REGULATION OF CELL FATE CHOICE IN THE MOUSE BLASTOCYST STAGE EMBRYO

Dissertation zur Erlangung des Doktorgrades
der Fakultät für Biologie
der Ludwig-Maximilians-Universität München

vorgelegt von

NADINE SCHRODE

September 2015

Diese Dissertation wurde im Zeitraum von Februar 2012 bis August 2015 am Sloan-Kettering Institute New York unter der Leitung von Prof. Dr. Anna-Katerina Hadjantonakis angefertigt. Vertreter der Arbeit vor der Fakultät für Biologie an der Ludwig-Maximilians-Universität München war Prof. Dr. Heinrich Leonhardt.

Erstgutachter: Prof. Dr. Heinrich Leonhardt

Zweitgutachter: Prof. Dr. Charles David

Tag der Abgabe: 08.09.2015

Tag der mündlichen Prüfung: 26.11.2015

CONTENT

LIST OF PUBLICATIONS

PUBLICATION I

Schrode N., Saiz N., Di Talia S., Hadjantonakis A.-K.

GATA6 levels modulate primitive endoderm cell fate choice and timing in the mouse blastocyst

Developmental Cell, 2014

PUBLICATION II

Niakan K.K., Schrode N., Cho L.T., Hadjantonakis A.K.

Derivation of extraembryonic endoderm stem (XEN) cells from mouse embryos and embryonic stem cells

Nature Protocols, 2013

PUBLICATION III

Czechanski A., Byers C., Greenstein I., Schrode N., Donahue L.R., Hadjantonakis A.K., Reinholdt L.G.

Derivation and characterization of mouse embryonic stem cells from permissive and nonpermissive strains

Nature Protocols, 2014

PUBLICATION IV

Kumar R., DiMenna L., Schrode N., Liu T.C., Franck P., Muñoz-Descalzo S., Hadjantonakis A.K., Zarrin A.A., Chaudhuri J., Elemento O., Evans T.

AID stabilizes stem-cell phenotype by removing epigenetic memory of pluripotency genes

Nature, 2013

ABBREVIATIONS

2i	2 inhibitor
3i	3 inhibitor
5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
AFP	Alpha fetoprotein
AID	Activation-induced cytidine deaminase
AKT	Thymoma viral proto-oncogene
AMOT	Angiomotin
aPKC	Atypical protein kinase C
APOBEC1	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1
BAC	Bacterial artificial chromosome
BER	Base excision repair
BMP	Bone morphogenetic protein
C. elegans	Caenorhabditis elegans
CDX2	Caudal type homeobox 2
ChIP-Seq	Chromatin immunoprecipitation - sequencing
cMYC	Myelocytomatosis oncogene
CSHL	Cold Spring Harbor Laboratories
DAB2	Disabled 2, mitogen-responsive phosphoprotein
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase (cytosine-5)
E	Embryonic day
ELF5	E74-like factor 5
EOMES	Eomesodermin
EPC	Ectoplacental cone
EPI	Epiblast
ERK	Mitogen-activated protein kinase 1/2 (MAPK)

ES cell	Embryonic stem cell
ESRRB	Estrogen related receptor, beta
ExE	Extraembryonic ectoderm
EZR	Ezrin
FACS	Fluorescence activated cell sorting
FBS	Fetal calf serum
FGF	Fibroblast growth factor
FGFR	Fibroblast growth factor receptor
FOXA2	Forkhead box A2
FRS2a	Fibroblast growth factor receptor substrate 2
GATA	GATA binding protein
GFP	Green fluorescent protein
GRB2	Growth factor receptor-bound protein 2
GSK3	Glycogen synthase kinase-3
H2B	Histone 2B
HMG	High mobility group
ICM	Inner cell mass
iPS cell	induced pluripotent stem cell
JAK	Janus kinase
KLF4	Kruppel-like factor 4
LATS1/2	Large tumor suppressor kinases 1 and 2
LIF	Leukemia inhibitory factor
MEF	Mouse embryonic fibroblast
MINS	Modular interactive nuclear segmentation
mRNA	messenger ribonucleic acid
NF2	Neurofibromin 2
OCT4	Octamer binding transcription factor 4 (= POU5F1)
PAR	Partitioning defective homolog
PCR	Polymerase chain reaction
PDGFRA	Platelet derived growth factor receptor, alpha polypeptide
PE	Parietal endoderm

PI3K	Phosphatidylinositol 3-kinase
PLC	Phosphoinositide phospholipase C
POU5F1	POU domain, class 5, transcription factor (= OCT4)
PrE	Primitive endoderm
RAS	Rat sarcoma virus oncogene
RT	Reverse transcription
SHP2	Src homology region 2 domain containing phosphatase 2
SMAD	Portmanteau of Mothers against decapentaplegic (MAD) and Smaller (SMA)
SMAD2,3	SMAD family member 2,3
SOS	Son of sevenless
SOX2	SRY (sex determining region Y)-box
SRY	Sex determining region of Chr Y
STAT	Signal transducer and activator of transcription
TE	Trophectoderm
TEAD4	TEA domain family member 4
TET	Ten-eleven translocation protein
TGFb	Transforming growth factor beta
TS cell	Trophoblast stem cell
VE	Visceral endoderm
WNT	Wingless-type MMTV integration site family, member 3A
XEN cell	extraembryonic endoderm stem cell
YAP1	Yes-associated gene 1

SUMMARY

During the second cell fate choice in the mammalian embryo, two lineages are specified, the extraembryonic primitive endoderm (PrE), which will give rise to the yolk sac, and the pluripotent epiblast (EPI), which will go on to form the embryo proper. This lineage choice takes place in the inner cell mass (ICM) of the blastocyst stage embryo. Cells of the ICM are initially totipotent and express markers for both, PrE and EPI, but gradually downregulate one or the other until their expression becomes mutually exclusive and the two lineages are committed. In this work computational analysis of cells in the ICM showed that the distribution of emerging EPI and PrE cells is stochastic and independent from each other.

The earliest markers for which the gradual segregation of the two lineages can be observed are the transcription factors NANOG, which becomes EPI-specific, and GATA6, which marks the PrE. NANOG is known to be involved in the maintenance of pluripotency in embryonic stem cells and in its absence the embryo fails to form the EPI lineage. The present work demonstrated that GATA6 is the master regulator of the PrE lineage. Embryos lacking GATA6 did not upregulate any successive PrE markers and failed to specify the PrE. Instead, all cells of the ICM expressed NANOG. In wildtype context the downregulation of NANOG in PrE-biased cells is mediated by FGF signaling. This process was revealed to be GATA6 dependent, thereby shedding light on a novel element of the network governing the second cell fate decision. GATA6 was also shown to act in a dosage dependent manner, with *Gata6* heterozygous embryos exhibiting not only reduced numbers of PrE cells, but also slowed lineage commitment, suggesting relative levels of NANOG and GATA6 may govern the segregation of EPI and PrE.

Stem cells can be derived from both lineages, with embryonic stem (ES) cells representing the EPI and extraembryonic endoderm stem (XEN) cells the PrE. This work demonstrated efficient protocols for the derivation of

either cell type, thereby facilitating the study of the properties of the two lineages.

Beyond the involvement of transcription factors and signaling pathways, epigenetic gene regulation is thought to be involved in the maintenance and loss of pluripotency. DNA methylation and demethylation are such mechanisms. While DNA methylation is important for the silencing of genes e.g. during differentiation, DNA demethylation takes place in the preimplantation embryo to remove such repressive marks. In this work, ES cells were derived from blastocysts lacking the DNA demethylation promoting deaminase AID. These ES cells exhibited lower derivation efficiencies, pointing to the possible involvement of DNA demethylation for the establishment of the pluripotent epiblast lineage. Furthermore, fibroblasts lacking AID were reprogrammed and it was revealed that the resulting induced pluripotent stem (iPS) cells failed to maintain their pluripotent state.

In conclusion, the present work provides insights into the establishment of pluripotency in vivo and in vitro by revealing new mechanisms for the lineage choice between EPI and PrE as well as the potential involvement of epigenetic regulation.

ZUSAMMENFASSUNG

Während der zweiten Zellschicksalsentscheidung im Säugetierembryo werden zwei Zelllinien gebildet, das extraembryonische primitive Endoderm (PrE), das sich später zum Dottersack entwickelt, und der pluripotente Epiblast (EPI), aus dem der Embryo hervorgeht. Diese Entscheidung findet in der inneren Zellmasse (ICM) der Blastozyste statt. Die Zellen der ICM sind zunächst totipotent und exprimieren Marker für PrE und EPI, regulieren jedoch jeweils einen Marker graduell herunter bis sich ihre Expression gegenseitig ausschließt und die Zelllinien differenziert sind. In der vorliegenden Arbeit konnte mit Hilfe von Computer unterstützten Berechnungen gezeigt werden, dass die Verteilung der beiden neu entstehenden Zelltypen stochastisch, und ihre Lokalisation nicht von einander abhängig ist.

Die frühesten Marker, durch die die graduelle Segregation der beiden Abstammungslinien beobachtet werden kann, sind die Transkriptionsfaktoren NANOG, der EPI-spezifisch wird, und GATA6, der PrE-spezifisch wird. NANOG ist durch seine Beteiligung an der Aufrechterhaltung der Pluripotenz in embryonalen Stammzellen bekannt und in seiner Abwesenheit im Embryo kann die EPI-Linie nicht gebildet werden. Die vorliegende Arbeit konnte zeigen, dass GATA6 der Hauptregulator der PrE-Linie ist. Embryonen ohne GATA6 aktivieren keine nachfolgenden PrE-Marker und können kein PrE bilden. Stattdessen exprimieren alle ICM-Zellen in diesen Embryonen NANOG. NANOG wird gewöhnlich in PrE-Vorläuferzellen durch den FGF Signalweg herunterreguliert. Diese Arbeit hat erstmals gezeigt, dass die FGF-vermittelte Repression von NANOG jedoch GATA6 abhängig ist und zeigt damit einen neuen Mechanismus auf, der die zweite Zellschicksalsentscheidung reguliert. Es wurde weiterhin gezeigt, dass GATA6 in einer Dosis-abhängigen Weise fungiert. In *Gata6* heterozygoten Embryonen war nicht nur die Anzahl von PrE-Zellen reduziert, sondern auch die Differenzierung dieser Zellen verlangsamt. Dies deutet darauf hin, dass

relative Level von NANOG und GATA6 die Segregation von EPI und PrE regulieren könnten.

Von beiden Abstammungslinien, EPI und PrE, können Stammzellen gewonnen werden und diese dienen als Zellkultur-Modell für die beiden Linien. Embryonale Stammzellen (ES-Zellen) repräsentieren den pluripotenten Epiblasten während sogenannte extraembryonale Endoderm-Stammzellen (XEN-Zellen) das primitive Endoderm repräsentieren. In der vorliegende Arbeit konnten effiziente Protokolle entwickelt werden um beide Stammzelltypen zu gewinnen, wodurch die Studie der Eigenschaften der beiden Linien erleichtert wird.

Abgesehen von der Beteiligung von Transkriptionsfaktoren und Signalwegen, spielt die epigenetische Genregulation eine wichtige Rolle in der Aufrechterhaltung und dem Verlust von Pluripotenz. DNA-Methylierung und Demethylierung sind solche Mechanismen. Während DNA-Methylierung wichtig für die Stilllegung von Genen, z.B. während der Differenzierung ist, findet während der Präimplantationsentwicklung DNA-Demethylierung statt, um solche repressiven Markierungen zu entfernen. In dieser Arbeit wurden ES-Zellen von Blastozysten gewonnen, denen die Deaminase AID fehlt, welche DNA-Demethylierung fördert. Solche ES Zellen konnten mit verringerter Effizienz gewonnen werden, was darauf hindeutet, dass DNA-Demethylierung in der Etablierung der pluripotenten EPI-Linie eine Rolle spielt. Darüber hinaus wurden Fibroblasten ohne AID zu induzierten pluripotenten Stammzellen (iPS-Zellen) reprogrammiert. Solche iPS-Zellen konnten ihren pluripotenten Status nicht aufrechterhalten.

Zusammenfassend gibt die vorliegende Arbeit Einblick in die Etablierung der Pluripotenz in vivo und in vitro, durch die Aufdeckung neuer Mechanismen für die Zellschicksalsentscheidung zwischen EPI und PrE, sowie der Beteiligung von epigenetischer Regulation.

1. INTRODUCTION

1.1. PREIMPLANTATION MOUSE DEVELOPMENT: AN OVERVIEW

The field of developmental biology aspires to understand the mechanisms that allow a single cell, the fertilized egg, to develop into a complex functioning organism, comprising a multitude of cells, tissues and organs that all have to interact proficiently on several levels. The process of development is therefore tightly regulated and timed. In mammals, where embryonic development takes place *in utero*, the initial days after fertilization are dedicated to laying the groundwork for implantation through formation of extraembryonic tissues that enable the absorption of maternal nutrients and by poising the embryonic lineage for the subsequent differentiation of the embryonic tissues.

In the mouse, preimplantation development spans the first five days after fertilization and is devoted to the specification of the first three lineages: the extraembryonic trophectoderm (TE) and primitive endoderm (PrE) lineages and the pluripotent epiblast (EPI) (Figure 1). During the first two days, from the fertilized egg, or zygote, to the 8-cell stage, cells (also called blastomeres) undergo a series of synchronous cleavage divisions, with all blastomeres remaining totipotent – i.e. retaining the potential to develop into all extraembryonic as well as embryonic cell types, reflected by their ability to give rise to each of the first 3 lineages (Hillman et al., 1972; Tarkowski and Wróblewska, 1967). At the 8-cell stage (~E2.5-2.75) the process of compaction begins. During this time blastomeres adhere tightly to each other via the interaction of extracellular E-Cadherin domains, which form adherens junctions (Ducibella and Anderson, 2003; Fierro-González et al., 2013; Hyafil et al., 1980; Johnson and Ziomek, 1981a; Larue et al., 1994; Pauken and Capco, 1999; Reeve and Ziomek, 1981). The compacted embryo now resembles a mulberry in shape and has therefore been named morula (from Latin *morum*). Due to the process of compaction, cells generated during the following cell division cycles acquire inside or outside positions. Cells

localized on the inside are called the inner cell mass (ICM) and initially (~ until the 6th cleavage division) retain the developmental potential of early blastomeres (Handyside, 1978; Rossant and Lis, 2003; Suwinska et al., 2008). Outside cells undergo polarization and epithelialization and thereby start the differentiation into TE, which will give rise to the ectoplacental cone and the chorion, the embryonic contribution to the placenta (Johnson and Ziomek, 1981b).

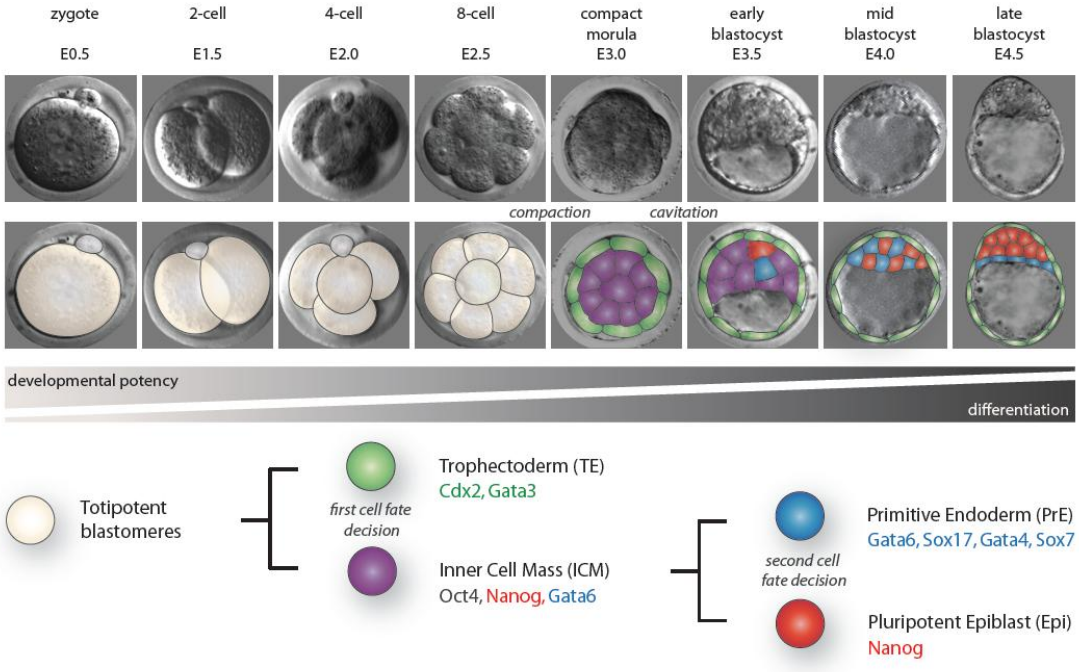


FIGURE 1. Preimplantation development. Preimplantation development is devoted to the zygote-to-embryo-transition, which is marked by a series of cell divisions leading to the formation of the blastocyst. During this time two consecutive cell fate decisions take place to specify the three lineages the implanting blastocyst is comprised of: the extraembryonic trophectoderm (green) and primitive endoderm (blue) lineages and the pluripotent epiblast (red). Modified from (Schrode et al., 2013).

Following compaction and epithelialization of outside cells, these TE cells start pumping ions into the intercellular space, thereby creating a fluid filled cavity called the blastocyst cavity (reviewed in (Watson et al., 2004)). Cavitation and growth of the embryo through two rounds of symmetric and asymmetric divisions concurrently leads to the formation the early blastocyst (~32 cells) (Motosugi et al., 2005; Smith and McLaren, 1977). A second extra-embryonic lineage now arises within the ICM, the primitive endoderm (also

referred to as hypoblast), which will give rise to the endoderm of the visceral and parietal yolk sacs and also possibly contribute cellular descendants to the gut endoderm (Kwon et al., 2008; Viotti et al., 2014). PrE cells form an epithelium lining the blastocyst cavity along the surface of the remaining inside cells, which are now referred to as the epiblast. Epithelial cells of the extraembryonic TE and the PrE now encompass the pluripotent EPI, which will ultimately give rise to the embryo proper, and all somatic cells. Having reached this stage between embryonic day 4 and 5 (~ 150-200 cells) the embryo is now ready to implant into the maternal uterine wall.

1.2. LEAVING BEHIND TOTIPOTENCY: THE FIRST CELL FATE DECISION

From fertilization to the 8-cell stage all blastomeres retain totipotency and importantly, they coexpress transcription factors that are needed for both following lineage decisions (Figure 3). At the late 8-cell stage (~ E2.75) the embryo compacts and increases E-Cadherin dependent cell-cell contacts, laying the groundwork for the first cell fate decision, which is highly influenced by spatial information (Johnson et al., 1986). Initial studies demonstrated that TE/ICM specification takes place according to an outside or inside position at the late morula stage (“inside-outside model”) (Balakier and Pedersen, 2003; Tarkowski and Wróblewska, 1967). Additionally polarity was linked to the first cell fate decision (Johnson and Ziomek, 1981a). While cells positioned on the inside of the compact morula, now called inner cell mass (ICM), remain apolar and totipotent (Rossant, 1975; Spindle, 2005), cells on the outside acquire apical-basal polarity and become biased towards the extraembryonic TE. Polarization of these cells takes places through asymmetric localization of polarity proteins, such as the complex formed by the serine/threonine atypical protein kinase C (aPKC) and the adapter proteins PAR3 and PAR6 (aPKC-PAR complex) as well as the cytoplasmic peripheral membrane protein Ezrin (EZR) (Dard et al., 2009; Louvet et al., 1996; Pauken and Capco, 2000; Plusa, 2005; Vinot et al., 2005). Following compaction, two rounds of

asymmetric cell divisions take place, eventually leading to the formation of the blastocyst.

Specification of the trophectoderm is thought to be regulated by a set of master regulators, including the transcriptional enhancer factor family member TEAD4, the caudal-related transcription factor CDX2 and the product of the T-box gene EOMES (Ciruna and Rossant, 1999; Hancock et al., 1999; Nishioka et al., 2008; Russ et al., 2000; Strumpf, 2005; Yagi et al., 2007). Of these transcription factors, deletion of *Tead4* shows the earliest defect with a failure to specify the TE and the blastocyst cavity (Nishioka et al., 2008; Yagi et al., 2007). This places TEAD4 at the top of the hierarchy, and it has been shown to initiate the TE program by activating the downstream lineage specific markers *Cdx2* and *Gata3* (Figure 3) (Nishioka et al., 2009; Ralston et al., 2010). *Tead4* is a component of the Hippo signaling pathway and is expressed in both inside and outside cells. The lineage specific activity of the Hippo signaling pathway stems from the nuclear localization of TEAD's binding partner and transcriptional co-activator YAP1 in outside cells. By contrast, in apolar inside cells, Hippo signaling through the kinase LATS1/2 phosphorylates YAP1, which inhibits its translocation to the nucleus and leads to its degradation, thereby preventing TEAD4 activity in the ICM (Nishioka et al., 2009; Ralston et al., 2010). Recent studies have linked this differential activity in inside and outside cells with cell polarity, implicating the Hippo pathway components Angiomotin (*Amot*) and Neurofibromatosis (*Nf2*) (Cockburn et al., 2013; Hirate et al., 2013; Leung and Zernicka-Goetz, 2013). AMOT is thought to form an adherens junction-associated complex with NF2 and LATS and cooperatively promote phosphorylation of YAP, leading to ICM identity. In outside cells, YAP phosphorylation would be prevented due to the sequestration of AMOT to their apical domain.

Regulation of *Cdx2* expression by the TEAD4-YAP nuclear complex in outside cells is accompanied by the concomitant downregulation of the transcription factor *Pou5f1* (*Oct4*), possibly through direct repression, since in the absence of *Oct4* most ICM cells differentiate into TE cells (Nichols et al.,

1998; Niwa et al., 2005). Reversely, in the absence of *Cdx2*, outer cells fail to form TE and to downregulate *Oct4* and *Nanog,* suggesting CDX2 represses *Oct4* and *Nanog* in TE cells (Niwa et al., 2005; Strumpf, 2005). In the ICM OCT4 in cooperation with the transcription factor SOX2 is thought to activate expression of the FGF signaling ligand *Fgf4* and the homeobox transcription factor *Nanog* as has been shown in ES cells (Ambrosetti et al., 2000; Basilico et al., 1997; Boyer et al., 2005). FGF4, secreted by ICM and later EPI cells, is an important paracrine factor, acting through the MEK/ERK signaling pathway to promote proliferation of the TE, which expresses the FGF receptor FGFR2 (Tanaka et al., 1998). On the other hand, CDX2 binds to an FGF4-responsive enhancer in the promoter of the growth factor gene *Bmp4*, which led to BMP4 secretion by TS cells (Murohashi et al., 2009). In combination with the finding that ES cell maintenance requires BMP4 (Kodaira et al., 2006; Qi et al., 2004; Ying et al., 2003) it could be inferred that TE cells secret BMP4, which may be important for ICM growth. A recent study inhibiting components of the BMP pathway suggested a role for BMP signaling in the formation of the PrE (Graham et al., 2014).

Specification of the trophectoderm is a gradual process with TE cells remaining somewhat plastic and becoming only fully committed by the mid to late blastocyst stage, accompanying the second cell fate decision (Anani et al., 2014; Rossant and Vijh, 1980; Watanabe et al., 2014).

1.3. ESTABLISHING THE PLURIPOTENT LINEAGE: THE SECOND CELL FATE DECISION

Pluripotent embryonic stem (ES) cells (Evans and Kaufman, 1981) hold great potential for regenerative medicine and have therefore been the subject of extensive research, most of it focusing on their ability to differentiate into all cell types of the body. To fully understand the potential they hold however, it is important to understand pluripotency in the context of their *in vivo* counterpart, the pluripotent EPI (Boroviak et al., 2014; Brook and Gardner, 1997). As discussed above the EPI arises during the second cell fate decision, taking place within the ICM of the blastocyst stage embryo.

Immunohistochemistry using antibodies directed against proteins, which represent lineage specific markers and live imaging using a PrE specific reporter, $Pdgfra^{H2B\text{-}GFP}$ (Hamilton et al., 2003), led to a three-phase model of lineage specification in the ICM (Figure 1 and 3) (Chazaud et al., 2006; Plusa et al., 2008). This model posits that the ICM of the early blastocyst (16-32 cells) comprises a homogenous cell population of bipotent cells, which coexpress markers for both EPI and PrE lineages. In a seemingly stochastic fashion, ICM cells then begin up- and downregulating lineage specific markers, leading to a mixed population of PrE and EPI biased cells after the mid blastocyst stage (>64 cells) (Figure 1 and 3). EPI and PrE biased cells at this stage are arranged in a salt and pepper distribution within the ICM and are marked by mutually exclusive expression of lineage specific transcription factors. These cells subsequently begin the process of sorting to their final adjacent layers and gain apical-basal polarity, which leads to the epithelialized PrE lining the blastocyst cavity by the late blastocyst stage (> 100 cells) (Figure 1) (Gerbe et al., 2008; Meilhac et al., 2009; Saiz et al., 2013). Sorting of PrE cells has been shown to involve a variety of different cell behaviors, including PrE cells relocating to the surface of the ICM, changing their lineage bias or undergoing apoptosis if neither occurs in time (Plusa et al., 2008).

While much is known about the cellular processes taking place during the second cell fate decision, there are still many open questions surrounding the gene regulatory network that governs it.

TRANSCRIPTION FACTORS

Two transcription factors, the zinc finger transcription factor GATA6 and the homeobox protein NANOG, are the earliest markers for PrE and EPI biased cells, respectively, and their interplay is thought to be at the center of the gene regulatory network controlling ICM fate choice (Chambers et al., 2003; Koutsourakis et al., 1999; Mitsui et al., 2003; Morrisey et al., 1998a). Both, *Gata6* and *Nanog*, along with the transcription factors *Oct4* and *Sox2* are co-expressed from the early morula stage onward (Figure 3) (Dietrich and

Hiiragi, 2007; Plusa et al., 2008). *Nanog*, *Sox2* and *Oct4* subsequently become restricted to the EPI lineage during the second cell fate decision (Avilion et al., 2003; Grabarek et al., 2012; Guo et al., 2010; Niwa et al., 2005; Palmieri et al., 1994; Strumpf, 2005). In contrast *Gata6* becomes confined to the PrE lineage, which subsequently activates additional PrE specific transcription factors, such as *Sox17* and *Gata4* (Artus et al., 2011; Kurimoto, 2006; Morris et al., 2010; Niakan et al., 2010; Plusa et al., 2008). Collectively these transcription factors are at the center of the second cell fate decision (Figure 3).

NANOG is a homeobox transcription factor, discovered for its importance in the maintenance of pluripotency (Chambers et al., 2003; Mitsui et al., 2003; Silva et al., 2009). Ectopic expression of *Nanog* in ES cells maintains their undifferentiated state even in the absence of leukemia inhibitory factor (LIF, see Introduction p. 27) (Chambers et al., 2003). *In vivo* NANOG is expressed in the early morula, in the ICM of blastocyst stage embryos, in the EPI of E6.5 and E7.5 embryos and in embryonic germ cells (Chambers et al., 2003; Hart et al., 2004; Hatano et al., 2005; Mitsui et al., 2003). During the second cell fate decision NANOG is the first transcription factor that becomes restricted to the EPI lineage (Guo et al., 2010). *Nanog* deficient embryos fail to specify the EPI and display pan-ICM expression of the PrE marker GATA6 (Frankenberg et al., 2011; Messerschmidt and Kemler, 2010; Mitsui et al., 2003; Silva et al., 2009). In both, ES cells as well as *in vivo, Nanog* expression is positively regulated by the transcription factors OCT4 and SOX2, which form a regulatory complex that binds to the *Nanog* promoter (Liang et al., 2008; Rodda, 2005; Wang et al., 2006).

SOX2 is a member of the SRY-related HMG-box (SOX) family of transcription factors and is involved in the self-renewal of cells of the developing central nervous system (Li et al., 1998; Zappone et al., 2000). In ES cells *Sox2* is regulated by an autoregulatory SOX2/OCT4 complex (Chew et al., 2005). It is involved in the regulation of other pluripotency genes like *Oct4* and *Nanog* and is therefore essential for the maintenance of

pluripotency (Chew et al., 2005; Kuroda et al., 2005; Masui et al., 2007; Okumura-Nakanishi et al., 2005). *In vivo*, *Sox2* is the earliest marker of ICM cells and its restriction was suggested to be regulated by HIPPO signaling (Avilion et al., 2003; Guo et al., 2010; Wicklow et al., 2014). During the second cell fate decision SOX2 becomes restricted to the EPI lineage (Avilion et al., 2003; Guo et al., 2010; Wicklow et al., 2014). Accordingly, embryos lacking *Sox2* are peri-implantation lethal with a failure to form ICM (Avilion et al., 2003; Keramari et al., 2010) and cannot yield ES cells (Avilion et al., 2003). It is one of the factors used to reprogram somatic cells to induced pluripotent stem (iPS) cells, underscoring its importance for the regulation of pluripotency (Takahashi and Yamanaka, 2006).

OCT4 is a transcription factor necessary, like *Sox2*, to maintain pluripotency in ES cells and *in vivo* (Nichols et al., 1998; Niwa et al., 2000). *In vivo* it is expressed in germ cells, oocytes and the ICM of the morula until early to mid blastocyst but then becomes restricted to the EPI (Palmieri et al., 1994; Pesce et al., 1998; Rosner et al., 1990; Schöler et al., 1990; Yeom et al., 1996). Homozygous null mutant embryos are peri-implantation lethal with a failure to specify PrE and the blastocyst eventually differentiates into trophectoderm (Frum et al., 2013; Le Bin et al., 2014; Nichols et al., 1998). OCT4 forms a trimeric complex with SOX2 to autoregulate their own expression as well as other pluripotency factors such as *Nanog* and *Fgf4* (Chew et al., 2005; Kuroda et al., 2005; Okumura-Nakanishi et al., 2005; Rodda, 2005; Yuan et al., 1995). In a complex with NANOG, OCT4 and SOX2 are also thought to repress differentiation-promoting genes (Sun et al., 2006).

GATA6 is a zinc finger transcription factor and member of the GATA family. In the postimplantation embryo it is involved in liver development and together with *Gata4* in heart and pancreas development (Koutsourakis et al., 1999; Morrisey et al., 1996; Xin et al., 2006; Xuan et al., 2012; Zhao et al., 2005; 2008). During preimplantation development *Gata6* is expressed in all cells in the early morula but becomes restricted to the PrE coincident with the second cell fate decision (Chazaud et al., 2006; Plusa et al., 2008). It is

thereafter involved in the formation of the extraembryonic endoderm lineages (Koutsourakis et al., 1999; Morrisey et al., 1998a). Accordingly, overexpression of *Gata6* in ES cells causes their conversion to extraembryonic endoderm (XEN, (Kunath, 2005)) stem cells (Fujikura et al., 2002; Shimosato et al., 2007). Embryos lacking *Gata6* were initially reported to be early postimplantation lethal with defects in PrE derivative lineages (Koutsourakis et al., 1999; Morrisey et al., 1998a) as well as with an absence of a PrE layer as early as E4.5 (Cai et al., 2008). However, in these previous studies of early *Gata6* null mutant embryo genotyping was performed using *in situ* hybridization or immunofluorescence on histological sections and did not provide robust unequivocal genotyping information as do the PCR genotyping methods that are used as a standard today (see (Morrisey et al., 1998a), Figure 5). The exact phenotype and thereby role of *Gata6* during preimplantation development therefore remained an open question.

Sox17 is a member of the SOX family of transcription factors and is involved in the development of gut endoderm, cardio-vascular and fetal (but not adult) hematopoietic stem cells (Francois et al., 2010; Kanai-Azuma et al., 2002; Kim et al., 2007; Matsui, 2006). During preimplantation development *Sox17* is expressed in the PrE (Morris et al., 2010; Niakan et al., 2010). Homozygous null mutants form a PrE layer but are postimplantation lethal with severe impairment in the formation of the definitive endoderm (Kanai-Azuma et al., 2002; Viotti et al., 2014). However, when implantation is artificially delayed through experimental induction of diapause, embryos lacking *Sox17* display defects in the epithelial integrity of the PrE (Artus et al., 2011), suggesting it is involved in the formation of the extracellular matrix secreted by the PrE during preimplantation development.

Gata4, like *Gata6* is a member of the GATA transcription factor family and is expressed in the heart, liver, pancreas and small intestine of the developing embryo (Arceci et al., 1993). *Gata4* and *Gata6* are speculated to have partially redundant roles, for example in pancreas and heart development (Xin et al., 2006; Xuan et al., 2012; Zhao et al., 2005). During

preimplantation development *Gata4* is activated specifically in PrE cells, and it is the first factor exclusively marking the PrE lineage (Kurimoto, 2006; Plusa et al., 2008). *Gata4* null mutant embryos display postimplantation lethality with defects in heart development and ventral morphogenesis (Kuo et al., 1997; Molkentin et al., 1997; Narita et al., 1997). Nevertheless, GATA4 is capable of inducing extraembryonic endoderm fate and *Gata6* expression when overexpressed in ES cells (Fujikura et al., 2002) and its interaction with GATA6 has been suggested to be necessary for PrE specification (Capo-Chichi et al., 2005; Fujikura et al., 2002; Soudais et al., 1995). These results support the important roles of GATA4 for PrE specification while the less severe early defects in single knockout embryos support the hypothesized redundancy of GATA6 and GATA4.

FGF SIGNALING

As discussed above, *Gata6* and *Nanog* are thought to be the transcription factors governing the events of the second cell fate decision. Their interplay is additionally impinged on by FGF signaling, potentially aiding communication between ICM cells during their lineage divergence.

LIGANDS AND RECEPTORS

The FGF family of growth factors encompasses 22 members, which signal through four ligand-dependent FGF receptor tyrosine kinases (Figure 2). These FGF receptors gain additional ligand specificity through alternative splicing of the extracellular binding domain and their isoforms are tissue-specifically regulated (reviewed in (Zhang et al., 2006)). Activation of the FGF signaling pathway at preimplantation stages seems to occur specifically through the FGF4/FGFR2 ligand/receptor complex (Arman et al., 1998; Feldman et al., 1995; Wilder et al., 1997) (Figure 2).

Fgf4, like *Nanog*, is under the direct transcriptional control of the OCT4/SOX2 regulatory complex (Ambrosetti et al., 2000; Avilion et al., 2003; Keramari et al., 2010; Nichols et al., 1998; Yuan et al., 1995) and is expressed highly in all cells of the ICM from the 8-cell stage onwards (Guo et

al., 2010). It becomes restricted to EPI-biased cells at the beginning of their specification (Guo et al., 2010). Null mutants for *Fgf4* fail to specify the PrE and display pan-ICM expression of *Nanog* (Kang et al., 2013; Krawchuk et al., 2013).

The FGF receptors FGFR1 and FGFR2 are both expressed in the ICM of the preimplantation embryo (Ohnishi et al., 2014), but only deletion of *Fgfr2* has been shown to be peri-implantation lethal (Arman et al., 1998). Furthermore, FGFR1 remains expressed in both PrE and EPI, while FGFR2 becomes restricted to the PrE during the second cell fate decision (Guo et al., 2010; Ohnishi et al., 2014).

SIGNAL TRANSDUCTION

FGF signaling is activated by the formation of an FGF ligand/FGF receptor (FGFR) complex, leading to dimerization of the receptor. Downstream signal transduction proceeds mainly through four different pathways, the JAK/STAT, the PLCγ, the PI3K and the ERK pathway (Figure 2) (Dailey et al., 2005). The latter two have been found to be active during preimplantation development. Both are relayed through formation of a complex between FGFR, FRS2a, SHP2 and GRB2, which in turn activates AKT via PI3K and ERK via the SOS/RAS pathway (Figure 2). In ES cells the FGF/AKT pathway acts upstream of extracellular matrix components like *Laminin* and *Collagen IV* (Li et al., 2001) and therefore seems likely to be involved in the epithelial maturation of the PrE rather than its formation. In contrast, the FGF/ERK pathway has been strongly implicated in the segregation of the EPI and PrE lineages. Perturbation of the FGF/ERK pathway in preimplantation stage embryos has been shown to cause severe effects on PrE/EPI lineage segregation. Null mutants for the ligand *Fgf4*, its putative receptor *Fgfr2*, the downstream adapter *Grb2* and inhibition of the mitogen-activated protein kinase ERK all lead to a failure to fully specify the PrE, which leads to an ICM comprised solely of EPI cells (Arman et al., 1998; Chazaud et al., 2006; Cheng et al., 1998; Feldman et al., 1995; Goldin and Papaioannou, 2003; Kang et al., 2013; Krawchuk et al., 2013; Nichols et al.,

2009; Ohnishi et al., 2014; Yamanaka et al., 2010). Conversely, treatment of early embryos with ectopic FGF4 directs all ICM cells towards the PrE lineage (Yamanaka et al., 2010).

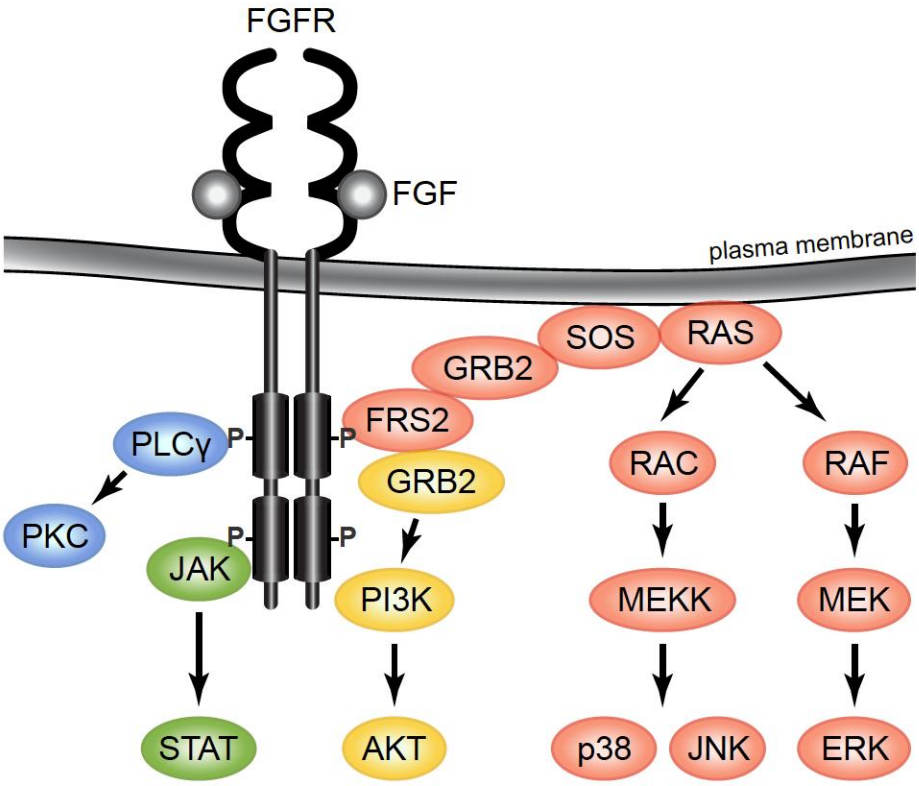


FIGURE 2. FGF signaling pathway. FGF signaling is relayed through 4 different pathways (JAK/STAT; PLCγ; PI3K/AKT; MEK/ERK). In ES cells the MEK/ERK pathway downregulates *Nanog*. Refer to text for further information. Arrows: positive regulation.

REGULATION OF PRE/EPI DIFFERENTIATION

TRANSCRIPTION FACTOR INTERACTIONS

As discussed above, many of the transcription factors as well as FGF signaling components that are involved in the second cell fate decision are coexpressed from the early morula stage onward (Dietrich and Hiiragi, 2007; Guo et al., 2010; Plusa et al., 2008). Lineage bias towards EPI or PrE is initiated around the early blastocyst stage when cells in the ICM start downregulating either *Gata6* or *Nanog*, with future EPI cells downregulating *Gata6* and future PrE cells *Nanog,* until they attain a mutually exclusive

expression pattern (Chazaud et al., 2006; Guo et al., 2010; Kurimoto, 2006; Plusa et al., 2008; Rossant et al., 2003).

From this stage onwards the two factors become the first markers for their respective lineages and are thought to govern their maturation. Indeed, overexpression of *Gata6* is sufficient to direct ES cells towards an extraembryonic endoderm like state and leads to the downregulation of *Nanog* (Fujikura et al., 2002; Shimosato et al., 2007). In ES cells NANOG has been shown to bind the *Gata6* promoter and possibly suppress it (Mitsui et al., 2003; Singh et al., 2007). These observations hint at a mutually repressive relationship between *Gata6* and *Nanog* and it is thought that early co-expression of the factors is maintained by a state of mutual repression until this balance is disturbed and lineage segregation initiated at the early blastocyst stage. Supporting this idea is the observation that all ICM cells of *Nanog* null mutants express GATA6 (Frankenberg et al., 2011; Messerschmidt and Kemler, 2010; Silva et al., 2009) while *Gata6* null mutants display defects in PrE derivative lineages (Cai et al., 2008; Koutsourakis et al., 1999; Morrisey et al., 1998a).

If the two factors indeed repress each other, the question remains how they maintain stable expression levels until lineage segregation commences. As has been shown for *Oct4* and *Sox2* (Jaenisch and Young, 2008), positive autoregulation has been speculated for *Gata6* (Frankenberg et al., 2011). The *Gata6* gene has been found to have GATA binding motifs of its own and could therefore regulate its own expression (Molkentin, 2000). *Nanog* is known to be positively regulated by SOX2 and OCT4 as well as through positive autoregulation (Jaenisch and Young, 2008; Rodda, 2005; Wu et al., 2006). However, repression through FGF signaling (Santostefano et al., 2012), as well as autorepression, has also been shown for *Nanog* in ES cells (Fidalgo et al., 2012; Navarro et al., 2012).

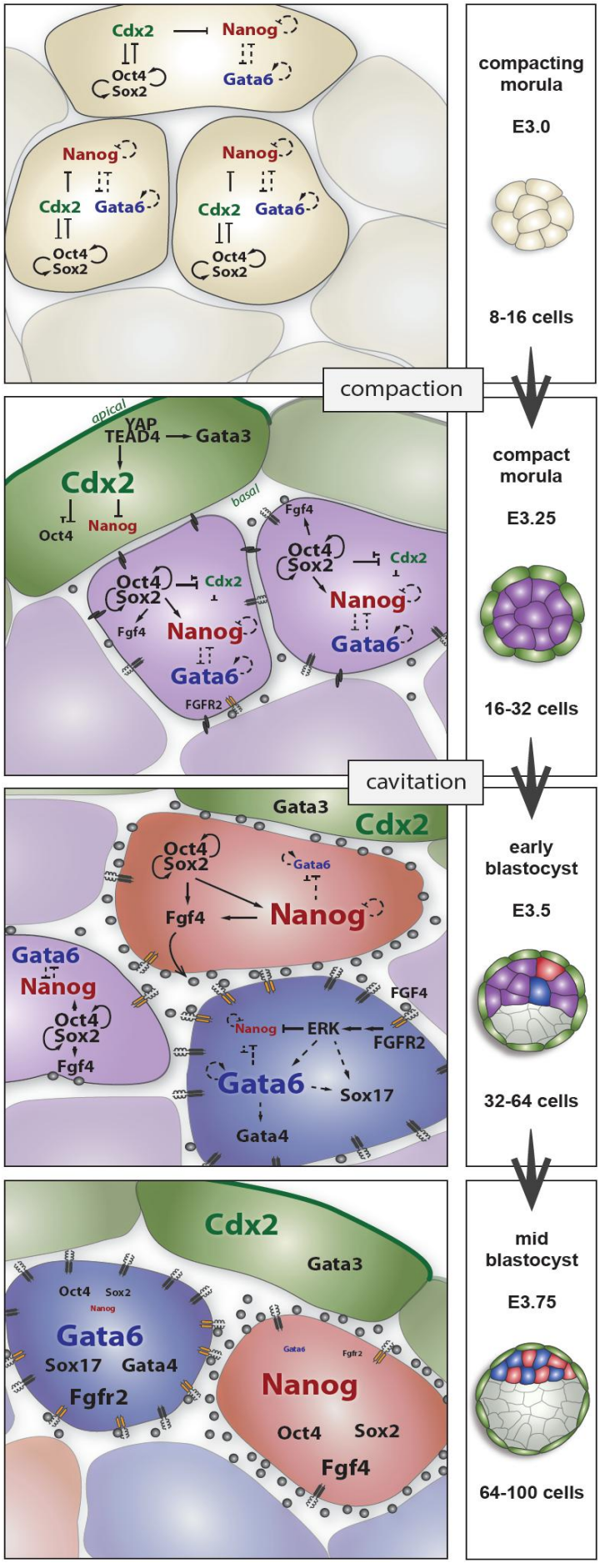


FIGURE 3. Gene regulatory networks governing pre-implantation development. The lineage specific transcription factors *Cdx2* (TE), *Gata6* (PrE) and *Nanog* (EPI) are coexpressed at the early morula stage but subsequently segregate to their respective lineages. Cdx2 becomes exclusive to the TE during the first cell fate decision at the morula stage and represses *Oct4* and *Nanog* in TE biased cells. In the inner cell mass *Nanog* as well as *Fgf4* are positively regulated by a OCT4/SOX2 regulatory complex. Additionally, *Gata6* and *Nanog* may autoregulate themselves. *Nanog* and *Gata6* become exclusive to EPI and PrE respectively during the second cell fate decision, starting at the early blastocyst stage. In EPI biased cells NANOG is thought to repress *Gata6* and vice versa. EPI and PrE specification is additionally regulated through FGF/ERK signaling, with EPI cells secreting FGF4. This may lead to repression of *Nanog* through the FGFR2/ERK pathway in PrE biased cells. Further PrE specific genes, such as *Sox17* and *Gata4,* may be regulated by GATA6 and FGF/ERK signaling. Refer to text for further information. Green: TE. Blue: PrE. Red: EPI. Purple: bipotent ICM. Arrows: positive regulation. Blocked arrows: repression. Dashed arrows: speculated interactions. Adapted from (Schrode 2013).

INFLUENCE OF FGF SIGNALING

Gata6 and *Nanog* display balanced co-expression until the early blastocyst stage when their divergence begins. FGF signaling is directly involved in this process by promoting PrE differentiation in a cell non-autonomous fashion. This effect is thought to be caused by downregulation of *Nanog* through FGF/ERK signaling in these cells, which in case of NANOG/GATA6 mutual repression would lead to upregulation of *Gata6*. Indeed, work in ES cells showed that homodimerization of FGFR2, leading to its constitutive activation is sufficient for transcriptional downregulation of *Nanog* (Santostefano et al., 2012).

In addition, the reciprocal expression of FGF ligand and receptor in the two lineages paralleling their segregation suggests the direct involvement of FGF4 and FGFR2 in the second cell fate decision. Moreover, Guo et al. showed that the inverse correlation of *Fgf4* and *Fgfr2* in the ICM precedes that of *Nanog* and *Gata6* (Guo et al., 2010). In a recent single cell expression profiling study of embryos during preimplantation development, Ohnishi et al. further demonstrated that *Fgf4* is indeed the first lineage specific factor that is differentially expressed in the ICM of early blastocysts and seems therefore to be a candidate to break homogeneity in the ICM (Ohnishi et al., 2014). These findings support a hypothesis where FGF4 initiates the second cell fate decision and provides the cue to break symmetry in the ICM.

Interestingly, recent studies investigating *Fgf4* mutant embryos, demonstrated a reduction of PrE cells in *Fgf4* zygotic or maternal zygotic heterozygous embryos, respectively (Kang et al., 2013; Krawchuk et al., 2013). Moreover, Kang et al. showed that the EPI/PrE salt and pepper conformation could not be rescued in *Fgf4* null embryos by treatment with exogenous FGF4, while Krawchuck et al. occasionally saw progressive modulation of lineage proportions through treatment with varying FGF doses. This suggests that FGF4 levels have to be tightly regulated in the embryo to ensure the differentiation of a correct ratio of EPI and PrE cells in the ICM and

it raises the question of how the initially homogenous cells of the ICM are able to react differentially to similar levels of FGF4 in the extracellular space.

ESTABLISHING THE INITIAL LINEGEAGE BIAS

Though above findings point to FGF4 as the initial cue that starts the second cell fate decision, it is still unclear what causes *Fgf4* to be expressed differentially in the first place.

One model assumes a stochastic event caused by expression heterogeneities breaks symmetry in the ICM and causes cells to acquire a lineage bias – especially *Nanog* shows very heterogeneous expression in ICM cells as well as ES cells (Chambers et al., 2007; Kalmar et al., 2009; Mitsui et al., 2003; Torres-Padilla and Chambers, 2014) and it has been speculated that NANOG acts upstream of *Fgf4* (Frankenberg et al., 2011). Such heterogeneities could initiate a bias in cells of the ICM.

On the other hand, a study by Morris et al. proposed that the time of internalization of ICM cells influences their future fate (Morris et al., 2010). Accordingly, cells internalized during the second wave of asymmetric divisions would be further along in their development and therefore express higher levels of FGF4's putative receptor FGFR2, which would increase their susceptibility to FGF signaling and increase the potential to differentiate towards PrE (Morris et al., 2013). However, an independent tracing study by Yamanaka et al. found no such correlation (Yamanaka et al., 2010). More studies are therefore necessary in this area to determine the initial mechanism(s) that break symmetry within the ICM.

CHANGES IN DEVELOPMENTAL PLASTICITY

While it is still unclear how the initial lineage bias is established and how expression of the master regulators *Gata6* and *Nanog* is regulated to govern this process, observations have been made about the events following the initiation of lineage specification.

Following the initial upregulation of *Fgf4* in future EPI cells, its putative receptor *Fgfr2* is upregulated in PrE precursor cells (Guo et al., 2010; Ohnishi

et al., 2014). PrE biased cells sequentially downregulate the pluripotency markers *Nanog*, *Sox2* and finally *Oct4* and in parallel activate additional PrE markers, namely *Sox17*, *Gata4* and *Sox7* (Figure 3) (Artus et al., 2011; Frankenberg et al., 2011; Grabarek et al., 2012; Guo et al., 2010; Kurimoto, 2006; Morris et al., 2010; Niakan et al., 2010; Plusa et al., 2008). Both FGF/ERK signaling, as well as GATA6, are suggested to be involved in this process. FGF/ERK signaling leads to the downregulation of *Nanog* and possibly upregulation of *Sox17* and *Gata4* (Frankenberg et al., 2011). Additionally, both, *Sox17* and *Gata4* posses GATA response elements and are therefore speculated to be regulated by GATA6 (Niakan et al., 2010; Wang and Song, 1996). However, in *Nanog* null mutant embryos, *Gata6*-positive ICM cells do not express *Sox17* and *Gata4*, suggesting that FGF signaling is indeed cooperatively involved in the activation of these genes (Frankenberg et al., 2011).

Simultaneously, pluripotency markers, which were homogenously expressed throughout the ICM until the early blastocyst stage, are being restricted to future EPI cells. *Sox2* is found solely in EPI cells around the mid blastocyst stage (Guo et al., 2010). Interestingly, *Oct4*, which is commonly regarded, and taken advantage of, as an important pluripotency marker in ES cells, is restricted to the EPI only by the late blastocyst stage (Grabarek et al., 2012) once the lineages have sorted. This points to ES cells representing mature EPI cells, which has recently been confirmed through gene expression profiling and functional cell-derivation assays (Boroviak et al., 2014). Around the time of implantation *Nanog* is progressively downregulated, priming the EPI for the subsequent differentiation into the embryonic lineages (Chambers et al., 2003; Hatano et al., 2005; Medvedev et al., 2008; Mitsui et al., 2003). It can be readily inferred that precursors of both lineages become increasingly locked in their fate as the differentiation program progresses and additional factors are activated. By the time the PrE and EPI lineages exhibit a clear salt and pepper distribution in the ICM, their cell fates were thought to be fully committed and to have lost their initial plasticity. However, experiments

involving treatment of embryos with exogenous FGF4, as well as ERK inhibitors, for various time frames during blastocyst maturation showed that both cell types retain their EPI/PrE bipotency until the late blastocyst stage, if forced by extreme modulation of FGF signaling (Yamanaka et al., 2010). Similarly, isolation of lineage biased cells from various stages and transplantation into early embryos showed varying plasticity of ICM cells up to the peri-implantation stage and intriguingly a higher plasticity of PrE than EPI cells (Grabarek et al., 2012). This raises the question when and how exactly cells of the ICM commit to either lineage and loose their ability to respond to FGF signaling cues, i.e. their plasticity.

Among the multitude of possible factors influencing plasticity, as well as fate stabilization, is epigenetic regulation.

DNA METHYLATION STATUS

The adult mammalian body consists of an astounding number of different cell types, forming complex tissues and organs that are in constant cross talk with each other to function correctly. The formation of the embryo proper entails the differentiation of the pluripotent EPI cells into a multitude of precursor cells and eventually differentiated somatic cell types. EPI cells are therefore subject to strict regulation to maintain pluripotency and simultaneously to be able to readily enter a primed state and initiate differentiation upon various cues (Chambers et al., 2003). Complex signaling and transcription factor cascades ensure directed differentiation of these cells while epigenetic modifications, such as histone marks and DNA methylation, secure their identity is maintained and prevent straying from their given differentiation path (Reik, 2007).

DNA methylation is considered a repressive mark and is one of the most important epigenetic mechanisms to ensure cell identity, for example by silencing pluripotency genes in differentiating cells (Tsumura et al., 2006). It occurs through methylation of the fifth carbon of cytosine, creating 5-methylcytosine (5mC) and mainly occurs in the context of symmetrical CpG sequences (Bird and Wolffe, 1999; Jones and Takai, 2001). So-called CpG

islands, DNA stretches with high CpG density, are often found near the promoter regions of genes. When these are hypomethylated it is considered a sign of active gene transcription (Cross and Bird, 1995; Larsen et al., 1992). De novo DNA methylation is mainly achieved through the DNA methyltransferases DNMT3a and DNMT3b. On the other hand, the maintenance of distinct methylation patterns in specialized cells and tissues is ensured through DNMT1, which recognizes hemimethylated DNA in daughter cells and restores the full methylation status (Bestor, 2000; Li et al., 1992; Margot et al., 2003; Okano et al., 1999).

At the blastocyst stage a hypomethylated genome enables pluripotent EPI cells to differentiate into any cell type and acquire their respective distinct DNA methylation pattern (Tsumura et al., 2006) (Figure 4). This is not the case at the zygote stage however, since sperm cells and oocytes are highly (~90 % and ~40 % respectively) methylated cell types (Kobayashi et al., 2012). As a result, global demethylation must occur after fertilization to reset the epigenome for the next generation (Figure 4).

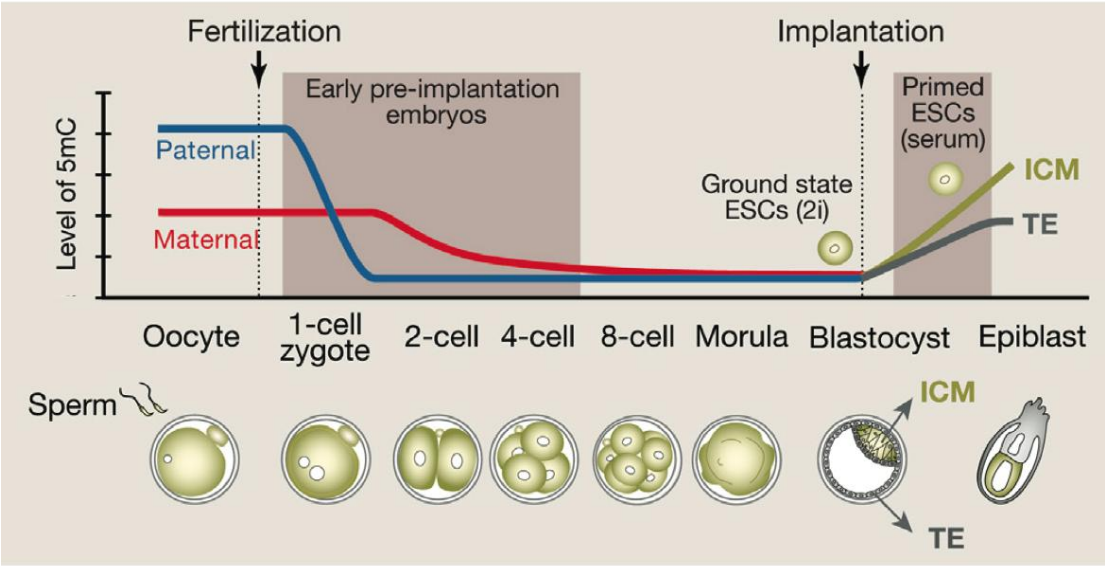


FIGURE 4. DNA methylation and demethylation during preimplantation development. The DNA of the fertilized oocyte, combining both parental genomes, is highly and asymmetrically methylated. The paternal genome is rapidly and probably actively demethylated shortly after fertilization. In contrast, the maternal genome is passively demethylated. DNA demethylation is completed by the blastocyst stage when the first cell fate decisions are taking place and de novo methylation of the embryonic genome is initiated. Adapted from (Wu and Zhang, 2014).

DNA methylation is considered a comparatively stable epigenetic mark and to date no enzyme is known to remove methyl groups directly (Cortázar et al., 2007). Alternatively however, methylation density can be diluted passively through replication in the absence of nuclear DNMT1, thereby causing gradual loss of the mark in daughter cells (Howell et al., 2001; Santos et al., 2002). This replication dependent demethylation is thought to occur in the maternal genome after fertilization and during preimplantation development (Dean et al., 2001; Howell et al., 2001; Oswald et al., 2000; Rougier et al., 1998; Santos et al., 2002). In contrast, the paternal genome exhibits the highest DNA methylation level of all murine cells and undergoes rapid and complete demethylation shortly after zygote formation (Dean et al., 2001; Mayer et al., 2000; Oswald et al., 2000; Popp et al., 2010; Santos et al., 2002). It has therefore been suggested that the paternal genome undergoes active demethylation. However, a mechanism to allow for such a process has proven elusive until very recently. Groundbreaking studies revealed that such active demethylation can take place by oxidation of 5-methylcytosine (5mC) through TET (ten-eleven translocation protein) enzymes into 5-hydroxymethylcytosine (5hmC) and into further oxidized derivatives 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (Ito et al., 2011; Kriaucionis and Heintz, 2009; Pfaffeneder et al., 2011; Tahiliani et al., 2009). These derivatives seem to compromise binding of the methylation maintenance machinery and thereby augment passive demethylation (reviewed in (Li et al., 2015)) (Figure 5). DNA demethylation through TET proteins may play a role in the first cell fate decision, where knockdown of *Tet1* favored cell specification towards the TE (Ito et al., 2010).

Another mechanism proposed to accomplish active demethylation is through deamination of 5-methylcytosine to thymidine (Morgan et al., 2004). This conversion triggers the base excision repair mechanism and results in restoration of unmethylated cytosine. Two members of the APOBEC family of cytosine deaminases that could fulfill such a function are the AID and APOBEC1 (Morgan et al., 2004; Rai et al., 2008). AID has been found to be

coexpressed with the pluripotency related factors *Nanog* and *Stella* in ES cells, oocytes and germ cells and has been implicated in the demethylation of *Nanog* and *Oct4* promoters in somatic cell reprogramming (Bhutani et al., 2010; Morgan et al., 2004). However, it is yet widely unknown if or how the enzyme plays a role in early demethylation events.

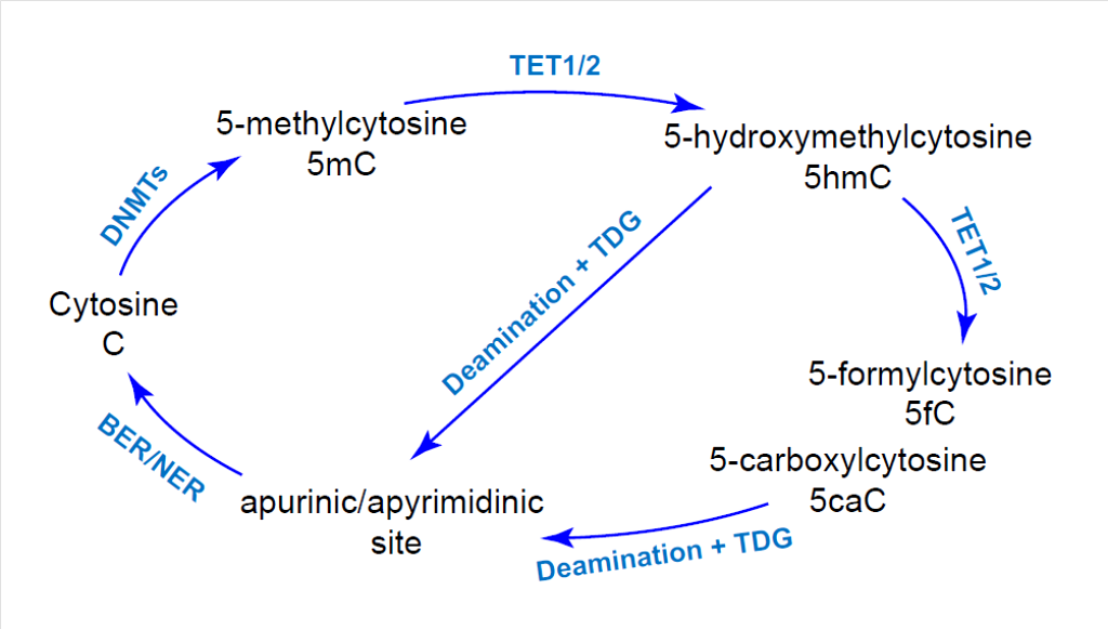


Figure 5. DNA methylation cycle. DNA methylation is carried out at the 5' carbon position of cytosine by DNMT enzymes and produces 5-methylcytosine (5mC). TET enzymes catalyze the oxidation of 5mC into 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and finally 5-carboxylcytosine (5caC). Compromised binding of the methylation machinery leads to enhanced passive demethylation. Alternatively, 5caC as well as 5hmC can be deaminated by deaminases such as AID or APOBEC1. The resulting thymidine is repaired through base excision repair (BER) and produces unmethylated cytosine. Refer to text for further information. TDG: Thymine DNA glycosylase.

