CAG-DEST-pgk-hph	Betschinger et al., 2013	N/A
pPB-CAG-3xFLAG-DEST-pgk-hph	Betschinger et al., 2013	N/A
pPB-TRE-DEST-rTA-HSV-neo	this paper	N/A
pPB-TRE-DEST-rTA-pgk-hph	this paper	N/A
Software and Algorithms		
oligo	https://bioconductor.org/packages/devel/bioc/manuals/oligo/man/oligo.pdf	package version 1.40.0
limma	https://www.bioconductor.org/packages/release/bioc/vignettes/limma/inst/doc/usersguide.pdf	package version 3.32.0
QuasR	Gaidatzis et al., 2015	package version 1.16.0
edgeR	https://www.bioconductor.org/packages/release/bioc/vignettes/edgeR/inst/doc/edgeRUsersGuide.pdf	package version 3.18.0
Bowtie	Langmead et al., 2009	version 4.4.7
MACS2	Zhang et al., 2008	package version 2.1.1.20160309
ComplexHeatmap	https://bioconductor.org/packages/release/bioc/manuals/ComplexHeatmap/man/ComplexHeatmap.pdf	package version 1.18.1
MaxQuant	http://www.coxdocs.org/doku.php?id=maxquant:start	version 1.5.3.8
FlowJo	FlowJo	version 9.9.6
Fiji	Fiji	version 1.0
ZEN	Zeiss	Version 14.0.12.201
Photoshop	Adobe	CS6

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to the Lead Contact, Joerg Betschinger (joerg.betschinger@fmi.ch).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mouse ESCs

Male RGd2 mouse embryonic stem cells (ESCs) containing a Rex1:GFPd2-IRES-Blasticidin reporter (Kalkan et al., 2017) are a gift from T.Kalkan (Stem Cell Centre, Cambridge, UK) and female RGd2 cells a gift from M. Leeb (Max Perutz laboratories, Vienna). Male O4GIP ESCs contain a GFP-IRES-Puromycin transgene under control of an *Oct4* regulatory element (Betschinger et al., 2013). Male E14 ESCs are a gift from M.Leeb.

Patients

We performed 26 trios whole exome sequencing on patient's affected skin-derived DNA and parental blood sample. The study was approved by the ethics committee of Dijon University Hospital and written informed consent from all subjects or their legal representatives was obtained.

METHOD DETAILS

Cell Culture

ESCs were cultured on plastic coated with gelatin or laminin (Sigma) in N2B27 medium (DMEM/F12 (Life Technologies), Neurobasal (GIBCO) supplemented with N2 (homemade) and B-27 Serum-Free Supplement (GIBCO), 1mM sodium pyruvate (GIBCO), 2mM L-glutamine (GIBCO), 0.1 mM 2-mercaptoethanol (Sigma)), and, where indicated, with 2i (3μM CHIR99021 and 1μM PD0325901 (Steward lab, Dresden)), KO serum replacement (Thermo Fisher), Rapa (Cell Signaling), AZD8055 (Santa Cruz), BafilomycinA (Merck), ConcanamycinA (Santa Cruz), Vacuolin-1 (Santa Cruz), Chloroquine (Sigma), 4-Hydroxytamoxifen (0.1μM, Sigma), Doxycycline (0.1μg/ml, Sigma), FGF and ActivinA (Smith lab, Cambridge). For differentiation, ESCs were plated on gelatin-coated plates at

1.5×10^4 cells/cm2 and, the following day, 2i was withdrawn. For EpiLC differentiation (Hayashi et al., 2011), KO serum replacement (1%), FGF (12ng/ml) and ActivinA (20ng/ml) were additionally added. For differentiation of starved cells, ESCs grown on plastic or glass plates were incubated in R-N2B27 (amino-acid and glucose free RPMI (homemade), with B27 and N2), and, as indicated, supplemented with 3.8mg/ml glucose (BioConcept), aas (GIBCO, MEM non-essential aas, MEM essential aas, 2mM L-glutamine), and differentiation was induced by 2i removal. For aa refeeding experiments, ESCs were incubated for 30 minutes in HBSS (Thermo Fisher) containing 3.8mg/ml glucose and then switched to N2B27 for another 30minutes. For lysosomal activity assays, cells were incubated for 1 hour with 5μg/ml DQ-BSA (DQ Red BSA, Life Technologies, D12051) and Lysotracker (LysoTracker Deep Red, Life Technologies, L12492) diluted 1:10000 in 2i medium. After washing in PBS, fluorescence was measured on a LSRII SORP Analyzer (Becton Dickinson) and quantified using FlowJo (FlowJo, LLC). siRNA (detailed in Table S3) and plasmid transfections using Lipofectamin RNAiMAX and 2000 (Thermo Fisher), respectively, and exit from self-renewal of O4GIP and RGd2 ESCs was performed as described before (Betschinger et al., 2013). Briefly, differentiating O4GIP cells were treated with 2i medium plus 1μg/ml Puromycin (GIBCO) and uncommitted cells quantified after 3 days by adding Alamar Blue (Invitrogen), diluted 1:10 in 2i medium, which was subsequently read out on a SpectraMax Gemini EM (Molecular Devices) microplate reader. Self-renewal of differentiating RGd2 cells was quantified by plating single cells on laminin (Sigma)-coated plates, with resulting clones stained for alkaline phosphatase activity (Sigma) according to the manufacturer's instruction and counted.

Immunostaining

Cell were seeded on laminin-coated 96well glass plates (Greiner Bio-One), fixed with 4% paraformaldehyde (Electron Microscopy Sciences) and immunostained as described (Betschinger et al., 2013). Images were captured using a LSM 710 scanning head confocal microscope (Zeiss) and handled using Fiji and Adobe Photoshop. Primary antibodies used were anti-FLAG (Sigma, F1804, 1:1000), anti-LAMP1 (DSHB, 1D4B, 1:500), anti-Lamtor1 (Cell Signaling, 8975S, 1:500), anti-RagB (Cell Signaling, 8150S, 1:100), anti-RagC (Cell Signaling, 9480S, 1:100) and anti-Tfe3 (Sigma, HPA023881, 1:1000). DNA was visualized with Hoechst33342 (Life Technologies). Quantification of nucleo-cytoplasmic Tfe3 ratios using CellProfiler (Broad Institute) was performed as described (Betschinger et al., 2013). Briefly, nuclei and cells were identified and the average fluorescence of nuclei and cytoplasm (cell-nuclei) determined. For the data presented in Figure S1D, 198 (negative siRNA), 208 (Tcf7l1 siRNAs), 165 (Tsc2 siRNAs), 97 (Flcn siRNAs), 116 (Fnip1/2 siRNAs), 153 (Lamtor1 siRNAs), 148 (Lamtor2 siRNAs), 272 (Lamtor3 siRNAs), 195 (RagA siRNAs), 257 (RagB siRNAs), 133 (RagA/B siRNAs), 264 (RagC siRNAs), 236 (RagD siRNAs), and 360 (RagC/D siRNAs) cells were quantified.

CRISPR/Cas9 screen

Lentiviral gRNA libraries (Addgene 50947) (Koike-Yusa et al., 2014) were generated in HEK293 cells. Stable Cas9 expressing female RGd2 ESC (gift from Martin Leeb, Max Perutz laboratories, Vienna) clones were derived by transfecting with pPB-LR5.1-EF1a-hph2ACas9 (derived from pPB-LR5.1-EF1a-puro2ACas9, gift of Kosuke Yusa, Wellcome Trust Sanger Institute) and 7.2x10^6 cells of two independent clones were infected with the lentiviral gRNA library at a 20-fold coverage and multiplicity of infection of 0.25. Stable integrations were selected for 6d in the presence of 0.1 μg/ml Puromycin (GIBCO) and 10^7 cells for each clone were plated on Laminin-coated dishes in the absence of 2i, while 2x10^7 were used for generation of input sequencing libraries. After 3d of differentiation, 2x10^7 cells were passaged in N2B27 containing 1μg/ml Blasticidin (GIBCO) to select for Rex1 expression, and cells at passage 4 and 5 taken for generation of experimental sequencing libraries. Uninfected control cells were lost after the second passage. Genomic DNA was isolated using Proteinase K (Roche) digestion and phenol-chloroform extraction (Sigma).

GENOME EDITING

For C-terminal 3xFLAG Flcn tagging, the TALEN was assembled using Golden Gate TALEN cloning kit (Cermak et al., 2011) (Addgene 1000000024) into acceptor vectors SV40-ELD and SV40-KKR (Flemr and Bühler, 2015) (FlcnC_F-Q3ELD: HD HD NN NI NI NN HD HD HD HD NI HD NI NN HD NG, FlcnC_R-Q3KKR: HD HD NI NN NI NI NN NN NI NN NG NG HD NG HD NN NN NI NN NG). Male E14 ESCs were transiently transfected with a recombination reporter (Flemr and Bühler, 2015), TALENs and Flcn tagging template (gtccaccgtccgaagccccacagctacagagtcacggagcGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGtgactccgagaactccttctggaaggtggtgtacagacca) and derivative clones were genotype for successful recombination. CRISPR/Cas9 genome editing was performed by transient transfection of hCas9 and U6-gRNA plasmids (Mali et al., 2013) (Addgene plasmids 41815 and 41824) into male RGd2 ESCs. Target gRNA sequences are detailed in Table S3. For KO ESC clone generation, a dsRed expression plasmid was cotransfected and single dsRed expressing cells were deposited into 96well plates 2d later. For screen validation, transfected cell pools were differentiated by 2i withdrawal after 2d and serially passaged in N2B27 containing 1μg/ml Blasticidin.

Molecular biology

Coding sequences were amplified from ESC cDNA and recombined into pDONR221 using Gateway technology (Thermo Fisher). The lysosomal targeting sequence of Lamtor1 (aas 1-39) (Nada et al., 2009) was inserted by In-Fusion cloning (Clontech) and point mutations (RagGDP (RagB: T54N, RagC: S75N, RagD: S77L), RagGTP (RagB: Q99L, RagC:Q120L, RagD:Q121L), otherwise indicated) were introduced by polymerase chain reaction (PCR). In Tfe3(10xA), S545,S550,S551,S553,S557,S560,S561,S564,S565 and S567 are mutated to alanine residues. Expression destination vectors were pPB-CAG-DEST-pgk-hph (Betschinger et al., 2013), pPB-CAG-3xFLAG-DEST-pgk-hph (Betschinger et al., 2013), pPB-TRE-DEST-rTA-HSV-neo and pPB-TRE-DEST-rTA-pgk-hph, and were selected for stable integration after co-transfection with pBASE (Betschinger et al., 2013) into ESCs in the presence of 150 μg/ml HygromycinB (Thermo Fisher) or 200 μg/ml G418 (Thermo Fisher). For mRNA quantification, total cellular RNA was isolated using RNeasy Mini kit (QIAGEN), complementary DNA generated using SuperScript III reverse transcriptase (Invitrogen), and quantitative PCR performed using TaqMan Fast Universal PCR master mix (Applied Biosystems) with gene specific primers and a GAPDH probe (Applied Biosystems) for normalization. We used the universal probe library (UPL, Roche) and Taqman system (Applied Biosystems). Oligonucleotide sequences and probes are specified in Table S3.

Protein methods

Immunoprecipitations using FLAG antibodies (Sigma, F1804) of wild-type ESCs expressing 3xFLAG-tagged Rag GTPase mutants, and Tfe3 antibodies (Sigma, HPA023881) in three biological replicates of *Tfe3*^{WT}, *Lamtor1*^{WT}, *Tfe3*$^{KO.6}$, *Lamtor1*$^{KO.1}$, *Flcn*$^{KO.1}$ and wild-type ESCs expressing empty vector, RagC^{GDP} and RagD^{GDP} were performed as described before (Betschinger et al., 2013) using Dynabeads (Thermo Fisher). Proteins were digested on the beads first with 0.2μg LysC (WAKO) for 6h at 37°C in 2.5M Guanidin-HCl, containing 20mM EPPS pH8.5, 10mM CAA and 5mM TCEP in a total volume of 6μl. Then, samples were diluted with 18μl 50mM HEPES pH 8.5 containing 0.2μg modified porcine Trypsin (Promega) and cleaved overnight at 37°C. Another 0.2μg Trypsin was added and the cleavage continued for 4 hours. The generated peptides were acidified with 1μl of 20% TFA and analyzed by capillary liquid chromatography tandem mass spectrometry with an EASY-nLC 1000 using the two-column set up (Thermo Fisher). The peptides were loaded with 0.1% formic acid, 2% acetonitrile in H2O onto a peptide trap (Acclaim PepMap 100, 75um x 2cm, C18, 3μm, 100Å) at a constant pressure of 800 bar. Peptides were separated, at a flow rate of 150 nl/min with a linear gradient of 2%–6% buffer B in buffer A in 3 minutes followed by a linear increase from 6 to 22% in 40 minutes, 22%–28% in 9 min, 28%–36% in 8min, 36%–80% in 1 min and the column was finally washed for 14 min at 80% B (Buffer A: 0.1% formic acid, buffer B: 0.1% formic acid in acetonitrile) on a 50μm x 15cm ES801 C18, 2μm, 100Å column mounted on a DPV ion source (New Objective) connected to a Orbitrap Fusion (Thermo Fisher). The data were acquired using 120000 resolution for the peptide measurements in the Orbitrap and a top T (3 s) method with HCD fragmentation for each precursor and fragment measurement in the LTQ according the recommendation of the manufacturer (Thermo Fisher).

Cell lysates for western blotting were generated in RIPA buffer (50 mM Tris, pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% Tx-100, 0.1% SDS). Primary antibodies were anti-FLAG (Sigma, F1804, 1:100), anti-GAPDH (Sigma, G8795, 1:1000), anti-4EBP1 (Cell Signaling, 9452S, 1:100), anti-S6 (Cell Signaling, 2217, 1:100), anti-S6K1 (Cell Signaling, 9202S, 1:100), anti-pS6 (S235/236) (Cell Signaling, 2211, 1:100), anti-pS6K1 (T389) (Cell Signaling, 9205, 1:100), anti-Lamtor1 (Cell Signaling, 8975S, 1:500), anti-Flcn (Baba et al., 2006), anti-Tfe3 (Sigma, HPA023881, 1:1000) and anti-Tsc2 (Cell Signaling, 4308S, 1:1000).

QUANTIFICATION AND STATISTICAL ANALYSIS

CRISPR/Cas9 screen quantification

For gRNA quantification, we developed an approach using unique molecular identifiers (UMIs). Biological replicates were divided into aliquots of 500ng genomic DNA (input samples: 8μg total; selected samples: 2μg total). An oligo nucleotide containing UMIs and the 3′ part of the Illumina P5 sequence (5′-TCCCTACACGACGCTCTTCCGATCTN(5-9)TCTTGTGGAAAGGACGAAACACC-3′) was integrated using Phusion DNA polymerase (Thermo Fisher) in a single round of denaturation, annealing (62°C) and extension. Then 2μl of exonuclease I (NEB) were added and incubated at 37°C for 45min, followed by heat inactivation and purification using AMPure XP beads (Agencourt) with a 1:1 ratio. gRNAs were amplified using Phusion DNA polymerase for 25 cycles (Fwd: 5′-TCCCTACAC GACGCTCTTCCGATCT-3′; Rev: 5′-GTTCAGACGTGTGCTCTTCCGATCTCTTGTGTAGCGCCAAGTGCC-3′; annealing temperature of 64°C)) and purified using AMPure XP beads (1:1 ratio, elution in 22μl). 5μl were used in an Illumina indexing PCR reaction using Phusion DNA polymerase (6cycles) and NEBnext multiplex oligos set 1 (NEB). Samples were purified using AMPure XP beads (1:1 ratio) and sequenced on one lane of HiSeq2500 (high output mode, 50bp single-end reads). UMI and gRNA sequences were extracted from individual reads by searching for exact matches to the vector backbone sequence at expected offsets (more than 76% of all reads). About 95% of extracted gRNA sequences corresponded to one of the expected sequences from the gRNA library (Koike-Yusa et al., 2014), and the number of UMIs was counted for each gRNA. Counts of UMIs were 2.8 to 7.7-fold lower than read counts, indicating a low to moderate level of PCR duplication. As read versus UMI counts displayed a linear relationship, no correction for saturation of highly abundant gRNAs was applied. gRNAs with at least one non-zero count in any sample were selected for further analysis (84,471 gRNAs, 96.6% of the library). Counts from individual replicates for each of the six libraries (2 input, 4 experimental) were summed and analyzed by MAGeCK (version 0.5.4) (Li et al., 2014) using parameters "–norm-method total–gene-test-fdr-threshold 0.25–adjust-method fdr–sort-criteria pos–remove-zero none" to identify gRNAs and genes significantly enriched in the experimental compared to the input libraries.

Protein identification and quantification

Relative quantification of mass-spec data was performed with MaxQuant version 1.5.3.8 using Andromeda as search engine and label free quantification as described (Ostapcuk et al., 2018). The mouse subset of the UniProt version 2015_01 combined with the contaminant DB from MaxQuant was searched and the protein and peptide FDR were set to 0.01. Only proteins identified with at least three peptides were plotted in Figures 6A and 6B.

Microarray data analysis

Raw data (CEL files) were downloaded (all publicly available sequencing data used in this study are indicated in Table S3), background corrected and normalized using rma from oligo (https://bioconductor.org/packages/devel/bioc/manuals/oligo/man/oligo.pdf) (package version 1.40.0). A single probeset was selected per gene using annotation from the Bioconductor package mouse4302.db (package version 3.2.3), selecting the probeset with the largest interquartile range for each gene. Differentially expressed genes between Myc inhibition and DMSO samples were identified using limma (https://www.bioconductor.org/packages/release/bioc/vignettes/limma/inst/doc/usersguide.pdf) (package version 3.32.0) with default parameters. Genes were considered as significantly deregulated with a minimum absolute fold-change of 2 and a p value smaller than 0.01.

RNaseq and GROseq data analysis

Cellular RNA from ESCs, cells 34h after 2i withdrawal, and long-term differentiated cells sorted for Rex1 expressing cells was purified using RNAeasy (QIAGEN), strand-specific RNA-seq libraries were generated using TruSeq mRNA Library preparation kit (Illumina), and libraries sequenced on an Illumina HiSeq2500 machine (50bp single-end reads). Reads were aligned using qAlign from the Bioconductor package QuasR (Gaidatzis et al., 2015) (package version 1.16.0) to mouse GRCm38/mm10 genome with default parameters except for splicedAlignment = TRUE. EpiLC data were sequenced as 36bp reads (Buecker et al., 2014), and therefore no spliced alignment could be performed. *Tsc2* KO and *Nprl2* KO ESC data (Li et al., 2018) as well as global run-on sequencing (GROseq) in ESCs (Dorighi et al., 2017) were sequenced as 75bp and 50bp paired end reads respectively, and therefore paired = ”fr” was used. For *in vivo* embryo data (Boroviak et al., 2015) pre-existing alignments to mouse GRCm38/mm10 genome downloaded from ArrayExpress (E-MTAB-2958) were used. Alignments were quantified with qCount from the Bioconductor package QuasR (Gaidatzis et al., 2015) (package version 1.16.0) for known UCSC genes obtained from the TxDb.Mmusculus.UCSC.mm10.knownGene package (package version 3.4.0) using default parameters (Table S1).

For identification of genes specifically regulated in either ESCs or during differentiation, we used generalized linear models. Only genes with at least 3 counts per million in at least two biological samples were considered (ESCs: 12,478 genes; differentiation: 12,709 genes). Statistically significantly deregulated genes were identified using edgeR (https://www.bioconductor.org/packages/release/bioc/vignettes/edgeR/inst/doc/edgeRUsersGuide.pdf) (package version 3.18.0) and these genes were fitted to five separate generalized linear models for different sets of samples:

a) ~*genotype*: *Flcn* KO or *Lamtor1* KO ESCs compared to respective control *Flcn*^{WT} and *Lamtor1*^{WT} ESCs.
b) ~*treatment*: 3h Tam or 34h Tam-treated Tfe3-ERT2 ESCs compared to 3h or 34h ethanol-treated Tfe3-ERT2 ESCs.
c) ~*time*: *Flcn*^{WT} and Ethanol-treated Tfe3-ERT2 ESCs and 34h differentiated progeny
d) ~*time* + *treatment*: 34h Tam-treated differentiated Tfe3-ERT2 cells compared to ethanol-treated Tfe3-ERT2 ESCs and 34h differentiated progeny. *treatment* therefore identifies differentiation-specific nuclear Tfe3 targets.
e) ~*time* + *genotype* + *time:genotype*: *Flcn*$^{KO.1}$ ESCs and differentiated progeny compared to *Flcn*^{WT} ESCs and differentiated progeny. *time:genotype* therefore identifies differentiation-specific Flcn targets.

Raw P values were corrected for multiple testing by calculating false discovery rates (FDR). Significant effects in each contrast were defined as changes with a minimum absolute fold-change of 2 and a FDR less than 0.01.

For visualization in Figure 2A and S2B, calculations of RNA log2 fold changes from edgeR were used and only significantly deregulated genes in at least one ESC condition (*genotype*, model (a) and *treatment*, model (b)) were considered (total: 2,483 genes). For visualization in Figure 2C, significantly deregulated genes in at least one differentiation condition (*time*, model (c), *treatment*, model (d) and *time:genotype*, model (e)) were considered (total: 3,370 genes). Read counts were normalized (dividing by the total number of aligned reads and multiplying with minimal library size, and adding a pseudocount of 8) and log2 transformed. Centered expression values were generated by calculating differences in log-space of each gene in each sample to its average over all samples. For log2 fold change comparisons shown in Figure S2A and S2B, ESC genes detected in this study (12,478 genes) and annotated in other datasets (total: 11,770 genes) were considered. Log2 fold changes were calculated using log2-transformed normalized read counts as described above, averaging biological replicates, and calculating differences (corresponding to fold-changes in linear space) between conditions. Boxplots (Figures S1D and S2B) were generated using the boxplot function in R with default parameters and outline = FALSE. Differentially expressed genes upon transient TFEB overexpression in HeLa cells (Sardiello et al., 2009) were converted to mouse homologs using HomoloGene (https://www.ncbi.nlm.nih.gov/homologene) allowing only 1:1 mapping of genes. Analyses of enriched gene sets (Figure 2C; Table S1) were performed using DAVID (https://david.ncifcrf.gov) for GO terms of biological processes, cellular components and molecular functions. Principal component analysis (Figure 2B) was performed on normalized, log2-transformed and mean-centered expression data in R using the prcomp function.

ChIPseq and DHSseq data analysis

Published ChIPseq and DNaseI hypersensitive site sequencing (DHSseq) reads were aligned using qAlign from the Bioconductor package QuasR (Gaidatzis et al., 2015) (package version 1.16.0) to mouse GRCm38/mm10 genome with default parameters. H3K4me3 and H3K27me3 ChIP-seq data in ESCs and EpiLCs (Kurimoto et al., 2015) were sequenced on a SOLiD 5500xl platform, and therefore Bowtie (Langmead et al., 2009) (version 4.4.7) with parameter -C for alignment of reads in colorspace was used. Alignments were sorted and indexed using SAMtools (package version 1.2), and quantified with qCount from the Bioconductor package

QuasR (Gaidatzis et al., 2015) (package version 1.16.0). IP enrichments were calculated using log2((nIP / tIP * c + p) / (nInput / tInput * c + p), where nIP and nInput are the read counts in a region in IP and input samples, respectively, tIP and tInput are the total number of aligned reads in the two samples, c is a constant scaling factor (the minimal number of aligned reads in any sample) and p is a pseudocount to stabilize enrichments that are based on low numbers of reads (p = 8).

Tfe3 peaks were called on Tfe3 ctrl and shFlcn ChIP in ESCs (Betschinger et al., 2013) using MACS2 (Zhang et al., 2008) (package version 2.1.1.20160309) with default parameters. Peaks with at least twofold enrichment in IP over IgG in at least one out of the two Tfe3 ChIP samples were considered (total: 13,512 peaks) (Table S1). Tfe3 peaks were assigned to genes by calculating distance of peak midpoint to the nearest transcriptional start site (TSS), using a set of non-redundant TSSs with a single start site randomly selected for each gene. Peaks with a maximum distance less than 2000bp to the nearest TSS were classified as proximal (total: 2,053 peaks) and with a greater distance than 2000bp as distal (total: 11,459). Reads of H3K27ac and H3K4me1 in ESCs and EpiLCs (Buecker et al., 2014) were counted in Tfe3 peak regions and distal sites with enrichment of H3K27ac in either ESCs or EpiLCs greater than twofold in IP over input or H3K4me1 in either ESCs or EpiLCs greater than 1.5 fold in IP over input were considered as putative Tfe3 enhancers (total: 2,387) (Table S1).

Genomic profiles for heatmaps centered on the Tfe3 peak midpoint (Figure S2E) were generated with qProfile from the Bioconductor package QuasR (Gaidatzis et al., 2015) (package version 1.16.0), and visualized using ComplexHeatmap (https://bioconductor.org/packages/release/bioc/manuals/ComplexHeatmap/man/ComplexHeatmap.pdf) (package version 1.18.1). Except for Tfe3 ChIP and GROseq samples, averaged replicates are represented.

Enrichment or depletion of specific genes sets in gene clusters (Figure 2D and Figure S2F) were quantified as standardized residuals, calculated using the chisq.test function in R.

Whole exome sequencing (WES)

Genomic DNA from fresh skin, cultured skin fibroblasts, and blood samples was extracted using the Gentra Puregene Blood and Tissue Extraction Kit (QIAGEN). Genomic DNA integrity and quantity was assessed by agarose gel electrophoresis, NanoDrop spectrophotometry, and Qubit fluorometry (Thermo Fisher). Exome capture and sequencing, except for Patients 3 and 4, was performed at Integragen (Evry, 209 France) from 1 μg of genomic DNA per individual using the Agilent SureSelect Human All Exon V5. Libraries were sequenced on a HiSeq platform (Illumina) using paired-end 75-bp reads. Sequences were aligned to the human genome reference sequence (GRCh37/hg19 212 build of UCSC Genome Browser), and single-nucleotide variants and small insertions/deletions were systematically detected as previously described (Allen et al., 1992). Candidate *de novo* mutational events were identified by focusing on protein-altering and splice-site changes: (1) supported by at least three reads and 10% of total reads in the patient; (2) absent in both parents, as defined by variant reads representing less than 5% of total reads; (3) at base-pair positions covered by at least four reads in the entire trio; and (4) present at a frequency less than 1% in dbSNP (build 147) and 0.1% in the Exome Aggregation Consortium (ExAC Browser, exac.broadinstitute.org).

WES for Patient 3 was enriched using the SureSelect XT Human All Exon V6 kit (Agilent) and sequenced (paired end 150bp) on a NextSeq500 sequencing system (Illumina) at a mean exome depth of 120X. The exome was defined as all coding exons of UCSC and Ensembl ± 20bp intron flanks. This coverage was > 99,5% for known intellectual disability genes, and 96.3% of the exome was covered at least 15X. Reads were aligned to hg19 using BWA (BWA-MEM v0.7.5a) and variants were called using the GATK haplotype caller (v3.4.46). Detected variants were annotated, filtered and prioritized using the Bench NGS Lab platform (Cartagenia, Leuven, Belgium), and confirmed by Sanger sequencing.

Libraries for WES of Patient 4 were prepared using the Kapa HTP kit (Illumina, San Diego, CA, USA) and capture was performed using the SeqCap EZ MedExome (Roche NimbleGen Madison, WI, USA). Sequencing was done on an Illumina HiSeq2500 HTv4 (Illumina, San Diego, CA, USA) with paired-end 125-bp reads. Read alignment to GRCh37 (hg19) and variant calling were done with a pipeline based on BWA-MEM0.7 and GATK 3.3.0. The median coverage of the captured target region was at least 100x. Variant annotation and prioritizing were done using Cartagenia Bench Lab NGS (Agilent Technologies). We excluded variants located outside the ± 6 nucleotide exons and intron/exon boundaries and variants with a minor allele frequency (MAF) of > 1% in control databases, including dbSNP137 (https://www.ncbi.nlm.nih.gov/projects/SNP), 1000 Genomes Project (release of February 2012), and Exome Variant Server (EVS), NHLBI Exome Sequencing ProjectNational Heart, Lung, and Blood Institute GO Exome Sequencing Project (ESP5400 release) (http://evs.gs.washington.edu/EVS/) and our in-house exome controls.

TFE3 sequencing

Regions of interest were amplified using custom intronic primers and long-range polymerase chain reactions with the PrimeSTAR GXL DNA Polymerase (Takara Bio, Saint-Germain-en-Laye, France). Polymerase chain reaction amplicons were pooled, purified, and quantified from each affected individual. Sequencing libraries were prepared using the Nextera XT DNA Sample Preparation kit (Illumina, Paris, France). Paired-end sequencing of 150-bp reads was performed on a MiSeq platform using 300-cycle reagent kits (v2; Illumina, Paris, France).

Complementary DNA analysis

Complementary DNA was sequenced from the first patient's fresh skin, and second and fifth patient's fibroblasts. Total RNA was isolated with TRIzol Reagent (Life Technologies) from fibroblasts, according to the manufacturer's instructions. 1μg RNA was transcribed into cDNA with the QuantiTect Reverse Transcription Kit (QIAGEN). Using PCR primers positioned in exons 2 and 6 in *TFE3* gene, complementary DNA was amplified by polymerase chain reaction (PCR) using Prime STAR GXL kit (Takara) according to the manufacturer's protocol. The libraries were prepared with the Nextera XT DNA Sample Preparation Kit (Illumina). Generated libraries were sequenced on a MiSeq instrument (Illumina) according to the manufacturer's recommendations for paired-end 150 bp reads.

Determination of the X chromosome inactivation pattern at the HUMARA locus

X chromosome inactivation (XCI) pattern was estimated at the *HUMARA* locus as described previously (Allen et al., 1992) on DNA extracted from uncultured fibroblasts. Fluorescent PCR products were analyzed by capillary electrophoresis on an ABI3130XL genetic analyzer (Applied Biosystems) and peak areas were generated with the GeneMapper software (Applied Biosystems). Skewing was defined as greater than 80% of one X allele active.

Details for quantification and statistical analysis in Figures

Details for quantification and statistical analysis are specified in the figure legends, including number of biological or technical replicates, and statistical tests. Data is presented as the average and standard deviation.

DATA AND SOFTWARE AVAILABILITY

The accession number for the data generated in this paper is Gene Expression Omnibus GEO:GSE108057. All other publicly available datasets used are specified in Table S3.

Cell Stem Cell, Volume *24*

Supplemental Information

Lysosomal Signaling Licenses Embryonic Stem Cell Differentiation via Inactivation of Tfe3

Florian Villegas, Daphné Lehalle, Daniela Mayer, Melanie Rittirsch, Michael B. Stadler, Marietta Zinner, Daniel Olivieri, Pierre Vabres, Laurence Duplomb-Jego, Eveline S.J.M. De Bont, Yannis Duffourd, Floor Duijkers, Magali Avila, David Geneviève, Nada Houcinat, Thibaud Jouan, Paul Kuentz, Klaske D. Lichtenbelt, Christel Thauvin-Robinet, Judith St-Onge, Julien Thevenon, Koen L.I. van Gassen, Mieke van Haelst, Silvana van Koningsbruggen, Daniel Hess, Sebastien A. Smallwood, Jean-Baptiste Rivière, Laurence Faivre, and Joerg Betschinger

Figure S1. Related to Figure 1

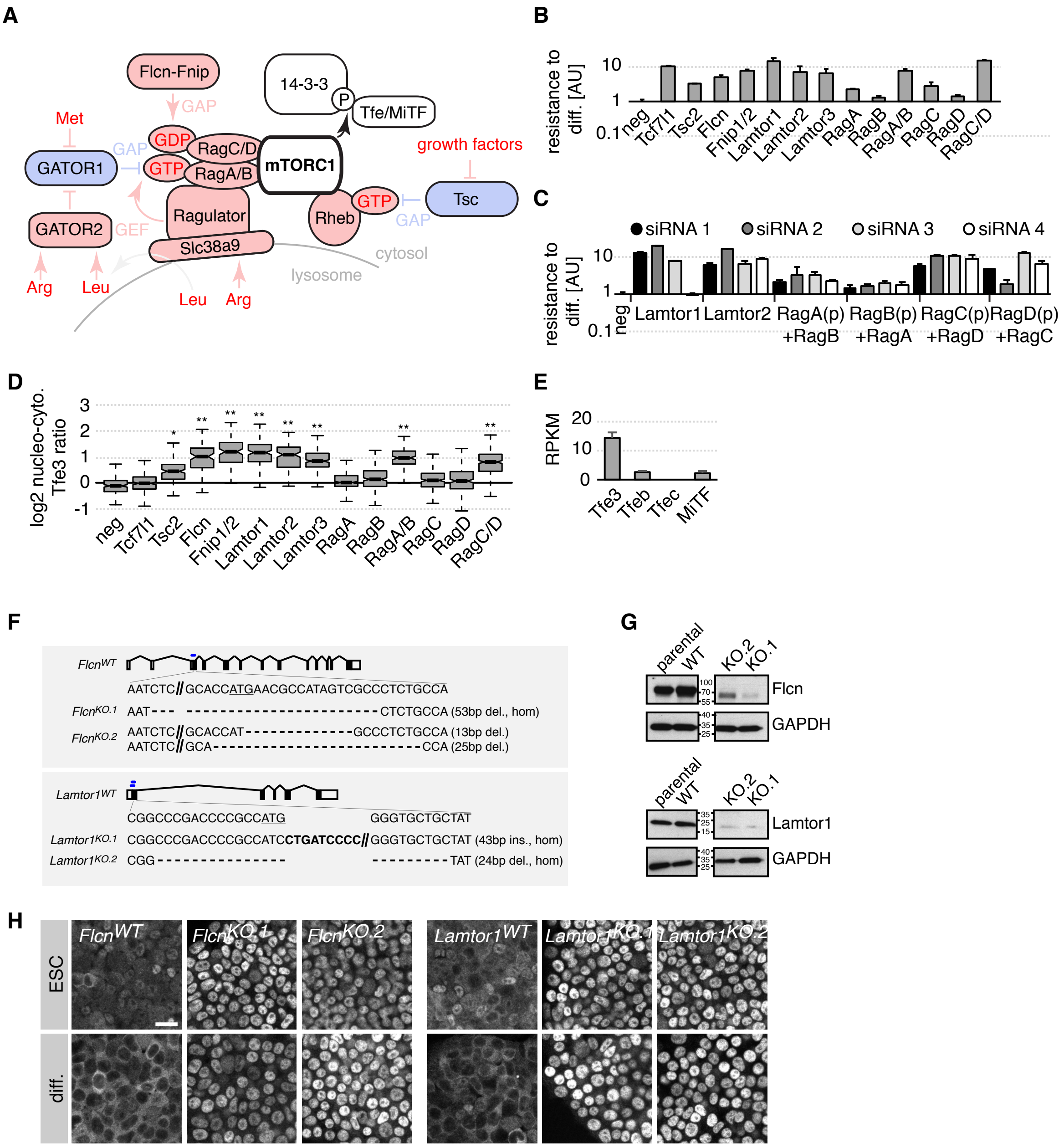


Figure S1, related to Figure 1: Characteristics of lysosomal signaling loss of function mutants.

(**A**) Model of amino acid-dependent mTORC1 regulation. Activators and inhibitors of mTORC1 are indicated in red and blue, respectively.

(**B**,**C**) Resistance to differentiation of O4GIP ESCs transfected with indicated siRNAs after 3 days of 2i withdrawal. Results were normalized to negative (neg) siRNA cells and are shown on a log(10)-scaled axis. Average and SD of 2 technical replicates. (p) denotes pool of 4 siRNAs. Note that the Tcf7l1, Tsc2, Flcn, Fnip1/2 and Lamtor3 siRNA pools have been deconvoluted before (Betschinger et al., 2013).

(**D**) log2 nucleo-cytoplasmic Tfe3 ratios in O4GIP ESCs transfected with indicated siRNAs. Significance was determined using a two-tailed unpaired t-test compared to neg control sample. (*)$<10^{-15}$ and (**)$<10^{-30}$.

(**E**) RPKM (reads per kilobase of transcript per million mapped reads) of MiTF/Tfe transcription family members in wildtype ESCs. Average and SD of two independent RNA-Seq experiments.

(**F**) *Flcn* and *Lamtor1* KO ESC lines used in this study. Location of gRNAs is indicated with blue bars. Top sequence track is the wildtype allele with the ATG start codon underscored. Mutant alleles lack the start codon.

(**G**) Western blot of indicated genotypes confirms absence/severe reduction of proteins.

(**H**) Tfe3 localization in ESCs and after 24 h of differentiation (diff.) in indicated genotypes. Scale bar is 20µm.

Figure S2. Related to Figure 2

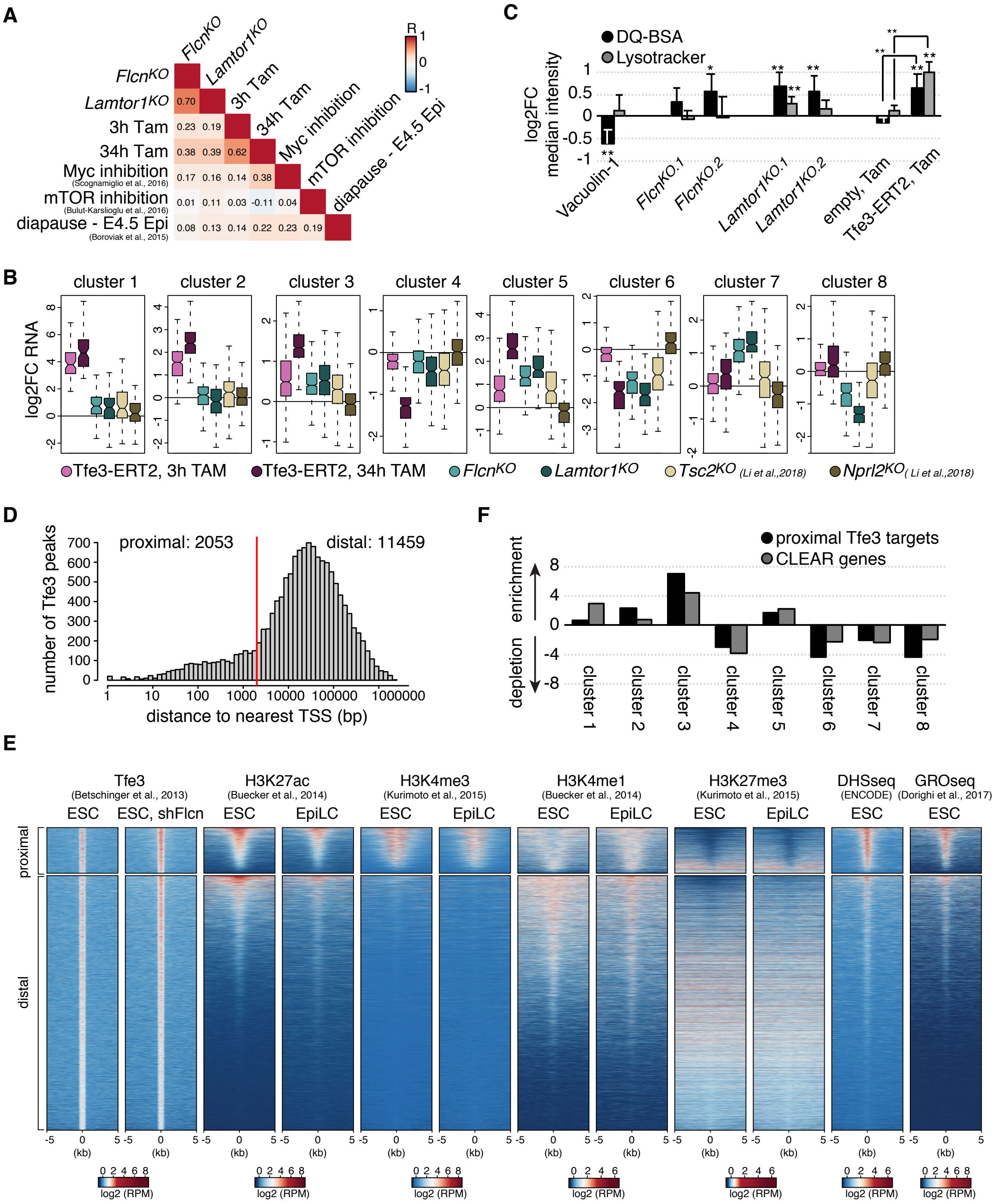


Figure S2, related to Figure 2: Analysis of RNA sequencing data.

(**A**) Pearson's correlation coefficients of pairwise comparisons between log2 fold expression changes of indicated samples. Epiblast (Epi).

(**B**) Boxplot of log2 fold changes of genes in clusters 1-8 of indicated genotypes. Fold change (FC).

(**C**) Log2 fold changes in median fluorescence intensity of DQ-BSA or Lysotracker staining relative to untreated (100nM Vacuolin-1, control and Tfe3-ERT2) and sibling wildtype genotypes (*Lamtor1* and *Flcn* KO ESC lines). Average and SD of 4 biological replicates. Significance was determined using a two-tailed unpaired t-test. (*)<0.1 and (**)<0.05.

(**D**) Distance of Tfe3 peaks to the nearest transcriptional start site (TSS). Cut-off for proximal and distal binding is shown as red line.

(**E**) Heatmap of Tfe3, histone 3 K27 acetylation (H3K27), histone H3 K4 trimethylation (H3K4me3), histone H3 K4 monomethylation (H3K4me1) and histone H3 K27 trimethylation (H3K27me3) ChIPseq read densities as well as DNAseI hypersensitive sites (DHS) and global run-on sequencing (GROseq) read densities across all proximal Tfe3 (top) and distal Tfe3 (bottom) binding sites. Each row represents a 10-kb window centred on the Tfe3 peak midpoint. Rows are sorted for H3K27ac ChIP enrichment in ESCs. Active cis-regulatory regions are transcribed (GROseq), in open chromatin (DHS) and negative for H3K27me3. Of those, promoters are marked by H3K27ac and H3K4me3, and enhancers by H3K27ac and H3K4me1. Reads per million (RPM).

(**F**) Enrichment and depletion (standardized residuals) of genes associated with proximal Tfe3 peaks and CLEAR genes (Sardiello et al., 2009) in indicated clusters.

Figure S3. Related to Figure 3

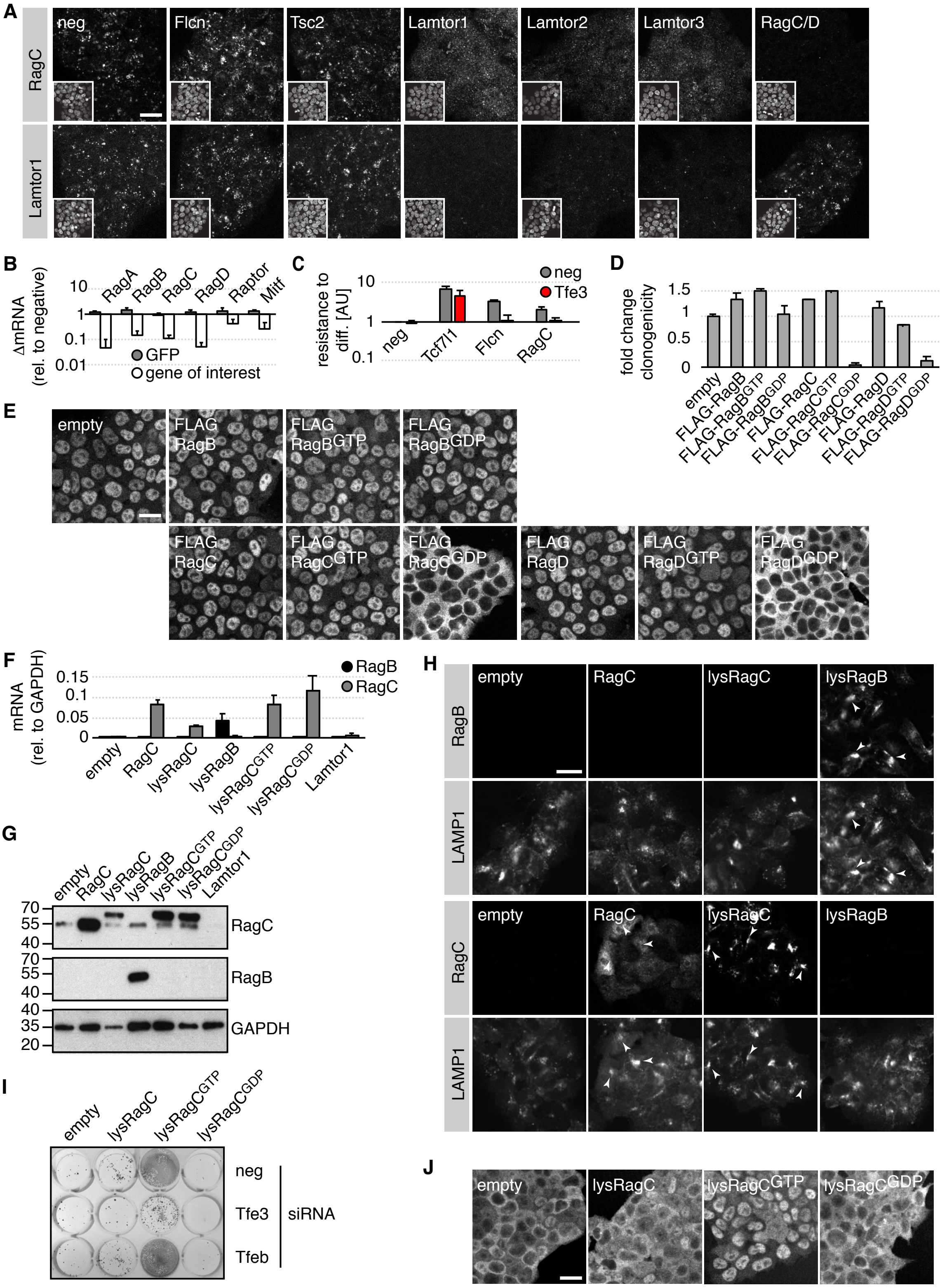


Figure S3, related to Figure 3: Functional analysis of Rag GTPase signaling in ESCs.

(**A**) Subcellular localization of RagC and Lamtor1 in O4GIP ESCs transfected with indicated siRNAs. Insets are DNA staining of the field of view. Note cytoplasmic delocalization of RagC upon knockdown of Lamtor1, 2 and 3 but not Flcn or Tsc2. Scale bar is 20µm.

(**B**) Knockdown of transcripts by indicated siRNAs. Results were normalized to negative siRNA transfected cells and are shown on a log(10)-scaled axis. Average and SD of 2 technical replicates.

(**C**) Resistance to differentiation of O4GIP ESCs transfected with indicated siRNA combinations after 3 days of 2i withdrawal. Results were normalized to negative (neg) siRNA cells and are shown on a log(10)-scaled axis. Average and standard deviation (SD) of 2 biological replicates.

(**D**,**E**) Retention of self-renewal relative to empty vector transfected cells after 3 days of 2i withdrawal (D) and Tfe3 localization (E) in *Flcn* $^{KO.1}$ cells expressing indicated Rag transgenes. Average and SD of 2 technical replicates. Scale bar is 20µm.

(**F**,**G,H**) mRNA relative to GAPDH (F), protein expression (G) and subcellular localization (H) of indicated transcripts/proteins in *Lamtor1*$^{KO.1}$ ESCs expressing indicated constructs. Colocalization of lysRag proteins and LAMP1 is indicated by arrowheads (H). Note that RagB antibodies failed to recognize endogeneous RagB and that the microscope amplification settings for overexpressed RagC proteins did not allow detection of endogeneous RagC (**Figure 3A**). Also note cytoplasmic localization of overexpressed RagC. Average and SD of 2 biological replicates. Scale bar is 20µm.

(**I**) Alkaline phosphatase staining of wildtype cells expressing indicated RagC transgenes that were transfected with indicated siRNAs and subjected to clonal self-renewal after 3 days of 2i withdrawal.

(**J**) Tfe3 localization in wildtype cells expressing indicated RagC transgenes. Scale bar is 20µm.

Figure S4. Related to Figure 4

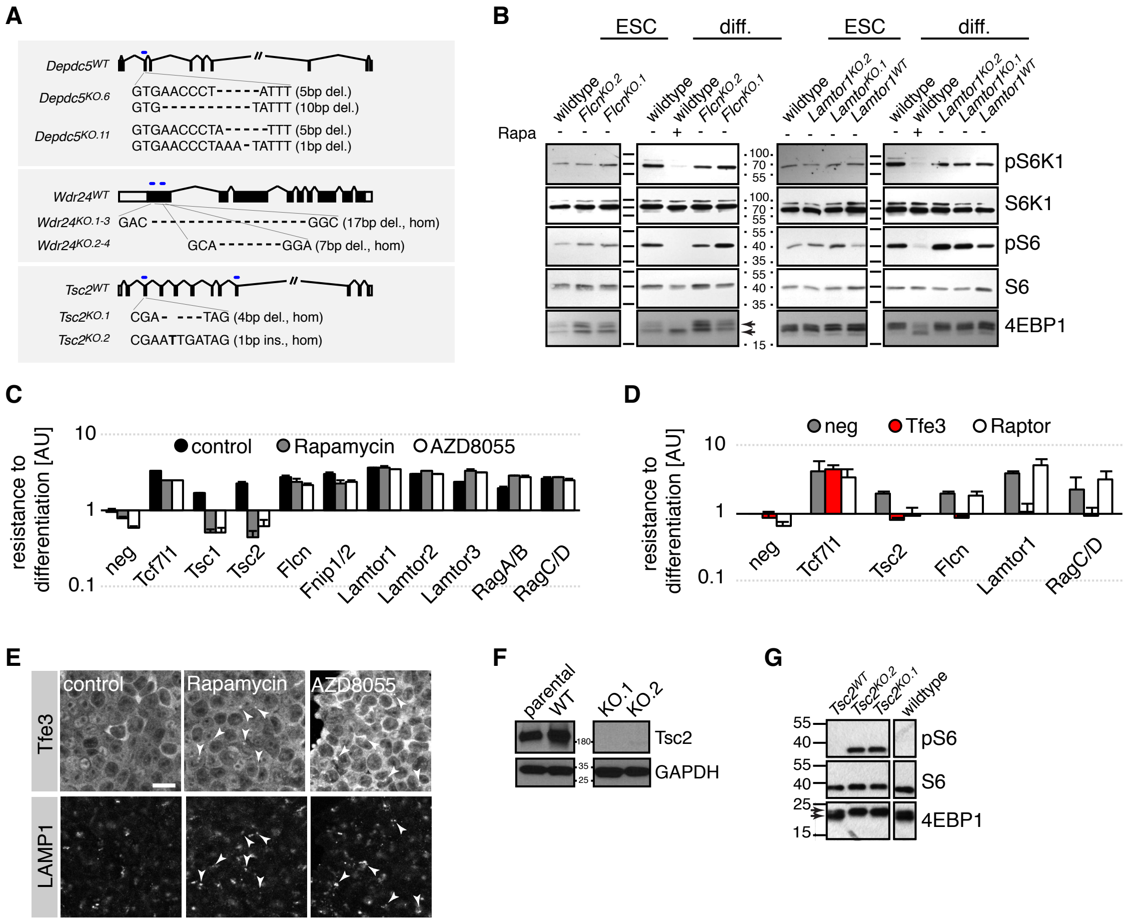


Figure S4, related to Figure 4: ESC-specific mTORC1 regulation.

(**A**) *Depdc5*, *Wdr24* and *Tsc2* KO ESC lines used in this study. Labeling is similar to **Figure S1F**.

(**B**,**G**) Western blot analysis of mTORC1 targets in ESCs and after 48 hours of differentiation of indicated genotypes. 20nM Rapamycin was added where indicated. Arrows mark fast (unphosphorylated) and slow migrating (phosphorylated) 4EBP1.

(**C**,**D**) Resistance to differentiation in O4GIP ESCs transfected with indicated siRNA combinations after 3 days of 2i withdrawal. 20nM Rapamycin and 50nM AZD8055 were added where indicated. Results were normalized to neg siRNA cells and are shown on a log(10)-scaled axis. Average and SD of 2 technical replicates.

(**E**) Tfe3 and LAMP1 localization in ESCs treated with 20nM Rapamycin or 50nM AZD8055 for 24 hours. Arrows indicate colocalization of Tfe3 and LAMP1. Scale bar is 20µm

(**F**) Absence of protein in *Tsc2* KO ESCs.

Figure S5. Related to Figure 5

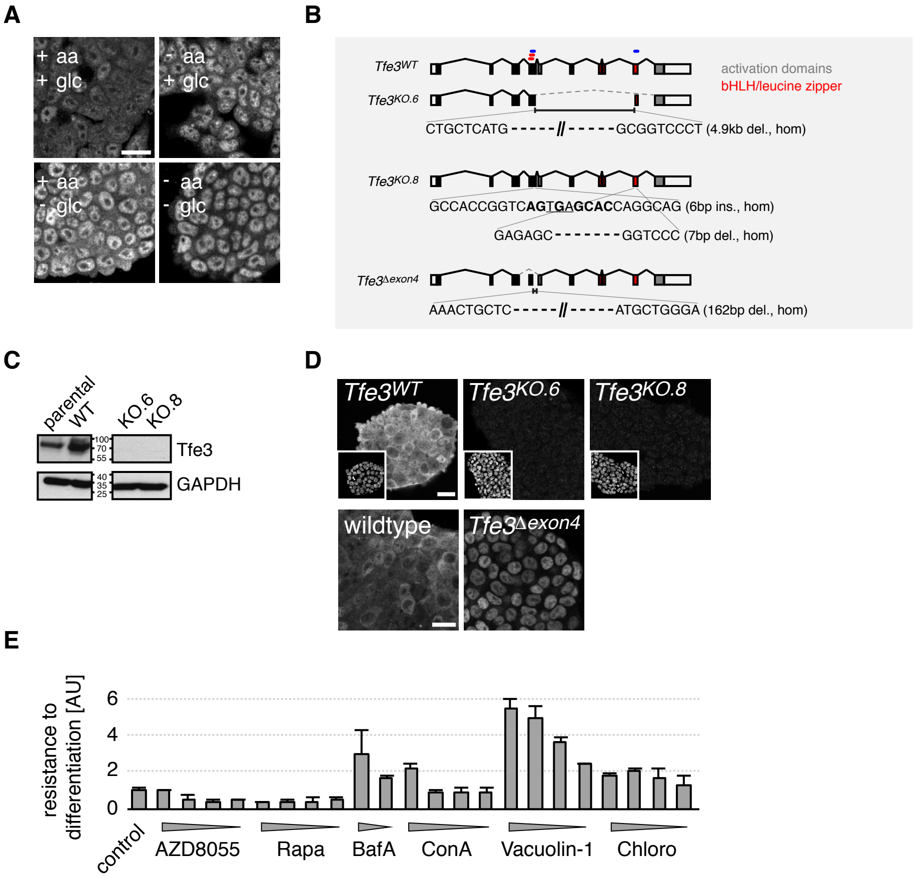


Figure S5, related to Figure 5: Catabolic sensing in ESCs.

(**A**) Tfe3 localization in wildtype ESCs cultured for 2d in indicated nutrient conditions. Scale bar is 20µm.

(**B**) *Tfe3* KO and $Tfe3^{\Delta exon4}$ ESC lines used in this study. Labeling similar to **Figure S1F**. Red bars indicate location of Tfe3 gRNAs recovered in the primary screen and dashed lines exon skipping. In frame stop codon is underscored in the $Tfe3^{KO.8}$ cell lines.

(**C**) Absence of Tfe3 protein in in KO cell lines.

(**D**) Absence of Tfe3 immunoreactivity in $Tfe3^{KO}$ ESC clones with insets showing DNA staining of the field of view (upper). Ectopic nuclear Tfe3 in $Tfe3^{\Delta exon4}$ cells (lower). Scale bar is 20µm.

(**E**) Resistance to differentiation of O4GIP ESCs differentiated for 3 days in the presence of AZD8055 (200nM, 100nM, 50nM, 25nM), Rapamycin (Rapa, 80nM, 40nM, 20nM, 10nM), BafilomycinA (BafA, 2.5nM, 1.25nM), ConcanamycinA (ConA, 20nM, 10nM, 5nM, 2.5nM), Vacuolin-1 (400nM, 200nM, 100nM, 50nM) and Chloroquine (Chloro, 20µM, 10µM, 5µM, 2.5µM). Results were normalized to control treated cells. Average and SD of 2 technical replicates.

Figure S6. Related to Figure 6

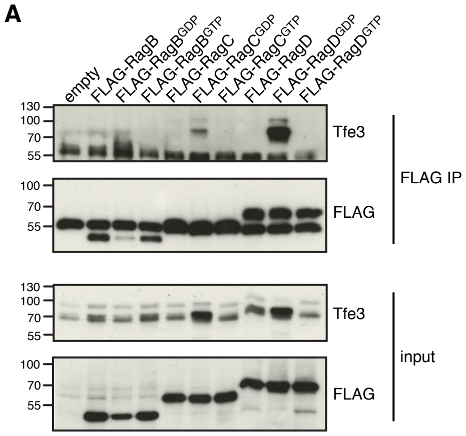


Figure S6, related to Figure 6: RagC/D$^{\mathrm{GDP}}$ interacts with Tfe3.

(**A**) FLAG immunoprecipitates of wildtype ESCs expressing indicated Rag transgenes probed for Tfe3 and FLAG.

Figure S7. Related to Figure 7

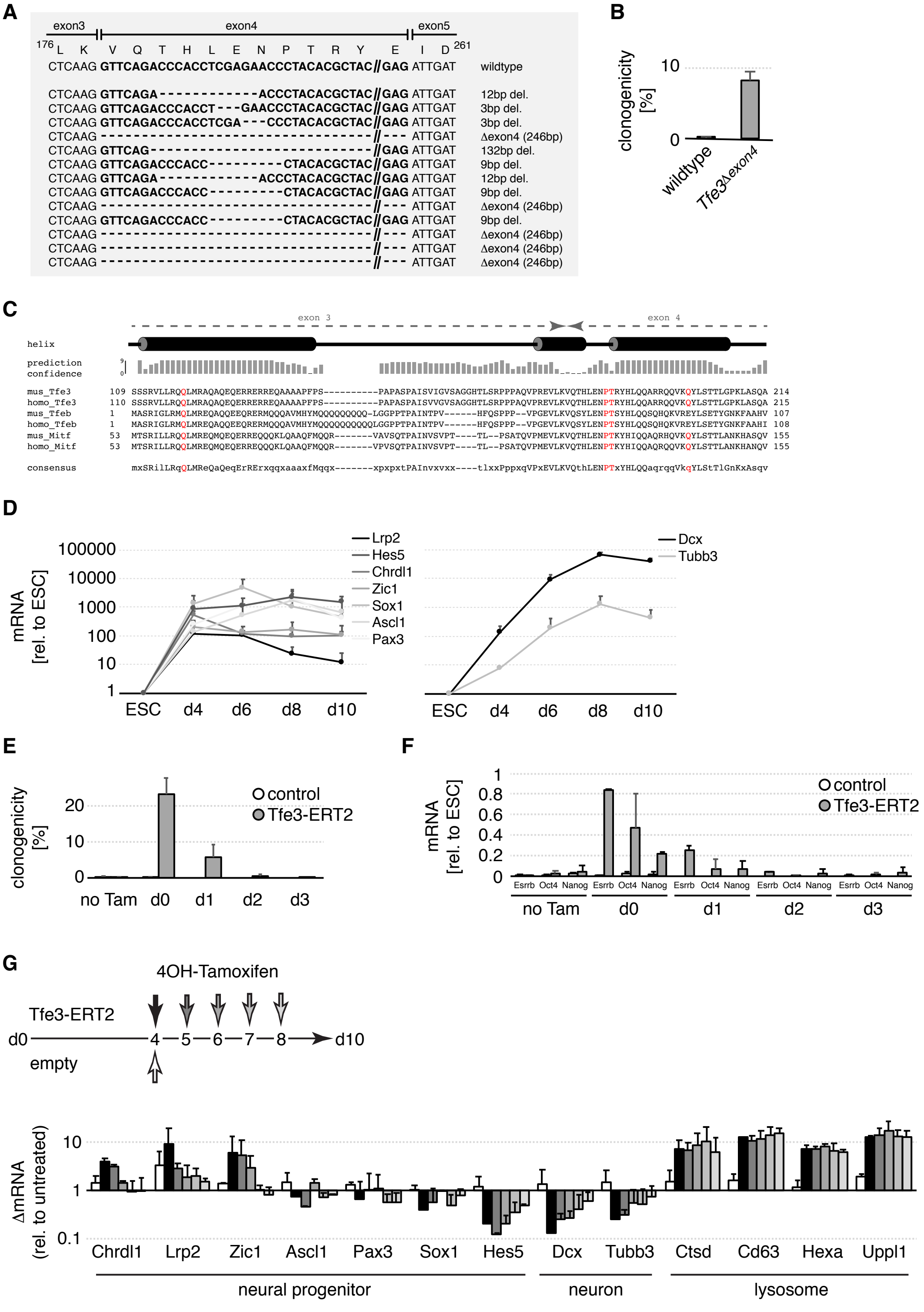


Figure S7, related to Figure 7: Tfe3 induces a neural differentiation delay.

(**A**) In frame Tfe3 mRNA deletions in 13 independent clones derived from differentiation-resistant Cas9-expressing cells transiently transfected with screen gRNAs targeting exon4 (Figure S5B). Exon boundaries and translation of wildtype *Tfe3* are indicated.

(**B**) Retention of self-renewal after 3 days of 2i withdrawal in $Tfe3^{\Delta exon4}$ cells. Average and SD of 2 technical replicates.

(**C**) Amino acid sequence alignment of the exon3-4 boundary in mouse and human MiTF/Tfe family members. I-Tasser (Yang et al., 2015) secondary structure prediction including confidence values is indicated. Patient mutation homologues are highlighted in red.

(**D**) Induction of neural progenitor (left) and neuronal (right) markers upon release from 2i. Fold changes relative to ESCs are shown on a log(10)-scaled axis. Average and SD of two biological replicates. Days (d).

(**E**,**F**) Self-renewal (E), and Essrb, Oct4 and Nanog expression (F) after 4 days of 2i withdrawal in cells of indicated genotypes treated with Tam starting from day 0, 1, 2 or 3 of differentiation. Average and SD of two biological replicates.

(**G**) Fold mRNA changes of empty vector and Tfe3-ERT2 expressing cells at day 10 of neuronal differentiation. Tam was added between day 4 and day 8. Relative changes to untreated cells of the same genotype are shown on a log(10)-scaled axis. Average and SD of two biological replicates.

Article

TRANSPARENT PROCESS

THE EMBO JOURNAL

Zfp281 orchestrates interconversion of pluripotent states by engaging Ehmt1 and Zic2

Daniela Mayer[1,2], Michael B Stadler[1,3], Melanie Rittirsch[1], Daniel Hess[1], Ilya Lukonin[1,2], Maria Winzi[4], Austin Smith[5], Frank Buchholz[4] & Joerg Betschinger[1,*]

Abstract

Developmental cell fate specification is a unidirectional process that can be reverted in response to injury or experimental reprogramming. Whether differentiation and de-differentiation trajectories intersect mechanistically is unclear. Here, we performed comparative screening in lineage-related mouse naïve embryonic stem cells (ESCs) and primed epiblast stem cells (EpiSCs), and identified the constitutively expressed zinc finger transcription factor (TF) Zfp281 as a bidirectional regulator of cell state interconversion. We showed that subtle chromatin binding changes in differentiated cells translate into activation of the histone H3 lysine 9 (H3K9) methyltransferase Ehmt1 and stabilization of the zinc finger TF Zic2 at enhancers and promoters. Genetic gain-of-function and loss-of-function experiments confirmed a critical role of Ehmt1 and Zic2 downstream of Zfp281 both in driving exit from the ESC state and in restricting reprogramming of EpiSCs. Our study reveals that cell type-invariant chromatin association of Zfp281 provides an interaction platform for remodeling the cis-regulatory network underlying cellular plasticity.