1.4. Stem cell culture models for blastocyst lineages

The study of preimplantation development is often hampered by the limitations the system presents. A single embryo consists of not more than 8 to 200 cells. If the lineages are to be analyzed separately they have to be painstakingly isolated, which further decreases sample size for high-throughput biochemical analyses. Single cell approaches have proven successful but are difficult and time-consuming. To be able to further study the molecular mechanisms underlying the specification and maintenance of

the different preimplantation lineages, tissue culture models can be derived from blastocyst stage embryos (Figure 6).

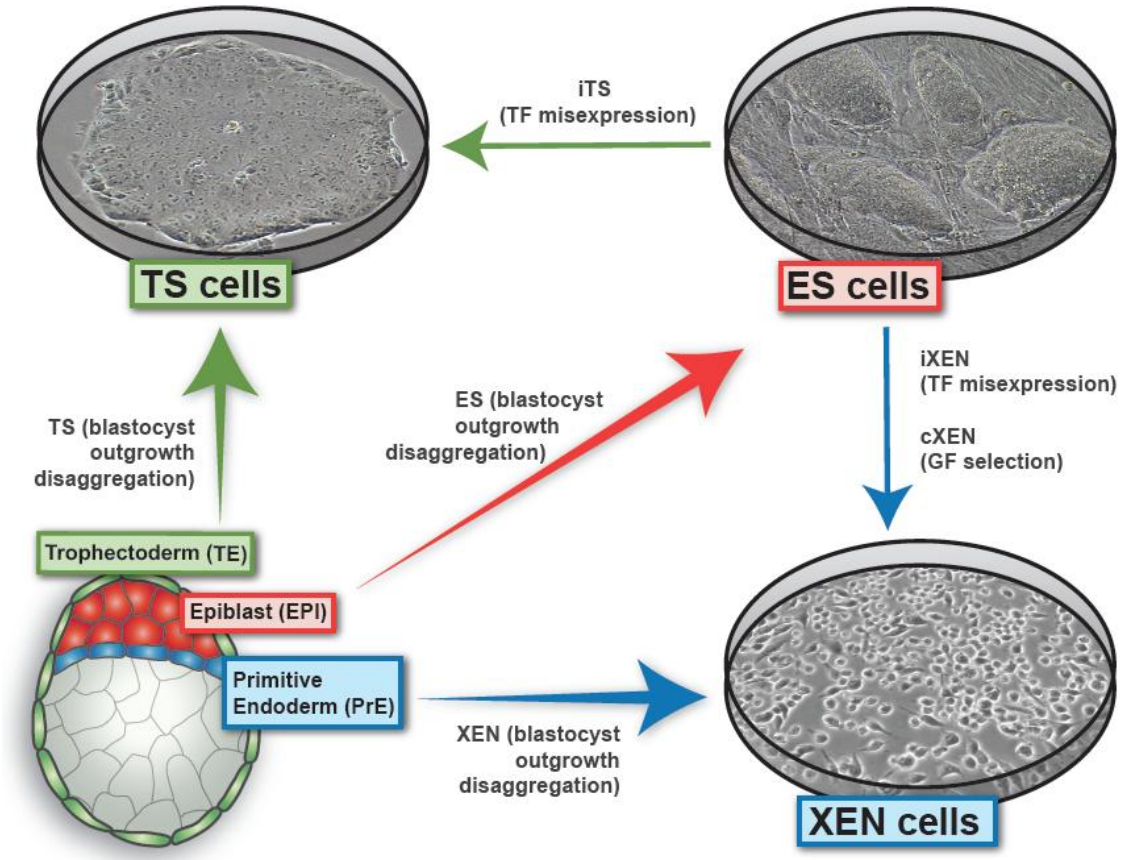


FIGURE 6. Stem cell types derived from the blastocyst. Stem cells can be derived from all three lineages of the late blastocyst stage embryo. Throphoblast stem (TS) cells are derived from the TE, extraembryonic endoderm stem (XEN) cells from the PrE and embryonic stem (ES) cells from the pluripotent EPI. Adapted from (Niakan et al., 2013).

EMBRYONIC STEM CELLS AND INDUCED PLURIPOTENT STEM CELLS

The most commonly known and widely studied cells derived from the blastocyst stage embryo are embryonic stem (ES) cells, foremost for their potential for regenerative medicine (reviewed in (Hochedlinger and Jaenisch, 2003; Lerou and Daley, 2005)). ES cells are derived from the ICM of blastocyst stage embryos and represent the *in vitro* counterpart of the EPI lineage (Boroviak et al., 2014; Brook and Gardner, 1997). As such, they are

pluripotent – i.e. they have the potential to differentiate into all somatic tissues and germ cells – and capable of self-renewal – i.e. they can divide indefinitely. They express markers of the pluripotent epiblast, such as *Nanog*, *Sox2*, *Oct4*, *Klf4* and *Esrrb*. Importantly, when injected into blastocysts they contribute solely to the EPI and all derivative germ layers in resulting chimeric embryos, the gold standard to test for developmental potential of any cell type (Bradley and Robertson, 1986).

The first murine ES cells were successfully derived in 1981 by Evans and Kaufman on mitotically inactivated mouse embryonic fibroblast (MEF) feeder cells and in serum containing medium (Evans and Kaufman, 1981), as well as by Martin using teratocarcinomas cell-conditioned medium (Martin, 1981). Advancements in understanding ES cell biology and pluripotency have since made it possible to derive and maintain ES cells in more defined conditions:

Feeder cells are thought to aid ES cells by providing cell-cell interactions and secreting beneficial growth factors (Eiselleova et al., 2008; Lim and Bodnar, 2002; Ma et al., 2012; Meng et al., 2008; Villa-Diaz et al., 2009), including leukemia inhibitory factor (LIF) (Hasegawa et al., 2011; Stewart et al., 1992). The benefit of adding LIF to ES culture media was found in 1988 and allows ES cell propagation in the absence of feeder cells (Williams et al., 1988). Similarly, it has been shown that teratocarcinoma stem cell-conditioned medium is equivalent to a 5-fold concentration of the LIF added to ES cell culture, explaining its effect in the original derivation process (Davies and Fairchild, 2012). In ES cells LIF activates the JAK/STAT pathway, which inhibits differentiation and promotes self-renewal activating core pluripotency genes, including *Nanog*, *Oct4* and *Klf4* (Hirai et al., 2011; Matsuda et al., 1999; Niwa et al., 1998; 2009; Smith et al., 1988). The effects of LIF in ES cell culture simulate activation of STAT3 in the preimplantation embryo *in vivo* through WNT/GSK3 signaling.

Fetal bovine serum (FBS) is a non-chemically defined supplement used almost universally for cell culture because it promotes cell growth and

proliferation. It contains nutrients and growth factors for ES cells, but also potential differentiation factors (Mannello and Tonti, 2007; Ogawa et al., 2004; Vanroose et al., 2001). More chemically defined serum replacement formulae were therefore developed, both commercially (Knockout Serum Replacement, KSR) and non-commercially (N2B27) to maintain and derive murine ES cells (Price et al., 1998; Ying et al., 2003; 2008). The withdrawal of serum alone however cannot maintain pluripotency, as ES cells produce autocrine factors that lead to differentiation. An essential growth factor for ES cell self-renewal, which is contained in serum, is bone morphogenetic protein 4 (BMP4), which induces the expression of Id genes and inhibits MEK signaling (Kodaira et al., 2006; Qi et al., 2004; Ying et al., 2003). LIF in synergy with BMP4 is therefore sufficient to maintain ES cell pluripotency and self-renewal in culture and can be used to substitute both feeder cells and serum or either one respectively.

As in the ICM of the blastocyst, the autocrine differentiation-promoting factor secreted by ES cells is FGF4, which activates ERK signaling. On this basis Ying and colleagues proposed that LIF and BMP signals act downstream of ERK to shield the pluripotent state. They found that inhibiting differentiation through ERK, while promoting growth capacity through inhibition of glycogen synthase kinase-3 (GSK3) could maintain pluripotency without requirement of any growth factors or cytokines (Ying et al., 2008). Initially this was achieved through small-molecule inhibitors blocking FGF receptor tyrosine kinases, the ERK cascade and GSK3 (3i). A potent inhibitor of ERK however was able to substitute for inhibitors of FGFR and the downstream ERK cascade. This two-inhibitor culture condition (or 2i) was able to more effectively maintain the pluripotent state and has since revolutionized the culture and derivation of ES cells.

Despite advancements in ES cell culture, their derivation remains often inefficient (10% - 30%) and can be highly strain dependent (Brook and Gardner, 1997; Kawase et al., 1994; McWhir et al., 1996; Robertson, 1997; Suzuki et al., 1999). For this reason researchers often employ additional strategies to increase yield, such as microsurgical isolation of the inner cell

mass, selective removal of differentiating cells or aggregation with outbred host embryos (Brook and Gardner, 1997; Gertsenstein et al., 2010; McWhir et al., 1996). Inhibition of GSK3 could further increase efficiency up to 76% for the permissible C57BL/6 strain and 31% for the less permissible BALB/c strain (Umehara et al., 2007).

While the derivation and maintenance of murine ES cells is improving continuously, the study of human ES cells, essential for future applications in regenerative medicine, has been hampered by ethical and moral concerns surrounding the usage of live embryos for their derivation. A groundbreaking study in 2006 showed that overexpression of the transcription factors *Oct4*, *Sox2*, *Klf4* and *cMyc* is sufficient to induce reprogramming of differentiated cells to a pluripotent state. These induced pluripotent stem (iPS) cells display morphological and functional characteristics of ES cells, including the expression of pluripotency markers and their developmental potential (Takahashi and Yamanaka, 2006). These iPS cells are derived from adult cells and could therefore be used to create patient specific pluripotent cells. Thus, great efforts are being made to study the mechanisms underlying cellular reprogramming, the actual similarity and potential disparities to ES cells as well as their safety and potential for medical use (for review, see (Buganim et al., 2013; Deng et al., 2015; Hanna et al., 2010; Inoue et al., 2014; Lowry, 2012; Simonson et al., 2015)).

TROPHOBLAST STEM CELLS

Trophoblast stem (TS) cells represent a cell culture model for the trophectoderm lineage and its derivatives and are useful for modeling placental development. Specifically, TS cells are thought to stem from the polar TE, the part of the trophectoderm in direct contact with the ICM. The polar TE is a stem cell pool, which is maintained until postimplantation stages and fuels the continued production of tissue. It is suggested to give rise to cells of the extraembryonic ectoderm (ExE), the ectoplacental cone (EPC) and secondary trophoblast giant cell (Simmons and Cross, 2005; Tanaka et al., 1998; Uy et al., 2002). Accordingly, when injected into blastocysts, TS

cells only give rise to derivatives of the trophoblast lineage, such as the ExE, EPC and eventually the placenta (Tanaka et al., 1998).

Trophoblast cells in vivo require FGF signaling through FGF4/FGFR2 for their proliferation. In turn, the derivation of TS cells can be accomplished from outgrowths of blastocysts or explanted ExE isolated from postimplantation embryos and by supplementing the derivation medium with FGF4, the FGF receptor binding enhancer heparin and mouse embryonic fibroblast-conditioned medium (Tanaka et al., 1998).

TS cells express characteristic markers of the trophoblast lineages, like *Cdx2, Gata3* and *Eomes* and are devoid of OCT4. Remarkably, overexpression of *Cdx2* or deletion of *Oct4* in ES cells leads to their conversion to TS like cells, underlining these factors' leading role in regulation of trophectoderm specification (Nichols et al., 1998; Niwa et al., 2000; 2005).

EXTRAEMBRYONIC ENDODERM STEM CELLS

Extraembryonic endoderm stem (XEN) cells, like ES cells, are derived from the ICM of blastocyst stage embryos. They are a morphologically heterogeneous cell type, comprising epithelial-like as well as refractile cells, and represent an *in vitro* model of the primitive endoderm lineage and its derivatives, the visceral (VE) and parietal endoderm (PE) (Kunath, 2005). They were first derived comparatively recently and express markers of PrE, VE and PE, like *Pdgfra*, *Foxa2*, *Afp*, *Gata6*, *Gata4*, *Sox17* and *Sox7* (Kunath, 2005). However, when injected into blastocysts, XEN cells mainly contribute to the parietal endoderm and only rarely to the visceral endoderm. Their developmental potential may therefore be biased towards the PE (Brown et al., 2010; Kunath, 2005). However, conversion of XEN cells to a visceral endoderm-like state is possible through treatment with the transforming growth factor beta (TGF□) ligand Nodal or the growth factor-Cripto/FRL-1/Cryptic co-receptor Cripto as well as with Bmp4, which *in vivo* are involved in the formation and maturation of multiple visceral endoderm derivatives (Artus et al., 2011; Julio et al., 2011).

Derivation of XEN cells from the blastocyst can be achieved in different ways leaning strongly on protocols described previously for ES and TS derivation (Kunath, 2005). Accordingly, either blastocyst outgrowths are allowed to form in culture similarly to ES and TS cell derivation procedures or ICMs are isolated through microsurgery as has been employed for ES cell derivations. Both culture conditions, the addition of LIF as for ES cells, or FGF4 as for TS cells can successfully be used to derive XEN cells. It will be important for future studies utilizing these cells, to establish more standardized and optimized protocols that point out modifications to select for XEN cell propagation over TS or ES cells in these cultures.

Additionally it has been shown that treatment of ES cells with retinoic acid or formation of embryoid bodies can induce ES cells to trans-differentiate into XEN-like cells (Artus et al., 2010; Capo-Chichi et al., 2005; Coucouvanis and Martin, 1999; Soprano et al., 2007). Importantly, overexpression of the transcription factor *Gata6* directs ES cells to a XEN like state, hinting at the role of this marker in the formation of the extraembryonic endoderm lineages) (Fujikura et al., 2002; Shimosato et al., 2007).

1.5. AIMS OF THIS WORK

Preimplantation development is a tightly timed and regulated process. By the time of implantation of the embryo into the uterine wall, the extraembryonic lineages, trophectoderm and primitive endoderm, need to be properly formed for the embryo to be able to implant and develop. Concomitantly, the epiblast has to be specified and prepared to allow ready commitment to the somatic tissues upon differentiation cues. The establishment of the pluripotent epiblast and the primitive endoderm during the second cell fate decision in the mammalian embryo is therefore extensively studied. However, an intricate and interdependent network of transcription factors and signaling components are involved in the regulation of this process and the exact mechanisms of their roles and interactions are yet to be elucidated.

One potential master regulator for the second cell fate decision, GATA6, was reported in discordant roles and the main aim of this work was to elucidate its role in the second cell fate decision and its integration with known regulatory mechanisms. To this end, we analyzed in detail the *Gata6* null mutant at preimplantation stages and developed a computational analysis pipeline, which enabled us to achieve unbiased, high throughput single cell quantitative protein expression measurments.

The initiation of the second cell fate decision involves progressive transition from a homogenous to a heterogeneous cell population. A further aim of this work was to gain insights into possible mechanisms that could break this symmetry. We therefore analyzed the influence of relative transcription factor levels on PrE/EPI cell fate choice as well as their spatial expression patterns inside the ICM.

This cell fate choice between EPI and PrE establishes the pluripotent state *in vivo*. Among the multitude of possible factors influencing the emergence as well as the stabilization of the pluripotent epiblast lineage, is epigenetic regulation. We specifically set out to determine the possible role of DNA demethylation through AID on the pluripotent state, by analyzing its influence on iPS reprogramming, as well as ES derivation efficiency.

To ease future studies on preimplantation development as well as studies on ES cells and the multitude of potential applications they hold, we further aimed to develop improved and standardized protocols for the derivation of XEN and ES cells from mouse blastocysts.

2. RESULTS

PUBLICATION I

GATA6 LEVELS MODULATE PRIMITIVE ENDODERM CELL FATE CHOICE AND TIMING IN THE MOUSE BLASTOCYST

Schrode N., Saiz N., Di Talia S., Hadjantonakis AK.

Developmental Cell
29(4):454-67
2014

CellPress

Developmental Cell
Article

GATA6 Levels Modulate Primitive Endoderm Cell Fate Choice and Timing in the Mouse Blastocyst

Nadine Schrode,[1] Néstor Saiz,[1] Stefano Di Talia,[2] and Anna-Katerina Hadjantonakis[1,*]
[1]Developmental Biology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
[2]Department of Cell Biology, Duke University Medical Center, Durham, NC 27710, USA
*Correspondence: hadj@mskcc.org
http://dx.doi.org/10.1016/j.devcel.2014.04.011

SUMMARY

Cells of the inner cell mass (ICM) of the mouse blastocyst differentiate into the pluripotent epiblast or the primitive endoderm (PrE), marked by the transcription factors NANOG and GATA6, respectively. To investigate the mechanistic regulation of this process, we applied an unbiased, quantitative, single-cell-resolution image analysis pipeline to analyze embryos lacking or exhibiting reduced levels of GATA6. We find that *Gata6* mutants exhibit a complete absence of PrE and demonstrate that GATA6 levels regulate the timing and speed of lineage commitment within the ICM. Furthermore, we show that GATA6 is necessary for PrE specification by FGF signaling and propose a model where interactions between NANOG, GATA6, and the FGF/ERK pathway determine ICM cell fate. This study provides a framework for quantitative analyses of mammalian embryos and establishes GATA6 as a nodal point in the gene regulatory network driving ICM lineage specification.

INTRODUCTION

The first cell differentiation events during mammalian development result in the segregation of two extraembryonic lineages, the trophectoderm (TE) and the primitive endoderm (PrE), from the pluripotent epiblast (EPI), the founder tissue of most of the embryo proper (Saiz and Plusa, 2013; Schrode et al., 2013). These three cell types, found in the late blastocyst, are thought to arise through two sequential rounds of binary cell fate decisions. In the mouse, the first cell fate decision begins at the 8- to 16-cell stage, when the morula undergoes compaction and cells on the surface acquire apicobasal polarity, eventually becoming TE (Johnson and Ziomek, 1981). The second decision involves scattered cell differentiation within the inner group of cells, referred to as the inner cell mass (ICM), followed by cell sorting, and results in the differentiation of PrE and EPI lineages.

GATA6 and NANOG are the earliest markers of the PrE and EPI lineages, respectively; however, they are coexpressed in all ICM cells at the early blastocyst stage (32–64 cells) (Plusa et al., 2008). As embryos develop, individual ICM cells acquire exclusive GATA6 or NANOG expression in an apparently stochastic manner, which is thought to reflect the specification of PrE and EPI fates (Chazaud et al., 2006; Plusa et al., 2008). This process, proposed to be mediated both by stimulation of the fibroblast growth factor (FGF)/extracellular signal-regulated protein kinase (ERK) pathway (Chazaud et al., 2006; Kang et al., 2013a; Krawchuk et al., 2013; Nichols et al., 2009; Yamanaka et al., 2010) and by reciprocal repression between GATA6 and NANOG (Singh et al., 2007), results in a salt-and-pepper distribution of PrE and EPI precursors in the ICM of mid blastocysts (64–100 cells) (Chazaud et al., 2006). This scattered lineage specification is followed by the sorting of PrE precursors to the surface of the ICM in the late blastocyst (>100 cells), where they come to form an epithelium separating the blastocyst cavity and the EPI (Meilhac et al., 2009; Plusa et al., 2008; Saiz et al., 2013). However, the molecular mechanisms and gene regulatory networks governing the specification of PrE and pluripotent EPI within the ICM of the early blastocyst are still poorly understood.

In the mouse, FGF signaling is critical for PrE specification (Chazaud et al., 2006; Goldin and Papaioannou, 2003; Kang et al., 2013a; Krawchuk et al., 2013; Nichols et al., 2009; Yamanaka et al., 2010). FGF4 and FGFR2 are reciprocally expressed in EPI- and PrE-biased cells, respectively. FGF4, produced by EPI-biased cells (Guo et al., 2010; Ohnishi et al., 2014), has been proposed to activate the FGF/ERK pathway in neighboring cells, leading to the downregulation of NANOG and induction of the PrE program. However, the mechanism has not yet been experimentally addressed. Importantly, single-cell gene expression profiling studies coupled with the analysis of mutants (Kang et al., 2013a; Ohnishi et al., 2014) have suggested that FGF signaling is not required for the initial establishment of the gene regulatory network (GRN) in ICM cells but is essential for cells to exit this immature multilineage priming state and differentiate into EPI or PrE.

GATA6, a member of the GATA family of zinc-finger transcription factors, is the earliest PrE marker expressed in the early mouse embryo (Morrisey et al., 1996). It is first detected at around the 8-cell stage in all blastomeres, and by the mid blastocyst (~64-cell stage) it is restricted to PrE progenitors (Chazaud et al., 2006; Plusa et al., 2008). Ectopic expression of GATA6 in mouse embryonic stem (ES) cells is sufficient to direct them to a PrE-like state (Artus et al., 2010; Fujikura et al., 2002; Shimosato et al., 2007). GATA6 therefore likely acts near the top of the hierarchy regulating PrE development. However, the position of GATA6 relative to NANOG and the FGF/ERK pathway in the GRN driving ICM cell fate specification in vivo remains to be established.

In this study, we have undertaken a quantitative, single-cell-resolution analysis to understand the process of PrE segregation from the pluripotent EPI and begin to mechanistically decipher

the networks in which GATA6 engages to regulate this event. To investigate the role of GATA6 in ICM development, we have analyzed a wild-type, heterozygote, and null mutant *Gata6* allelic series (Sodhi et al., 2006) using automated nuclear segmentation (Lou et al., 2014) followed by single-cell-resolution quantitative three-dimensional (3D) image analyses.

Our results demonstrate that the early spatial pattern of differentiation of PrE versus EPI precursors is stochastic, and that spatial order emerges gradually at later stages. GATA6 is required for PrE cell fate specification and for the execution of the PrE program. *Gata6* null mutant embryos lack a PrE entirely, and exhibit pan-ICM expression of the pluripotency-associated factors NANOG, OCT4, and SOX2. In *Gata6* heterozygotes the proportion of ICM cells adopting a PrE fate is reduced and their commitment is decelerated, such that the period of time over which ICM cells make a PrE fate choice is extended. Exposure to exogenous FGF4 failed to restore PrE precursors within *Gata6* null mutant embryos, indicating that GATA6 is required for activation of the PrE program and the concomitant down-regulation of *Nanog* induced by FGF4. Collectively, our findings place GATA6 at the top of the hierarchy regulating PrE specification.

RESULTS

Cell fate choice is, in large part, determined by the action of key lineage-specific transcription factors. PrE and EPI lineage specification within the ICM of the mouse blastocyst appears to be undertaken in a stochastic manner. A sequence of events involving lineage specification and subsequent positional segregation has been defined. It involves the initial coexpression of factors within all ICM cells and progressive restriction of gene expression to lineage precursors, followed by a combination of cell sorting and cell death to refine their position (Artus et al., 2013; Chazaud et al., 2006; Gerbe et al., 2008; Meilhac et al., 2009; Plusa et al., 2008). Within this emergent mechanistic framework, GATA6 is the earliest expressed PrE-specific transcription factor, whereas NANOG is the earliest expressed EPI-specific transcription factor. However, these factors are initially coexpressed within the ICM and so are only markers once they become mutually exclusive, and thus this initiation and transition in marker localization are likely to be key to understanding the establishment of respective PrE and EPI fates.

A Pipeline for Single-Cell-Resolution Quantitative Analysis of Expression and Position: Progressive Distribution of GATA6 and NANOG

A rigorous mechanistic understanding of how single cells can operate coordinately to produce global effects relies on methods to resolve single-cell-resolution information in the context of a population. Thus far, attempts at single-cell analysis of cell fate decisions in preimplantation mammalian embryos have been hindered by time-consuming manual data processing at a small scale. To decipher the details of the GRN operating within the ICM, we assembled an unbiased single-cell-resolution analysis pipeline. This pipeline comprised software specifically developed for automated nuclear segmentation of 3D image data of mouse preimplantation-stage embryos (Lou et al., 2014), followed by quantitative fluorescence and spatial data analyses. The highly accurate segmentation afforded by our pipeline facilitates single-cell-resolution, large-scale comparisons of protein concentrations represented by fluorescence intensities after immunostaining and confocal imaging (Figure 1A). In this way, an analysis could be undertaken at the level of the entire ICM, taking into account all cells within each embryo analyzed.

A Method for Unbiased Assignment of Cell Fate

We performed immunohistochemistry using antibodies directed against GATA6 and NANOG proteins on wild-type embryos at several successive stages. As expected, wild-type morulae and early blastocysts exhibited coexpression of GATA6 and NANOG, which resolved in a mutually exclusive expression pattern at the mid blastocyst and in sorted lineages by the late blastocyst (Figures 1B and 2A). Applying our segmentation pipeline, we extracted fluorescence intensities of GATA6 and NANOG in each nucleus, estimated nuclear concentration by subtracting background fluorescence, and corrected for attenuation of fluorescence signals by tissue depth. Analysis of nuclear concentrations as a function of stage (32–64 cells, 64–128 cells, and >128 cells) revealed that at the early stage most cells expressed high levels of both GATA6 and NANOG. At the 64- to 128-cell stage, however, most cells expressed either high GATA6 and low NANOG or vice versa. At the last stage analyzed, all cells were either GATA6 positive and NANOG negative or NANOG positive and GATA6 negative (Figures 1B and 1C). We used the nuclear concentrations at this stage to empirically deduce a range over which cells could be defined as either EPI or PrE. This procedure provided an unbiased, reproducible, and quantitative method to assign fates to each cell of any given embryo. It allowed us to assign cell fate (Figure 1D), based on the automatic threshold procedure we developed (Figure 1B), in the absence of arbitrary factors, such as human user error.

The Early Spatial Pattern of Differentiation of PrE versus EPI Precursors Is Stochastic, whereas Spatial Order Emerges Gradually at Later Stages

We analyzed the pattern of emergence of cell differentiation to determine the importance of stochastic effects versus cell-cell communication or positional effects. We sought to determine whether there are spatial correlations between the fates of cells and/or the expression of NANOG and GATA6. Such correlations would suggest that cell-cell interaction/communication or position within the embryo plays a prominent role in early differentiation. Alternatively, a lack of correlation would suggest that stochastic, cell-autonomous processes are the main initial determinants of fate choice. To distinguish between these two scenarios, we computed the Pearson's correlation coefficient as a function of cell-cell distance. The coefficient measures the strength of the relationship between two variables (for example, cell fates). The coefficient can assume values between −1 (a perfect negative relationship) and 1 (a perfect positive relationship). A correlation coefficient of 0 implies that the two variables are independent, so that the status of one variable does not inform the status of the other. To compute the correlation coefficient, we binned data based on cell-cell distance choosing bin sizes of about 10 μm, an estimate of cell diameter, ensuring several data points present in each bin.

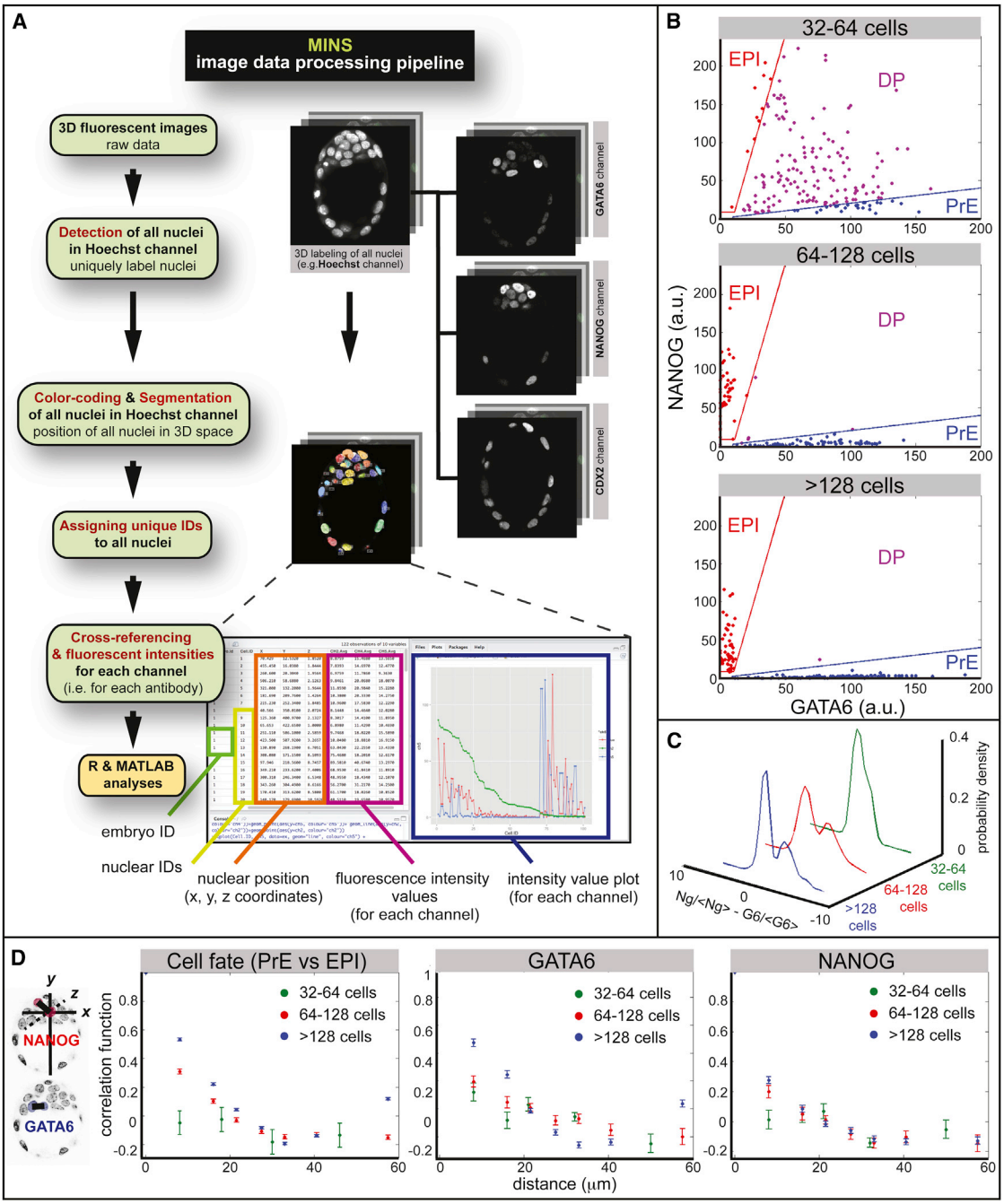


Figure 1. Quantitative Analysis of GATA6 and NANOG Expression

(A) Image data-processing pipeline incorporating MINS software.

(B) Nuclear concentration of GATA6 and NANOG at different stages. Blue and red lines show the threshold function used to define cells as PrE and EPI precursors, respectively. Cells that do not fall into either group are coexpressing NANOG and GATA6, and are classified as DP (double-positive, uncommitted ICM) cells. a.u., arbitrary units.

(C) Probability density of normalized NANOG concentration minus normalized GATA6 concentration (<Ng> and <G6> indicate the average concentrations of NANOG and GATA6, respectively). Earliest stage (defined by total cell number) cells express both GATA6 and NANOG (unimodal distribution), but at the "salt-and-pepper" stage, cells are either GATA6 positive or NANOG positive, resulting in a bimodal distribution.

(D) Correlation coefficient of cell fates and levels of GATA6 and NANOG as a function of cell-cell distance for three different stages. Correlation = 0, cells are making their fate choice independent of other cells, and no fate prediction can be made; correlation = 1, certain prediction that cells with close distance are alike; correlation = −1, certain prediction that cells with close distance are unalike. Error bars represent standard errors, estimated using Fisher transformation.

Cell fates (EPI, PrE, or undifferentiated ICM) were assigned in an automatic unbiased fashion (Figure 1B). We computed the correlation coefficient as a function of distance for differentiated cells (i.e., EPI and PrE), with undifferentiated cells excluded from the calculation. We found that at the earliest stage (~32–64 cells) there was no correlation in the spatial pattern of cell differentiation (Figure 1D), as expected for a random cell-autonomous decision (Figure S1A available online). Consistently, the position of cells that differentiated early was independent of the embryo's coordinates, such as distance from the blastocyst cavity (Figure S1B). This random pattern of early differentiation likely reflected a lack of spatial correlation in the levels of NANOG and GATA6 at this stage (measured across all cells), and suggested that the earliest stages of cell specification are dominated by cell-autonomous stochastic factors. At the salt-and-pepper stage (~64–128 cells), a pattern of correlation began to emerge in both cell fate and NANOG and GATA6 levels, so that adjacent cells had a slightly positive chance of exhibiting the same fate (Figure 1D). This could, in principle, originate from the previous uncorrelated pattern, if cells retained similar expression levels and/or fate upon division. At the last stage analyzed (>128 cells), after cells had sorted, a clear spatial correlation was present between cell fates and the levels of both GATA6 and NANOG (Figure 1D). Importantly, the dependency of the correlation coefficient on cell-cell distance at this stage was similar to that expected, given the geometry of the embryo and sorted cell fates (see the comparison between a "virtual" embryo and our data; Figure S1). Collectively, these results are consistent with a model in which the earliest steps of cell differentiation are dominated by stochastic fluctuations of NANOG and GATA6, which later resolve into a clear spatial pattern as a result of cell sorting. Finally, we observed that the correlation in NANOG was slightly lower than for GATA6 (many EPI cells at the >128-cell stage were already downregulating NANOG expression), probably as a result of its highly transient expression and its extinguishment coincident with embryo implantation.

Gata6 Mutants Exhibit Pan-ICM NANOG Expression

GATA6, like NANOG, is initially expressed by all ICM cells, but thereafter becomes exclusive to the PrE (Plusa et al., 2008). An antagonistic interaction of GATA6 with NANOG has been proposed to be at the core of the GRN regulating PrE versus EPI specification (Frankenberg et al., 2011). Thus, we hypothesized that if GATA6 acts near the top of the hierarchy regulating PrE cell fate specification, its elimination must cause a profound defect in lineage specification within the ICM. We therefore determined the phenotype arising from the complete loss, as well as reduction, of GATA6 in *Gata6*$^{-/-}$ mutants and *Gata6*$^{+/-}$ heterozygotes, respectively. By contrast to wild-type stage-matched embryos, *Gata6*$^{-/-}$ embryos expressed NANOG starting at the 8-cell stage in all cells, and thereafter throughout the ICM (Figure 2A). Notably, GATA6 was not detected at any stage analyzed, indicating that the allele is a protein null and that the zygotic ablation of *Gata6* produces no protein, and suggesting no maternal mRNA is transcribed and/or translated into detectable protein (Figure 2A). To compound this observation, we analyzed *Gata6* maternal zygotic mutant embryos. These exhibited an equivalent phenotype to *Gata6* zygotic mutant embryos, referred to as *Gata6*$^{-/-}$ throughout the text (Figure S2A).

Because *Gata6* mutant embryos have been reported as exhibiting a defect at postimplantation stages, affecting the cardiac mesoderm or visceral endoderm (Koutsourakis et al., 1999; Morrisey et al., 1996), we collected embryos from *Gata6*$^{+/-}$ intercrosses at embryonic day (E)5.5. We failed to recover any *Gata6*$^{-/-}$ embryos among a total of 42 embryos obtained from seven litters at E5.5 (Table S1). Immunofluorescence revealed that both wild-type and heterozygous embryos exhibited a normal morphology, with a GATA4-positive visceral endoderm layer (Figure S2B). The apparent discrepancy between our observation of a preimplantation defect, which never gave rise to egg cylinder-stage embryos, and previous studies reporting early postimplantation defects could be attributed to allele- or strain-specific differences (Koutsourakis et al., 1999; Morrisey et al., 1996).

Cells in *Gata6* Mutant ICMs Do Not Upregulate NANOG Expression

Previous studies have proposed that GATA6 and NANOG function, at least in part, through mutual repression (Frankenberg et al., 2011; Singh et al., 2007). Such a mutual antagonism could explain the mutually exclusive expression of markers and the salt-and-pepper distribution of cells committed to PrE versus EPI lineages. In accordance with such a model of mutual repression, the absence of GATA6 could relieve *Nanog* from GATA6-mediated repression and result in elevated levels of NANOG. To determine whether this is the case in vivo, we quantified protein levels in all cells of wild-type, *Gata6*$^{+/-}$, and *Gata6*$^{-/-}$ embryos by measuring fluorescence intensities after immunostaining and confocal imaging using the analysis pipeline we constructed (Figure 1A).

Levels of NANOG protein were not elevated in the ICM cells of *Gata6*$^{-/-}$ embryos (n = 364 cells/12 embryos), compared to NANOG-positive ICM cells in *Gata6*$^{+/-}$ or wild-type embryos (n = 476 cells/25 embryos and n = 323 cells/17 embryos, respectively) (Figure 2B), suggesting that some distinct or additional mechanism must control NANOG levels within the ICM in vivo. Interestingly, we noted that in wild-type embryos, the levels of NANOG exhibited a broad dynamic range in the early blastocyst but became more homogeneous as development proceeded, accompanying the maturation of the EPI lineage (Figure 2B). By contrast, *Gata6*$^{-/-}$ embryos exhibited a decreased dynamic range of NANOG levels earlier, which was comparable to that observed in mature EPI cells in wild-type late blastocysts. These data therefore suggest that in the absence of GATA6, all ICM cells prematurely commit to an EPI fate.

GATA6 Is Involved in NANOG Repression in TE Cells

In addition to pan-ICM expression of NANOG, we also noted a significantly increased number of NANOG-positive cells in the TE of *Gata6*$^{-/-}$ embryos compared to wild-type embryos at early and mid blastocyst stages, which appeared to recover by the late blastocyst stage (Figure 2C). These findings suggest that GATA6 functions, at least in part, to actively repress NANOG in TE cells. TE cells ectopically expressing NANOG also coexpressed TE markers such as CDX2, GATA3, and EOMES (Figure 2C). These TE cells appeared morphologically normal, and blastocyst outgrowths of *Gata6*$^{-/-}$ embryos showed expansion of the TE and differentiation of trophoblast giant cells

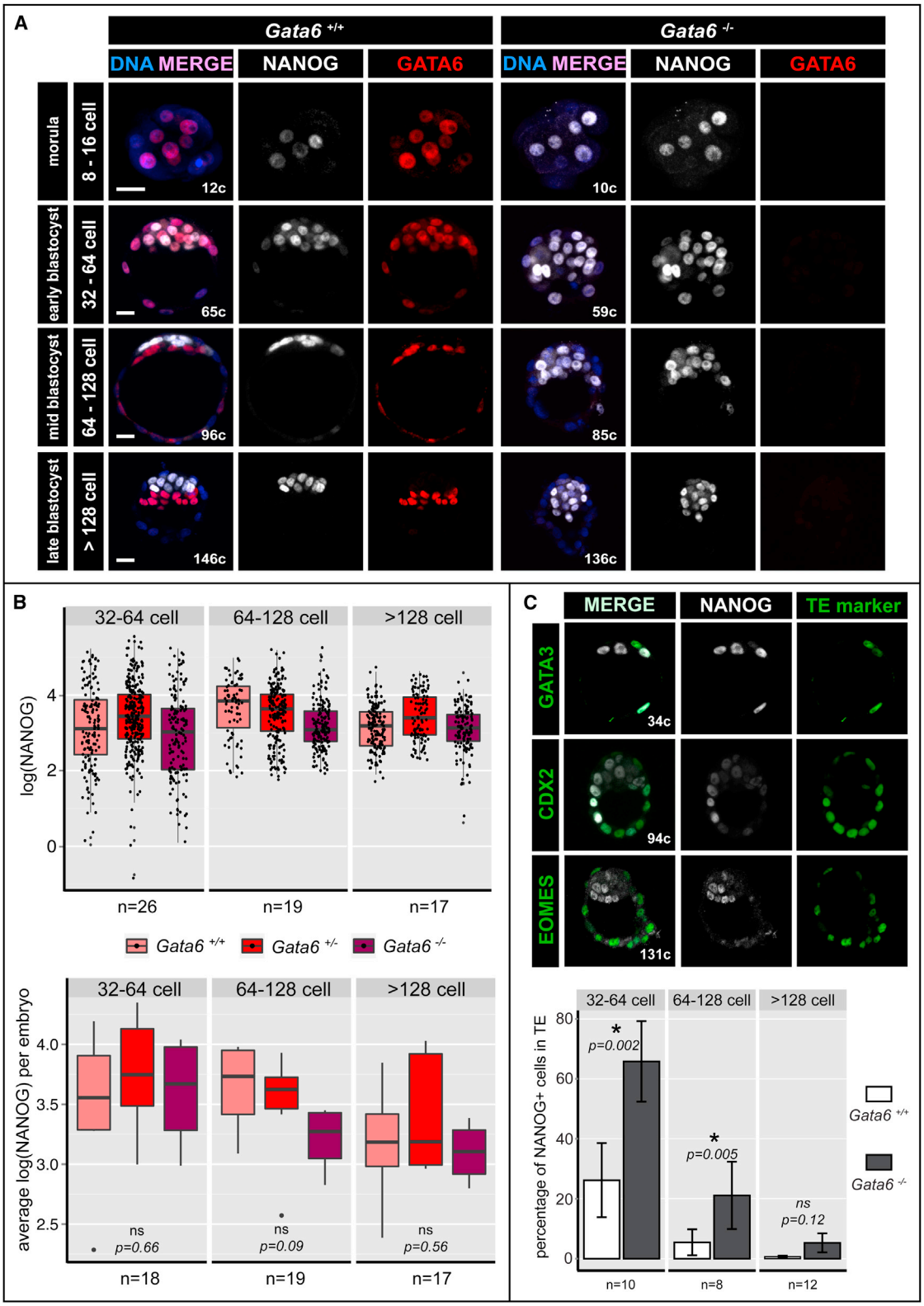


Figure 2. NANOG Is Expressed Ectopically in *Gata6* Mutants

(A) Localization and distribution of NANOG (white) and GATA6 (red) in wild-type (*Gata6*$^{+/+}$) and *Gata6* mutant mouse embryos from the morula to late/implanting blastocyst stage. Blue in merge: Hoechst. The scale bars represent 20 μm. c, cells.

(legend continued on next page)

(Figure S2C). Additionally, although postimplantation *Gata6*$^{-/-}$ embryos could not be recovered, we noted Mendelian ratios of empty implantation sites in uteri dissected after E5.5. Collectively, these data suggest that, despite the ectopic expression of NANOG, the TE is functional in *Gata6*$^{-/-}$ embryos.

The PrE Lineage Is Not Specified in the Absence of GATA6

PrE formation is accompanied by the sequential activation of a number of factors regulating cell fate specification, proliferation, and polarity, such as PDGFRA, SOX17, and GATA4, among others (Artus et al., 2011; Gerbe et al., 2008; Kurimoto et al., 2006; Morris et al., 2010; Niakan et al., 2010; Plusa et al., 2008; Saiz et al., 2013). Having determined that *Gata6* mutants exhibit pan-ICM NANOG expression, we wanted to investigate whether the PrE program is activated in the absence of GATA6. To this end, we determined the expression of these secondary PrE markers in *Gata6*$^{+/+}$, *Gata6*$^{+/-}$, and *Gata6*$^{-/-}$ embryos. Embryos lacking GATA6 failed to activate *Pdgfra* expression, as assessed in *Gata6*$^{-/-}$*; Pdgfra*$^{H2B\text{-}GFP/+}$ embryos (Movie S1), and exhibited no SOX17 (n = 13 embryos) or GATA4 (n = 4 embryos) expression at either mid (64–100 cells) or late blastocyst (>128 cells) stages (Figure 3A).

At the late blastocyst stage, PrE cells polarize and form an epithelium (Smyth et al., 1999). Markers of apicobasal polarity, including aPKC and DAB2, become localized to the apical surface of PrE cells, revealing their progressive epithelial maturation (Gerbe et al., 2008; Saiz et al., 2013). By contrast to wild-type embryos, we found that *Gata6* mutants failed to apically express both aPKC and DAB2 (Figure 3A), indicating that their ICM cells not only failed to upregulate PrE markers but also were unable to form an epithelium, one of the defining features of the PrE (Saiz et al., 2013).

Because all ICM cells in *Gata6*$^{-/-}$ embryos expressed NANOG, we assessed whether ICM cells in these embryos adopted an EPI identity in the absence of GATA6. We found that the pluripotency-associated factors SOX2 and OCT4 were expressed in all ICM cells in *Gata6*$^{-/-}$ embryos at all blastocyst stages, whereas they were restricted to the EPI in control *Gata6*$^{+/+}$ or *Gata6*$^{+/-}$ late blastocysts (Figure 3B), thus suggesting that all ICM cells had acquired EPI identity in the absence of GATA6.

Of note, in 9 out of 20 *Gata6*$^{-/-}$ blastocysts analyzed, we identified between 1 and 3 cells on the surface of the ICM, out of a total of between 15 and 40 cells, that did not express NANOG, and thus were GATA6/NANOG double negative. Surprisingly, these cells expressed the TE marker CDX2 (Figure 3C), which could suggest that in the absence of both NANOG and GATA6, ICM cells might acquire an alternative cell fate resulting in their upregulation of TE markers such as CDX2 if positioned adjacent to the blastocyst cavity. However, when embryos were allowed to develop in vitro this cell population did not expand over time, suggesting that CDX2-positive GATA6/NANOG-negative cells were either transient or the result of delayed CDX2 downregulation in some ICM cells, although the latter might also be expected to be CDX2/GATA6 and/or NANOG positive.

One mechanism to ensure balanced lineage representation and proper allocation of progenitors within the ICM of the blastocyst is apoptosis (Plusa et al., 2008; Artus et al., 2013). To determine whether the lack of PrE cells was due to selective apoptosis at an early stage or to a cell fate switch, we calculated the average total cell number per embryo and the relative distribution of cells to the TE and ICM compartments. We noted that neither embryo size nor relative ICM size differed significantly between wild-type (n = 17) and *Gata6*$^{-/-}$ embryos (n = 12) at any stage (Figures 3D and 3E), suggesting there was no increase in cell death in the absence of GATA6.

Collectively, our data demonstrate that, in the absence of GATA6, ICM cells failed to adopt a PrE fate, and a PrE epithelial layer was altogether absent. Absence of GATA6 results in absence of the PrE lineage coupled to acquisition of EPI fate.

GATA6 Levels Influence Both the Timing and Frequency of PrE Cell Specification

Because GATA6 is necessary for PrE specification, we wanted to explore whether the amount of GATA6 protein present in individual ICM cells has any influence on their fate choice. To address this question, we quantified PrE and EPI cell numbers as well as GATA6 expression levels in wild-type and *Gata6*$^{+/-}$ embryos. Quantification of fluorescence intensity levels revealed a significant reduction in GATA6 protein in *Gata6*$^{+/-}$ PrE cells (n = 444 cells/25 embryos) compared to wild-type cells (n = 470 cells/17 embryos) (Figure 4A).

Interestingly, *Gata6*$^{+/-}$ embryos displayed coexpression of GATA6 and NANOG within a subpopulation of ICM cells until later stages than wild-type embryos (Figure 4B), suggesting that cell fate specification is delayed or less abrupt when GATA6 levels are reduced. To distinguish between these two scenarios, and to further characterize the effects of reduced levels of GATA6 on the dynamics of cell fate specification, we quantified the frequency of PrE and EPI cells as a function of cell number. This analysis revealed that in wild-type embryos the specification of both EPI and PrE cells occurs over a short period of time (until ~100-cell stage). In *Gata6*$^{+/-}$ embryos, specification of PrE cells was significantly retarded, indicating that wild-type GATA6 levels are required for rapid commitment to PrE cell fate (Figure 4C). Interestingly, specification of the EPI exhibited a similar abruptness, but it occurred at a slightly earlier time (Figure 4C). As a consequence of the slow commitment of PrE cells, undifferentiated ICM cells were observed at later stages than in wild-type embryos (Figures 4B and 4C). Furthermore, the number and proportion of PrE cells in *Gata6*$^{+/-}$ embryos were significantly reduced (n = 17) compared to

(B) Distribution of NANOG protein levels as measures of logarithm-transformed fluorescence intensity of each EPI cell (upper panel) and averaged for each embryo (lower panel) in NANOG-positive ICM cells of *Gata6*$^{+/+}$, *Gata6*$^{+/-}$, and *Gata6*$^{-/-}$ embryos of the stages indicated. Data are represented as Tukey box plots. ns, not significant.

(C) Localization and distribution of NANOG (white) and TE markers GATA3, CDX2, and EOMES (green) in *Gata6* mutant embryos at early, mid, and late blastocyst stages (upper panel). Percentage of NANOG-positive cells in the TE of *Gata6*$^{+/+}$ and *Gata6*$^{-/-}$ embryos at early, mid, and late blastocyst stages (lower panel). Error bars represent standard errors. Asterisks indicate statistical significance: *$p < 0.05$.

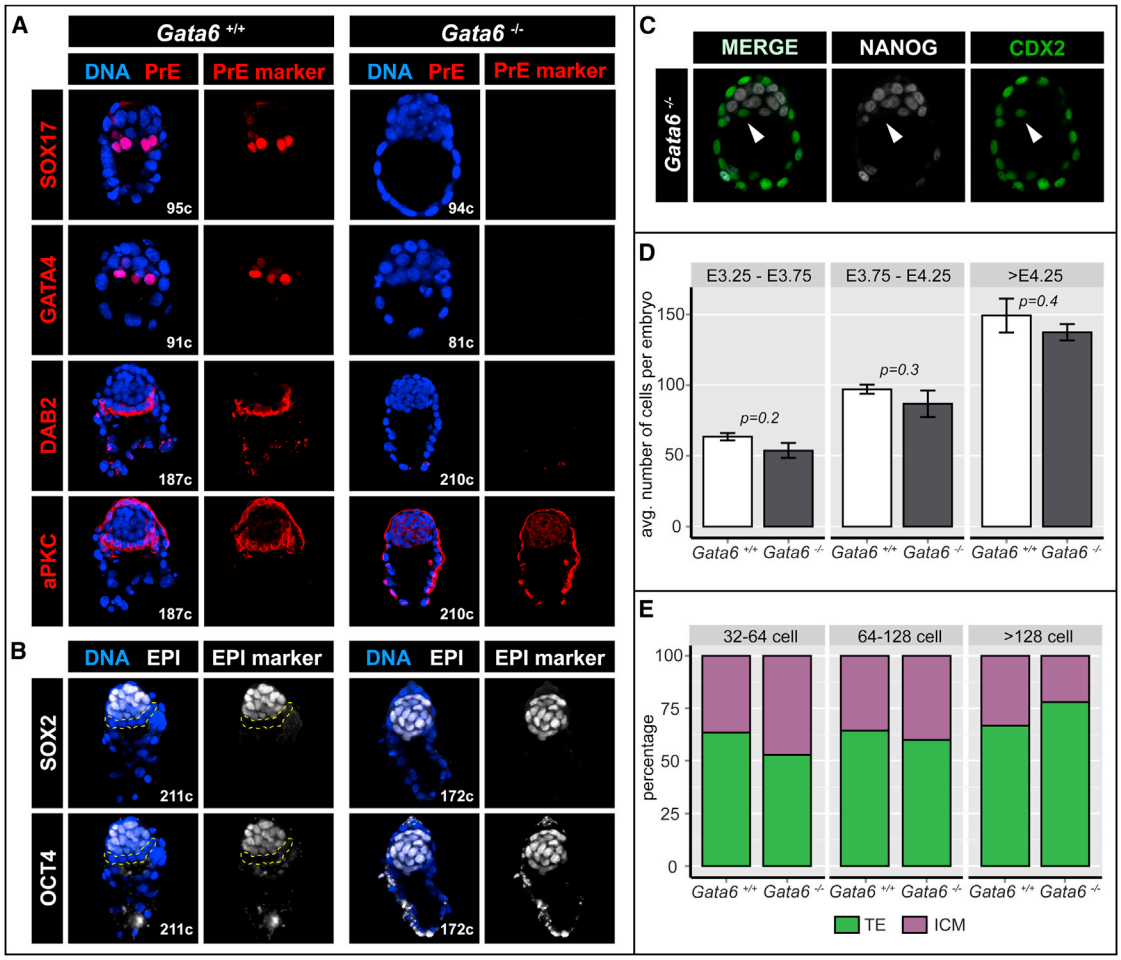


Figure 3. PrE Cannot Be Specified in the Absence of GATA6

(A) Localization and distribution of PrE markers SOX17 and GATA4 and epithelial markers DAB2 and aPKC (red) in wild-type (*Gata6*$^{+/+}$) and *Gata6*$^{-/-}$ embryos at the 64- to 128-cell and >128-cell stage, respectively. Blue in merge: Hoechst.

(B) Localization and distribution of EPI markers OCT4 and SOX2 (white) in wild-type (*Gata6*$^{+/+}$) and *Gata6*$^{-/-}$ embryos at the >128-cell stage. OCT4 staining in TE cells is cytoplasmic. Dashed yellow lines indicate the PrE layer. Blue in merge: Hoechst.

(C) Localization and distribution of NANOG (white) and CDX2 (green) in *Gata6* mutant mouse embryos at the 64- to 128-cell stage. Arrowheads indicate CDX2-expressing cells in the ICM.

(D) Average number of total cells in *Gata6*$^{+/+}$ and *Gata6*$^{-/-}$ embryos at early, mid, and late blastocyst stages. Error bars represent standard errors.

(E) Cell composition in *Gata6*$^{+/+}$, *Gata6*$^{+/-}$, and *Gata6*$^{-/-}$ embryos at the stages indicated. Cell numbers are shown as percentages of the total number of cells per embryo.

wild-type embryos (n = 25), whereas the number of EPI cells was increased (Figure 4B). Collectively, these data indicate that GATA6 protein concentration is important for PrE fate specification, and that a reduction leads to a deceleration in the rate of PrE fate specification as well as a reduction in the number of cells adopting a PrE fate within the ICM. Because *Gata6*$^{+/-}$ animals were viable, this reduction in the number of PrE cells did not appear to affect the function of the PrE lineage. It is possible that PrE cell numbers either recovered at postimplantation stages or that the PrE might tolerate cell number variability.

FGF Signaling Is Not Sufficient to Direct PrE Specification in the Absence of GATA6

There is increasing evidence that differential activation of the FGF/mitogen-activated protein kinase pathway in individual ICM cells determines EPI or PrE cell fate choice in mice (Chazaud et al., 2006; Nichols et al., 2009; Yamanaka et al., 2010). Our experiments have revealed a failure in PrE cell specification in *Gata6*$^{-/-}$ embryos, a phenotype that resembles that observed in mutants in which FGF signaling components have been perturbed (Chazaud et al., 2006; Kang et al., 2013a; Krawchuk et al., 2013) or in wild-type embryos treated with inhibitors of FGF signaling (Nichols et al., 2009; Yamanaka et al., 2010). In our previous analysis of *Fgf4* mutant embryos, we noted that GATA6 expression was initiated at the early blastocyst stage but failed to be maintained in the absence of FGF4, resulting in a lack of PrE cells (Kang et al., 2013a).

To determine the relative positions of FGF signaling and GATA6 in the hierarchy governing PrE specification, we treated

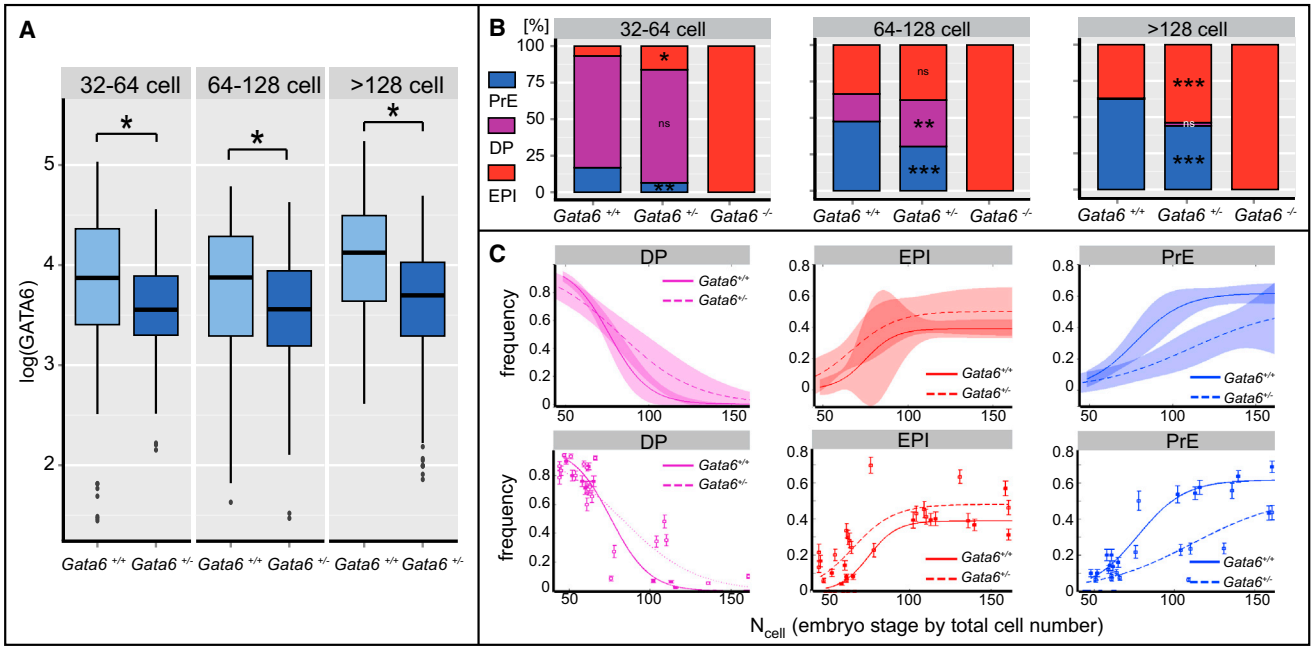


Figure 4. Protein-Level Analysis Reveals Dosage Effect in *Gata6*$^{+/-}$ Embryos

(A) GATA6 protein levels as measures of logarithmically transformed fluorescence intensity in GATA6-positive ICM cells of *Gata6*$^{+/+}$ and *Gata6*$^{+/-}$ embryos at the indicated stages. Data are represented as Tukey box plots.

(B) ICM composition in *Gata6*$^{+/+}$, *Gata6*$^{+/-}$, and *Gata6*$^{-/-}$ embryos. Cell numbers are shown as percentages of total ICM number. PrE: GATA6+; DP: NANOG+, GATA6+; EPI: NANOG+. Asterisks indicate statistical significance: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

(C) Frequency of cell fate specification as a function of cell number for undifferentiated ICM (DP) cells, EPI precursors, and PrE precursors. Probability functions were estimated by logistic regression. Shaded areas show 95% confidence intervals (upper panel). Data points were fitted using logistic regression. Error bars represent standard deviation estimated using binomial distribution. p values were computed using the χ^2 test. PrE induction is slower ($p < 10^{-5}$) and loss of ICM is also slower ($p < 10^{-4}$), whereas induction of EPI is faster ($p = 0.01$).

Gata6$^{-/-}$ embryos with saturating doses of FGF, which have been previously shown to induce a PrE identity throughout the ICM in wild-type embryos (Yamanaka et al., 2010) and *Fgf4* mutants (Kang et al., 2013a). Because timed treatments have revealed variations in the sensitivity to FGF signaling through preimplantation development (Frankenberg et al., 2011; Grabarek et al., 2012; Yamanaka et al., 2010), we used two regimes in a parallel set of experiments (Figures 5A and 6A). First, we applied exogenous FGF for 48 hr starting at E2.5 (8- to 16-cell stage), prior to the initiation of any differentiation, when NANOG and GATA6 are coexpressed (Figure 5A). *Gata6*$^{-/-}$ embryos failed to respond to doses of between 500 and 1,000 ng/ml of either FGF4 (Figures 5A and 5B) or FGF2 (Figure S3). They maintained NANOG expression throughout the ICM, with levels comparable to those of untreated embryos (Figures 5B and 5D). By contrast, wild-type embryos displayed pan-ICM GATA6 or SOX17 expression and complete absence of NANOG expression (Figures 5B–5D), as previously described (Saiz et al., 2013; Yamanaka et al., 2010). *Gata6*$^{+/-}$ embryos exhibited a comparable response to wild-type embryos, although some cells expressing NANOG could be identified (Figure 5C), suggesting that the kinetics of the effect produced upon FGF treatment was altered by a reduction in the dose of GATA6. FGF treatment also failed to induce expression of later PrE markers, such as SOX17, in *Gata6*$^{-/-}$ embryos (Figure 5B).

Collectively, these data led us to hypothesize that whereas FGF4 acts as an instructive signal to activate the PrE program, GATA6 is necessary to execute it. Thus, in the absence of GATA6, FGF treatment of embryos fails to activate the PrE program. Alternatively, *Gata6*$^{-/-}$ embryos may exhibit aberrant expression of FGF receptors on their ICM cells and fail to mount a response to an FGF signal.