Keywords cell state transition; cellular plasticity; differentiation; pluripotency; reprogramming

Subject Categories Chromatin, Transcription & Genomics; Stem Cells & Regenerative Medicine; Transcription

DOI 10.15252/embj.2019102591 | Received 3 June 2019 | Revised 21 October 2019 | Accepted 24 October 2019

The EMBO Journal (2019) e102591

Introduction

Mammalian development is a hierarchical process that coordinates organismal growth with increasing cellular differentiation. The lineage progression of the few pluripotent cells in the blastocyst toward the many specialized cell types in the mature embryo is by and large unidirectional. However, fully differentiated cells can be de-differentiated into induced pluripotent stem cells (iPSCs) by ectopic expression of the transcription factors (TFs) Oct4, Sox2, Klf4, and c-Myc (Takahashi & Yamanaka, 2006). Reprogramming of somatic cells into iPSCs requires erasure of the entire developmental history of a somatic cell, but whether this depends on the reversal of developmental hierarchies is unclear (Ladewig *et al*, 2013; Takahashi & Yamanaka, 2015).

Transcriptional and epigenomic profiling of the reprogramming process has revealed an ordered series of events that include the transient and sequential activation of late and early developmental genes (Takahashi *et al*, 2014; Cacchiarelli *et al*, 2015; Amlani *et al*, 2018). Although the specific trajectory is dictated by the identity of the starting somatic cell type (Jackson *et al*, 2016; Nefzger *et al*, 2017) and the experimental regime (Chantzoura *et al*, 2015; Stuart *et al*, 2019), iPSC formation may involve the reversion of natural developmental mechanisms (Takahashi & Yamanaka, 2015). Consistent with this possibility, a mesenchymal-to-epithelial transition is necessary for iPSC formation (Li *et al*, 2010; Samavarchi-Tehrani *et al*, 2010), while the converse epithelial-to-mesenchymal transition is crucial for embryogenesis, e.g., during gastrulation and neural crest formation (Acloque *et al*, 2009). Although it is debated whether these observations reflect a shared developmental intermediate (Raab *et al*, 2017), they suggest that de-differentiation and differentiation employ common mechanisms in opposite directions. Here, we systematically and functionally examine this concept using naïve pluripotent embryonic stem cells (ESCs) and primed pluripotent epiblast stem cells (EpiSCs; Smith, 2017).

Embryonic stem cells and EpiSCs are developmentally related derivatives of mouse embryonic day (E) 3.75–4.5 blastocysts (Boroviak *et al*, 2014) and E5.5–8.0 embryos (Brons *et al*, 2007; Tesar *et al*, 2007; Osorno *et al*, 2012), respectively. ESCs cultured in the presence of two inhibitors (2i) resemble naïve pluripotent cells of the preimplantation epiblast (Boroviak *et al*, 2014), while primed pluripotent EpiSCs cultured in the presence of FGF2 and activin A (FA) resemble cells of the late gastrula (Osorno *et al*, 2012; Kojima *et al*, 2014; Tsakiridis *et al*, 2014). Upon *in vitro* differentiation, ESCs progress through a transient post-implantation epiblast-like (EpiLC) cell state that is amenable to EpiSC derivation (Zhang *et al*, 2010; Hayashi *et al*, 2011). Conversely, activation of just one TF,

---

1 Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland
2 Faculty of Sciences, University of Basel, Basel, Switzerland
3 Swiss Institute of Bioinformatics, Basel, Switzerland
4 Medical Systems Biology, UCC, Medical Faculty Carl Gustav Carus, TU Dresden, Dresden, Germany
5 Wellcome-MRC Cambridge Stem Cell Institute and Department of Biochemistry, University of Cambridge, Cambridge, UK
*Corresponding author. Tel: +41 79 500 9513; E-mail: joerg.betschinger@fmi.ch

such as Stat3, Klf4, or Esrrb, is sufficient to reprogram EpiSCs into naïve pluripotent EpiSC-derived iPSCs (Epi-iPSCs) in the presence of 2i (Guo *et al*, 2009; Yang *et al*, 2010; Festuccia *et al*, 2012). The interconvertibility of ESCs and EpiSCs thus provides an experimental system to explore whether de-differentiation includes the reversion of differentiation mechanisms.

Using a large-scale loss-of-function reprogramming screen in sensitized EpiSCs, we identify the zinc finger TF Zfp281 as a prominent bidirectional ESC-EpiSC transition regulator. We show that Zfp281 exhibits stable chromatin association and drives ESC progression through differentiation-specific interaction with Ehmt1 and Zic2. Genomic analysis revealed activation of Ehmt1 and enrichment of Zic2 at Zfp281-bound cis-regulatory elements (CREs) that are associated with developmental transcription in EpiLCs and EpiSCs. Zfp281 therefore establishes and stabilizes cell fate commitment to safeguard the unidirectionality of pluripotent state transitions.

Results

Zfp281 is a bidirectional ESC-EpiSC transition regulator

We hypothesized that mechanisms common to differentiation and de-differentiation may be encoded in genes that both promote exit from the naïve ESC state and impair reprogramming of EpiSCs. ESC differentiation drivers have been determined in several genetic loss-of-function screens (Guo *et al*, 2011; Westerman *et al*, 2011; Betschinger *et al*, 2013; Leeb *et al*, 2014; Li *et al*, 2018), but it is unknown whether those also inhibit reprogramming of EpiSCs into naïve pluripotency. We therefore set out to systematically identify reprogramming roadblocks using a large-scale endoribonuclease-prepared small interfering RNA (esiRNA) loss-of-function screen (Ding *et al*, 2009). We made use of O4GIP^{GY118F} EpiSCs expressing green fluorescent protein (GFP) and puromycin *N*-acetyltransferase under the regulatory sequences of the *Oct4* gene (Guo *et al*, 2009), and a Stat3-activating receptor (GY118F) responsive to granulocyte colony-stimulating factor (Gcsf) driven by a constitutive promoter (Yang *et al*, 2010). Upon exposure to Gcsf and 2i for 4 days (d), O4GIP^{GY118F} EpiSCs gave rise to self-renewing Epi-iPSCs at an efficiency of roughly 0.1% (Fig EV1A), thus providing a sensitized setup to identify reprogramming inhibitors. O4GIP^{GY118F} EpiSCs were transfected with esiRNAs targeting 9,540 transcripts and control esiRNAs targeting Luciferase (Luc) and the GY118F downstream effector Stat3 in 384-well plates (Fig 1A). The next day, reprogramming was induced by changing to 2i and Gcsf. After 4 days, we selected Epi-iPSCs in the presence of puromycin and quantified viability with a fluorescent assay after 3–4 days. The screen was performed in duplicate, and *Z* scores were calculated per plate (Table EV1). Positive (Stat3 esiRNA), but not negative (non-targeting Luc esiRNA and no esiRNA), controls induced negative *Z* scores (Fig EV1B). Screen hits with average *Z* scores < -2 included ribosome and proteasome subunits, Stat3 and Oct4 (Fig 1B), and were strongly enriched for functions associated with RNA maturation and translation using gene ontology (GO) analysis (Fig EV1C). These therefore contain genes required for reprogramming and/or cell survival. Screen hits with positive *Z* scores, conversely, are expected to inhibit reprogramming and/or proliferation. Among the 146 hits with an average *Z* score > 2, the zinc finger TF *Zfp281* and the E3 ubiquitin ligase *Fbxw7* scored highest. Zfp281 and Fbxw7 have previously been shown to restrict iPSC generation from somatic cells (Buckley *et al*, 2012; Fidalgo *et al*, 2012, 2016; Okita *et al*, 2012), thus suggesting successful identification of reprogramming roadblocks.

To determine whether any of the 146 genes also drive ESC differentiation, we compared our screen hits with those from two previous large-scale ESC differentiation loss-of-function studies (Betschinger *et al*, 2013; Li *et al*, 2018; Fig 1C). Zfp281 and the cytochrome *c* oxidase subunits Cox5a and Cox6c scored strongest in all screens. For validation, we depleted each of them by siRNA transfection in independent GY118F-expressing Oct4 reporter 796.4 EpiSCs (Yang *et al*, 2010) and also included siRNAs targeting Fbxw7 and Tcf7l1 as controls (Fig 1B and C). Knockdown of Fbxw7 and Zfp281, but not of Cox5a, Cox6c, or Tcf7l1, increased reprogramming (Fig 1D and Appendix Fig S1A). Therefore, Cox5a and Cox6c are false-positive or cell line-dependent screen hits, and we focused our further efforts on Zfp281. Consistent with previous findings (Fidalgo *et al*, 2016), Epi-iPSCs derived by Zfp281 depletion expressed the naïve TFs Esrrb, Klf4, Nr0b1, and Tbx3, and reduced levels of the primed markers Oct6, Fgf5, Sox3, and Dnmt3b (Fig EV1D), suggesting successful reversion to the pluripotent ground state.

To quantify the dynamics of this process, we used self-renewal in 2i as a proxy for acquisition of Epi-iPSC identity. Compared to controls, Zfp281 depletion dramatically increased the colony-forming capacity of single cells after 2 and 4 days of Gcsf addition (Fig 1E), indicating accelerated and more efficient reprogramming of EpiSCs. Gcsf supplementation was essential, and co-depletion of Stat3 abolished Epi-iPSC formation from O4GIP^{GY118F} EpiSCs in the presence of Gcsf (Fig EV1E). Similarly, Leukemia inhibitory factor (Lif), which activates Stat3 in EpiSCs (Yang *et al*, 2010), was required for reprogramming of Zfp281-depleted OEC2 EpiSCs (Fig 1F). To test whether Zfp281 acts only in the context of active Stat3, we used conditional expression of Esrrb or Klf4 in O4GIP EpiSCs through addition of doxycycline (Dox) to induce reprogramming. In the absence of extrinsic Lif, we observed an increase in Dox-induced Epi-iPSC colonies upon knockdown of Zfp281 (Fig 1G), suggesting that Zfp281 functions independent of the specific reprogramming regime. Taken together, these findings demonstrate that the vast majority of cell state transition regulators act unidirectionally. Zfp281, in contrast, acts bidirectionally as it drives ESC differentiation and inhibits reprogramming of EpiSCs. Notably, this is inverse to the activity of reprogramming TFs, e.g., Klf4 and Esrrb, that induce and consolidate the naïve ESC state (Guo *et al*, 2009; Niwa *et al*, 2009; Festuccia *et al*, 2012; Martello *et al*, 2012; Yamane *et al*, 2018).

Zfp281 promotes exit from naïve pluripotency independent of Tet1 and Tet2

To characterize the function of Zfp281 in ESC differentiation, we inactivated the gene in naïve RGd2 ESCs that contain a destabilized GFP protein downstream of the *Rex1* (*Zfp42*) promoter (Appendix Fig S2A and B), which allows near real-time tracking of cell state transition (Kalkan *et al*, 2017): GFP is homogeneously expressed in 2i and up to 16 hours (h) after 2i withdrawal (GFP^{high};

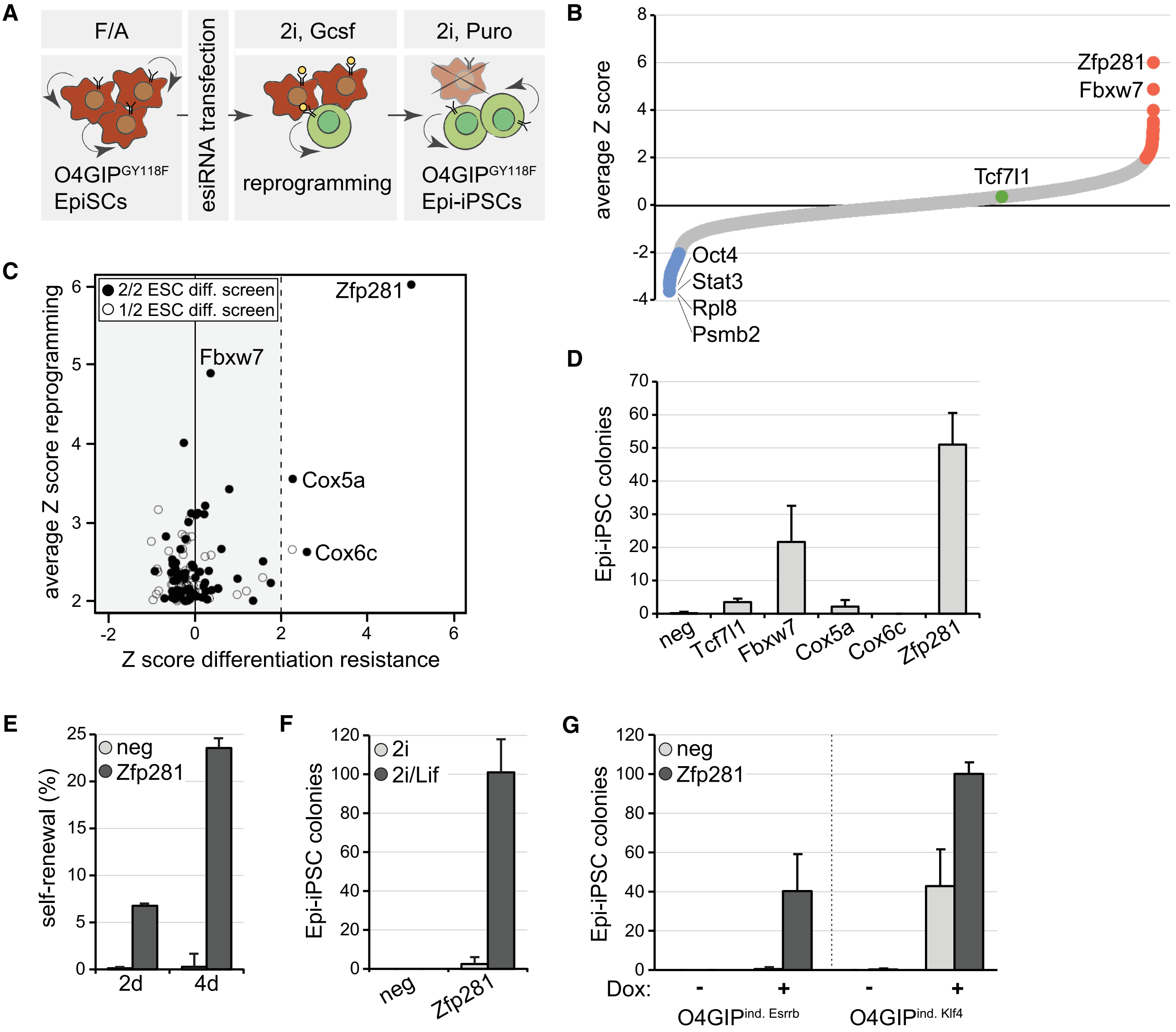


Figure 1. Zfp281 inhibits reprogramming of EpiSCs.

A Schematic outline of the reprogramming screen. Red indicates O4GiP^{GY118F} EpiSCs and green O4GIP^{GY118F} Epi-iPSCs.

B Average *Z* scores of the two screen replicates. Note that esiRNAs targeting Mll1 (Zhang *et al*, 2016a) and Mbd3 (Rais *et al*, 2013) were not included in our library and that Otx2 (Acampora *et al*, 2013) scored below the significance threshold. Screen hits with negative (blue) and positive (red) *Z* scores (red), and *Tcf7l1* (green) are highlighted.

C Comparison of reprogramming screen hits with two ESC differentiation screens (Betschinger *et al*, 2013; Li *et al*, 2018). Empty and full circles indicate genes recovered in one and both ESC differentiation screens, respectively.

D Number of Epi-iPSC colonies derived from 796.4 EpiSCs transfected with indicated siRNAs, stimulated with Gcsf and 2i for 4 days, and selected with puromycin. Average and standard deviation (SD) of three experiments performed in duplicates. Negative siRNA (neg).

E Self-renewal of O4GIP^{GY118F} reprogramming intermediates after 2 days or 4 days of stimulation with Gcsf and 2i following transfection with indicated siRNAs. Average and SD of two experiments performed in duplicates.

F Number of Epi-iPSC colonies derived from OEC2 EpiSCs transfected with indicated siRNAs, treated for 4 days in 2i or 2i/Lif medium, and selected with puromycin. Average and SD of two experiments performed in duplicates.

G Number of Epi-iPSC colonies derived from O4GIP EpiSCs carrying Dox-inducible Esrrb or Klf4 transgenes after transfection with indicated siRNAs, stimulation with or without Dox for 2 days, and selection with puromycin. Average and SD of two experiments performed in duplicates.

Kalkan *et al*, 2017) before becoming progressively downregulated (GFP^{low}) as ESCs exit from self-renewal (Fig EV2A). In 2i, reporter expression in two independent *Zfp281* knockout (KO) clones was similar to the parental wild-type cell line (*WT*) and an untargeted wild-type sibling clone (*Zfp281* WT) (Fig EV2B). In contrast, 32 and 72 h after 2i withdrawal, 30 and < 1% of *WT* cells were GFP^{high}, while 75 and 10% of *Zfp281* KO cells maintained high GFP expression, respectively. Consistent with impaired exit from the ESC state, 10% of *Zfp281* KO cells formed colonies in 2i after 72 h of differentiation (Fig 2A). This phenotype was reverted by transgenic Zfp281 expression (Fig 2B). Resistance to exit self-renewal was also observed in KO cells generated in a different ESC lines (Appendix Fig S2A and B, Fig EV2C), and in EpiLC (Hayashi *et al*, 2011) and embryoid body (EB) differentiation regimes (Figs 2A and EV2B). *Zfp281* mutant cells maintained *Rex1* reporter expression and self-renewal even after lengthy periods in the absence of 2i (Figs 2A and EV2B), demonstrating that differentiation resistance is persistent.

Differentiating *Zfp281* KO cells expressed varied levels of the *Rex1* reporter (Fig EV2B) and formed colonies in 2i less efficiently than naïve pluripotent ESCs. For example, 32 h after 2i withdrawal, mutant cells displayed only 40% of the self-renewal capacity of ESCs (Fig EV2D). To test whether this reduction is linked to population heterogeneity, we purified GFP^{high} and GFP^{low} cells at 32 h using fluorescence-activated cell sorting. As expected (Kalkan *et al*, 2017), sorted *WT* GFP^{low} cells were largely committed to differentiation and unable to generate clones in 2i (Fig 2C). In contrast, *Zfp281* KO GFP^{low} cells formed colonies almost as efficiently as GFP^{high} cells. Rex1 downregulation and exit from the ESC state is, thus, disconnected in *Zfp281* mutants. However, the efficiency with which GFP^{high} cells formed colonies after 32 h of 2i withdrawal was lower than of GFP^{high} cells after 24 h (Kalkan *et al*, 2017) and of ESCs (Fig 2C). This was irrespective of genotype, suggesting a gradual decline in self-renewal during differentiation both in *Zfp281* mutant and in *WT* GFP^{high} cells. The reduced clonogenicity of *Zfp281* KO populations compared to ESCs may therefore be consequential to impaired progression of an advanced cell state with limited self-renewal capacity and independent of population heterogeneity. In fact, GFP^{low} cells in long-term differentiated *Zfp281* mutants re-established GFP^{high} expression within a few days (Fig EV2E), revealing reversibility of the GFP^{low} state in the absence of *Zfp281*. To test sufficiency, we generated naïve RGd2 cells conditionally overexpressing Zfp281 under Dox regulation (Fig EV2F). Dox treatment in the presence of 2i induced silencing of the *Rex1* reporter and loss of self-renewal in a subset of cells (Figs 2D and EV2G). Zfp281 is therefore required and sufficient for exit from naïve pluripotency.

A previous study showed that differentiation of Serum/Lif-cultured ESCs is accompanied by upregulation of Zfp281, which in turn destabilizes metastable pluripotency by binding to the methylcytosine hydroxylase Tet1 and transcriptionally suppressing Tet2 (Fidalgo *et al*, 2016). If the same mechanisms were to regulate exit from naïve pluripotency downstream of Zfp281, loss of *Tet1* should induce the same phenotype as loss of *Zfp281*, and loss of *Tet2*, the opposite. We therefore generated *Tet1*, *Tet2*, and *Zfp281*/*Tet2* KO alleles in naïve RGd2 ESCs (Appendix Fig S2C and D). Strikingly, the extinction of the *Rex1* reporter and self-renewal was similar in differentiating *WT* cells and *Tet1* and *Tet2* mutants, while the absence of *Tet2* in *Zfp281*/*Tet2* KO cells did not revert resistance to differentiation caused by absence of *Zfp281* alone (Figs 2E and EV2H). We furthermore noted only modest changes in Zfp281 mRNA or protein during ESC differentiation, and across existing RNA sequencing (RNA-seq) datasets of EpiLC and EpiSC differentiation (Buecker *et al*, 2014; Factor *et al*, 2014; Bao *et al*, 2018) and epiblast development (Boroviak *et al*, 2015; Figs 2F and EV2I). Zfp281 has also been reported to repress *Nanog* transcription through interacting with the NuRD complex in Serum/Lif-cultured ESCs (Fidalgo *et al*, 2012). However, Nanog mRNA was unchanged in naïve *Zfp281* KO ESCs or EpiSCs depleted of Zfp281 by siRNA transfection (Fig EV2J and K). Furthermore, knockdown of Nanog did not restore differentiation in ESCs depleted of Zfp281 while it partially did so, as expected (Pereira *et al*, 2006), in ESCs depleted of Tcf7l1 (Fig EV2L). Taken together, these results suggest that Zfp281 is expressed fairly constantly during exit from naïve pluripotency and drives differentiation independent of Tet1, Tet2, and Nanog.

Zfp281 acts independent of cell state-exclusive chromatin association

To identify the transcriptional defects causing differentiation resistance, we performed RNA-seq of *WT* and *Zfp281* KO cells in 2i, and 16 and 32 h after 2i withdrawal ($WT^{2i,16\ h,32\ h}$ and $Zfp281^{2i,16\ h,32\ h}$; Table EV2). The expression of several naïve and primed pluripotency markers was perturbed in $Zfp281^{16\ h}$ and $Zfp281^{32\ h}$ cells (Fig EV3A), confirming impaired silencing of naïve identity in *Zfp281* mutants. *k*-means clustering of mRNAs that significantly changed during *WT* differentiation or in *Zfp281* KO cells (2,495 genes) identified six gene clusters (Fig 3A and B): Clusters 1–4 (1,898 genes) contain the majority of genes that were differentially transcribed in $WT^{32\ h}$ cells and of which a subset was already regulated in $WT^{16\ h}$ cells. Comparison with external EpiLC (Buecker *et al*, 2014) and EpiSC (Factor *et al*, 2014; Bao *et al*, 2018) expression datasets, which were not employed in the clustering analysis, revealed persistence of the bulk transcriptional changes established in $WT^{32\ h}$ cells, suggesting that clusters 1–4 contribute to pluripotent cell state progression. Clusters 5 and 6 (597 genes), in contrast, contain genes that were mostly unchanged in $WT^{32\ h}$ cells, but transiently regulated in $WT^{16\ h}$ cells and differentially expressed in EpiSCs. Clusters 5 and 6 may therefore act in gastrulation stage epiblast development and/or EpiSCs.

Clusters 1 and 2 were largely unaffected in differentiating *Zfp281* KO cells, whereas the repression and induction, respectively, of cluster 3 and 4 genes were blunted in both $Zfp281^{16\ h}$ and $Zfp281^{32\ h}$ cells (Fig 3A and B). Cell state-specific comparison revealed that this was predominantly due to deregulation during differentiation (Fig 3C). Although we cannot exclude that the transcriptional defects in $Zfp281^{32\ h}$ cells were influenced by cell state heterogeneity (Fig EV2D), perturbed expression of cluster 3 and 4 genes in $Zfp281^{16\ h}$ cells, a time point at which downregulation of Rex1 reporter expression (Kalkan *et al*, 2017) and exit from self-renewal (Fig EV2A and D) has not yet commenced, suggests a direct role of Zfp281 in regulating these genes. Zfp281 may therefore drive exit from naïve pluripotency through controlling gene clusters 3 and 4, which contain the naïve pluripotency TFs *Klf4*, *Klf5*, and *Nr0b1*, and the primed markers *Sox3* and *Dnmt3b*, respectively (Table EV2), and are enriched for generic developmental terms using GO analysis (Fig EV3B). Conversely, cluster 5 and 6 genes were similarly mis-expressed in $Zfp281^{2i}$, $Zfp281^{16\ h}$, and $Zfp281^{32\ h}$ cells, notably with directionalities that are inverse to the changes observed in EpiSCs. GO analysis revealed significant enrichment of regulators of cell adhesion, which is critical for cell

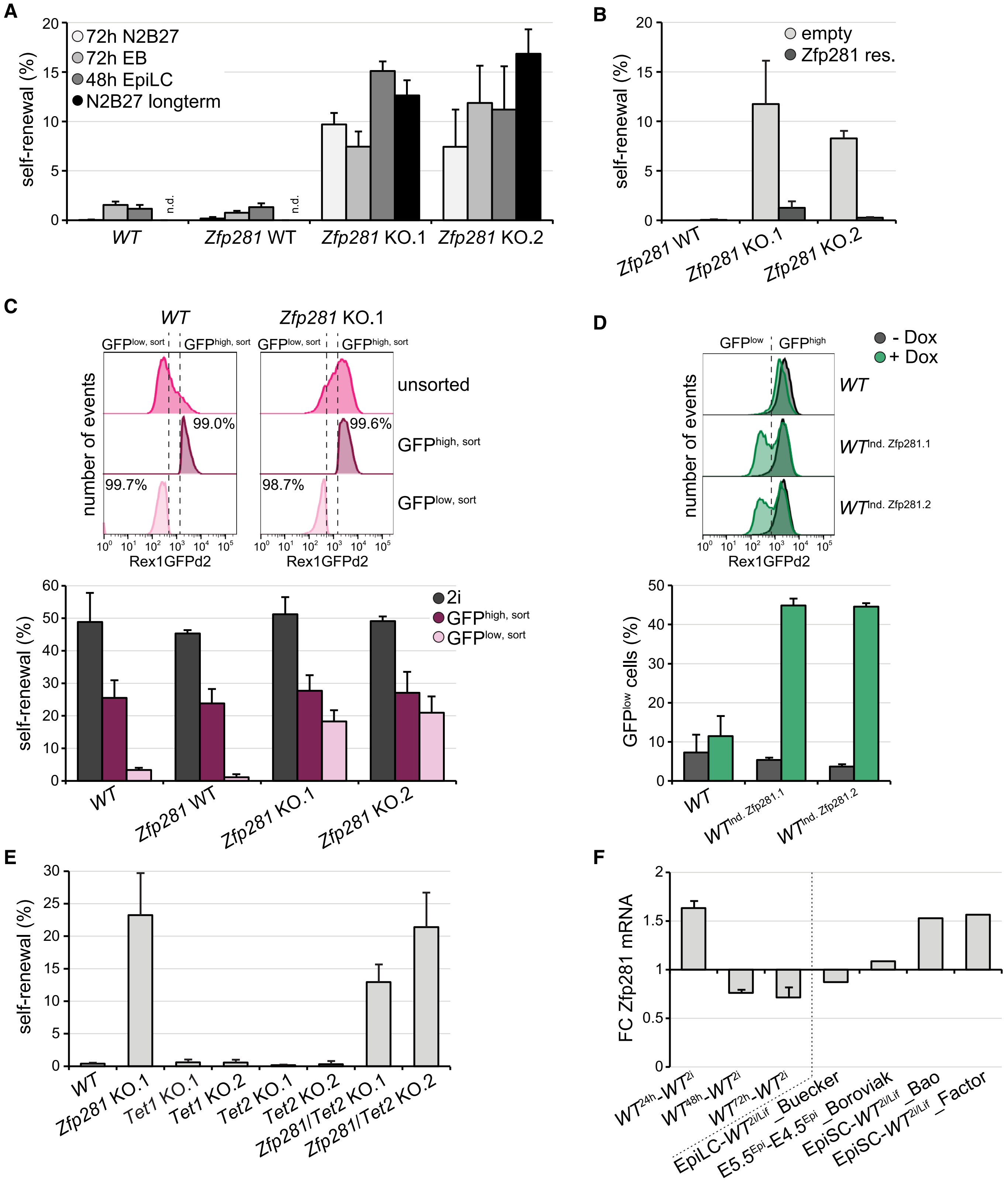

A
self-renewal (%)
20
15
10
5
0
○ 72h N2B27
○ 72h EB
● 48h EpiLC
● N2B27 longterm
n.d.
n.d.
WT
Zfp281 WT
Zfp281 KO.1
Zfp281 KO.2
B
self-renewal (%)
○ empty
● Zfp281 res.
Zfp281 WT
Zfp281 KO.1
Zfp281 KO.2
C
WT
Zfp281 KO.1
GFPlow, sort
GFPhigh, sort
number of events
unsorted
GFPhigh, sort
GFPlow, sort
99.0%
99.6%
99.7%
98.7%
Rex1GFPd2
self-renewal (%)
60
50
40
30
20
10
0
● 2i
● GFPhigh, sort
○ GFPlow, sort
WT
Zfp281 WT
Zfp281 KO.1
Zfp281 KO.2
D
GFPlow
GFPhigh
● - Dox
● + Dox
number of events
WT
WTInd. Zfp281.1
WTInd. Zfp281.2
Rex1GFPd2
GFPlow cells (%)
50
40
30
20
10
0
E
self-renewal (%)
30
25
20
15
10
5
0
WT
Zfp281 KO.1
Tet1 KO.1
Tet1 KO.2
Tet2 KO.1
Tet2 KO.2
Zfp281/Tet2 KO.1
Zfp281/Tet2 KO.2
F
FC Zfp281 mRNA
2
1.5
1
0.5
0
WT24h-WT2i
WT48h-WT2i
WT72h-WT2i
EpiLC-WT2i/Lif_Buecker
E5.5Epi-E4.5Epi_Boroviak
EpiSC-WT2i/Lif_Bao
EpiSC-WT2i/Lif_Factor

Figure 2.

Figure 2. Zfp281 drives exit from naïve pluripotency independent of Tet enzymes.

A, B Self-renewal in RGd2 ESCs of specified genotypes expressing indicated transgenes (B) after differentiation in indicated conditions (A) or 72 h in N2B27 (B). Average and SD of two experiments performed in duplicates. Note that control cells were lost during continuous passaging in N2B27 (A). Not determined (n.d.).

C Representative flow cytometry profiles of *WT* and *Zfp281* KO.1 cells after 32 h of 2i withdrawal before (unsorted) and after purification of cells with indicated GFP expression (top panel). Self-renewal of undifferentiated (2i) and sorted GFP^{high,sort} and GFP^{low,sort} cells of indicated genotypes (bottom panel). Average and SD of three experiments performed in duplicates.

D Representative flow cytometry profiles of control and Zfp281-inducible ESCs (top panel) and quantification of GFP^{low} cells (bottom panel) after 2 days in 2i and in the presence (green) or absence (black) of Dox. Average and SD of two experiments.

E Self-renewal in RGd2 ESCs of specified genotypes after 72 h in N2B27. Average and SD of two experiments performed in duplicates.