ERK Inhibition Promotes NANOG Upregulation in the Absence of GATA6

Upon FGF/ERK stimulation of wild-type E2.5 embryos, NANOG is repressed and GATA6 is activated in all ICM cells (Yamanaka et al., 2010). If GATA6 were the factor directly mediating NANOG repression upon FGF/ERK stimulation, we would hypothesize that inhibition of the pathway in *Gata6*$^{-/-}$ embryos would have no effect on NANOG expression compared to untreated, *Gata6*$^{-/-}$ embryos. Culture in the presence of 1 μM ERK1/2 inhibitor PD0325901 (Nichols et al., 2009) resulted in a marked upregulation of NANOG throughout the ICM in all embryos, regardless of their genotype (Figure 5D). These data indicate that, whereas GATA6 is necessary for NANOG downregulation, an additional ERK1/2-mediated mechanism operates to maintain NANOG levels within EPI cells.

Collectively, these results demonstrate that GATA6 is necessary to cell-autonomously mediate FGF/ERK signaling upstream of ERK1/2, which in turn mediates NANOG repression.

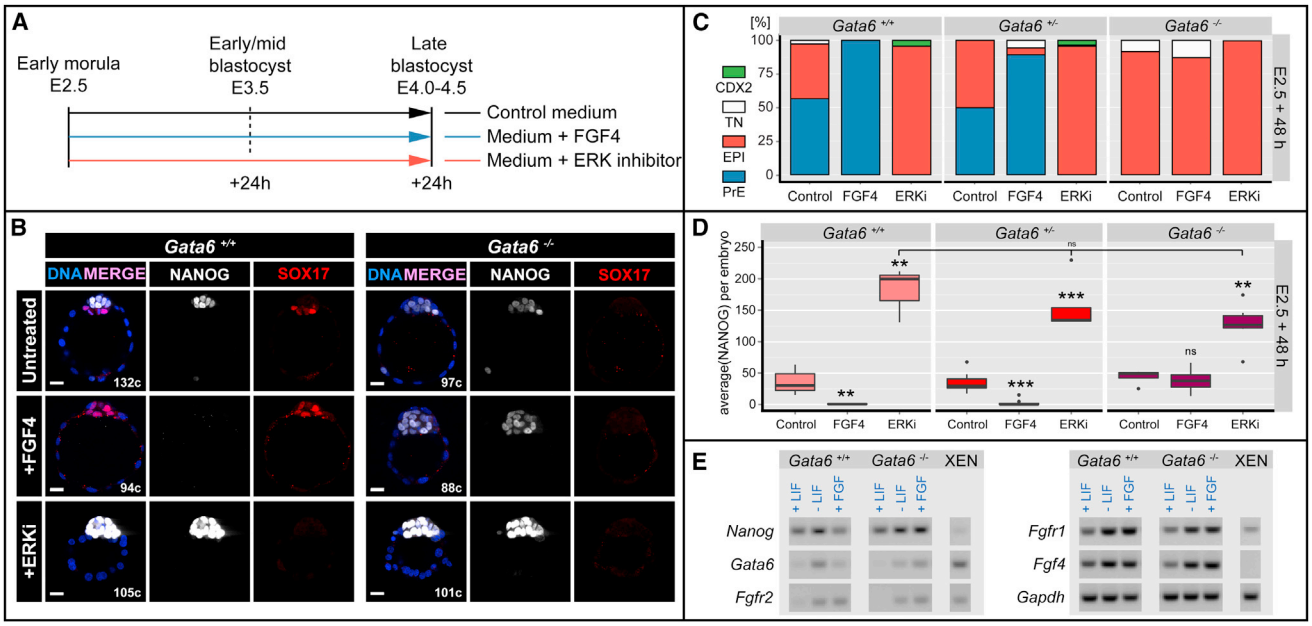


Figure 5. FGF4 Is Not Sufficient to Direct PrE Specification in the Absence of GATA6, whereas ERK Inhibition Results in NANOG Upregulation

(A) Experimental timeline.

(B) Representative immunofluorescence images of *Gata6*$^{+/+}$ and *Gata6*$^{-/-}$ embryos cultured from the morula (E2.5) stage for 48 hr in culture medium alone (untreated), medium + 1 μg/ml FGF4 + 1 μg/ml heparin (+FGF4), or medium + 1 μM PD0325901 (+ERKi). Blue: Hoechst; white: NANOG; red: SOX17. Scale bars represent 20 μm.

(C) ICM composition in *Gata6*$^{+/+}$, *Gata6*$^{+/-}$, and *Gata6*$^{-/-}$ embryos after culture as in (A). Cell numbers are shown as percentages of the total ICM number. TN, triple negative for NANOG, PrE marker, and CDX2.

(D) Fluorescence levels of NANOG as the measure of the average per embryo in *Gata6*$^{+/+}$, *Gata6*$^{+/-}$, and *Gata6*$^{-/-}$ ICMs after culture as in (A). Data are represented as Tukey box plots. Asterisks indicate statistical significance: **p < 0.01, ***p < 0.001. ns, not significant.

(E) RT-PCR on *Gata6*$^{+/+}$ and *Gata6*$^{-/-}$ ESCs after 48 hr culture in standard LIF-containing medium (+LIF), after LIF withdrawal (−LIF), and in the presence of 250 ng/ml FGF (+FGF).

Furthermore, the effects of ERK inhibition reveal a distinct role for FGF signaling in EPI cells, potentially acting to inhibit levels of NANOG so as to maintain them within a physiological range.

Modulation of FGF Signaling Cannot Induce an ICM Fate Switch after PrE and EPI Specification

Next, we cultured embryos in the presence of FGF4 for 24 hr from E3.5 (early-to-mid blastocyst stage), when PrE and EPI specification has already begun (Figure 6A). By contrast to FGF4 treatment experiments from E2.5, exposure to 500 ng/ml of FGF4 from E3.5 did not result in pan-ICM expression of GATA6 in either wild-type or *Gata6*$^{+/-}$ embryos (Figures 6B and 6C). Although we did observe a moderate expansion of the PrE compartment, most wild-type embryos retained a NANOG-positive population (Figure 6C). Similarly, *Gata6*$^{-/-}$ embryos exhibited no reduction in the size of the NANOG-positive compartment upon FGF4 treatment (Figures 6B and 6C). We observed a proportion of ICM cells that expressed neither NANOG nor PrE markers (GATA6 or SOX17), regardless of their genotype (triple negative). This population, albeit larger in *Gata6*$^{+/-}$ and *Gata6*$^{-/-}$ embryos, did not change significantly upon FGF treatment, suggesting it resulted from developmental downregulation of NANOG as the EPI matures, rather than a consequence of FGF treatment.

In agreement with these results, we observed that inhibition of ERK1/2 at E3.5 with 1 μM PD0325901 did not result in downregulation of GATA6 in wild-type or *Gata6*$^{+/-}$ embryos, as previously described (Nichols et al., 2009). On the other hand, consistent with our previous observations from incubations initiated at E2.5, ERK1/2 inhibition resulted in upregulation of NANOG throughout the EPI regardless of embryo genotype (Figure 6B).

These results suggest that desensitizing cells to FGF may be a mechanism operating to prevent changes in ICM lineage composition upon sustained FGF production. Such a model is supported by the fact that XEN stem cells, which represent the primitive endoderm lineage, can be derived from *Fgf4* mutants, demonstrating a temporally restricted requirement for FGF/ERK signaling within the PrE lineage (Kang et al., 2013b).

The Induction of *Gata6* Expression Is Independent of NANOG Repression

It has been proposed that NANOG directly represses *Gata6* (Mitsui et al., 2003; Singh et al., 2007) and that downregulation of NANOG through FGF signaling alleviates *Gata6* suppression, leading to its upregulation and differentiation of cells toward a PrE fate. It has been shown in embryonic stem cells (ESCs) that activation of FGF/ERK signaling directly represses *Nanog* transcription (Santostefano et al., 2012). To test whether this

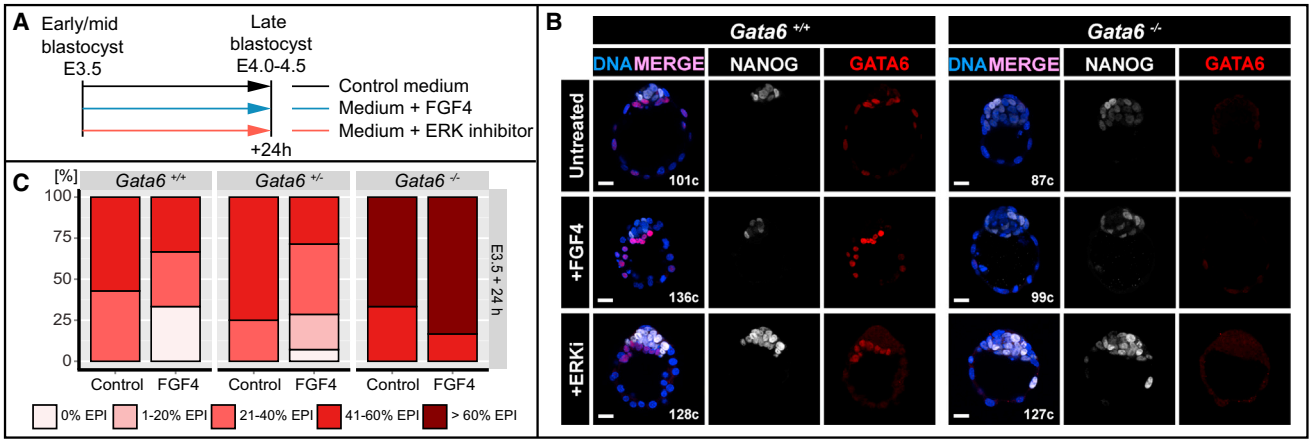


Figure 6. Changes in FGF Signaling Cannot Induce Cell Fate Switch after PrE and EPI Specification

(A) Experimental timeline.

(B) Representative immunofluorescence images of *Gata6*$^{+/+}$ and *Gata6*$^{-/-}$ embryos cultured from the salt-and-pepper (E3.5) stage for 24 hr in medium without (untreated) or with 500 ng/ml FGF4 + 1 μg/ml heparin (+FGF4) or medium + 1 μM PD0325901 (+ERKi). Blue: Hoechst; white: NANOG; red: GATA6. Scale bars represent 20 μm.

(C) Percentages of embryos falling into one of five FGF-response categories: 0% EPI, 1%–20% EPI, 21%–40% EPI, 41%–60% EPI, and >60% EPI in ICMs of embryos treated as in (A).

downregulation of NANOG leads to an upregulation of *Gata6* or whether FGF signaling can directly influence *Gata6* expression, we derived *Gata6*$^{+/+}$ and *Gata6*$^{-/-}$ ESCs (Czechanski et al., 2014) and analyzed gene expression after culture in the presence of FGF. Whereas undifferentiated *Gata6*$^{+/+}$ and *Gata6*$^{-/-}$ ESCs cultured in leukemia inhibitory factor (LIF) exhibited little or no expression of *Gata6*, *Gata6* transcripts were detected in *Gata6*$^{+/+}$ and *Gata6*$^{-/-}$ ESCs maintained in the presence of FGF (Figure 5E). As a control, *Gata6*$^{+/+}$ and *Gata6*$^{-/-}$ ESCs were also allowed to differentiate in the absence of LIF and without addition of FGF. These cells exhibited an upregulation of *Gata6* comparable to cells treated with FGF. These data suggest that *Gata6* is activated upon differentiation of ESCs in a GATA6-independent manner. Additionally, the absence of a downregulation of *Nanog* in the absence of GATA6 (Figures 5D and 5E) suggests that, if FGF signaling is involved in the upregulation of *Gata6*, it does so directly and independent of NANOG repression. However, because *Gata6* was detected both in the presence and absence of FGF, we were unable to exclude the possibility that this repression occurred independently of FGF signaling.

DISCUSSION

With the aim of obtaining unbiased quantitative information on individual cells to understand how they direct the behavior of a population within early mammalian embryos, we developed a pipeline involving single-cell-resolution, quantitative immunofluorescence image analysis, coupled with Cartesian coordinate information of cellular position. This allowed us to account for how the salt-and-pepper distribution of emergent PrE and EPI lineage progenitors arises within the ICM of the mouse blastocyst. The method we developed extends previous findings using single-cell microarray-based expression analyses (Ohnishi et al., 2014), as it also incorporates positional information for individual cells. Our results suggest that the early pattern of cell differentiation is purely stochastic, and might be dominated by cell-autonomous noisy expression of NANOG and GATA6. Following this initial phase, order starts to emerge, most likely as the result of cell division, cell motility, and the acquisition of apicobasal polarity (Gerbe et al., 2008; Meilhac et al., 2009; Saiz et al., 2013) as well as apoptosis of misplaced cells (Artus et al., 2013; Plusa et al., 2008). In future studies, it will be important to determine whether the early stochastic differentiation is propagated to later stages, or whether corrective feedback mechanisms operate to ensure that the embryo ends up with the correct number of EPI and PrE progenitors, despite an early phase of stochastic differentiation.

The distribution of lineage progenitors is reflected by the expression of the key lineage-specific transcription factors GATA6 and NANOG. Because it is the earliest expressed PrE factor, we reasoned that GATA6 might be responsible for directing ICM cell fate choice. To gain insight into the mechanisms driving cell fate decision away from pluripotency, we analyzed embryos exhibiting a reduction (heterozygotes) or absence (null mutants) of *Gata6*. Our studies reveal a strict GATA6 dose dependence in both the rate and probability of PrE specification within the ICM, which is manifest as a reduction in PrE cells in *Gata6* heterozygotes and an absence of a PrE lineage in *Gata6* mutants. The phenotype observed in *Gata6* mutants appears more specific, and more severe, with respect to the ICM than mutants in FGF signaling components, such as *Fgf4*, which also exhibit a lack of PrE (Kang et al., 2013a). By contrast to *Gata6* mutants, which fail to activate any PrE markers, *Fgf4* mutants initially express but fail to maintain GATA6, as well as later PrE-specific factors, such as PDGFRA (Artus et al., 2010; Kang et al., 2013a; Plusa et al., 2008). Furthermore, and unlike *Gata6* mutants, *Fgf4* mutant embryos exhibit a reduction in embryo cell number, as well as defects in the TE, due to a mitogenic role of FGF/ERK signaling in this lineage (Kang et al., 2013a).

Although *Gata6* mutants display ectopic expression of NANOG in TE cells, this did not appear to impair TE function. Rather, it suggests that GATA6 may be involved in the repression of NANOG in these cells. Interestingly, the mechanism by which this happens appears to be distinct from that in ICM cells, because *Fgf4* mutant embryos exhibit expression of NANOG in the TE (Kang et al., 2013a), suggesting this process is independent of FGF signaling.

All cells within *Gata6* mutant ICMs express the pluripotency-associated factors NANOG, SOX2, and OCT4 at all blastocyst stages. Indeed, NANOG has been proposed to interact with GATA6 in a mutually repressive manner (Frankenberg et al., 2013). In the absence of GATA6-mediated repression, one might predict an increase in NANOG levels. However, NANOG levels were not observed to increase in *Gata6* mutant embryos, suggesting that GATA6 itself may be regulating the on/off switch of NANOG, rather than fine-tuning its expression. On the other hand, NANOG expression was decreased and its range was narrowed earlier in $Gata6^{-/-}$ than in $Gata6^{+/+}$ ICM cells (64- to 128-cell stage, compared to >128-cell stage). NANOG levels therefore appeared more comparable to mature EPI cells observed in late blastocyst-stage wild-type embryos. We speculate that the absence of GATA6 does not necessarily result in a novel ICM cell type; rather, cells maintain NANOG levels comparable to those of mature EPI cells of wild-type embryos, where GATA6 is repressed without a concomitant increase in NANOG levels. One could argue that cells in $Gata6^{-/-}$ early ICMs are less plastic than cells in stage-matched wild-type ICMs, and so may already be committed to an EPI fate. Relative, rather than absolute, protein concentrations in any given cell might ensure this, and one might envisage a scenario in which GATA6 and NANOG, acting through mutual repression, maintain an equilibrium that restrains cells from committing to either EPI or PrE fate until an appropriate time. At some point the balance is stochastically disturbed, disequilibrium ensues, and one of these two factors predominates. This event drives the onset of cell lineage choice and commitment to PrE or EPI fates. In $Gata6^{-/-}$ embryos, this initial balance of transcription factors is never established, and consequently all ICM cells prematurely differentiate to a mature EPI state.

In heterozygous embryos, GATA6 protein levels were reduced compared to stage-matched wild-type embryos. Indeed, $Gata6^{+/-}$ embryos exhibited a significantly lower number of PrE cells and a slower commitment, so that at later stages more GATA6/NANOG double-positive (DP) ICM cells were observed in an uncommitted (NANOG+;GATA6+) state. The commitment to an EPI fate appeared slightly faster, although more data are required to confirm this hypothesis.

Given these data, the GRN driving ICM lineage choice can be subject to geometric reasoning (Corson and Siggia, 2012). Our observations are compatible with a simple geometric model of cell differentiation (Waddington landscape). We propose that GATA6 and NANOG mutual repression generates a system with two stable states (valleys): NANOG-positive;GATA6-negative and NANOG-negative;GATA6-positive. Stochastic fluctuations in the levels of GATA6 and NANOG push the cells toward one or the other fate. The time required to reach a stable state is controlled by how much the relative levels of the two factors deviate from the unstable region in which cells remain DP (Figure 1B) ICM (the peak in the landscape). In the absence of GATA6 the landscape changes, so that the only stable state is the NANOG-positive;GATA6-negative state, which should, therefore, display similar levels of NANOG to wild-type EPI cells. In $Gata6^{+/-}$ embryos, levels of GATA6 are reduced and, therefore, a higher number of cells end up in the EPI state, as a result of the relatively higher levels of NANOG compared to GATA6. In cells in which levels of GATA6 are still sufficient to drive the PrE fate, the GATA6:NANOG ratio should be reduced and, therefore, result in a slower transition toward the PrE fate. The model predicts that the opposite would be observed in $Nanog^{+/-}$ embryos. Additional experiments are required to test this model and determine the importance of relative versus absolute levels of GATA6 and NANOG for ICM cell differentiation.

How the transcriptional network involving this interaction of GATA6 and NANOG is integrated with signaling cues to regulate ICM cell fate specification remains an open question. Modulation of FGF signaling, in mutants or using inhibitors or growth factors, affects NANOG and GATA6 expression in wild-type blastocysts (Frankenberg et al., 2011; Kang et al., 2013a; Nichols et al., 2009; Yamanaka et al., 2010). However, treatment of *Gata6* mutants with exogenous FGF4 failed to direct cells toward a PrE fate. These results show that GATA6 is required for FGF-mediated downregulation of NANOG. However, inhibition of ERK in *Gata6* mutants resulted in robust upregulation of NANOG, suggesting that FGF/ERK signaling can directly regulate NANOG levels. These results reveal that FGF signaling is necessary but not, as previously assumed, sufficient to repress NANOG and induce PrE differentiation. Instead, this process is GATA6 dependent, with GATA6 acting upstream of ERK. We suggest a tunable system with three nodes, consisting of GATA6, FGF receptor (FGFR), and NANOG, in which GATA6 is upregulating and possibly mediating stabilization of FGFR (or another upstream element in the cascade), thereby cell-autonomously allowing for elevated stimulation or attenuation of the FGF/ERK pathway by FGF4. Increased ERK signaling would then be able to repress NANOG, which in turn would lead to decreased GATA6 repression by NANOG (Figure 7). Through this extended feedback loop a PrE fate would be reinforced. In EPI cells, very low levels of GATA6 would be insufficient to upregulate FGFR, and baseline levels of the receptor would only allow for marginal induction of FGF/ERK signaling, which would not be able to repress but instead merely moderate NANOG levels, as observed in wild-type EPI cells (Figure 7). Single-cell gene expression profiling demonstrated expression of *Fgfr1* in all ICM cells, and notably a correlation of *Gata6* and *Fgfr2* mRNA levels in blastocysts (Ohnishi et al., 2014). Further studies will be needed to determine whether there is a direct link between GATA6 and FGFR expression, but such a mechanism could produce a general effect on receptor tyrosine kinases (RTKs), as our results also demonstrate the absence of the related RTK PDGFRA in the absence of GATA6.

Unexpectedly, we also noted that, independent of genotype, stimulation of the FGF/ERK pathway at blastocyst stages failed to induce a PrE fate in cells that had presumably already acquired an EPI state. These results suggest that a subpopulation of EPI cells may be lineage committed soon after the establishment of a salt-and-pepper distribution of cells within the ICM and less responsive to FGF signaling cues, in agreement with

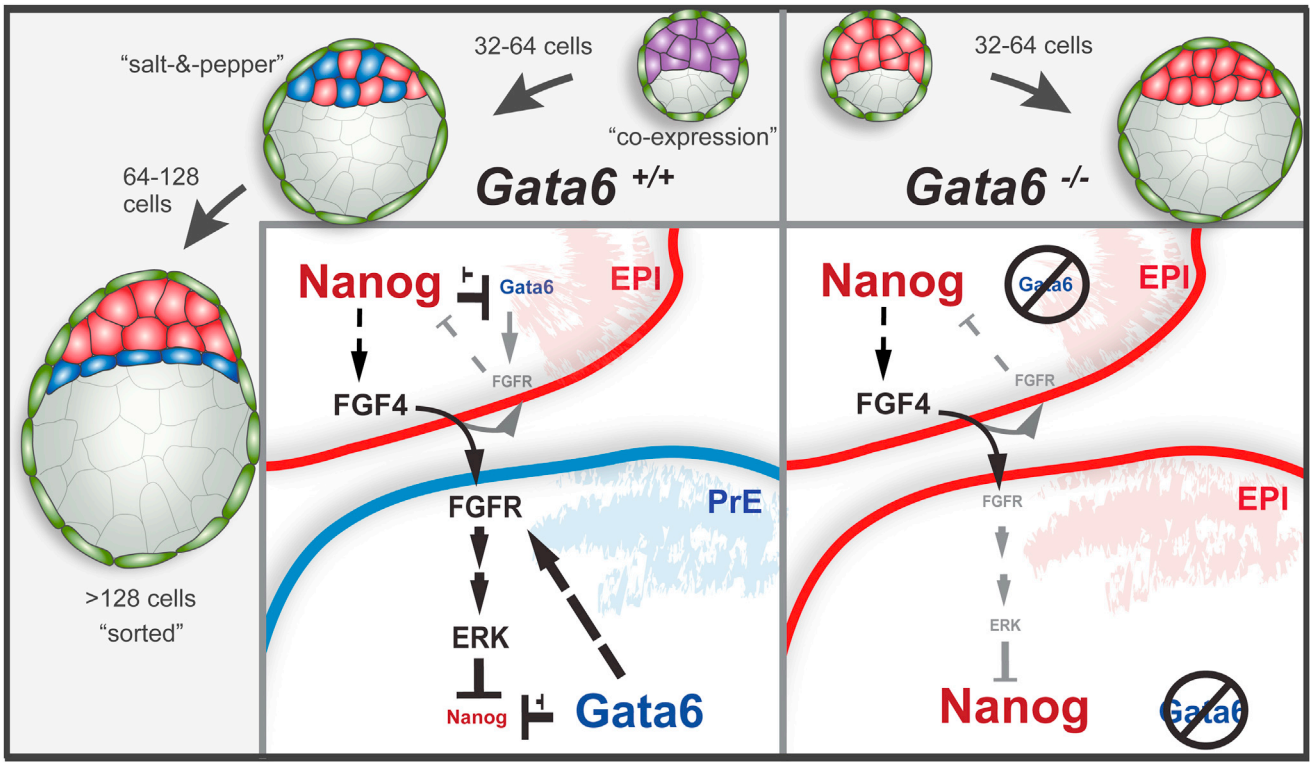


Figure 7. Working Model

GATA6 potentiates the upregulation of FGFR2, thereby cell-autonomously stimulating the FGF/ERK pathway activated by FGF4. Increased ERK signaling leads to repression of NANOG, which in turn releases GATA6 repression. This feedback loop reinforces a PrE fate. In EPI cells, very low levels of GATA6 maintain but do not upregulate FGFR2. Baseline levels of the receptor allow for marginal induction of FGF/ERK signaling, which in turn moderates NANOG levels. Within the insets: lines with arrowheads, positive regulation; lines with blocks, negative regulation; dashed lines, hypothesized regulation; grey lines, weak effects due to low protein concentrations.

previous reports (Grabarek et al., 2012). Because FGF signaling fails to repress NANOG in the absence of GATA6 (Figure 5), this phenomenon can be explained if, in mature EPI cells, GATA6 is repressed and they concomitantly cease to respond to FGF signaling. Such a mechanism could operate to prevent committed EPI cells from switching their fate in the FGF-rich environment of the sorted EPI compartment of the late blastocyst. However, additional experiments will be required to elucidate the mechanistic details behind these observations.

Importantly, recent studies in other mammals (Kuijk et al., 2012; Niakan and Eggan, 2013; Roode et al., 2012; Van der Jeught et al., 2013) have revealed that GATA6 and NANOG are expressed early within the ICM and exhibit a sequence of expression that is comparable to that of the mouse (Chazaud et al., 2006; Dietrich and Hiiragi, 2007; Plusa et al., 2008), suggesting that some aspects of the GRN driving ICM lineage choice are evolutionarily conserved. Surprisingly, it was recently noted that by contrast to rodents (Kang et al., 2013a; Nichols et al., 2009; Roode et al., 2012; Yamanaka et al., 2010), modulation of FGF signaling does not elicit an effect on ICM lineage commitment in human and bovine blastocysts (Kuijk et al., 2012; Roode et al., 2012; Van der Jeught et al., 2013). If the role of FGF signaling in ICM lineage specification is a rodent species-specific adaptation, our studies lead us to speculate that GATA6 might be a key evolutionarily conserved transcription factor driving PrE lineage specification across eutherian mammals.

EXPERIMENTAL PROCEDURES

Mouse Husbandry

Mouse strains used were *Gata6*$^{cKO/cKO}$ (Sodhi et al., 2006), *CAG:Cre* (Sakai and Miyazaki, 1997), *Zp3:Cre* (de Vries et al., 2000), *Pdgfra*$^{H2B\text{-}GFP/+}$ (Hamilton et al., 2003), and wild-type CD1 (Taconic). *Gata6*$^{KO/+}$ animals were of a mixed (CD1/B6/129) strain background and were generated by crossing *Gata6*$^{cKO/cKO}$ to a *CAG:Cre* strain. Embryos with maternal and zygotic ablation of *Gata6* were obtained by breeding *Gata6*$^{cKO/cKO}$; *Zp3:Cre*$^{Tg/+}$ females with *Gata6*$^{cKO/cKO}$ males. Work on mice was subject to approval by, and carried out in accordance with guidelines from, the Memorial Sloan Kettering Cancer Center Institutional Animal Care and Use Committee.

Embryo Collection and Handling

Mice were maintained under a 12 hr light cycle. Preimplantation embryos were flushed from uteri or oviducts in flushing holding medium (FHM, HEPES-buffered potassium simplex optimized medium [KSOM] based; Millipore) as described (Nagy et al., 2003). The zona pellucida was removed from blastocysts by brief incubation in acid Tyrode's solution (Sigma). Embryos were fixed in 4% paraformaldehyde in PBS for 10 min at room temperature and stored in PBS at 4°C.

Embryo In Vitro Culture

Embryos were cultured in drops of KSOM-AA (KSOM supplemented with amino acids; Millipore) under mineral oil (Sigma) at 37°C in a 5% CO_2

atmosphere. For live imaging, embryos were cultured on glass-bottom dishes (MatTek) in an environmental chamber (Solent Scientific) as described previously (Piliszek et al., 2011). Embryos at postcavitation stages were cultured without zona pellucidae. For FGF treatment experiments, FGF4 or FGF2 (R&D Systems) was added to culture medium at 500–2,000 ng/ml in the presence of 1 μg/ml heparin (Sigma). For ERK inhibition experiments, 1 μM ERK1/2 inhibitor PD0325901 (Stemgen) was added to the culture. Embryos were cultured from E2.5 or E3.5 for various periods of time.

Immunostaining

Immunofluorescence staining was performed as previously described (Frankenberg et al., 2013). The following primary antibodies were used: goat anti-SOX2 (R&D Systems) at a dilution of 1:50; mouse anti-OCT4 (Santa Cruz), goat anti-GATA4 (Santa Cruz), goat anti-GATA6 (R&D Systems), and goat anti-SOX17 (R&D Systems) at 1:100; rabbit anti-PKCζ (Santa Cruz), mouse anti-DAB2 (BD Transduction Laboratories), and mouse anti-CDX2 (BioGenex) at 1:200; rabbit anti-NANOG (Cosmo Bio), mouse anti-GATA3 (BioLegend), and rabbit anti-EOMES (Abcam) at 1:500. Secondary Alexa Fluor-conjugated antibodies (Invitrogen) were used at a dilution of 1:500. DNA was visualized using 5 μg/ml Hoechst 33342 (Invitrogen) in PBS.

Image Data Acquisition and Processing

Laser-scanning confocal images were acquired on a Zeiss LSM 510 META. Embryos were mounted in PBS on glass-bottom dishes. Fluorescence was excited with a 405 nm laser diode (Hoechst), 488 nm argon laser (GFP, Alexa Fluor 488), 543 nm HeNe laser (Alexa Fluor 543, 555), and 633 nm HeNe laser (Alexa Fluor 633 and 647). Images were acquired using a Plan-Neofluar 40×/1.3 Oil DIC objective, with an optical section thickness of 1–1.2 μm. Raw data were processed using ZEN (Carl Zeiss Microsystems), ImageJ (NIH), and MINS (Lou et al., 2014) software. Three-dimensional time-lapse imaging was performed on a Zeiss LSM 510 META. Embryos were placed in drops of KSOM under mineral oil on glass-bottom dishes (MatTek). Time intervals between z stacks were 15 min, for a 15–20 hr total. Raw data were processed using Zeiss ZEN software.

Image Segmentation and Quantitative Fluorescence Intensity Analysis Pipeline

For cell counting and quantification, MINS software (http://katlab-tools.org) was used (Lou et al., 2014). Details of the analyses are provided in Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, one table, and one movie and can be found with this article online at http://dx.doi.org/10.1016/j.devcel.2014.04.011.

ACKNOWLEDGMENTS

We thank S. Duncan and P. Soriano for *Gata6*^{cKO} and *Pdgfra*$^{H2B\text{-}GFP}$ mouse lines, respectively; V. Seshan of the Memorial Sloan Kettering Cancer Center Biostatistics Core Facility for advice on fluorescence intensity data transformations; and M. Kang, J. Nichols, S. Nowotschin, B. Plusa, C. Schroeter, and P. Xenopoulos for discussions and comments on the manuscript. Our work is supported by NIH-R01-HD052115, NIH-R01-DK084391, and NYSTEM (to A.-K.H.) and NIH-K99-HD074670 (to S.D.T.).

Received: December 14, 2013
Revised: April 3, 2014
Accepted: April 10, 2014
Published: May 15, 2014

REFERENCES

Artus, J., Panthier, J.J., and Hadjantonakis, A.K. (2010). A role for PDGF signaling in expansion of the extra-embryonic endoderm lineage of the mouse blastocyst. Development *137*, 3361–3372.

Artus, J., Piliszek, A., and Hadjantonakis, A.K. (2011). The primitive endoderm lineage of the mouse blastocyst: sequential transcription factor activation and regulation of differentiation by Sox17. Dev. Biol. *350*, 393–404.

Artus, J., Kang, M., Cohen-Tannoudji, M., and Hadjantonakis, A.K. (2013). PDGF signaling is required for primitive endoderm cell survival in the inner cell mass of the mouse blastocyst. Stem Cells *31*, 1932–1941.

Chazaud, C., Yamanaka, Y., Pawson, T., and Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Dev. Cell *10*, 615–624.

Corson, F., and Siggia, E.D. (2012). Geometry, epistasis, and developmental patterning. Proc. Natl. Acad. Sci. USA *109*, 5568–5575.

Czechanski, A., Byers, C., Greenstein, I., Schrode, N., Donahue, L.R., Hadjantonakis, A.K., and Reinholdt, L.G. (2014). Derivation and characterization of mouse embryonic stem cells from permissive and nonpermissive strains. Nat. Protoc. *9*, 559–574.

de Vries, W.N., Binns, L.T., Fancher, K.S., Dean, J., Moore, R., Kemler, R., and Knowles, B.B. (2000). Expression of Cre recombinase in mouse oocytes: a means to study maternal effect genes. Genesis *26* (2), 110–112.

Dietrich, J.E., and Hiiragi, T. (2007). Stochastic patterning in the mouse preimplantation embryo. Development *134*, 4219–4231.

Frankenberg, S., Gerbe, F., Bessonnard, S., Belville, C., Pouchin, P., Bardot, O., and Chazaud, C. (2011). Primitive endoderm differentiates via a three-step mechanism involving Nanog and RTK signaling. Dev. Cell *21*, 1005–1013.

Frankenberg, S., Shaw, G., Freyer, C., Pask, A.J., and Renfree, M.B. (2013). Early cell lineage specification in a marsupial: a case for diverse mechanisms among mammals. Development *140*, 965–975.

Fujikura, J., Yamato, E., Yonemura, S., Hosoda, K., Masui, S., Nakao, K., Miyazaki, J.-i., and Niwa, H. (2002). Differentiation of embryonic stem cells is induced by GATA factors. Genes Dev. *16*, 784–789.

Gerbe, F., Cox, B., Rossant, J., and Chazaud, C. (2008). Dynamic expression of Lrp2 pathway members reveals progressive epithelial differentiation of primitive endoderm in mouse blastocyst. Dev. Biol. *313*, 594–602.

Goldin, S.N., and Papaioannou, V.E. (2003). Paracrine action of FGF4 during periimplantation development maintains trophectoderm and primitive endoderm. Genesis *36*, 40–47.

Grabarek, J.B., Zyzyńska, K., Saiz, N., Piliszek, A., Frankenberg, S., Nichols, J., Hadjantonakis, A.K., and Plusa, B. (2012). Differential plasticity of epiblast and primitive endoderm precursors within the ICM of the early mouse embryo. Development *139*, 129–139.

Guo, G., Huss, M., Tong, G.Q., Wang, C., Li Sun, L., Clarke, N.D., and Robson, P. (2010). Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. Dev. Cell *18*, 675–685.

Hamilton, T.G., Klinghoffer, R.A., Corrin, P.D., and Soriano, P. (2003). Evolutionary divergence of platelet-derived growth factor α receptor signaling mechanisms. Mol. Cell. Biol. *23*, 4013–4025.

Johnson, M.H., and Ziomek, C.A. (1981). Induction of polarity in mouse 8-cell blastomeres: specificity, geometry, and stability. J. Cell Biol. *91*, 303–308.

Kang, M., Piliszek, A., Artus, J., and Hadjantonakis, A.K. (2013a). FGF4 is required for lineage restriction and salt-and-pepper distribution of primitive endoderm factors but not their initial expression in the mouse. Development *140*, 267–279.

Kang, M., Xenopoulos, P., Muñoz-Descalzo, S., Lou, X., and Hadjantonakis, A.K. (2013b). Live imaging, identifying, and tracking single cells in complex populations in vivo and ex vivo. Methods Mol. Biol. *1052*, 109–123.

Koutsourakis, M., Langeveld, A., Patient, R., Beddington, R., and Grosveld, F. (1999). The transcription factor GATA6 is essential for early extraembryonic development. Development *126*, 723–732.

Krawchuk, D., Honma-Yamanaka, N., Anani, S., and Yamanaka, Y. (2013). FGF4 is a limiting factor controlling the proportions of primitive endoderm and epiblast in the ICM of the mouse blastocyst. Dev. Biol. *384*, 65–71.

Kuijk, E.W., van Tol, L.T., Van de Velde, H., Wubbolts, R., Welling, M., Geijsen, N., and Roelen, B.A. (2012). The roles of FGF and MAP kinase signaling in the segregation of the epiblast and hypoblast cell lineages in bovine and human embryos. Development *139*, 871–882.

Kurimoto, K., Yabuta, Y., Ohinata, Y., Ono, Y., Uno, K.D., Yamada, R.G., Ueda, H.R., and Saitou, M. (2006). An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis. Nucleic Acids Res. *34*, e42.

Lou, X., Kang, M., Xenopoulos, P., Muñoz-Descalzo, S., and Hadjantonakis, A.K. (2014). A rapid and efficient 2D/3D nuclear segmentation method for analysis of early mouse embryo and stem cell image data. Stem Cell Reports *2*, 382–397.

Meilhac, S.M., Adams, R.J., Morris, S.A., Danckaert, A., Le Garrec, J.F., and Zernicka-Goetz, M. (2009). Active cell movements coupled to positional induction are involved in lineage segregation in the mouse blastocyst. Dev. Biol. *331*, 210–221.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631–642.

Morris, S.A., Teo, R.T., Li, H., Robson, P., Glover, D.M., and Zernicka-Goetz, M. (2010). Origin and formation of the first two distinct cell types of the inner cell mass in the mouse embryo. Proc. Natl. Acad. Sci. USA *107*, 6364–6369.

Morrisey, E.E., Ip, H.S., Lu, M.M., and Parmacek, M.S. (1996). GATA-6: a zinc finger transcription factor that is expressed in multiple cell lineages derived from lateral mesoderm. Dev. Biol. *177*, 309–322.

Nagy, A., Gertsenstein, M., Vinterstein, K., and Behringer, R. (2003). Manipulating the Mouse Embryo. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Niakan, K.K., and Eggan, K. (2013). Analysis of human embryos from zygote to blastocyst reveals distinct gene expression patterns relative to the mouse. Dev. Biol. *375*, 54–64.

Niakan, K.K., Ji, H., Maehr, R., Vokes, S.A., Rodolfa, K.T., Sherwood, R.I., Yamaki, M., Dimos, J.T., Chen, A.E., Melton, D.A., et al. (2010). Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating extraembryonic gene expression and indirectly antagonizing self-renewal. Genes Dev. *24*, 312–326.

Nichols, J., Silva, J., Roode, M., and Smith, A. (2009). Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. Development *136*, 3215–3222.

Ohnishi, Y., Huber, W., Tsumura, A., Kang, M., Xenopoulos, P., Kurimoto, K., Oleś, A.K., Araúzo-Bravo, M.J., Saitou, M., Hadjantonakis, A.K., and Hiiragi, T. (2014). Cell-to-cell expression variability followed by signal reinforcement progressively segregates early mouse lineages. Nat. Cell Biol. *16*, 27–37.

Piliszek, A., Kwon, G.S., and Hadjantonakis, A.K. (2011). Ex utero culture and live imaging of mouse embryos. Methods Mol. Biol. *770*, 243–257.

Plusa, B., Piliszek, A., Frankenberg, S., Artus, J., and Hadjantonakis, A.K. (2008). Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. Development *135*, 3081–3091.

Roode, M., Blair, K., Snell, P., Elder, K., Marchant, S., Smith, A., and Nichols, J. (2012). Human hypoblast formation is not dependent on FGF signalling. Dev. Biol. *361*, 358–363.

Saiz, N., and Plusa, B. (2013). Early cell fate decisions in the mouse embryo. Reproduction *145*, R65–R80.

Saiz, N., Grabarek, J.B., Sabherwal, N., Papalopulu, N., and Plusa, B. (2013). Atypical protein kinase C couples cell sorting with primitive endoderm maturation in the mouse blastocyst. Development *140*, 4311–4322.

Sakai, K., and Miyazaki, J.-i. (1997). A transgenic mouse line that retains Cre recombinase activity in mature oocytes irrespective of the cre transgene transmission. Biochem. Biophys. Res. Commun. *237*, 318–324.

Santostefano, K.E., Hamazaki, T., Pardo, C.E., Kladde, M.P., and Terada, N. (2012). Fibroblast growth factor receptor 2 homodimerization rapidly reduces transcription of the pluripotency gene Nanog without dissociation of activating transcription factors. J. Biol. Chem. *287*, 30507–30517.

Schrode, N., Xenopoulos, P., Piliszek, A., Frankenberg, S., Plusa, B., and Hadjantonakis, A.K. (2013). Anatomy of a blastocyst: cell behaviors driving cell fate choice and morphogenesis in the early mouse embryo. Genesis *51*, 219–233.

Shimosato, D., Shiki, M., and Niwa, H. (2007). Extra-embryonic endoderm cells derived from ES cells induced by GATA factors acquire the character of XEN cells. BMC Dev. Biol. *7*, 80.

Singh, A.M., Hamazaki, T., Hankowski, K.E., and Terada, N. (2007). A heterogeneous expression pattern for Nanog in embryonic stem cells. Stem Cells *25*, 2534–2542.

Smyth, N., Vatansever, H.S., Murray, P., Meyer, M., Frie, C., Paulsson, M., and Edgar, D. (1999). Absence of basement membranes after targeting the LAMC1 gene results in embryonic lethality due to failure of endoderm differentiation. J. Cell Biol. *144*, 151–160.

Sodhi, C.P., Li, J., and Duncan, S.A. (2006). Generation of mice harbouring a conditional loss-of-function allele of Gata6. BMC Dev. Biol. *6*, 19.

Van der Jeught, M., O'Leary, T., Ghimire, S., Lierman, S., Duggal, G., Versieren, K., Deforce, D., Chuva de Sousa Lopes, S., Heindryckx, B., and De Sutter, P. (2013). The combination of inhibitors of FGF/MEK/Erk and GSK3β signaling increases the number of OCT3/4- and NANOG-positive cells in the human inner cell mass, but does not improve stem cell derivation. Stem Cells Dev. *22*, 296–306.

Yamanaka, Y., Lanner, F., and Rossant, J. (2010). FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. Development *137*, 715–724.

Developmental Cell, Volume *29*

Supplemental Information

GATA6 Levels Modulate Primitive Endoderm Cell Fate Choice and Timing in the Mouse Blastocyst

Nadine Schrode, Nestor Saiz, Stefano Di Talia, and Anna-Katerina Hadjantonakis

SUPPLEMENTAL FIGURES AND LEGENDS

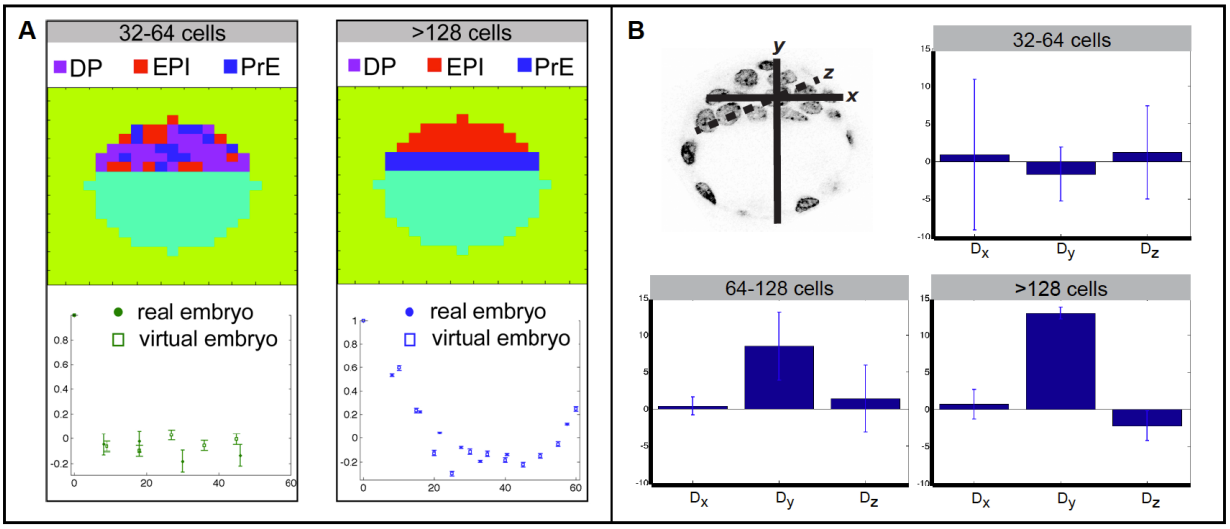


Figure S1, related to Figure 1: Observed correlations in cell fates are consistent with an early random commitment and geometrical constraints at later stages.

A simplified 2D geometry mimicking the mouse blastocyst was generated on a uniformly spaced lattice. Cell fates were assigned either randomly (32-64 cell stage) or in a pattern resembling the embryo after sorting of PrE and EPI cells. The Pearson's correlation coefficient of cell fates as a function of cell distance was computed for virtual embryos and compared to the measured data of real embryos. The simulated dependency of the correlation function is in good agreement with the measured correlations, suggesting that early cell fate specification is controlled by random commitment, and that the correlations observed in late embryos are controlled by their geometry. **B** Average difference in the (x,y,z) coordinates of EPI and PrE cells as a function of developmental stage (see cartoon for information on coordinates; at the 32-64 cells stage, this computation was performed on the few cells that have already committed to EPI or PrE fate- see data in Figure 1B). At all stages, embryos do not show asymmetries in x and z, as expected. At the 32-64 cells stage, there is no asymmetry in y, suggesting that early commitment is controlled by cell autonomous stochastic factors. Asymmetries in y become evident at the 64-128 cells stage and further increase at the >128 cells stage. These observations suggest that the movement of EPI cells away from the cavity and PrE cells towards the cavity is subsequent to the random cell fate commitment observed at the earlier stages.

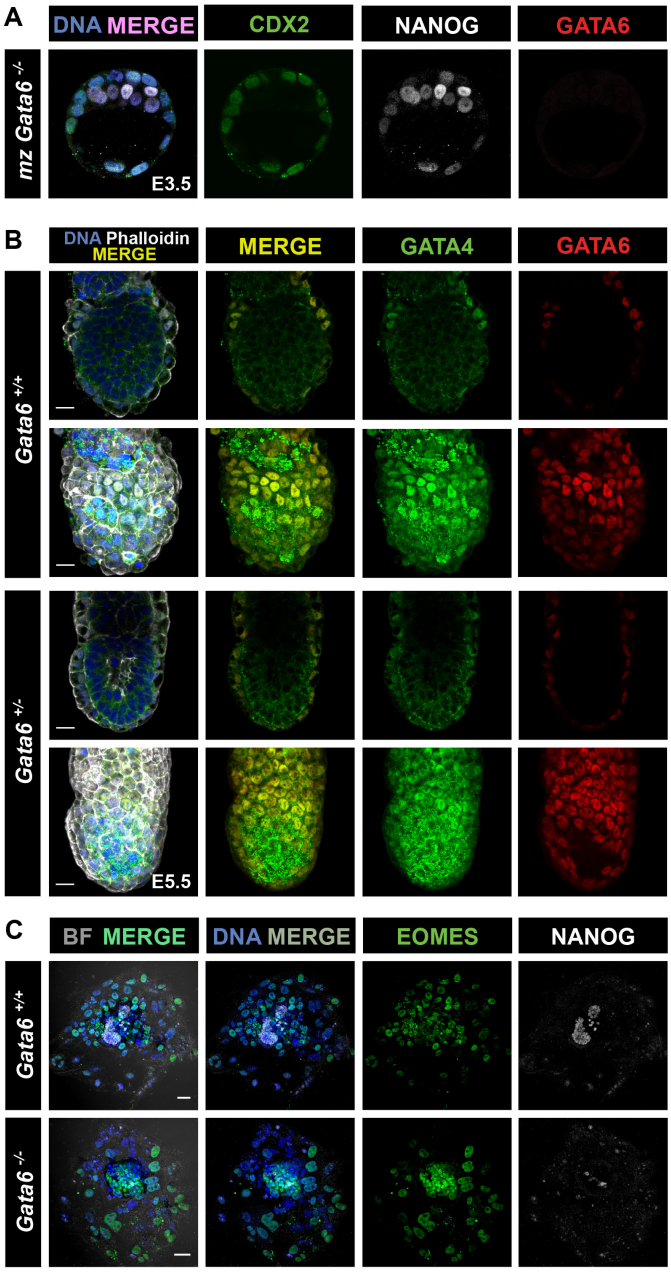


Figure S2, related to Figure 2: Marker expression in maternal zygotic Gata6 mutant, wild-type and heterozygote embryos at different stages and in blastocyst outgrowths.

A Localization and distribution of CDX2 (green), NANOG (white) and GATA6 (red) in maternal zygotic E3.5 *Gata6* mutant blastocysts. Blue in merge: Hoechst. **B** Localization and distribution of GATA4 (green) and GATA6 (red) in E5.5 *Gata6*$^{+/+}$ and *Gata6*$^{+/-}$ embryos. Blue in merge: Hoechst. White in merge: Phalloidin. C Localization and distribution of EOMES (green) and NANOG (white) in outgrowths of *Gata6*$^{+/+}$ and *Gata6*$^{-/-}$ embryos 4 days after plating. Blue in merge: Hoechst. BF: Brightfield.

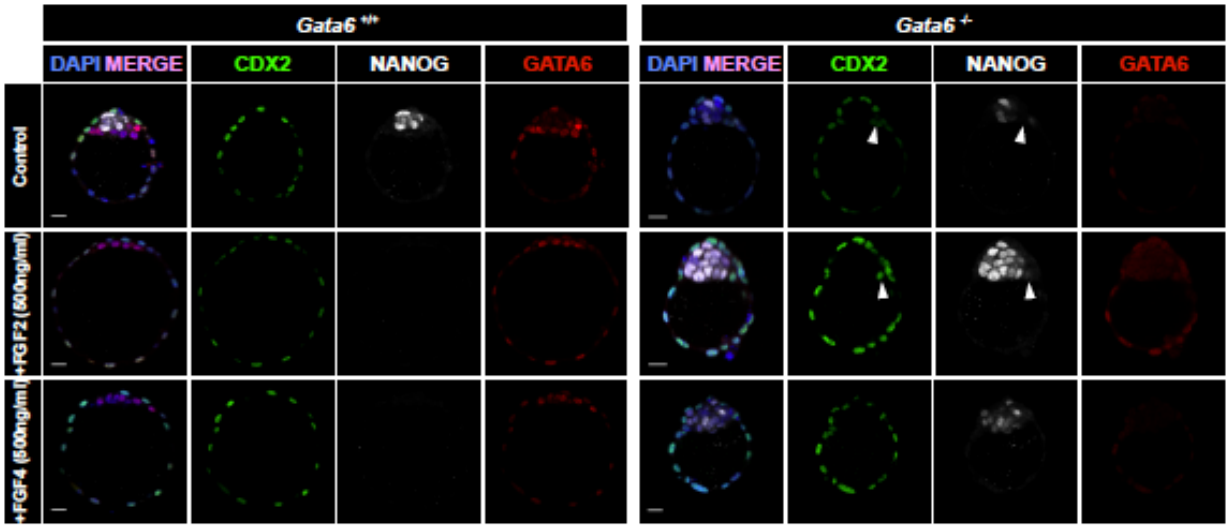


Figure S3, related to Figure 3: CDX2 expression in ICM cells after prolonged culture in FGF.

Localization and distribution of CDX2 (green), NANOG (white) and GATA6 (red) in *Gata6* wild-type and mutant mouse embryos. Embryos were cultured from the 8-cell stage (E2.5) for 72h to assess whether the population of CDX2+ cells in the ICM expanded over time as a consequence of exposure to FGF. Arrowheads indicate CDX2 expressing cells in the ICM.

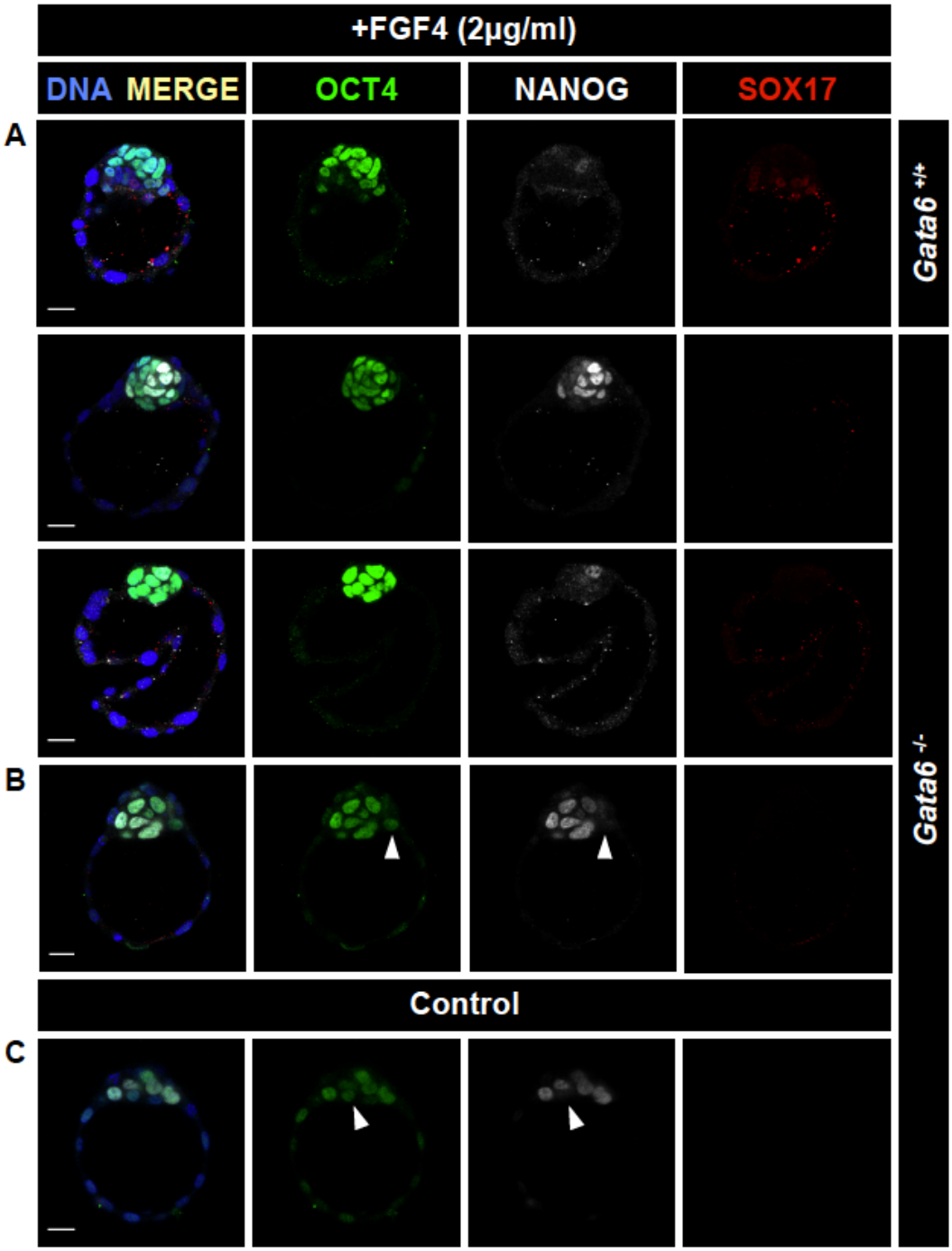


Figure S4, related to Figure 5: NANOG/GATA6 negative cells in the ICM of FGF-treated embryos are OCT4 positive.

Localization and distribution of OCT4 (green), NANOG (white) and SOX17 (red) in *Gata6* wild-type and mutant blastocysts after culture in 2 µg/mL FGF from E3.5 for 24 h **(A)**, from E2.5 for 48 h **(B)** and untreated **(C)**. Blue in merge: Hoechst.

SUPPLEMENTAL TABLES

	Genotype			
Stage of embryo	***Gata6*$^{+/+}$**	***Gata6*$^{+/-}$**	***Gata6*$^{-/-}$**	**Total no. of embryos recovered**
<64 cells **early blastocyst**	28	68	34	130
64-128 cells **mid-blastocyst**	19	40	20	79
>128 cells **late blastocyst**	10	10	12	32
E5.5 **post-implantation**	8	34	0	42

Table S1, related to Figure 2: Genotypes of embryos recovered from *Gata6* heterozygous intercrosses at various stages.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Image data acquisition and processing

Laser scanning confocal images of immunostained embryos were acquired on a Zeiss LSM 510 META. Embryos were mounted in PBS on glass-bottom dishes. Fluorescence was excited with a 405-nm laser diode (Hoechst), a 488-nm Argon laser (GFP, Alexa Fluor 488), a 543-nm HeNe laser (Alexa Fluor 543, 555) and a 633-nm HeNe laser (Alexa Fluor 633 and 647). Images were acquired using a Plan-Neofluar 40×/1.3 Oil DIC objective, with an optical section thickness of 1–1.2 µm. Raw data were processed using ZEN (Carl Zeiss Microsystems), ImageJ (NIH) and MINS (Lou et al., 2014) software. 3D time-lapse imaging was performed on a Zeiss LSM 510 META. Embryos were placed in drops of KSOM under mineral oil on glass-bottom dishes (MatTek Corp). Time intervals between z-stacks were 15 minutes, for 15-20 hours total. Raw data were processed using Zeiss ZEN software (Carl Zeiss Microsystems).

Image segmentation and quantitative fluorescence intensity analysis pipeline

For cell counting and quantification MINS software (http://katlab-tools.org/) was used (Lou et al., 2014). Confocal images in *.lsm format were loaded into the MINS pipeline, which detects nuclei, accurately segments them and exports results of nuclear count, size, position and fluorescence intensity of all channels in table format, and as images of numbered nuclei (Figure 1). The MINS output was checked for over- or under-segmentation and tables were corrected manually. Mitotic cells were also found and excluded from the analysis. TE cells were defined molecularly and positionally, as nuclei positive for CDX2 in cells located on the surface of the embryo. For representation of fluorescence intensity distribution throughout ICMs, data was subject to logarithm transformation to reduce the technical variability due to the loss of fluorescence along the z-axis. For statistical comparison of intensity distributions, these values were plotted as the average log intensity per embryo rather than per cell. A Wilcoxon Mann-Whitney Test was performed to compare two populations, a Kruskal-Wallis Test for more than two populations. All analysis of fluorescence intensity levels and statistics were performed using R (http://www.r-project.org/). For correlation and probability functions nuclear concentrations were extracted from nuclear intensity measurements by first subtracting background fluorescence, and then correcting for the decrease in fluorescence intensities due to tissue depth in z-stacks of image data. Background intensities for both NANOG and GATA6

were estimated from the average fluorescence intensity values (at the >128 cell stage) of half of the TE cells having the lowest intensities. The value obtained for GATA6 was comparable to the one obtained by imaging *Gata6*$^{-/-}$ embryos and the value for NANOG was very low and similar to the one observed in PrE cells at late stages, thereby validating our procedure. A correction factor as a function of axial position (z coordinate) in the embryo was estimated by determining how the signal intensity of the DAPI channel scaled with z. The DAPI intensities of all nuclei in an embryo were fitted as a function of z using the Smoothing Splines function in MATLAB (http://www.mathworks.com/products/matlab/), and the fit was then used to correct the fluorescence intensities. The number of PrE and EPI cells was established using the threshold functions shown in Figure 1B. Correlation functions were computed by evaluating the Pearson's correlation coefficient for each pair of cells as a function of the distance between them. The probabilities of cell specification (as well as their 95% confidence intervals) as a function of cell number were estimated through logistic regression using the Curve Fitting Toolbox in MATLAB.

Genotyping

Following image acquisition embryos were genotyped as described previously (Kang et al., 2013a) using the following primers: *Gata6cKO*-fw: 5′-GTGGTTGTAAGGCGGTTTGT-3′, *Gata6cKO*-rev: 5′-ACGCGAGCTCCAGAAAAAGT-3′, Gata6KO-fw: 5′-AGTCTCCCTGTCATTCTTCCTGCTC-3′, *Gata6KO*-rev: 5′-TGATCAAACCTGGGTCTACACTCCTA-3′, *GFP*-fw: 5′-AAGTTCATCTGCACCACCG-3′, *GFP*-rev: 5′-TGCTCAGGTAGTGGTTGTCG-3′.

Embryo-derived stem cell derivation and culture

ES cell lines were derived from embryos collected from *Gata6*$^{+/-}$ intercrosses according to standard procedures (Czechanski et al., 2014). Embryonic day (E) 3.5 embryos were on mitotically inactivated murine embryonic fibroblast (MEF) feeder cells for 9 days. Outgrowths were disaggregated and passaged into 24-well dishes. Medium was replaced every 2 days. Emerging ES cell colonies were passaged into new 24-well plates until confluence. Cells were then cultured in the absence of MEFs and genotyped. All cells were grown at 37°C in 5% CO2. ES cells were routinely cultured as described previously (Kang et al., 2013a).

RT-PCR

RNA from *Gata6*$^{+/+}$ and *Gata6*$^{-/-}$ ES cells was isolated using Trizol Reagent (Life Technologies) according to the manufacturer's instructions. Reverse transcription (RT) and genomic DNA digestion was performed using QuantiTect Reverse Transcrption Kit (Qiagen) according to the manufacturer's instructions. PCR was performed on the prepared cDNA using the following primer sequences: *Gapdh*-fw: 5'-GTGTTCCTACCCCCAATGTGT-3', *Gapdh*-rev: 5'-ATTGTCATACCAGGAAATGAGCTT-3', *Gata6*-fw: 5'-AGACGGCACCGGTCATTACC-3', *Gata6*-rev: 5'-TCACCCTCAGCATTTCTACGCC-3', *Nanog*-fw: 5'-TGCGGACTGTGTTCTCTCAGG-3', *Nanog*-rev: 5'-CCACTGGTTTTTCTGCCACCG-3', *Fgf4*-fw: 5'-ACTACCTGCTGGGCCTCAAA-3', *Fgf4*-rev: 5'-AAGGCACACCGAAGAGCTTG-3', *Fgfr1*-fw: 5'-CACCTGCATCGTGGAGAATG-3', *Fgfr1*-rev: 5'-CAAGTTGTCTGGCCCGATCT-3', *Fgfr2*-fw: 5'-CCTTCGGGGTGTTAATGTGG-3', *Fgfr2*-rev: 5'-GGGTACAGCATGCCAGCAAT-3'. Primers were designed for amplicons to span exon borders to avoid possible amplification of genomic DNA. *Gata6* primers were designed to bind outside of Exon2, which is deleted in *Gata6* mutant cells.

Movie S1. PdgfraH2B-GFP PrE Reporter Is Not Activated during Gata6 Mutant Blastocyst Development, Related to Figure 3.

Time series of early to late blastocyst stage embryos from a *Gata6*$^{+/-}$; PdgfraH2BGFP × *Gata6*$^{+/-}$ cross. Time total: 20 hr; time interval: 15 min; z-total: 62 μm; z-interval: 2 μm. Labels indicate genotypes.

Location:
http://www.sciencedirect.com/science/article/pii/S1534580714002317
with Supplemental Information

Download:
http://www.sciencedirect.com/science/MiamiMultiMediaURL/1-s2.0-S1534580714002317/1-s2.0-S1534580714002317-mmc2.mp4/272236/html/S1534580714002317/ac5f3215afa970b1112cd155487d008a/mmc2.mp4

ADDITIONAL RESULTS

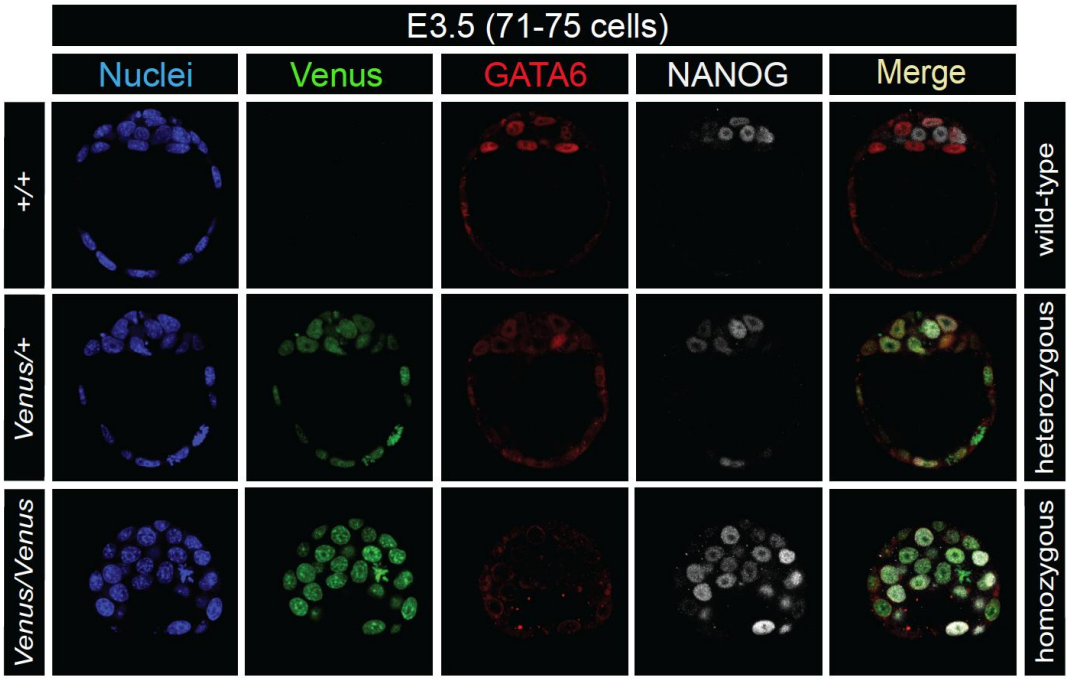


Figure 7. Gata6$^{H2B\text{-}Venus}$ reporter expression in blastocyst stage embryos and immunohistochemistry for GATA6 and NANOG protein. Gata6$^{H2B\text{-}Venus}$ knock-in reporter expression creates embryos, which phenocopy *Gata6* heterozygous and null mutant embryos. Gata6$^{H2B\text{-}Venus}$ displays expression in the absence of GATA6 protein. (Freyer et al., submitted).

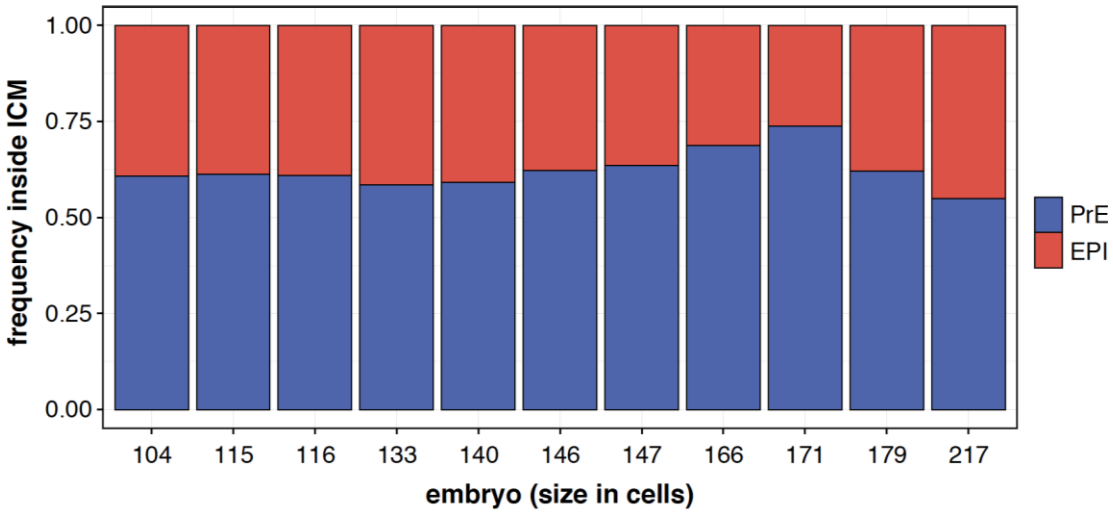


Figure 8. ICM composition in wildtype late blastocyst stage embryos. Each stacked bar represents one wildtype embryo ICM, showing that proportions of epiblast (EPI) and primitive endoderm (PrE) are relatively stable at the mid to late blastocyst stage, with around 40 % and 60 %, respectively.

PUBLICATION II

DERIVATION OF EXTRAEMBRYONIC ENDODERM STEM (XEN) CELLS FROM MOUSE EMBRYOS AND EMBRYONIC STEM CELLS

Niakan K.K., Schrode N., Cho L.T., Hadjantonakis A.K.

Nature Protocols
8(6):1028-41
2013

Derivation of extraembryonic endoderm stem (XEN) cells from mouse embryos and embryonic stem cells

Kathy K Niakan[1,2,5], Nadine Schrode[3], Lily T Y Cho[1,4] & Anna-Katerina Hadjantonakis[3]

[1]The Anne McLaren Laboratory for Regenerative Medicine, University of Cambridge, Cambridge, UK. [2]Centre for Trophoblast Research, Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK. [3]Developmental Biology Program, Sloan-Kettering Institute, New York, New York, USA. [4]Department of Surgery, University of Cambridge, Cambridge, UK. [5]Present address: Division of Stem Cell Biology and Developmental Genetics, Medical Research Council National Institute for Medical Research, London, UK. Correspondence should be addressed to K.K.N. (kniakan@nimr.mrc.ac.uk) or A.-K.H. (hadj@mskcc.org).

Published online 2 May 2013; doi:10.1038/nprot.2013.049

At the time of implantation in the maternal uterus, the mouse blastocyst possesses an inner cell mass comprising two lineages: epiblast (Epi) and primitive endoderm (PrE). Representative stem cells derived from these two cell lineages can be expanded and maintained indefinitely *in vitro* as either embryonic stem (ES) or XEN cells, respectively. Here we describe protocols that can be used to establish XEN cell lines. These include the establishment of XEN cells from blastocyst-stage embryos in either standard embryonic or trophoblast stem (TS) cell culture conditions. We also describe protocols for establishing XEN cells directly from ES cells by either retinoic acid and activin-based conversion or by overexpression of the GATA transcription factor Gata6. XEN cells are a useful model of PrE cells, with which they share gene expression, differentiation potential and lineage restriction. The robust protocols for deriving XEN cells described here can be completed within 2–3 weeks.

INTRODUCTION

The mouse embryo ~3.5 d after fertilization forms a blastocyst comprising three lineages[1]: the extraembryonic trophectoderm (TE), the PrE and the pluripotent Epi (**Fig. 1**) from which cognate *ex vivo* stem cells can be derived. TS cells are derived from the TE[2], XEN cells from the PrE[3] and ES cells from the Epi (refs. 4,5; **Fig. 2**) (reviewed in ref. 6). Notably, each of these stem cell lines is a useful model of the blastocyst cell lineage that they represent. Mouse ES and TS cells have been used successfully for many years to model Epi or TE biology, including the mechanisms of pluripotency maintenance and placental development, respectively. Recently derived XEN cell lines have the distinctive characteristic of cells with at least two morphologies: they are highly refractile as well as epithelial-like[3] (**Fig. 2**), and they are only beginning to be used to understand the mechanisms of PrE development with significance for stem cell and developmental biology.

Mouse ES cells can be directed to differentiate into extraembryonic lineages by the overexpression of single transcription factors, such as the caudal-related homeodomain transcription factor Cdx2 (to derive TS cells)[7] or the GATA transcription factor Gata6 (to derive XEN cells)[8]. Retinoic acid treatment of mouse ES cells[9–11] or embryoid body aggregation[12] has been shown to promote a heterogeneous mixture of XEN-differentiated cells. Notably, these cells have not been demonstrated to self-renew indefinitely, unlike bona fide XEN cell lines. We have recently demonstrated that mouse ES cells can be converted to stable XEN cell lines using retinoic acid together with activin[13]. In this protocol, we focus on the derivation of XEN cells from embryos and ES cells.

The molecular mechanisms underlying XEN cell establishment and maintenance are beginning to be understood. Robust methods for XEN cell derivation from embryos and ES cells, as well as the concomitant availability of XEN cell lines, will further facilitate and improve our understanding of the key fate decisions that occur within the early embryo, including unraveling mechanisms underlying cellular differentiation and pluripotency[14,15]. As a stem cell type that can be derived from both embryos and ES cells, XEN cells are emerging as a valuable tool for modeling the XEN lineage.

Applications of XEN cells

XEN cell derivation can be used as a phenotypic tool to assess the requirement of genes for XEN cell specification, maintenance or expansion, as we have previously demonstrated for SRY-box containing gene 17 *Sox17*)[4], platelet-derived growth factor receptor alpha (*Pdgfra*)[11] and fibroblast growth factor 4 (Fgf4)[16]. Notably, established XEN cell lines serve as a paradigm for XEN biology and for the differentiation of the PrE into derivatives such as visceral[17–19] and parietal endoderm.

Given the fact that reciprocal inductive interactions between the pluripotent Epi and its adjacent extraembryonic tissues, including the visceral endoderm (reviewed in refs. 20 and 21), are crucial to patterning the early embryo, XEN cells can be used as an *ex vivo* tool for teasing apart the underlying mechanisms and for identifying the key molecules involved[12]. XEN cells can be used as an important *in vitro* tool for elucidating details of additional patterning activities of the extraembryonic endoderm, such as identifying factors involved in cardiac induction[22–24]. Moreover, as they can be propagated in large quantities and do not require growth factor supplements to culture media, these cells are a cost-effective, attractive and tractable system for high-throughput analyses. They can be used in screens for PrE-differentiating factors or in proteomics analyses to identify secreted factors that mediate tissue patterning (for example, during cardiac induction)[22–24]. XEN cells show paternally imprinted X-chromosome inactivation[3], and thus they serve as a useful model for understanding the molecular basis of post-translational and epigentic modifications[25]. Furthermore, as the extraembryonic endoderm has recently been shown to contribute to the gut endoderm *in vivo*[26], XEN cells may also serve as an alternative self-renewing source of definitive endoderm-derived cells and tissues for regenerative medicine. Intriguingly, XEN cells rapidly silence retroviral transcription compared with mouse ES and TS cells and may therefore also represent a useful model to study how viral gene expression is regulated[27].

XEN cell establishment from embryos

Several protocols describe the derivation of XEN cells from mouse embryos, including the direct plating of blastocyst stage embryos[3]

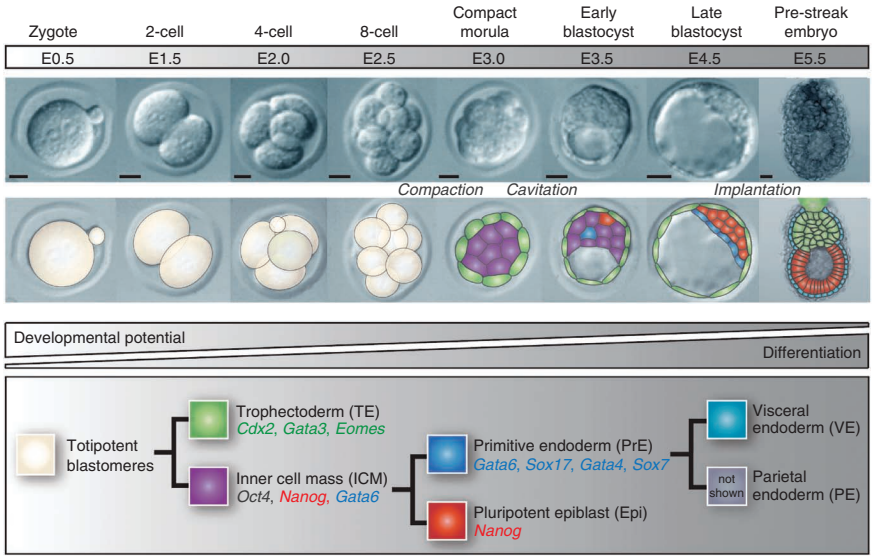


Figure 1 | Overview of early embryonic development. Proper lineage segregation before implantation is ensured by two cell-fate decisions, with the first giving rise to trophectoderm and inner cell mass, and the second leading to the allocation of primitive endoderm and epiblast. Lineage-associated gene expression is noted below each cell type. After implantation, the PrE differentiates into visceral and parietal endoderm. E: embryonic day. Scale bars, 50 μm.

or immunosurgery, to isolate and to plate the inner cell mass[28,29]. In 2005, Kunath *et al.*[3] reported the first isolation and characterization of XEN cells using methods based on existing protocols of deriving ES[28] and TS cells[2,30,31]. In this case, embryonic day 3.5 blastocysts were cultured on mouse embryonic fibroblast (MEF) cells until they formed an outgrowth that, as in routinely used ES and TS cell derivation protocols, was subsequently disaggregated to promote the proliferation of the different cell types present in the blastocyst. In both procedures, XEN cell outgrowths could be observed alongside ES or TS cells, respectively, suggesting that these conditions are also permissive for XEN cell derivation.

Our experience has been that slight modifications to these protocols (especially media conditions and time of disaggregation) can favor the propagation of XEN cells over that of ES or TS cells. We have also found that, in our hands, TS cell conditions are generally less efficient for the derivation of XEN cell lines compared with ES cell conditions (21% and 56% efficiency, respectively). ES cells start appearing from the outgrowths first and are less resilient, which might facilitate the ability of XEN cells to outcompete them in a culture dish. TS cells, in contrast, reach their proliferative peak around the same time as XEN cells and tend to outcompete XEN cells present in the same cultures, especially in the presence of recombinant FGF4 (ref. 2) or FGF2 (bFGF), the latter of which can elicit the same effect[31] and is slightly more cost effective.

XEN cell establishment by transcription factor overexpression

In classical experiments, Davis *et al.*[32] demonstrated that the expression of a single transcription factor, MyoD, was sufficient to convert fibroblasts into myogenic cells. More recently, transcription factor–mediated cell-fate switches have been demonstrated for several other cell types, including the reprogramming of fibroblasts into induced pluripotent stem cells via ectopic expression of the POU domain, class 5, transcription factor 1 (Pou5f1, also known as Oct4), Klf4, Sox2 and Myc (also known as c-Myc)[33].

Ectopic transcription factor expression has also been shown to induce the conversion of XEN cells from ES cells (reviewed in ref. 6). Niwa and colleagues[8,34] demonstrated that the expression of Gata4 or Gata6 alone is sufficient to induce the conversion of ES cells into XEN cells. Notably, these GATA-derived XEN cells share the molecular and functional characteristics of embryo-derived XEN cells, including contribution to PrE lineages in chimeric embryos[34]. XEN-like cells have also been generated by ectopic expression of Sox17 (refs. 14,35–37). It is unclear whether Sox17 alone is sufficient to drive XEN cell commitment, as these cells retain the expression of ES cell–associated genes, such as *Nanog* and *Oct4* (refs. 14,35,38), and their contribution to chimeras has not yet been reported. Moreover, it is unclear whether alternative PrE-associated transcription factors can be used to convert ES to XEN cells.

The simplest method used to derive XEN cells from ES cells is the expression of *Gata6* either by transfection of a circular plasmid or a linearized DNA using chemical or nonchemical (e.g., electroporation) methods. There are several chemical-based transfection approaches. The most widely used methods for ES cells are lipid-based lipopolyamines[39] and cationic polymer–based transfection (reviewed in ref. 40). Lipopolyamines (e.g., lipofection) work by coating a *Gata6*-cDNA plasmid with a cationic lipid, allowing the DNA to cross the negatively charged phospholipid ES cell membrane by endocytosis[41,42]. Cationic polymers (e.g., Xfect) act by binding DNA and condensing it, thereby facilitating DNA entry into the cell cytoplasm[40,43]. An advantage of lipopolyamines and cationic polymers is that they do not require specialized equipment, although the reagents may be cytotoxic and the ratio of DNA to transfection reagent needs to be optimized to the specific cell type[42] (reviewed in ref. 44).

Electroporation is the principal alternative to chemical-based transfection approaches and is routinely used for ES cell transfection[45–47]. Electroporation allows for permeabilization of the ES cell membrane and transfection of a *Gata6* cDNA via the application of a transient electrical field. Depending on the purity of the DNA preparation, this can be a highly efficient method and is most commonly used for gene targeting of ES cells. The choice of whether the *Gata6* cDNA plasmid is introduced as a circular plasmid or as a linearized DNA can affect transfection efficiency, as supercoiled or open-circular DNA is optimal for transient expression, whereas linearized DNA is more recombinogenic and best suited for stable transfection[48–50].

Continued passage of transiently transfected cells will result in the dilution of the exogenous DNA and may restrict the time window for subsequent analysis. Therefore, although it is initially more time consuming, the generation of stable transfected ES cells that have integrated a linearized *Gata6* cDNA into the genome can be experimentally more consistent compared with transient transfection. By including a drug-resistance gene in the *Gata6* cDNA–targeting plasmid, drug selection can be used to expand only targeted ES cells that retain the ability to overexpress the *Gata6* cDNA.

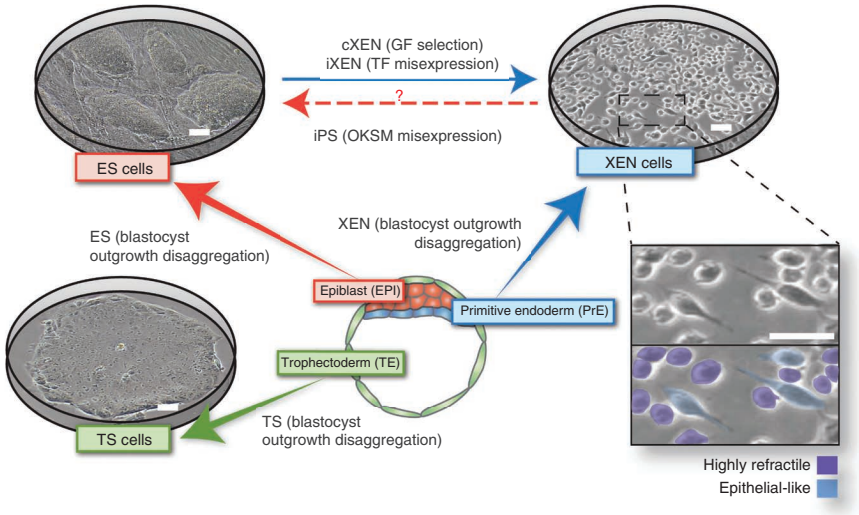


Figure 2 | Stem cell types that can be derived and propagated in culture representing the three blastocyst lineages. Embryonic stem (ES) cells represent the epiblast, trophoblast stem (TS) cells represent the trophectoderm and extraembryonic endoderm (XEN) cells represent the primitive endoderm cell lineage. Heterogeneities in XEN cell morphology are indicated: highly refractile phase-bright and epithelial-like. Cognate embryo–derived stem cells retain the expression of key lineage-associated genes. GF, growth factor; iPS, induced pluripotent stem; OKSM, Oct4, Sox2, Klf4 and c-Myc. Scale bars, 100 μm.

Another important consideration is the use of an efficient and robust promoter to drive exogenous *Gata6* expression. In mouse ES cells, the cytomegalovirus (*CMV*) promoter may be silenced[51]. In contrast, the ubiquitin C (*UBC*) promoter, elongation factor 1α (*EF1A*) promoter, and the chicken β-actin (*ACTB*) promoter coupled with the CMV early enhancer (*CAGG*) have been reported to be more robustly expressed[51].

Although these are not described here, alternative approaches for *Gata6* overexpression include viral transduction. Owing to potential toxicity, viral production and transduction require at least a category II tissue culture facility, and the silencing of retroviral transcription in XEN cells may obviate this approach[27]. Moreover, reversible overexpression of *Gata6* can be engineered using an inducible expression system containing a tetracycline-responsive element (controlled by tetracycline or doxycycline), a mutant estrogen-responsive element (controlled by 4-hydroxytamoxifen) or a glucocorticoid-responsive element (controlled by dexamethasone). When combined with drug selection, an inducible system is an effective means of overexpressing *Gata6* simultaneously in all cells. Although overexpression of *Gata6* may initially be a simpler method to drive XEN differentiation, a major note of caution is that its ubiquitous expression may hinder the ability of converted XEN cells to differentiate into subtypes of XEN cells. One major limitation may be that constitutive *Gata6* expression blocks differentiation of induced XEN cells into visceral endoderm subtypes, as *in vivo* these cells lack *Gata6* expression while retaining the expression of *Gata4* (refs. 52,53). Here we describe a method to constitutively express *Gata6* in mouse ES cells, which can be modified to overexpress other genes such as *Gata4* or to introduce an inducible expression system as has been described previously[34,54,55]. Moreover, transposon systems such as *piggyBac*[56] may also be used, although these methods have not yet been shown to convert ES to XEN cells.

XEN cell establishment by growth factor conversion of ES cells

Although ES cells largely contribute to the Epi in chimera embryos, it has been observed that committed XEN and TS cells also arise within pluripotent culture conditions[14,57,58]. The presence of TS- and XEN-like cells within ES cell cultures suggests that ES cells may have a broader cell fate potential and are therefore able to differentiate into stable extraembryonic stem cell lines directly. Indeed, ES cell aggregation results in the formation of embryoid bodies with an outer layer of XEN cells[12], and growth factors such as retinoic acid have been shown to differentiate embryonic carcinoma and ES cells into XEN-like cells[9–11,59]. Moreover, the culture of ES cells in serum-free medium on fibronectin-coated dishes at high density has been suggested to promote visceral endoderm differentiation[60]. However, it was unclear whether self-renewing XEN cells could be differentiated directly from mouse ES cells without gene manipulation. We have recently expanded on these studies by developing a technique to convert ES cells directly into stable XEN (cXEN) cell lines that are equivalent to embryo-derived XEN cells[13]. We previously demonstrated that cXEN cells are molecularly indistinguishable from embryo-derived XEN and iXEN (i.e., transcription factor induced) cells and are equivalently responsive to differentiation-promoting factors (i.e., bone morphogenetic protein (BMP)-induced visceral endoderm differentiation). Our efforts to generate chimera embryos from cXEN cells resulted in contribution to the parietal endoderm at early postimplantation stages (K.K.N., M. Kang and A.-K.H., unpublished observations), as has been reported for XEN cells derived from embryos[3]. Here we describe our method for converting mouse ES cells into stable cXEN cell lines using growth factors. Specifically, we use retinoic acid together with activin, which has been demonstrated to promote primitive endoderm development *in vivo* and *in vitro*[17,61]. Notably, although retinoic acid and activin have been shown to differentiate ES cells into neurons and definitive endoderm, respectively, we find that the combination of both factors promotes XEN cell differentiation. We previously tested a range of retinoic acid concentrations from 0 to 100 μM and a range of activin concentrations from 0 to 20 ng ml^{-1}. Although XEN-like cells can emerge in the absence of exogenous retinoic acid and activin[14], the proportion of XEN-like cells is enhanced (40.3%) in 0.01 μM retinoic acid and 10 ng ml^{-1} activin[13]. We therefore use the lowest dose of retinoic acid and activin that gives us the highest efficacy of XEN cell emergence, although it is presently unclear whether different concentrations of these growth factors influence the identity of the XEN cells obtained. Thus far, we have been able to derive XEN cell lines from any mouse ES cell line and mouse strain unless the cells are mutant for a gene required for XEN establishment and/or maintenance[13], although the speed of conversion can vary with any given cell line. An important consideration is the initial heterogeneity that emerges after growth factor treatment. As with the protocol for XEN cell derivation from embryos, selecting cells with XEN cell morphology facilitates the expansion of stable cXEN cells that are able to self-renew indefinitely. This ES cell conversion protocol is particularly useful for determining the genetic requirement for XEN cell function using mutant mouse ES cells without the necessity for mutant mice.

MATERIALS

REAGENTS

Derivation from embryos using modified ES cell conditions

- Pregnant mouse, 3.5 days post coitum (d.p.c.) ! CAUTION Experiments involving rodents must conform to all relevant institutional and governmental regulations.
- Mitotically inactivated MEF cells (Reagent Setup)
- MEF medium (DMEM supplemented with FBS (Gibco, cat. no. 10082147), L-glutamine, β-mercaptoethanol, MEM non-essential amino acids and sodium pyruvate. Optionally, add penicillin-streptomycin; see Reagent Setup)
- M2 medium (Millipore, cat. no. MR-015-D)
- ES cell medium (DMEM supplemented with FBS (ES cell qualified; Gibco, cat. no. 16141079), L-glutamine (Gibco, cat. no. 25030-081), β-mercaptoethanol (Sigma-Aldrich, cat. no. M6250), MEM non-essential amino acids (Gibco, cat. no. 11140035), sodium pyruvate (Gibco, cat. no. 11360-070), leukemia inhibitory factor (LIF; Millipore, cat. no. ESG1106); optional: penicillin-streptomycin (Invitrogen, cat. no. 15140122); see Reagent Setup)
- Ethanol, 70% (vol/vol; Sigma-Aldrich, cat. no. E7023) ! CAUTION Ethanol is flammable and may be harmful if inhaled or ingested.
- PBS, 1×, sterile (Ca^{2+} and Mg^{2+} free; Gibco, cat. no. 14190-094)
- Trypsin-EDTA, 0.25% (wt/vol; Gibco, cat. no. 25200-056)

Derivation from embryos using modified TS cell conditions

- Pregnant mouse, 3.5 d.p.c.
- Mitotically inactivated MEF cells (Reagent Setup)
- TS cell medium (includes advanced RPMI 1640 (Gibco, 12633-012), FBS (Gibco, cat. no. 10082147), L-glutamine, β-mercaptoethanol, sodium pyruvate; optional: 1% (vol/vol) penicillin-streptomycin. Aliquots are supplemented with recombinant FGF2 or FGF4 (Sigma-Aldrich, cat. no. F8424) and heparin (Sigma-Aldrich, cat. no. H3393); see Reagent Setup)
- MEF-conditioned medium (Reagent Setup)
- M2 medium (Millipore, cat. no. MR-015-D)
- Ethanol, 70% (vol/vol; Sigma-Aldrich, cat. no. E7023) ! CAUTION Ethanol is flammable and may be harmful if inhaled or ingested.
- PBS, 1×, sterile (Ca^{2+} and Mg^{2+} free; Gibco, cat. no. 14190-094)
- Trypsin-EDTA, 0.25% (wt/vol; Gibco, cat. no. 25200-056)

Derivation from mouse ES cells by *Gata6* overexpression

- Mitotically inactivated MEF cells (Reagent Setup)
- MEF medium (Reagent Setup)
- ES cells (Reagent Setup)
- ES cell medium (Reagent Setup)
- Ethanol, 70% (vol/vol; Sigma-Aldrich, cat. no. E7023) ! CAUTION Ethanol is flammable and may be harmful if inhaled or ingested.
- Gelatin, 0.1% (wt/vol; Sigma-Aldrich, cat. no. G1890) made up in sterile dH_2O (Sigma-Aldrich, cat. no. W1503) and autoclaved to dissolve
- PBS, 1×, sterile (Ca^{2+} and Mg^{2+} free; Gibco, cat. no. 14190-094)
- Trypsin-EDTA, 0.05% (wt/vol; Gibco, cat. no. 25300054)
- Xfect (Clontech, cat. no. 631320) or Lipofectamine 2000 (Invitrogen, cat. no. 11668030)
- Opti-MEM I (Invitrogen, cat. no. 51985)
- Bio-Rad Gene Pulser II Xcell eukaryotic system (Bio-Rad, cat. no. 377267)
- Cuvette, 4 mm (Cell Projects, cat. no. EP-104)
- Agarose (Sigma-Aldrich, cat. no. A9414)
- Maxi-prep kit (Qiagen, cat. no. 12262)
- *Gata6* cDNA vector (Origene, cat. no. MC219384)
- Isopropanol (Sigma-Aldrich, cat. no. I9516)
- Sodium acetate (Fluka, cat. no. 71196)
- Wizard SV genomic DNA purification system (Promega, cat. no. A2360)

Derivation from mouse ES cells using growth factors

- cXEN derivation medium (standard XEN medium supplemented with all-*trans* retinoic acid (Sigma-Aldrich, cat. no. R2625) dissolved in DMSO plus activin A (R & D Systems, cat. no. 338-AC-010); Reagent Setup)
- Mitotically inactivated MEF cells (Reagent Setup)
- MEF medium (Reagent Setup)
- ES cells (Reagent Setup)
- ES cell medium (Reagent Setup)
- Ethanol, 70% (vol/vol; Sigma-Aldrich, cat. no. E7023) ! CAUTION Ethanol is flammable and may be harmful if inhaled or ingested.
- Gelatin, 0.1% (wt/vol; Sigma-Aldrich, cat. no. G1890) made up in sterile dH_2O (Sigma-Aldrich, cat. no. W1503) and autoclaved to dissolve
- PBS, 1×, sterile (Ca^{2+} and Mg^{2+} free; Gibco, cat. no. 14190-094)
- Trypsin-EDTA, 0.05% (wt/vol; Gibco, cat. no. 25300054)

EQUIPMENT

- Conical tubes (15 ml; BD Falcon, cat. no. 352095)
- Conical tubes (50 ml; BD Falcon, cat. no. 352070)
- Tissue culture–treated plate (six well; Corning, cat. no. 3516)
- Tissue culture–treated plate (12 well; Corning, cat. no. 3512)
- Tissue culture–treated plate (24 well; Corning, cat. no. 3524)
- Tissue culture–treated dish (100 mm; Corning, cat. no. 430167)
- Tissue culture–treated dish (150 mm; Corning, cat. no. 430599)
- Tissue culture–treated plate (96-well flat bottom; Corning, cat. no. 3595)
- Tissue culture–treated plate (96-well round bottom; Corning, cat. no. 3799)
- Plate (four well; Nunc, cat. no. 12566300)
- Cryotube vials (Nunc, cat. no. 144444)
- Filter, 0.22 μm (Corning, cat. no. 431097)
- Flame-pulled glass Pasteur pipettes (unpulled; e.g., Fisher Scientific, 13-678-20D)
- Dissecting microscope (e.g., Leica M80)
- Inverted microscope (e.g., Olympus CKX41)
- Water bath (e.g., Grant JB Aqua 12)
- Centrifuge (e.g., Eppendorf Centrifuge 5804)
- Hemocytometer (e.g., Hausser scientific, cat. no. 3110)
- CO_2 incubator (e.g., Sanyo)

Dissection tools

- Scissors (e.g., Roboz, cat. no. RS-5910)
- Forceps (e.g., Dumont no. 5, Roboz; cat. no. RS-4976)
- Syringes with 27-gauge needles (10 ml; BD Biosciences, cat. no. 309623)

Mouth-controlled pipette (Reagent Setup)

- Plastic mouthpiece (MedTech, cat. no. 1501P-B4036-2)
- Rubber tubing (Fisher Scientific, cat. no. 22-362-772)
- Filtered pipette tip, P1000 (USA Scientific, cat. no. 1126-7810)

REAGENT SETUP

Mitotically inactivated MEF cells Mitotically inactivated MEFs should be plated on pregelatinized cell culture–treated plates 1 d before initiating XEN derivation MEFs can be derived from CF1 mice that are dissected 12.5–13.5 d.p.c.; cells are passaged 3–5 times before inactivation, either by γirradiation or by mitomycin C treatment.

ES cells ES cells should be thawed ~1–2 weeks before initiating XEN differentiation and passaged at least twice. ES cells are typically maintained on MEFs.

ES cell medium Prepare a solution of the following: DMEM supplemented with 15% (vol/vol) FBS (ES cell qualified), 2 mM L-glutamine, 0.1 mM β-mercaptoethanol, 0.1 mM MEM non-essential amino acids, 1 mM sodium pyruvate and 10^3 IU leukemia inhibitory factor. Optionally, add 1% (vol/vol) penicillin-streptomycin. Medium should be prepared in a sterile tissue culture hood and filter sterilized. The medium can be stored for up to 1 month at 4 °C.

MEF medium Prepare a solution of the following: DMEM supplemented with 15% (vol/vol) FBS, 2 mM L-glutamine, 0.1 mM β-mercaptoethanol, 0.1 mM MEM non-essential amino acids and 1 mM sodium pyruvate. Optionally, add 1% (vol/vol) penicillin-streptomycin. Medium should be prepared in a sterile tissue culture hood and filter sterilized. The medium can be stored for up to 1 month at 4 °C.

TS cell medium Prepare a solution of the following: advanced RPMI 1640 supplemented with 20% (vol/vol) FBS, 2 mM L-glutamine , 0.1 mM β-mercaptoethanol, 1 mM sodium pyruvate; optional: 1% (vol/vol) penicillin-streptomycin. Medium should be prepared in a sterile tissue culture hood and filter sterilized. The medium can be stored for up to 1 month at 4 °C. Prepare a freshly made aliquot supplemented with 24 ng ml^{-1} recombinant FGF2 or FGF4 and 1 μg ml^{-1} heparin.

MEF-conditioned medium Preparation of MEF-conditioned medium should begin at least 10 d before initiating the XEN derivation protocol B. Thaw and culture three 150-mm tissue culture plates of mitotically inactivated MEFs in 25 ml of standard TS cell medium (not supplemented with FGF2/FGF4 and heparin) for 3 d. Collect the medium in 50-ml conical tubes and store the medium at −20 °C. Add another 25 ml of TS cell medium to each plate and repeat the procedure twice until 225 ml of MEF-conditioned medium have been collected. Thaw the frozen batches and centrifuge the

Box 1 | Preparation of *Gata6* cDNA for transfection or electroporation

● TIMING ~2 d

1. In order to ensure sufficient quantity of plasmid DNA, a large-scale preparation of *Gata6* cDNA plasmid by maxiprep is required for transfection. A circular *Gata6* cDNA plasmid is generally used for chemical-based transfection, whereas for electroporation, the plasmid should be linearized using a unique restriction enzyme at a non-essential region, which will not affect the function of the gene and drug selection.
2. If linearized DNA is required then cut one time with a unique restriction enzyme that is in a region of the plasmid that will not affect the function. Linearize 100 µg of *Gata6* cDNA plasmid by restriction enzyme digestion overnight at the optimal incubation temperature. Run a small aliquot on a 0.8% (wt/vol) agarose gel to ensure linearization.
3. For high-purity plasmid preparation, precipitate the linearized *Gata6* cDNA with a 1/10-volume of salt (3 M sodium acetate) and 2.5 volumes of 100% ethanol; store at −20 °C overnight.
4. The next day, spin the DNA at 16.1 × 10^3*g* for 5 min at 4 °C and wash the pellet once with 500 µl of 70% (vol/vol) ethanol.
5. Remove the supernatant and allow any remaining ethanol to evaporate in the hood for several min.

▲ CRITICAL STEP To ensure sterility, work in a sterile tissue culture hood for this and subsequent steps.

6. Resuspend the DNA pellet in 100 µl of sterile PBS without Mg^{2+} or Ca^{2+} (a final DNA concentration of 1 µg $µl^{-1}$).

▲ CRITICAL STEP The presence of salts could be detrimental to transfection efficiency and interfere with electroporation.

■ PAUSE POINT Plasmid DNA preparation can be done any time before transfection. The resuspended DNA can be kept at −20 °C for several months.

medium at 2,300*g* for 20 min at room temperature (20–25 °C) to remove debris. Filter-sterilize the medium and store it at 4 °C.

Standard XEN medium Prepare a solution of advanced RPMI 1640 supplemented with 15% (vol/vol) FBS and 0.1 mM β-mercaptoethanol. Optionally, add 1% (vol/vol) penicillin-streptomycin. Medium should be prepared in a sterile tissue culture hood and filter sterilized. The medium can be stored for up to 1 month at 4 °C. Advanced RPMI 1640 contains glutamine; alternatively, RPMI medium can be used and requires the addition of 2 mM L-glutamine.

cXEN derivation medium Freshly prepare a solution of standard XEN medium supplemented with 0.01 µM all-*trans* retinoic acid dissolved in DMSO plus 10 ng ml^{-1} activin A. A wide range of retinoic acid concentrations can be used, ranging from 0.01 to 10 µM. The derivation medium can also be supplemented with 24 ng ml^{-1} recombinant FGF2 or FGF4 and 1 µg ml^{-1} heparin, although this is not required if endogenous *Fgf4* is intact.

! CAUTION Retinoic acid is light sensitive and may be harmful if ingested or absorbed.

Freezing medium Freshly prepare a solution of 10% (vol/vol) DMSO (Sigma-Aldrich, cat. no. D2650) and 90% (vol/vol) FBS. Keep the solution on ice or at 4 °C until immediately before use.

***Gata6* cDNA for transfection** Prepared as described in **Box 1**.

EQUIPMENT SETUP

Mouth-controlled pipette Assemble a mouth-controlled pipette. This is used for handling embryos and for dissociating blastocyst *in vitro*–matured outgrowths. Mouth-controlled pipettes can be assembled in a variety of designs (for variations see Nagy *et al.*[28]). A simple mouth pipettor can be made by inserting the pointed end of a P1000 filter tip into one end of a cut rubber tube and attaching a plastic mouthpiece to the other. A pulled glass Pasteur pipette can now be attached to the filter tip. Pasteur pipettes can be pulled either over a flame or by using a micropipette puller (**Fig. 3**).

PROCEDURE

Preparation of MEF feeders ● TIMING 15–30 min

1| Pregelatinize tissue culture plates for >10 min with 1–2 ml of 0.1% (wt/vol) gelatin. Aspirate the gelatin before seeding the cells.

? TROUBLESHOOTING

2| One day before starting the derivation, place a frozen vial of mitotically inactivated MEFs in a 37 °C water bath until nearly thawed and transfer the contents into a 15-ml Falcon tube containing 5 ml of MEF medium. Centrifuge the medium at 200*g* for 4 min at room temperature and aspirate the supernatant.

? TROUBLESHOOTING

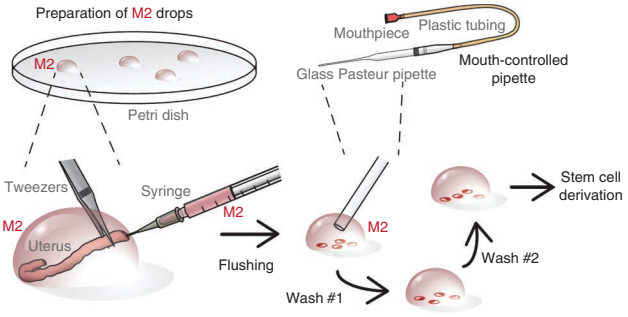


Figure 3 | Recovery of blastocyst-stage embryos from uteri of adult female mice. Several drops of M2 medium are prepared on the lid of a 100-mm dish. The dissected and cleaned uterine horn of a pregnant female (3.5 d.p.c.) is placed in one drop. While securing the uterine horn with forceps, the needle of a 1-ml syringe is inserted and the uterine horn is flushed with ~0.2 ml of M2 medium. The flushed blastocysts are located under high magnification and transferred to a fresh drop of M2 medium using a mouth-controlled pipette. This transfer is repeated at least twice to wash away debris, lipid drops and blood cells.

3| Resuspend the cell pellet in MEF medium and plate it at a density of 6 × 10^4 MEFs per well for a four-well plate, 1.3 × 10^5 MEFs per well for a six-well plate or 7.8 × 10^5 cells per 100-mm tissue culture plate, and then incubate the plates overnight at 37 °C. The number of plates to prepare depends on the number of XEN cell derivations to be performed. Generally, four plates (4-well plates) per pregnant mouse for Steps 6A and 6B are sufficient, or a six-well plate for Steps 6C–F.
? TROUBLESHOOTING

4| The morning before starting the derivation, aspirate the medium on the MEF feeders and replace it with ES cell medium for derivation using ES cell conditions (Step 6A), with TS cell medium for derivation using TS cell conditions (Step 6B) or with XEN cell medium for derivation from ES cells (Steps 6C–F). When required, plate MEFs that carry drug-resistance genes, such as DR4 MEFs (resistant to neomycin, hygromycin, puromycin and Zeocin) (**Table 1**).
? TROUBLESHOOTING

Cell preparation

5| If you are deriving cells using ES or TS cell conditions, follow option A to collect E3.5 blastocysts. If you are deriving XEN cells from ES cells, follow option B.

(A) Collecting E3.5 blastocysts (flushing)

(i) Dissect out the uterine horns of a pregnant mouse at 3.5 d.p.c. and place them on the lid of a 100-mm culture dish in a drop of M2 medium.

(ii) Under a dissection microscope (approximately ×10 total magnification), remove large fat pads and cut away the ovaries and oviducts. Place each of the cleaned uterine horns in a fresh drop of M2 medium.

(iii) With a 1-ml syringe, take up 0.5 ml of M2 medium and attach a needle (27 gauge). While securing the uterine horn with forceps, insert the needle and flush the uterine horn with ~0.2 ml of M2 medium (**Fig. 3**) and remove the uterine horn from the drop. Repeat this process with the second uterine horn.

(iv) Under higher magnification (~×30–×40), find the flushed blastocysts in both drops.
? TROUBLESHOOTING

(v) By using a mouth-controlled pipette, transfer the blastocysts to a new drop of M2 medium. Repeat the step twice to wash away debris, lipid drops and blood cells.

(vi) Proceed with the derivation of XEN cells using either option A (ES cell conditions) or option B (TS cell conditions) in Step 6.

(B) ES cell maintenance in medium with serum

(i) Thaw the ES cells rapidly in a 37 °C water bath and transfer them into a 15-ml Falcon tube containing 5 ml of MEF medium.

(ii) Centrifuge the tube at 200*g* for 4 min at room temperature, aspirate the supernatant and resuspend the ES cell pellet using 1 ml of ES cell medium.

(iii) Aspirate the medium on the irradiated MEFs and replace it with ES cell medium. Plate ~3 × 10^5 ES cells onto MEFs in one well of a six-well plate or 2 × 10^6 ES cells on a 100-mm tissue culture dish. Incubate the cells at 37 °C.
▲ **CRITICAL STEP** All 37 °C incubations should be carried out in a 5% CO_2 humidified incubator.

(iv) Change the ES cell medium daily to maintain the cell line and monitor the cells. Proceed to the next step (passage) when they reach ~80% confluency; this typically occurs every 2–3 d.

(v) To passage the ES cells, wash the well that contains the cells with 1 ml of 1× PBS to remove residual proteins. Aspirate the PBS and add 0.25 ml of 0.05% (wt/vol) trypsin for a six-well plate. Rock the plate back and forth to dissociate the cells. Incubate the plate at room temperature for no more than 5 min until all the cells are nearly detached. Do not leave the cells in trypsin for longer than necessary.

(vi) Neutralize the trypsin with 1 ml of ES cell medium. Pipette the solution up and down with a P1000 pipette approximately ten times to dissociate the cells into small clumps and single cells.

(vii) With a P1000 pipette, add 50 µl of the dissociated ES cells onto a freshly prepared MEF well in 2 ml of fresh ES cell medium to continue growing the cell line at a ~1/25 dilution (roughly, a 1/10 dilution should be ready for passaging in 2 d, a 1/25 dilution in 3 d and a 1/40 dilution in 4 d). The trace amount of trypsin on the first day will not adversely affect the cells.
▲ **CRITICAL STEP** The use of a pipette smaller than P1000 can damage the cells and should be avoided, except to pick colonies or count cells on a hemocytometer. In order to ensure healthy recovery after thawing, ES cells should be passaged at least once before proceeding with the steps below.

(viii) If you are electroporating the cells for XEN derivation, passage the cells the day before. If the cells are grown on MEF, MEF-deplete the cells at the same time (see Step 5B(ix)).

(ix) MEF-deplete the ES cells. This step is unnecessary if the ES cells are grown in MEF-free conditions. Plate the cells onto a pregelatinized plate in ES cell medium the day before proceeding to XEN derivation. Alternatively, ES cells may be MEF-depleted on the day of derivation by plating them on a pregelatinized 100-mm plate containing 7–10 ml of ES medium (no LIF) for 30 min. After 30 min, the majority of MEFs should have adhered to the pregelatinized plate. Tilt the 100-mm plate to collect the supernatant medium plus ES cells and transfer it into a 15-ml Falcon tube. Do not wash the plate with fresh medium as this may dislodge the attached MEFs.

XEN derivation ● TIMING 15–20 d

6| If you are deriving XEN cells under ES cell conditions, follow option A. If you are deriving XEN cells under TS cell conditions, follow option B. If you are chemically transfecting ES cells with a *Gata6* cDNA plasmid using cationic polymer, follow option C. If you are chemically transfecting ES cells with a *Gata6* cDNA plasmid using lipopolyamine, follow option D. If you are electroporating ES cells with a *Gata6* cDNA plasmid, follow option E. If you are deriving XEN cells from ES cells using growth factors, follow option F.

? TROUBLESHOOTING

(A) XEN cell derivation under ES cell conditions

(i) *Day 1*. Thoroughly clean a dissection microscope with 70% (vol/vol) ethanol. By using a mouth-controlled pipette, place one blastocyst per feeder-covered well in the prepared four-well plates (**Fig. 4**).
▲ **CRITICAL STEP** The plates should be incubated in a 5% CO_2 humidified incubator at 37 °C, except while performing Step 6A(i–iv), which can be performed under ambient conditions.
? TROUBLESHOOTING

(ii) *Day 2*. Observe the plates under a microscope. The blastocysts should have hatched from the zona pellucida and attached to the feeder layer (**Fig. 4**).

(iii) *Day 3*. Observe the plates under a microscope. The blastocysts should have started to form an outgrowth (**Fig. 4**). Carefully aspirate the medium and replace it with fresh ES cell medium.
? TROUBLESHOOTING

(iv) *Day 4*. Thoroughly clean a dissection microscope with 70% (vol/vol) ethanol. Dilute 0.25% (wt/vol) trypsin to a sufficient volume of 0.01% (wt/vol) trypsin/1× PBS. Wash the cells with 500 µl of 1× PBS. Aspirate the PBS and add 100 µl of 0.1%

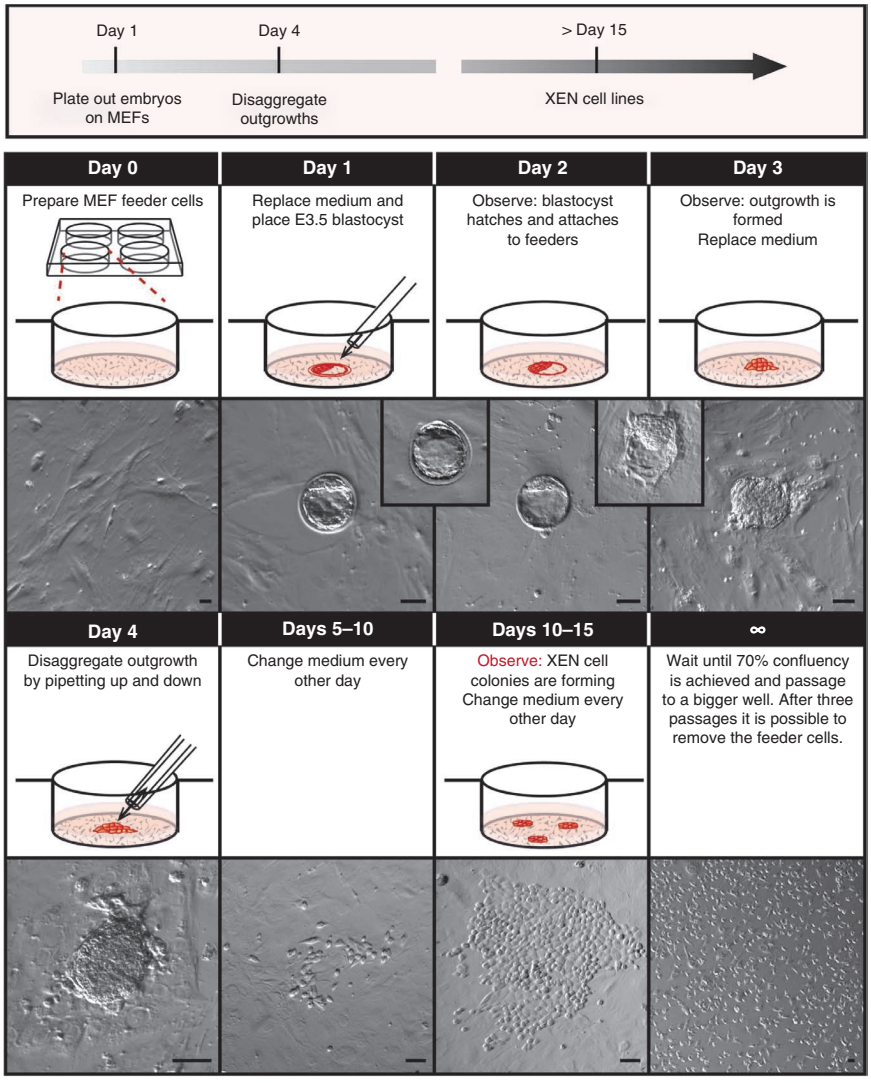


Figure 4 | Timeline for XEN cell derivation from mouse blastocysts. Protocol for the derivation of XEN cells from blastocysts. Day 0: feeder cell–coated four-well plates are prepared. Day 1: medium is replaced and a freshly flushed E3.5 blastocyst is placed in the center of the well. Day 2: the blastocyst hatches and attaches to the feeder cells. Day 3: the blastocyst has formed an outgrowth. Medium is replaced. Day 4: the prominent outgrowth is trypsinized and disaggregated with a P20 pipette. Days 5–10: the medium is replaced every other day and the plate is observed daily. Days 10–15: XEN-like cells with stellate and refractile morphology will emerge. The medium is replaced every other day and the plate is observed daily. Day 15 and onward: After 70% confluency is attained, the XEN cells are passaged onto a well in a six-well plate and subsequently into a 100-mm dish. After three more passages they can be MEF-depleted and the XEN cell line can be frozen. Scale bars, 100 µm.

(wt/vol) trypsin to each well. Incubate the plates at 37 °C for 5 min. While observing the cells under a cleaned dissection microscope, use a P20 pipette to disaggregate the outgrowth by pipetting up and down several times. Add 400 µl of fresh ES cell medium.

(v) *Day 5*. Carefully replace the medium with fresh ES cell medium.

(vi) *Days 6–14*. Replace the medium every other day until XEN cell colonies can be observed (**Fig. 4**).

(vii) *Day 15–indefinite*. When 70% confluency is achieved, passage the cells onto a feeder-covered well of a six-well plate. From this point onward, the cells can be cultured in standard XEN medium. Repeat the step, but passage the cells onto a feeder-covered 100-mm dish. Passage the cells two more times. If desired, remove feeders and freeze the XEN cell line as described in Step 7.

(B) XEN cell derivation under TS cell conditions

(i) *Day 1*. Thoroughly clean a dissection microscope with 70% (vol/vol) ethanol. By using a mouth-controlled pipette, place one blastocyst per feeder-covered well in the prepared four-well plates (**Fig. 4**).
▲ **CRITICAL STEP** The plates should be incubated in a 5% CO_2 humidified incubator at 37 °C, except while performing (Step 6B(i–iv)), which can be performed under ambient conditions.
? TROUBLESHOOTING

(ii) *Day 2*. Observe the plates under a microscope. The blastocysts should have hatched from the zona pellucida and attached to the MEF feeder layer (**Fig. 4**).

(iii) *Day 3*. Observe the plates under a microscope. The blastocysts should have started to form an outgrowth (**Fig. 4**). Carefully aspirate the medium and replace it with fresh TS cell medium.
? TROUBLESHOOTING

(iv) *Day 4*. Thoroughly clean a dissection microscope by wiping it with 70% (vol/vol) ethanol. Dilute 0.05% (wt/vol) trypsin to a sufficient volume of 0.01% (wt/vol) trypsin/1× PBS. Wash the cells with 500 µl of 1× PBS. Aspirate the PBS and add 100 µl of 0.01% (wt/vol) trypsin to each well. Incubate the plates at 37 °C for 5 min. While observing the cells under a dissection microscope, use a P20 pipette to disaggregate the outgrowth by pipetting up and down several times. Add 400 µl of 30% TS cell medium/70% MEF-conditioned medium.

(v) *Day 5*. Carefully replace the medium with fresh 30% TS cell medium/70% MEF-conditioned medium.

(vi) *Days 6–14*. Replace the medium every other day until XEN cell colonies are observed (**Fig. 4**).

(vii) *Day 15–indefinite*. When cells reach 70% confluency, passage the cells onto a feeder-covered well of a six-well plate. From this point onward, the cells can be cultured in standard XEN medium. Repeat the step, but passage the cells onto a feeder-covered 100-mm dish. Passage the cells two more times. If desired, remove the feeders and freeze the XEN cell line as described in Step 7.

(C) Chemical transfection of ES cells with a *Gata6* cDNA plasmid using a cationic polymer (e.g., Xfect) for transient or stable transfection

(i) *Day 1*. Re-plate MEF-depleted cells onto pregelatinized plates. Cells plated at a density of 7×10^4 per well of a 24-well plate will be ready for transfection ~12 h after plating.

(ii) *Day 2*. On the day of transfection, ensure that the ES cells are ~80% confluent. All subsequent volumes are for one well of a 24-well plate and may be scaled up or down according to the manufacturer's recommendations (Clontech).

(iii) Prepare two tubes: the first tube contains a 25-µl solution of 1 µg of *Gata6* cDNA plasmid diluted with Xfect reaction buffer in a sterile 1.5-µl Eppendorf tube; the second tube contains 0.5 µl of Xfect polymer (prevortexed) plus 24.5 µl of Xfect reaction buffer in a separate 1.5-µl Eppendorf tube. Use an empty plasmid (i.e., typically the same plasmid containing a constitutively active fluorescent reporter instead of *Gata6*) or sterile dH_2O as negative controls.

(iv) Vortex each tube, and then add the polymer solution to the DNA solution (total volume = 50 µl) and vortex again. Incubate the DNA-polymer solution for 10 min at room temperature.

(v) To one well of the 24-well plate, add the 50 µl of DNA-polymer solution dropwise. Mix the solution gently by tilting the plate back and forth, and then incubate the plate in a 5% CO_2 humidified incubator at 37 °C for 3 h.

(vi) Aspirate the medium containing the DNA-polymer complex from the ES cells, replace it with 0.5 ml of fresh ES cell medium and incubate the plate in a 5% CO_2 humidified incubator at 37 °C overnight.
? TROUBLESHOOTING

(vii) *Day 3*. Replace the ES medium in each well with 0.5 ml of XEN medium and incubate the plate overnight in a 5% CO_2 humidified incubator at 37 °C.

(viii) *Days 4–10*. If the *Gata6* cDNA plasmid has been engineered to also express a drug-resistance gene, then drug selection can be used to expand successfully transfected cells. Replace with XEN cell medium supplemented with the appropriate drug selection (**Table 1**). Replace with fresh XEN medium containing drug selection daily for ~10 d. XEN-like cells should appear within 5 d.

(ix) *Day 11–indefinite*. Enzymatically passage the *Gata6* cDNA–transfected cells with 0.05% (wt/vol) trypsin onto a pregelatinized 100-mm dish. Continue selection with drugs in order to establish a XEN cell line. Collect a fraction of the cells for genotyping and freeze the remaining XEN cells as described in Step 7.

(D) Chemical transfection of ES cells with a *Gata6*-cDNA plasmid using lipopolyamine (e.g., Lipofectamine 2000) for transient or stable transfection

(i) *Day 1*. Re-plate MEF-depleted cells onto pregelatinized plates. Cells plated at a density of 7×10^4 per well of a 24-well plate will be ready for transfection ~12 h after plating.

(ii) *Day 2*. On the day of transfection, ensure that the ES cells are ~80% confluent. All subsequent volumes are for one well of a 24-well plate and may be scaled up or down according to the manufacturer's recommendations (Invitrogen).

(iii) Prepare two tubes: the first tube contains a 50-µl solution containing 1 µg of Gata6 cDNA plasmid diluted with Opti-MEM I in a sterile 1.5-µl centrifuge tube; the second tube contains 4.5 µl of Lipofectamine 2000 plus 45.5 µl of Opti-MEM I in a separate 1.5-µl centrifuge tube. Use an empty plasmid (i.e., typically the same plasmid containing a constitutively active fluorescent reporter instead of *Gata6*) or sterile dH_2O as negative controls. Incubate the tubes for 5 min at room temperature.

(iv) Gently mix the diluted plasmids (*Gata6* or control) with the diluted Lipofectamine reagent (total volume = 100 µl) and incubate the tubes for 20 min at room temperature to allow for the formation of DNA-liposome complexes. During the 20-min incubation period, replace the medium of the plated cells with 0.5 ml of fresh ES cell medium without antibiotics.

(v) To one well of the 24-well cell culture dish, add 100 µl of DNA-liposome complex dropwise. Mix gently by tilting the plate back and forth, and then incubate the plate overnight in a 5% CO_2 humidified incubator at 37 °C.
? TROUBLESHOOTING

(vi) *Day 3*. Replace the ES medium in each well with 0.5 ml of XEN medium and incubate the plate overnight in a 5% CO_2 humidified incubator at 37 °C.

(vii) *Days 4–10*. If the *Gata6*-cDNA plasmid has been engineered to also express a drug-resistance gene, then drug selection can be used to expand successfully transfected cells. Replace the medium with XEN cell medium supplemented with the appropriate drug selection (**Table 1**). Replace with fresh XEN medium containing drug selection daily for ~10 d. XEN-like cells should appear within 5 d.

(viii) *Day 11–indefinite*. Enzymatically passage the *Gata6* cDNA–transfected cells with 0.05% (wt/vol) trypsin onto a pregelatinized 100-mm dish. Continue selection with drugs to establish a XEN cell line. Collect a fraction of the cells for genotyping and freeze the remaining XEN cells as described in Step 7.

TABLE 1 | Drug selection.

Marker	Gene product	Selection method
Neo	Aminoglycoside phosphotransferase; *neo* gene from the bacterial transposon *Tn5*	Select cells in G418 (0.1–1.0 mg ml^{-1}), an aminoglycoside that blocks protein synthesis and is similar to kanamycin
Hyg	Hygromycin-B-transferase; *hyg* gene from *Escherichia coli*	Select cells in hygromycin-B (10–300 mg ml^{-1}), an amino-cyclitol that inhibits protein synthesis
Pac/Puro	Puromycin-N-acetyl transferase; *pac* gene from *Streptomyces alboniger*	Select cells in puromycin (0.5–5 mg ml^{-1}), an antibiotic that inhibits protein synthesis
Zeo	Bleomycin-binding protein; a *zeo* gene (a.k.a. blaR) is located on the bacterial transposon *Tn5*	Select cells in bleomycin or commercially available Zeo (50–500 mg ml^{-1}), an antibiotic that binds DNA and blocks RNA synthesis

(E) Electroporation of ES cells with a *Gata6* cDNA plasmid for transient or stable transfection

(i) *Day 1*. On the day of electroporation, ensure that the ES cells are ~80% confluent in one 100-mm tissue culture plate. Aspirate the medium from the 100-mm culture plate and wash the cells with 5 ml of PBS.

(ii) Aspirate the PBS and trypsinize the cells with 1 ml of 0.05% (wt/vol) trypsin. Quench the trypsinized cells with 3 ml of MEF medium.

(iii) Transfer the ES cells to a 15-ml Falcon tube and pellet the cells at 200*g* for 4 min at room temperature. Aspirate the MEF medium and resuspend the ES cell pellet with 5 ml of PBS.

(iv) Pellet the cells at 200*g* for 4 min at room temperature. Aspirate the PBS and resuspend the ES cells in 0.7 ml of PBS.

(v) Gently mix the linearized *Gata6* cDNA sample with the resuspended cells and incubate the mixture for 5 min at room temperature.

(vi) Transfer the DNA-cell mixture to a 4-mm electroporation cuvette. Electroporate the mixture at 500 µF/230 V. The time constant should be between 6 and 8 ms.

(vii) Allow the ES cells to recover in the cuvette for 5 min at room temperature and then transfer the cells into 10 ml of ES cell medium; plate the entire volume onto a MEF-coated 100-mm plate. Mix gently by tilting the plate back and forth, and then incubate the plate overnight at 37 °C.

(viii) *Day 2*. Exchange the medium with 10 ml of standard ES cell medium and incubate the plate overnight in a 5% CO_2 humidified incubator at 37 °C.