F Zfp281 mRNA changes during ESC differentiation detected by quantitative PCR (left) and extracted from published RNA-seq datasets (Buecker *et al*, 2014; Factor *et al*, 2014; Boroviak *et al*, 2015; Bao *et al*, 2018) (right). Average and SD of two technical replicates (left).

polarization (Ebnet *et al*, 2018) that initiates lumenogenesis after exit from naïve pluripotency (Shahbazi *et al*, 2017). To test whether Zfp281 controls polarization, we generated spheroids in Matrigel as described before (Shahbazi *et al*, 2017). *WT* ESCs formed polarized spheroids with expanded lumens that were encircled by apical F-actin, while *Zfp281* KO cells grew as unpolarized and disorganized cellular aggregates that were morphologically similar to ESCs (Fig 3D). Although we cannot exclude that this is consequential to impaired exit from self-renewal, regulation of cluster 5 and 6 genes by Zfp281 may therefore contribute to cell polarization and cavity formation during ESC differentiation.

Oct4, similar to Zfp281, is expressed at equal levels in ESCs and EpiLCs, but occupies distinct CREs in the two cell states (Buecker *et al*, 2014). To determine whether Zfp281 acts through cell state-specific chromatin association, we profiled its genome localization in WT^{2i} and $WT^{32\ h}$ cells using chromatin immunoprecipitation (ChIP) coupled to deep sequencing (ChIP-seq) (Table EV2). *De novo* motif finding identified the consensus CCCCTCCCCC motif in 82.4% of 23,756 peaks (Fig EV3C), which is similar to results obtained in Serum/Lif ESCs (Fidalgo *et al*, 2016). Surprisingly, Zfp281 occupancy in WT^{2i} and $WT^{32\ h}$ cells was as highly correlated (Pearson's correlation coefficient $R = 0.81$) as between replicates ($R = 0.84$ and $R = 0.85$, respectively), with only few peaks exclusively detected in any of the two cell states (Figs 3E and EV3D). A lower correlation was observed between $WT^{32\ h}$ cells and published data for EpiSCs (Huang *et al*, 2017; $R = 0.69$) and trophoblast stem cells (TSCs; Ishiuchi *et al*, 2019; $R = 0.55$), but binding at peaks associated with cluster 1–6 genes was largely unchanged (Fig EV3E and F), suggesting stable chromatin association also during later pluripotency progression and in lineage-unrelated TSCs. To determine whether Zfp281 binds to CREs, we profiled histone H3 lysine 27 acetylation (H3K27ac), a chromatin mark associated with active promoters and enhancers. Comparison of our Zfp281 and H3K27ac with published histone mark ChIP-seq data (Kurimoto *et al*, 2015; Buecker *et al*, 2014; Fig EV3G) identified 7,697 Zfp281 peaks proximal to transcriptional start sites (TSSs), of which 54% were at active promoters (co-localization with H3K27ac and H3K4 tri-methylation), and 16,059 distal Zfp281 peaks of which 62% were at putative enhancers (co-localization with H3K4 mono-methylation). 38% of the latter were also enriched for H3K27ac, qualifying them as active enhancers. Notably, peaks with slightly increased Zfp281 binding in $WT^{32\ h}$ cells gained H3K27ac and expression of associated genes during differentiation, while decreased binding was associated with reduced H3K27ac and transcription (Fig 3F). Despite stable occupancy of target sites, quantitative binding changes in Zfp281 at these sites therefore parallel differences in CRE activity. However, this was similar at peaks linked to Zfp281-insensitive cluster 1/2 and Zfp281-sensitive cluster 3/4 genes (Fig 3G), showing that differential binding strength at CREs correlates with differentiation-specific gene expression, but only partially with transcriptionally regulated targets.

Zfp281 interacts with Ehmt1 and Zic2 during ESC differentiation

Since chromatin occupancy was largely unchanged, we reasoned that Zfp281 may control transcription through cell state-specific protein interaction partners. To test this, we performed Zfp281 immunoprecipitations (IPs) coupled to semi-quantitative mass spectrometry (MS) in nuclear extracts of WT^{2i} and $WT^{40\ h}$ cells, including *Zfp281* KO lysates to control for antibody specificity (Table EV3). Stringent selection criteria identified the previously reported interactor Nanog (Fidalgo *et al*, 2012) in WT^{2i} cells and several proteins specifically enriched in $WT^{40\ h}$ cells (Fig 4A). Strikingly, the latter were transcriptionally induced and the former repressed during differentiation (Fig 4B), suggesting that differential binding to Zfp281 may, at least in part, be driven by protein abundance. To determine functional downstream effectors, we decided to probe genetic interaction of differentiation-specific interactors with Zfp281 in our conditionally overexpressing ESCs (Fig 2D). To this end, we depleted selected binding partners using siRNA transfection, induced Zfp281 by Dox treatment, and quantified *Rex1* reporter distribution after 32 h in 2i (Fig 4C). As controls, we included siRNAs targeting Zfp281 itself and Tcf7l1. Transfection of Zfp281 but not Tcf7l1 siRNAs blocked emergence of GFP^{low} cells (76% reduction) (Figs 4C and EV4A), thus confirming suitability for identifying genetic Zfp281 interactors. Of all candidates tested individually, only knockdown of Ehmt1 and Zic2 reduced the fraction of GFP^{low} cells (34 and 32%, respectively), an effect enhanced by simultaneous depletion of both (63%). Conversely, conditional overexpression of Ehmt1 and Zic2 in RGd2 ESCs (Fig EV4B and C), similar to Zfp281, induced downregulation of the *Rex1* reporter in a subset of cells (Fig EV4D). The *de novo* DNA methyltransferases Dnmt3a and Dnmt3b have overlapping functions during embryogenesis (Okano *et al*, 1999) and, hence, may act redundantly. However, simultaneous depletion of Dnmt3a and Dnmt3b by siRNAs or in compound *Dnmt3a/3b* KO cells (Appendix Fig S2E and Fig EV4E) did not impair Zfp281-induced reporter repression (Figs 4C, and EV4A and F), demonstrating that Zfp281 drives differentiation independent of Dnmt3a and Dnmt3b.

Zic2 is a zinc finger TF that represses poised developmental enhancers in Serum/Lif ESCs (Luo *et al*, 2015). Ehmt1 (GLP) is a methyltransferase that can be found in a complex with Ehmt2

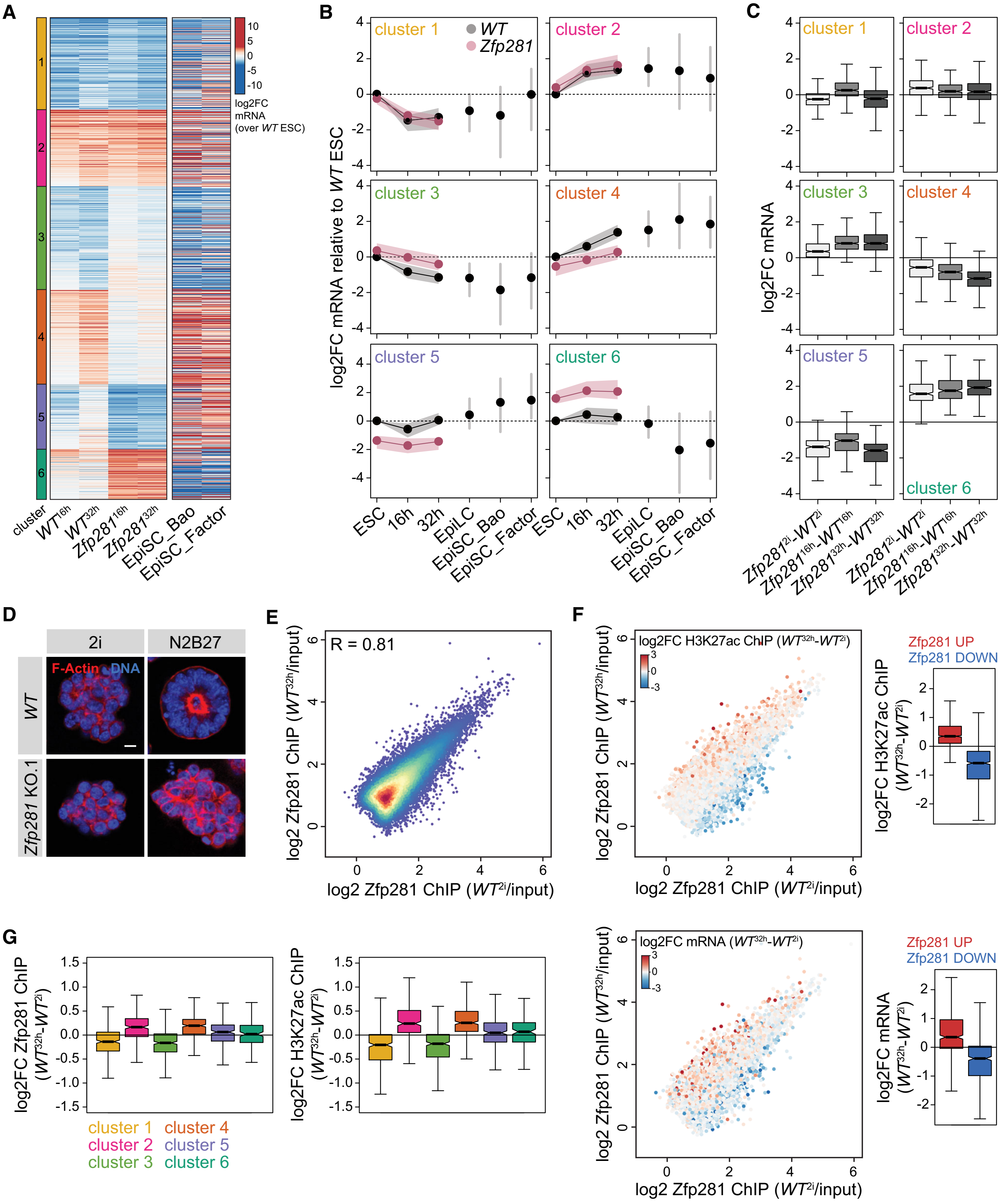

A
1
2
3
4
5
6
10
5
0
-5
-10
log2FC mRNA (over WT ESC)
cluster
WT16h
WT32h
Zfp28116h
Zfp28132h
EpiSC_Bao
EpiSC_Factor
B
cluster 1
WT
Zfp281
cluster 2
cluster 3
cluster 4
cluster 5
cluster 6
log2FC mRNA relative to WT ESC
ESC
16h
32h
EpiLC
EpiSC_Bao
EpiSC_Factor
C
log2FC mRNA
Zfp2812i-WT2i
Zfp28116h-WT16h
Zfp28132h-WT32h
D
2i
N2B27
F-Actin
DNA
WT
Zfp281 KO.1
E
R = 0.81
log2 Zfp281 ChIP (WT32h/input)
log2 Zfp281 ChIP (WT2i/input)
F
log2FC H3K27ac ChIP (WT32h-WT2i)
3
0
-3
Zfp281 UP
Zfp281 DOWN
log2FC H3K27ac ChIP (WT32h-WT2i)
G
log2FC Zfp281 ChIP (WT32h-WT2i)
log2FC H3K27ac ChIP (WT32h-WT2i)
cluster 1
cluster 2
cluster 3
cluster 4
cluster 5
cluster 6
log2FC mRNA (WT32h-WT2i)
log2FC mRNA (WT32h-WT2i)
Zfp281 UP
Zfp281 DOWN

Figure 3.

Figure 3. Zfp281 directs sequential gene expression despite stable occupancy of target sites.

A mRNA $\log_2$ fold changes ($\log_2$FC) in *WT*$^{16\ \mathrm{h}}$, *WT*$^{32\ \mathrm{h}}$, *Zfp281*$^{16\ \mathrm{h}}$, and *Zfp281*$^{32\ \mathrm{h}}$ samples relative to *WT*$^{2\mathrm{i}}$ cells, and in EpiSCs relative to *WT*$^{2\mathrm{i/Lif}}$ cells (Factor *et al*, 2014; Bao *et al*, 2018). *Zfp281*$^{2\mathrm{i}}$, *Zfp281*$^{16\ \mathrm{h}}$, and *Zfp281*$^{32\ \mathrm{h}}$ and *WT*$^{16\ \mathrm{h}}$ and *WT*$^{32\ \mathrm{h}}$ samples were used for *k*-means clustering.

B, C Quantification of (A) including mRNA $\log_2$FC in EpiLCs relative to *WT*$^{2\mathrm{i/Lif}}$ (Buecker *et al*, 2014) and as indicated (C). (B) Dots represent the median, and shades, the lower and upper quartile. (C) Middle line represents median; notches, 95% confidence interval of the median; box, interquartile range; and whiskers, 1.5 times the interquartile range. Number of datapoints: 477 (cluster 1), 397 (cluster 2), 534 (cluster 3), 490 (cluster 4), 337 (cluster 5), and 260 (cluster 6).

D Representative immunofluorescence staining of spheroids in Matrigel derived from *WT* or *Zfp281* KO.1 ESCs grown in 2i or N2B27 for 3 days. Blue: DNA. Red: F-actin. Scale bar is 10 μm.

E Scatter plot comparing Zfp281 $\log_2$ ChIP enrichment relative to matched inputs in *WT*$^{2\mathrm{i}}$ and *WT*$^{32\ \mathrm{h}}$ cells.

F Same as in (E) with dots colored according to H3K27ac ChIP $\log_2$FC at the same peaks (top left), and to gene expression $\log_2$FC associated with peaks by nearest distance to TSS (bottom left) in *WT*$^{32\ \mathrm{h}}$ relative to *WT*$^{2\mathrm{i}}$ cells. Quantification of H3K27ac ChIP (top right) and mRNA (bottom right) $\log_2$FC at top 1,000 Zfp281 peaks with increased (red) or decreased (blue) Zfp281 binding during ESC differentiation. Boxes as in Fig 3C for 1,000 datapoints each.

G Quantification of Zfp281 (left) and H3K27ac (right) ChIP $\log_2$FC in *WT*$^{32\ \mathrm{h}}$ compared to *WT*$^{2\mathrm{i}}$ cells at Zfp281 peaks assigned to gene clusters 1–6. Boxes and number of datapoints as in Fig 3C.

(G9a), which both mediate mono- and di-methylation of histone H3 lysine 9 (H3K9me1 and H3K9me2; Tachibana *et al*, 2005). The genetic interactions in naïve ESCs (Fig 4C) together with the preferential binding during differentiation observed in both nuclear extracts using IP-MS (Fig 4A) and in whole-cell lysates using IP–Western blot (Fig EV4G) suggest that Ehmt1 and Zic2 are functional downstream effectors of Zfp281.

Ehmt1 and Zic2 regulate ESC differentiation and reprogramming of EpiSCs

We therefore generated individual and compound *Ehmt1* and *Zic2* KO RGd2 ESCs (Appendix Fig S2F and G). In contrast to wild-type or *Zic2* KO cells, *Ehmt1* and *Ehmt1*/*Zic2* KO cells were spindle-shaped (Fig EV4H) and proliferated slowly (Fig EV4I). They were not arrested at a specific cell cycle stage (Fig EV4J) and did not exhibit downregulation of the *Rex1* reporter in 2i (Fig EV4K). After 2i withdrawal for 32 h (or 72 h), 75% (9%) of *Ehmt1* and 55% (7%) of *Zic2* KO cells maintained GFP expression, increasing to 90% (35%) in *Ehmt1*/*Zic2* compound KO cells (Fig EV4K), while 30% (1%) of cells from untargeted sibling clones (*Zic2* WT and *Ehmt1* WT) were GFP$^{\mathrm{high}}$. Correspondingly, 5, 4, and 12% of *Zic2*, *Ehmt1*, and *Ehmt1*/*Zic2* KO cells retained self-renewal after 72 h of differentiation (Fig 5A). *Ehmt1* and *Ehmt1*/*Zic2*, but not *Zic2* KO ESCs, were unable to form polarized spheroids in Matrigel (Fig 5B). Quantification of this effect was similar to *Zfp281* mutants (Fig EV4L). We therefore conclude that *Ehmt1* is required for polarization and that *Ehmt1* and *Zic2* promote exit from self-renewal independently of each other.

In the absence of *Ehmt1*, H3K9me2 was limited to DAPI-rich speckles (Fig EV4M), which is reminiscent of the depletion of euchromatic H3K9me2 and its enrichment at pericentric heterochromatin in *Ehmt2* mutants (Tachibana *et al*, 2002). Since exposure to the Ehmt inhibitors A-366 and UNC0642 induced dose-dependent cell lethality (data not shown), we decided to test Ehmt1's enzymatic activity by expressing specific loss-of-function alleles in *Ehmt1* KO cells: An Ehmt1 protein with mutations in the ankyrin domain (Ehmt1$^{\mathrm{ank}}$), responsible for binding to methylated H3K9 *in vitro* (Collins *et al*, 2008), reverted nuclear H3K9me2 distribution (Fig EV4M) and resistance to *Rex1* downregulation (Figs 5C and EV4N) to a similar extent as the wild-type protein did. In contrast, substitutions in the SET domain (Ehmt1$^{\mathrm{NH\text{-}LE}}$) that perturb Ehmt1 methyltransferase *in vitro* (Tachibana *et al*, 2008) rescued only partially and a small deletion in the SET domain (Ehmt1$^{\Delta\mathrm{NHHC}}$) that additionally ablates binding to Ehmt2 completely abolished rescue. Therefore, both catalytic activity of Ehmt1 and formation of a larger methyltransferase protein complex are implicated in ESC transition.

Chemical inhibition and knockdown of Ehmt enzymes in somatic cells enhances reprogramming (Shi *et al*, 2008; Sridharan *et al*, 2013; Rodriguez-Madoz *et al*, 2017). Consistently, depletion of Ehmt1 but also of Zic2 in 796.4 and O4GIP$^{\mathrm{GY118F}}$ EpiSCs increased Epi-iPSC formation in the presence of Gcsf (Fig 5D and Appendix Fig S1B). The effect was modest, but enhanced by the combined knockdown of both. Taken together, these results suggest that Ehmt1 and Zic2, similar to Zfp281, drive exit from naïve pluripotency and restrain reprogramming of EpiSCs. Notably, phenotypes induced by co-depletion of Zic2 and Ehmt1 were weaker than elimination of Zfp281, suggesting the existence of additional Zfp281 effectors that may include other histone-modifying complexes (Huang *et al*, 2017; Zhou *et al*, 2017; Ishiuchi *et al*, 2019).

Overlapping transcriptional functions of Zfp281 and Ehmt1/Zic2

To test whether the biochemical and functional interaction with Zfp281 is reflected in similar transcriptional outputs, we profiled mRNA expression in *Ehmt1*, *Zic2,* and *Ehmt1*/*Zic2* KO cells in 2i and after 32 h of differentiation (Table EV2). Principal component (PC) analysis, including *Zfp281* KO, EpiLC (Buecker *et al*, 2014), and EpiSC (Bao *et al*, 2018) datasets, of changes relative to *WT* ESCs identified PC1 to discriminate developmental timing and to separate differentiated cells from ESCs (Fig 6A). Mutant and wild-type ESCs projected similarly onto PC1 and expressed pluripotency marker genes at similar levels (Fig EV5A), confirming their naïve identity. PC2, in contrast, segregated *WT* from *Zfp281* and, in particular, *Ehmt1* KO genotypes. We, indeed, observed 1,274 deregulated genes in *Ehmt1*$^{2\mathrm{i}}$ cells that were unchanged in *Zfp281*$^{2\mathrm{i}}$ cells (Fig EV5B). These were enriched for homeostatic and cell adhesion GO terms (Fig EV5B) and likely contribute to the cellular and polarization phenotypes in *Ehmt1* KO cells.

Progression of *Zfp281*$^{16\ \mathrm{h}/32\ \mathrm{h}}$, *Ehmt1*$^{32\ \mathrm{h}}$, *and Ehmt1*/*Zic2*$^{32\ \mathrm{h}}$ cells along PC1 was impaired when compared to matching *WT* controls (Fig 6A), which we also observed in PC analysis using blastocyst development datasets (Boroviak *et al*, 2015; Fig 6B). In fact, alterations in *Ehmt1* and *Zfp281* KO cells correlated during differentiation (Fig EV5C), suggesting similar defects in developmental

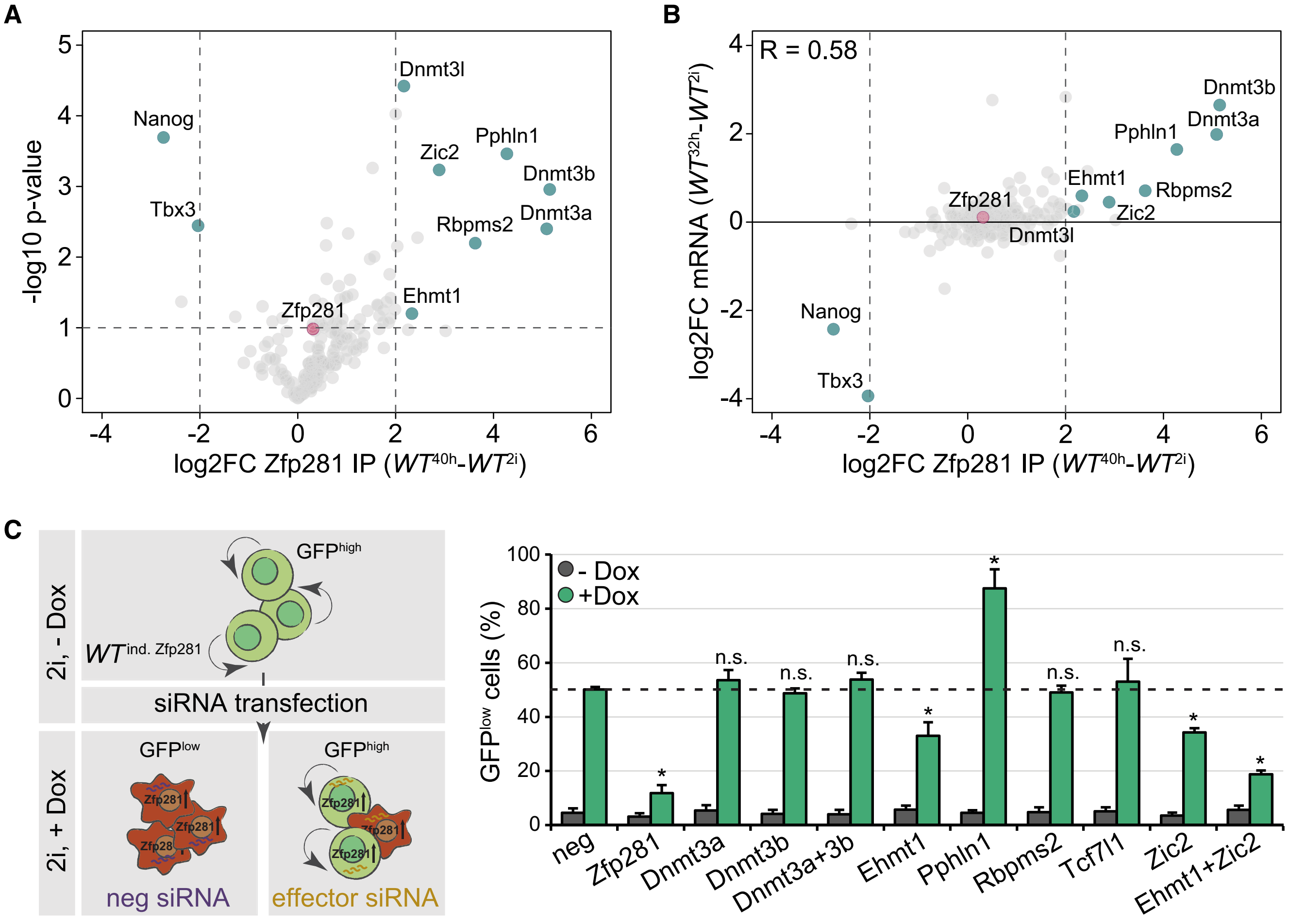


Figure 4. Zfp281 acts by associating with Ehmt1 and Zic2.

A Cell state-specific Zfp281 interactors in WT^{2i} and $WT^{40\ h}$ cells. Pink and cyan mark Zfp281 and selected binding partners, respectively. Quantification is based on three biological replicates.

B Same as (A) with mRNA log$_2$FC of differential binding partners during ESC differentiation instead of *P*-values.

C Procedure to identify Zfp281 effectors in naïve Zfp281-inducible RGd2 cells (left). Quantification of GFP^{low} $WT^{ind.\ Zfp281}$ cells transfected with indicated siRNAs and incubated for 32 h in 2i in the presence (green) or absence (black) of Dox (right). Dashed line marks fraction of GFP^{low} cells in control cells exposed to Dox. Significance was determined using a Wilcoxon–Mann–Whitney rank sum test compared to neg control sample. *$P < 0.05$; not significant (n.s.). Average and SD of four experiments.

transcription. This correlation was not strong ($R = 0.44$), but increased ($R = 0.57$) when only considering cluster 1–6 gene expression (Fig EV5D). Although we also noted slight deregulation of clusters 1 and 2, transcriptional defects in clusters 3–6 were similar in *Ehmt1*$^{32\ h}$ and *Zfp281*$^{32\ h}$ cells (Fig 6C and D). Ehmt2-dependent H3K9 di-methylation is associated with gene silencing (Zylicz *et al*, 2015), consistent with the majority of genes showing increased RNA levels in *Ehmt1*^{2i} cells (Fig EV5B). Nevertheless, 38% of target genes were downregulated. Changes in the absence of *Ehmt1* are therefore likely consequential to both direct and indirect effects and may also include the contribution of non-histone Ehmt1 substrates (Sim *et al*, 2017) to transcription.

Based on mRNA levels, *Zic2*$^{32\ h}$ cells were not separated from matching control cells (Fig 6A–D). This was surprising, since *Zic2* KO cells appeared similarly impaired in exiting self-renewal as *Ehmt1* KO cells (Fig 5A). However, loss of *Zic2* in *Ehmt1*$^{32\ h}$ cells enhanced the deregulation of clusters 1–4 during differentiation (Fig 6D) and induced a shift along PC1 (Fig 6A and B). Linear regression revealed that perturbations in *Ehmt1/Zic2* KO cells were predominantly the sum of alterations in single mutants rather than synergistic (Fig EV5E), implying subtle, but functionally relevant, changes in *Zic2* KO cells. Ehmt1 and Zic2 therefore regulate transcription independently of each other, aligning with their additive loss-of-function phenotypes in differentiation and reprogramming (Fig 5A and D). Taken together, this analysis demonstrates connected functions of Zfp281 and Ehmt1/Zic2 in gene expression during cell state transition.

Ehmt1 and Zic2 act downstream of Zfp281 on chromatin

To identify direct targets and to explore how those relate to the physical interaction with Zfp281 in differentiated cells, we

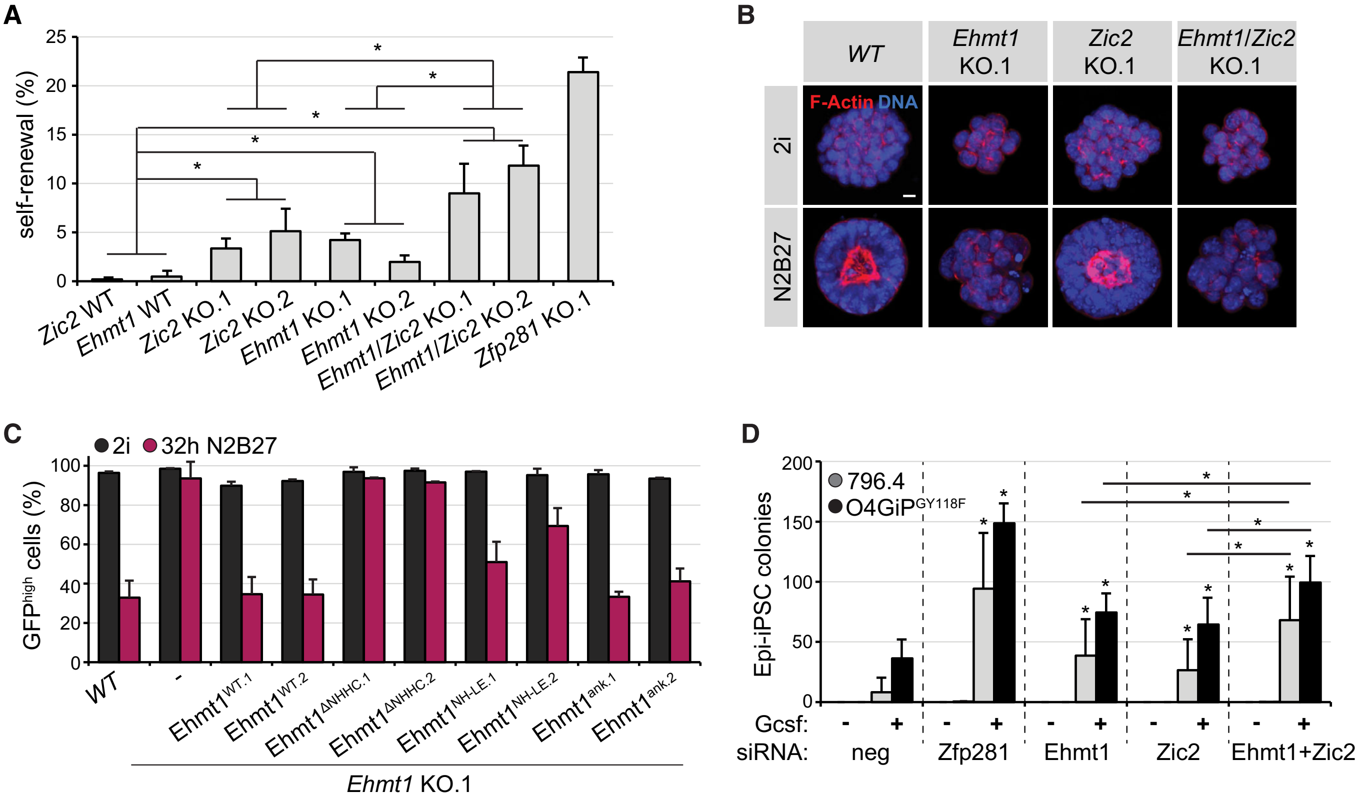


Figure 5. Ehmt1 and Zic2 drive exit from the ESC state and restrict reprogramming of EpiSCs.

A Self-renewal in cells with indicated genotypes 3 days after 2i withdrawal. Significance was determined using a Wilcoxon–Mann–Whitney rank sum test comparing the specified genotype groups. $*P < 0.05$. Average and SD of two experiments performed in duplicates.

B Representative immunofluorescence staining of spheroids in Matrigel derived from indicated genotypes in 2i or N2B27 for 4 days. Blue: DNA. Red: F-actin. Scale bar is 10 μm.

C Quantification of GFP^{high} cells in *WT* cells or *Ehmt1* KO clones expressing indicated transgenes in 2i (black) or 32 h after 2i withdrawal (pink). Average and SD of two experiments.

D Number of Epi-iPSC colonies derived from 796.4 (gray) and O4GIP^{GY118F} (black) EpiSCs transfected with indicated siRNAs, stimulated with Gcsf and 2i for 4 days, and selected with puromycin. Significance was determined using a Wilcoxon–Mann–Whitney rank sum test compared to neg control sample of the respective cell line, or comparing Zic2 and Ehmt1 to Ehmt1/Zic2 depletion. $*P < 0.05$. Average and SD of five experiments performed in duplicates.

performed Zic2 and Ehmt1 ChIP-seq, and profiled H3K9me2 as a proxy for Ehmt1 activity. Due to absence of ChIP-seq compatible Ehmt1 antibodies, we inserted an N-terminal Flag-Avi tag at both *Ehmt1* alleles in ESCs expressing the BirA biotin ligase (Fig EV5F). This did not perturb exit from self-renewal (Fig EV5G), indicating expression of a functional Ehmt1 fusion protein (Bio-Ehmt1). ChIP-seq using Streptavidin beads identified broad Ehmt1-occupied chromatin domains that, consistent with an enzyme–substrate relationship, scaled with H3K9me2 genome-wide (Fig EV5H). To determine overlap with Zfp281, we quantified Ehmt1 and H3K9me2 enrichment at Zfp281-bound and matching unbound control windows (see Materials and Methods for details). Ehmt1 localization at both sets of regions was unchanged during differentiation or in *Zfp281* KO cells (Fig EV5I and J), indicating that Zfp281 is not required for Ehmt1 localization on chromatin. In contrast to Ehmt1, H3K9me2 increased in *WT*$^{32\ h}$ cells and EpiLCs (Kurimoto *et al*, 2015), with a more pronounced increase at Zfp281-occupied loci than control windows (Fig 7A). To test whether these dynamics require Zfp281, we performed immunofluorescence staining, which revealed that mutant cells failed to gain H3K9me2 by 32 h after 2i withdrawal (Fig EV5K). H3K9me2 ChIP-seq in *Zfp281* mutants confirmed that H3K9me2 levels were unaffected in *Zfp281*^{2i} cells but did not increase in *Zfp281*$^{32\ h}$ cells (Figs 7A and EV5J). Impaired gain of H3K9me2 was observed at both Zfp281-bound and Zfp281-unbound sites, suggesting that Zfp281 is a differentiation-specific pervasive activator of Ehmt1 during ESC transition. Zic2 ChIP-seq identified 28,495 peaks, of which approximately 30% overlapped with Zfp281 (Fig 7B). These were closer to promoters and enriched for H3K27ac compared to Zic2-only and Zfp281-only peaks (Fig EV5L and M), suggesting co-localization of Zfp281 and Zic2 at CREs. In *WT*$^{32\ h}$ cells, Zic2 increased predominantly at co-bound peaks (Fig 7B). Although we noted a general reduction of Zic2 on chromatin in *Zfp281* mutants, Zic2 localization was particularly perturbed at co-bound sites in *Zfp281*$^{32\ h}$ cells (Figs 7B and EV5N). Our findings therefore suggest that Zfp281 engages Ehmt1 and Zic2 during ESC differentiation through chromatin co-occupancy-dependent and chromatin co-occupancy-independent mechanisms.

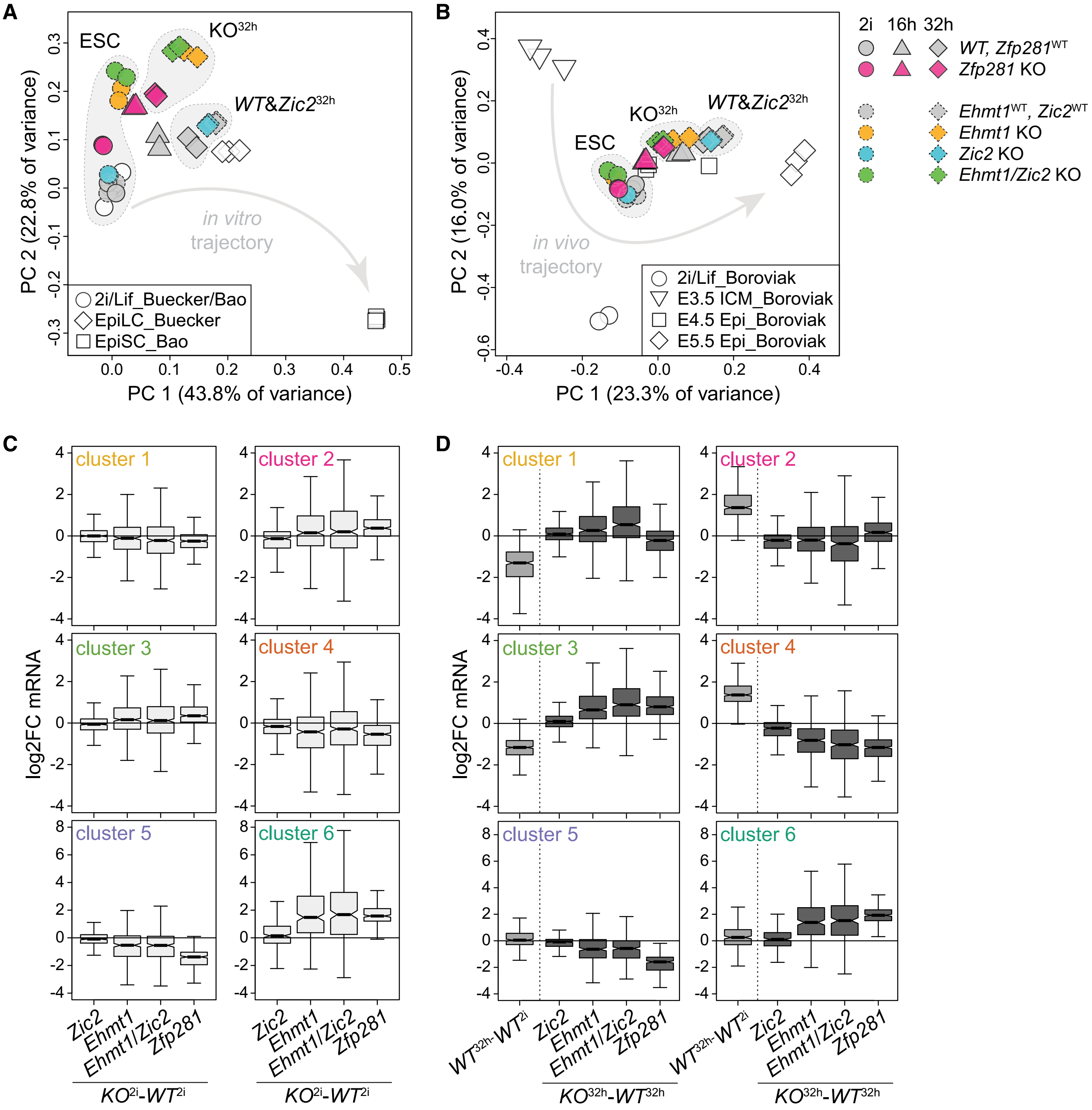


Figure 6. Related transcriptional defects in *Ehmt1/Zic2* and *Zfp281* KO cells.

A, B PC analysis of indicated samples normalized to *WT* ESCs (A) and all samples (B) within each dataset. Arrows indicate developmental trajectories. Full and dashed outlines indicate independent RNA-seq experiments.

C, D Quantification of cell state-specific mRNA $\log_2$FC of gene clusters 1–6 between indicated genotypes in 2i (C) and 32 h after 2i withdrawal (D). Boxes and number of datapoints as in Fig 3C.

At Zfp281-bound peaks, the gain of H3K9me2 and Zic2 occurred mostly at mutually exclusive sets of genomic loci with reduced and increased Zfp281 binding during differentiation, respectively (Fig 7C and D), indicating that these are sites of direct physical interaction that control transcription of nearby genes. We therefore stratified H3K9me2 and Zic2 binding dynamics at Zfp281 peaks by gene cluster association. H3K9me2 increased predominantly at peaks linked to repressed clusters 1 and 3, and Zic2 at peaks belonging to induced clusters 2 and 4 (Fig 7E). Surprisingly, the gain in *WT*$^{32\ h}$ cells and reduction in *Zfp281*$^{32\ h}$ cells were indistinguishable between clusters 1 and 3 (for H3K9me2) and clusters 2 and 4 (for Zic2). Hence, Zfp281 catalyzes H3K9me2 and Zic2 deposition at transition-associated CREs genome-wide and without any qualitative or quantitative specificity for its transcriptional targets in clusters 3 and 4. Why cluster 1 and 2 gene expression is insensitive to perturbation of H3K9me2 and Zic2 dynamics in *Zfp281* KO cells (Figs 3C and 7E) remains to be determined. Additional chromatin regulators may be involved, since transcription of clusters 1 and 2 is also less sensitive to *Ehmt1* depletion than of clusters 3 and 4 (Fig 6D). We therefore propose that Zfp281 drives and stabilizes transition-specific transcription, at least in part, through activation of Ehmt1

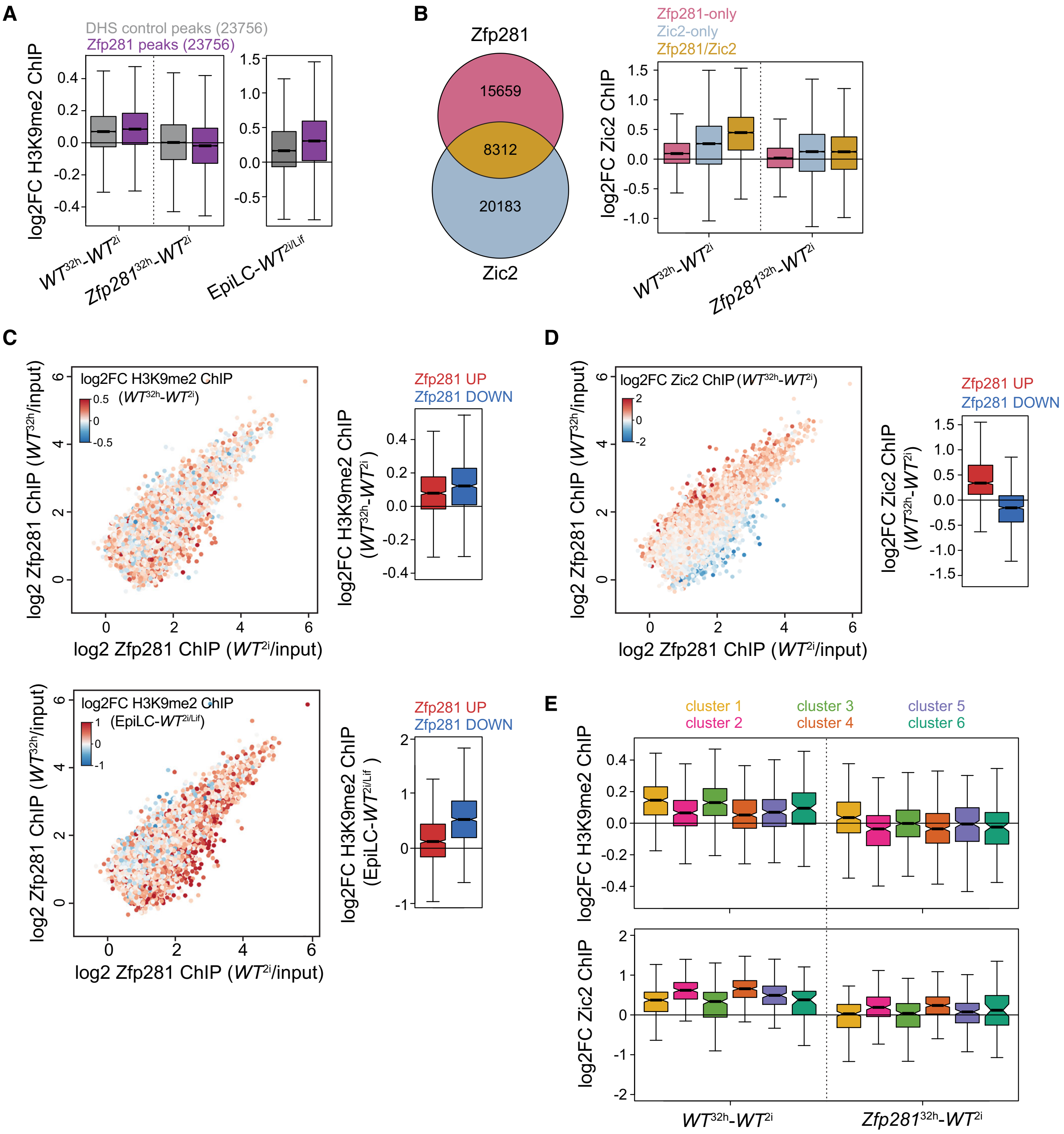


Figure 7.

◀ **Figure 7. Zfp281 engages with Ehmt1 and Zic2 at developmental CREs.**

A H3K9me2 ChIP $\log_2$FC between indicated cell states and genotypes at 10-kb windows surrounding Zfp281-bound (purple) or matching DNase-hypersensitive site (DHS) control peaks (gray). Boxes as in Fig 3C for 23,756 datapoints each.

B Overlap of Zfp281 and Zic2 ChIP peaks (left) and Zic2 ChIP $\log_2$FC between specified cell states and genotypes at indicated peak subsets (right). Boxes as in Fig 3C for 15,659 (Zfp281-only), 20,183 (Zic2-only), and 8,312 (Zfp281/Zic2) datapoints.

C, D Same as in Fig 3F. Coloring is according to H3K9me2 ChIP $\log_2$FC between $WT^{32\ h}$ and WT^{2i} cells (C, top left) and between EpiLCs and $WT^{2i/Lif}$ cells (C, bottom left) at Zfp281 peaks extended to 10-kb windows, and according to Zic2 ChIP $\log_2$FC between $WT^{32\ h}$ and WT^{2i} cells (D, left). Quantification of corresponding ChIP changes at top 1,000 Zfp281 peaks with increased (red) or decreased (blue) Zfp281 binding during ESC differentiation (right). Boxes and number of datapoints as in Fig 3F.

E H3K9me2 (top) and Zic2 (bottom) ChIP $\log_2$FC between indicated cell states and genotypes at all Zfp281 peaks extended to 10-kb windows (top) or Zfp281/Zic2 co-bound peaks (bottom) associated with nearest TSSs of cluster 1–6 genes. Boxes and number of datapoints as in Fig 3C.

at cluster 3 CREs and recruitment or stabilization of Zic2 at cluster 4 CREs.

Discussion

Cellular plasticity in response to injury *in vivo* or TF overexpression *in vitro* is frequently accompanied by the reversal of cellular specialization (Merrell & Stanger, 2016). Although single-cell profiling has shown that this process is not a strict inversion of natural development (Treutlein *et al*, 2016; Gerber *et al*, 2018), differentiation and de-differentiation trajectories may mechanistically intersect. We aimed to uncover such plasticity regulators in pluripotent cells and performed loss-of-function screening for genes that both drive exit from ESC self-renewal and shield EpiSCs from reprogramming into the pluripotent ground state. Within the experimental limitations of this approach, we identified only one gene, the TF *Zfp281*. Such exclusivity suggests a prominent role in establishing and maintaining the unidirectionality of pluripotent cell state progression *in vitro*. The former is consistent with perturbed epiblast maturation in *Zfp281* mutant embryos (Huang *et al*, 2017), but whether Zfp281 protects cellular identities against de-differentiation *in vivo* remains to be determined. We, however, note that Zfp281 restrains iPSC formation from fibroblasts at a late pre-iPSC stage (Fidalgo *et al*, 2012), supporting the notion that resetting of EpiSCs into naïve pluripotency recapitulates a late phase of somatic cell reprogramming (Dunn *et al*, 2019). Other factors that, similar to Zfp281, drive differentiation and inhibit de-differentiation of cell states not represented in our ESC-EpiSC conversion system are likely to exist.

We showed that Zfp281 is important for robust ESC differentiation. This is reminiscent of lineage-specifying TFs that are specifically expressed in the lineages they instruct (Graf & Enver, 2009). In adult mice, Zfp281 is indeed transcribed strongest in heart tissue and its overexpression in fibroblasts enhances cardiac reprogramming (Zhou *et al*, 2017). During ESC differentiation, however, Zfp281 neither changes expression nor occupies distinct genomic sites, indicating a facilitating, rather than specifying, function. Using biochemical, genetic, and genomic experiments, we provide evidence that Zfp281 directs sequential gene expression through permissive and instructive mechanisms involving physical interaction with Ehmt1 and Zic2. Cluster 5 and 6 genes are differentially expressed in EpiSCs, but transcriptionally primed by Zfp281 throughout differentiation, including in the naïve pluripotent starting cell state. Cluster 5 contains modulators of cell adhesion that may contribute to polarization of post-implantation epiblast cells. However, although *Zfp281* null embryos degenerate during gastrulation, they do form an epithelial egg cylinder (Huang *et al*, 2017), suggesting operation of compensatory mechanisms *in vivo*. Cluster 3 and 4 genes are, in contrast, regulated by Zfp281 predominantly during transition, enriched for developmental functions, and therefore likely responsible for extinguishing ESC identity. Strikingly, Zfp281 quantitatively decreases at CREs associated with repressed cluster 3 genes and, *vice versa*, increases at CREs associated with induced cluster 4 genes. Concomitant gain of H3K9me2 and Zic2 at cluster 3 and 4 CREs, respectively, suggests that subtle Zfp281 chromatin binding differences are converted into CRE activity changes by reciprocal activation of Ehmt1 and Zic2. Consistently, embryonic arrest in *Ehmt1* and *Ehmt2* mutant mouse embryos (Tachibana *et al*, 2002, 2005) has been attributed to reduced H3K9me2 deposition at and impaired silencing of developmental enhancers (Zylicz *et al*, 2015), while Zic2 triggers neural plate gene expression in EpiSCs through enhancer activation (Iwafuchi-Doi *et al*, 2012). The molecular mechanisms inducing differential binding of Zfp281 to Ehmt1 and Zic2 remain to be elucidated. Because of similar overexpression phenotypes in ESCs, we surmise that protein ratios are rate-determining. Zic2 protein levels, despite unchanged mRNA (Fig 4B), indeed increase during differentiation, while Ehmt1 persists (Fig EV4G). However, Ehmt1 and Ehmt2 interact with additional zinc finger TFs (Shinkai & Tachibana, 2011) that may compete with Zfp281 for binding to Ehmt1 specifically in ground state ESCs. Taken together, we suggest that Zfp281 promotes stable cell state transition by permissively marking cluster 5 and 6 genes for forthcoming developmental regulation, and instructing cluster 3 and 4 gene expression for elimination of naïve pluripotency, the latter through differential engagement with Ehmt1 or Zic2 at CREs.

Zfp281 drives differentiation to and inhibits reprogramming of EpiSCs, and therefore qualifies as a bidirectional cell state regulator with antipodal functions during developmental progression and reversion. Whether this is because Zfp281-Ehmt1/Zic2 control the same cell state transition and act through the same target genes in mutual directions remains to be determined. We showed that *Zfp281* KO cells exposed to differentiation-promoting conditions are less clonogenic than ESCs and give rise to an equilibrium of $\text{Rex1}^{\text{high}}$ and Rex1^{low} cell states that features defective coupling of Rex1 downregulation with exit from self-renewal. It is therefore conceivable that Zfp281 acts after an initial commitment step to induce and stabilize irreversible silencing of naïve identity. During reprogramming, Zfp281 may inversely antagonize induction of naïve pluripotency prior to formation of a $\text{Rex1}^{\text{high}}$ state, which has recently been shown to mark transition intermediates with full ESC self-renewal activity and destined to give rise to Epi-iPSCs (Stuart *et al*, 2019). Acting at a late stage of reprogramming also reconciles our finding of Zfp281 impairing Epi-iPSC formation by STAT3 activation and Esrrb overexpression (Figs 1G and EV1E) with the notion that both regimes proceed along distinct transcriptional trajectories before converging on a Rex1^{high} cell state (Stuart *et al*, 2019). However, Zfp281 depletion in EpiSCs also causes cell cycle arrest and apoptosis (Fidalgo *et al*, 2016), but not upregulation of naïve pluripotent TFs such as Nanog (Fig EV2K), and is not alone sufficient for reprogramming to Epi-iPSCs in the presence of 2i (Fig 1F). Zfp281 activates transcription of Nodal signaling components, such as the cluster 5 genes *Lefty1* and *Lefty2*, in ESCs, EpiSCs, and the post-implantation epiblast (Huang *et al*, 2017). Nodal signaling is required for EpiSC maintenance (Vallier *et al*, 2009), but dispensable for exit from ESC self-renewal (Mulas *et al*, 2017) and somatic cell reprogramming (Ruetz *et al*, 2017). We therefore speculate that Zfp281 controls the EpiSC state and reprogramming through different effector genes, the former through regulation of cluster 5/6 genes, e.g., *Lefty1*, *Lefty2* or cell polarity regulators, and the latter through stabilization of cell state transition-specific cluster 3/4 genes.

In metastable serum/Lif ESCs, Zfp281 is reported to promote DNA methylation and differentiation by recruiting Tet1 and silencing Tet2 (Fidalgo *et al*, 2016). Upon exit from the naïve ESC state, Zfp281 binds to Dnmt3a, Dnmt3b, and Dnmt3l (Fig 4A), suggesting that Zfp281 controls 5-methylcytosine turnover through regulating DNA methylating and demethylating enzymes. Furthermore, Dnmt1, Dnmt3a, and Dnmt3b interact with the Ehmt1-Ehmt2 heterodimer (Estève *et al*, 2006; Epsztejn-Litman *et al*, 2008), H3K9me2 and 5-methylcytosine marks overlap genome-wide (von Meyenn *et al*, 2016), and DNA is hypomethylated upon depletion or catalytic inhibition of Ehmt2 in ESCs (Zhang *et al*, 2016b). We, however, showed that Zfp281 drives exit from naïve pluripotency independent of Dnmt3a, Dnmt3b, Tet1, and Tet2. Although the specific contributions of 5-methylcytosine, H3K9me2, and other Ehmt substrates (Sim *et al*, 2017) to pluripotent cell plasticity remain to be determined, our findings suggest that resolution of naïve pluripotency *in vitro* is masked or mechanistically distinct in heterogeneous Serum/Lif ESC cultures. Similarly, Zic2 has previously been reported to act as a repressor in metastable serum/Lif ESCs (Luo *et al*, 2015), but we detect only minor transcriptional defects in naïve *Zic2* KO ESCs.

Using the paradigm of pluripotent cell state conversion *in vitro*, we here identify, in Zfp281, a regulator of cellular plasticity that modulates CRE activity and transcription without reliance on cell type-restricted expression or chromatin occupancy. We propose that the persistence of Zfp281 at developmental CREs throughout cell state transition provides a molecular platform for ordered remodeling of the cis-regulatory architecture and further consolidation by lineage-specifying TFs. These findings demonstrate that differential gene expression is not a necessary criterion of cellular plasticity regulators, and we suggest that this feature may not be limited to pluripotent cells.

Materials and Methods

Cell culture

Embryonic stem cells (male and female RGd2 cells containing a Rex1:GFPd2-IRES-blasticidin (Wray *et al*, 2011), male O4GIP ESCs containing a GFP-IRES-puromycin transgene under control of an Oct4 regulatory element (Betschinger *et al*, 2013), and male E14 cells) were cultured on plastic coated with gelatin or laminin (Sigma) in N2B27 medium [DMEM/F12 (Life Technologies), Neurobasal (Gibco) supplemented with N2 (homemade) and B-27 serum-free supplement (Gibco), 2 mM L-glutamine (Gibco), and 0.1 mM 2-mercaptoethanol (Sigma)] with 2i [3 μM CHIR99021 and 1 μM PD0325901 (Steward Lab, Dresden)] and, where indicated, with 1 μg/ml doxycycline (Sigma). EpiSCs [O4GIP and OEC2 (Guo *et al*, 2009) and 796.4 (Yang *et al*, 2010)] were cultured on plastic coated with fibronectin (Millipore) in N2B27 with bFGF (12 ng/ml) and activin A (20 ng/ml) (FA) (Smith Lab, Cambridge).

For monolayer differentiation, ESCs were seeded on gelatin-coated plates at 1.5 × 10^4 cells/cm^2 in 2i, and the following day, 2i was withdrawn. Cells were incubated in N2B27 for 32 h or 72 h, or continuously passaged on laminin-coated plates, as indicated. For EpiLC differentiation for 48 h (Hayashi *et al*, 2011), medium was adjusted to FA and 1% knockout serum replacement (Thermo Fisher). For embryoid body (EB) differentiation, ESCs were seeded at 2.5 × 10^4 cells/cm^2 on ultra-low attachment plates (Corning) in serum media [GMEM (Sigma) supplemented with 10% fetal bovine serum (Sigma), 1 mM sodium pyruvate (Gibco), 2 mM L-glutamine (Gibco), 0.1 mM non-essential amino acids (Gibco), and 0.1 mM 2-mercaptoethanol (Sigma)] for 72 h. Exit from pluripotency was quantified by measuring GFP fluorescence in RGd2 cells on a LSRII SORP Analyzer (Becton Dickinson) and analyzed using FlowJo (FlowJo, LLC), and by performing self-renewal and commitment assays as described before (Betschinger *et al*, 2013). Briefly, differentiated RGd2 and E14 cells at indicated time points were plated at clonal density on laminin-coated plates in 2i medium, and resulting colonies were stained for alkaline phosphatase activity (Sigma) according to the manufacturer's instruction and counted. RGd2 cells were additionally selected with 10 μg/ml blasticidin (Thermo Fisher). O4GIP cells were differentiated for 72 h and treated with 2i medium containing 1 μg/ml puromycin (Gibco), and uncommitted cells were quantified after 3 days by adding 1:10 diluted Alamar Blue (Invitrogen) in 2i medium, following by readout on a SpectraMax Gemini EM (Molecular Devices) microplate reader. For cell cycle analysis, cells were fixed in cold 70% ethanol for 30 min at 4°C, washed twice with PBS and 0.1% BSA (Sigma), treated with 5 μg RNaseA (Thermo Fisher) for 15 min at room temperature (RT), and stained with 10 μg propidium iodide (Sigma). Cells were analyzed on a LSRII SORP Analyzer, and cell cycle distributions were determined using FlowJo.

siRNA transfections were performed as described (Betschinger *et al*, 2013) using 16.7 nM siRNA (detailed in Table EV4) and transfection mixes in OptiMEM (Invitrogen) containing Lipofectamine 2000 or RNAiMAX (Thermo Fisher) for ESCs or EpiSCs, respectively.

For EpiSC reprogramming, cells were plated at 1.5 × 10^4 cells/cm^2 on fibronectin-coated plates in N2B27 with FA. The next day, medium was changed to 2i and, as indicated, supplemented with 30 ng/ml granulocyte colony-stimulating factor (Gcsf) (PeproTech), 10 ng/ml Lif (Smith lab, Cambridge), or 1 μg/ml doxycycline. After 4 days (d), medium was changed to 2i with 1 μg/ml puromycin, and Epi-iPSC colonies were stained for alkaline phosphatase activity and counted. For experiments shown in Figs 1E and EV1A, cells were subjected to self-renewal assays in 2i after 2 and 4 days of Gcsf supplementation. For experiments shown in Fig EV1D, individual Epi-iPSC colonies were picked and expanded in 2i with puromycin for further experiments.

Spheroid formation of ESCs was performed as described before (Shahbazi *et al*, 2017). Briefly, ESCs grown in 2i medium were washed in N2B27, and 7.5×10^3 cells were resuspended in 25 µl ice-cold growth factor reduced Matrigel (Corning, 356231), plated dropwise on uncoated 96-well glass plates (Greiner Bio-One), and aggregated in N2B27 with or without 2i for 3 or 4 days.

EpiSC screen

O4GIP^{GY118F} EpiSCs were reverse-transfected in fibronectin-coated 384-well plates using mixtures of 50 ng esiRNA and 0.075 µl Lipofectamine 2000 in 10 µl OptiMEM. EpiSCs were plated at a density of 5,000 cells/well in 80 µl N2B27 with FA. Each plate included three negative (Luciferase esiRNA) and two positive (Stat3 esiRNA) control wells. The next day, medium was changed to 2i containing 30 ng/ml Gcsf and 4 days later to 2i containing 1 µg/ml puromycin. After 3–4 days, medium was changed to 2i containing 1/10 vol Alamar Blue (Invitrogen) and cell survival quantified on a SpectraMax M2 (Molecular Devices).

Genome editing

CRISPR/Cas9 genome editing was performed by transient co-transfection of hCas9 and U6-gRNA plasmids (Mali *et al*, 2013) (Addgene plasmids 41,815 and 41,824) and a dsRed expression plasmid into E14 or female RGd2 ESCs. Two days later, single dsRed-positive cells were sorted into gelatin-coated 96-well plates containing serum media supplemented with 10 ng/ml Lif and 2i. Clones were genotyped by sequencing amplified target loci and by confirming protein absence in Western blots. For generation of knockout (KO) cell lines, two independent KO clones (specified in Appendix Fig S2) and, in the case of *Zfp281*, *Ehmt1*, and *Zic2* targeting, one untargeted wild-type sibling clone were kept for further analysis.

N-terminal Flag-Avi tagging of Ehmt1 was performed in *WT* or *Zfp281* KO ESCs constitutively expressing the BirA biotin ligase (see below). The recombination template was generated by cloning homology arms (548 bp upstream and 618 bp downstream of the Ehmt1 transcription start site) into pDONR221 using Gateway technology (Thermo Fisher) and inserting the Flag-Avi sequence by Seamless Cloning (Thermo Fisher). Targeting was performed as above with hCas9, U6-gRNA, and dsRed expression plasmids, but included co-transfection of the recombination template. Targeting was confirmed by sequencing the target locus and by the expression of biotinylated Ehmt1 in Western blots.

gRNA sequences, genotyping primers, and the Ehmt1 recombination template are specified in Table EV4. gRNA sequences targeting *Tet1* and *Tet2* (Wang *et al*, 2013) and *Dnmt3a* and *Dnmt3b* (Domcke *et al*, 2015) have been described.

Immunostaining

Cells seeded on laminin-coated 96-well glass plates were fixed with 4% paraformaldehyde (Electron Microscopy Sciences) for 10 min and spheroids grown in Matrigel for 20 min at RT. Samples were blocked for 1 h in blocking solution (PBS, 0.1% Triton X-100 (Sigma), 3% donkey serum (Sigma), and 1% BSA) and incubated overnight with primary antibodies [Ehmt1 (Abcam, ab41969, 1:300) and H3K9me2 (Abcam, ab1220, 1:300)] at 4°C. After three washes in washing solution (PBS, 0.1% Triton X-100), secondary antibodies were added, DNA was stained with Hoechst33342 (Life Technologies) and, where indicated, incubated with Alexa Fluor 488 Phalloidin (Life Technologies, A12379, 1:40) for 20 min at RT. Images were acquired using a LSM 710 scanning head confocal microscope (Zeiss) at 20× magnification and handled using Fiji and Adobe Photoshop (Adobe). Imaging of spheroids used for quantification in Fig EV4L was performed with a Yokogawa CV7000S high-throughput confocal microscope at 20× magnification. Images were acquired in confocal mode as *z*-stack multiplane images over *z* distance of 50 µm with a 5 µm step size, and maximum intensity projections were stored, representatives of which are shown in Fig 5B.

Molecular biology

Coding sequences for Ehmt1, Esrrb, Klf4, and Zfp281 were amplified from ESC complementary DNA (cDNA) and for BirA biotin ligase from a plasmid (gift of Matyas Flemr, Friedrich Miescher Institute, Basel). For Zic2, the coding sequence was synthesized as a double-stranded gBlock (IDT). Polynucleotides were recombined into pDONR221 using Gateway technology. Ehmt1 point mutations [Ehmt1$^{\Delta NHHC}$: NHHC1198-1201del (Tachibana *et al*, 2008), Ehmt1$^{NH\text{-}LE}$: NH1198-1199LE (Tachibana *et al*, 2008), and Ehmt1^{ank}: W872A, W877A, and E880A (Collins *et al*, 2008)] were introduced by polymerase chain reaction (PCR). Expression destination vectors were pPB-CAG-DEST-pgk-hph (Betschinger *et al*, 2013) and pPB-TRE-DEST-rTA-pgk-hph (Villegas *et al*, 2019), and GY118F expression vector as described (Yang *et al*, 2010). Stable integration into ESCs or EpiSCs after co-transfection with pBASE (Betschinger *et al*, 2013) was selected in the presence of 150 µg/ml hygromycin B (Thermo Fisher).

For relative mRNA quantification, total RNA was isolated from indicated samples using RNeasy Mini Kit (Qiagen) and cDNA prepared using SuperScript III reverse transcriptase (Invitrogen). Quantitative PCR was performed using TaqMan Fast Universal PCR Master Mix (Applied Biosystems) with gene-specific primers, using either the universal probe library (UPL, Roche) or TaqMan system (Applied Biosystems), and a GAPDH probe (Applied Biosystems) for normalization. Oligonucleotide sequences and probes are listed in Table EV4.