(ix) *Days 3–9*. If the *Gata6*-cDNA plasmid has been engineered to also express a drug-resistance gene, then drug selection can be used to expand successfully transfected cells. Exchange the medium with 10 ml of standard XEN cell medium

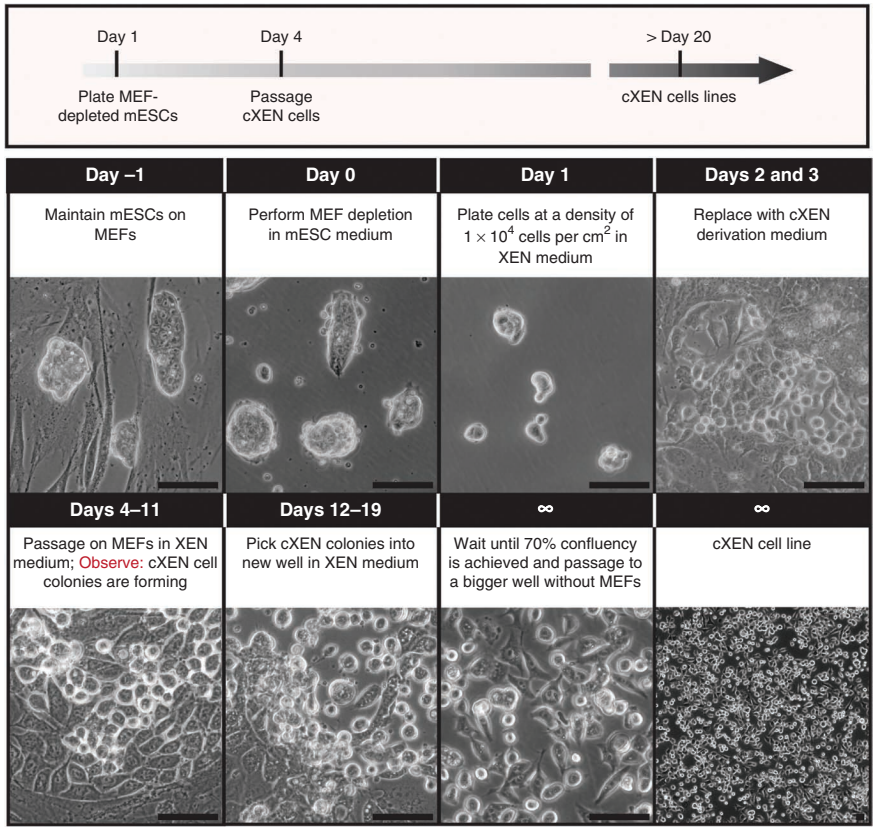


Figure 5 | Timeline for cXEN cell derivation from mouse ES cells. Protocol for the conversion of ES cells to cXEN cells using growth factors. Day −1: ES cells are maintained in ES medium on MEFs. Day 0: ES cells are passaged onto a pregelatinized plate for MEF depletion. Day 1: ES cells are enzymatically passaged with 0.05% (wt/vol) trypsin and plated at a density of 1×10^4 cells per cm^2 in standard XEN medium. Days 2 and 3: the medium is replaced with cXEN derivation medium daily (0.01–10 μM retinoic acid plus 10 ng ml^{-1} activin A). Days 4–11: cells are enzymatically passaged onto MEFs in XEN medium. The medium is replaced every day or every other day depending on confluency. XEN-like cells with stellate and refractile morphology will emerge within ~5 d. Days 12–19: XEN-like cells are picked manually and placed in a MEF-coated or pregelatinized plate in XEN medium. They are passaged two more times onto pregelatinized plates and the cXEN cell line is frozen. Scale bars, 100 μm.

supplemented with the appropriate drug selection (**Table 1**). Replace the medium with fresh XEN medium containing drug selection daily for ~10 d. XEN-like cells should appear within 5 d.

? TROUBLESHOOTING

(x) *Day 10–indefinite*. Enzymatically passage the *Gata6* cDNA–transfected cells with 0.05% (wt/vol) trypsin onto a pregelatinized 100-mm dish. Continue selection with drugs to establish a XEN cell line. Collect a fraction of the cells for genotyping and freeze the remaining XEN cells as described in Step 7.

(F) cXEN cell derivation from ES cells

(i) *Day 1*. Enzymatically passage MEF-depleted ES cells with 0.05% (wt/vol) trypsin. Inhibit the trypsin with XEN medium.

(ii) Centrifuge the ES cells at 200*g* for 4 min at room temperature and aspirate the supernatant. Be stringent in aspirating (without touching the pellet). This step removes ES medium, which is important as lingering LIF potentially inhibits cXEN derivation.

(iii) Resuspend the pellet in ~5 ml of standard XEN medium. Mix the tube well and immediately remove two 20-μl aliquots to count MEF-depleted ES cells on a hemocytometer. Take the average (n = 3) number of cells in one big square (16 little squares), multiply by 10^4 and multiply by the number of milliliters of medium, which will be equal to the total number of cells.

(iv) Remove the gelatin from the six-well cXEN derivation plate and add fresh standard XEN medium. Add the appropriate volume of trypsinized ES cells to start the cXEN derivation procedure for a final density of 1×10^4 cells per cm^2 (i.e., 9.6×10^4 cells per well of a six-well plate; **Fig. 5**). Incubate the plate overnight at 37 °C in standard XEN medium. The cell density is an important determinant of differentiation efficiency, and as different ES cell lines proliferate at different rates, we recommend determining the optimal density for each line.

(v) *Day 2*. Twenty-four hours after the initial plating, aspirate the standard XEN medium and replace it with 2 ml of cXEN derivation medium per well (**Fig. 5**). Incubate the plate overnight in a 5% CO_2 humidified incubator at 37 °C.

(vi) *Day 3*. Aspirate the cXEN derivation medium and replace it with fresh cXEN derivation medium. Incubate the plate overnight in a 5% CO_2 humidified incubator at 37 °C.

(vii) *Day 4*. Enzymatically passage the cells with 0.05% (wt/vol) trypsin. Quench the trypsin with 0.5 ml of standard XEN medium. Aspirate the MEF medium in a MEF-coated six-well plate and replace it with 2 ml of standard XEN medium per well. Transfer the entire contents of the differentiated cells and distribute the cells evenly into the well of the MEF-coated plate (1:1 dilution; **Fig. 5**).

(viii) *Days 5–11*. Aspirate the medium from the cXEN cells and replace it with fresh standard XEN medium every day or every other day depending on confluency. XEN-like cells with stellate and refractile morphology will emerge ~5 d after the first passage on MEFs; however, the well will also contain cells of different morphologies. Use standard XEN medium hereafter. During the first (and sometimes second) passage, the cXEN cells recover better when MEFs are present, although they are not a requirement. FGF2/FGF4 and heparin can also be added during the derivation, but they too are not a requirement as long as endogenous *Fgf4* is intact.

(ix) *Days 12–19*. Manually pick XEN-like cells using a 20-µl pipette under a microscope to facilitate the ES to cXEN cell derivation (**Fig. 5**). Depending on the density, place the isolated cells in a MEF-coated (low density of XEN-like cells) or pregelatinized plate (high density of XEN-like cells) in XEN medium and continue feeding with standard XEN medium. **? TROUBLESHOOTING**

(x) *Day 20–indefinite*. Once the cells reach 70–90% confluency, passage the cells enzymatically with 0.05% (wt/vol) trypsin and plate them onto a pregelatinized or MEF-coated 100-mm plate or several wells of a six-well plate. The majority of cells should morphologically resemble XEN cells. If there are any contaminating cells that are not XEN-like cells, aspirate, pick away or reciprocally pick the XEN-like cells to enrich for this population in a new plate or well. Passage two more times onto pregelatinized plate or well and freeze the cXEN cell line (Step 7).

DNA extraction and freezing cells ● TIMING 1–2 d

7| cXEN or XEN cells can be frozen and thawed using conventional stem cell freeze-thaw protocols in freezing medium. Passage as above and resuspend the trypsinized cell pellet in prechilled (4 °C) freezing medium as described in option A. If you wish to extract DNA from *Gata6* cDNA–targeted ES cells, there are a number of different extraction protocols that may be used. Option B describes the Promega SV Wizard purification system.

(A) Freezing cells

(i) Transfer the resuspended cells into prelabeled cryotube vials.

(ii) Quickly transfer the cryotube vials to −80 °C freezer overnight covered with Styrofoam.

■ PAUSE POINT For long-term storage, keep cells frozen in a liquid nitrogen dewar.

(B) Lysing cells for genomic DNA extraction

(i) Enzymatically passage cells as above and resuspend the cell pellet in the Promega Wizard SV lysis buffer according to the manufacturer's instructions, and then mix the cell lysate by pipetting.

■ PAUSE POINT The cell lysates can be frozen at −80 °C until needed.

(ii) Transfer each sample lysate from the cell culture plate to a separate Wizard SV minicolumn assembly.

(iii) Spin the assembly at 16.1 × 10^3*g* for 3 min at room temperature.

(iv) Remove the minicolumn from the assembly and discard the liquid in the collection tube. Replace the minicolumn into the collection tube.

(v) Add 650 µl of Wizard SV wash solution (with 95% (vol/vol) ethanol added) to each assembly. Centrifuge the tube at 16.1 × 10^3*g* for 1 min at room temperature. Discard the liquid from the collection tube.

(vi) Repeat Step 7B(v) for a total of four washes.

(vii) Discard the liquid from the collection tube and reassemble the minicolumn assembly. Centrifuge for 2 min at 16.1 × 10^3*g* at room temperature to dry the binding matrix.

(viii) Transfer the Wizard SV minicolumn to a new 1.5-ml tube. Add 50 µl of room-temperature, nuclease-free water. Incubate the minicolumn for 2 min at room temperature.

(ix) Centrifuge the minicolumn at 16.1 × 10^3*g* for 1 min at room temperature.

(x) Remove the minicolumn and store the purified DNA at −20 or −70 °C for several months or years, respectively.

? TROUBLESHOOTING

Troubleshooting advice can be found in **Table 2**.

TABLE 2 | Troubleshooting table.

Step	Problem	Possible reason	Solution
1–6	Contamination	Poor sterile technique	Ensure that all surfaces are thoroughly cleaned with 70% (vol/vol) ethanol before use and observe proper tissue culture technique
5A(iv)	Finding blastocysts	Lipid drops, blood cells and debris are impeding the view	Adjust magnification and mirror angle to identify blastocysts by their distinctive surrounding zona pellucida
6A(i), 6B(i)	Differentiation	Blastocyst misattachment	Blastocysts should be placed in the center of the well to avoid attachment to the sides of the well

(continued)

TABLE 2 | Troubleshooting table (continued).

Step	Problem	Possible reason	Solution
6A(iii), 6B(iii)	Small outgrowth	Insufficient time to develop	If no prominent outgrowth can be observed, wait an additional day before proceeding with the next steps
6	Low cell density	Low viability and/or contamination	It is imperative to closely observe the cells daily. XEN cell colonies can be observed as early as 10 d but might take up to 15 d to appear. Check for possible contamination
6B	TS cells outcompeting XEN cells	Addition of FGF and heparin; dissociation of the outgrowth	TS cells are more likely to outcompete the XEN cells later on. If this is observed, desist from adding FGF and heparin to the TS cell medium, and dissociate the outgrowth less rigorously. In some cases, it can be beneficial to entirely refrain from dissociating the outgrowth
6C, 6D, 6E	Insufficient expression of exogenous *Gata6*	*CMV* promoter may be silenced in ES cells	We recommend using a plasmid with an *EF1A* or *CAGG* promoter to ensure robust expression
6D	Low transfection efficiency	Serum in the medium may interfere with DNA-liposome complex formation	Opti-MEM I can be used during the overnight incubation instead of ES cell medium without penicillin or streptomycin to improve the transfection efficiency; however, this may compromise ES cell viability

● **TIMING**

Steps 1–5, preparation of MEF feeders: 15–30 min
Step 6, XEN derivation: 15–20 d
Step 7, DNA extraction and freezing cells: 1–2 d
Box 1, preparation of *Gata6* cDNA for transfection or electroporation: ~2 d

ANTICIPATED RESULTS

We have used the protocols described here to establish XEN cell lines from either mouse blastocysts or ES cells. XEN cell lines exhibit characteristic heterogeneous morphology with both highly refractile phase-bright and epithelial-like cells (**Figs. 2**, **4** and **5**). XEN cells can be distinguished from mouse ES cells as the latter form dome-shaped clusters of cells with characteristic high nuclear-to-cytoplasmic ratio[4,5], whereas the former grow as individual cells and the nucleus of each cell is not clearly distinguishable (compare **Fig. 2**). XEN cells are also distinct from TS cells that grow as epithelial colonies comprising cells with distinct nuclei and can differentiate into multinucleated trophoblast giant cells *in vitro*[2] (compare **Fig. 2**). It was previously demonstrated that XEN cells oscillate between these two morphologies[3]. All XEN cell lines retain the expression of key XEN-associated genes including the GATA transcription factor Gata4, the SOX factor Sox7 and Disabled homolog 2 Dab2 (**Supplementary Fig. 1**). Notably, these genes are not homogeneously expressed in all cells within a XEN cell line (**Supplementary Fig. 1**), and it remains unclear whether heterogeneity in morphology and gene expression reflects a fixed or oscillating heterogeneity representing distinct XEN cell types or substrates present in culture. Notably, mouse XEN cells usually lack the expression of ES cell–associated genes including octamer-binding transcription factor Oct4 (**Supplementary Fig. 1**) and Nanog. XEN cells self-renew indefinitely in culture in the absence of exogenous growth factors such as FGF2 (refs. 13,16). Moreover, XEN cells can be directed to differentiate into α-fetoprotein (Afp)-expressing visceral endoderm-like cells with the addition of BMP4 (refs. 18,19) and are committed to primitive endoderm-derived lineages in chimeric embryos[3,34]. Although not reported in the literature, in our own experience the karyotype of XEN cells can change over time and cell lines can acquire karyotypic anomalies with extended passage in culture (N.S. and A.-K.H., unpublished observations). It is therefore preferable to work with cells that are of as low a passage as possible. It should be noted that even clonal XEN cell lines exhibit some degree of variability, for example, in the ratio of cells exhibiting different morphologies and/or in the expression of molecular markers. It is currently not clear whether this inherent variability reflects the cell of origin within an embryo or ES cell culture, cell culture history or an unidentified stochastic or deterministic factor. Furthermore, although there is no 'gold standard' XEN cell line that is used across laboratories, as a point of reference,

especially when first establishing methods for XEN cell culture within a laboratory, it is advisable to obtain an established and characterized XEN cell line from another laboratory. In contrast to ES cells, no clear strain biases have been reported for the derivation of XEN cells from embryos or conversion from ES cells. This is likely because the protocols for XEN cell derivation, as provided here, are very efficient as compared with most non-inhibitor protocols for ES cell derivation. Notably, even though XEN cells are a relatively new stem cell type, which can be used in a number of applications, in the future, it will be important to compare the differentiation efficiency and molecular identity of XEN cells derived under distinct culture conditions to determine whether there may be biases in potential and/or differences in gene expression. In all, XEN cells are emerging as a useful stem cell model for understanding the convergence of signaling and transcriptional control during XEN cell–fate specification and differentiation.

Note: Supplementary information is available in the online version of the paper.

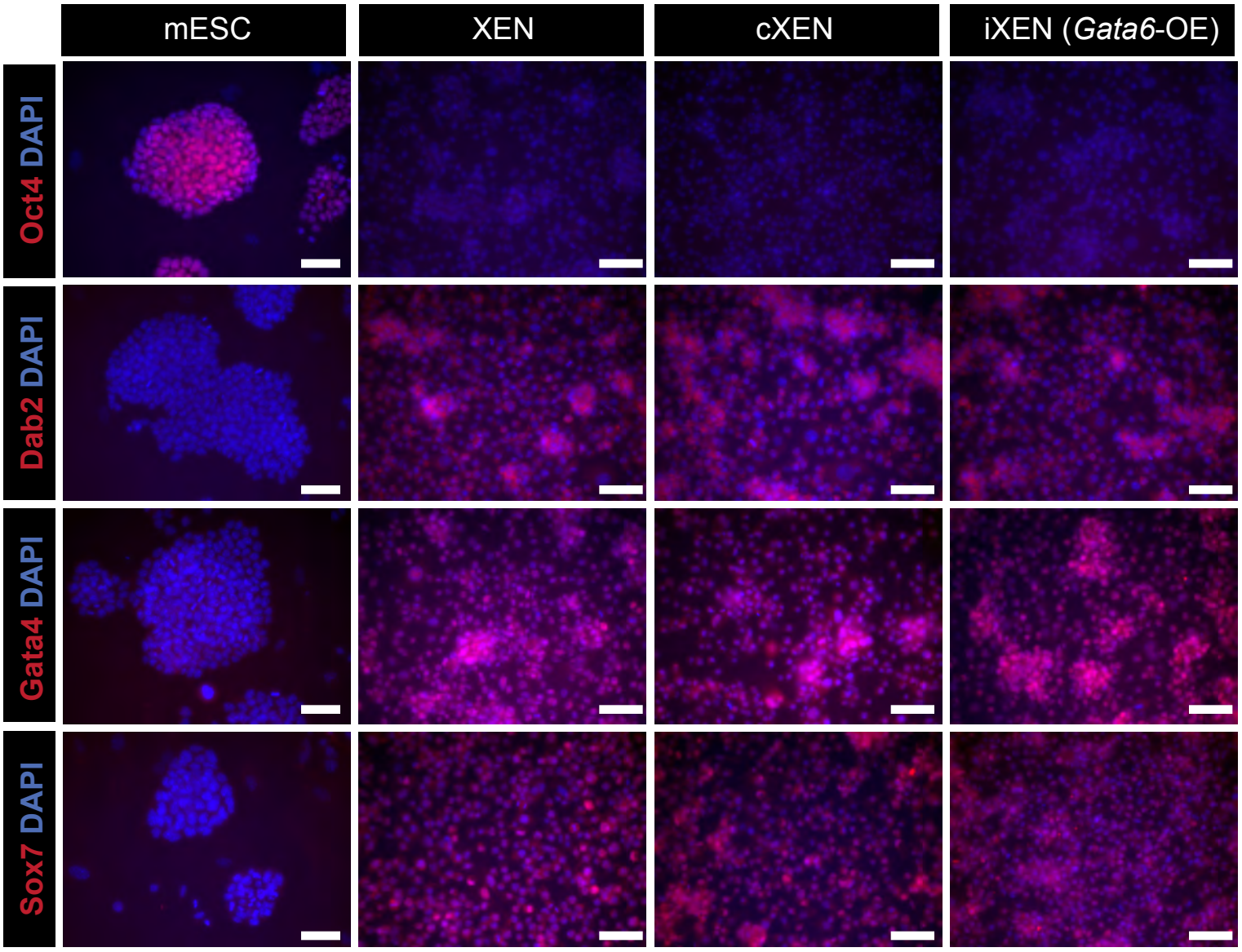
ACKNOWLEDGMENTS We thank A. Foley, M. Kang and S. Wamaitha for comments on this protocol. Work in our laboratories is supported by the Human Frontier Science Program, US National Institutes of Health (NIH) grants RO1-HD052115 and RO1-DK084391 to A.-K.H.; the New York State Department of Health NYSTEM IDEA grant C024318 to A.-K.H.; and the March of Dimes Foundation Research Grant 1-FY11-436 to K.K.N. L.T.Y.C. is supported by a Medical Research Council (MRC) capacity-building studentship. K.K.N. is supported by a Centre for Trophoblast Research Next Generation Fellowship.

AUTHOR CONTRIBUTIONS K.K.N. and A.-K.H. outlined the protocol. K.K.N., N.S. and A.-K.H. wrote the protocol manuscript with the help of L.T.Y.C.

COMPETING FINANCIAL INTERESTS The authors declare no competing financial interests.

1. Schrode, N. *et al.* Anatomy of a blastocyst: cell behaviors driving cell fate choice and morphogenesis in the early mouse embryo. *Genesis* Published online; http://dx.doi.org/10.1002/dvg.22368 (25 February 2013).
2. Tanaka, S., Kunath, T., Hadjantonakis, A.K., Nagy, A. & Rossant, J. Promotion of trophoblast stem cell proliferation by FGF4. *Science* **282**, 2072–2075 (1998).
3. Kunath, T. *et al.* Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. *Development* **132**, 1649–1661 (2005).
4. Evans, M.J. & Kaufman, M.H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154–156 (1981).
5. Martin, G.R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. USA* **78**, 7634–7638 (1981).
6. Artus, J. & Hadjantonakis, A.K. Troika of the mouse blastocyst: lineage segregation and stem cells. *Curr. Stem Cell Res. Ther.* **7**, 78–91 (2012).
7. Niwa, H. *et al.* Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* **123**, 917–929 (2005).
8. Fujikura, J. *et al.* Differentiation of embryonic stem cells is induced by GATA factors. *Genes Dev.* **16**, 784–789 (2002).
9. Capo-Chichi, C.D. *et al.* Perception of differentiation cues by GATA factors in primitive endoderm lineage determination of mouse embryonic stem cells. *Dev. Biol.* **286**, 574–586 (2005).
10. Soprano, D.R., Teets, B.W. & Soprano, K.J. Role of retinoic acid in the differentiation of embryonal carcinoma and embryonic stem cells. *Vitam. Horm.* **75**, 69–95 (2007).
11. Artus, J., Panthier, J.J. & Hadjantonakis, A.K. A role for PDGF signaling in expansion of the extra-embryonic endoderm lineage of the mouse blastocyst. *Development* **137**, 3361–3372 (2010).
12. Coucouvanis, E. & Martin, G.R. Signals for death and survival: a two-step mechanism for cavitation in the vertebrate embryo. *Cell* **83**, 279–287 (1995).
13. Cho, L.T. *et al.* Conversion from mouse embryonic to extra-embryonic endoderm stem cells reveals distinct differentiation capacities of pluripotent stem cell states. *Development* **139**, 2866–2877 (2012).
14. Niakan, K.K. *et al.* Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating extraembryonic gene expression and indirectly antagonizing self-renewal. *Genes Dev.* **24**, 312–326 (2010).
15. Lim, C.Y. *et al.* Sall4 regulates distinct transcription circuitries in different blastocyst-derived stem cell lineages. *Cell Stem Cell* **3**, 543–554 (2008).
16. Kang, M., Piliszek, A., Artus, J. & Hadjantonakis, A.K. FGF4 is required for lineage restriction and salt-and-pepper distribution of primitive endoderm factors but not their initial expression in the mouse. *Development* **140**, 267–279 (2013).
17. Kruithof-de, Julio M. *et al.* Regulation of extra-embryonic endoderm stem cell differentiation by Nodal and Cripto signaling. *Development* **138**, 3885–3895 (2011).
18. Paca, A. *et al.* BMP signaling induces visceral endoderm differentiation of XEN cells and parietal endoderm. *Dev. Biol.* **361**, 90–102 (2012).
19. Artus, J. *et al.* BMP4 signaling directs primitive endoderm-derived XEN cells to an extraembryonic visceral endoderm identity. *Dev. Biol.* **361**, 245–262 (2012).
20. Beddington, R.S. & Robertson, E.J. Anterior patterning in mouse. *Trends Genet.* **14**, 277–284 (1998).
21. Beddington, R.S. & Robertson, E.J. Axis development and early asymmetry in mammals. *Cell* **96**, 195–209 (1999).
22. Brown, K. *et al.* eXtraembryonic ENdoderm (XEN) stem cells produce factors that activate heart formation. *PloS ONE* **5**, e13446 (2010).
23. Liu, W., Brown, K., Legros, S. & Foley, A.C. Nodal mutant eXtraembryonic ENdoderm (XEN) stem cells upregulate markers for the anterior visceral endoderm and impact the timing of cardiac differentiation in mouse embryoid bodies. *Biol. Open* **1**, 208–219 (2012).
24. Holtzinger, A., Rosenfeld, G.E. & Evans, T. *Gata4* directs development of cardiac-inducing endoderm from ES cells. *Dev. Biol.* **337**, 63–73 (2010).
25. Senner, C.E. *et al.* DNA methylation profiles define stem cell identity and reveal a tight embryonic-extraembryonic lineage boundary. *Stem Cells* **30**, 2732–2745 (2012).
26. Kwon, G.S., Viotti, M. & Hadjantonakis, A.K. The endoderm of the mouse embryo arises by dynamic widespread intercalation of embryonic and extraembryonic lineages. *Dev. Cell.* **15**, 509–520 (2008).
27. Golding, M.C., Zhang, L. & Mann, M.R. Multiple epigenetic modifiers induce aggressive viral extinction in extraembryonic endoderm stem cells. *Cell Stem Cell* **6**, 457–467 (2010).
28. Nagy, A., Gertsenstein, M., Vintersten, K. & Behringer, R. *Manipulating the Mouse Embryo: a Laboratory Manual,* 3rd edn (Cold Spring Harbor Laboratory Press, 2003).
29. Lin, T.P. Microsurgery of inner cell mass of mouse blastocysts. *Nature* **222**, 480–481 (1969).
30. Himeno, E., Tanaka, S. & Kunath, T. Isolation and manipulation of mouse trophoblast stem cells. *Curr. Protoc. Stem Cell Biol.* **7**, 1E.4.1–1E.4.27 (2008).
31. Uy, G.D., Downs, K.M. & Gardner, R.L. Inhibition of trophoblast stem cell potential in chorionic ectoderm coincides with occlusion of the ectoplacental cavity in the mouse. *Development* **129**, 3913–3924 (2002).
32. Davis, R.L., Weintraub, H. & Lassar, A.B. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**, 987–1000 (1987).
33. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
34. Shimosato, D., Shiki, M. & Niwa, H. Extra-embryonic endoderm cells derived from ES cells induced by GATA factors acquire the character of XEN cells. *BMC Dev. Biol.* **7**, 80 (2007).
35. Rugg-Gunn, P.J. *et al.* Cell-surface proteomics identifies lineage-specific markers of embryo-derived stem cells. *Dev. Cell* **22**, 887–901 (2012).
36. Qu, X.B., Pan, J., Zhang, C. & Huang, S.Y. Sox17 facilitates the differentiation of mouse embryonic stem cells into primitive and definitive endoderm *in vitro*. *Dev., Growth Differ.* **50**, 585–593 (2008).

37. Shimoda, M. *et al.* Sox17 plays a substantial role in late-stage differentiation of the extraembryonic endoderm *in vitro*. *J. Cell Sci.* **120**, 3859–3869 (2007).
38. Aksoy, I. *et al.* Oct4 switches partnering from Sox2 to Sox17 to reinterpret the enhancer code and specify endoderm. *EMBO J.* **32**, 938–953 (2013).
39. Felgner, P.L. *et al.* Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proc. Natl. Acad. Sci. USA* **84**, 7413–7417 (1987).
40. De Smedt, S.C., Demeester, J. & Hennink, W.E. Cationic polymer–based gene delivery systems. *Pharm. Res.* **17**, 113–126 (2000).
41. Friend, D.S., Papahadjopoulos, D. & Debs, R.J. Endocytosis and intracellular processing accompanying transfection mediated by cationic liposomes. *Biochim. Biophys. Acta* **1278**, 41–50 (1996).
42. Gao, X. & Huang, L. Cationic liposome-mediated gene transfer. *Gene Ther.* **2**, 710–722 (1995).
43. Anderson, D.G., Lynn, D.M. & Langer, R. Semi-automated synthesis and screening of a large library of degradable cationic polymers for gene delivery. *Angew. Chem. Int. Ed. Engl.* **42**, 3153–3158 (2003).
44. Li, W. & Szoka, F.C. Jr. Lipid-based nanoparticles for nucleic acid delivery. *Pharm. Res.* **24**, 438–449 (2007).
45. Doetschman, T. *et al.* Targeted correction of a mutant *HPRT* gene in mouse embryonic stem cells. *Nature* **330**, 576–578 (1987).
46. Andreason, G.L. & Evans, G.A. Introduction and expression of DNA molecules in eukaryotic cells by electroporation. *Biotechniques* **6**, 650–660 (1988).
47. Shigekawa, K. & Dower, W.J. Electroporation of eukaryotes and prokaryotes: a general approach to the introduction of macromolecules into cells. *Biotechniques* **6**, 742–751 (1988).
48. Cherng, J.Y. *et al.* Effect of DNA topology on the transfection efficiency of poly((2-dimethylamino)ethyl methacrylate)-plasmid complexes. *J. Control. Release* **60**, 343–353 (1999).
49. McNally, M.A., Lebkowski, J.S., Okarma, T.B. & Lerch, L.B. Optimizing electroporation parameters for a variety of human hematopoietic cell lines. *Biotechniques* **6**, 882–886 (1988).
50. Stuchbury, G. & Munch, G. Optimizing the generation of stable neuronal cell lines via pre-transfection restriction enzyme digestion of plasmid DNA. *Cytotechnology* **62**, 189–194 (2010).
51. Chung, S. *et al.* Analysis of different promoter systems for efficient transgene expression in mouse embryonic stem cell lines. *Stem Cells* **20**, 139–145 (2002).
52. Koutsourakis, M., Langeveld, A., Patient, R., Beddington, R. & Grosveld, F. The transcription factor GATA6 is essential for early extraembryonic development. *Development* **126**, 723–732 (1999).
53. Morrisey, E.E., Ip, H.S., Lu, M.M. & Parmacek, M.S. GATA-6: a zinc finger transcription factor that is expressed in multiple cell lineages derived from lateral mesoderm. *Dev. Biol.* **177**, 309–322 (1996).
54. Iacovino, M. *et al.* Inducible cassette exchange: a rapid and efficient system enabling conditional gene expression in embryonic stem and primary cells. *Stem Cells* **29**, 1580–1588 (2011).
55. Ting, D.T., Kyba, M. & Daley, G.Q. Inducible transgene expression in mouse stem cells. *Methods Mol. Med.* **105**, 23–46 (2005).
56. Ding, S. *et al.* Efficient transposition of the *piggyBac* (*PB*) transposon in mammalian cells and mice. *Cell* **122**, 473–483 (2005).
57. Beddington, R.S. & Robertson, E.J. An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. *Development* **105**, 733–737 (1989).
58. Macfarlan, T.S. *et al.* Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57–63 (2012).
59. Strickland, S., Smith, K.K. & Marotti, K.R. Hormonal induction of differentiation in teratocarcinoma stem cells: generation of parietal endoderm by retinoic acid and dibutyryl cAMP. *Cell* **21**, 347–355 (1980).
60. Yasunaga, M. *et al.* Induction and monitoring of definitive and visceral endoderm differentiation of mouse ES cells. *Nat. Biotechnol.* **23**, 1542–1550 (2005).
61. Mesnard, D., Guzman-Ayala, M. & Constam, D.B. Nodal specifies embryonic visceral endoderm and sustains pluripotent cells in the epiblast before overt axial patterning. *Development* **133**, 2497–2505 (2006).



Molecular characteristics of XEN cell lines. Immunofluorescence analysis of mouse ES cells (mESC), embryo-derived XEN cells, cXEN cells and iXEN cells generated by *Gata6*-overexpression (OE) in mESCs after 6 days. Cell lines were analyzed for the expression of Oct4 (sc-5279, Santa Cruz Biotech, 1:500), Dab2 (SC-13982, Santa Cruz Biotech, 1:500), Gata4 (SC-9053, Santa Cruz Biotech, 1:500) or Sox7 (MAB2766, R&D, 1:500) (red) with DAPI (blue) merge using the method described in reference 14. Scale bars: 100 μm.

Nature Protocols: doi: 10.1038/nprot.2013.049

PUBLICATION III

DERIVATION AND CHARACTERIZATION OF MOUSE EMBRYONIC STEM CELLS FROM PERMISSIVE AND NONPERMISSIVE STRAINS

Czechanski A., Byers C., Greenstein I., Schrode N., Donahue L.R., Hadjantonakis A.K., Reinholdt L.G.

Nature Protocols
9(3):559-74
2014

Derivation and characterization of mouse embryonic stem cells from permissive and nonpermissive strains

Anne Czechanski[1], Candice Byers[1], Ian Greenstein[1], Nadine Schrode[2], Leah Rae Donahue[1], Anna-Katerina Hadjantonakis[2] & Laura G Reinholdt[1]

[1]Genetic Resource Science, The Jackson Laboratory, Bar Harbor, Maine, USA. [2]Developmental Biology Program, Sloan-Kettering Institute, New York, New York, USA. Correspondence should be addressed to L.G.R. (laura.reinholdt@jax.org).

Published online 6 February 2014; doi:10.1038/nprot.2014.030

Mouse embryonic stem cells (mESCs) are key tools for genetic engineering, development of stem cell–based therapies and basic research on pluripotency and early lineage commitment. However, successful derivation of germline-competent embryonic stem cell lines has, until recently, been limited to a small number of inbred mouse strains. Recently, there have been considerable advances in the field of embryonic stem cell biology, particularly in the area of pluripotency maintenance in the epiblast from which the mESCs are derived. Here we describe a protocol for efficient derivation of germline-competent mESCs from any mouse strain, including strains previously deemed nonpermissive. We provide a protocol that is generally applicable to most inbred strains, as well as a variant for nonpermissive strains. By using this protocol, mESCs can be derived in 3 weeks and fully characterized after an additional 12 weeks, at efficiencies as high as 90% and in any strain background.

INTRODUCTION

Embryonic stem cells (ESCs) are the *ex vivo* equivalent of the epiblast lineage of the blastocyst, and therefore they share the same developmental potential to differentiate into any one of the three primary germ layers: mesoderm, definitive endoderm and ectoderm (**Fig. 1**). This developmental pluripotency combined with a high capacity for self-renewal *in vitro* are the defining features of ESCs. Mouse embryonic stem cells (mESCs) are derived from preimplantation-stage embryos[1,2]. The progenitor cells that give rise to mESCs reside in the epiblast of the late blastocyst (~4 d *post coitum* (d.p.c.)) and express several pluripotency-associated factors, including OCT3/4 (officially known as POU5F1) and NANOG[3]. In addition to their capacity for self-renewal and stable pluripotency *in vitro*, mESCs have the defining capacity to populate the germline after microinjection into, or aggregation with, host embryos, making mESCs essential tools for genetic engineering[4]. The Nobel prize–winning discovery that genes could be genetically modified in mice by using mESCs was published over 30 years ago, and since then nearly 50,000 genetically modified alleles have been created by individual investigators around the world and by the International Knockout Mouse Consortium (IKMC, http://www.mousephenotype.org/martsearch_ikmc_project) which endeavors to create null and/or conditional null alleles for every gene in the mouse genome[5,6]. mESCs are also used for basic research on pluripotency and for the development of stem cell–based therapies as the starting material for directed differentiation of enriched, defined cell types *in vitro*.

Derivation of mESCs

Despite knowledge of the basic requirements for mESCs to maintain pluripotency, derivation of mESCs remained inefficient, and it was limited to just a few mouse strains for many years[7]. These so-called permissive strains included 129 substrains, as well as the most commonly used inbred mouse strain C57BL/6. Early protocols showed the requirement of leukemia inhibitory factor (LIF) to activate signal transducer and activator of transcription 3 (STAT3), bone morphogenetic protein (BMP) (or serum) and mitotically inactivated feeder layers, preferably mouse embryonic fibroblasts (MEFs), to prevent the differentiation of mESCs *in vitro*. However, derivation efficiency in permissive strains was at best 30%, as determined by the percentage of embryos giving rise to stable mESC lines[8,9]. Moreover, in nonpermissive mouse strains, such as CBA, NOD, DBA and others, derivation efficiency was either extraordinarily low or nonexistent[7]. Therefore, for mammalian geneticists who rely on specific inbred strain backgrounds for human disease modeling, genetically engineered alleles created in 129, C57BL/6 or hybrid ES cells required 10–20 backcross generations (up to 2 years) to create the desired genetic background. Moreover, for mammalian species other than the mouse, genetic engineering was simply not possible owing to an inability to derive legitimate ES cells despite considerable effort over many years. Recent advances in site-specific nuclease technologies (e.g., zinc-finger nucleases (ZFNs), transcription activator–like effector nucleases (TALENs) and clustered regularly interspersed short palindromic repeats (CRISPRs)/Cas)) are enabling direct, targeted deletion and targeted, sequential gene modifications via pronuclear injection of mouse embryos[10] and other species, including rat[11,12]. For genetic engineering, these technologies circumvent the need for ES cells. However, the applicability for multifaceted genomic modifications via homologous recombination with large inserts, across a variety of strains, has not yet been demonstrated.

Derivation of mESCs from nonpermissive strains

To overcome strain and species limitations to ES cell derivation, a variety of approaches have been used. For example, on the basis of the premise that the presence of primitive endoderm caused loss of pluripotency in mES cell progenitors within the inner cell mass (ICM), careful excision of the epiblast by biopsy or immunosurgery was shown to improve the derivation efficiency[3,13]. In addition, for many years, delayed implantation or diapause induction by ovariectomy or tamoxifen injection was also used to promote derivation efficiency, possibly via developmental stasis, during which epiblasts have an opportunity to expand. Finally, as knowledge of the genes required for early lineage specification

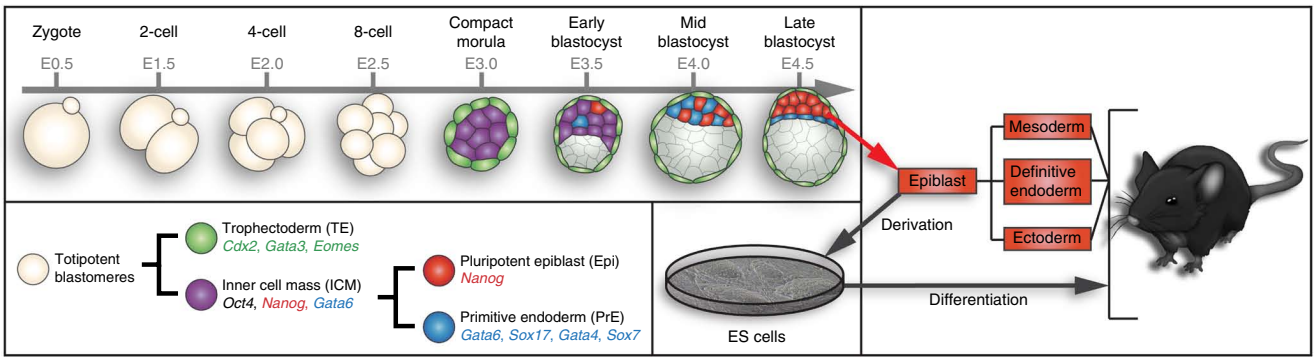


Figure 1 | Overview of preimplantation development in mice. The pre-epiblast lineage in the early embryo is defined by lineage-restricted expression of the *Oct3/4*, *Nanog* and *Gata6* genes. As early lineage specification proceeds, the pluripotent epiblast lineage is defined by *Nanog* expression. The epiblast lineage will give rise to all three definitive germ layers of the embryo-proper, namely all somatic cells and germ cells, and it is the population from which mESCs are derived. mESC cell lines retain the developmental potential of the epiblast lineage, and as such they can contribute to all three germ layers and the germ line of host blastocyst- or morula-stage embryos.

and pluripotency has grown, protocols for the efficient derivation of mES cells by promoting or inhibiting the expression of specific genetic pathways were developed. *Oct4 (Pou5f1)* is a transcription factor that is essential for the maintenance of pluripotency in cells of the ICM, the epiblast and in mES cell lines. Importantly, loss of *Oct4* was shown to be a feature of cultured embryos that failed to give rise to stable ES cell lines[14]. On the basis of this discovery, culture conditions that promote *Oct4* expression, namely inhibition of the MAP kinase pathway, were introduced. However, successful derivation of mES cells from the recalcitrant strain background, CBA, still required a combination of diapause induction, epiblast excision and inhibition of MEK kinase via PD98059 (ref. 14). In the context of these modifications to traditional ES cell derivation protocols, derivation efficiency in the CBA strain was ~25%, which is a significant advance for a nonpermissive strain[14].

The pluripotent ground state and overcoming barriers to mESC derivation

The discovery that self-renewal and pluripotency are intrinsic properties of mESCs was demonstrated by Smith and colleagues[14], who showed that the inhibition of MEK/ERK and glycogen synthase kinase-3 (GSK3) signaling (three inhibitors (3i): PD184352, PD173074/SU5402 and CHIR99021, respectively) were together sufficient, combined with activation of STAT3 by LIF (3i/LIF), to promote the pluripotent ground state of emergent ESCs from mice and from rats[15–17]. These laboratories went on to show that the inhibition of FGF receptor signaling is dispensable in the context of more potent inhibition of MEK signaling (2i: CHIR99021 to inhibit GSK3β and PD0325901 to inhibit MEK1/2)[16]. Both 3i/LIF and, subsequently, 2i/LIF culture conditions have since been successfully applied for efficient (50–70%) derivation of germline-competent mESCs from recalcitrant strains such as NOD, CBA and DBA[18–21]. Moreover, these culture conditions have been used to successfully derive germline-competent rat ESCs from rat embryos[16,17], which is an accomplishment that quickly led to the creation of the first rat gene knockout by homologous recombination in rESCs[22]. The successful derivation of ESCs from recalcitrant strains and from rat by using 2i/LIF culture conditions suggests that emergent ESCs from these strains or species are unable to maintain a pluripotent ground state under traditional ESC culture conditions (serum + LIF). In fact, it was later shown that, unlike emergent ESCs from permissive strain background (e.g., 129), emergent ESCs from nonpermissive strain backgrounds (e.g., NOD) are unstable and differentiate to a more advanced epiSC (postimplantation epiblast stem cell) state, which has been termed a primed pluripotent state, in the absence of exogenously provided inhibitors of ERK signaling[23].

Although the basis of strain and species recalcitrance to ESC derivation is not yet fully understood, these results suggest that inhibition of the pathways responsible for differentiation of ICM epiblast cells to postimplantation epiblast cells might be sufficient to overcome barriers to mESC derivation in all inbred strain backgrounds. This new model of the pluripotent, ground state of ESCs is an important advance in our understanding of early lineage commitment, and it has informed our mESC derivation protocol, which is highly efficient, regardless of strain background.

Experimental design

We previously published efficient derivation of germline-competent mESC lines from the recalcitrant strain DBA/2J (ref. 20). Crucial to the success of this protocol was the exclusion of serum during the outgrowth phase, combined with the inhibition of MEK/ERK (1i: PD98059) signaling during the outgrowth phase and during subsequent culture of emergent ES cell lines (3i: CHIR99021, PD173074 and PD0325901). As published data later showed the FGF receptor inhibitor PD173074 to be dispensable and the MEK inhibitor PD98059 to be redundant, in the context of the more potent MEK inhibitor PD0325901 (ref. 16), our current protocol uses the now standard 2i combination (CHIR99021 and PD0325901) to achieve the same exogenous inhibition with simpler media formulae.

Our protocol begins with the collection and culture of late blastocyst–stage embryos, which can be generated by natural mating or by *in vitro* fertilization. These embryos are then cultured in derivation medium to allow for ICM outgrowth. Unlike traditional ESC derivation medium, which contains serum, our derivation medium uses serum replacement in the form of an artificial serum replacement or, in the case of nonpermissive strain backgrounds, we use defined serum-free

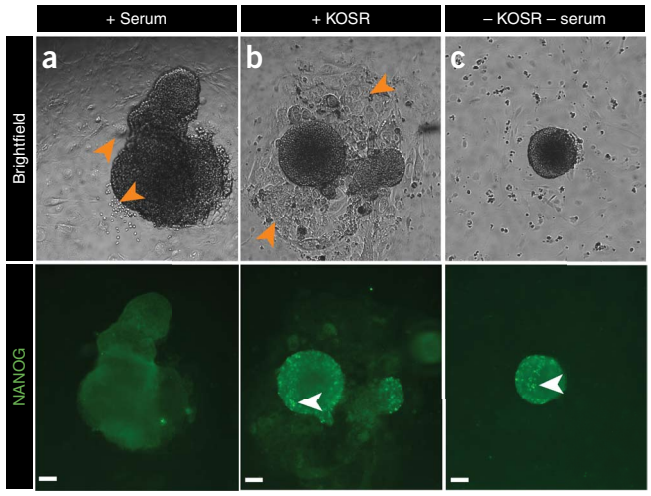


Figure 2 | NANOG immunolabeling in ICM outgrowths grown in the presence of 2i/LIF with or without serum or KOSR. (**a**,**b**) ICM outgrowths grown in traditional derivation medium containing serum have minimal NANOG expression (**a**) and exhibit differentiation (orange arrowheads) (**b**). (**b**) Outgrowths grown in the presence of KOSR contain NANOG-expressing cells (white arrowhead) but also exhibit differentiation (**b**, orange arrowheads). (**c**) Outgrowths grown in serum-free medium exhibit NANOG-expressing cells (white arrowhead) with minimal differentiation. Scale bars, 50 µm. All procedures involving mice were approved by The Jackson Laboratory's and Sloan-Kettering Institute's Institutional Animal Care and Use Committees and were performed in accordance with NIH guidelines for the care and use of animals in research.

medium[24]. The exclusion of serum from the derivation medium was previously shown to promote mESC derivation efficiency[9,25], and we have found that it promotes NANOG expression in ICM outgrowths (**Fig. 2**). NANOG, which is essential for the acquisition and maintenance of pluripotency, is a biomarker of mESC progenitor cells and is a more reliable readout of cell potency than OCT3/4 (refs. 14,26). Upon ICM disaggregation, incipient mESC lines are cultured either in traditional ESC medium + 2i/LIF (variant A; see Step 10A) or in defined serum-free medium + 2i/LIF (as described by Silva *et al.*[18], Ying and Smith[24] and this protocol, variant B (Step 10B)). We and others have found the latter culture conditions to be essential for robust derivation of ES cell lines from the recalcitrant NOD and its derivative strains (NSG and NRG) (**Table 1**)[19]. Although the original defined, serum-free + 2i/LIF conditions were created for feeder-free derivation and culture, feeder layers improve derivation efficiency, promote colony attachment (which is favorable for subcloning, a key manipulation in gene-targeting experiments) and may provide enhanced karyotypic stability. Therefore, we use feeder layers of mitotically inactivated embryonic fibroblasts. Importantly, this protocol provides all the steps necessary to derive and establish low-passage (passage (P)3) ES cell lines from any inbred strain background, as well as detailed instructions for quality assurance and characterization. Regardless of the culture conditions, pluripotency and euploidy degrade with increasing passage number in mESCs (P20 and higher). Therefore, newly established mESC lines should be maintained with an eye toward maximizing and preserving low-passage stocks.

TABLE 1 | Derivation efficiencies in selected strain backgrounds using traditional derivation conditions versus this protocol.

Strain background	Reported efficiency in standard ES cell derivation conditions (serum, LIF[31])	Ref.	Efficiency achieved using this protocol (2i, LIF)
DBA	1% (DBA/1lacJ)	32	60% (ref. 20) (DBA/2J)
BALB/c	2.4%	7	90%[a] (BALB/cJ)
C3H	3%	33	90%[a] (C3H/HeJ)
NOD	1.5%	34	75%[a,b] (NOD/ShiLtJ)
NSG	Unreported[c]	34	65% [a,b] (NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ)
A/J	Unreported	NA	90%[a]
PWD/J	Unreported	NA	60%[a]
C57BL6	17–80%	3,7,35	100%[a] (C57BL/6J)
CBA	0%	14	Not attempted

Substrains are provided if known/reported. Efficiency is the percentage of disaggregated embryos that gave rise to stable mESC lines, as defined by morphology and survival through passaging. In the literature, efficiency is defined by the number of stable mESCs generated from the number of embryos collected. With this protocol, the vast majority (90%) of embryos give rise to an outgrowth, but not every outgrowth is disaggregated due to budget constraints. Therefore an efficiency calculation based on the number of embryos collected would be a gross underestimate.

[a]Our unpublished data. [b]Required serum-free culture conditions (as described by Ying and Smith[24] and this protocol, variant B). [c]Presumed similar permissiveness to NOD, as NSG is a derivative strain.

MATERIALS

REAGENTS

Derivation of mES cells

- Pregnant mouse, 3.5 d.p.c. ! CAUTION Experiments involving rodents must conform to all relevant governmental and institutional (i.e., institutional animal care and use committee (IACUC)) standard operating procedures and regulations.
- MEFs (see Reagent Setup)
- Mitomycin C (Sigma, cat. no. M0503; prepare 1 mg/ml stock in PBS and store it at 4 °C for up to 6 months)
- 2-Mercaptoethanol, 55 mM (Invitrogen, cat. no. 21985-023) ! CAUTION This reagent is a biohazard; adequate safety instructions should be taken when handling it.
- B27 supplement (Invitrogen, cat. no. 17504-044)
- CHIR99021 (Stemgent, cat. no. 04-0004; see Reagent Setup)
- Defined trypsin inhibitor (Invitrogen, cat. no. R-007-100)
- DMEM, high glucose, no glutamine, no sodium pyruvate (Invitrogen, cat. no. 11960069)
- DMEM/F-12 medium (Invitrogen, cat. no. 11320-033)
- DMSO (Sigma-Aldrich, cat. no. D2650) ! CAUTION This reagent is a biohazard; adequate safety instructions should be taken when handling it.
- Ethanol, 70% (vol/vol) (Sigma-Aldrich, cat. no. E7023) ! CAUTION This reagent is highly flammable, and it is a skin and eye irritant.
- FBS, ES cell grade (Lonza, cat. no. 14-501F) ▲ CRITICAL FBS should be lot-tested for maximum ES cell viability and growth. Filter it through a 0.45-µm filter, and store it in 25-ml single-use aliquots at −20 °C for up to 1 year.
- Gelatin, 0.1% (wt/vol) (STEMCELL Technologies, cat. no. 07903)
- GlutaMAX (Invitrogen, cat. no. 35050061)
- KnockOut serum replacement (KOSR; Invitrogen, cat. no. 10828-028); store 25-ml single-use aliquots of the reagent at −20 °C for up to 1 year
- EmbryoMax potassium simplex optimized medium (KSOM) embryo culture (1×) medium, powder (Millipore, cat. no. MR-020-P)
- LIF, 10^7 units (Millipore, cat. no. ESG1107)
- EmbryoMax M2 medium (1×), powder (Millipore, cat. no. MR-015-D)
- MEM nonessential amino acid solution (NEAA; Invitrogen, cat. no. 11140050)
- N-2 supplement (Invitrogen, cat. no. 17502-048)
- Neurobasal medium (Invitrogen, cat. no. 21103-049)
- PBS without calcium and magnesium (Invitrogen, cat. no. 20012027)
- PD0325901 (Stemgent, cat. no. 04-0006; see Reagent Setup)
- Penicillin-streptomycin, 100× (Invitrogen, cat. no. 15140122)
- Sodium pyruvate, 100 mM (Invitrogen, cat. no. 11360070)
- Trypsin-EDTA, 0.05% (wt/vol) (Invitrogen, cat. no. 25300054)

Karyotyping

- Colcemid, 10 µg/ml (Invitrogen, cat. no. 15212-012)
- Glacial acetic acid (Sigma-Aldrich, cat. no. A6283) ! CAUTION This reagent is a biohazard and a severe irritant; adequate safety instructions should be taken and all work with this chemical should be conducted in a fume hood.
- KCl, 0.56% (vol/vol) (Sigma-Aldrich, cat. no. P9541)
- Methanol (Sigma-Aldrich, cat. no. 494437) ! CAUTION This reagent is a biohazard; adequate safety instructions should be taken when handling it.
- Vectashield hard-set mounting medium with DAPI (Vector Laboratories, cat. no. H-1500)

Immunolabeling

- DAPI (Invitrogen, cat. no. D1306; see Reagent Setup) or Hoechst (Invitrogen, cat. no. 33342)
- FBS (Lonza, cat. no. 14-501F)
- Goat anti-mouse IgG Alexa Fluor 488 (Invitrogen, cat. no. A11029)
- Goat anti-rabbit IgG Alexa Fluor 488 (Invitrogen, cat. no. A11034)
- Goat anti-mouse IgM Alexa Fluor 488 (Invitrogen, cat. no. A21042) ▲ CRITICAL Secondary antibodies conjugated to fluorescent dyes are light-sensitive. A broad array of Alexa fluorophores (Invitrogen) are available depending on the desired excitation/emission spectra. Moreover, alternatives to Alexa Fluor dyes are widely available, including FITC, Texas Red, rhodamine and others.
- Mouse IgG anti-OCT3/4 antibody (Santa Cruz Biologicals, cat. no. sc-5279; see Reagent Setup)
- Mouse IgM anti-SSEA-1 antibody (Santa Cruz Biologicals, cat. no. sc-21702; see Reagent Setup)
- Rabbit IgG anti-Nanog antibody (Abcam, cat. no. ab21603)
- PBS without calcium and magnesium (Invitrogen, cat. no. 20012027)
- Paraformaldehyde, 4% (wt/vol) (PFA; Electron Microscopy Sciences, cat. no. 157-4) ! CAUTION This reagent is a biohazard; adequate safety instructions should be taken when handling it.
- Triton X-100 (Promega, cat. no. H5142)

SNP panel or genotyping

- DNeasy blood and tissue kit (Qiagen, cat. no. 69506)
- Tris, 1 M, pH 8.0 (Invitrogen, cat. no. AM9856)

Pathogen testing

- Sabouraud dextrose agar deep-fill plates (BD, cat. no. 221180)
- Tryptose phosphate broth (Sigma-Aldrich, cat. no. T8159)

EQUIPMENT

- Filter unit such as Stericup-GV, 0.22 µm (Millipore, cat. no. SCGVU05RE)
- Biosafety cabinet (e.g., Nuaire, model NU-425-400 or equivalent)
- Calibrated, glass micropipettes, 50 µl (Drummond Scientific, cat. no. 2-000-050). Note that the aspirator apparatus required for the mouth-controlled embryo transfer pipette is supplied with the micropipettes (see Equipment Setup). Alternatively, 9-inch glass Pasteur pipettes, pulled over a flame, can be assembled with a plastic mouthpiece and P1000 pipette tip, as shown in **Figure 3**.
- Cell counter, automatic (e.g., Nexcelom Cellometer Auto T4) or manual hemocytometer (e.g., Reichert Bright-Line, Fisher Scientific, cat. no. 02-671-5)
- Conical tubes, 15 ml (USA Scientific, cat. no. 1475-0511)
- CoolCell LX (Biocision, cat. no. BCS-405)
- Cover glass, 24 × 60 mm (Fisher Scientific, cat. no. 12-548-5P)
- Cryovials, 2 ml (USA Scientific, cat. no. 5612-2263)
- Micro-dissecting scissors (e.g., Roboz, cat. no. RS-5610)
- Forceps, two pairs (e.g., Roboz, cat. no. RS-4984)
- Microscope, inverted (e.g., Nikon, model TS100 or equivalent)
- Microscope, stereo (e.g., Leica, model MZ12.5 or equivalent)
- Microscope slides (Fisher Scientific, cat. no. 12-550-15)
- Pasteur pipettes, borosilicate, 9 inches, unplugged (Fisher Scientific, cat. no. 13-678-20C)
- Serological pipettes, 5 ml (USA Scientific, cat. no. 1075-0810)
- Serological pipettes, 10 ml (USA Scientific, cat. no. 1071-0810)
- Syringes, 5 ml (BD, cat. no. 309646) with 27-G needle (BD, cat. no. 305109)
- Tissue culture–treated plate, four wells (Corning Life Sciences, cat. no. 353654)
- Tissue culture–treated plate, six wells (USA Scientific, cat. no. 5665-7160)
- Tissue culture–treated dish, 35 mm (USA Scientific, cat. no. CC7682-3340)
- Tissue culture–treated dish, 60 mm (USA Scientific, cat. no. CC7682-3359)
- Tissue culture–treated dish, 100 mm (USA Scientific, cat. no. 5666-4160)
- Vacuum pump (Fisher Scientific, cat. no. 13-878-40 or equivalent)
- Water bath (Thermo Scientific, model 2864 or equivalent)
- Water-jacketed CO_2 incubator, 37 °C, 5% CO_2 with HEPA filtration and 95% humidity (Thermo Scientific, model 3120 or equivalent)

REAGENT SETUP

Cryopreservation medium Cryopreservation medium is 80% culture medium (FBS or serum-free ES cell medium, depending on the medium used for the mESC line), 10% (vol/vol) FBS and 10% (vol/vol) DMSO. The medium can be stored at 4 °C for up to 1 week.

EmbryoMax KSOM and M2 embryo culture media The EmbryoMax KSOM and M2 embryo culture media are freshly made according to the manufacturer's instructions on the day of embryo collection.

▲ CRITICAL The reagents must be freshly prepared.

KOSR ES medium KOSR ES medium is DMEM-high glucose supplemented with 15% (vol/vol) KOSR, 1× penicillin-streptomycin, 2 mM GlutaMAX, 1 mM sodium pyruvate, 0.1 mM MEM NEAA, 0.1 mM 2-mercaptoethanol, 10^3 IU of LIF, 1 µM PD0325901 and 3 µM CHIR99021. Sterilize the medium through a 0.22-µm filter and store it at 4 °C for up to 1 week, protected from light.

FBS ES medium FBS ES medium is DMEM-high glucose supplemented with 15% (vol/vol) FBS, 1× penicillin-streptomycin, 2 mM GlutaMAX, 1 mM

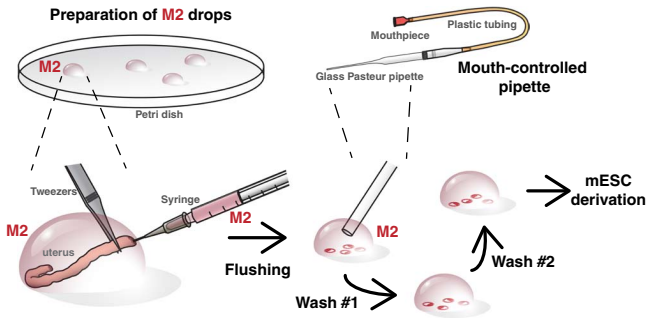


Figure 3 | Overview of blastocyst-stage embryo collection. Embryos are flushed from the uteri at 3.5–3.75 d.p.c. by using M2 medium. Embryos are pooled and washed through a series of M2 drops by using a mouth-controlled pipette before plating for ES cell derivation. It is important to get rid of tissue debris before embryo plating. It also provides an opportunity to assess the stage and quality of embryos recovered for stem cell derivation. All procedures involving mice were approved by The Jackson Laboratory's and Sloan-Kettering Institute's Institutional Animal Care and Use Committees and were performed in accordance with NIH guidelines for the care and use of animals in research.

sodium pyruvate, 0.1 mM MEM NEAA, 0.1 mM 2-mercaptoethanol, 10^3 IU of LIF, 1 µM PD0325901 and 3 µM CHIR99021. Sterilize the medium through a 0.22-µm filter and store it at 4 °C for up to 1 week, protected from light.

Serum-free ES medium Serum-free ES medium is a 1:1 mixture of DMEM-F12/N2 (DMEM-F12 supplemented with N-2) and Neurobasal/B27 (Neurobasal supplemented with B27) with 1× (vol/vol) penicillin-streptomycin, 1 mM GlutaMAX, 0.5 mM sodium pyruvate, 0.1 mM MEM NEAA, 0.1 mM 2-mercaptoethanol, 10^3 IU of LIF, 1 µM PD0325901 and 3 µM CHIR99021. Sterilize the medium through a 0.22-µm filter and store it at 4 °C for up to 1 week, protected from light.

MEF medium MEF medium is DMEM-high glucose supplemented with 10% (vol/vol) FBS, 1% (vol/vol) penicillin-streptomycin and 2 mM GlutaMAX. Sterilize the medium through a 0.22-µm filter and store it at 4 °C for up to 1 week, protected from light.

CHIR99021 Prepare a stock solution of 30 mM by resuspending 2 mg in 140 µl of DMSO. Divide the solution into single-use aliquots and store them at −20 °C for up to 6 months.

PD0325901 Prepare a stock solution of 10 mM by resuspending 2 mg in 414 µl of DMSO. Divide the solution into single-use aliquots and store them at −20 °C for up to 6 months.

Methanol:glacial acetic acid, 3:1 In a fume hood, mix methanol and glacial acetic acid at a ratio of 3:1. Use 10 ml of fresh fixative per cell line. Keep the reagent on ice until it is ready for use. ▲ **CRITICAL** The reagent must be freshly prepared.

Blocking buffer Combine PBS with 5% (vol/vol) FBS and 0.1% (vol/vol) Triton X-100. Store the buffer at 4 °C for up to 1 week.

DAPI stock solution Prepare a stock solution by dissolving 10 mg of DAPI in 2 ml of sterile molecular biology–grade water. Store this solution at 4 °C for up to 1 year, protected from light. Use it at a dilution of 1:2,500.

Primary antibody solution Dilute the primary antibodies to 1:250 in PBS with 1% (wt/vol) FBS. The working concentrations for antibodies can vary depending on the supplier and/or the lot, and therefore they may need to be determined empirically. We usually use dilutions of 1:250 for the anti-OCT3/4, 1:250 for the anti-SSEA-1 and 1:250 for the anti-NANOG antibodies listed in the Reagents section. Store the diluted antibodies at 4 °C for up to 1 month. For longer-term storage, divide the diluted antibody into single-use aliquots and store them at −20 °C.

Secondary antibody solution Dilute the secondary antibodies to 1:1,000 in PBS with 1% (vol/vol) FBS. The working concentrations for antibodies can vary depending on the supplier and/or the lot, and therefore they may need to be determined empirically. We usually use dilutions of 1:1,000 for the anti-mouse IgG, 1:1,000 for the anti-rabbit IgG and 1:1,000 for the anti-mouse IgM antibodies listed in reagents. Store the diluted antibodies at 4 °C for up to 1 month. For longer-term storage, divide the diluted antibody into single-use aliquots and store them at −20 °C.

Mitomycin C Prepare a 1 mg/ml stock in PBS and store it at 4 °C for up to 6 months.