Protein methods

Cell lysates for Western blotting were generated in RIPA buffer (50 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, and 0.1% SDS). Primary antibodies were anti-GAPDH (Sigma, G8795, 1:5,000), anti-Ehmt1 (Abcam, ab41969, 1:500), anti-Tet1 (Millipore, 09-872, 1:1,000), anti-Tet2 (Abcam, ab124297, 1:300), anti-Zic2 (Abcam, ab150404, 1:500), anti-Zfp281 (Bethyl Laboratories, A303-118A, 1:500), and anti-Streptavidin coupled to HRP (Sigma, EV2438, 1:1,000).

Nuclear immunoprecipitations (IPs) for Zfp281 were performed in three biological replicates using *WT*^{2i}, *Zfp281*^{2i}, *WT*$^{40\ h}$, and *Zfp281*$^{40\ h}$ cells. Cells were washed with cold PBS, resuspended in five packed cell volume (pcv) of buffer A (10 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, and 10 mM KCl), incubated for 10 min on ice, and broke open using a Dounce homogenizer. Nuclei were pelleted at 3,300 *g* for 15 min at 4°C, resuspended in 3 pcv of buffer B (20 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 0.2 mM EDTA, and 20% glycerol) supplemented with 420 mM NaCl, Complete Mini protease, PhosSTOP phosphatase inhibitors (Roche), and 250 U/ml Benzonase (Sigma), and incubated for 30 min at 4°C on a rotating wheel. Insoluble material was pelleted at 25,000 *g* for 30 min at 4°C, and the supernatant diluted with buffer B to a final concentration of 150 mM NaCl and including 0.02% NP40. 1% of the supernatant was kept as input sample and the remainder incubated with 10 μl Dynabeads Protein G (Invitrogen) and 1 μg Zfp281 antibody (Bethyl Laboratories, A303-118A) for 1 h at 4°C on a rotating wheel. Beads were collected on a magnetic rack for 2–3 min to remove the supernatant, and washed four times in 1 ml buffer B containing 150 mM NaCl and 0.02% NP40 for 10 min each at 4 °C on a rotating wheel. For mass spectrometry, proteins were digested on the beads as described before (Villegas *et al*, 2019).

Whole-cell lysate IPs using 1 μg Zfp281 antibody (Bethyl Laboratories, A303-118A) were performed as described before (Villegas *et al*, 2019) using Dynabeads.

Chromatin immunoprecipitation

For ChIP of endogenous proteins or histone modifications, 8×10^6 cells per IP were fixed for 10 min with 1.1% formaldehyde in fixing solution (0.1 M NaCl, 1 mM EDTA, 0.5 mM EGTA, and 50 mM HEPES pH 7.5) at RT on a rotating wheel, and neutralized with glycine to a final concentration of 0.125 M for 5 min at RT. Cells were washed three times with ice-cold PBS by spinning at 1,600 *g* for 5 min at 4°C, incubated for 10 min at 4°C on a rotating wheel with 1 ml lysis buffer 1 (50 mM HEPES pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, and 0.25% Triton X-100), pelleted, and incubated for a further 10 min at 4°C in 1 ml lysis buffer 2 (10 mM Tris pH 8.0, 200 mM NaCl, 1 mM EDTA, and 0.5 mM EGTA). Nuclei were pelleted, resuspended in 140 μl shearing buffer (50 mM Tris pH 8.0, 10 mM EDTA, and 1% SDS), and sonicated in Diagenode 15-ml Falcon tubes for 25 cycles (30 s ON, 30 s OFF) in ice-cold water using a Bioruptor Plus (Diagenode). 10% of sonicated DNA was kept as input sample. Lysates were further pelleted at 14,000 *g* for 10 min at 4°C, and the supernatant diluted 1:10 with ChIP dilution buffer (50 mM Tris pH 8.0, 167 mM NaCl, 1.1% Triton X-100, and 0.11% Na-deoxycholate). Lysates were precleared over 10 μl Dynabeads for 2 h and incubated overnight at 4°C on a rotating wheel with the following antibodies: 2 μg H3K27ac (Active Motif, 39135), 2 μg H3K9me2 (Abcam, ab1220), 2 μg Zic2 (Abcam, ab150404), or 2 μg Zfp281 (Bethyl Laboratories, A303-118A). The next day, 10 μl Dynabeads were added and incubated with lysates for 1 h at 4°C on a rotating wheel. Beads were washed with 1 ml of the following buffers for 5 min each at 4°C: twice with wash buffer 1 (50 mM Tris pH 8.0, 0.1% SDS, 0.1% Na-deoxycholate, 1% Triton X-100, 150 mM NaCl, 1 mM EDTA, and 0.5 mM EGTA), once with wash buffer 2 (50 mM Tris pH 8.0, 0.1% SDS, 0.1% Na-deoxycholate, 1% Triton X-100, 500 mM NaCl, 1 mM EDTA, and 0.5 mM EGTA), once with wash buffer 3 (50 mM Tris pH 8.0, 250 mM LiCl, 0.5% Na-deoxycholate, 0.5% NP40, 1 mM EDTA, and 0.5 mM EGTA), and twice with wash buffer 4 (50 mM Tris pH 8.0, 10 mM EDTA, and 5 mM EGTA). Finally, beads were eluted twice with 100 μl elution buffer (0.1 M $NaHCO_3$, and 1% SDS) for 15 min at RT in a shaker at maximum speed, and combined supernatants de-crosslinked overnight by supplementation to 200 mM NaCl and continuous shaking at maximum speed at 65°C. The same procedure was followed for input samples by adjusting the total volume of elution buffer to 200 μl and 200 mM NaCl. The next day, DNA was purified using MinElute PCR Purification Kit (Qiagen).

Bio-ChIP for Flag-Avi-tagged Ehmt1 was performed as described before (Ostapcuk *et al*, 2018) with minor modifications. Briefly, 8×10^6 cells per IP were fixed for 8 min with 1% formaldehyde in PBS at RT on a rotating wheel, and neutralized with adjusting glycine to a final concentration of 0.125 M and incubation for 1 min at RT and for 5 min on ice. Cells were washed three times with ice-cold PBS and pelleted at 1,000 *g* for 5 min at 4°C. Cells were lysed in lysis buffers 1 and 2 as described above. Nuclei were washed once in 5 ml NUC buffer (15 mM HEPES pH 7.5, 60 mM KCl, 15 mM NaCl, and 0.32 mM sucrose) and resuspended in 1 ml NUC buffer supplemented with Complete Mini protease inhibitors, 3.3 μl 1 M $CaCl_2$, and 2–3 μl Micrococcal Nuclease (Cell Signaling, 10011S). Enzymatic activity was induced for 15 min at 37°C and shaking at 1,000 rpm, and stopped by addition of 50 μl of STOP solution (250 mM EDTA and 500 mM EGTA) and 110 μl of 10× ChIP buffer (167 mM Tris pH 8.0, 1.67 M NaCl, 12 mM EDTA, 10% Triton X-100, and 0.1% SDS) with a further incubation for 5 min on ice. Nuclei were gently disrupted by sonication in Diagenode 15-ml Falcon tubes for eight cycles (5 s ON, 5 s OFF) in ice-cold water using a Bioruptor Plus. Lysates were centrifuged at 14,000 *g* for 5 min at 4°C, 5% of the supernatant was kept as input sample, and the remaining supernatant precleared for 2 h over 10 μl Dynabeads at 4°C on a rotating wheel. Chromatin was incubated for 1 h with M-280 Streptavidin-coupled Dynabeads (Invitrogen) at 4°C on a rotating wheel, and washed with 1 ml of the following buffers for 5 min each at 4°C: twice with TE buffer (10 mM Tris pH 8.0, and 1 mM EDTA) supplemented with 2% SDS, once with high-salt buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, and 0.1% Na-deoxycholate), once with wash buffer 3 (see above), and twice with TE buffer. Beads were eluted in 60 μl elution buffer (see above) supplemented with 2 μl RNaseA (10 mg/ml stock) and incubated for 30 min at 37°C while mixing. After supplementation to 10 mM EDTA, 10 mM Tris pH 8.0, and 2 μl proteinase K (10 mg/ml, Promega), the bead suspension was further incubated for 3 h at 55°C and overnight at 65°C while shaking. The same procedure was followed for input samples, including RNaseA and proteinase K digestion. DNA was purified using AMPure XP beads (Beckman Coulter).

Sequencing libraries

RNA from ESCs grown in 2i and 16 and 32 h after 2i withdrawal was isolated using RNeasy kit (Qiagen). For *Zfp281* KO and corresponding *WT* cells, total RNA was subjected to ribosomal RNA depletion using Ribozero removal kit (Illumina) followed by library construction using ScriptSeq V2 library preparation kit (Illumina). For *Ehmt1*, *Zic2*, *Ehmt1/Zic2* KO, and corresponding *WT* cells, RNA-seq libraries were prepared using TruSeq mRNA Library Preparation Kit (Illumina). ChIP-seq libraries were prepared using NEBNext Ultra kit (New England BioLabs) following the manufacturer's recommendations. Sequencing was performed on an Illumina HiSeq2500 machine (50-bp single-end reads).

Screen analysis

For the EpiSC reprogramming screen, *Z* scores were calculated for each plate, excluding the two outer most columns and rows (Table EV1). Screen replicates are presented in Fig EV1B. Average *Z* scores > 2 were considered as screen hits, identifying 146 genes (Fig 1B). We quantified their role in exit from the ESC state by extracting primary data for these 146 genes from two previous ESC differentiation studies (Betschinger *et al*, 2013; Li *et al*, 2018), and computing *Z* scores on this subset (Table EV1). Sixty-seven and 129 of the 146 genes were mapped on results from Betschinger *et al* (2013) and Li *et al* (2018), respectively, resulting in mapping of 130 genes in total. Average *Z* scores from both ESC differentiation screens are presented in Fig 1C.

Quantification of immunostaining

Quantification of H3K9me2 was performed in CellProfiler3 (Broad Institute). Nuclei were identified using Hoechst33342 staining and average H3K9me2 and DNA fluorescence intensity per nucleus determined: WT^{2i} (453 cells), $Zfp281^{2i}$ (574 cells), $WT^{32\ h}$ (465 cells), $Zfp281^{32\ h}$ (792 cells).

For quantification of d3 and d4 spheroids from two biological replicates, images were stitched to generate a single image per channel and per well, and used for object segmentation with MATLAB (MathWorks). Segmented object outlines were exported and used for feature extraction with software package CellProfiler3. Extracted features (84 features) describing object area, shape, and intensity were normalized within corresponding assay plates using *Z* score transformation and unified into a cross-comparable dataset. Normalized features were used to extract 10 principal components (PC), which were further used for unsupervised clustering with the software package PhenoGraph (Levine *et al*, 2015). Unsupervised clustering returned 13 clusters, and objects belonging to the two sparsest clusters were discarded as outliers based on extreme variance in measured features in the respective classes. To describe radial distribution of F-actin signal intensity, segmented objects were divided into 20 concentric regions of same width. Distribution of F-actin staining intensity was described as mean fraction of intensity per region (MeanFrac) whereby the total intensity in the given region was normalized to the fraction of object area corresponding to the region. Mean fractions of intensity per region were not *Z* scored as these values were normalized per object *ad initio*. Visualization in Fig EV4L shows the mean of F-actin signal in a given concentric ring of all d4 spheroids per genotype and medium condition: WT^{2i} (559 spheroids), $Ehmt1^{2i}$ (349 spheroids), $Zic2^{2i}$ (292 spheroids), $Ehmt1/Zic2^{2i}$ (164 spheroids), $Zfp281^{2i}$ (324 spheroids), $WT^{96\ h}$ (437 spheroids), $Ehmt1^{96\ h}$ (228 spheroids), $Zic2^{96\ h}$ (202 spheroids), $Ehmt1/Zic2^{96\ h}$ (233 spheroids), $Zfp281^{96\ h}$ (547 spheroids). Heatmap was generated using aheatmap function from the Bioconductor package NMF (Gaujoux & Seoighe, 2010) (package version 0.21.0).

Protein identification and quantification

Relative quantification of mass-spec data from three biological replicates (Table EV3) was performed with MaxQuant (version 1.5.3.8) using Andromeda as search engine and label-free quantification as described (Villegas *et al*, 2019). Briefly, the mouse subset of the UniProt version 2015_01 combined with the contaminant DB from MaxQuant was searched, and the protein and peptide FDR were set to 0.01. For Fig 4A and B, proteins were considered, which passed an interaction threshold of an at least twofold enrichment in IPs from WT^{2i} or $WT^{40\ h}$ lysates compared to matched Zfp281 KO lysates with a significant *P*-value < 0.1, and were quantifiable with at least two unique razor peptides.

RNA-seq and GRO-seq analysis

RNA-seq reads from wild-type and mutant cells in 2i, and 16 and 32 h after 2i withdrawal were aligned to the mouse GRCm38/mm10 genome using qAlign from the Bioconductor package QuasR (Gaidatzis *et al*, 2015) (package version 1.22.0) with default parameters except for splicedAlignment=TRUE. Published RNA-seq from ESCs cultured in 2i/Lif and EpiLC (Buecker *et al*, 2014) was 36-bp reads, and therefore, no spliced alignment could be performed. RNA-seq from ESCs cultured in 2i/Lif and EpiSCs (Factor *et al*, 2014; Bao *et al*, 2018), and global run-on sequencing (GRO-seq) data from 2i/Lif-cultured ESCs (Dorighi *et al*, 2017) were 100-bp and 50-bp paired-end reads, respectively, and therefore, paired="fr" was used. For *in vivo* embryo data (Boroviak *et al*, 2015), preexisting alignments to mouse GRCm38/mm10 genome were downloaded from ArrayExpress (E-MTAB-2958) and used. Alignments were quantified for known UCSC genes obtained from the TxDb.Mmusculus.UCSC.mm10.knownGene package (package version 3.4.4) using qCount from the Bioconductor package QuasR with default parameters (Table EV2).

Only transcripts with at least three counts per million in at least two biological samples from this study were considered as expressed genes (total: 13,096 genes). For identification of significantly deregulated genes, edgeR (Robinson & Oshlack, 2010; package version 3.24.0) was used and detected genes were fitted to two generalized linear models:

(a) ~ *time* + *genotype* + *time:genotype*: This model uses WT^{2i}, $WT^{16\ h}$, $WT^{32\ h}$, $Zfp281^{2i}$, $Zfp281^{16\ h}$, and $Zfp281^{32\ h}$ expression datasets. Genes with a significant *time* coefficient are genes that change either between WT^{2i} and $WT^{16\ h}$ or WT^{2i} and $WT^{32\ h}$ cells, genes with a significant *genotype* coefficient are differentially expressed between $Zfp281^{2i}$ and WT^{2i} cells, and genes with a significant interaction term *time:genotype* are deregulated in *Zfp281* KO cells specifically during 16-h or 32-h differentiation.

(b) ~ *genotype*: This model identifies genes that differ between $Ehmt1^{2i}$ and WT^{2i} cells.

Raw *P*-values were corrected for multiple testing by calculating false discovery rates (FDR). Significant genes were identified as genes with an absolute $\log_2$ fold change > 1.0 and an FDR of < 0.01.

For visualization of RNA-seq data, except principal component analysis (PCA) in Fig 6A and B and heatmap of selected markers in

Fig EV5A, $\log_2$ fold-change values were used that were obtained from edgeR by fitting the indicated datasets to the following models:

(c) ~ *time* (EpiLCs or EpiSCs): EpiLCs (Buecker *et al*, 2014) or EpiSCs (Factor *et al*, 2014) compared to $WT^{2i/Lif}$ (Buecker *et al*, 2014) and EpiSC compared to $WT^{2i/Lif}$ (Bao *et al*, 2018; used for Fig 3A and B).

(d) ~ *genotime*: $Zfp281^{16\ h}$ or $Zfp281^{32\ h}$ compared to WT^{2i} (used for Figs 3A and B, and EV3A), where *genotime* is the combination of genotype and time.

(e) ~ *genotype* (cell state-specific): KO cells in 2i compared to WT^{2i}, or KO cells 16 h or 32 h after 2i withdrawal compared to $WT^{16\ h}$ or $WT^{32\ h}$, respectively (used for Figs 3C and 6C and D, and EV5B–E).

For heatmap visualization in Fig 3A, only significantly deregulated genes in at least one condition were considered (model (a): *time*, *genotype*, or *time:genotype*; total: 2,495 genes) and the following $\log_2$ contrast was used for clustering: $WT^{16\ h}$-WT^{2i}, $WT^{32\ h}$-WT^{2i}, $Zfp281^{2i}$-WT^{2i}, $Zfp281^{16\ h}$-WT^{2i}, $Zfp281^{32\ h}$- WT^{2i}, $Zfp281^{16\ h}$-$WT^{16\ h}$, and $Zfp281^{32\ h}$-$WT^{32\ h}$. For heatmap visualization in Fig EV5A, RNA-seq read counts were normalized (divided by the total number of aligned reads (library size), multiplied with minimal library size, and added with a pseudocount of 8) and $\log_2$-transformed, and the mean of biological replicates was plotted. For visualization in Fig EV5B, significantly deregulated genes in $Ehmt1^{2i}$ (model (b): *genotype*) but not $Zfp281^{2i}$ (model (a): *genotype*) are highlighted. In Fig 3B, dots represent the median, and shades, the lower and upper quartile of indicated samples. Boxplots were generated using the boxplot function in R with default parameters except outline=FALSE. Correlation plots (Fig EV5C and D) were generated using corrplot function from the Bioconductor package corrplot (https://github.com/taiyun/corrplot) (package version 0.84). Heatmaps (Figs 3A and EV3A, and EV5A) were generated using aheatmap function from the Bioconductor package NMF. For PCA represented in Fig 6A, normalized (see above) and $\log_2$-transformed read counts were centered by subtracting the average of *WT* ESCs within each of the following four RNA-seq datasets: (i) *Zfp281* KO and corresponding *WT* samples from this study; (ii) *Ehmt1*, *Zic2, Ehmt1/Zic2* KO, and corresponding *WT* samples from this study; (iii) $WT^{2i/Lif}$ cells and EpiLCs (Buecker *et al*, 2014); and (iv) $WT^{2i/Lif}$ cells and EpiSCs (Bao *et al*, 2018). For PCA represented in Fig 6B, normalized (see above) and $\log_2$-transformed read counts were centered by subtracting the average over all samples within each of the following three RNA-seq datasets: (i) *Zfp281* KO and corresponding *WT* samples from this study; (ii) *Ehmt1*, *Zic2, Ehmt1/Zic2* KO, and corresponding *WT* samples from this study; and (iii) $WT^{2i/Lif}$ cells and embryo samples (Boroviak *et al*, 2015). Centered read counts for each PCA are provided in Table EV2, and the detailed R code is provided in Dataset EV1. PCA was performed using the prcomp function in R. Analyses of enriched gene sets (Figs EV1C, EV3B and EV5B) were performed using DAVID (Huang *et al*, 2009) (version 6.8) for GO terms of biological processes.

The linear model to estimate synergistic transcriptional effects in *Ehmt1* and *Zic2* KO cells (Fig EV5E) takes into account all (13,096) expressed genes, and was fitted using lm function in R:

$$\Delta Ehmt1/Zic2_{\mathrm{i}} \sim \beta_{\mathrm{E}}\Delta Ehmt1_{\mathrm{i}} + \beta_{\mathrm{Z}}\Delta Zic2_{\mathrm{i}} + \beta_{\mathrm{int}}\Delta Ehmt1{:}Zic2_{\mathrm{i}}$$

Δ: transcriptional difference (KO—WT) of gene $_{\mathrm{i}}$

β: regression coefficient:

$_{\mathrm{E}}$: Ehmt1

$_{\mathrm{Z}}$: Zic2

$_{\mathrm{int}}$: non-additive interaction

ChIP-seq and DHS-seq analysis

ChIP-seq data from this study, published datasets (Buecker *et al*, 2014; Huang *et al*, 2017; Ishiuchi *et al*, 2019), and DNase I hypersensitive site sequencing (DHS-seq) (Encode; accession number: ENCSR000CMW) reads were aligned to mouse GRCm38/mm10 genome using qAlign from the Bioconductor package QuasR with default parameters. Published H3K4me3, H3K27me3, and H3K9me2 ChIP-seq data (Kurimoto *et al*, 2015) were aligned using Bowtie (Langmead *et al*, 2009) (version 4.4.7) with parameter -C in colorspace. Alignments were sorted and indexed using SAMtools (package version 1.2), and all ChIP-seq data were quantified with qCount from the Bioconductor package QuasR. Read counts were normalized (divided by the total number of aligned reads (library size), multiplied with minimal library size, and added with a pseudocount of 8) and $\log_2$-transformed. For DHS-seq, reads per million (RPM) were calculated by dividing the total number of aligned reads, multiplying with one million, adding a pseudocount of 8, and $\log_2$ transforming the data.

We observed a non-linear relationship in Zfp281 ChIP-seq data when comparing read counts in Zfp281 peaks between ChIP-seq replicates in WT^{2i}, and therefore performed loess regression using normalizeBetweenArrays function of the Bioconductor package limma (Ritchie *et al*, 2015) (package version 3.38.2) with method = cyclingloess.

In Ehmt1 ChIP-seq data, we detected a variable dependency of read counts in genomic tiles on the tile's GC composition (GC bias), which was most pronounced in input samples. In order to reduce this bias, we used a loess-based normalization method: First, reads were counted in each sample in 10-kb windows (either genome-tiling windows or windows centered on Zfp281 peaks and corresponding control windows; see below). The counts were then scaled [divided by the total number of aligned reads (library size)], multiplied with minimal library size, added with a pseudocount of 8, and $\log_2$-transformed. A loess curve was fit to the $\log_2$-transformed counts as a function of the fraction of G+C bases in the window using the R function loess with span = 0.3. This fit robustly captures the global signal dependency on the underlying GC composition. GC-corrected $\log_2$ read counts were then obtained by subtracting the values predicted by the loess fit from the observed $\log_2$ read counts (residuals of the fit).

Zfp281 peaks were called on Zfp281 ChIP-seq reads in WT^{2i} and $WT^{32\ h}$ cells using Macs2 (Zhang *et al*, 2008; version 2.1.1.20160309) with default parameters. Peaks that were at least twofold enriched (IP over respective inputs) in at least one of the four Zfp281 ChIP samples were considered (total: 23,756 peaks; Table EV2). For comparison of Zfp281 ChIP samples from this study to ChIP in EpiSCs (Huang *et al*, 2017) and TSCs (Ishiuchi *et al*, 2019), peaks were called on Zfp281 ChIP-seq reads in WT^{2i}, $WT^{32\ h}$, EpiSCs, and TSCs. Peaks that were at least twofold enriched (IP over respective inputs) in at least one of the four Zfp281 ChIP samples from this study, in the one Zfp281 ChIP sample in EpiSCs, or in one of the two Zfp281 ChIP samples in

TSCs were considered (total: 27,435 peaks) and used for plotting (Fig EV3E and F). Zic2 peaks were called on Zic2 ChIP-seq reads in WT^{2i}, $WT^{32\ h}$, $Zfp281^{2i}$, and $Zfp281^{32\ h}$ cells. To quantitatively compare Zfp281 and Zic2 binding, both peak sets were combined and overlapping peak regions were merged using the function reduce from Bioconductor package GenomicRanges (Lawrence *et al*, 2013) (package version 1.34.0). Fused peaks were classified into single- or co-bound as follows: Peaks that were enriched at least twofold (IP over respective inputs) in at least one of the four Zfp281 ChIP samples and in at least one of the eight Zic2 ChIP samples were considered as co-bound (total: 8,312 peaks), while those if detected only in Zfp281 ChIP or only in Zic2 ChIP samples were considered as Zfp281-only (total: 15,659 peaks) or Zic2-only (total: 20,183 peaks), respectively (Table EV2). Reads of H3K27ac ChIP-seq datasets were counted in Zfp281, Zic2, or Zfp281/Zic2 co-bound peaks and normalized as described above.

Zfp281 or Zfp281/Zic2 co-bound peaks were assigned to genes by calculating the distances of peak midpoint to the nearest transcriptional start site (TSS) using a set of non-redundant TSSs with a single start site randomly selected for each gene. Zfp281 peaks were classified as proximal if the distance to the nearest TSS was less than 2,000 bp (7,697 peaks) and as distal otherwise (16,059 peaks; Fig EV3G). H3K27ac reads in WT^{2i} and $WT^{32\ h}$ cells, and H3K4me1 (Buecker *et al*, 2014) and H3K4me3 (Kurimoto *et al*, 2015) reads in $WT^{2i/\mathrm{Lif}}$ and EpiLCs were counted in Zfp281 peak regions and normalized as described above. Proximal Zfp281 binding sites with at least 1.5-fold enrichment of H3K27ac over respective inputs in either WT^{2i} or $WT^{32\ h}$ cells and with at least twofold enrichment of H3K4me3 over respective inputs in either $WT^{2i/\mathrm{Lif}}$ cells or EpiLCs were considered as associated with active promoters (54% of proximal peaks; total: 4,128). Distal Zfp281 binding sites with at least 1.5-fold enrichment of H3K4me1 over respective inputs in either $WT^{2i/\mathrm{Lif}}$ cells or EpiLCs were considered as putative enhancers (62% of distal peaks; total: 9,990), of which sites additionally enriched at least 1.5-fold over respective inputs in H3K27ac in either WT^{2i} or $WT^{32\ h}$ cells were classified as active enhancers (38% of putative enhancers; total: 3,818). For quantification at target sites differentially bound by Zfp281 during differentiation (Figs 3F, and 7C and D), the 1,000 binding sites with strongest increase (Zfp281 UP) and decrease (Zfp281 DOWN) in Zfp281 ChIP signal in $WT^{32\ h}$ compared to WT^{2i} were considered.

Due to the broad chromatin distribution of Ehmt1 and H3K9me2, ChIP-seq reads were first quantified in genome-tiling windows of 10 kb, which were generated using tileGenome function from Bioconductor package GenomicRanges with tilewidth=10000 and cut.last.tile.in.chrom=TRUE. In Fig EV5H, all 10-kb genome-tiling windows were separated in five bins with equal number of tiles but increasing Ehmt1 ChIP $\log_2$ enrichment over respective input in WT^{2i} cells. To quantify Ehmt1 and H3K9me2 enrichment at Zfp281 binding sites, Zfp281 peak regions were extended to 10 kb centered on the peak midpoint using the function resize from the Bioconductor package GenomicRanges. As a control set, DHS-seq peaks (peak annotation downloaded from ENCODE; accession number: ENCSR000CMW) were extended to 10 kb centered on the peak midpoint and only regions non-overlapping with 10-kb extended Zfp281 peaks were considered. The final set of control regions was obtained by randomly sampling one DHS 10-kb peak per Zfp281 10-kb peak, such that the distributions of DHS-seq signal ($\log_2$ RPM) and GC content between Zfp281 10-kb extended peaks and the selected control regions matched closely.

Boxplots were generated using the boxplot function in R with default parameters except outline=FALSE. Genomic profiles for heatmaps centered on the Zfp281 peak midpoint (Fig EV3G) were generated with qProfile from the Bioconductor package QuasR, and visualized using ComplexHeatmap (package version 1.20.0) (Gu *et al*, 2016). Except for Zfp281 ChIP-seq and GRO-seq (Dorighi *et al*, 2017), the averages of two replicates are shown.

Details for quantification and statistical analysis

Details of experiments are specified in figure legends, including the number of biological or technical replicates and the number of datapoints, and are further detailed in the respective Materials and Methods sections. In brief, experimental data quantification is represented as the mean and standard deviation (SD). Significances in Figs 4C, and 5A and D were determined using a Wilcoxon–Mann–Whitney rank sum test using R function wilcox_test. In Fig 2B, the computational analysis shows dots as the median and shades as the lower and upper quartile. In Figs 3C, F, G, 6C, D, and 7A–E, and EV5H–K and M, N, the computational analysis shows the median as middle line, the 95% confidence interval of the median as notches, the interquartile range as box, and 1.5 times the interquartile range as whiskers. In Fig EV5E, the estimated regression coefficients and standard errors are derived from the computational linear model.

Data availability

Next-generation sequencing data reported in this study have been deposited at the Gene Expression Omnibus with accession number GSE131017; http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE131017.

Expanded View for this article is available online.

Acknowledgements

We would like to thank S. Dessus-Babus, K. Jacobeit, E. Pandini Figueiredo Moreno, T. Roloff, and S. Smallwood (FMI) for processing sequencing samples; L. Gelman, S. Bourke, and M. Zinner (FMI) for imaging assistance; H. Kohler (FMI) for cell sorting; M. Flemr and P. Knuckles (FMI) for advice on genome editing and ChIP; and P.A. Ginno, L. Hoerner, and A.H.F.M. Peters (FMI) for providing reagents. We are grateful to S. Gasser, M. Michalski, F. Mohn, and A.H.F.M. Peters (FMI), and S. Stricker (Helmholtz Centre) for comments on the manuscript. This work was supported by the EU Seventh Framework Programme Integrated Project SyBoSS (to A.S. and F.B.) and the Novartis Research Foundation (to J.B.). A.S is a Medical Research Council Professor.

Author contributions

DM, MR, and JB performed experiments. DM and MBS performed bioinformatical analysis. DH performed mass spectrometry. IL performed spheroid image analysis. MW, AS, and FB assisted the esiRNA screen. DM and JB wrote the paper.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Acampora D, Di Giovannantonio LG, Simeone A (2013) Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition. *Development* 140: 43 – 55

Acloque H, Adams MS, Fishwick K, Bronner-Fraser M, Nieto MA (2009) Epithelial-mesenchymal transitions: the importance of changing cell state in development and disease. *J Clin Invest* 119: 1438 – 1449

Amlani B, Liu Y, Chen T, Ee L-S, Lopez P, Heguy A, Apostolou E, Kim SY, Stadtfeld M (2018) Nascent induced pluripotent stem cells efficiently generate entirely iPSC-derived mice while expressing differentiation-associated genes. *Cell Rep* 22: 876 – 884

Bao S, Tang WW, Wu B, Kim S, Li J, Li L, Kobayashi T, Lee C, Chen Y, Wei M *et al* (2018) Derivation of hypermethylated pluripotent embryonic stem cells with high potency. *Cell Res* 28: 22 – 34

Betschinger J, Nichols J, Dietmann S, Corrin PD, Paddison PJ, Smith A (2013) Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. *Cell* 153: 335 – 347

Boroviak T, Loos R, Bertone P, Smith A, Nichols J (2014) The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nat Cell Biol* 16: 516 – 528

Boroviak T, Loos R, Lombard P, Okahara J, Behr R, Sasaki E, Nichols J, Smith A, Bertone P (2015) Lineage-specific profiling delineates the emergence and progression of naive pluripotency in mammalian embryogenesis. *Dev Cell* 35: 366 – 382

Brons IGM, Smithers LE, Trotter MWB, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA *et al* (2007) Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* 448: 191 – 195

Buckley SM, Aranda-Orgilles B, Strikoudis A, Apostolou E, Loizou E, Moran-Crusio K, Farnsworth CL, Koller AA, Dasgupta R, Silva JC *et al* (2012) Regulation of pluripotency and cellular reprogramming by the ubiquitin-proteasome system. *Cell Stem Cell* 11: 783 – 798

Buecker C, Srinivasan R, Wu Z, Calo E, Acampora D, Faial T, Simeone A, Tan M, Swigut T, Wysocka J (2014) Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell Stem Cell* 14: 838 – 853

Cacchiarelli D, Trapnell C, Ziller MJ, Soumillon M, Cesana M, Karnik R, Donaghey J, Smith ZD, Ratanasirintrawoot S, Zhang X *et al* (2015) Integrative analyses of human reprogramming reveal dynamic nature of induced pluripotency. *Cell* 162: 412 – 424

Chantzoura E, Skylaki S, Menendez S, Kim S-I, Johnsson A, Linnarsson S, Woltjen K, Chambers I, Kaji K (2015) Reprogramming roadblocks are system dependent. *Stem Cell Reports* 5: 350 – 364

Collins RE, Northrop JP, Horton JR, Lee DY, Zhang X, Stallcup MR, Cheng X (2008) The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules. *Nat Struct Mol Biol* 15: 245 – 250

Ding L, Paszkowski-Rogacz M, Nitzsche A, Slabicki MM, Heninger A-K, de Vries I, Kittler R, Junqueira M, Shevchenko A, Schulz H *et al* (2009) A genome-scale RNAi screen for Oct4 modulators defines a role of the Paf1 complex for embryonic stem cell identity. *Cell Stem Cell* 4: 403 – 415

Domcke S, Bardet AF, Adrian Ginno P, Hartl D, Burger L, Schübeler D (2015) Competition between DNA methylation and transcription factors determines binding of NRF1. *Nature* 528: 575 – 579

Dorighi KM, Swigut T, Henriques T, Bhanu NV, Scruggs BS, Nady N, Still CD, Garcia BA, Adelman K, Wysocka J (2017) Mll3 and Mll4 facilitate enhancer RNA synthesis and transcription from promoters independently of H3K4 monomethylation. *Mol Cell* 66: 568 – 576.e4

Dunn S-J, Li MA, Carbognin E, Smith A, Martello G (2019) A common molecular logic determines embryonic stem cell self-renewal and reprogramming. *EMBO J* 38: e100003

Ebnet K, Kummer D, Steinbacher T, Singh A, Nakayama M, Matis M (2018) Regulation of cell polarity by cell adhesion receptors. *Semin Cell Dev Biol* 81: 2 – 12

Epsztejn-Litman S, Feldman N, Abu-Remaileh M, Shufaro Y, Gerson A, Ueda J, Deplus R, Fuks F, Shinkai Y, Cedar H *et al* (2008) *De novo* DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat Struct Mol Biol* 15: 1176 – 1183

Estève P-O, Chin HG, Smallwood A, Feehery GR, Gangisetty O, Karpf AR, Carey MF, Pradhan S (2006) Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. *Genes Dev* 20: 3089 – 3103

Factor DC, Corradin O, Zentner GE, Saiakhova A, Song L, Chenoweth JG, McKay RD, Crawford GE, Scacheri PC, Tesar PJ (2014) Epigenomic comparison reveals activation of 'seed' enhancers during transition from naive to primed pluripotency. *Cell Stem Cell* 14: 854 – 863

Festuccia N, Osorno R, Halbritter F, Karwacki-Neisius V, Navarro P, Colby D, Wong F, Yates A, Tomlinson SR, Chambers I (2012) Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. *Cell Stem Cell* 11: 477 – 490

Fidalgo M, Faiola F, Pereira C-F, Ding J, Saunders A, Gingold J, Schaniel C, Lemischka IR, Silva JCR, Wang J (2012) Zfp281 mediates Nanog autorepression through recruitment of the NuRD complex and inhibits somatic cell reprogramming. *Proc Natl Acad Sci USA* 109: 16202 – 16207

Fidalgo M, Huang X, Guallar D, Sanchez-Priego C, Valdes VJ, Saunders A, Ding J, Wu W-S, Clavel C, Wang J (2016) Zfp281 coordinates opposing functions of Tet1 and Tet2 in pluripotent states. *Cell Stem Cell* 19: 355 – 369

Gaidatzis D, Lerch A, Hahne F, Stadler MB (2015) QuasR: quantification and annotation of short reads in R. *Bioinformatics* 31: 1130 – 1132

Gaujoux R, Seoighe C (2010) A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics* 11: 367

Gerber T, Murawala P, Knapp D, Masselink W, Schuez M, Hermann S, Gac-Santel M, Nowoshilow S, Kageyama J, Khattak S *et al* (2018) Single-cell analysis uncovers convergence of cell identities during axolotl limb regeneration. *Science* 362: eaaq0681

Graf T, Enver T (2009) Forcing cells to change lineages. *Nature* 462: 587 – 594

Gu Z, Eils R, Schlesner M (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32: 2847 – 2849

Guo G, Yang J, Nichols J, Hall JS, Eyres I, Mansfield W, Smith A (2009) Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* 136: 1063 – 1069

Guo G, Huang Y, Humphreys P, Wang X, Smith A (2011) A PiggyBac-based recessive screening method to identify pluripotency regulators. *PLoS One* 6: e18189

Hayashi K, Ohta H, Kurimoto K, Aramaki S, Saitou M (2011) Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* 146: 519 – 532

Huang DW, Sherman BT, Lempicki RA (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 37: 1 – 13

Huang X, Balmer S, Yang F, Fidalgo M, Li D, Guallar D, Hadjantonakis A-K, Wang J (2017) Zfp281 is essential for mouse epiblast maturation through transcriptional and epigenetic control of Nodal signaling. *Elife* 6: 243

Ishiuchi T, Ohishi H, Sato T, Kamimura S, Yorino M, Abe S, Suzuki A, Wakayama T, Suyama M, Sasaki H (2019) Zfp281 shapes the transcriptome of trophoblast stem cells and is essential for placental development. *Cell Rep* 27: 1742 – 1754.e6

Iwafuchi-Doi M, Matsuda K, Murakami K, Niwa H, Tesar PJ, Aruga J, Matsuo I, Kondoh H (2012) Transcriptional regulatory networks in epiblast cells and during anterior neural plate development as modeled in epiblast stem cells. *Development* 139: 3926 – 3937

Jackson SA, Olufs ZPG, Tran KA, Zaidan NZ, Sridharan R (2016) Alternative routes to induced pluripotent stem cells revealed by reprogramming of the neural lineage. *Stem Cell Reports* 6: 302 – 311

Kalkan T, Olova N, Roode M, Mulas C, Lee HJ, Nett I, Marks H, Walker R, Stunnenberg HG, Lilley KS *et al* (2017) Tracking the embryonic stem cell transition from ground state pluripotency. *Development* 144: 1221 – 1234

Kojima Y, Kaufman-Francis K, Studdert JB, Steiner KA, Power MD, Loebel DAF, Jones V, Hor A, de Alencastro G, Logan GJ *et al* (2014) The transcriptional and functional properties of mouse epiblast stem cells resemble the anterior primitive streak. *Cell Stem Cell* 14: 107 – 120

Kurimoto K, Yabuta Y, Hayashi K, Ohta H, Kiyonari H, Mitani T, Moritoki Y, Kohri K, Kimura H, Yamamoto T *et al* (2015) Quantitative dynamics of chromatin remodeling during germ cell specification from mouse embryonic stem cells. *Cell Stem Cell* 16: 517 – 532

Ladewig J, Koch P, Brüstle O (2013) Leveling Waddington: the emergence of direct programming and the loss of cell fate hierarchies. *Nat Rev Mol Cell Biol* 14: 225 – 236

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

Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, Morgan MT, Carey VJ (2013) Software for computing and annotating genomic ranges. *PLoS Comput Biol* 9: e1003118

Leeb M, Dietmann S, Paramor M, Niwa H, Smith A (2014) Genetic exploration of the exit from self-renewal using haploid embryonic stem cells. *Cell Stem Cell* 14: 385 – 393

Levine JH, Simonds EF, Bendall SC, Davis KL, Amir E-AD, Tadmor MD, Litvin O, Fienberg HG, Jager A, Zunder ER *et al* (2015) Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. *Cell* 162: 184 – 197

Li R, Liang J, Ni S, Zhou T, Qing X, Li H, He W, Chen J, Li F, Zhuang Q *et al* (2010) A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell* 7: 51 – 63

Li M, Yu JSL, Tilgner K, Ong SH, Koike-Yusa H, Yusa K (2018) Genome-wide CRISPR-KO screen uncovers mTORC1-mediated Gsk3 regulation in naive pluripotency maintenance and dissolution. *Cell Rep* 24: 489 – 502

Luo Z, Gao X, Lin C, Smith ER, Marshall SA, Swanson SK, Florens L, Washburn MP, Shilatifard A (2015) Zic2 is an enhancer-binding factor required for embryonic stem cell specification. *Mol Cell* 57: 685 – 694

Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM (2013) RNA-guided human genome engineering via Cas9. *Science* 339: 823 – 826

Martello G, Sugimoto T, Diamanti E, Joshi A, Hannah R, Ohtsuka S, Göttgens B, Niwa H, Smith A (2012) Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. *Cell Stem Cell* 11: 491 – 504

Merrell AJ, Stanger BZ (2016) Adult cell plasticity *in vivo*: de-differentiation and transdifferentiation are back in style. *Nat Rev Mol Cell Biol* 17: 413 – 425

von Meyenn F, Iurlaro M, Habibi E, Liu NQ, Salehzadeh-Yazdi A, Santos F, Petrini E, Milagre I, Yu M, Xie Z *et al* (2016) Impairment of DNA methylation maintenance is the main cause of global demethylation in naive embryonic stem cells. *Mol Cell* 62: 848 – 861

Mulas C, Kalkan T, Smith A (2017) NODAL secures pluripotency upon embryonic stem cell progression from the ground state. *Stem Cell Reports* 9: 77 – 91

Nefzger CM, Rossello FJ, Chen J, Liu X, Knaupp AS, Firas J, Paynter JM, Pflueger J, Buckberry S, Lim SM *et al* (2017) Cell type of origin dictates the route to pluripotency. *Cell Rep* 21: 2649 – 2660

Niwa H, Ogawa K, Shimosato D, Adachi K (2009) A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature* 460: 118 – 122

Okano M, Bell DW, Haber DA, Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell* 99: 247 – 257

Okita Y, Matsumoto A, Yumimoto K, Isoshita R, Nakayama KI (2012) Increased efficiency in the generation of induced pluripotent stem cells by Fbxw7 ablation. *Genes Cells* 17: 768 – 777

Osorno R, Tsakiridis A, Wong F, Cambray N, Economou C, Wilkie R, Blin G, Scotting PJ, Chambers I, Wilson V (2012) The developmental dismantling of pluripotency is reversed by ectopic Oct4 expression. *Development* 139: 2288 – 2298

Ostapcuk V, Mohn F, Carl SH, Basters A, Hess D, Iesmantavicius V, Lampersberger L, Flemr M, Pandey A, Thomä NH *et al* (2018) Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes. *Nature* 557: 739 – 743

Pereira L, Yi F, Merrill BJ (2006) Repression of Nanog gene transcription by Tcf3 limits embryonic stem cell self-renewal. *Mol Cell Biol* 26: 7479 – 7491

Raab S, Klingenstein M, Möller A, Illing A, Tosic J, Breunig M, Kuales G, Linta L, Seufferlein T, Arnold SJ *et al* (2017) Reprogramming to pluripotency does not require transition through a primitive streak-like state. *Sci Rep* 7: 16543

Rais Y, Zviran A, Geula S, Gafni O, Chomsky E, Viukov S, Mansour AA, Caspi I, Krupalnik V, Zerbib M *et al* (2013) Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* 502: 65 – 70

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43: e47

Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* 11: R25

Rodriguez-Madoz JR, San Jose-Eneriz E, Rabal O, Zapata-Linares N, Miranda E, Rodriguez S, Porciuncula A, Vilas-Zornoza A, Garate L, Segura V *et al* (2017) Reversible dual inhibitor against G9a and DNMT1 improves human iPSC derivation enhancing MET and facilitating transcription factor engagement to the genome. *PLoS One* 12: e0190275

Ruetz T, Pfisterer U, Di Stefano B, Ashmore J, Beniazza M, Tian TV, Kaemena DF, Tosti L, Tan W, Manning JR *et al* (2017) Constitutively active SMAD2/3 are broad-scope potentiators of transcription-factor-mediated cellular reprogramming. *Cell Stem Cell* 21: 791 – 805.e9

Samavarchi-Tehrani P, Golipour A, David L, Sung H-K, Beyer TA, Datti A, Woltjen K, Nagy A, Wrana JL (2010) Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* 7: 64 – 77

Shahbazi MN, Scialdone A, Skorupska N, Weberling A, Recher G, Zhu M, Jedrusik A, Devito LG, Noli L, Macaulay IC *et al* (2017) Pluripotent state transitions coordinate morphogenesis in mouse and human embryos. *Nature* 552: 239 – 243

Shi Y, Desponts C, Do JT, Hahm HS, Schöler HR, 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

Shinkai Y, Tachibana M (2011) H3K9 methyltransferase G9a and the related molecule GLP. *Genes Dev* 25: 781 – 788

Sim Y-J, Kim M-S, Nayfeh A, Yun Y-J, Kim S-J, Park K-T, Kim C-H, Kim K-S (2017) 2i maintains a naive ground state in ESCs through two distinct epigenetic mechanisms. *Stem Cell Reports* 8: 1312 – 1328

Smith A (2017) Formative pluripotency: the executive phase in a developmental continuum. *Development* 144: 365 – 373

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

Stuart HT, Stirparo GG, Lohoff T, Bates LE, Kinoshita M, Lim CY, Sousa EJ, Maskalenka K, Radzisheuskaya A, Malcolm AA *et al* (2019) Distinct molecular trajectories converge to induce naive pluripotency. *Cell Stem Cell* 25: 388 – 406

Tachibana M, Sugimoto K, Nozaki M, Ueda J, Ohta T, Ohki M, Fukuda M, Takeda N, Niida H, Kato H *et al* (2002) G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev* 16: 1779 – 1791

Tachibana M, Ueda J, Fukuda M, Takeda N, Ohta T, Iwanari H, Sakihama T, Kodama T, Hamakubo T, Shinkai Y (2005) Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes Dev* 19: 815 – 826

Tachibana M, Matsumura Y, Fukuda M, Kimura H, Shinkai Y (2008) G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *EMBO J* 27: 2681 – 2690

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

Takahashi K, Yamanaka S (2015) A developmental framework for induced pluripotency. *Development* 142: 3274 – 3285

Tesar PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, Gardner RL, McKay RDG (2007) New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 448: 196 – 199

Treutlein B, Lee QY, Camp JG, Mall M, Koh W, Shariati SAM, Sim S, Neff NF, Skotheim JM, Wernig M *et al* (2016) Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. *Nature* 534: 391 – 395

Tsakiridis A, Huang Y, Blin G, Skylaki S, Wymeersch F, Osorno R, Economou C, Karagianni E, Zhao S, Lowell S *et al* (2014) Distinct Wnt-driven primitive streak-like populations reflect *in vivo* lineage precursors. *Development* 141: 1209 – 1221

Vallier L, Mendjan S, Brown S, Chng Z, Teo A, Smithers LE, Trotter MWB, Cho CH-H, Martinez A, Rugg-Gunn P *et al* (2009) Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. *Development* 136: 1339 – 1349

Villegas F, Lehalle D, Mayer D, Rittirsch M, Stadler MB, Zinner M, Olivieri D, Vabres P, Duplomb-Jego L, De Bont ESJM *et al* (2019) Lysosomal signaling licenses embryonic stem cell differentiation via inactivation of Tfe3. *Cell Stem Cell* 24: 257 – 270

Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, Jaenisch R (2013) One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153: 910 – 918

Westerman BA, Braat AK, Taub N, Potman M, Vissers JHA, Blom M, Verhoeven E, Stoop H, Gillis A, Velds A *et al* (2011) A genome-wide RNAi screen in mouse embryonic stem cells identifies Mp1 as a key mediator of differentiation. *J Exp Med* 208: 2675 – 2689

Wray J, Kalkan T, Gomez-Lopez S, Eckardt D, Cook A, Kemler R, Smith A (2011) Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nat Cell Biol* 13: 838 – 845

Yamane M, Ohtsuka S, Matsuura K, Nakamura A, Niwa H (2018) Overlapping functions of Krüppel-like factor family members: targeting multiple transcription factors to maintain the naïve pluripotency of mouse embryonic stem cells. *Development* 145: dev162404

Yang J, van Oosten AL, Theunissen TW, Guo G, Silva JCR, Smith A (2010) Stat3 activation is limiting for reprogramming to ground state pluripotency. *Cell Stem Cell* 7: 319 – 328

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W *et al* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9: R137

Zhang K, Li L, Huang C, Shen C, Tan F, Xia C, Liu P, Rossant J, Jing N (2010) Distinct functions of BMP4 during different stages of mouse ES cell neural commitment. *Development* 137: 2095 – 2105

Zhang H, Gayen S, Xiong J, Zhou B, Shanmugam AK, Sun Y, Karatas H, Liu L, Rao RC, Wang S *et al* (2016a) MLL1 inhibition reprograms epiblast stem cells to naive pluripotency. *Cell Stem Cell* 18: 481 – 494

Zhang T, Termanis A, Özkan B, Bao XX, Culley J, de Lima Alves F, Rappsilber J, Ramsahoye B, Stancheva I (2016b) G9a/GLP complex maintains imprinted DNA methylation in embryonic stem cells. *Cell Rep* 15: 77 – 85

Zhou H, Morales MG, Hashimoto H, Dickson ME, Song K, Ye W, Kim MS, Niederstrasser H, Wang Z, Chen B *et al* (2017) ZNF281 enhances cardiac reprogramming by modulating cardiac and inflammatory gene expression. *Genes Dev* 31: 1770 – 1783

Zylicz JJ, Dietmann S, Günesdogan U, Hackett JA, Cougot D, Lee C, Surani MA (2015) Chromatin dynamics and the role of G9a in gene regulation and enhancer silencing during early mouse development. *Elife* 4: 717

Expanded View Figures

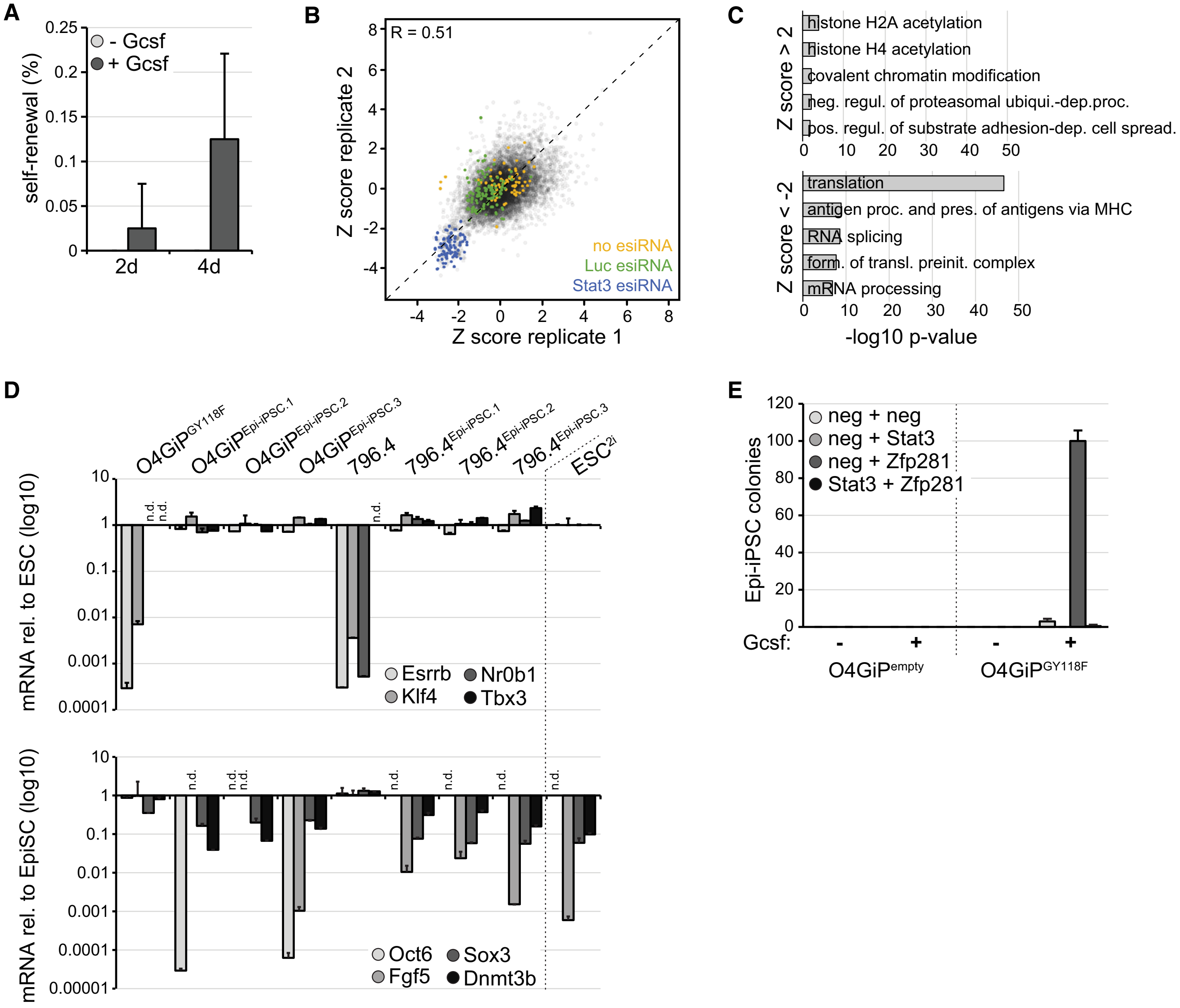


Figure EV1. Enhanced reprogramming of EpiSCs in the absence of Zfp281.

A Self-renewal of O4GIP^{GY118F} reprogramming intermediates after 2 or 4 days in 2i in the presence or absence of Gcsf. Average and SD of 2 experiments performed in duplicates.

B Scatter plot of *Z* scores between screen replicates. Negative controls (no esiRNA and non-targeting Luc esiRNA) are marked in yellow and green, respectively, and positive controls (Stat3 esiRNA) in blue. Pearson's correlation coefficient (R).

C Top 5 GO terms enriched in screen hits with *Z* scores > 2 (top) and < -2 (bottom).

D Induction of naïve (top) and repression of primed (bottom) pluripotency markers in Epi-iPSCs derived from Zfp281-depleted and Gcsf-stimulated O4GIP^{GY118F} and 796.4 EpiSCs. mRNA fold changes relative to ESCs (top) and EpiSCs (bottom) are shown on a log(10)-scaled axis. Average and SD of two technical replicates. Not detected (n.d.).

E Epi-iPSC colonies derived from O4GIP^{empty} and O4GiP^{GY118F} EpiSCs transfected with indicated siRNAs, incubated for 4 days in 2i in the presence or absence of Gcsf, and selected with puromycin. Average and SD of two technical replicates.

Figure EV2. Characterization of Zfp281 and Tet enzymes in ESC differentiation.

A Self-renewal in RGd2 cells at indicated time points of 2i withdrawal. Average and SD of two experiments performed in duplicates.

B Representative flow cytometry profiles of RGd2 ESCs of specified genotypes, at indicated time points and in indicated conditions. Numbers are average and SD of GFP^{high} cells in two experiments.

C, D Self-renewal in RGd2 cells of indicated genotypes after 3 days (C) or indicated time points (D) of 2i withdrawal. Average and SD of two experiments performed in duplicates. E denotes E14 parental cell line origin.

E Flow cytometry profiles (left panel) of long-term differentiated *Zfp281* KO.2 cells in N2B27 and indicating GFP sorting gates (left), and of unsorted or sorted $\text{GFP}^{\text{low,sort}}$ and $\text{GFP}^{\text{high,sort}}$ cells after an additional 2–3 days of culture in N2B27 and indicating gates used for quantification of GFP distribution (right). Please note that profiles shown on the right were recorded on a different instrument than the profile presented on the left. Quantification of GFP distribution (right panel) in N2B27 cultures derived from indicated sorted cells of specified genotypes. Average and SD of 2 experiments.

F *Zfp281* transcription relative to untreated WT^{2i} cells in Zfp281-inducible ESC clones after 48 h in 2i and in the presence or absence of Dox. Average and SD of two technical replicates.

G Self-renewal of sorted $\text{GFP}^{\text{high,sort}}$ and $\text{GFP}^{\text{low,sort}}$ populations of indicated genotypes after exposure to Dox (green) or control conditions (black) for 48 h. Green circles on *x*-axis mark Dox-treated non-self-renewing samples. Average and SD of two technical replicates.

H Representative flow cytometry profiles of RGd2 ESCs of specified genotypes at indicated time points and in indicated conditions. Numbers are average and SD of GFP^{high} cells in two experiments.

I Western blot showing Zfp281 protein levels during ESC progression.

J, K Nanog (J, K) and Zfp281 (K) mRNA levels relative to WT^{2i} cells in ESCs of specified genotypes (J) and in indicated EpiSCs 24 h after transfection of neg and Zfp281 siRNAs (K). Average and SD of 5 (J) and 2 (K) experiments performed in duplicates.

L Resistance to differentiation in O4GIP ESCs transfected with indicated siRNA combinations after 3 days of 2i withdrawal relative to untransfected cells. Average and SD of two experiments performed in duplicates.

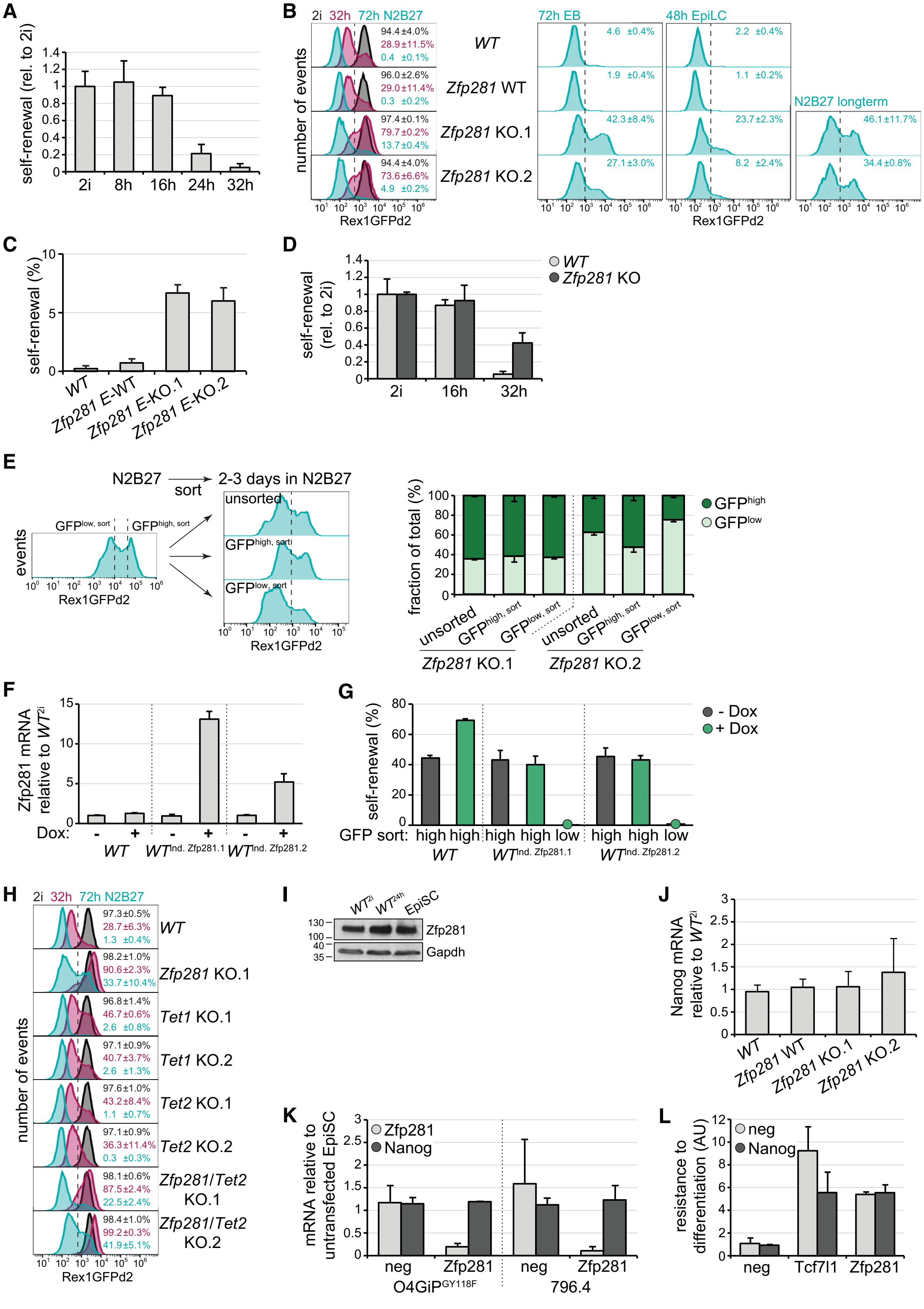

A
self-renewal (rel. to 2i)
1.4
1.2
1
0.8
0.6
0.4
0.2
0
2i
8h
16h
24h
32h
B
2i 32h 72h N2B27
number of events
94.4±4.0%
28.9±11.5%
0.4 ±0.1%
96.0±2.6%
29.0±11.4%
0.3 ±0.2%
97.4±0.1%
79.7±0.2%
13.7±0.4%
94.4±4.0%
73.6±6.6%
4.9 ±0.2%
WT
Zfp281 WT
Zfp281 KO.1
Zfp281 KO.2
72h EB
4.6 ±0.4%
1.9 ±0.4%
42.3±8.4%
27.1±3.0%
48h EpiLC
2.2 ±0.4%
1.1 ±0.2%
23.7±2.3%
8.2 ±2.4%
N2B27 longterm
46.1±11.7%
34.4±0.8%
Rex1GFPd2
C
self-renewal (%)
10
5
0
WT
Zfp281 E-WT
Zfp281 E-KO.1
Zfp281 E-KO.2
D
self-renewal (rel. to 2i)
1.4
1.2
1
0.8
0.6
0.4
0.2
0
2i
16h
32h
WT
Zfp281 KO
E
N2B27
sort
2-3 days in N2B27
events
GFPlow, sort
GFPhigh, sort
Rex1GFPd2
unsorted
GFPhigh, sort
GFPlow, sort
fraction of total (%)
100
80
60
40
20
0
unsorted
GFPhigh, sort
GFPlow, sort
unsorted
GFPhigh, sort
GFPlow, sort
Zfp281 KO.1
Zfp281 KO.2
GFPhigh
GFPlow
F
Zfp281 mRNA relative to WT2i
15
10
5
0
Dox:
-
+
-
+
-
+
WT
WTInd. Zfp281.1
WTInd. Zfp281.2
G
self-renewal (%)
80
60
40
20
0
GFP sort:
high
high
high
high
low
high
high
low
WT
WTInd. Zfp281.1
WTInd. Zfp281.2
- Dox
+ Dox
H
2i 32h 72h N2B27
number of events
97.3±0.5%
28.7±6.3%
1.3 ±0.4%
WT
98.2±1.0%
90.6±2.3%
33.7±10.4%
Zfp281 KO.1
96.8±1.4%
46.7±0.6%
2.6 ±0.8%
Tet1 KO.1
97.1±0.9%
40.7±3.7%
2.6 ±1.3%
Tet1 KO.2
97.6±1.0%
43.2±8.4%
1.1 ±0.7%
Tet2 KO.1
97.1±0.9%
36.3±11.4%
0.3 ±0.3%
Tet2 KO.2
98.1±0.6%
87.5±2.4%
22.5±2.4%
Zfp281/Tet2 KO.1
98.4±1.0%
99.2±0.3%
41.9±5.1%
Zfp281/Tet2 KO.2
Rex1GFPd2
I
WT2i
WT24h
EpiSC
130
100
40
35
Zfp281
Gapdh
J
Nanog mRNA relative to WT2i
3
2.5
2
1.5
1
0.5
0
WT
Zfp281 WT
Zfp281 KO.1
Zfp281 KO.2
K
mRNA relative to untransfected EpiSC
3
2.5
2
1.5
1
0.5
0
Zfp281
Nanog
neg
Zfp281
neg
Zfp281
O4GiPGY118F
796.4
L
resistance to differentiation (AU)
12
10
8
6
4
2
0
neg
Nanog
neg
Tcf7l1
Zfp281

Figure EV2.