MEFs Derive MEFs from embryonic day (E) 12.5–14.5 embryos that have been decapitated and eviscerated. Collect embryos from a pregnant mouse 12.5–14.5 d.p.c. Wash the decapitated and eviscerated embryos in PBS and place them in a clean 100-mm culture dish containing 3–5 ml of 0.05% (wt/vol) trypsin-EDTA. Coarsely mince the embryos with sterile forceps and a razor blade. Draw the resulting fragments into a 10-ml syringe and pass them through a 16-G needle two or three times to reduce the tissue to a slurry. Add the slurry to a 50-ml conical tube containing an equal volume of MEF medium and allow any remaining large fragments to settle. Transfer the contents (avoiding settled material) to a new 50-ml conical tube, and centrifuge at 150*g* for 5 min. Resuspend the pellet in MEF medium, count the cells and plate them onto a 100-mm, gelatinized dish (expect a yield of ~1×10^6 cells per embryo and plate at ~1×10^5 cells per cm^2, which is approximately one 100-mm dish per embryo). When the cells are confluent (~2 d), passage 1:3 by using 0.05% (wt/vol) trypsin-EDTA. After three passages, mitotically inactivate the MEFs by using mitomycin C or γ-irradiation. ! **CAUTION** Experiments involving rodents must conform to all relevant governmental and institutional (IACUC) standard operating procedures and regulations.

Mitotic inactivation of MEFs using mitomycin C Grow MEFs until they are confluent, and then replace the medium with fresh medium containing 10 µg/ml mitomycin C. Return the plates to the incubator for 2–3 h. Wash the MEFs extensively (two or three times) with PBS, collect them using 0.05% (wt/vol) trypsin-EDTA and then count the cells. Resuspend the cells in cryopreservation medium at a concentration of ~1×10^7 cells per ml, quickly divide the suspension into 1-ml aliquots in cryovials and temporarily store the aliquots at −80 °C in a CoolCell LX or an equivalent container, which will allow for an optimum freeze rate of −1 °C per minute. Transfer the frozen cells to liquid nitrogen storage within 1 week. MEFs can be stored in liquid nitrogen indefinitely. To prepare feeder layers, thaw one vial and resuspend it in ~100 ml of MEF medium and plate it onto gelatinized dishes at a density of 1×10^5 cells per cm^2.

Mitotic inactivation of MEFs using γ-irradiation Expose confluent MEFs to γ-rays and achieve an exposure of 5,000–10,000 rads. Collect and resuspend the MEFs in cryopreservation medium, as described for mitomycin-treated MEFs.

EQUIPMENT SETUP

Mouth pipette Heat the center of a glass, calibrated micropipette over a gas microburner, stretch the pipette by hand and then break the pipette in the center. Assess the broken end under a stereomicroscope. Keep only those pipettes that are thin, even and that have a diameter 1.5-2× that of a standard mouse blastocyst. Insert the pulled pipette into an aspirator apparatus (**Fig. 3**). If mouth pipetting is not possible or desirable, micropipette aids are commercially available (e.g., Drummond, Captrol III, cat. no. 3-000-752), and they can be used in place of the mouth-controlled aspirator apparatus shown in **Figure 3**.

Gelatinized plates Gelatinize all tissue culture plates and dishes to promote cell attachment. Pipet a sufficient amount of 0.1% (wt/vol) gelatin to cover the surface of a tissue culture–treated plate. Incubate the plate for at least 15 min at room temperature (20–25 °C). Aspirate the gelatin and dry it briefly.

MEF feeder cell plates Thaw and resuspend MEFs in MEF medium at a seeding density of ~1×10^5 cells per ml, and use the appropriate volume for ~1×10^5 cells per cm^2 according to the growth area of the required plate or dishes. Prepare MEF feeder plates at least 1 d before use. MEF feeders may be used up to 1 week after plating. Before use, rinse the MEF feeder plates with PBS. ▲ **CRITICAL** MEF viability is highly variable in serum-free culture medium, and generally low-passage (P2–P3) MEFs are more tolerant. Test higher-passage (<P3) MEFs empirically for long-term survival (5–10 d) in serum-free medium.

Surface sterilization In accordance with aseptic technique, sterilize all surfaces, bottles, racks, pipette aids and the other equipment by spraying or swabbing with 70% (vol/vol) ethanol before all procedures.

PROCEDURE

Preparation of MEF feeder plates ● TIMING 1 d

1| One day before collecting embryos, thaw and plate inactivated MEFs onto four-well plates. The number of wells needed is dependent on the number of pregnant females used. The number of embryos received per female is highly variable, but it generally averages between six and eight. Blastocysts are plated in individual wells. Thus, for one female, 6–8 wells should be needed.

2| On the day of embryo collection, remove the MEF medium and rinse the four-well feeder plates with PBS. Replace the medium with either KOSR ES medium (variant A; see Step 10A, permissive strains) or serum-free ES medium (variant B; see Step 10B, nonpermissive strains). Return the plates to the incubator until the embryos are gathered. ▲ CRITICAL For permissive strains (129, B6) and for most of the nonpermissive strains listed in **Table 1,** use KOSR ES medium for derivation (variant A; Step 10A). However, if you are deriving ES cells from a known nonpermissive strain such as NOD or a NOD-derivative strain, use serum-free ES cell medium (variant B; Step 10B). If the strain has not been previously characterized as nonpermissive or permissive, follow Step 10A initially and then repeat with Step 10B if the first attempt is unsatisfactory.

Collecting E3.5 blastocysts ● TIMING 15 min per pregnant female

3| Remove the uterus of a 3.5-d.p.c. pregnant mouse. Trim away the excess fat.

4| With a 5-ml syringe, collect 4 ml of M2 medium and attach a 27-G needle. While securing the uterine horn with forceps, insert the needle and flush the uterine horn with 2 ml of M2 medium into a Petri dish. Repeat this process with the second uterine horn (**Fig. 3**).

5| By using a mouth pipette or alternative micropipetting aid, collect the flushed embryos and wash through several drops of M2 before placing into a final drop of M2 medium (**Fig. 3**).

6| If embryos are not fully expanded, blastocysts with clearly discernible blastocoels (**Fig. 1**, E 4.0), culture the embryos overnight in a 35-mm dish of KSOM at 37 °C, 5% CO_2. If blastocysts are fully expanded at the time of collection, continue to Step 5.

▲ CRITICAL STEP If the blastocysts are not fully expanded when proceeding to Step 5, the percentage of embryos that hatch and attach to the feeders layers after plating may be very low (>50%). Embryos may also be collected at 3.75 d.p.c. to ensure a higher percentage of expanded blastocysts.

▲ CRITICAL STEP Traditional mESC derivation protocols recommend the removal of the zona pellucida before plating. Although this step is unnecessary (viable embryos will hatch naturally under appropriate culture conditions) and can lead to embryo loss, there are some scenarios that may warrant zona removal. For a detailed protocol on zona removal by using acid Tyrode's solution, see Nagy *et al.*[27].

Plating and early culture ● TIMING 7–9 d

7| By using a mouth pipette (or alternative micropipette aid), plate blastocysts in either KOSR ES medium (see Step 10A) or serum-free ES medium (see Step 10B), one blastocyst per well of a four-well MEF feeder plate (**Fig. 4**, day 1) containing the appropriate medium for derivation.

8| Incubate the plates at 37 °C, 5% CO_2. Do not disturb the plates for 48 h to allow for the blastocysts to hatch and attach to the feeder layer. The majority (80–90%) of blastocysts should hatch and attach to the feeder layer. On day 3 after plating, replace half the medium with fresh medium, by using either KOSR ES medium or serum-free ES medium depending on the desired experimental variant (Step 10A or Step 10B). Continue to feed the cells in this manner every other day until disaggregation. After the initial 48 h of incubation, monitor the growth of the outgrowth under a microscope daily (**Fig. 4**, days 2 and 3). If outgrowths are prominent 1 week after plating, proceed to the next step. If not, wait for 1 or 2 more days before proceeding to Step 9. Do not allow the outgrowths to grow so large that the center becomes dark. (Note that dark areas in outgrowths shown in **Figure 2** are also depleted of NANOG).

? TROUBLESHOOTING

9| Disaggregate outgrowths mechanically (option A) or enzymatically (option B).

(A) Mechanical disaggregation

(i) By using a mouth pipette, transfer the outgrowth onto a new feeder-covered four-well plate in either FBS ES medium (for variant A; Step 10A) or serum-free ES medium (for Step 10B).

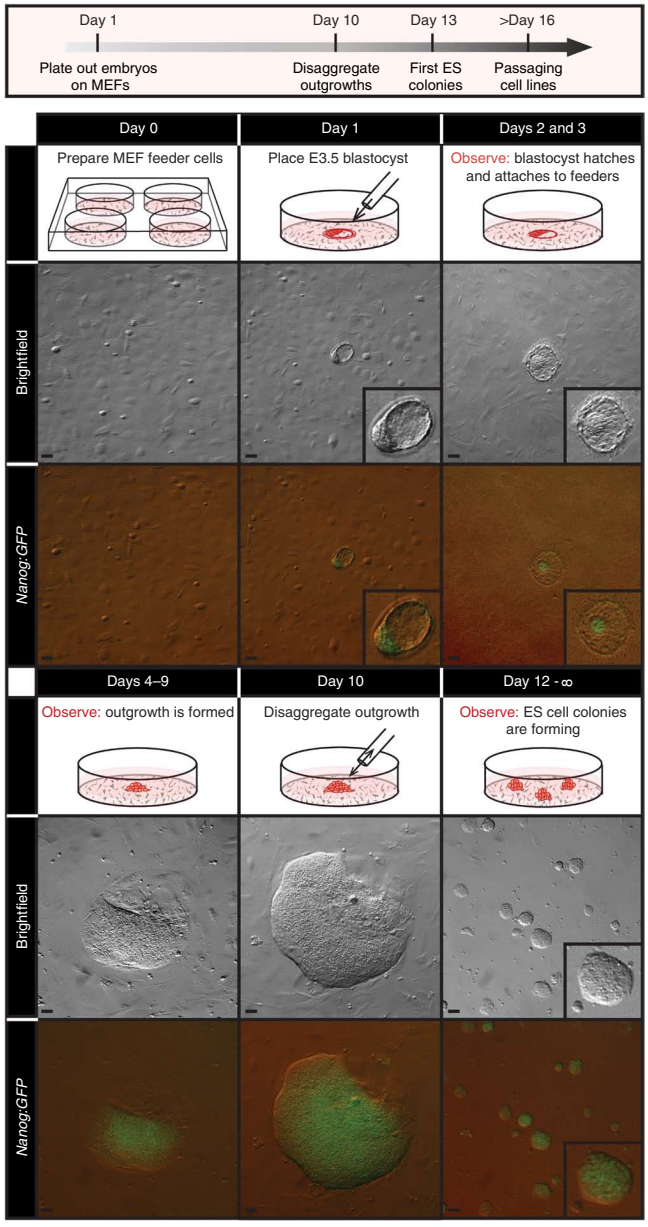


Figure 4 | Overview of ES cell derivation process. Blastocyst-stage embryos are plated on mitotically inactivated MEFs and allowed to hatch from the zona pellucida. Shown here are transgenic embryos carrying a GFP reporter cassette driven by the *Nanog* promoter (*Nanog:GFP*). By days 2–3, hatched embryos attach to the feeder layer, and over the next several days (4–9) the *Nanog*-positive ICM forms a large *Nanog*-positive outgrowth from which mESCs are extracted by mechanical disaggregation and/or trypsinization. Scale bars, 50 µm. All procedures involving mice were approved by The Jackson Laboratory's and Sloan-Kettering Institute's Institutional Animal Care and Use Committees and were performed in accordance with NIH guidelines for the care and use of animals in research.

▲ **CRITICAL STEP** At this step, if you are following variant A (see Step 10A), outgrowths are now moved from KOSR ES medium to FBS ES cell medium. If you are following variant B (see Step 10B), continue culture in serum-free ES medium.

(ii) By using tweezers that have been sterilized with 70% (vol/vol) ethanol, break up the outgrowth into small fragments.

(iii) Pipet the fragments several times with the pulled pipette to achieve further disaggregation.

(iv) Replace the medium with the appropriate medium, as described in Step 9A(i) above, and observe the wells daily. Once the colonies have emerged (2–3 d later), enzymatically passage 1:1 as described in option B. The goal is not to reach confluency, but rather to dissociate cells before large colonies become dark. The incipient ES cell line is now at P1 (passage 1). Continue to passage as described in option B, and increase the passage number by 1 with each subsequent exposure to trypsin and carefully track the passage number by labeling the plates and tubes. See Step 10 for details on ES cell passage.

(B) Enzymatic disaggregation

(i) Carefully wash the wells with 500 µl of PBS.

(ii) Aspirate the PBS and add 100 µl of 0.05% (wt/vol) trypsin to each well. Incubate the cells at 37 °C for 5 min.

(iii) By using a mouth-controlled pipette, disaggregate the outgrowth by pipetting up and down vigorously several times.

(iv) If you are following variant B, add 100 µl of defined trypsin inhibitor to inactivate the trypsin. Pipet by gentle trituration.

▲ **CRITICAL STEP** This step is not necessary if you are following variant A, as the FBS in the medium inactivates the trypsin.

(v) Add 500 µl of either FBS ES medium (for variant A) or serum-free ES medium (for variant B). Transfer onto a new feeder-covered four-well plate.

▲ **CRITICAL STEP** At this step, if you are following variant A, outgrowths are now moved from KOSR ES medium to FBS ES cell medium. If you are following variant B, continue culture in serum-free ES medium.

(vi) The incipient ES cell line is now at P1 (passage 1). Increase the passage number by 1 with each subsequent exposure to trypsin and carefully track the passage number by labeling the plates and tubes. See Step 10 for details on ES cell passage.

(vii) Replace the medium with appropriate medium, as described in Step 9B(v), and observe daily. ES cell colonies will form, and they should be visible 3–4 d after disaggregation.

ES cell maintenance ● TIMING 9 d

10| Replace the medium and observe the cells daily. Continue to use FBS ES cell medium if you are following variant A or serum-free medium if you are following variant B. Monitor the growth and morphology and take careful notes. Mixed differentiation within the first 1–2 passages is not unusual; however, continued culture should promote typical ES cell morphology (**Fig. 5**). If poor morphology persists after the first two passages, the incipient ES cell line may be unstable. Expect variable growth rates among emergent lines and be aware that unusually rapid growth could be indicative of karyotypic instability (for example, trisomy for chromosome 8 often results in unusually rapid growth in mESCs). In addition, be alert for signs of deterioration, which include vacuolated cytoplasm, detachment of cells from colonies and cellular debris in the medium. If deterioration continues through passages, discard the cell line. Finally, be aware of signs of contamination, including sudden change in pH, turbidity, small round particles (yeast) and filaments (fungi). Discard contaminated cell lines and quarantine the remaining lines if possible. After disaggregation, passage the cells 1:4 (seeding density is ~1–4 × 10^5 cells per cm^2 or ~3–5 × 10^5 cells per well) onto a six-well MEF feeder plate (or 35-mm dish) according to option A if you are following variant A and according to option B if you are following variant B.

? TROUBLESHOOTING

(A) Passaging cells, variant A, FBS ES cell medium

(i) Rinse the cells with room-temperature PBS.

(ii) Add a quantity of 0.05% (wt/vol) trypsin to cover the cell layer, usually one-third of the amount used to culture the cells.

(iii) Return the cells to the incubator for 4–6 min. Observe cell dissociation with an inverted microscope, with occasional swirling.

(iv) As soon as the colonies begin to dissociate, inactivate the trypsin by adding an equal amount of ES cell medium.

(v) Collect the cells and then remove any remaining cells by rinsing the culture dish with additional ES cell medium. Collect both aliquots (trypsinized cells and the rinse) in one tube and create a single-cell suspension by gentle trituration with a pipette.

(vi) Centrifuge for 5 min at 150*g* at room temperature. Resuspend the pellet in 2.5 ml of ES medium (for a 35-mm plate), creating a single-cell suspension through gentle trituration.

(B) Passaging cells, variant B, serum-free ES cell medium

(i) Rinse the cells with room-temperature PBS.

(ii) Add a quantity of 0.05% (wt/vol) trypsin to cover the layer of cells, usually one-fifth to one-third of the amount used to culture the cells.

(iii) Return the cells to the incubator for 3–5 min.
▲ **CRITICAL STEP** Cells cultured in serum-free culture medium will dissociate more quickly in the presence of trypsin. Avoid unnecessary exposure to trypsin by continually observing cell dissociation with an inverted microscope.

(iv) As soon as colonies begin to dissociate, inactivate the trypsin by adding an equal amount of defined trypsin inhibitor.

(v) Collect the cells and then remove any remaining cells by rinsing the culture dish with serum-free ES medium. Collect both aliquots (trypsinized cells and the rinse) in one tube and create a single-cell suspension by gentle trituration with a pipette.

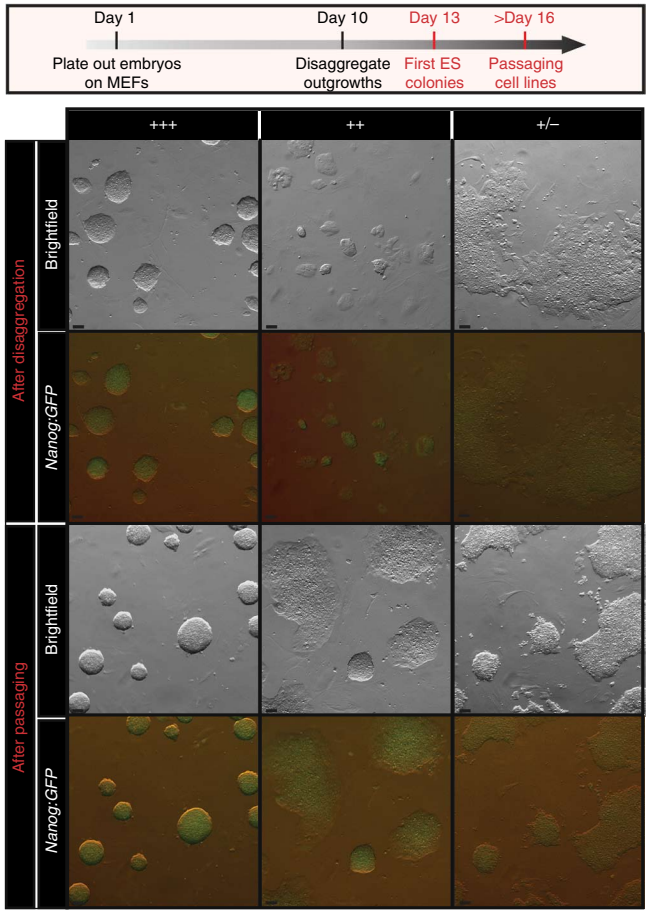


Figure 5 | Emergent mESC morphology. Healthy, emergent mESC lines grow as tightly formed, refractile colonies throughout passaging (+++), and they have robust *Nanog* expression. Lines that show persistent differentiation, flat colony morphology or slow growth through early passaging (+ and +/–) are less desirable and are not likely to emerge as robust, germline-competent mESC lines. Scale bars, 50 µm. All procedures involving mice were approved by The Jackson Laboratory's and Sloan-Kettering Institute's Institutional Animal Care and Use Committees and were performed in accordance with NIH guidelines for the care and use of animals in research.

(vi) Centrifuge the tube for 5 min at 150*g* at room temperature. Resuspend the pellet in 2.5 ml of serum-free ES medium, creating a single-cell suspension through gentle trituration.

11| Continue to culture the cells, replacing the medium with the appropriate medium (FBS ES medium for variant A and serum-free medium for variant B), and observe the cells daily. Once 70% confluency is reached (2–3 d later), passage cells 1:4 onto a 100-mm feeder-covered dish. Alternatively, if genotyping is required, cells can instead be passaged onto two 60-mm dishes, setting aside one for genotyping (Step 19C).
? TROUBLESHOOTING

12| Freeze some of the cells in culture when they reach 70% confluency. Change the medium the morning before a freeze to minimize stress on the cells. Follow the passaging protocol in Step 10 above and resuspend in 1 ml of the appropriate medium (FBS ES medium for variant A and serum-free medium for variant B). Count the cells by using either a manual or automatic cell counter. Centrifuge the cells for 5 min at 150*g* at room temperature. Resuspend the cells in cryopreservation medium at a concentration of 3×10^6 cells per ml. Add 1 ml of medium per cryovial. Transfer each vial to a CoolCell (or other container that will allow for a cool rate of –1 °C/min) and place it at –80 °C. After 24 h, transfer the cells to liquid nitrogen. The frozen cells are now at P3–P4 (see Step 9B(vi) regarding passage number assignment), depending on whether the initial disaggregation (Step 9) was enzymatic (P4) or mechanical (P3).
▲ **CRITICAL STEP** It is best practice to freeze cells as soon as possible to create a stock of ultra-low-passage vials.
? TROUBLESHOOTING
■ **PAUSE POINT** Cells can now be stored in liquid nitrogen indefinitely. Carefully maintain these stocks and plan subsequent thaws and expansions to preserve low-passage vials.

ESC characterization ● TIMING ~3 weeks
13| One day before seeding ES cells, thaw and plate mitotically inactivated MEFs. For the full characterization of one cell line, two 60-mm dishes and six wells on a 12-well plate are needed. In addition, for a full characterization, two gelatinized 60-mm plates without feeders are needed. Incubate MEF feeder plates or dishes at 37 °C, 5% CO_2.

14| The day after MEF plating, remove a vial of frozen ESCs from liquid nitrogen. Place the vial in a 37 °C water bath until the cells begin to thaw.

15| Sterilize the vial with 70% (vol/vol) ethanol and then transfer the cells to 10 ml of prewarmed (37 °C) medium in a 15-ml conical tube.

16| Centrifuge the cells for 5 min at 150*g* at room temperature.

17| Resuspend the pellet in 5 ml of medium and create a single-cell suspension by gentle trituration with a serological pipette. Plate the cells onto a 60-mm MEF feeder plate.

18| Once the cells have reached 70% confluency, trypsinize the cells as detailed in Step 10, resuspend them in 4 ml of ES medium and divide them into four 1-ml aliquots in separate 15-ml conical tubes.

19| Seed the cells according to the instructions below and proceed with immunolabeling for markers of pluripotency (option A), mitotic chromosome counting (option B), DNA extraction for genotyping (option C) and pathogen testing (option D). We recommend that users perform all of these options to enable complete mESC characterization.
(A) Immunolabeling for markers of pluripotency ● TIMING 5 h
(i) Add an additional 4 ml of ES medium to one of the 1-ml aliquots of cell suspension. Seed 0.5 ml of the resulting suspension into each of six wells of a 12-well plate MEF feeder plate.
(ii) Incubate the plate at 37 °C, 5% CO_2 for 1–2 d to allow cells to reach 50% confluency. Change the medium on the morning of the prep.
(iii) Rinse the cells with room-temperature PBS.
(iv) Fix the cells in 4% (wt/vol) PFA for 20 min at room temperature.
(v) Wash the cells with PBS three times for 5 min each.
■ **PAUSE POINT** Fixed cells may be kept at 4 °C in PBS for up to 2 weeks.
(vi) Incubate the cells in all wells with blocking buffer for 15 min on a rocking platform at room temperature.
(vii) Incubate the cells in primary antibody solution for at least 1 h at room temperature on a rocking platform (one well for each primary antibody, anti-OCT3/4, anti-NANOG and anti-SSEA1, and one well for each 'no primary' control option).

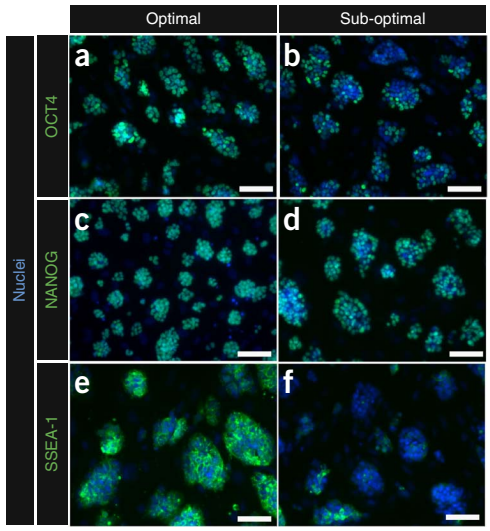


Figure 6 | mESC characterization by immunolabeling of pluripotency biomarkers. The pluripotency markers OCT3/4, NANOG and SSEA-1 can be used to assess pluripotency in new mESC lines. (**a**–**f**) Heterogeneous labeling (**b**,**d**,**f**) is associated with suboptimal GLT when compared with lines that show optimal, consistent labeling (**a**,**c**,**e**) for these markers. Scale bars, 50 µm. All procedures involving mice were approved by The Jackson Laboratory's and Sloan-Kettering Institute's Institutional Animal Care and Use Committees and were performed in accordance with NIH guidelines for the care and use of animals in research.

(viii) Wash the cells with PBS three times for 5 min each.

(ix) Incubate the cells in appropriate secondary antibody solution (e.g., goat anti-mouse IgG Alexa Fluor 488, goat anti-rabbit IgG Alexa Fluor 488 or goat anti-mouse IgM Alexa Fluor 488) for at least 1 h at room temperature on a rocking platform.

(x) Wash the cells with PBS three times for 5 min each.

(xi) Incubate the cells with DAPI for 10 min.

(xii) Wash the cells with PBS once for 5 min.

(xiii) Image the cells with an inverted epifluorescence microscope. Antibodies specific to NANOG and OCT3/4 should label the nuclei of ES cells, whereas SSEA-1–specific antibody will label the surface of ES cells. Proceed with ES cell lines that show maximum distribution of labeling, as shown in **Figure 6** (optimal versus suboptimal). If desired, quantitative assessment of labeling may also be performed by using flow cytometry[20].

▲ **CRITICAL STEP** Even robust, germline-competent mESCs are heterogeneous in their pluripotency marker expression, and in poorly performing mESC lines there can still be some pluripotency marker expression. Here, the goal of assessing pluripotency marker expression is to identify those lines with the most homogeneous and consistent expression of each maker, especially NANOG[26].

(B) Mitotic chromosome counting ● **TIMING 6 h**

(i) Take 1 ml of the cell suspension and add an additional 4 ml of ES medium (**Fig. 7**). Seed the cells onto a 60-mm gelatinized dish.

(ii) Incubate the dish at 37 °C, 5% CO_2 for 1–2 d to allow the cells to reach 50–70% confluency.

(iii) Add 10 µl of 10 µg/ml colcemid to 5 ml of fresh, warmed ES medium.

(iv) Remove the spent ES cell medium from one of the cultures growing on a gelatinized 60-mm plate and add the colcemid-containing medium to the culture.

(v) Incubate the dish for 3 h at 37 °C.

(vi) Trypsinize the cells as detailed in Step 10.

(vii) Aspirate and discard the supernatant, leaving about 500 µl of medium in the tube. Gently pipet up and down to resuspend the pellet.

(viii) Set a water bath to 75 °C.

(ix) By using a P1000 pipette, add 5 ml of 0.56% (wt/vol) KCl to the cells slowly.

▲ **CRITICAL STEP** KCl should be added in a drop-by-drop manner to prevent lysis. Carefully invert the cells to mix.

(x) Incubate the mixture for 60 min at room temperature.

(xi) Centrifuge the tube for 5 min at 150*g*.

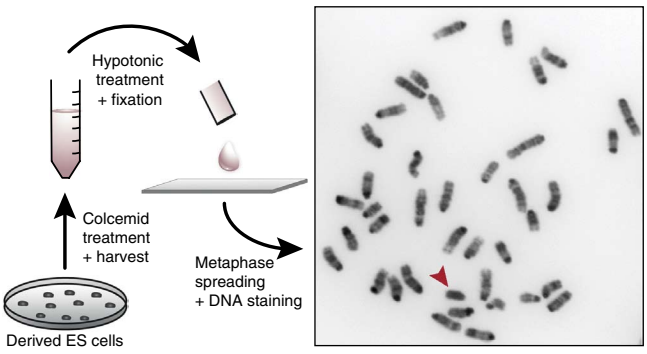


Figure 7 | Karyotypic analysis of ES cells. mESC lines are treated with colcemid to arrest cells in metaphase. Cells are fixed and ruptured by hypotonic treatment and nuclei are spread on glass slides. Mitotic chromosomes are stained with DAPI, which is a DNA-intercalating dye that preferentially binds A-T–rich DNA, resulting in a reverse G-banding pattern. For genetic engineering, male mESC lines are preferable for breeding efficiency. The Y chromosome (arrowhead) stains homogeneously with DAPI, and it is clearly distinguishable from the other chromosomes, which show discrete banding patterns. All procedures involving mice were approved by The Jackson Laboratory's and Sloan-Kettering Institute's Institutional Animal Care and Use Committees and were performed in accordance with NIH guidelines for the care and use of animals in research.

(xii) Aspirate and discard the supernatant, leaving 1 ml of the hypotonic solution. Resuspend the pellet by gentle pipetting.
(xiii) In a drop-by-drop manner, slowly add 1 ml of 3:1 methanol:glacial acetic acid. Swirl the tube as the fixative is added.
(xiv) Centrifuge the tube for 5 min at 150*g*.
(xv) Aspirate and discard the supernatant, leaving 1 ml of fixative. Resuspend the pellet by gentle pipetting.
(xvi) Slowly add 3 ml of fixative, by swirling to mix.
(xvii) Centrifuge the tube for 5 min at 150*g* at room temperature.
(xviii) Repeat Step 19B(xiv–xvi) twice.
(xix) Resuspend the pellet in 1 ml of fixative.
■ **PAUSE POINT** Cells may be stored in fixative at –20 °C for several years.
(xx) Steam a freshly cleaned slide over a water bath.
(xxi) Immediately pipet 75 µl of the cell suspension onto the steamed slide from a height of 6 inches. Tilt the slide and blow gently to disperse the cells.
(xxii) Place the slide on a hot plate to dry.
(xxiii) Store the slide for 24–48 h on a slide warmer at 37 °C.
(xxiv) Apply a drop of Vectashield hard-set mounting medium with DAPI to the slide, and then apply a coverslip.
(xxv) Acquire images of metaphase spreads at ×60 magnification. Analyze at least 20 spreads per cell line by counting the total chromosome number and assessing sex chromosomes (XX or XY), as shown in **Figure 7**. Determine the percentage of spreads with 40 chromosomes. For example, if 19 of 20 spreads have 40 chromosomes and a Y chromosome is observed in every spread, the cell line is 95% euploid, 40 XY. To select for mESC lines that will be used for genetic engineering, male lines are preferable for maximum breeding efficiency.

(C) DNA extraction for SNP panel and/or genotyping

● **TIMING 1 h**

(i) Take 1 ml of the cell suspension and add an additional 4 ml of ES medium. Seed the suspension onto a 60-mm gelatinized dish.
▲ **CRITICAL STEP** For optimal and accurate mESC-specific SNP genotyping, it is important to exclude feeders.

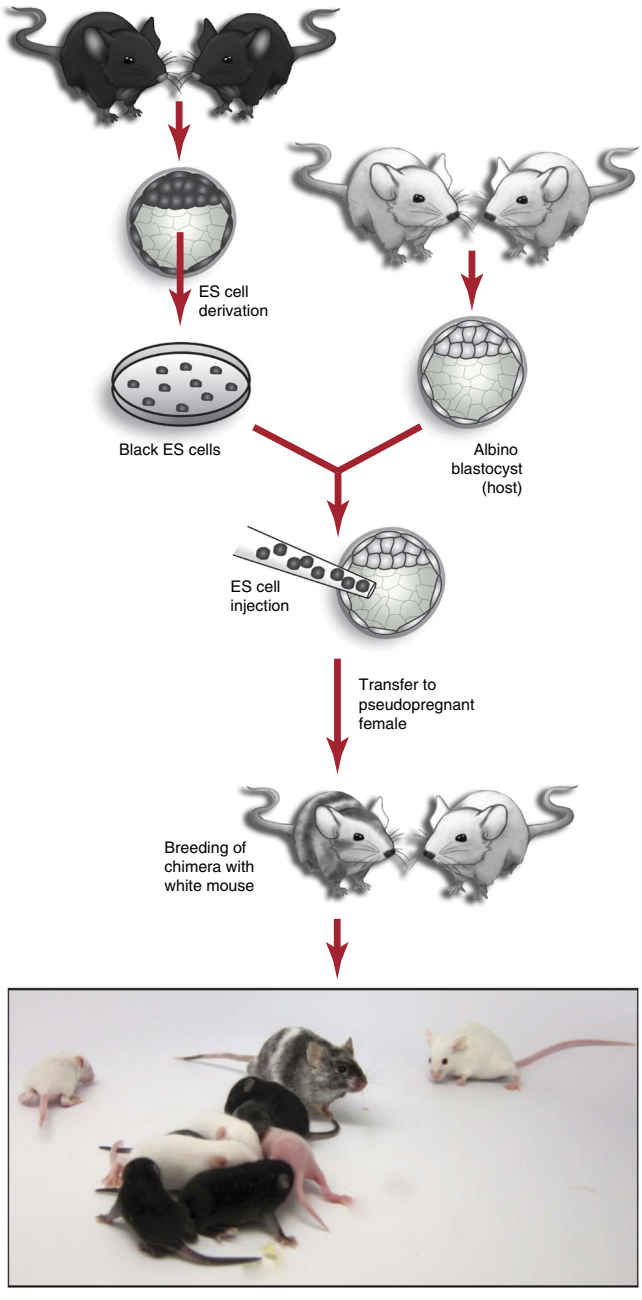


Figure 8 | Germline testing of mESC lines. mESC lines are injected into host blastocyst embryos, and mESC contribution to resulting chimeras is assessed by coat color. In this example, mESCs derived from a nonalbino strain (e.g., C57BL/6J, genotype *a/a*, Tyr^{+}/Tyr^{+}) are injected into host blastocysts from an albino strain (e.g., C57BL-Tyr^{c-2J}/J, genotype a/a, Tyr^{c}/Tyr^{c}). The resulting chimeras will have patches of fur that are albino (contributed by the host blast) and nonalbino (in this example, black (*a/a*) pigmentation is contributed by ES cells). Contribution from the mESC line is evidenced by the overall percentage of nonalbino coat color. Generally, high contribution by a male ES cell line will also produce predominantly male chimeras. Female chimeras are possible, but they are less productive. The male chimeras are then bred with an albino strain (in this case, a/a, Tyr^{c}/Tyr^{c}) to ensure that germ line transmission can be assessed by coat color. In this example, pigmentation is only possible in offspring that are *a/a*, Tyr^{+}/Tyr^{c} (black pups in image) where the wild-type *Tyr* and an *a* allele are contributed by sperm derived from the ES cell line. Albino offspring are *a/a*, Tyr^{c}/Tyr^{c} where the Tyr^{c} allele and an *a* allele are contributed by sperm derived from the host blastocyst. Pups shown are from two litters, 7 and 11 d postpartum. All procedures involving mice were approved by The Jackson Laboratory's and Sloan-Kettering Institute's Institutional Animal Care and Use Committees and were performed in accordance with NIH guidelines for the care and use of animals in research.

(ii) Incubate the cells at 37 °C, 5% CO_2 for 1–2 d to allow the cells to reach 50–70% confluency.
(iii) Extract DNA from a confluent culture with a Qiagen DNeasy kit, according to the manufacturer's instructions, and then perform PCR as needed for genotyping.
(iv) To confirm the genetic background of the established cell line, submit DNA to a SNP panel, e.g., the KBioscience mouse SNP panel[28].

(D) Pathogen testing ● TIMING 2 weeks

(i) Take 1 ml of the cell suspension and centrifuge it again at 150*g* for 5 min at room temperature, wash the pellet in PBS and resuspend it in 5 ml of antibiotic-free ES medium.
▲ **CRITICAL STEP** Culture the cells in the absence of antibiotics for the remainder of the pathogen test.
(ii) Seed the cells on a 60-mm MEF feeder dish.
(iii) Incubate the cells at 37 °C, 5% CO_2 for 1–2 d until they are fully confluent. Remove and reserve the 5 ml of used medium.
▲ **CRITICAL STEP** Contaminants can be harbored in the medium, and thus it is important that the used medium, in addition to the cells, be tested.
(iv) Trypsinize the cells as detailed in Step 10.
(v) Resuspend the cells in the used medium, by dividing them into three tubes: 0.5 ml for bacterial testing, 2 ml for fungal testing and 2.5 ml for *Mycoplasma* testing.
(vi) For bacterial testing, add the 0.5-ml aliquot from Step 19D(iv) to 4.5 ml of tryptose phosphate broth. Incubate the mixture at 37 °C for 48 h. Check for growth.
(vii) For fungal testing, inoculate a Sabouraud plate with the 2 ml of culture. Culture it for 14 d at room temperature. Check for growth.
(viii) For *Mycoplasma* testing, centrifuge the 2.5 ml of culture at 500*g* for 5 min at room temperature and aspirate until 200 µl of medium remains. Extract DNA by using a DNeasy kit according to the manufacturer's instructions. Test for the presence of *Mycoplasma* by using a PCR-based detection method (e.g., van Kuppeveld *et al.*[29]).
(ix) If desired, submit sample cells for a molecular mouse antibody production (MAP) test to screen for mouse pathogens in accordance with institutional policies.

Germline transmission (GLT) of parental lines

● TIMING 8 weeks

20| Thaw a well-characterized ES cell line onto a 60-mm MEF plate as detailed in Steps 14–17.
▲ **CRITICAL STEP** To maximize the success of GLT, the mES cell line should robustly express pluripotency markers and have a stable euploid karyotype of at least 70%, 40 XY.

21| Allow the cells to reach 70% confluency.

22| Dissociate the cells with trypsin.

TABLE 2 | Data collected to assess mESC suitability for chimera contribution and GLT.

Development	
No. of embryos transferred	50
No. of pups born	9–25
% born	18–50%
Chimeras	
No. of chimeras	3–14
% chimeras	30–60%
No. of male chimeras	2–14
% male chimeras	80–100%
Coat color contribution	
100%	0–8
>80%	0–5
40–80%	0–7
<40%	0–8
Fertility	
No. set up	2–14
No. productive	1–10
% productive	30–70%
GLT	
No. of chimeras with GLT	1–10
% GLT	50–100%
Rates of GLT	
100%	0–10
>50%	0–10
10–50%	0
<10%	0

Acceptable ranges are provided on the basis of our accumulated data from microinjection of 10, low passage (P6–P11), parental (mESCs that have not been genetically engineered) mESCs from a variety of strains, including BALB/cByJ, C3H/HeJ, PWD/PhJ, NOD/ShiLtJ and C57BL/6J. Each ES cell line was injected into 50 blastocyst stage host embryos from C57BL/6J or C57BL-*Tyrc-2J*/J strains. The percentage born is the number of viable offspring born divided by the number of embryos transferred and can vary substantially depending on the microinjection facility as well as mESC quality. The percentage of chimeras is the number of chimeras born divided by the total number born. The percentage of male chimeras is the number of chimeric males divided by the total males born. Coat color contribution is based on a subjective assessment of mESC contribution to the host embryo based on coat color in wean age animals. This assessment could vary based on microinjection facility and operator. The percentage of GLT is the number of animals demonstrating GLT divided by the number of productive animals. To assess fertility, chimeras are housed with proven breeders (typically as trios) at the age of 8–12 weeks and are kept for up to 12 weeks (to generate 20 offspring per chimera). Chimeras that do not produce offspring within this timeframe are considered nonproductive. GLT rate is the number of ES cell-derived offspring (by coat color) divided by the total offspring (~20).

23| Inactivate the trypsin and add an additional 3 ml of medium to the dish.

24| Return the cells to the incubator for 20 min to allow feeder cells to settle to the bottom of the dish.

25| Carefully remove the supernatant to a 15-ml conical tube.

26| Centrifuge the mixture for 5 min at 150*g* at room temperature.

27| Wash the cells with PBS.

28| Centrifuge the mixture for 5 min at 150*g* at room temperature.

29| Resuspend the cells in 1 ml of medium buffered with 10 mM HEPES.

30| Deliver cells to the microinjection facility. Request host embryos from a strain that will allow for determination of mESC contribution by coat color, as depicted in **Figure 8**. For a review of mouse coat-color loci, see Jackson[30].
▲ **CRITICAL STEP** Microinjection of mESCs into host blastocyst embryos requires specialized equipment and expertise and is usually provided by a core facility. For laboratories interested in establishing in-house microinjection/chimera production capability, detailed protocols are available from a variety of sources, including Nagy *et al.*[27]).

31| Collect the relevant data for 50–100 injections, as shown in **Table 2**. Generate ~20 offspring per chimera, assess GLT by using coat color as a guide (**Fig. 8**) and reference the data shown in **Table 2**. Up to 60% of the viable offspring from microinjection of a robust, parental male mES cell line should be chimeric, and 80–100% of these should be male. Moreover, at least 50% of the chimeric males should have GLT rates of 50% or more.
? TROUBLESHOOTING

? TROUBLESHOOTING
Troubleshooting advice can be found in **Table 3**.

TABLE 3 | Troubleshooting table.

Step	Problem	Possible reason	Solution
8	Blastocyst failure to attach or hatch	Blastocyst not fully developed	Plate only fully expanded blastocysts
		Unhealthy MEF feeder layer	Verify that MEFs are seeded at correct density and test for ability to support ES cell viability and growth with an established mESC line
		Embryos are unable to hatch from the zona pellucida	Dissolve the zona pellucida using acid Tyrode's before plating the blastocyst onto a feeder layer
10–12	Differentiation	Improper composition of culture media	Verify that media were made correctly and check dates on all media components
		Blastocyst attachment to edge of feeder layer/well	Blastocysts should be placed in the center of the well to avoid attachment to the sides of the well. Dishes can be rotated clockwise and/or counterclockwise to swirl the embryos into the center of each circular well
		Overgrowth of cultures	Outgrowths and colonies should be dissociated before the center of the outgrowth becomes dark
		Strain is nonpermissive	Repeat the derivation according to variant B

(continued)

TABLE 3 | Troubleshooting table (continued).

Step	Problem	Possible reason	Solution
10–12	Poor growth, and/or sudden change in media color, cloudiness in medium, presence of fungi, yeast, bacteria	Contamination	Clean and sterilize the cell culture incubator and review proper aseptic technique. Discard all contaminated cultures and quarantine any cultures growing in the same incubator. Discard and replace all media
	Excessive cell death, poor survival after passage	Improper composition of culture media	Ensure that media were prepared correctly and check all expiration dates. All culture media should be stored at 4 °C and should be protected from light. Media should be discarded and replaced after 1–2 weeks
		Excessive exposure to trypsin	Inhibit trypsin as soon as cells begin to dissociate
11	Poor post-thaw viability	Improper storage conditions	Cells should be initially frozen at –80 °C in a container that allows for a controlled freeze (1 °C per minute) and then rapidly transferred to liquid N_2 after 1–7 d. Also, be sure to check the integrity of cryovials
		Improper medium composition	Verify that cryopreservation medium was made correctly
		Poor handling	Always minimize the exposure of thawed cells to the cryopreservation medium, which contains DMSO. DMSO is an effective adjuvant but is also toxic to cells at ambient temperatures. Cells require a slow freeze and a rapid thaw. To thaw rapidly, use a 37 °C water or bead bath. Sterilize the outside of the vial and immediately transfer thawed contents to medium immediately after thawing
31	GLT is not detected	Selected breeding scheme does not allow for assessment of GLT by coat color	Verify that the coat color genotype of the ES cell line, the host blastocyst and breeder strain will allow for assessment of GLT by coat color. If the wrong strains have been selected, it may be possible to genotype offspring using strain-specific SNPs. Otherwise, select a different strain to breed chimeras or repeat with a different host blastocycst
		Cells are karyotypically abnormal	Verify the karyotype of the cell line; subclone or re-derive if necessary
		Pluripotency is limited, cells have poor viability or poor proliferative potential	If a high-passage vial was used, choose a lower passage for testing. Alternatively (and more likely), this result may indicate that the mESC line is inadequate. In this case, test a different line from the same derivation, or repeat the derivation according to variant B

● TIMING

Steps 1 and 2, preparation of MEF feeder plates: 1 d
Steps 3–6, collecting E3.5 blastocysts: 15 min per pregnant female
Steps 7–9, plating and early culture: 7–9 d
Steps 10–12, ES cell maintenance: 9 d
Steps 13–19, ESC characterization: 3 weeks (including all Step 19 options)
Step 19A, immunolabeling for markers of pluripotency: 5 h
Step 19B, mitotic chromosome counting: 6 h
Step 19C, DNA extraction for SNP panel and/or genotyping: 1 h
Step 19D, pathogen testing: 2 weeks
Steps 20–31, GLT of parental lines: 8 weeks

ANTICIPATED RESULTS

By using this protocol, we have produced germline-competent ES cell lines from a variety of mouse strains, including those thought to be nonpermissive, such as DBA[20], PWD and NOD (A.C. and L.G.R., unpublished data). The careful selection of healthy, blastocyst-stage embryos, combined with minimal manipulations before plating promotes a high yield of ICM outgrowths from embryos that hatch naturally from the zona pellucida and adhere to MEF feeders. Under the derivation conditions described here, up to 90% of ICM outgrowths will yield incipient ES cell lines, with some variation depending on the strain background and the experience of the technicians. During the outgrowth phase, exclusion of serum from the derivation medium minimizes trophectoderm differentiation and 2i/LIF culture promotes expansion of the epiblast.

In incipient ES cell lines, we have found that serum-free 2i/LIF medium (variant B) promotes robust NANOG expression and, ultimately, higher yield of coat-color chimeras through germline testing[20]. Although serum-free culture conditions are typically used in feeder-free culture systems, we have found that the use of feeders promotes more efficient derivation (ratio of the number of embryos giving rise to ES cell lines/number of embryos plated) and also yields a higher percentage of adherent ES cell lines, which are more favorable for gene-targeting experiments. However, not surprisingly, MEF viability in the absence of serum is highly variable. This can be overcome with careful selection of low-passage feeder MEFs that have been previously tested for survival in serum-free conditions and careful monitoring of feeder layers during culture. In addition, in the context of a gene-targeting experiment, serum can be temporarily introduced during drug selection to promote MEF survival after electroporation. Given these complexities, if a particular mouse strain has not been previously characterized as nonpermissive (**Table 1**), derivation of ES cells in the presence of serum (variant A) should be attempted first. If robust pluripotent, karyotypically stable ES cell lines are not achieved with variant A, then serum-free derivation should follow.

It is important to closely monitor the growth of incipient mES cell lines with careful observation of morphology and density, and to replenish the medium daily. The timing of passages is dictated by the confluency of the culture (and not convenience), and therefore it will vary from one incipient cell line to the next. Cell lines should be cryopreserved at P3–P4, and subsequent culturing and freezing should be carefully planned to ensure ample frozen stocks of low-passage cells. Generally, as the passage number increases, genomic stability and pluripotency decline. Moreover, there is emerging evidence (though still anecdotal) that although 2i/LIF conditions promote pluripotency, prolonged culture may promote aneuploidy; therefore, it is important to consider weaning newly derived, characterized mESC lines from 2i. With the exception of lines created by variant B, we maintain low-passage stocks in 2i/LIF and then create higher-passage working stocks that are weaned from 2i. To do this, cells are essentially brought through several passages with decreasing doses of 2i (e.g., dividing the dose by two at each passage), and then working stock vials are cryopreserved.

Before proceeding with germline testing or *in vitro* differentiation, newly derived ESC lines should be extensively characterized. mES cell lines should be free of *Mycoplasma*, bacteria, fungi and mouse pathogens. SNP paneling is recommended to confirm the strain of origin of mES cells, and karyotypic assessment should be used to determine sex and percentage euploidy. For gene targeting, male ES cell lines that are 70% euploid (40 XY) or higher are acceptable, but <90% is ideal. Karyotypic stability is stochastic in cultured cells, and it should be reassessed after extensive passaging, as in a gene-targeting experiment. Cell lines that have developed an unstable karyotype may be subcloned; however, given the efficiency with which ES cells can be derived, re-derivation is preferable. Finally, to ensure maximum developmental potential reflected in robust chimera contribution, highly euploid mESCs should robustly express well-accepted markers of the pluripotent state such as NANOG, OCT4, SOX2 and SSEA-1.

ACKNOWLEDGMENTS This work was supported in part by the US National Institutes of Health (NIH), Office of Research Infrastructure Programs (grant nos. U42-OD011102, U42-OD010921). Work in A.-K.H.'s laboratory is supported by the Human Frontiers Sciences Program and the NIH (grant nos. RO1-HD052115 and RO1-DK084391). We are grateful for the excellent microinjection services provided by the Cell Biology and Microinjection core at The Jackson Laboratory.

AUTHOR CONTRIBUTIONS A.C. and L.G.R. developed the protocol and wrote the manuscript. C.B., I.G. and L.R.D. contributed to and/or supported the development of the protocol. N.S. and A.-K.H. contributed to and expanded the protocol and contributed to the writing manuscript.

COMPETING FINANCIAL INTERESTS The authors declare no competing financial interests.

1. Evans, M.J. & Kaufman, M.H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154–156 (1981).
2. Martin, G.R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. USA* **78**, 7634–7638 (1981).
3. Brook, F.A. & Gardner, R.L. The origin and efficient derivation of embryonic stem cells in the mouse. *Proc. Natl. Acad. Sci. USA* **94**, 5709–5712 (1997).
4. Bradley, A., Hasty, P., Davis, A. & Ramirez-Solis, R. Modifying the mouse: design and desire. *Biotechnology* **10**, 534–539 (1992).
5. Eppig, J.T., Blake, J.A., Bult, C.J., Kadin, J.A. & Richardson, J.E. The Mouse Genome Database (MGD): comprehensive resource for genetics and genomics of the laboratory mouse. *Nucleic Acids Res.* **40**, D881–D886 (2012).
6. Bradley, A. *et al.* The mammalian gene function resource: the International Knockout Mouse Consortium. *Mamm. Genome* **23**, 580–586 (2012).

7. Kawase, E. *et al.* Strain difference in establishment of mouse embryonic stem (ES) cell lines. *Int. J. Dev. Biol.* **38**, 385–390 (1994).
8. Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W. & Roder, J.C. Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proc. Natl. Acad. Sci. USA* **90**, 8424–8428 (1993).
9. Buehr, M. *et al.* Rapid loss of Oct-4 and pluripotency in cultured rodent blastocysts and derivative cell lines. *Biol. Reprod.* **68**, 222–229 (2003).
10. Yang, H. *et al.* One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* **154**, 1370–1379 (2013).
11. Zschemisch, N.H. *et al.* Zinc-finger nuclease mediated disruption of Rag1 in the LEW/Ztm rat. *BMC Immunol.* **13**, 60 (2012).
12. Hauschild, J. *et al.* Efficient generation of a biallelic knockout in pigs using zinc-finger nucleases. *Proc. Natl. Acad. Sci. USA* **108**, 12013–12017 (2011).
13. Handyside, A.H. *et al.* Biopsy of human preimplantation embryos and sexing by DNA amplification. *Lancet* **1**, 347–349 (1989).
14. Buehr, M. & Smith, A. Genesis of embryonic stem cells. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **358**, 1397–1402: discussion 1402 (2003).
15. Ying, Q.L. *et al.* The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519–523 (2008).
16. Buehr, M. *et al.* Capture of authentic embryonic stem cells from rat blastocysts. *Cell* **135**, 1287–1298 (2008).
17. Li, P. *et al.* Germline competent embryonic stem cells derived from rat blastocysts. *Cell* **135**, 1299–1310 (2008).
18. Silva, J. *et al.* Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS Biol.* **6**, e253 (2008).
19. Hanna, J. *et al.* Metastable pluripotent states in NOD-mouse-derived ESCs. *Cell Stem Cell* **4**, 513–524 (2009).
20. Reinholdt, L.G. *et al.* Generating embryonic stem cells from the inbred mouse strain DBA/2J, a model of glaucoma and other complex diseases. *PLoS One* **7**, e50081 (2012).
21. Nichols, J. *et al.* Validated germline-competent embryonic stem cell lines from nonobese diabetic mice. *Nat. Med.* **15**, 814–818 (2009).
22. Tong, C., Li, P., Wu, N.L., Yan, Y. & Ying, Q.L. Production of *p53* gene knockout rats by homologous recombination in embryonic stem cells. *Nature* **467**, 211–213 (2010).
23. Nichols, J., Silva, J., Roode, M. & Smith, A. Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. *Development* **136**, 3215–3222 (2009).
24. Ying, Q.L. & Smith, A.G. Defined conditions for neural commitment and differentiation. *Methods Enzymol.* **365**, 327–341 (2003).
25. Markoulaki, S., Meissner, A. & Jaenisch, R. Somatic cell nuclear transfer and derivation of embryonic stem cells in the mouse. *Methods* **45**, 101–114 (2008).
26. Silva, J. *et al.* Nanog is the gateway to the pluripotent ground state. *Cell* **138**, 722–737 (2009).
27. Nagy, A., Gertsenstein, M., Vintersten, K. & Behringer, R. Production of chimeras. in *Manipulating the Mouse Embryo: a Laboratory Manual* 453–506 (Cold Spring Harbor Laboratory Press, 2003).
28. Petkov, P.M. *et al.* An efficient SNP system for mouse genome scanning and elucidating strain relationships. *Genome Res.* **14**, 1806–1811 (2004).
29. van Kuppeveld, F.J. *et al.* 16S rRNA-based polymerase chain reaction compared with culture and serological methods for diagnosis of *Mycoplasma pneumoniae* infection. *Eur. J. Clin. Microbiol. Infect. Dis.* **13**, 401–405 (1994).
30. Jackson, I.J. Molecular and developmental genetics of mouse coat color. *Annu. Rev. Genet.* **28**, 189–217 (1994).
31. Roach, M.L. & McNeish, J.D. Methods for the isolation and maintenance of murine embryonic stem cells. *Methods Mol. Biol.* **185**, 1–16 (2002).
32. Roach, M.L., Stock, J.L., Byrum, R., Koller, B.H. & McNeish, J.D. A new embryonic stem cell line from DBA/1lacJ mice allows genetic modification in a murine model of human inflammation. *Exp. Cell Res.* **221**, 520–525 (1995).
33. Brown, D.G., Willington, M.A., Findlay, I. & Muggleton-Harris, A.L. Criteria that optimize the potential of murine embryonic stem cells for *in vitro* and *in vivo* developmental studies. *In Vitro Cell. Dev. Biol.* **28A**, 773–778 (1992).
34. Nagafuchi, S. *et al.* Establishment of an embryonic stem (ES) cell line derived from a non-obese diabetic (NOD) mouse: *in vivo* differentiation into lymphocytes and potential for germ line transmission. *FEBS Lett.* **455**, 101–104 (1999).
35. Gardner, R.L. & Brook, F.A. Reflections on the biology of embryonic stem (ES) cells. *Int. J. Dev. Biol.* **41**, 235–243 (1997).

PUBLICATION IV

AID STABILIZES STEM-CELL PHENOTYPE BY REMOVING EPIGENETIC MEMORY OF PLURIPOTENCY GENES

Kumar R., DiMenna L., Schrode N., Liu T.C., Franck P., Muñoz-Descalzo S., Hadjantonakis A.K., Zarrin A.A., Chaudhuri J., Elemento O., Evans T.

Nature
500(7460):89-92
2013

LETTER

doi:10.1038/nature12299

AID stabilizes stem-cell phenotype by removing epigenetic memory of pluripotency genes

Ritu Kumar[1], Lauren DiMenna[2], Nadine Schrode[3], Ting-Chun Liu[1], Philipp Franck[1], Silvia Muñoz-Descalzo[3], Anna-Katerina Hadjantonakis[3], Ali A. Zarrin[4], Jayanta Chaudhuri[2], Olivier Elemento[5] & Todd Evans[1]

The activation-induced cytidine deaminase (AID; also known as AICDA) enzyme is required for somatic hypermutation and class switch recombination at the immunoglobulin locus[1]. In germinal-centre B cells, AID is highly expressed, and has an inherent mutator activity that helps generate antibody diversity[2]. However, AID may also regulate gene expression epigenetically by directly deaminating 5-methylcytosine in concert with base-excision repair to exchange cytosine[3]. This pathway promotes gene demethylation, thereby removing epigenetic memory. For example, AID promotes active demethylation of the genome in primordial germ cells[4]. However, different studies have suggested either a requirement[5] or a lack of function[6] for AID in promoting pluripotency in somatic nuclei after fusion with embryonic stem cells. Here we tested directly whether AID regulates epigenetic memory by comparing the relative ability of cells lacking AID to reprogram from a differentiated murine cell type to an induced pluripotent stem cell. We show that *Aid*-null cells are transiently hyper-responsive to the reprogramming process. Although they initiate expression of pluripotency genes, they fail to stabilize in the pluripotent state. The genome of *Aid*-null cells remains hypermethylated in reprogramming cells, and hypermethylated genes associated with pluripotency fail to be stably upregulated, including many MYC target genes. Recent studies identified a late step of reprogramming associated with methylation status[7], and implicated a secondary set of pluripotency network components[8]. AID regulates this late step, removing epigenetic memory to stabilize the pluripotent state.

The path to pluripotency involves multiple steps that can be more or less efficient, and this can be modelled in transcription-factor-induced reprogramming[9]. Reprogrammed induced pluripotent stem cell (iPSC) clones initially retain methylation patterns that may reflect an epigenetic memory of the source lineage[10,11]. To test directly the function of AID in reprogramming, we prepared tail-tip fibroblasts (TTFs) from *Aid*-knockout or control sibling mice. Before reprogramming, the fibroblasts appeared morphologically identical (Supplementary Fig. 1a), and were transduced equivalently with lentivirus encoding the four human 'Yamanaka' reprogramming transcription factors[12]: OCT4, KLF4, SOX2 and cMYC (OKSM; Supplementary Fig. 1b). Expression levels of exogenous human OCT4 72 h after infection (before induction of the endogenous murine gene) were equivalent, showing that lack of AID does not affect expression vector function (Supplementary Fig. 1c). Exogenous transcription factor expression was subsequently equivalently repressed in wild-type and *Aid*-null transduced cells (Supplementary Fig. 1d). Using quantitative polymerase chain reaction with reverse transcription (RT–qPCR) assays, transcript levels for *Aid* were not reliably detected above background in wild-type TTFs. However, after 1 week of reprogramming, transcripts were readily measured in wild-type cells, increasing as much as tenfold during the reprogramming process (Supplementary Fig. 2).

Initial reprogramming steps include the induction of proliferation and a change in the morphology of fibroblasts to smaller and rounder cells[13]. We found that cells lacking AID are initially hyper-responsive to reprogramming factors. The change in morphology was more rapid in $Aid^{-/-}$ cells, beginning at 2 days after transduction. After 4 days, $Aid^{-/-}$ cells were rounded and smaller than $Aid^{+/+}$ cells (Fig. 1a). A higher fraction of $Aid^{-/-}$ cells stained positive for SSEA1, an early marker for pluripotency[14] (Fig. 1b). At day 7, more *Aid*-null cells expressed NANOG compared with controls (Fig. 1c), correlating with modestly higher transcript levels for several pluripotency genes before the first week of reprogramming (Fig. 1d). Consistent with the molecular data, at 2 weeks the *Aid*-null cells consistently generated more early colonies than wild-type cells (sixfold more colonies on average over six independent experiments). Thus, *Aid*-null fibroblasts are hyper-responsive to reprogramming, even though the growth curves for uninfected wild-type and *Aid*-null cells were indistinguishable (Supplementary Fig. 3a). Rather, the enhanced expression of pluripotency markers in $Aid^{-/-}$ cells suggests that AID normally helps to stabilize the differentiated state, creating a barrier to the initial process of reprogramming. When the $Aid^{-/-}$ fibroblasts were passaged before transduction, this hyper-responsiveness was no longer seen (Supplementary Fig. 3b), suggesting

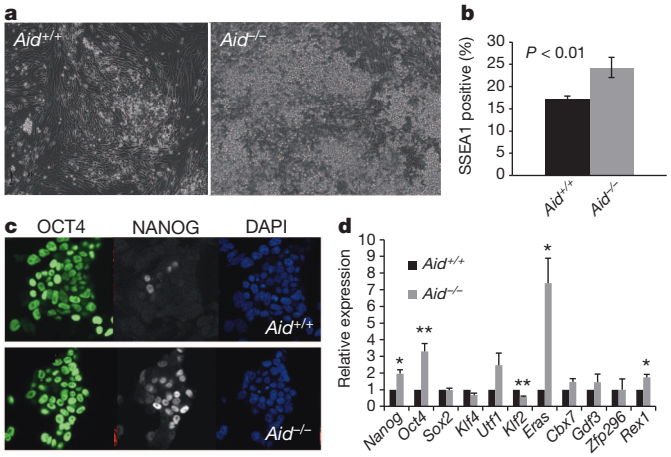


Figure 1 | Cells lacking AID are initially hyper-responsive to transcription-factor-based reprogramming. a, $Aid^{+/+}$ and $Aid^{-/-}$ fibroblasts after 4 days of OKSM transduction. Note the more rounded appearance of null cells. Original magnification, ×10. **b**, Cells positive for SSEA1 after 4 days of OKSM transduction, determined by flow cytometry. **c**, Immunostaining with anti-OCT4 and anti-NANOG antibodies after 1 week of OKSM transduction. Original magnification, ×40. **d**, Relative transcript levels of pluripotency genes determined by qPCR after 4 days of OKSM transduction. For each gene, transcript levels were normalized to $Aid^{+/+}$ cells set to a value of 1. Data represent the mean ± standard error of the mean (s.e.m.) from three independent experiments (*$P < 0.05$, **$P < 0.01$).