Figure EV3. Genomics of Zfp281.

A mRNA $\log_2$FC relative to WT^{2i} cells of selected core, naïve, and primed pluripotency markers in differentiating between *WT* and *Zfp281* KO cells at indicated time points.

B Top 5 enriched GO terms in clusters 1–6.

C Sequence logo from the *de novo* identified binding motif in 82.4% of 23,756 Zfp281 peaks.

D Scatter plot comparing $\log_2$ Zfp281 peak ChIP enrichment over matched inputs between replicates in WT^{2i} (top) and $WT^{32\ h}$ (bottom) cells.

E, F Scatter plot comparing $\log_2$ Zfp281 ChIP enrichment over matched inputs in $WT^{32\ h}$ cells and EpiSCs (Huang *et al*, 2017) (E) and TSCs (Ishiuchi *et al*, 2019) (F). Peaks were assigned to closest transcriptional start sites (TSSs) and colored according to association with gene clusters 1–6.

G Heatmap of Zfp281, H3K27ac, histone H3K4 tri-methylation (H3K4me3), H3K4 mono-methylation (H3K4me1), and histone H3K27 tri-methylation (H3K27me3) ChIP-seq, DNase-hypersensitive sites (DHS), and global run-on sequencing (GRO-seq) read densities across all proximal (± 2 kb of TSS) Zfp281 (top) and distal Zfp281 (bottom) peaks. Each row represents a 10-kb window centered on the peak mid of Zfp281. Rows are sorted for H3K27ac ChIP read densities in ESCs. Reads per million (RPM).

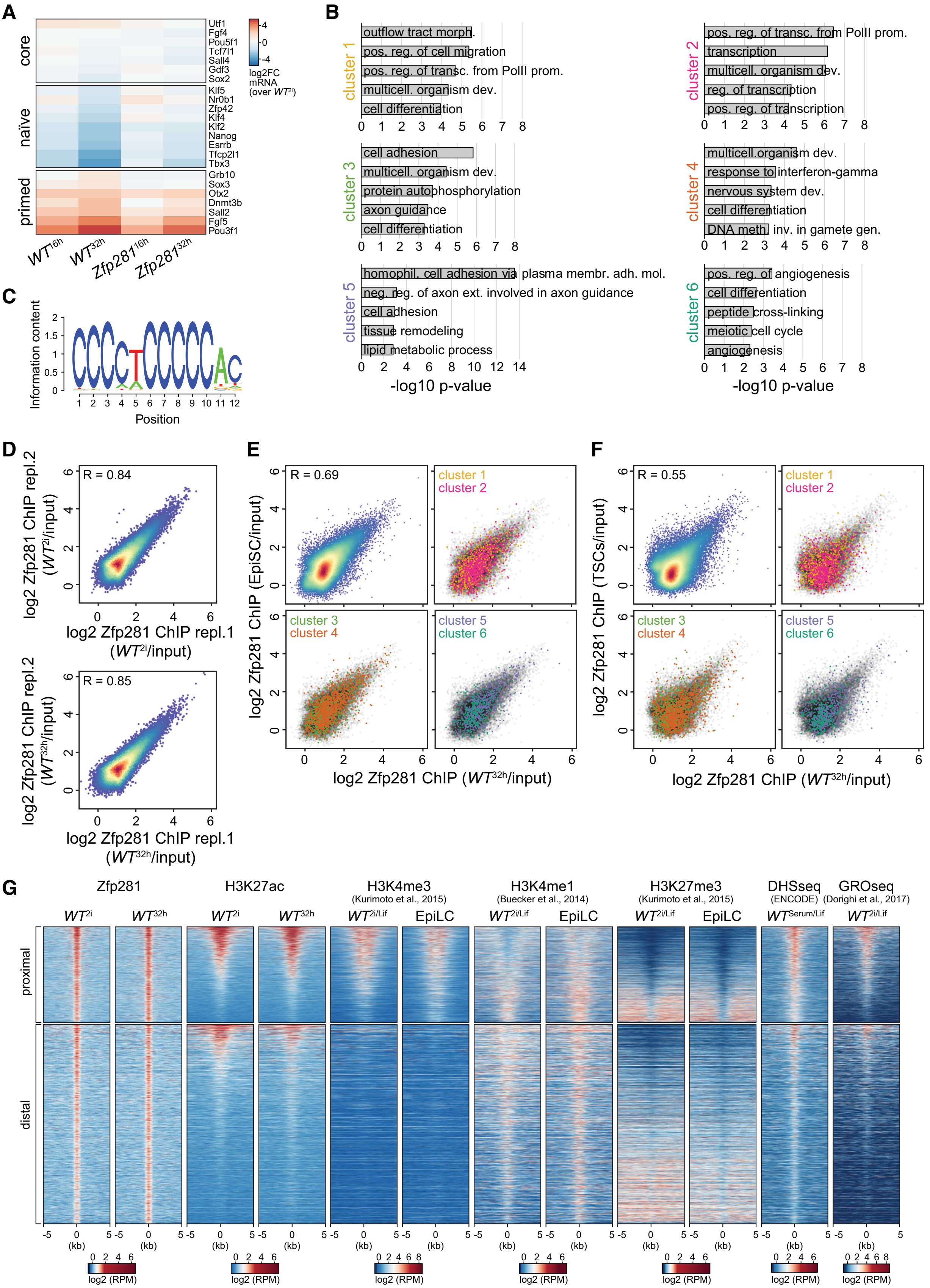

A
core
naive
primed
Utf1
Fgf4
Pou5f1
Tcf7l1
Sall4
Gdf3
Sox2
Klf5
Nr0b1
Zfp42
Klf4
Klf2
Nanog
Esrrb
Tfcp2l1
Tbx3
Grb10
Sox3
Otx2
Dnmt3b
Sall2
Fgf5
Pou3f1
4
0
-4
log2FC mRNA (over WT2i)
WT16h
WT32h
Zfp28116h
Zfp28132h
B
cluster 1
outflow tract morph.
pos. reg. of cell migration
pos. reg. of transc. from PolII prom.
multicell. organism dev.
cell differentiation
0 1 2 3 4 5 6 7 8
cluster 2
pos. reg. of transc. from PolII prom.
transcription
multicell. organism dev.
reg. of transcription
pos. reg. of transcription
0 1 2 3 4 5 6 7 8
cluster 3
cell adhesion
multicell. organism dev.
protein autophosphorylation
axon guidance
cell differentiation
0 1 2 3 4 5 6 7 8
cluster 4
multicell.organism dev.
response to interferon-gamma
nervous system dev.
cell differentiation
DNA meth. inv. in gamete gen.
0 1 2 3 4 5 6 7 8
cluster 5
homophil. cell adhesion via plasma membr. adh. mol.
neg. reg. of axon ext. involved in axon guidance
cell adhesion
tissue remodeling
lipid metabolic process
0 2 4 6 8 10 12 14
-log10 p-value
cluster 6
pos. reg. of angiogenesis
cell differentiation
peptide cross-linking
meiotic cell cycle
angiogenesis
0 1 2 3 4 5 6 7 8
-log10 p-value
C
Information content
2
1.5
1
0.5
0
1 2 3 4 5 6 7 8 9 10 11 12
Position
D
R = 0.84
log2 Zfp281 ChIP repl.2 (WT2i/input)
log2 Zfp281 ChIP repl.1 (WT2i/input)
R = 0.85
log2 Zfp281 ChIP repl.2 (WT32h/input)
log2 Zfp281 ChIP repl.1 (WT32h/input)
0 2 4 6
E
R = 0.69
cluster 1
cluster 2
cluster 3
cluster 4
cluster 5
cluster 6
log2 Zfp281 ChIP (EpiSC/input)
log2 Zfp281 ChIP (WT32h/input)
F
R = 0.55
cluster 1
cluster 2
cluster 3
cluster 4
cluster 5
cluster 6
log2 Zfp281 ChIP (TSCs/input)
log2 Zfp281 ChIP (WT32h/input)
G
Zfp281
H3K27ac
H3K4me3 (Kurimoto et al., 2015)
H3K4me1 (Buecker et al., 2014)
H3K27me3 (Kurimoto et al., 2015)
DHSseq (ENCODE)
GROseq (Dorighi et al., 2017)
WT2i
WT32h
WT2i
WT32h
WT2i/Lif
EpiLC
WT2i/Lif
EpiLC
WT2i/Lif
EpiLC
WTSerum/Lif
WT2i/Lif
proximal
distal
-5 0 5
(kb)
0 2 4 6
0 2 4 6 8
log2 (RPM)

Figure EV3.

Figure EV4. Characterization of Ehmt1 and Zic2 downstream of Zfp281.

A Representative flow cytometry profiles of Zfp281-inducible RGd2 ESCs transfected with indicated siRNAs after 32 h in 2i and in the presence (green) or absence (black) of Dox.

B, C FC Ehmt1 (B) and Zic2 (C) transcription relative to untreated *WT*^{2i} cells in indicated ESC clones after 32 h in 2i in the presence or absence of Dox. Average and SD of two technical replicates.

D Representative flow cytometry profiles of RGd2 ESCs with conditional Ehmt1 and Zic2 expression after 32 h in 2i and in the presence (green) or absence (black) of Dox. Numbers are the average and SD of GFP^{low} cells in two experiments.

E Zfp281 transcription relative to untreated *WT*^{2i} cells in indicated ESC clones after 32 h in 2i in the presence or absence of Dox. Average and SD of two technical replicates.

F Representative flow cytometry profiles of *Dnmt3a/3b* compound KO RGd2 ESCs with conditional Zfp281 expression after 32 h in 2i and in the presence (green) or absence (black) of Dox. Significance was determined using a Wilcoxon–Mann–Whitney rank sum test compared to *WT*$^{ind.\ Zfp281.1}$ Dox-treated cells. Not significant (n.s.) ≥ 0.05. Numbers are the average and SD of GFP^{low} cells in four experiments.

G Whole-cell lysate Zfp281 IPs in *WT* and *Zfp281* KO cells in 2i or 40 h after 2i withdrawal, and probed for indicated proteins. Input (left) and Zfp281 IP (right). (*) Ig heavy chain.

H–J Cell morphologies (H), growth curves (I), and cell cycle analyses using propidium iodide staining (J) of indicated genotypes in 2i. Average and SD of three experiments (I, J). Scale bar is 50 μm (H).

K Representative flow cytometry profiles of indicated genotypes in 2i, and after 32 and 72 h of 2i withdrawal. Numbers are the average and SD of GFP^{high} cells in two experiments.

L Quantification and hierarchical clustering of normalized F-actin intensity in 20 concentric rings (from center to circumference) in spheroids derived from ESCs with indicated genotypes in 2i or N2B27 for 4 days. Intensity is color-coded and illustrates central F-actin accumulation and, hence, polarization of *WT* and *Zic2* KO cells during differentiation.

M Representative immunofluorescence staining of *WT* or *Ehmt1* KO ESCs expressing the indicated transgenes. Top: H3K9me2 and DAPI. Bottom: Ehmt1. Co-localization of H3K9me2 with DAPI-rich speckles in *Ehmt1*^{2i} cells expressing no transgene, the ΔNHHC, or NH-LE alleles is indicated by arrowheads. Please note the absence of nuclear Ehmt1 staining in *Ehmt1*^{2i} cells and restoration by Ehmt1 transgenes. Scale bar is 10 μm.

N Representative flow cytometry profiles of indicated genotypes in 2i and 32 h after 2i withdrawal.

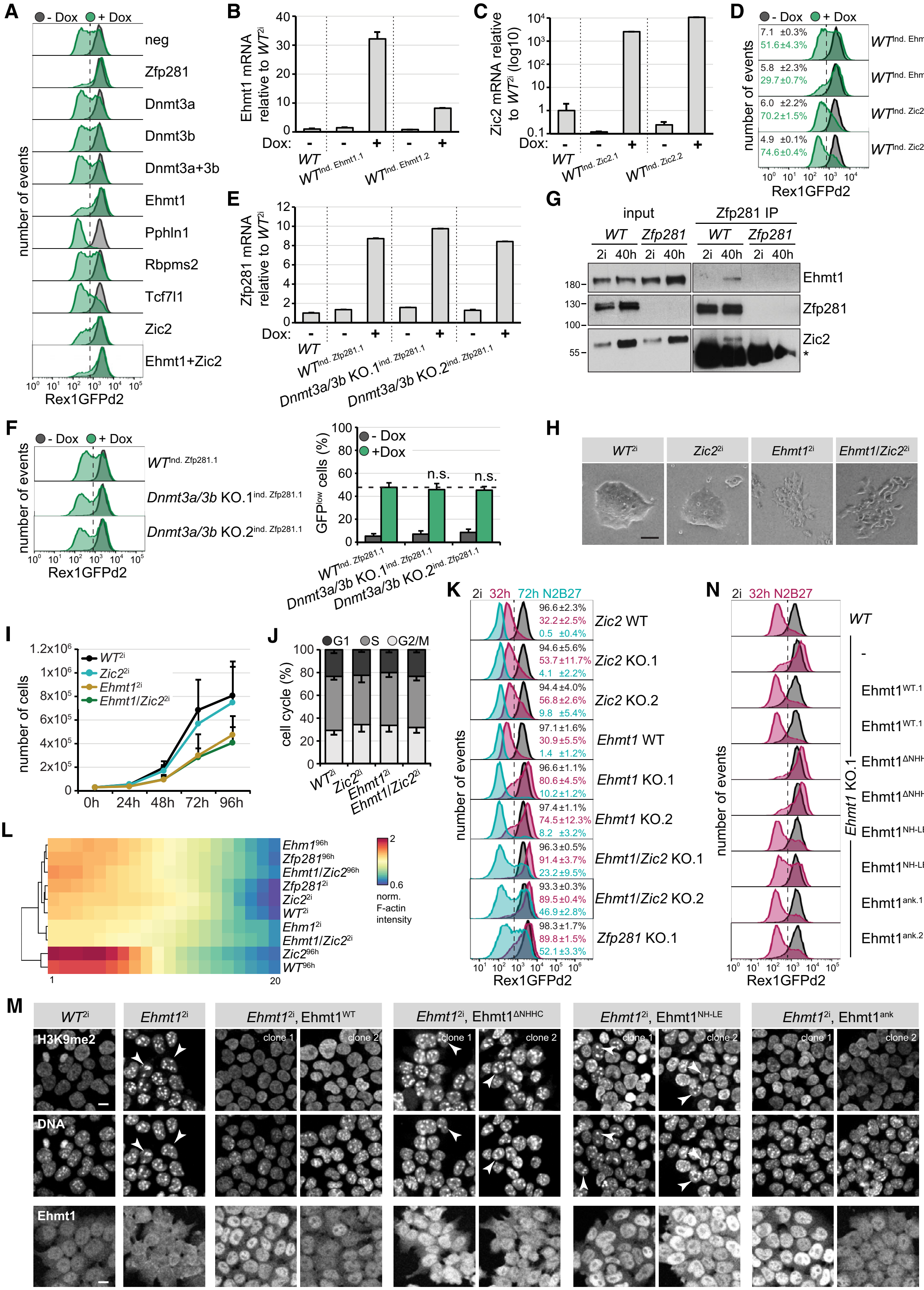

A
- Dox + Dox
neg
Zfp281
Dnmt3a
Dnmt3b
Dnmt3a+3b
Ehmt1
Pphln1
Rbpms2
Tcf7l1
Zic2
Ehmt1+Zic2
number of events
Rex1GFPd2
B
Ehmt1 mRNA relative to WT2i
Dox:
WT
WTind. Ehmt1.1
WTind. Ehmt1.2
C
Zic2 mRNA relative to WT2i (log10)
WTind. Zic2.1
WTind. Zic2.2
D
7.1 ±0.3%
51.6±4.3%
5.8 ±2.3%
29.7±0.7%
6.0 ±2.2%
70.2±1.5%
4.9 ±0.1%
74.6±0.4%
WTInd. Ehmt1.1
WTInd. Ehmt1.2
WTInd. Zic2.1
WTInd. Zic2.2
E
Zfp281 mRNA relative to WT2i
WTind. Zfp281.1
Dnmt3a/3b KO.1ind. Zfp281.1
Dnmt3a/3b KO.2ind. Zfp281.1
G
input
Zfp281 IP
WT
Zfp281
2i
40h
180
130
100
55
Ehmt1
Zfp281
Zic2
*
F
GFPlow cells (%)
n.s.
n.s.
H
WT2i
Zic22i
Ehmt12i
Ehmt1/Zic22i
I
number of cells
WT2i
Zic22i
Ehmt12i
Ehmt1/Zic22i
0h
24h
48h
72h
96h
J
G1
S
G2/M
cell cycle (%)
K
2i 32h 72h N2B27
96.6±2.3%
32.2±2.5%
0.5 ±0.4%
Zic2 WT
94.6±5.6%
53.7±11.7%
4.1 ±2.2%
Zic2 KO.1
94.4±4.0%
56.8±2.6%
9.8 ±5.4%
Zic2 KO.2
97.1±1.6%
30.9±5.5%
1.4 ±1.2%
Ehmt1 WT
96.6±1.1%
80.6±4.5%
10.2±1.2%
Ehmt1 KO.1
97.4±1.1%
74.5±12.3%
8.2 ±3.2%
Ehmt1 KO.2
96.3±0.5%
91.4±3.7%
23.2±9.5%
Ehmt1/Zic2 KO.1
93.3±0.3%
89.5±0.4%
46.9±2.8%
Ehmt1/Zic2 KO.2
98.3±1.7%
89.8±1.5%
52.1±3.3%
Zfp281 KO.1
L
Ehm196h
Zfp28196h
Ehmt1/Zic296h
Zfp2812i
Zic22i
WT2i
Ehm12i
Ehmt1/Zic22i
Zic296h
WT96h
norm. F-actin intensity
N
2i 32h N2B27
WT
-
Ehmt1WT.1
Ehmt1WT.1
Ehmt1ΔNHHC.1
Ehmt1ΔNHHC.2
Ehmt1NH-LE.1
Ehmt1NH-LE.2
Ehmt1ank.1
Ehmt1ank.2
Ehmt1 KO.1
M
WT2i
Ehmt12i
Ehmt12i, Ehmt1WT
Ehmt12i, Ehmt1ΔNHHC
Ehmt12i, Ehmt1NH-LE
Ehmt12i, Ehmt1ank
clone 1
clone 2
H3K9me2
DNA
Ehmt1

Figure EV4.

▶

Figure EV5. Ehmt1 and Zic2 transcriptomics and genomics.

A Log$_2$-normalized read counts of selected core, naïve, and primed pluripotency markers in ESCs of indicated genotypes.

B Scatter plot of mRNA log$_2$FC in *Zfp281*^{2i} and *Ehmt1*^{2i} cells (left). Top 5 GO terms enriched in genes upregulated (red) and downregulated (blue) specifically in *Ehmt1*^{2i} cells (right).

C, D Pairwise Pearson correlation coefficients of mRNA changes between indicated differentiated cells considering all detected transcripts (C) or gene cluster 1–6 transcripts (D).

E Estimated regression coefficients and standard errors for the contribution of *Ehmt1* (Δ*Ehmt1*), *Zic2* (Δ*Zic2*), and their interaction (Δ*Ehmt1:Zic2*) to cell state-specific gene expression changes in *Ehmt1*/*Zic2* compound KO cells. All expressed genes (13,096) are taken into account.

F Western blot confirming Ehmt1 biotinylation [probed with Streptavidin (Strep)] in ESCs of indicated genotypes expressing the BirA ligase.

G ESC self-renewal of indicated genotypes after 3 days of 2i withdrawal. Average and SD of three experiments performed in duplicates.

H Log$_2$ Ehmt1 and H3K9me2 ChIP enrichment in ESCs over matched inputs at five classes of 10-kb genome-wide windows binned by increasing Ehmt1 chromatin association. Boxes as in Fig 3C for 42,887 datapoints each.

I, J Ehmt1 (I, J) and H3K9me2 (J) ChIP log$_2$FC between indicated cell states and genotypes at Zfp281 peaks (purple) or matching and non-overlapping DHS control peaks (gray) extended to 10-kb windows. Boxes and number of datapoints as in Fig 7A.

K Representative immunofluorescence staining of H3K9me2 (left) and quantification relative to DNA (right) in indicated genotypes and conditions. Scale bar is 10 μm. Boxes as in Fig 3C for 453 (*WT*^{2}), 465 (*WT*$^{32\ h}$), 574 (*Zfp281*^{2}), and 792 (*Zfp281*$^{32\ h}$) datapoints.

L Density plot showing distance of Zfp281-only (pink), Zic2-only (blue), and Zfp281/Zic2 co-bound peaks (yellow) to nearest TSS.

M Zfp281 (left), Zic2 (middle), and H3K27ac (right) log$_2$ ChIP enrichment over matched inputs in ESCs at Zfp281-only (pink), Zic2-only (blue), and Zfp281/Zic2 co-bound (yellow) peaks. Boxes and number of datapoints as in Fig 7B.

N Cell state-specific Zic2 ChIP log$_2$FC between indicated genotypes and cell states at Zfp281-only (pink), Zic2-only (blue), and Zfp281/Zic2 co-bound (yellow) peaks. Boxes and number of datapoints as in Fig 7B.

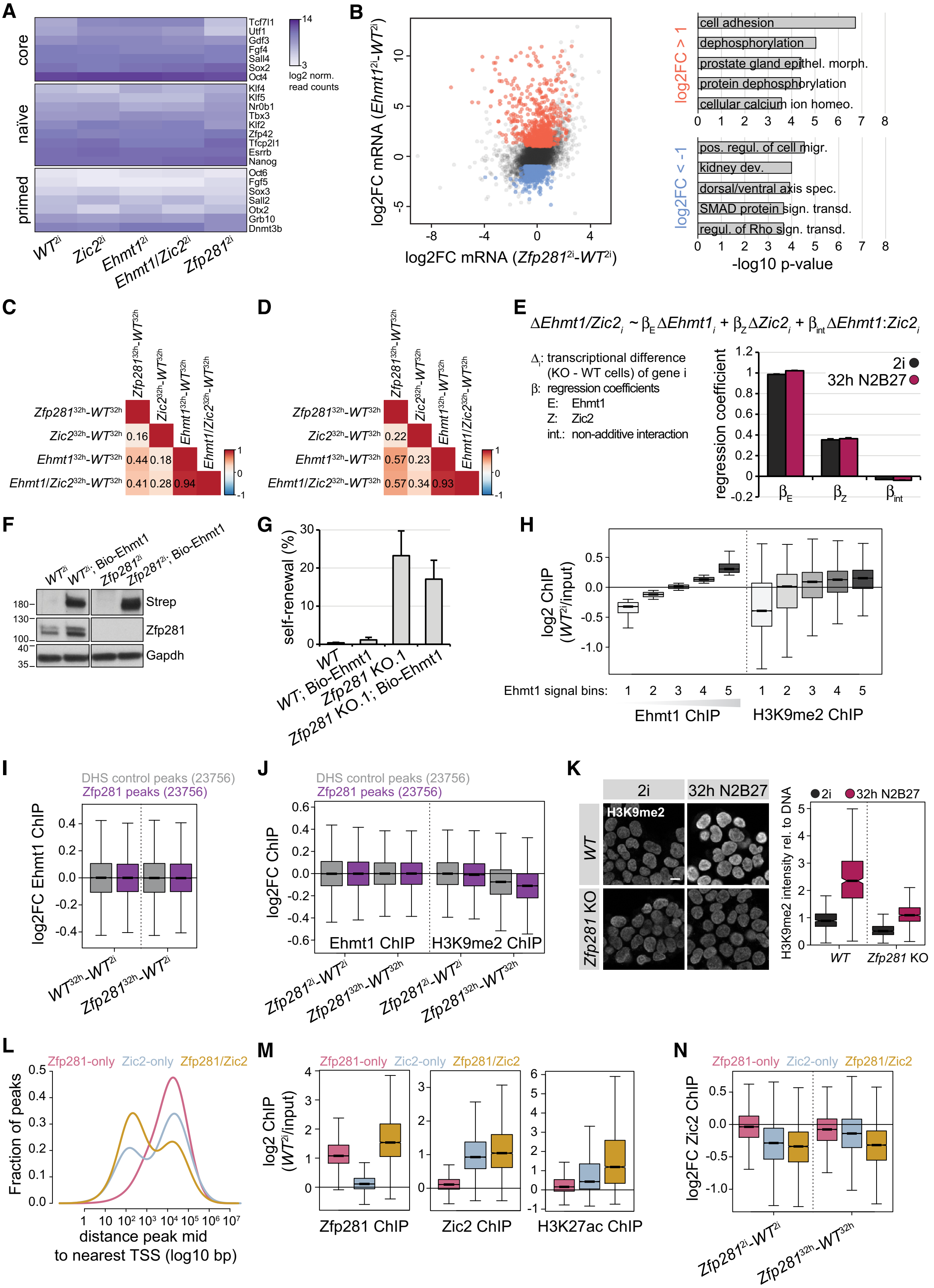

A
core
naïve
primed
Tcf7l1
Utf1
Gdf3
Fgf4
Sall4
Sox2
Oct4
Klf4
Klf5
Nr0b1
Tbx3
Klf2
Zfp42
Tfcp2l1
Esrrb
Nanog
Oct6
Fgf5
Sox3
Sall2
Otx2
Grb10
Dnmt3b
14
3
log2 norm. read counts
WT2i
Zic22i
Ehmt12i
Ehmt1/Zic22i
Zfp2812i
B
log2FC mRNA (Ehmt12i-WT2i)
log2FC mRNA (Zfp2812i-WT2i)
-8
-4
0
4
-5
0
5
10
log2FC > 1
cell adhesion
dephosphorylation
prostate gland epithel. morph.
protein dephosphorylation
cellular calcium ion homeo.
log2FC < -1
pos. regul. of cell migr.
kidney dev.
dorsal/ventral axis spec.
SMAD protein sign. transd.
regul. of Rho sign. transd.
0 1 2 3 4 5 6 7 8
-log10 p-value
C
Zfp28132h-WT32h
Zic232h-WT32h
Ehmt132h-WT32h
Ehmt1/Zic232h-WT32h
0.16
0.44
0.18
0.41
0.28
0.94
1
0
-1
D
Zfp28132h-WT32h
Zic232h-WT32h
Ehmt132h-WT32h
Ehmt1/Zic232h-WT32h
0.22
0.57
0.23
0.57
0.34
0.93
1
0
-1
E
ΔEhmt1/Zic2i ~ βE ΔEhmt1i + βZ ΔZic2i + βint ΔEhmt1:Zic2i
Δi: transcriptional difference (KO - WT cells) of gene i
β: regression coefficients
E: Ehmt1
Z: Zic2
int.: non-additive interaction
regression coefficient
2i
32h N2B27
βE
βZ
βint
-0.2
0
0.2
0.4
0.6
0.8
1
1.2
F
WT2i
WT2i; Bio-Ehmt1
Zfp2812i
Zfp2812i; Bio-Ehmt1
180
130
100
40
35
Strep
Zfp281
Gapdh
G
self-renewal (%)
0
10
20
30
WT
WT; Bio-Ehmt1
Zfp281 KO.1
Zfp281 KO.1; Bio-Ehmt1
H
log2 ChIP (WT2i/input)
0.5
0.0
-0.5
-1.0
Ehmt1 signal bins:
1 2 3 4 5
1 2 3 4 5
Ehmt1 ChIP
H3K9me2 ChIP
I
DHS control peaks (23756)
Zfp281 peaks (23756)
log2FC Ehmt1 ChIP
0.4
0.2
0.0
-0.2
-0.4
WT32h-WT2i
Zfp28132h-WT2i
J
DHS control peaks (23756)
Zfp281 peaks (23756)
log2FC ChIP
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
Ehmt1 ChIP
H3K9me2 ChIP
Zfp2812i-WT2i
Zfp28132h-WT32h
Zfp2812i-WT2i
Zfp28132h-WT32h
K
2i
32h N2B27
H3K9me2
WT
Zfp281 KO
2i
32h N2B27
H3K9me2 intensity rel. to DNA
0
1
2
3
4
5
WT
Zfp281 KO
L
Zfp281-only
Zic2-only
Zfp281/Zic2
Fraction of peaks
0.5
0.4
0.3
0.2
0.1
0.0
1 10 10² 10³ 10⁴ 10⁵ 10⁶ 10⁷
distance peak mid to nearest TSS (log10 bp)
M
Zfp281-only
Zic2-only
Zfp281/Zic2
log2 ChIP (WT2i/input)
Zfp281 ChIP
Zic2 ChIP
H3K27ac ChIP
N
Zfp281-only
Zic2-only
Zfp281/Zic2
log2FC Zic2 ChIP
0.5
0.0
-0.5
-1.0
Zfp2812i-WT2i
Zfp28132h-WT32h

Figure EV5.

APPENDIX

Zfp281 orchestrates interconversion of pluripotent states by engaging Ehmt1 and Zic2

Daniela Mayer, Michael B. Stadler, Melanie Rittirsch, Daniel Hess, Ilya Lukonin, Maria Winzi, Austin Smith, Frank Buchholz, Joerg Betschinger

TABLE OF CONTENTS

Appendix Figure S1: Deconvolution of siRNA pools (related to Figures 1 and 5).

(**A**, **B**) Epi-iPSC colonies derived from 796.4 EpiSCs transfected with indicated siRNAs (individual siRNAs or pools), stimulated for 4d with Gcsf and 2i, and selected with Puromycin. Average and SD of 3 experiments performed in duplicates.

Appendix Figure S2: Details of genome-edited ESC lines (related to Figures 2 and 5, and EV4).

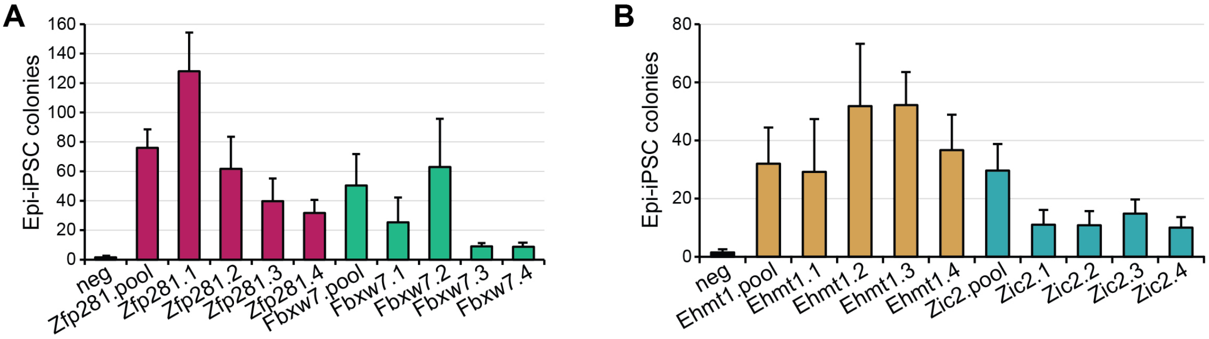
(**A**, **B**) Sequence of genome-edited *Zfp281* locus (**A**) and absence of protein (**B**) in KO cells. E denotes E14 parental cell line origin.

(**C**, **D**) Sequence of genome-edited *Tet1* and *Tet2* loci (**C**) and absence of proteins (**D**) in KO cells.

(**E**) Sequences of genome-edited *Dnmt3a* and *Dnmt3b* loci in $WT^{\text{ind. Zfp281.1}}$ ESCs.

(**F**, **G**) Sequence of genome-edited *Ehmt1* and *Zic2* loci (**F**) and absence of proteins (**G**) in KO cells.

Mayer et al., Appendix Figure S1



Mayer et al., Appendix Figure S2