[1]Department of Surgery, Weill Cornell Medical College, New York, New York 10065, USA. [2]Immunology Program, Sloan-Kettering Institute, Memorial Sloan Kettering Cancer Institute, New York, New York 10065, USA. [3]Developmental Biology Program, Sloan-Kettering Institute, Memorial Sloan Kettering Cancer Institute, New York, New York 10065, USA. [4]Genentech, South San Francisco, California 94080, USA. [5]Institute for Computational Biomedicine, Weill Cornell Medical College, New York, New York 10065, USA.

that passive removal of epigenetic marks through DNA replication can normalize the initial response to OKSM.

Although there were more cells, no obvious differences were observed at 2 weeks in the morphologies of iPSC-like colonies derived from $Aid^{-/-}$ cells compared with $Aid^{+/+}$ cells, and both stained positive for the pluripotency marker NANOG (Supplementary Fig. 4). However, by 4 weeks the colonies derived from $Aid^{-/-}$ fibroblasts were flattened with less defined edges (Fig. 2a). We tracked individual iPSC-like colonies and observed many $Aid^{-/-}$ colonies that appeared pluripotent at 3 weeks but showed a differentiated morphology at 4 weeks (Fig. 2b). At 3 weeks, the $Aid^{-/-}$ colonies showed a 'patchy' NANOG pattern, and by 4 weeks most colonies were differentiated, with few NANOG$^+$ cells (Fig. 2c). This highly reproducible phenomenon was never observed for colonies derived from $Aid^{+/+}$ cells, which retained their iPSC morphology and NANOG expression throughout the 4 weeks (Fig. 2b, c). The colonies derived from $Aid^{-/-}$ cells showed a progressive decline in the frequency of cells expressing SSEA1 (Fig. 2d) and OCT4 (Fig. 2e). The same results were observed using mouse embryonic fibroblasts (MEFs), as the *Aid*-null cells failed to maintain a pluripotent iPSC phenotype (Supplementary Fig. 5). The stabilization of pluripotency was equally effective in wild-type cells (and ineffective in *Aid*-null cells) regardless of whether constitutively expressed or doxycycline-inducible OSKM cassettes were used (Supplementary Fig. 6a).

Using either wild-type or *Aid*-null cells, the expression of OKSM using a doxycycline-inducible vector for only 6 days was insufficient to generate pluripotent colonies. OKSM expression for 9, 12, 21 or 28 days was sufficient to generate pluripotent colonies, showing a 'dose response' in the wild-type fibroblasts, but was equally ineffective for generating stable pluripotent cells from *Aid*-null cells at the 4-week time point (Supplementary Fig. 6b, c). When AID was expressed along with OKSM during initial reprogramming stages, this failed to rescue pluripotency (data not shown). However, when OKSM-transduced *Aid*-null cells were secondarily infected with an AID-expressing retrovirus after 1 week (Supplementary Fig. 7a, b), there was a partial but statistically significant rescue of pluripotency markers stabilized at the 4-week time point. Because the rescue was not complete we cannot rule out additional molecular mechanisms, beyond AID, that also contribute to the stabilization of pluripotency. Expression of AID in wild-type cells further enhanced the stable expression of pluripotency markers at 4 weeks (Supplementary Fig. 7c). As *Nanog* is syntenic with the *Aid* locus, an independent set of experiments used fibroblasts derived from an *Aid*-knockout strain in which the *neoR* gene was deleted[15]. These 'clean' knockout cells also failed to derive stable iPSC colonies (data not shown).

It was previously reported that iPSCs derived from different cell types retain an epigenetic memory of their somatic phenotype[10,11]. We tested whether genome replication through passaging could stabilize a pluripotent phenotype even in the absence of AID. We generated iPSC-like colonies and isolated 13 $Aid^{+/+}$ and 12 $Aid^{-/-}$ colonies (clones) at 3 weeks that appeared pluripotent by morphology. After three passages (p3) all the clones, regardless of genotype, retained iPSC-like morphology and stained positive for OCT4 and NANOG (Fig. 3a). However, between p7 and p10, 5 out of 12 $Aid^{-/-}$ clones failed to retain pluripotency and differentiated (Fig. 3b). All 13 $Aid^{+/+}$ colonies and the other 7 $Aid^{-/-}$ clones that were stable beyond p10 retained pluripotency until at least p50. All iPSC clones that retained morphology beyond p10 formed embryoid bodies equivalently and could differentiate into cells expressing smooth muscle actin (mesoderm), βIII tubulin (ectoderm) and α-fetoprotein (endoderm), regardless of *Aid* genotype (Fig. 3c). In an independent set of experiments, wild-type and *Aid*-null clones were picked and either maintained in culture as colonies or passaged. The non-passaged wild-type-derived clones maintained a pluripotent morphology, whereas the *Aid*-null cells differentiated. With passaging, 60% of the *Aid*-null clones were able to maintain a pluripotent morphology (Supplementary Fig. 8).

Thus, although AID is not essential for reprogramming, an important transition occurs at around 3 weeks that is assisted by AID to stabilize the pluripotent phenotype. We considered whether the previously described DNA-demethylating role of AID might be at least partially responsible for this phenotypic stabilization. We profiled the epigenome by carrying out reduced-representation bisulphite sequencing (RRBS) in *Aid*-null and control cells after 3 weeks of

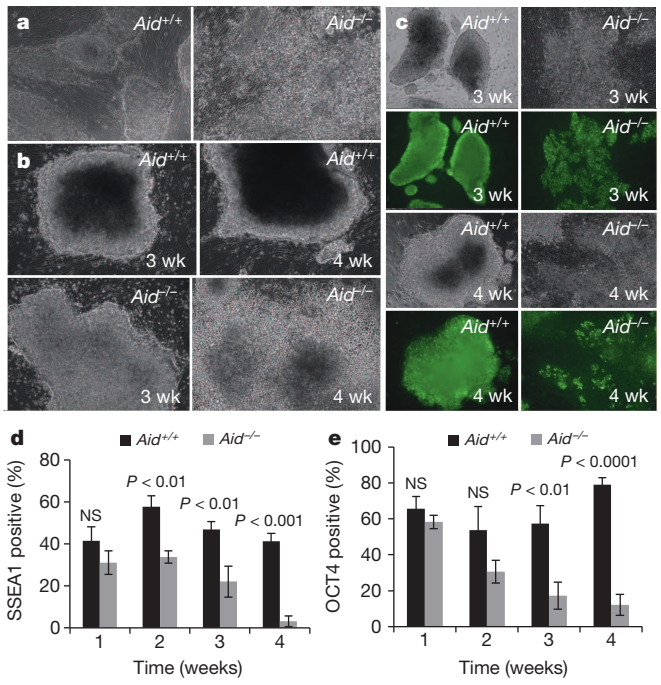


Figure 2 | AID stabilizes pluripotency. **a**, $Aid^{-/-}$ cells lose ES-cell-like morphological characteristics 4 weeks after OKSM transduction. **b**, $Aid^{-/-}$ iPSC-like colonies progressively lose this phenotype. wk, weeks. **c**, Mutant cells lose NANOG expression. **a**–**c**, Original magnification, ×10. **d**, **e**, Cells that stained positive for SSEA1 (**d**) or OCT4 (**e**) were measured by flow cytometry after 1, 2, 3 or 4 weeks of OKSM transduction; $n = 3$ independent experiments, error bars denote standard deviation. NS, not significant.

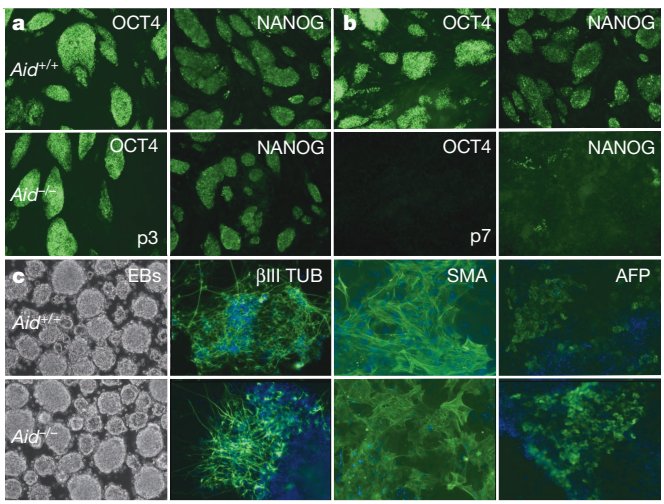


Figure 3 | Cells lacking AID reprogram inefficiently, but those that do are pluripotent. **a**, **b**, Immunostaining of iPSC colonies isolated 3 weeks after OKSM transduction and passaged three times (**a**) or seven times (**b**) before staining. In this example, the null cells failed to from stable iPSCs. **c**, Stable iPSC clones were passaged more than ten times. In this case (unlike in **b**) the passaged *Aid*-null cells maintained NANOG expression (data not shown) and generate embryoid bodies (EBs) that differentiate into cells positive for smooth muscle actin (SMA), βIII tubulin (βIII TUB) and α-fetoprotein (AFP). Original magnification, ×10.

reprogramming. We observed global DNA hypermethylation in the *Aid*-null cells (Supplementary Fig. 9a), when we considered either total CpGs or differentially methylated regions (DMRs). DNA hypermethylation occurs preferentially near RGYW motifs (Supplementary Fig. 9b), the characteristic DNA-targeting sites for AID (45% of the hypermethylated CpGs). Notably, as AID targets gene bodies, the hypermethylated regions are enriched in gene bodies, even though RRBS is biased towards capturing CpG-rich promoter and enhancer regions (Supplementary Fig. 9c).

Transcript profiles comparing the original, pre-transduced fibroblasts from $Aid^{+/+}$ and $Aid^{-/-}$ embryos closely match (Supplementary Fig. 10a). By contrast, many genes fail to be upregulated in the mutant cells during reprogramming (Fig. 4a, b and Supplementary Fig. 10b). The genes that fail to be upregulated during reprogramming are highly enriched ($P < 1 \times 10^{-10}$) in the gene set showing hypermethylation in *Aid*-null cells (Fig. 4c). Of note, genes that are upregulated during reprogramming in both wild-type and *Aid*-null cells are also enriched in hypermethylated genes ($P < 1 \times 10^{-14}$), suggesting that many genes overcome the loss of AID during reprogramming. Among the set of hypermethylated genes, in *Aid*-null cells there is a failure to upregulate a set of secondary pluripotency genes, including *Rex1* (also known as *Zfp42*), *Gdf3*, *Dnmt3l*, *Cbx7*, *Zfp296*, *Dnmt1*, *Apobec1* and *Tet3* (Fig. 4d). This expression data correlates well with the RBBS data, which was validated by MassArray bisulphite sequencing (Supplementary Fig. 11), and is consistent with a failure of this downstream network to be demethylated at late stages of reprogramming. Comparing the gene set of hypermethylated underexpressed genes to public embryonic stem (ES)-cell chromatin immunoprecipitation and massively parallel DNA sequencing (ChIP-seq) data sets revealed a marked enrichment of cMYC target genes in reprogramming cells lacking AID (Supplementary Fig. 12). According to qPCR data there was no difference in *cMyc* transcript levels in wild-type or *Aid*-null cells at any point in the reprogramming process (data not shown). Furthermore, enhanced levels of cMYC, achieved by retroviral transduction starting at 1 week of reprogramming, failed to stabilize pluripotency in *Aid*-null cells (Supplementary Fig. 13), suggesting that cMYC access to key target pluripotency genes (due to hypermethylation) rather than cMYC levels *per se* may be a limiting step for stabilization of the network.

Individual iPSC-like clones were isolated at 2 weeks and expanded to evaluate global methylation patterns for each individual clone. Hypermethylation patterns in the *Aid*-null clones were highly consistent with data obtained from bulk colony analysis. On the basis of genome-wide DNA methylation profiles (with the genome binned for 100 kb regions), *Aid*-null or wild-type clones cluster according to genotype (Supplementary Fig. 14a). Regions that were hypermethylated in the *Aid*-null bulk analysis tended to be hypermethylated in the *Aid*-null clones (Supplementary Fig. 14b). In fact, 66% of bulk hypermethylated DMRs were also hypermethylated in DNA derived from the isolated clones ($P < 1 \times 10^{-151}$). Methylation differences found previously for the secondary pluripotency genes were largely validated by MassArray bisulphite sequencing of DNA (Supplementary Fig. 14c), and most of the hypermethylated genes were expressed in the *Aid*-null clones at significantly lower levels compared with levels in wild-type clones (Supplementary Fig. 14d).

Altogether, our results show that active demethylation through AID-dependent cytosine deamination is important for stabilizing a genetic network controlling stem-cell phenotype. AID is not essential for reprogramming, because cells lacking AID can form stable pluripotent iPSCs, either through a passive demethylation process facilitated by DNA replication, or perhaps through compensation from related members of the APOBEC family[16]. A recent study suggested that loss of AID affects transcription-factor-induced reprogramming only by an early acute short hairpin RNA (shRNA)-dependent depletion[17], but did not report the 'iPSC' phenotype beyond 3 weeks, when we show that *Aid*-null cells eventually fail to stabilize as pluripotent. Unless passaged, *Aid*-null cells always fail to reprogram. RNA-seq profiles of *Aid*-null cells after 4 weeks of reprogramming cluster with fibroblast samples (data not shown), and qPCR analysis validates significant upregulation of fibroblast-associated genes in the *Aid*-null samples (Supplementary Fig. 15), consistent with an epigenetic memory for fibroblast fate that fails to be fully removed during reprogramming in

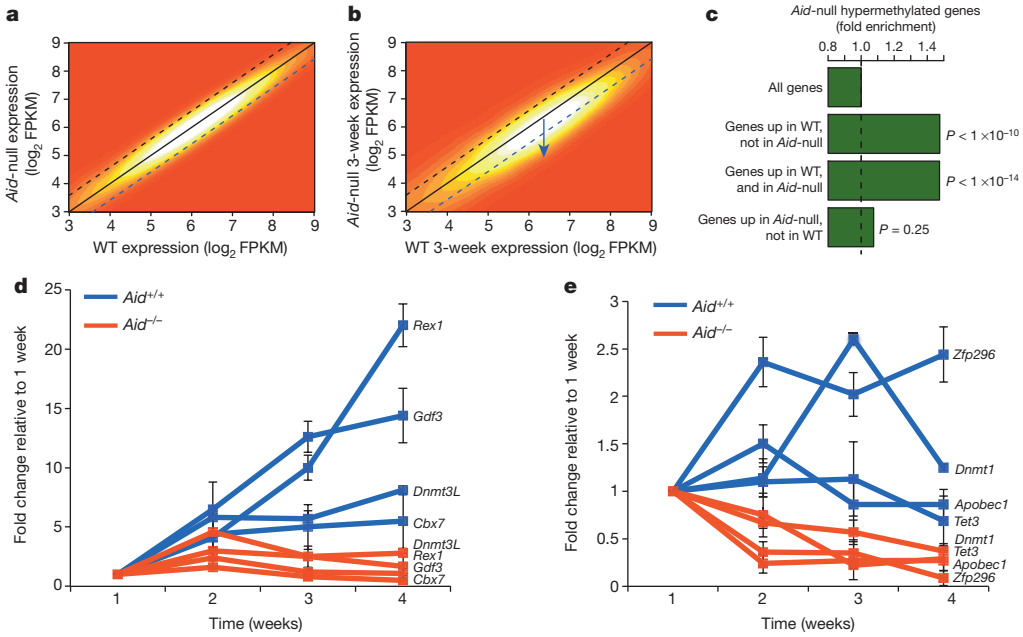


Figure 4 | Cells lacking AID fail to activate expression of hypermethylated pluripotency genes. **a**, In the original, pre-transduced fibroblasts few differentially expressed genes are off the diagonal. FPKM, fragments per kilobase of transcript per million fragments mapped; WT, wild type. **b**, Many genes are underexpressed in *Aid*-null cells 3 weeks after transduction. **c**, Fold enrichment for hypermethylated genes in *Aid*-null cells. All *P* values were calculated using the hypergeometric distribution. **d**, Transcript levels for hypermethylated pluripotency genes during reprogramming, comparing wild-type cells (blue lines) with *Aid*-null cells (red lines). Data represent the mean ± s.e.m. from three independent experiments.

the absence of AID. The function of AID for transition to a stable pluripotent phenotype may be relevant during embryogenesis, as *Aid*-knockout strains have small litters. We found that ES-cell lines could be established from *Aid*-null blastocysts, and these appeared by morphology and staining patterns to be normal (Supplementary Fig. 16). However, the efficiency of ES-cell-line derivation was markedly reduced (17/28 embryos, 60.7%) compared with congenic wild-type embryos (21/24, 87.5%; $P < 0.03$ by chi-squared analysis).

Although there is much promise for the use of iPSCs for disease modelling and cellular therapies[18], there remains concern about whether iPSC genomes are damaged through the process of reprogramming[19–21]. AID, which has natural mutator activity, is activated during reprogramming, and mediates demethylation in gene bodies including secondary pluripotency genes that encode proto-oncogenes. Although inefficient, we showed that iPSCs could be generated in the absence of AID, removing epigenetic memory marks through an AID-independent mechanism. Retaining epigenetic memory, in the absence of AID, might be useful for promoting efficiency of differentiation towards parental lineage fate. Furthermore, if AID-independent reprogramming lessens the mutation load, it could provide a safer strategy for the generation of iPSCs for cellular therapies.

METHODS SUMMARY

TTFs were prepared using 0.2% collagenase; MEFs were prepared from embryonic day (E)13.5 embryos. Intracellular immunostaining was performed after fixing the cells with 4% paraformaldehyde. Surface staining was done without fixing the cells. Complementary DNA was prepared using Superscript III and qPCR was performed (primers listed in Supplementary Table 1) using SYBR green and a Roche LightCycler 480 II. All data were from at least three independent experiments. Enhanced RRBS was performed as described[22] and bisulphite reads were aligned to the bisulphite-converted mm9 genome using Bismark[23]. We achieved very high coverage by sequencing one full Illumina lane (~200 M, 51 bp reads) per sample. For bulk colonies, the wild-type and *Aid*-null samples (at 3 weeks) were analysed with RRBS that covered 1.96 M and 1.82 M CpGs, respectively, with at least 10× coverage (with >1.62 M CpGs in common between the two runs). 1.52 M and 1.59 M CpGs had ≥20× coverage in wild-type and $Aid^{-/-}$ cells, respectively. 1.17 M and 1.32 M CpGs had ≥50× coverage, respectively. Differentially methylated CpGs were identified using the Fisher exact test with Benjamini–Hochberg correction for multiple testing. DMRs were defined as containing at least five differentially methylated CpGs and a total methylation difference of more than 10%. Paired-end RNA-seq libraries were constructed as described[24] and sequenced using an Illumina HiSeq2000. Reads were aligned to mm9 using TopHat and gene expression levels were quantified using Cufflinks, using upper-quartile and GC-content normalizations. Fold changes of 1.5 with fragments per kilobase of transcript per million fragments mapped (FPKM) >5 in at least one condition were used to derive differentially expressed genes. MassArray EpiTYPER analysis was performed on selected regions identified as hypermethylated by RRBS. PCR primers (Supplementary Table 2) were designed to probe amplicons using EpiDesigner (http://www.epidesigner.com/). Bisulphite conversion and MassArray analysis were performed as previously described[22]. Metadata for all RRBS experiments are given in Supplementary Table 3. Methylation differences were calculated at each CpG and boxplot analysis was performed at each interrogated region. All statistical analyses were performed using the R software package.

Full Methods and any associated references are available in the online version of the paper.

Received 14 August 2012; accepted 13 May 2013.
Published online 26 June 2013.

1. Muramatsu, M. *et al.* Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102,** 553–563 (2000).
2. Chaudhuri, J. *et al.* Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature* **422,** 726–730 (2003).
3. Bhutani, N., Burns, D. M. & Blau, H. M. DNA demethylation dynamics. *Cell* **146,** 866–872 (2011).
4. Popp, C. *et al.* Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* **463,** 1101–1105 (2010).
5. Bhutani, N. *et al.* Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* **463,** 1042–1047 (2010).
6. Foshay, K. M. *et al.* Embryonic stem cells induce pluripotency in somatic cell fusion through biphasic reprogramming. *Mol. Cell* **46,** 159–170 (2012).
7. Polo, J. M. *et al.* A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell* **151,** 1617–1632 (2012).
8. Buganim, Y. *et al.* Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* **150,** 1209–1222 (2012).
9. Meissner, A. Epigenetic modifications in pluripotent and differentiated cells. *Nature Biotechnol.* **28,** 1079–1088 (2010).
10. Kim, K. *et al.* Epigenetic memory in induced pluripotent stem cells. *Nature* **467,** 285–290 (2010).
11. Polo, J. M. *et al.* Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nature Biotechnol.* **28,** 848–855 (2010).
12. Sommer, C. A. *et al.* Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. *Stem Cells* **27,** 543–549 (2009).
13. Smith, Z. D., Nachman, I., Regev, A. & Meissner, A. Dynamic single-cell imaging of direct reprogramming reveals an early specifying event. *Nature Biotechnol.* **28,** 521–526 (2010).
14. Brambrink, T. *et al.* Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell* **2,** 151–159 (2008).
15. Sun, Y. *et al.* Critical role of activation induced cytidine deaminase in experimental autoimmune encephalomyelitis. *Autoimmunity* **46,** 157–167 (2013).
16. Conticello, S. G. The AID/APOBEC family of nucleic acid mutators. *Genome Biol.* **9,** 229 (2008).
17. Bhutani, N. *et al.* A critical role for AID in the initiation of reprogramming to induced pluripotent stem cells. *FASEB J.* **27,** 1107–1113 (2012).
18. Robinton, D. A. & Daley, G. Q. The promise of induced pluripotent stem cells in research and therapy. *Nature* **481,** 295–305 (2012).
19. Gore, A. *et al.* Somatic coding mutations in human induced pluripotent stem cells. *Nature* **471,** 63–67 (2011).
20. Hussein, S. M. *et al.* Copy number variation and selection during reprogramming to pluripotency. *Nature* **471,** 58–62 (2011).
21. Mayshar, Y. *et al.* Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. *Cell Stem Cell* **7,** 521–531 (2010).
22. Akalin, A. *et al.* Base-pair resolution DNA methylation sequencing reveals profoundly divergent epigenetic landscapes in acute myeloid leukemia. *PLoS Genet.* **8,** e1002781 (2012).
23. Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27,** 1571–1572 (2011).
24. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* **5,** 621–628 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank G. Mostoslavsky for the gift of reprogramming vectors. A. Melnick provided advice and consultation. We are grateful to T. Honjo for AID-knockout mice, and N. Papavasiliou for assistance in designing genotyping PCR primers. The Epigenomics Core Facility of Weill Cornell Medical College carried out the bisulphite sequencing and provided consultation. This study was supported by National Institutes of Health grant HL056182 (T.E.), AI072194 (J.C.) and National Science Foundation CAREER grant 1054964 (O.E.).

Author Contributions R.K. conceived the study, carried out experiments, and wrote the manuscript. L.D., N.S., T.-C.L., P.F. and S.M.-D. carried out experiments. A.A.Z. and A.-K.H. provided essential reagents and expertise. J.C. conceived the study and wrote the manuscript. O.E. conceived the study, carried out computational and informatics analyses, and wrote the manuscript. T.E. conceived the study and wrote the manuscript.

Author Information All RNA-seq and RRBS data have been deposited in the Gene Expression Omnibus under accession GSE46700. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.E. (tre2003@med.cornell.edu).

METHODS

Summary. Intracellular immunostaining was performed after fixing the cells with 4% paraformaldehyde. Surface staining was done without fixing the cells. All data were from at least three independent experiments. Enhanced RRBS was performed as described[22] and bisulphite reads were aligned to the bisulphite-converted mm9 genome using Bismark[23]. We achieved very high coverage by sequencing one full Illumina lane (~200 M, 51 bp reads) per sample. For bulk colonies, the wild-type and *Aid*-null samples (at 3 weeks) were analysed with RRBS that covered 1.96 M and 1.82 M CpGs, respectively, with at least 10× coverage (with >1.62 M CpGs in common between the two runs). 1.52 M and 1.59 M CpGs had ≥20× coverage in wild-type and $Aid^{-/-}$ cells, respectively. 1.17 M and 1.32 M CpGs had ≥50× coverage, respectively. Differentially methylated CpGs were identified using the Fisher exact test with Benjamini–Hochberg correction for multiple testing. DMRs were defined as containing at least five differentially methylated CpGs and a total methylation difference of more than 10%. Paired-end RNA-seq libraries were constructed as described[24] and sequenced using an Illumina HiSeq2000. MassArray EpiTYPER analysis was performed on selected regions identified as hypermethylated by RRBS. PCR primers (Supplementary Table 2) were designed to probe amplicons using EpiDesigner (http://www.epidesigner.com/). Bisulphite conversion and MassArray analysis were performed as previously described[22]. Metadata for all RRBS experiments are given in Supplementary Table 3. Methylation differences were calculated at each CpG and boxplot analysis was performed at each interrogated region. All statistical analyses were performed using the R software package.

Mice. AID-deficient ($Aid^{-/-}$) mice were a gift from T. Honjo. Some of the reprogramming experiments were performed on the $Aid^{-/-}$ and $Aid^{+/+}$ littermates derived by crossing $Aid^{+/-}$ mice, and for some of the experiments wild-type BALB/c mice were purchased from the Jackson Laboratory. Reprogramming experiments were performed on tail fibroblasts obtained from seven $Aid^{-/-}$ and seven $Aid^{+/+}$ age-matched mice (four littermates and three non-littermates, each). We observed similar results when comparing littermates or the purchased mice. Mice or embryos used for MEFs were genotyped by PCR. All animals were maintained according to the guidelines for animal welfare of the Memorial Sloan-Kettering Research Animal Resource Center.

Preparation of fibroblasts. For preparation of tail fibroblasts either whole tails or tail tips were collected from age-matched adult $Aid^{+/+}$ and $Aid^{-/-}$ mice (age range, 3–8 weeks). Tails or tail tips were minced using a razor blade/scalpel after washing with PBS. For preparation of MEFs, $Aid^{+/-}$ males and $Aid^{+/-}$ females were time mated and E13.5 embryos were collected. Embryos were washed several times and extra care was taken to prevent maternal cross contamination. After removing the head and all internal organs embryos were minced using a razor blade/scalpel. Minced tails or embryos were washed twice with DMEM (Cellgro) containing 10% FBS (Gemini Bio Products), 1 mM L-glutamine (Cellgro), 100 U ml^{-1} penicillin (Cellgro) and 10 μg ml^{-1} streptomycin (Cellgro), and dissociated using 0.2% collagenase from *Clostridium histolyticum* (Sigma) at 37 °C for 4 h (tails) or 2 h (embryos) with continuous shaking. Dissociated cells were filtered through a 70-μm cell strainer (BD falcon) and washed two times with DMEM containing 10% FBS and plated on 0.1% gelatin- (Cellgro) coated tissue culture plates. Fibroblasts were cultured in DMEM containing 10% FBS, 1 mM L-glutamine, 100 U ml^{-1} penicillin and 100 μg ml^{-1} streptomycin and media was replaced daily. After passaging two times, fibroblasts were frozen in media containing 10% DMSO, 40% FBS and 50% DMEM. Generally, p2 fibroblasts were used for the reprogramming experiments, unless indicated otherwise. Genotype was confirmed by PCR. Primers for wild type, F: 5′-GGTCCCAGTCTGAGATGTA; R: 5′-CAACGTGGCGTCCAAACAGGC; for knockout, F: 5′-CTGCCAAACCTGA TGTCTTGA; R: 5′-AACCAAGCCTATGCCTACAGC.

Production of lentiviruses and retroviruses. Reprogramming virus was produced by transfection of a single lentiviral stem cell cassette (STEMCCA) in HEK293T cells using polyethylenimine (PEI) (Polysciences). STEMCCA (constitutive) expresses all four human factors (OCT4, KLF4, SOX2, cMYC), using a single elongation factor 1α (EF1α) promoter within a lentiviral vector using a combination of 2A peptide and internal ribosome entry site (IRES) strategies. Both constitutive and doxycycline-inducible STEMCCA vectors were used for the reprogramming studies. For the doxycycline-inducible system, fibroblasts were treated with doxycycline (1 μg ml^{-1}) starting 1 day after transduction for 2 weeks. Lentiviruses were produced using a five-plasmid transfection system with STEMCCA transfected together with four expression vectors encoding the packaging proteins Gag-Pol, Rev and Tat, and the G protein of the vesicular stomatitis virus (VSV). Five vectors were incubated in DMEM with 0.06% PEI for 20 min and then transferred to the plates containing 80% confluent HEK293T cells. Transfection was performed in FBS (10%) containing media. After 5 h media was replaced with fresh media. Viral supernatant was collected after 48 h and spun at 3,000 r.p.m. to remove dead cells and filtered through a sterile syringe filter with 0.45 μm polyethersulfone membrane (VWR international). Ectopic expression of wild-type AID or catalytically mutant AID was achieved using retroviral vectors as described previously[25]. Retroviruses were produced by co-transfecting the vector with the packaging plasmid pCL-ECO in HEK293T cells. The cMYC overexpression experiments were achieved using retroviral vector pMXs-c-Myc (Addgene, 13375). pMXs-c-Myc was co-transfected with vector encoding VSVg in GP2-293 cells to collect functional virus.

Transduction with lentiviruses. Five-hundred-thousand fibroblasts were plated in one well of a gelatin-coated 6-well tissue culture plate. After 6 h, 1 ml of viral supernatant containing 8 μg ml^{-1} polybrene (Millipore) was added to the fibroblast-containing wells. For every experiment, $Aid^{-/-}$ and $Aid^{+/+}$ fibroblasts were infected at the same time with an identical titre of virus, prepared for each experiment from the same batch. Viral supernatant was removed after 12 h of infection and fresh media containing DMEM, 10% FBS, 1 mM L-glutamine, 100 U ml^{-1} penicillin and 10 μg ml^{-1} streptomycin was added. After 1 day media was replaced by mouse embryonic stem-cell (MES) media containing DMEM, 20% ES-cell-compatible FBS (Gemini Bio Products), LIF (2% conditioned medium) and 1.5×10^{-4} M monothioglycerol (MTG; Sigma), 1 mM L-glutamine, 100 U ml^{-1} penicillin and 10 μg ml^{-1} streptomycin. Cells were maintained in the same media. In some of the experiments cells were transferred to mitotically inactivated MEF feeders after 2 days of transduction. Approximately 30,000–50,000 cells were transferred to one well of a feeder-containing 6-well plate. Isolated iPSC colonies were always cultured on feeders. All the cells were kept at 37 °C in a humidified environment at 5% CO_2.

Immunostaining. Immunostaining was performed on fixed cells (4% PFA in BBS with 1 mM $CaCl_2$, 15 min) washed and blocked for 30 min in BBT-BSA buffer (BBS with 0.5% BSA, 0.1% Triton and 1 mM $CaCl_2$). Cells with primary antibodies were incubated overnight at 4 °C at the following dilutions: anti-NANOG (eBiosciences 14-5761, 1:100, 5 μg ml^{-1}) and anti-OCT4 (Santa Cruz 5279, 1:100, 2 μg ml^{-1}). Cells were washed and blocked in BBT-BSA and then incubated with Alexa-conjugated secondary antibodies (1:500, from Molecular Probes). Vectashield-DAPI was used as a mounting medium. Images were acquired using a Zeiss LSM 510-Meta confocal microscope or a Zeiss epifluorescence microscope with AxioVision software. For flow cytometric analysis cells were trypsinized, fixed with 4% PFA for 20 min, blocked for 1 h and then stained in suspension. SSEA1 (Santa Cruz, 1:100) staining was performed on unfixed cells. Cells were analysed on a BD-Accuri C6 flow cytometer (BD Biosciences) using CFlow Plus software.

RT–qPCR. Cells were trypsinized and collected in Trizol reagent (Life technologies). Total RNA extraction was done using the RNeasy Mini Kit (Qiagen). The cDNA was synthesized using the SuperScript III First-Strand Synthesis System (Invitrogen). RT–qPCR was performed on three experimental replicates using SYBR green Master I (Roche). Data were generated on a LightCycler 480 II (Roche) and analysed using LightCycler 480 software. qPCR data was calculated based on the mean of three experimental replicates. All quantifications were normalized to an endogenous *Gapdh* control. The relative quantification value for each target gene was compared to the calibrator for that target gene. All the primers used for qPCR spanned introns (Supplementary Table 1).

***In vitro* differentiation.** For embryoid body formation, iPSCs were passaged two times on feeder-free gelatin-coated culture dishes. 5×10^5 iPSCs were plated on a 10-cm low attachment dish. Embryoid bodies were cultured for 6 days in MES media without LIF, and media was changed every other day. For mesodermal differentiation, embryoid bodies were replated on gelatin-coated dishes and cultured for another 7 days in the same media. For neural differentiation, embryoid bodies were cultured in media containing DMEM F-12 (Cellgro), 0.5% N2 (Gibco) and 0.5% B27 (Gibco) supplements, 1 mM L-glutamine, 1% non-essential amino acids (Gibco) and 1.5×10^{-4} M monothioglycerol for another 3 days, after which embryoid bodies were seeded onto gelatin-coated dishes with the same media, plus 10 μM retinoic acid (Sigma) for another 4 days. For endoderm differentiation, embryoid bodies were replated onto gelatin-coated dishes in MES media lacking LIF but containing 0.5% FBS and 50 ng ml^{-1} Activin (R&D) for 7 days. For all differentiation cultures, media was changed daily.

ES-cell derivation from blastcyst-stage mouse embryos. $Aid^{-/-}$ males and $Aid^{-/-}$ females (or congenic wild-type pairs as controls) were time mated and E3.5 blastocyst-staged embryos were collected in M2 media according to standard protocols[26]. Embryos were placed on mitotically inactivated MEFs[27] in knockout DMEM (Gibco) containing 15% knockout serum replacement (KSR; Gibco), 2 mM L-glutamine, 0.1 mM β-mercaptoethanol, 0.1 mM non-essential amino acids, 1 mM sodium pyruvate, 100 U ml^{-1} penicillin and 100 μg ml^{-1} streptomycin, 1,000 U ml^{-1} LIF, 1 μM PD0325901 (ERK inhibitor) and 3 μM CHIR99021 (inhibitor of GSK3B). After embryos attached, media was replaced every other day. After 10–11 days an outgrowth could be observed and was dissociated in 0.25% trypsin/EDTA using a mouth-controlled drawn glass pipette. Trypsin was inactivated by adding knockout

DMEM (Gibco) containing 15% ES-cell-compatible FBS, 2 mM L-glutamine, 0.1 mM β-mercaptoethanol, 0.1 mM non-essential amino acids, 1 mM sodium pyruvate, 100 U ml^{-1} penicillin and 100 μg ml^{-1} streptomycin, 1,000 U ml^{-1} LIF, 1 μM PD0325901 and 3 μM CHIR99021. From this point onwards the culture was maintained in the same media on mitotically inactivated MEFs. Media was replaced every other day. After 3–5 days ES-cell-like colonies were observed and passaged at 70% confluency. Stable ES-cell lines were validated by passaging a minimum of ten times without loss of pluripotent morphology or marker expression.

Genomic analysis. RRBS was performed as described[22] and bisulphite reads were aligned to the bisulphite-converted mm9 using Bismark[23]. Differentially methylated CpGs were identified using the Fisher exact test with Benjamini–Hochberg correction for multiple testing. We defined DMRs as regions containing at least five differentially methylated CpGs, where contiguous differentially methylated CpGs are separated by 250 bp or less, and for which the total methylation change between wild-type and *Aid*-null cells is 10% or more (calculated using all CpGs within the considered region including those that were not called as differentially methylated). The distribution of DMR lengths was as follows. For hypermethylated DMRs, average, 222 bp; median, 190 bp; minimum, 10 bp; maximum, 4,081 bp. For the lower number of hypomethylated DMRs, we observed the following statistics: average, 229 bp; median, 188 bp; minimum, 19 bp; maximum, 1,250 bp. Paired-end RNA-seq libraries were constructed as previously described[24] and sequenced using an Illumina HiSeq2000. Reads were aligned to mm9 using TopHat and gene expression levels were quantified using Cufflinks, using upper-quartile and GC-content normalizations. Twofold changes with FPKM $>$ 5 in at least one condition were used to derive differentially expressed genes. All statistical analyses were performed using the R software package. The meta-data for all RBBS data are provided in Supplementary Table 3.

Quantitative DNA methylation analysis by mass spectrometry. The level of DNA methylation for specific genes was measured using a MALDI-TOF mass spectrometry based method (EpiTYPER; Sequenom) as previously described[28]. Briefly, 1 μg of DNA was treated with sodium bisulphite using the EZ methylation kit (Zymo-Research). The treatment converts non-methylated cytosines into uracil, leaving methylated cytosines unchanged. PCR amplification, addition of SAP solution and Transcription/RNase A cocktails were performed according to the protocol provided by Sequenom and the mass spectra were quantified by the EpiTYPER analyser. Amplicons probed are given in Supplementary Table 2.

25. Vuong, B. Q. *et al.* Specific recruitment of protein kinase A to the immunoglobulin locus regulates class-switch recombination. *Nature Immunol.* **10,** 420–426 (2009).
26. Nagy, A., Gertsenstein, M., Vintersten, K. & Behringer, R. *Manipulating the Mouse Embryo: A Laboratory Manual* 3rd edn (Cold Spring Harbor, 2003).
27. Joyner, A. L. *Gene Targeting: A Practical Approach* 2nd edn (Oxford Univ. Press, 2000).
28. Ehrich, M. *et al.* Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. *Proc. Natl Acad. Sci. USA* **102,** 15785–15790 (2005).

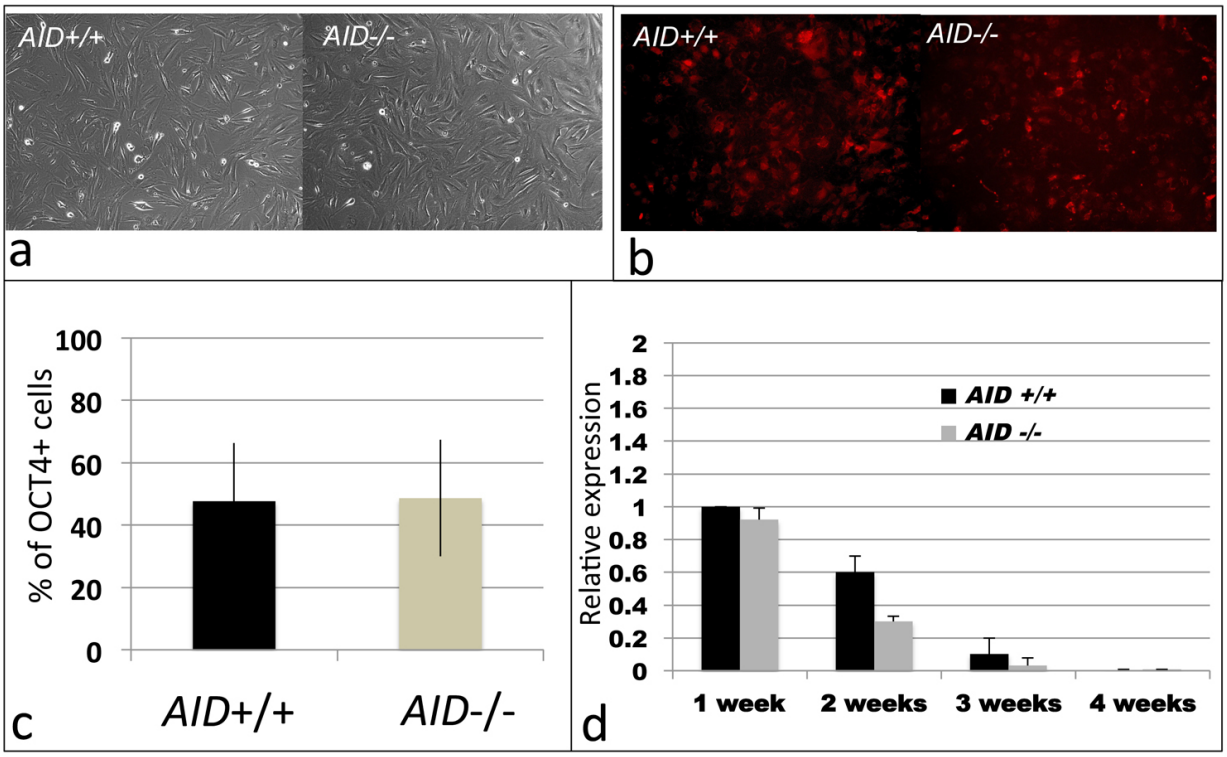
doi:10.1038/nature12299

Abbreviations used in the text

5hmC – 5-hyroxymethycytosine
5mC – 5-methylcytosine
AFP – Alpha-fetoprotein
AID – Activation induced cytidine deaminase
APOBEC1 - Apolipoprotein B mRNA editing enzyme, catalytic polypeptide 1
CBX7 - Chromobox 7
DMR – Differentially methylated regions
DNMT1 - DNA methyltransferase (cytosine-5) 1
DNMT3L - DNA (cytosine-5-)-methyltransferase 3-like
EB – Embryoid bodies
ESC – Embryonic stem cells
FPKM - Fragments per kilobase of exon per million fragments mapped
GC- Germinal Center
GDF3 - Growth differentiation factor 3
iPSC – Induced pluripotent stem cells
KLF4 - Kruppel-like factor 4
MYC- myelocytomatosis oncogene
neoR – Neomycin resistance
OCT4 - Octamer-binding transcription factor 4
RRBS – Reduced representation bisulphite sequencing
SMA – Smooth muscle actin
SOX2 - SRY-box containing gene 2
SSEA1 – stage-specific embryonic antigen 1
TET3 - Tet methylcytosine dioxygenase 3
TF – Transcription factor
TSS – Transcriptional start site
TTF – Tail tip fibroblast
ZFP296 - Zinc finger protein 296
βIII TUB – beta-III tubulin
UTF1 - Undifferentiated embryonic cell transcription factor 1
KLF2 - Kruppel-like factor 2
ERAS - ES cell-expressed Ras

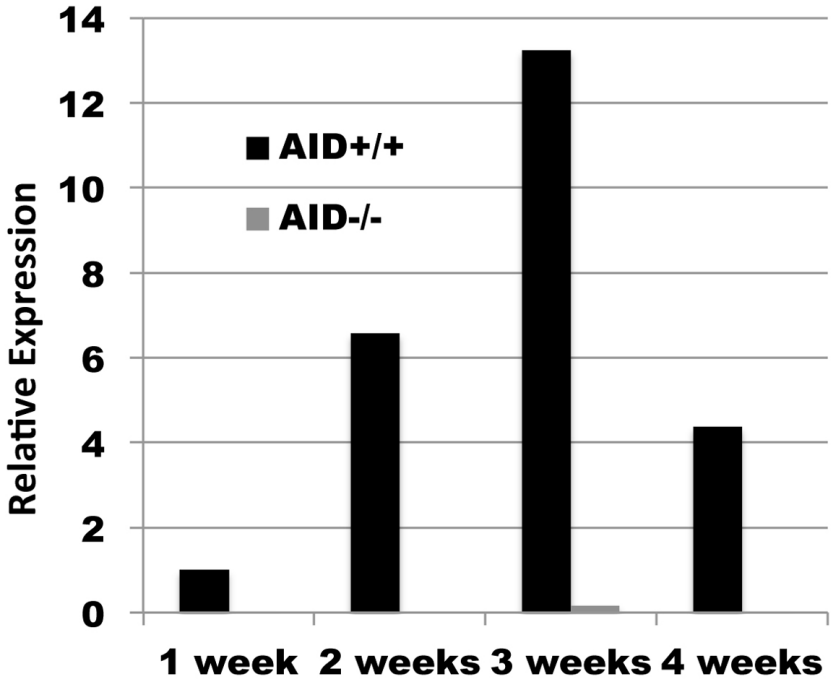
Supplemental Figures

Fig. S1



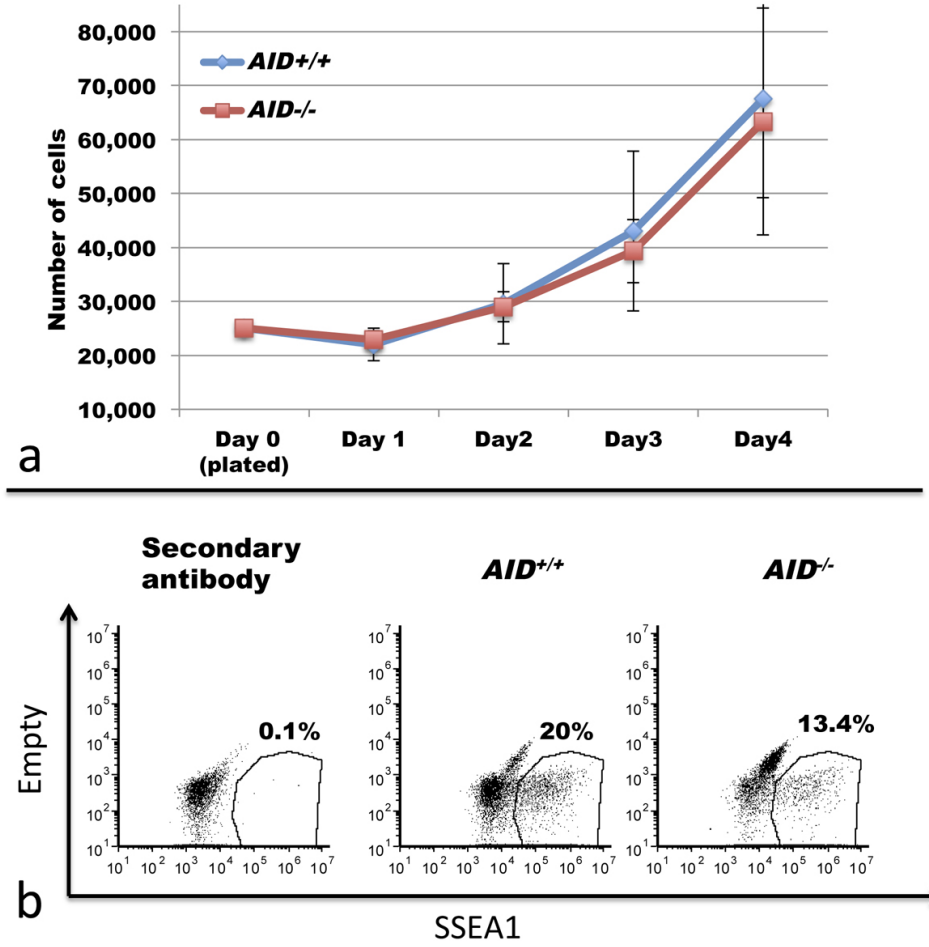
Supplemental Figure S1. AID-null fibroblasts appear normal in culture, are equally transduced by the reprogramming vector, and silence reprogramming factors equivalent to wildtype cells. a, Tail-tip fibroblast cultures derived either from wildtype or knockout AID mice. There are no discernable morphological differences in these cultures. b, Fluorescent images demonstrating equivalent transduction and expression of mCherry from the lentiviral expression vectors used for TF-induced reprogramming. c, Percentage of wildtype or AID-null cells positive for exogenous human OCT4, measured by flow cytometry, three days after transduction with the reprogramming vector. Bars show the average of three independent experiments, and there is no significant difference. This demonstrates that lack of AID does not impact expression vector function. d, Relative expression of exogenous human OCT4 gene transcripts in wildtype and AID-null fibroblasts after 1, 2, 3, or 4 weeks of reprogramming following OSKM transduction. Shown are averages of the mean +/- standard error from three independent experiments, which shows no statistical difference at each time point.

Fig. S2



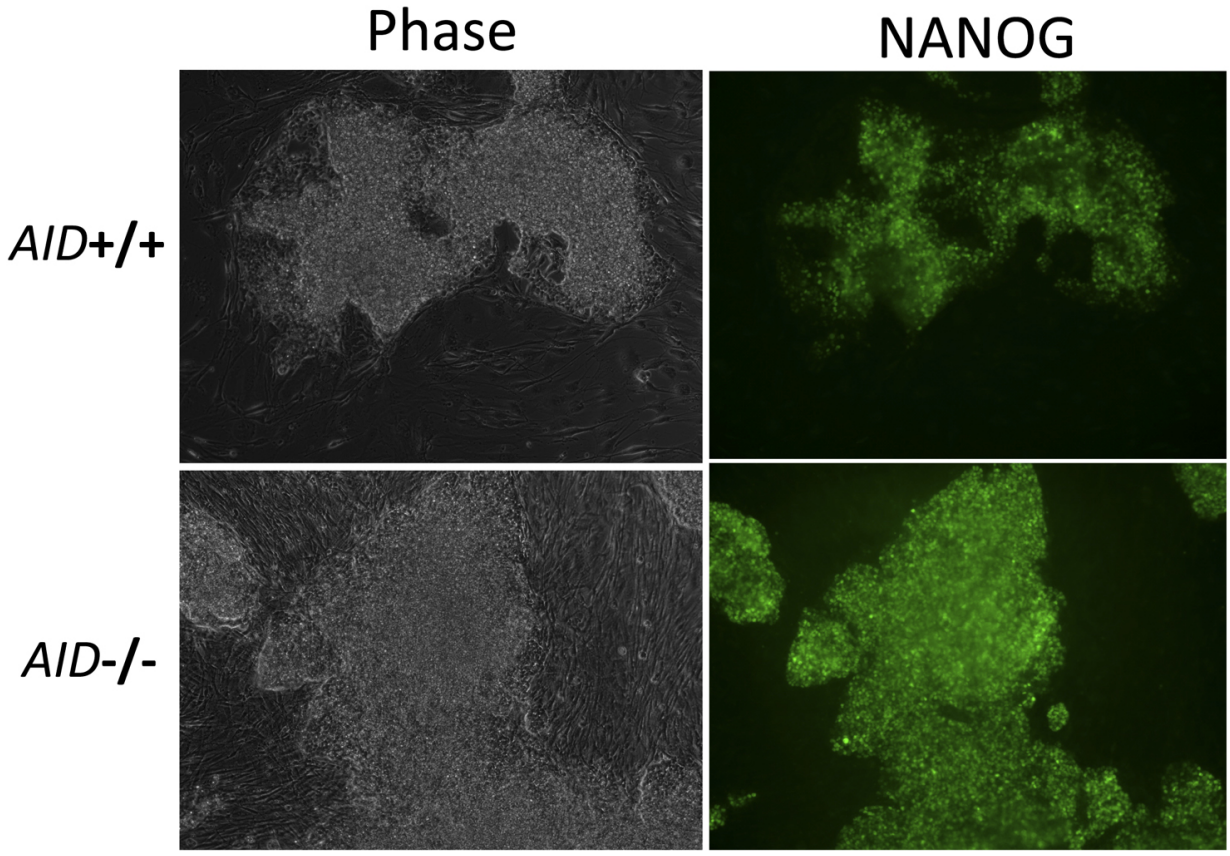
Supplemental Figure S2. AID is up-regulated in wildtype cells during the reprogramming process. Shown are qPCR results from a representative reprogramming experiment evaluating the expression levels for AID at 1, 2, 3, or 4 weeks. The black bars indicate levels of AID transcripts in the wildtype cells transduced with lentivirus, relative to the levels detected at 1 week, given an arbitrary value of 1. AID transcripts were not detected in fibroblasts prior to induction or reliably measured prior to 1 week. Nor were they detectable in knockout cells, as expected.

Fig. S3



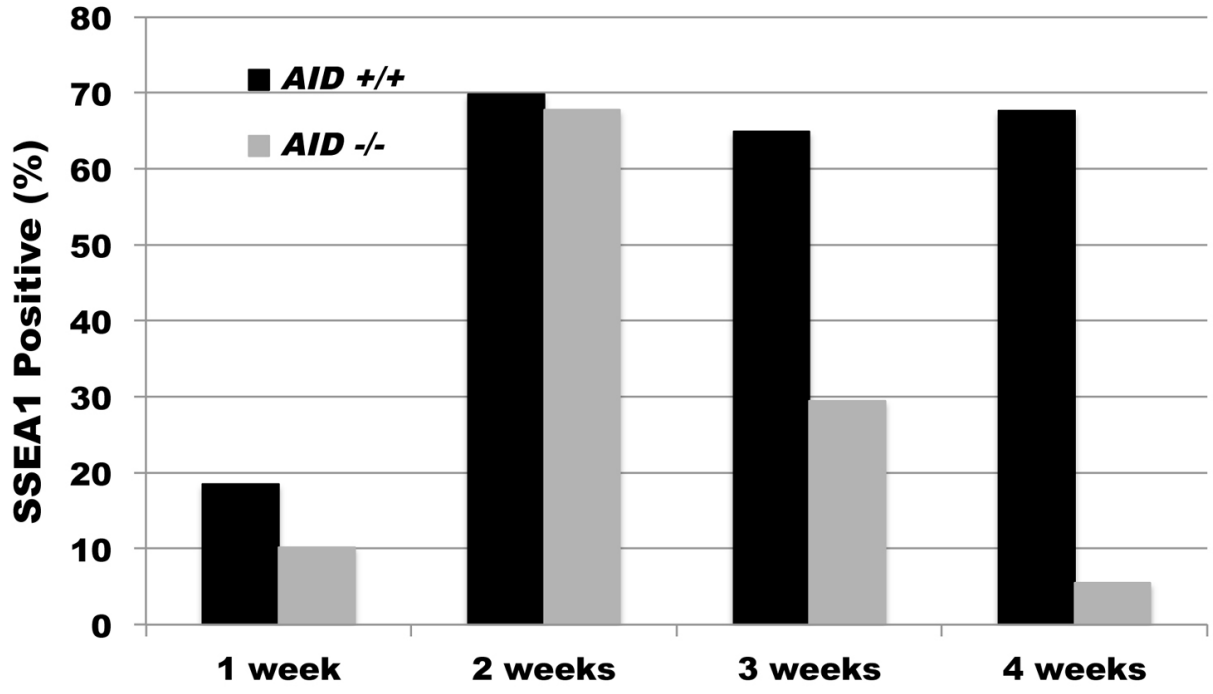
Supplemental Figure S3. Hyper-responsiveness to reprogramming is not due to enhanced proliferation of AID-null fibroblasts and passaging results in loss of this phenotype. a, Growth curves were plotted for AID-null and wildtype fibroblasts as indicated. 25,000 fibroblasts were plated and counted each day for 4 days. Each count occurred 24 hr following the previous count. Data represents mean of three individual experiments +/- standard deviation. There was no statistically significant difference in the growth rate of AID-null and wildtype fibroblasts. Eventually both sets of TTFs undergo senescence. b, Compared to the experiments represented in Fig. 1 using p2 primary fibroblasts, cells were passaged an additional four times prior to transduction. Plotted is the percentage of SSEA1 positive cells four days after transduction of the OKSM vector in $AID^{+/+}$ or $AID^{-/-}$ fibroblasts. The left panel is a staining control.

Fig. S4



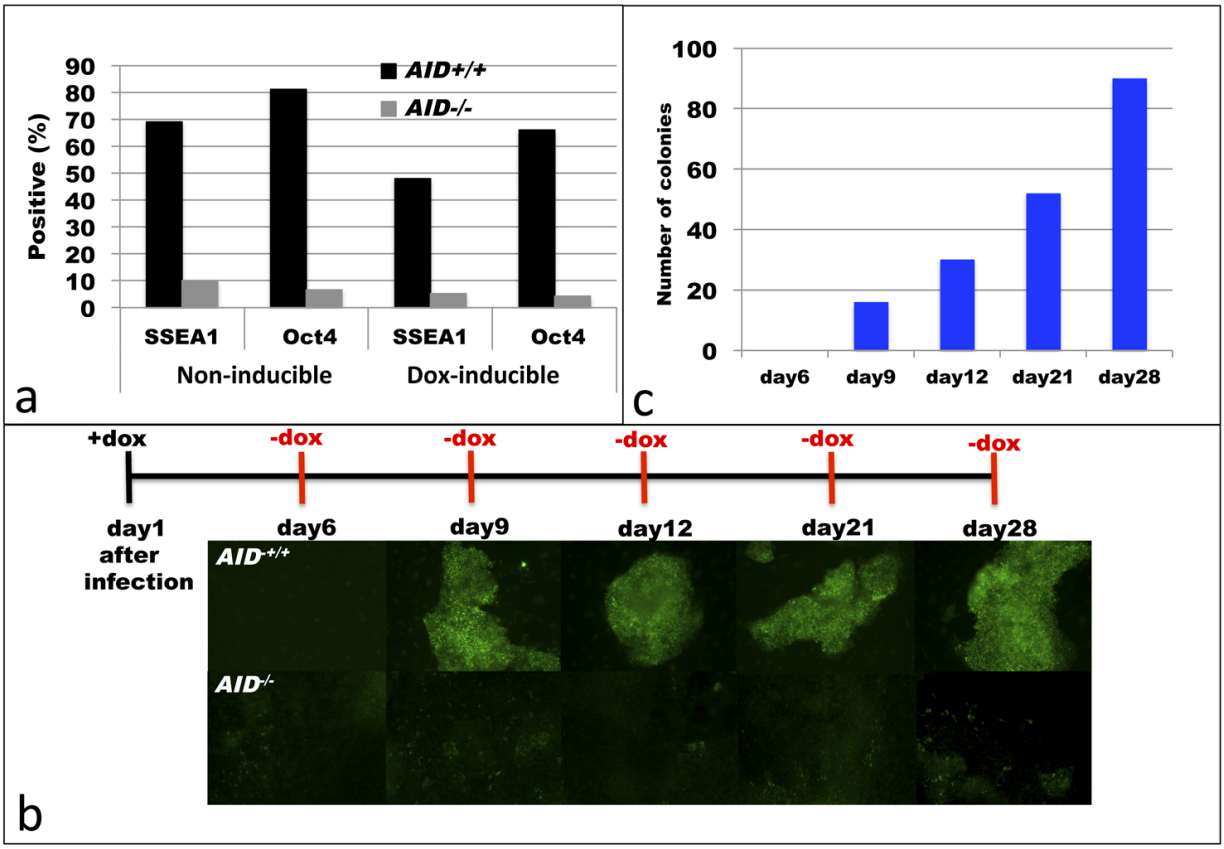
Supplemental Figure S4. AID wildtype and null cells generate NANOG+ iPSC-like colonies at 2 weeks of reprogramming. Shown are representative iPSC-like colonies after 2 weeks of reprogramming. Brightfield phase images are on the left. Both the wildtype and AID-null cells are positive for NANOG at this point in the process.

Fig. S5



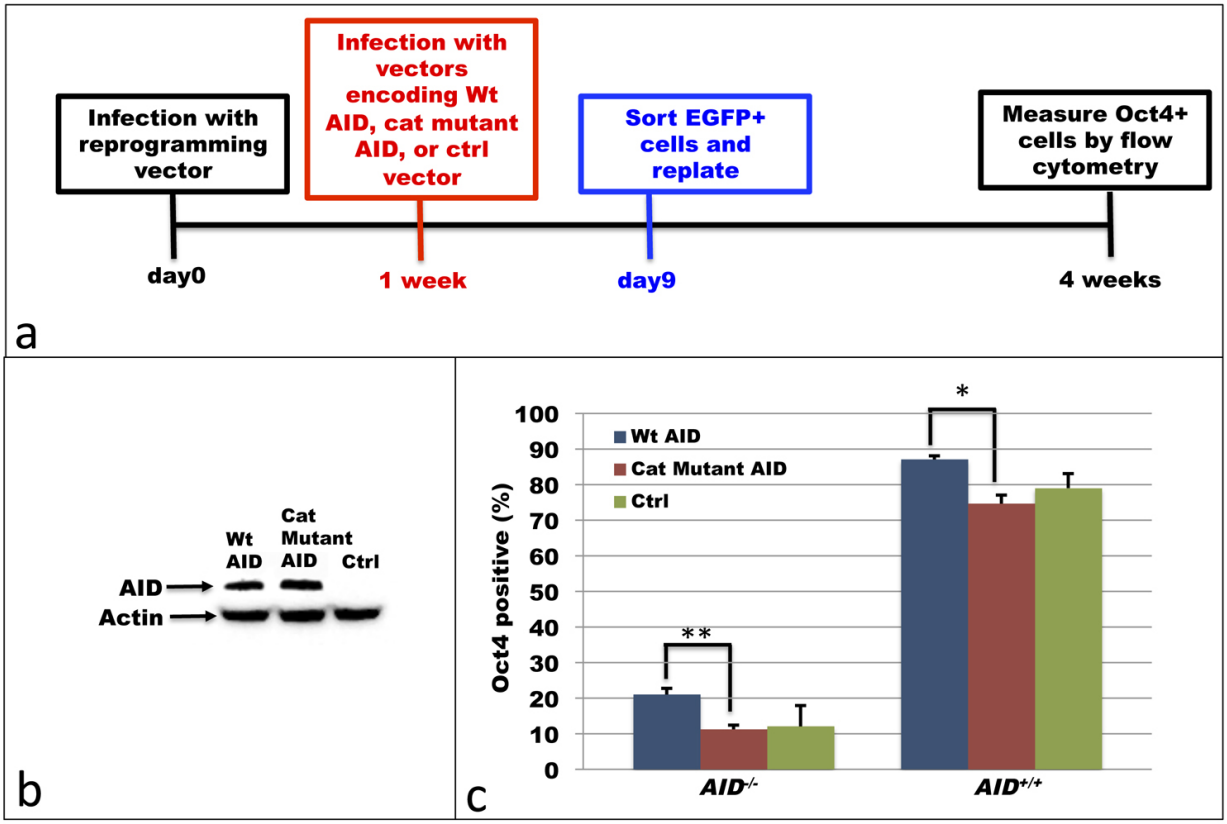
Supplemental Figure S5. AID stabilizes pluripotency during reprogramming of mouse embryonic fibroblasts. Shown is a representative flow cytometry experiment quantifying the relative percentage of SSEA1 positive cells in $AID^{+/+}$ and $AID^{-/-}$ cells after 1, 2, 3 or 4 weeks, following OSKM transduction. Equivalent results were obtained in three independent experiments. In each case the paired cells from a common stock were transduced at the same time with a common viral preparation.

Fig. S6



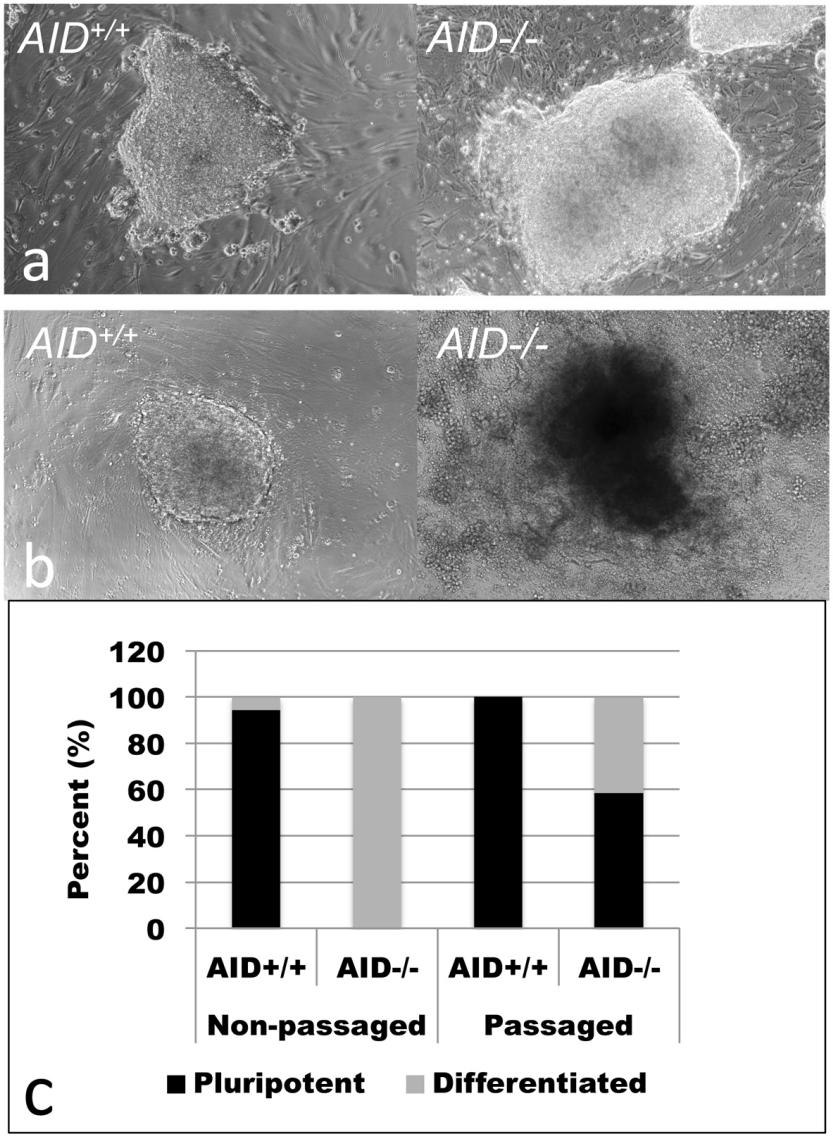
Supplemental Figure S6. AID-null cells fail to stably reprogram regardless of constitutive or inducible OSKM expression. a, Shown are results from a representative experiment comparing directly the relative percentage of SSEA1 positive cells obtained after 4 weeks of transduction using either constitutive or doxycycline-inducible OSKM STEMCCA lentiviruses, as indicated. b, $AID^{+/+}$ and $AID^{-/-}$ fibroblasts were transduced with doxycycline inducible reprogramming vectors. As indicated in the schematic (top) doxycycline was added one day after infection, replenished daily with media change, and was withdrawn after 6, 9, 12, 21 or 28 days. Colonies were stained with anti-NANOG antibody at day 28 (examples shown below). Withdrawal of doxycycline at different time points failed to stabilize AID-null colonies, suggesting that aberrant silencing or continuous expression of reprogramming factors has no role in the destabilization of AID-null colonies. c, In contrast to the AID-null cells, there is a progressive increase in the number of NANOG positive colonies correlating with extended period of doxycycline treatment in the transduced wildtype cells.

Fig. S7



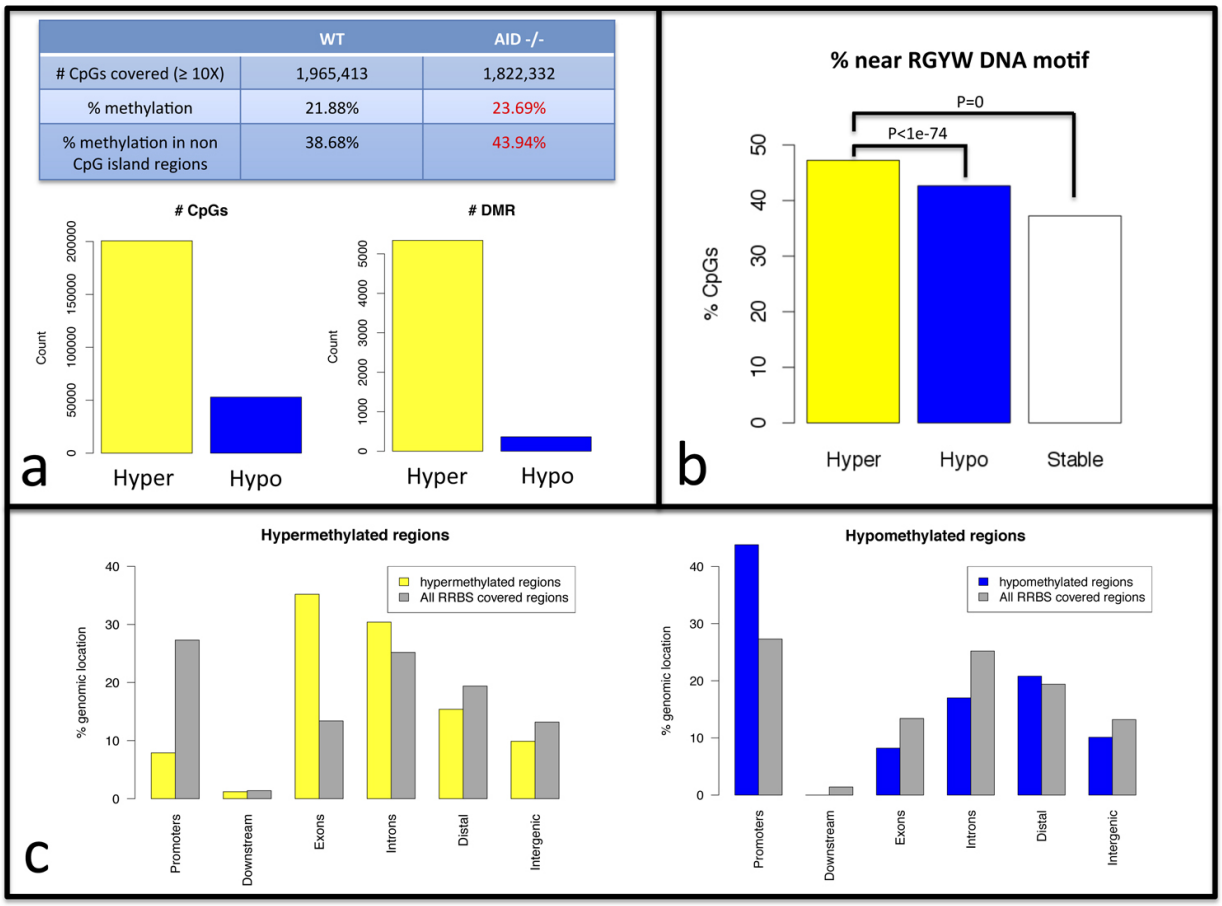
Supplemental Figure S7. Partial rescue of pluripotency can be achieved by ectopic expression of wildtype AID in AID-null fibroblast. a, Shown is a schematic of experimental strategy. After 1 week of transduction with the reprogramming vector, AID-null or wildtype fibroblasts were infected with retroviral vectors encoding wildtype AID, catalytically mutant AID, or control empty vector. Each of these vectors also expressed EGFP. After 48hrs of infection (day 9) EGFP+ cells were sorted and replated on gelatin-coated dishes with mouse ESC media. After culturing for an additional 3 weeks (4 weeks total after initial transduction with reprogramming vectors) cells were analyzed for OCT4 expression by flow cytometry. b, Western blotting experiments demonstrate equivalant expression of wildtype and catalytically mutant AID 48 hr after retroviral infection. Beta Actin serves as a loading control. c, Shown is the percentage of OCT4 positive cells quantified by flow cytometry. Ectopic expression of wildtype AID in AID-null cells resulted in a statistically significant increase in the percentage of OCT4 positive cells (left set, **$p<0.01$). Overexpression of AID in wildtype fibroblasts also resulted in an increase in OCT4 positive cells (right set, *$p<0.05$). Data represents mean of three individual experiments +/- standard deviation. The same results were obtained when SSEA1+GFP+ cells were sorted at day 9.

Fig. S8



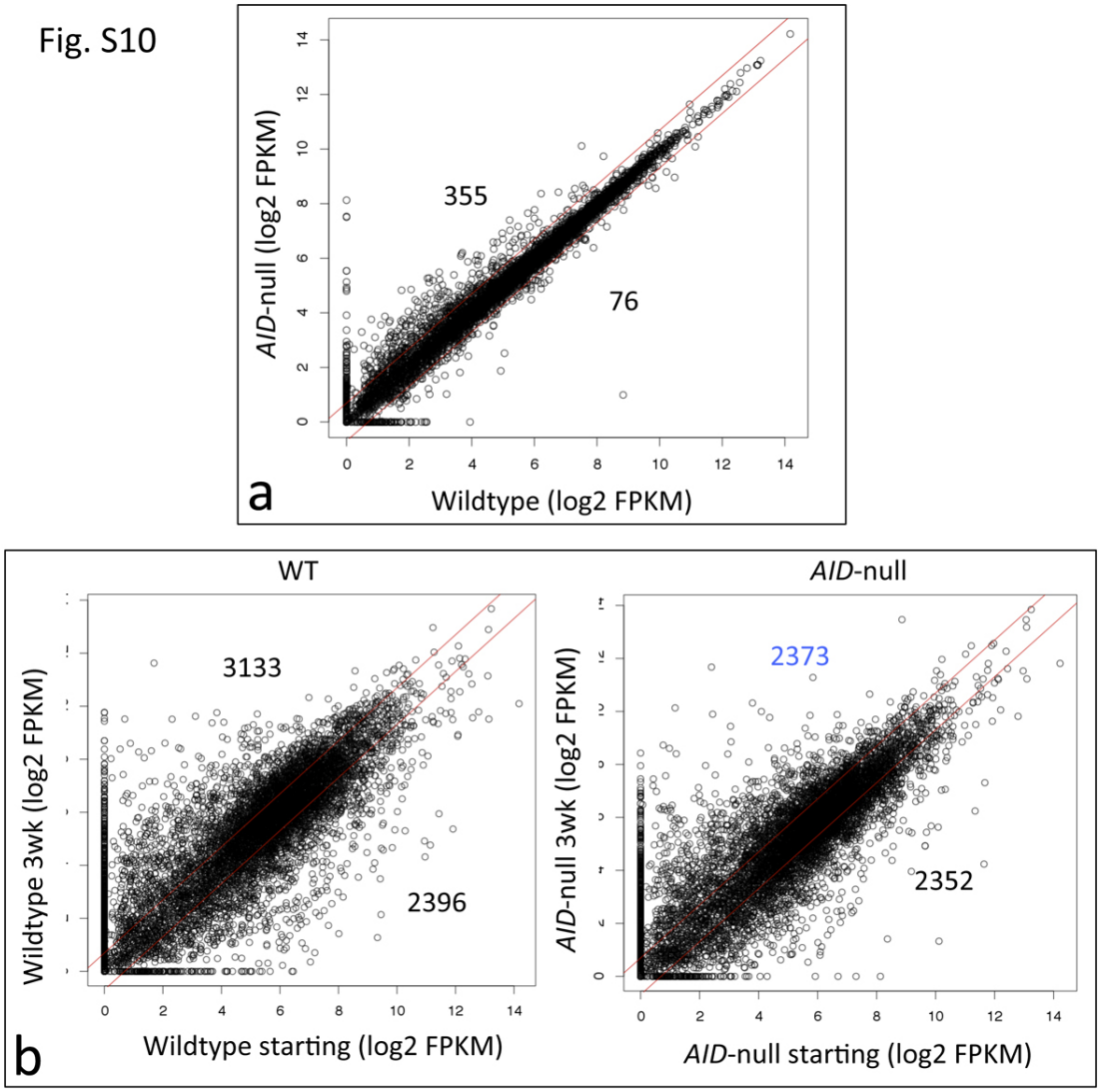
Supplemental Figure S8. Passaging is required for some AID-null clones to stabilize pluripotency. a, Morphology of colonies derived from wildtype or AID-null fibroblasts isolated at two weeks post-transduction. b, Morphology of isolated colonies at 3 weeks (one week after they had been isolated), when they were not passaged. Note the loss of characteristic iPSC morphology in the AID-null cells. c, Relative percentage of clones that remained pluripotent or differentiated, with or without passaging.

Fig. S9



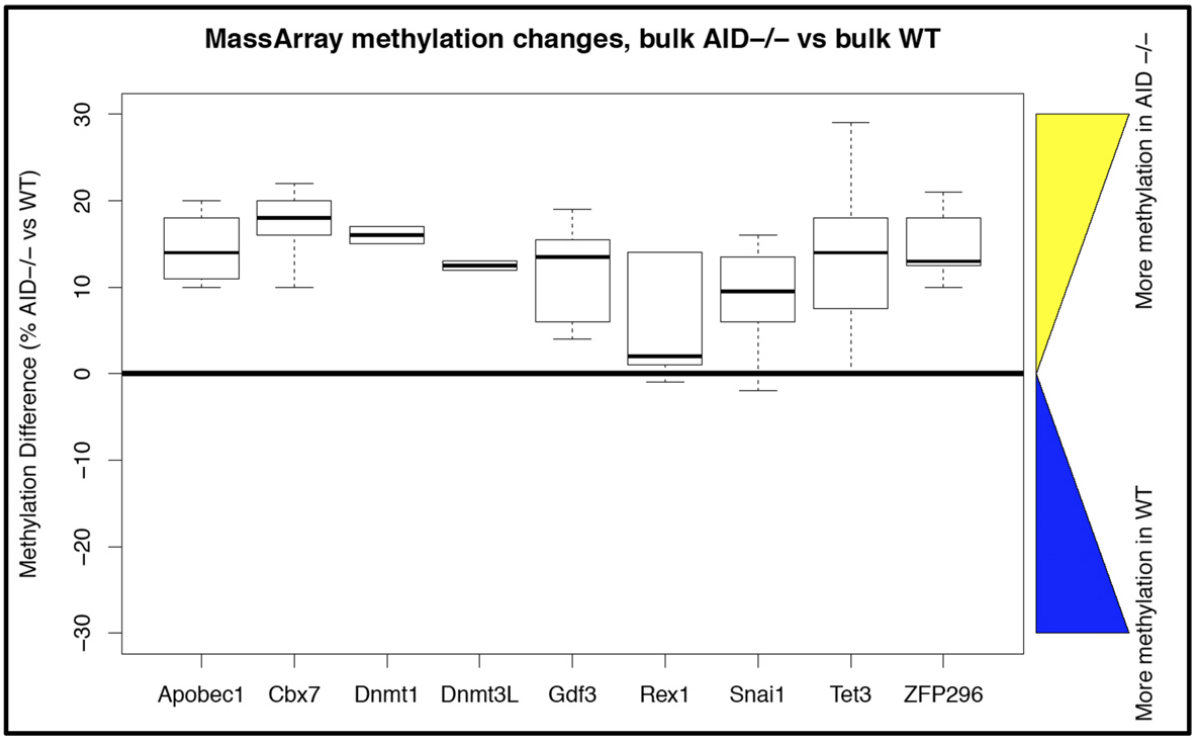
Supplemental Figure S9. Gene bodies are hyper-methylated in AID-null cells during the reprogramming process. Results shown in these figures were obtained after performing reduced representation bisulfite sequencing (RRBS) of wildtype and AID-null cells at the 3 week time point. Short bisulfite converted reads obtained by RRBS were aligned to the bisulfite converted reference mouse genome (mm9) using the Bismark program. a, >1.8-1.9M CpGs were covered with 10 or more reads (>10X) in each experiment. Globally, the total fraction of methylated CpGs was increased from 21.88% to 23.69% in $AID^{-/-}$ cells. When only counting CpGs not localized to CpG islands, the increase was even further pronounced (38.68% to 43.94% in $AID^{-/-}$). These numbers were obtained by dividing the total number of non-converted CpGs in bisulfite treated reads to the total number of CpGs. Compared to the wildtype cells, the AID-null cells show considerably more hyper-methylated CpGs. In this analysis, differential methylation was evaluated using Fisher exact tests comparing methylated vs total CpGs read counts in wildtype (WT) vs null ($AID^{-/-}$) and further adjusted for multiple testing, controlling the false discovery rate to 20%. Similar results were obtained at more stringent thresholds (not shown). Correspondingly, when we identified differentially methylated regions (DMRs) we observed vastly more hyper-methylated regions (Hyper) than hypo-methylated regions (Hypo). DMR were identified as described in Methods. b, Characteristic for AID is the preferred localization of hyper-methylated CpGs sites near (< 5 bp away) an RGYW hot-spot sequence, which is the known DNA recognition sites of AID (R=A/G, Y=C/T, W=A/T). In this figure, the fractions of differentially methylated CpGs were compared using Chi-square tests. c, (left) A larger fraction than expected of hyper-methylated regions in $AID^{-/-}$ vs WT (yellow) is located in gene bodies (exons and introns). The grey bars indicate the fraction of RRBS covered regions in each genomic compartment in the mouse genome (RRBS is biased towards CpG islands, hence the increased coverage of promoter regions that frequently contain or overlap with CpG islands). In contrast, hypo-methylated regions (right) are biased more to promoter regions.

Fig. S10



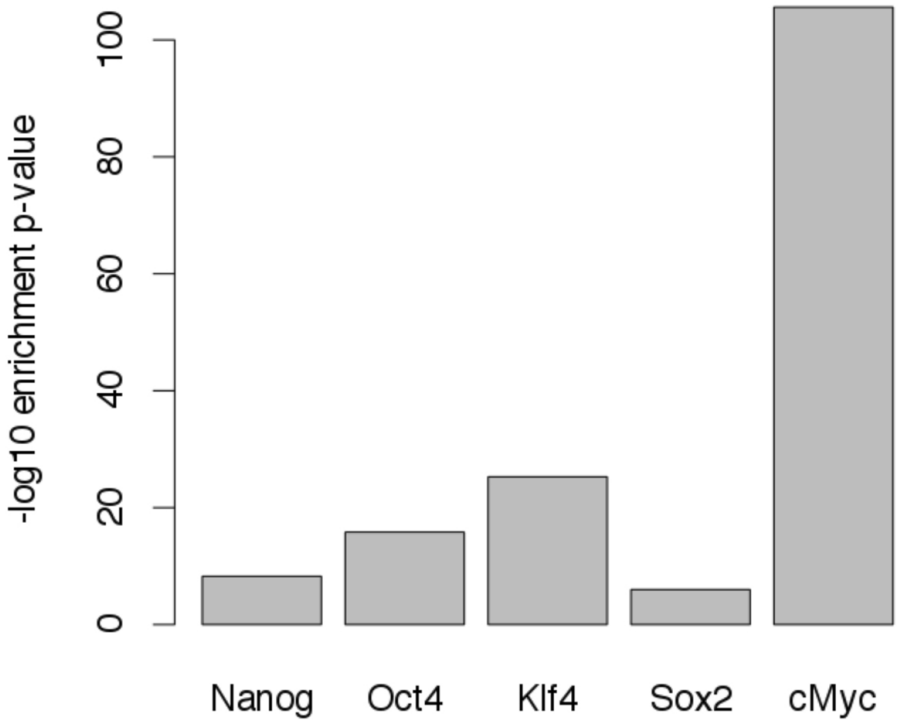
Supplemental Figure S10. Plots showing gene expression profile in AID-null and wildtype cells obtained by RNA-seq analysis. a, Plots show expression levels (FPKM, log2-transformed) from RNA-seq data in pre-transduced fibroblasts. FPKMs were calculated using TopHat for alignment and CuffLinks for quantification, as described in Methods. Each dot in these figures represents a gene. With a cut-off of 2-fold (red lines) and considering only genes expressed above 5 (FPKM) in at least one of the samples, 76 genes are relatively activated in wildtype cells, and 355 are relatively activated in AID-null cells. b, Plots show expression levels (log2 FPKM) from RNA-seq data from cells taken at 3 weeks of the reprogramming process relative to expression levels in pre-transduced fibroblasts. With a cut-off of 2-fold and requiring an FPKM of at least 5 in one sample, 3,133 genes are activated in wildtype (WT) cells, but only 2,373 are activated in AID-null cells. In contrast, the number of down-regulated genes is approximately the same in WT and AID-null cells.

Fig. S11



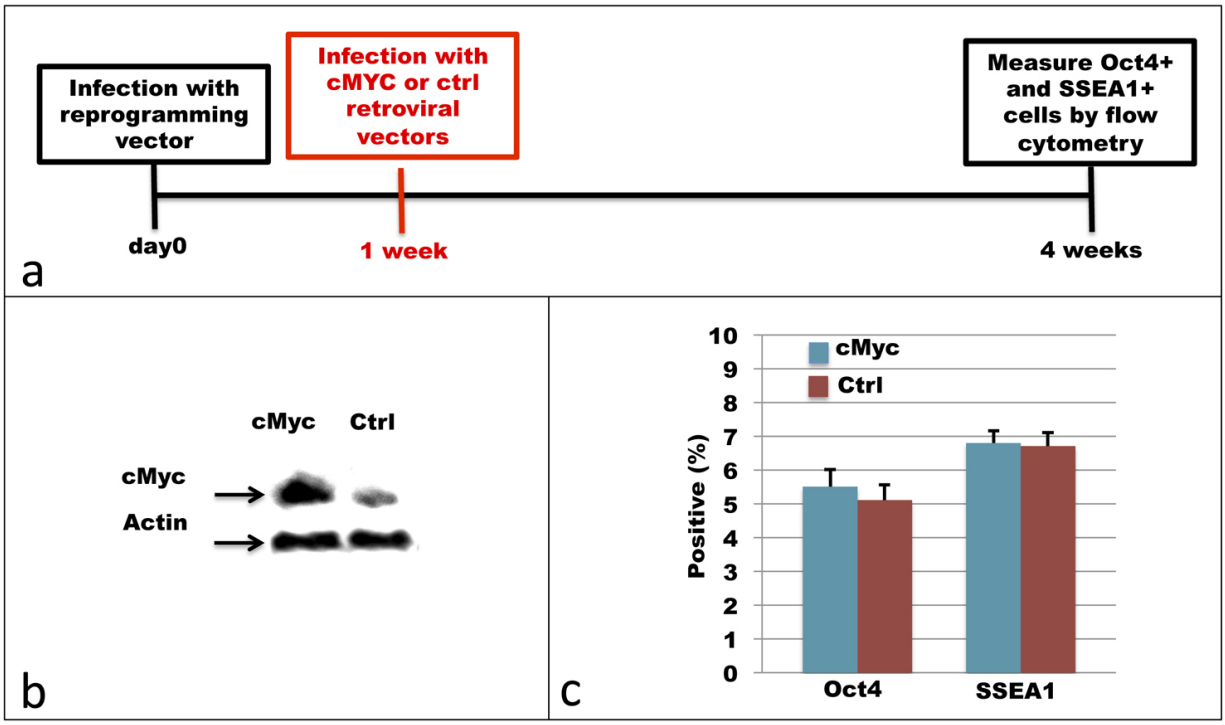
Supplemental Figure S11. MassArray validation of hyper-methylated regions in AID-null cells. MassArray analysis was performed on DMR regions identified by RRBS and located within each of the secondary pluripotency genes. For each CpG within each region, the wildtype methylation level was subtracted from the AID-null methylation level, and the CpG differences were plotted using boxplots. As shown in this panel, the vast majority of CpGs from regions identified as hyper-methylated by RRBS are also hyper-methylated according to the MassArray analysis.

Fig. S12



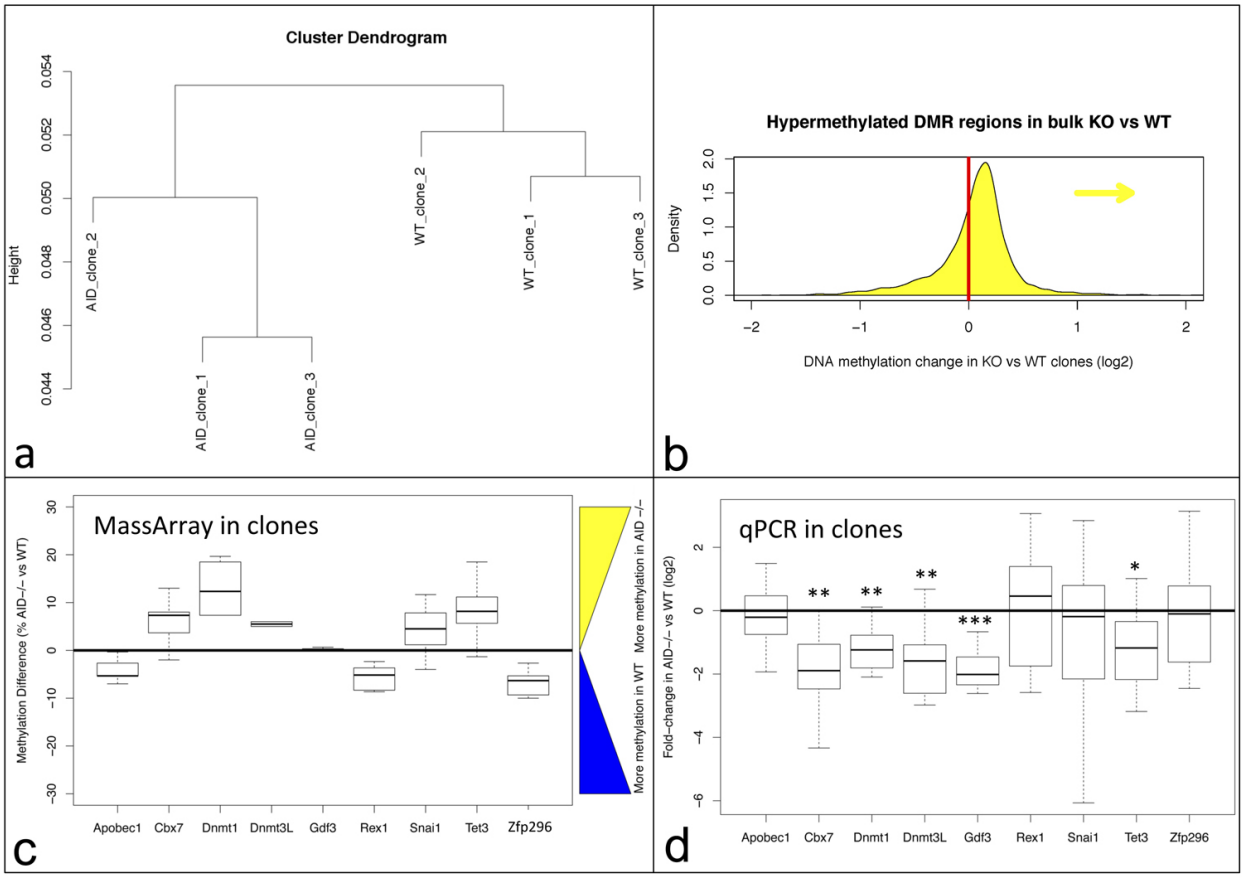
Supplemental Figure S12. cMyc target genes fail to be up-regulated during reprogramming of AID-null cells. The chart represents the level of overlap between genes that fail to be up-regulated in AID-null cells relative to wildtype cells during reprograming and genes that are bound by core pluripotency factors in ESCs, identified through publicly available ChIP-seq datasets (Chen et al. Cell 133, 1106-1117, 2008). Genes that fail to upregulate in AID-null cells were defined as those that upregulate 2-fold or more during reprogramming in WT cells, but not in AID-null cells. For this analysis, ChIP-seq reads were downloaded from GEO/SRA, aligned to the mm9 genome using BWA; peak calling was performed using ChIPseeqer (Giannopoulo & Elemento. BMC Bioinformatics 12, 277, 2011). Peak annotation was performed using RefSeq genes using the ChIPseeqerAnnotate program. Overlap was assessed using the hypergeometric distribution. Significance of enrichment (negative log10 hypergeometric p-value) is graphed on the Y-axis. Of note, all p-values were significant; however the overlap with cMyc target genes was vastly more significant than the other overlaps (even though other factors e.g. Klf4 have more target genes).

Fig. S13



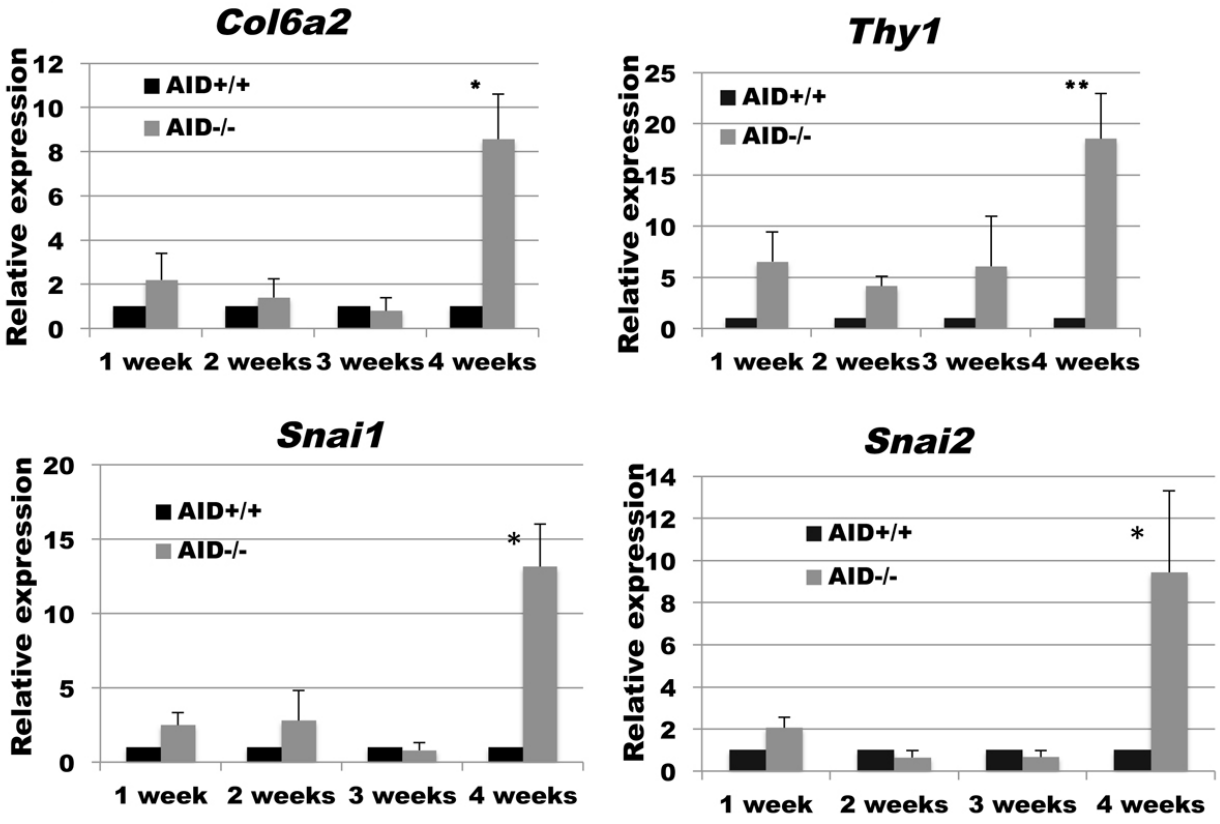
Supplemental Figure S13. Forced cMYC expression is not sufficient to rescue AID-dependent pluripotency. a, As indicated in the schematic, 1 week after transduction with reprogramming vectors, transduced AID-null cells were infected with a cMYC-expression vector. b, Western blotting experiments demonstrate over-expression of cMYC in AID-null cells. c, However, no difference was observed in the percentage of OCT4 or SSEA1 positive cells at 4 weeks, as measured by flow cytometry.

Fig. S14



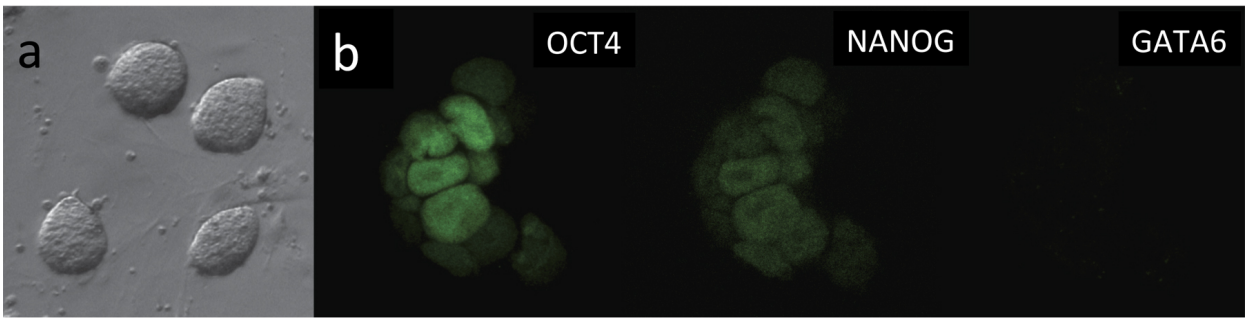
Supplemental Figure S14. Methylation studies and gene expression profiling of individual clones confirms a role of AID in stabilizing induced pluripotency through DNA demethylation. a, Isolated clones of iPSCs derived from AID-null or wildtype cells cluster with their own genotype according to methylome. Cluster dendograms are based on genome-wide DNA methylation patterns determined by RRBS global analysis. Briefly, the genome was binned in 100 kb windows and DNA methylation levels were evaluated in each window, for each clone. Methylation levels were obtained by averaging the methylation levels of all CpGs within each 100kb window. Average linkage hierarchical clustering of the clones was then performed using their methylation profiles and 1 – Pearson correlation as distance. b, Changes in methylation at the genome-wide level identified in bulk reprogramming cultures of wildtype or AID-null cells are validated in isolated clones. Following RRBS of DNA from 3 randomly isolated clones derived (each) from AID-null or wildtype fibroblasts, the average methylation (KO/WT) ratio (log2-transformed) was determined at all differentially methylated regions identified in bulk colonies. Hyper-methylated regions identified in the bulk samples (n=5,341) tend to be hyper-methylated in the knockout clones compared to wildtype clones, indicated by the rightward shift of the area under the curve (p<1e-151, paired Wilcoxon test). c, MassArray bisulfite sequencing validates hyper-methylation phenotype of most secondary pluripotency genes. Analysis was carried out as in Fig. 4e, except using DNA from isolated wildtype or AID-null reprogrammed clones, isolated initially 2 weeks into reprogramming, and evaluated at approximately 3 weeks. Shown is the average for three AID-null clones compared to the average for three wildtype (WT) clones. 6 of the 9 genes validated at the clonal level. That 3 of 9 did not might reflect the passaging of these cells required to generate sufficient materials. d, qPCR analysis of hyper-methylated secondary pluripotency genes showed significantly lower expression levels in AID-null clones compared to wildtype clones (*p<0.05, **p<0.01, ***p<001). 12 individual clones were picked from AID-null or wildtype cells 2 weeks following initiation of reprogramming in two separate experiments, and evaluated at approximately 3 weeks.

Fig. S15



Supplemental Figure S15. Fibroblast-associated genes are expressed at relatively high levels in AID-null cells that are failing to stabilize pluripotency after 4 weeks of reprogramming. Shown are the relative expression levels of several fibroblast-associated genes in $AID^{+/+}$ and $AID^{-/-}$ fibroblasts quantified by qPCR 1, 2, 3 or 4 weeks following OSKM transduction. Data is derived from three independent experiments and represents mean +/- standard error (*p,0.05, **p<0.01).

Fig. S16



Supplemental Figure S16. Embryonic stem cells can be derived from AID$^{-/-}$ embryos, although less efficiently compared to wildtype embryos. a, Representative phase contrast image of ES cell colonies derived from an AID$^{-/-}$ blastocyst. Without exception, AID–null ESCs show characteristic ESC morphology with sharp-edged, round colonies shown here at passage 9 following derivation. b, Representative immuno-staining of an AID$^{-/-}$ ESC colony stain positive for pluripotent markers using with anti-OCT4 and anti-NANOG antibodies, and do not express the differentiated lineage marker GATA6, using anti-GATA6 antibodies. These representative colonies were stained at passage 6 after derivation, and all clones are stable for pluripotency markers beyond passage 10.

Supplemental Tables

	Primer sequence (Forward)	Primer sequence (Reverse)	Accession number
mAid	GCCACCTTCGCAACAAGTCT	CCGGGCACAGTCATAGCAC	NM_009645
mNanog	GGACAGGTTTCAGAAGCAGAAG	ACCATTGCTAGTCTTCAACCAC	NM_028016
mOct4	ACCTCCTCTGAGCCCTGTGCC	CACCTTTCCAAAGAGAACGCCCA	NM_013633
mSox2	GCGGAGTGGAAACTTTTGTCC	CGGGAAGCGTGTACTTATCCTT	NM_011443
mKlf4	TGCCAGACCAGATGCAGTCAC	GTAGTGCCTGGTCAGTTCATC	NM_010637
mUtf1	GTCCCTCTCCGCGTTAGC	GGGGCAGGTTCGTCATTT	NM_009482
mKlf2	CTAAAGGCGCATCTGCGTA	TAGTGGCGGGTAAGCTCGT	NM_008452
mEras	GCCCCTCATCAGACTGCTAC	GCAGCTCAAGGAAGAGGTGT	NM_181548
mCbx7	TGCGGAAGGGCAAAGTTGAAT	ACAAGGCGAGGGTCCAAGA	NM_144811
mGdf3	GGGCCTCGCAGGACTTATG	TGGTCGCAGGTTATAGTAGGAC	NM_008108
mZfp296	CGAGGGACATGCCCATATCG	GGCGGTCGGAAGACTTAGAAC	NM_022409
mRex1	ACAAGGCGAGGGTCCAAGA	TATGACTCACTTCCAGGGGGCACT	NM_009556
mApobec1	ACCACACGGATCAGCGAAAC	GGGGTAGTTGACGAAATTCCTC	NM_001134391
mDnmt1	CCGTGGCTACGAGGAGAAC	TTGGGTTTCCGTTTAGTGGGG	NM_001199431
mDnmt3l	TGGAGACTTCCGACAGCTCTA	AGGGCTGGGGAGGATTTCA	NM_001081695
mTet3	TGCGATTGTGTCGAACAAATAGT	TCCATACCGATCCTCCATGAG	NM_183138
mcMyc	GATCAGCTCTCCTGAAAAGA	TCGAGGTCATAGTTCCTGTT	NM_001177352
hOCT4	TGGGCTCGAGAAGGATGTG	GCATAGTCGCTGCTTGATCG	NM_002701
hKLF4	CCCAATTACCCATCCTTCCT	ACGATCGTCTTCCCCTCTTT	NM_004235

Supplemental Table ST1. Primers used for qPCR analyses.

Gene	Amplicon Sequence
Rex1	GGGTGCAAGAAGAAGCTGAGGGATAAAACCGCCCTGAGGAAGCACATGCTTGTCCACG GGCCCCGGCGGCACGTGTGTGCAGAGTGTGGCAAAGCCTTCA
Dnmt1	CCAAGGTGGCTGTCAAGTGGTGCGATGCATGTTTGAGCAAATAGAATTGATTAAAATAGA TTTAAAATGGCACTCGTCTGGCTTTGTGTCTACCAGGCATCT
Dnmt3l	GGGAGAAAGGCTAAGCCCCTACTCACCTGAAGGAAGCGGGTAGTTGTCTCTTGGTCATC CTCAGTCAGCAGCAGATTGTCCATGAATATCCAGAAGAAGGGC
Zfp296	CAGCAGTGGCAGTGGTGGCGCCAGTTGAAAGCAAGCCCCCTCCAGTGAGCAGTGCTGC CCGGCGGAGCCCCACCTGTGATGTGTGCAAGAAGACCCTCAGCTCTTTCAGCAACCTCA AGGTCCACATGCGTTCCCACACTGGTGAGCGACCCTACTCCTGTGACCAGTGTTCCTAT GCCTGCGCTCAGAGCAGCAAGCTCAACAGACACAAGAAGACCCATCGGCAGCTGGCAC CCGGGAGCCCCTCCACTTCTGCC
Apobec1	CCCTGCAGGTGGCTGTGGGCACACCTGAGGAAACAAAGTCCGGCACACAGCTGGCAGA AGCCATCGCAGCAACATAAGCTCCCGAGGAAGGCGTCCAGAGACAGAGCAAGATGAGTT CCGAGACAGGTAAAGGGATGCCGGGTTCTAGCAAGGTGTTGAGTTGA
Gdf3	AAATTAAATTCTGCTAAACCGGCACCGCGGTAGTTAGACTAGCGTGGAGAGGTGATAAGA ACACGTGCCTTCCTCTGGGCTGATCTTGGAACCAAGACAGAATGTGCAAGATCCAAGGC CAGACTCCGGCCAGCCATGGTGTCCATCTGTATCCTCCTCCCTGCTCCACAGCACCCAC TTGAGATTTTTCTCTGATAAACCCTCGGGTTCCTTACCCTGGTCTGGGAGAAGCTGAAGC AGGTTCCCACGAACACCCAGCTCCTTCACGTAGCATAAGTCCTGCGAGGCCCCACTGGC TGCAGCGGCTTCTCGAGCCCGGATGATTTTCCTTA
Cbx7	GTGTGCATGTTGGGGTGTCCGGATGTGTTCATGGGCATCTCTACACAGGCCAATGGCAG GGAGACAGCACACTCGGAGGAGGCTTGGTGGGCAGCCTCTTCCCTGGTCCAGCCCGTG CCGGTTGTCCTGATCCTGTCCCCGGTGGCCAAGGCCGGCCTCTGTGTCCCACCACAAG GCTGGAGTGGAAAGTTCATAGTAAAGACCCACGGCAAGGGACAGGGAAGGGACACCAC CAGAAAGCTGGGGTGACTGCCTGAGGCAAGGTCTGGACCTCCTGGGCTCCTCCGGCTC CACGGATCAGCCTCTG
Snai1	TGGCTGTGCTTGGGCTGACCGGGCCCTGTCCAGGCCTTCCTAACCCTGCTGGTATCTCT CCCCAGGTGAGAAGCCATTCTCCTGCTCCCACTGCAACCGTGCTTTTGCTGACCGCTCC AACCTGCGTGCCCACCTCCAAACCCACTCGGATGTGAAGAGATACCAGTGCCAGGCCTG TGCCCGAACCTTCTCCCGCATGTCCTTGCTCCACAAGCACCAAGAGTCTGGCTGCTCCG GAGGCCCTCGCTGACCCTGCTACCTCCCCATCCTCGCTGGCATCTTCCCGGAGCTCACC CTCCTCCTCAC
Tet3	1)TTTCCCTCCGGGTTTTTTCCGCTTGCCCGTCGGCTCGGCTCATTTCCCAGGACACAGC AAAGTGGTAGGCATTTAGTAAACCGTCTGCCAAATGAAGGGGCTGTAGAAACCCCAGCC CAGCAGGCTGCTCCACACCCTTCTCTAACACTCAAACGGACAGACTTACCGACACAATCG CAGGTAGGGAACTCGGACTGAGCCTTCTTGGCCGGTGTGTCCAGCAGACTCTTAGTAGG TGTGTCTAGGTACTTTAGAGGTGACTCCAAGAAGCCACTGAGGGTTGGTGTGAGTGGGT TCTCAGCCTTGGTGGGCGTGGCCTCCTGTCCCCCCTCTTCTGAATGGAAGCAGGTAGTT GAGAGCACAGTCACGGCCCCTGAGGACTCGATCTTGATCTTCTTGGGGGAGCGGGTGG CAAAAGGGCTCTCCGGAGCTGGGAGACATGTTGATTGGTTTTCAGGTTCTTGAATGGGC ACCGAAGGTGGGCCGGGAAGCCCAAAGCTATCCCCAAATTCAGCCTCAAATTGCCGGAT GAGCTCCTCCAGCTTGTCATCGGCAGGGGGAAGGGGGCCACCAGTGGAGCCTGACTGC AGGGCTACCATGGGGCTGGGGGATCTCATTTCCTGAGTAGGGGGCAGCAGGCTGTCCC CGGAGGGACTAGGTGCAAATAG 2)TGGGACTGGGAGGCTGCCTCAGGCCCCACTCACGCTGCTCGTCTGGAAGATGCCCAC GACCTGGTGGCCTTTTCGGCCGTGGCCGAAGCTGTGTCATCTTACGGGGCCCTTAGTAC CCGGCTCTATGAAACCTTCAACCGTGAGATGAGTCGTGAGGCTGGGAGCAACGGCAGG GGCCCCCGGCCTGAGAGCTGCTCTGAGGGCAGTGAAGACCTGGACA

Supplemental Table ST2. Shown are the amplicons that were sequenced by MassArray analysis to validate the RRBS data. Note that two replicons were probed for Tet3.

Sample	Number of Reads	Number of Unique Aligned Reads	Alignment rate	Number of CpGs covered	Mean CpG coverage	Conversion rate
AID-null clone 1	156476796	102307140	65.40%	3111280	77.75	99.65%
AID-null clone 2	154731507	102016359	65.90%	3536375	66.74	99.70%
AID-null clone 3	165603203	111544028	67.40%	4214110	72.05	99.46%
WT clone 1	158138041	105320482	66.60%	3408245	84.65	99.57%
WT clone 2	155146448	104404463	67.30%	2933321	96.77	99.56%
WT clone 3	159087745	107087549	67.30%	3641092	80.52	99.57%
AID-null 3 wks	164635077	108625223	66.00%	1822333	161.98	99.50%
WT 3 wks	157653137	106427858	67.50%	1965414	134.04	99.70%
AID-null starting fibroblasts	152209632	98631148	64.80%	2629236	86.21	99.81%
WT starting fibroblasts	153679092	99364775	64.70%	2782197	84.69	99.81%

Supplemental Table ST3. Meta-data for the enhanced RRBS analysis on individual AID-null or wildtype (WT) clones or bulk colonies evaluated three weeks after reprogramming was initiated, or in starting fibroblasts.

3. DISCUSSION

3.1. TOOLS FOR STUDYING PREIMPLANTATION DEVELOPMENT

TISSUE CULTURE MODELS

To understand how tissues and organs function individually and in concert with each other it is imperative to understand how these highly functional units arise during development. An important branch of research in this direction is the investigation of the function and roles of individual genes and gene regulatory networks in different tissues. This has long been achieved by reverse genetics, deleting single or multiple genes to infer their original function and possible interactions with other factors. These studies rely on the availability or generation of suitable animal models for direct *in vivo* analysis. In many instances, specific tissues have to be isolated for further study, which is often technically difficult and/or quantitatively insufficient. For this reason cell culture models have been developed for many cell types. These are essential for high-throughput and quantitative biochemical assays. The *in vitro* culture of stem cells has proven especially useful for the study of development, as they can be kept in culture indefinitely. Importantly, stem cells hold the developmental potential to differentiate into various cell types, thereby recapitulating differentiation processes observed *in vivo*, making them suitable for the safe testing of drugs and other therapies.

Biochemical studies of preimplantation development are challenging due to the small size of these early embryos and their lineage compartments. Embryos at the late blastocyst stage reach a size of 100-200 cells and encompass two extraembryonic lineages, the TE and the PrE, as well as the embryonic lineage, the pluripotent EPI (see Introduction p. 5-7). Not only is the isolation of these lineages highly difficult but they also yield cell numbers for downstream biochemical applications. Stem cell culture models have therefore been successfully established for all three lineages. To gain insight into the second cell fate decision and the nature of the resulting lineages,

embryonic stem (ES) cells are derived from the EPI and extraembryonic endoderm stem (XEN) cells are derived from the PrE.

ES CELLS

The first successful *in vitro* culture of mouse ES cells was established in 1981 (Evans and Kaufman, 1981; Martin, 1981). Ever since, ES cell biology and methods to derive and maintain these cells have been studied extensively. Derivation efficiency has been a major problem for researchers in the field. Implementations of highly technical methods like isolation of the ICM and improvements in ES cell culture, such as addition of LIF, have improved ES cell derivation over the years (see Introduction p. 26-29). The development of the **2** **i**nhibitor (2i) culture conditions, utilizing inhibitors against the differentiation promoting FGF/ERK pathway and GSK3, has caused a leap in the understanding of the nature of pluripotency and the maintenance of ES cells in a state of so called naïve pluripotency (see Introduction p 28).

We have developed a protocol for the derivation of ES cells, taking advantage of the 2i system (Publication III). Using this protocol we achieved derivation efficiencies of 60% - 99%, including for previously considered non-permissive strains such as NOD or DBA. Combining the pluripotency promoting properties of 2i during the blastocyst outgrowth phase with the growth enhancing effects of serum during the colony-forming phase could be the decisive factor for the success of this protocol. Culture in 2i would thereby select for stably pluripotent epiblast cells inside the ICM early on, when it is imperative to allow only pluripotent EPI cells rather than PrE or TE cells to grow. Once this EPI outgrowth is disaggregated, cells are growing in single-cell culture and are therefore highly sensitive. Addition of nutrient and growth factor rich serum at this stage could enhance their survival and proliferation capacity.

The described protocol allows for routine, highly efficient and straightforward ES cell derivation, without the requirement of immunosurgical removal of TE or addition of other factors and produced germ line competent

ES cells possessing normal karyotypes. Our protocol has since been employed successfully in several studies (Bangs et al., 2015; Li et al., 2014; Schrode et al., 2014; Wen et al., 2014; Xenopoulos et al., 2015) and has allowed various laboratories the ready derivation of ES cell lines from a multitude of mouse models (unpublished data, students of the CSHL mouse genetics course 2013). Considering the high success rate, it could be imagined that this protocol or slight variations of it can be used for the derivation of ES cells from other species, as has been achieved similarly for rat and naïve ES cells from human (Buehr et al., 2008; Li et al., 2008; Ware et al., 2014).

Our protocol makes ES cell derivation more feasible, even for less experienced researchers. This could encourage the study of cell culture models in parallel to *in vivo* studies, facilitate the investigation of factors influencing the pluripotent state as well as differentiation processes, and allow the ready utilization of ES cells where *in vivo* studies are not suitable.

As mentioned above, ES cells are mostly used to model differentiation into somatic tissues to study development, organ function and disease. The study of preimplantation development however exploits the high degree of similarity of ES cells to the pluripotent epiblast in the blastocyst stage ICM. This cell culture system has advanced the understanding of the mechanisms and processes involved in the second cell fate decision, which cannot be otherwise studied *in vivo*. For instance, the effect of FGF/ERK signaling on the second cell fate decision has been hypothesized through reverse genetics but the mechanism through which this takes place, the transcriptional repression of *Nanog,* was confirmed in ES cells (see Introduction p. 19).

XEN CELLS

The derivation of XEN cells as a model for PrE development has only recently been achieved (Kunath, 2005). It is therefore a comparatively novel stem cell type, whose derivation and utilization is still in a phase of trial and optimization. We have created a collection of protocols for the derivation of XEN cells from blastocysts as well as ES cells (Publication II). XEN cells can

be derived from blastocyst outgrowths using ES or TS derivation conditions, i.e. in the presence of LIF or FGF4/Heparin respectively (see Introduction, p. 31). This intriguing observation hints at imbrication of pathways acting during preimplantation development and a more complex picture of regulation of the second cell fate decision.

For ES derivation conditions, LIF is added to the culture. In ES cells LIF acts through the JAK/STAT3 pathway to promote self-renewal (see Introduction p. 27). However, LIF also activates the MEK/ERK pathway, which is necessary for PrE development (see Introduction p. 15-16). This might explain how PrE cells are maintained in blastocysts and their outgrowths when cultured in LIF containing media. Slight modulation of the derivation protocol must then be used to obtain XEN cells rather than ES cells under these conditions. Longer culture phases before disaggregation of the outgrowth might provide additional time to fully establish the PrE. Gentle dissociation may also favor XEN cells. Epithelialized PrE cells tend to proliferate to encompass EPI cells when culturing isolated ICMs of the blastocyst stage (Handyside, 1978; Rossant, 1975). By creating several partial outgrowths through light disaggregation the overall surface for PrE cells to proliferate on is increased, potentially favoring their propagation over ES cells.

For TS derivation conditions, FGF4 and Heparin are added to the derivation media since MEK/ERK signaling is promoting TE proliferation (see Introduction p. 9). As discussed above, FGF/ERK signaling also plays an important role in PrE formation. This could explain why TS derivation conditions are also supportive for the derivation of XEN cells. However, since FGF signaling has a proliferative effect on TS cells they might be more likely to outcompete XEN cells in these culture conditions. Accordingly, we found that XEN derivation efficiencies are higher using ES conditions than TS conditions. Active ERK signaling needed for PrE formation might be sufficiently provided by LIF and through FGF4 secreted by EPI cells.

Moreover, exogenous FGF4 does not enhance reprogramming of ES cells to XEN cells through *Gata6* overexpression (Wamaitha et al., 2015).

XEN cells can also be derived by overexpression of the master regulator *Gata6* in ES cells, underscoring the relevance of this transcription factor for PrE lineages (see Introduction p. 13). Interestingly, this conversion is also possible by treatment with growth factors such as retinoic acid supplemented with activin A (Cho et al., 2012). Activins are members of the transforming growth factor beta family and signal via Smad2 to activate downstream transcription factors (Abe et al., 2004). Interestingly, it has been suggested that Smad2,3 signaling is involved in parietal endoderm formation (Roelen et al., 1998), which could elucidate the supporting effect of Activin on XEN cell conversion from ES cells. Retinoic acid on the other hand has previously been shown to facilitate PrE formation in ES cells and embryoid bodies (Capo-Chichi et al., 2005; Mummery et al., 1990; Rohwedel et al., 1999). Indeed, retinoic acid might positively regulate *Bmp2* (Heller et al., 1999; Rogers et al., 1992), which in turn can activate *Gata4* in heart development (Arceci et al., 1993; Ghatpande et al., 2006; Kostetskii et al., 1999). Overexpression of *Gata4* in ES cells, similarly to *Gata6*, has been shown to lead to XEN cell conversion (Fujikura et al., 2002). Accordingly, activation of *Gata4* through retinoic acid could explain its ability to convert ES cells to XEN cells.

As mentioned before, XEN cells are a comparatively recent stem cell type and are not fully characterized. Different ways to derive them efficiently will help further understand their biology and facilitate their use as a stem cell model (reviewed in (Moerkamp et al., 2013)). For instance, they represent an interesting tool for the study of X-inactivation, since the paternal X chromosome remains silenced in the PrE endoderm and its derivatives, in contrast to the EPI (Kunath, 2005; Okamoto et al., 2004).

Our protocol has proven useful in several studies (Bangs et al., 2015; Blij et al., 2015; Kropp et al., 2014; Loh et al., 2015; Lou et al., 2014; Lu et al., 2015; Merzouk et al., 2014), including very recent work investigating the role

and possible functions of *Gata6* in extraembryonic endoderm development (Wamaitha et al., 2015).

Limitations for the use of stem cell models

While the use of cell culture models to study the second cell fate decision can be useful, it is important to keep in mind that these cells likely do not fully recapitulate the *in vivo* state. The maintenance of cells in culture might introduce expression artifacts that do not exist *in vivo* at the putative stage. Additionally, cultured cells lack the context of their *in vivo* environment, which is essential for effects of local signaling and cell-cell contacts.

Both ES and XEN cells resemble closely derivatives of the EPI and PrE rather then these early lineages themselves. ES cells have recently been found to closest resemble the E4.5 EPI, shortly before implantation (Boroviak et al., 2014). At this stage *in vivo*, the EPI lineage is fully specified and has lost its plasticity to convert to the PrE lineage (see Introduction p. 21-22). ES cells therefore may not be suitable for modeling earlier phases of the second cell fate decision. Similarly, XEN cells are thought to represent the parietal endoderm (PE) lineage (Brown et al., 2010; Kunath, 2005), which is specified after implantation from PrE cells that undergo epithelial to mesenchymal transition and migrate along the basement membrane of the trophectoderm layer to eventually form the parietal yolk sac (Kadokawa et al., 1987). XEN cells might therefore have undergone dramatic changes from their lineage of origin, the PrE.

Single cell analysis of intact embryos

Tissue culture models have proven very useful and have allowed biochemical studies that are not achievable *in vivo*. However, their application has limitations. As discussed above, cells in culture might not faithfully recapitulate many aspects of cells growing *in vivo* necessitating careful interpretations of the results obtained in cell culture experiments.

These issues are well recognized in the field and in recent years new steps have been taken to find solutions. Several groups have created

methods to isolate single cells from early embryos and perform single cell gene expression analysis, including RT-PCR and microarrays (Boroviak et al., 2014; Guo et al., 2010; Ohnishi et al., 2014; Xue et al., 2013). These analyses have proven a useful tool to analyze transcriptional controls of genes through mRNA levels and produced considerable amounts of new information. It will be interesting to see similar work in null mutant embryos for genes important in the second cell fate decision as was done by Ohnishi et al. (Ohnishi et al., 2014). However, transcriptome analysis does not provide a readout for protein levels, which can differ due to post-transcriptional and post-translational regulation. Unfortunately, biochemical assays to measure protein levels, such as Western Blotting, require substantial cell numbers for their completion. These limitations could be overcome using FACS (Fluorescence Activated Cell Sorting), which could be adopted to isolate single cells for the quantification of protein levels.

It is increasingly evident that the *in vivo* context can play crucial roles in biological processes, especially during development. During preimplantation development, neighboring cells seem to influence each other through secretion of FGF4 and how cells react differentially to signaling from their neighbors still remains elusive. Additionally, cells acquire their fate in a seemingly random fashion inside the ICM. Therefore, retaining spatial information is essential. Fixation of whole embryo samples followed by immunohistochemistry and confocal microscopy has been used widely to study development. This methodology retains all spatial information and readily informs about the presence of a protein. However, it generally only provides binary information about the absence or presence of a protein of interest since it is not possible to quantitatively assess differing fluorescence levels by eye.

To achieve unbiased, high throughput single cell quantitative protein expression measurements while retaining spatial information, we created a semi-automated pipeline (Publication I, Figure 1 A, B). MINS (Modular Interactive Nuclear Segmentation) was used to perform highly accurate

nuclear segmentation on 3D microscopy images of preimplantation embryos (Lou et al., 2014). The resulting data comprised fluorescence intensity measurements of all nuclei in the embryo as well as positional information of all cells. To ensure faithful analysis we computationally corrected for variation of fluorescence intensities caused by tissue depth and created a threshold for unbiased assignment of cell fates. In this way, we were able to analyze the entire ICM, taking into account all cells within each embryo analyzed.

To gain a holistic picture of the processes and mechanisms governing development it will be important to utilize methods determining both single cell transcriptional activity as well as protein level analysis (reviewed in (Saiz et al., 2015)).

3.2. THE SECOND CELL FATE DECISION: TO BE OR NOT TO BE PLURIPOTENT

BREAKING SYMMETRY IN THE ICM

After specifying the TE at the morula stage, the ICM comprises a rather homogenous population of totipotent cells, coexpressing factors important for the imminent second cell fate decision. Starting around the early blastocyst stage, expression of the transcription factors *Gata6* and *Nanog* becomes increasingly heterogeneous until they display mutually exclusive expression with *Gata6* marking the PrE and *Nanog* the EPI. What causes the initial disruption of balanced co-expression is yet unknown and poses an important question in the field. Several interesting ideas have been brought forward to answer this question, mostly falling into one of two basic assumptions: predetermination of blastomeres versus stochastic adoption of cell fate through transcription factor heterogeneities (see Introduction p. 20). A satisfying and reproducible answer has not yet crystallized.

STOCHASTIC DIFFERENTIATION IN THE ICM

To gain more insight into this question we utilized our computational analysis pipeline for GATA6 and NANOG and determined cell fates inside the ICM. Since our approach allows for retention of spatial information we were

able to calculate the Pearson's correlation coefficient as a function of cell-cell distance (Publication I, Figure 1 and Figure S1). We found that there was no correlation in the spatial pattern of cell fates at the early blastocyst stage. Instead correlation started to emerge at the mid blastocyst stage, with neighboring cells having a slightly higher chance to exhibit the same fate. These results could be explained if early lineage biases result from stochastic cell-autonomous fluctuations. Some of these progenitors could then proliferate, thereby explaining the slightly higher chance of adjacent cells displaying the same identity later on.

Indeed, cellular noise, often in the form of transcriptional heterogeneity or fluctuation due to the unsynchronized phases of biochemical reactions or epigenetic mechanisms, can lead to bimodal states. Studies showed that controlled transcriptional noise in a reporter gene can lead to cellular variability in bacteria and yeast, with the latter generating bistable protein expression states (Blake et al., 2003; Ozbudak et al., 2002). In *C. elegans* intestinal specification was shown to be highly regulated. However, incompletely penetrant mutations for loss of intestinal cells caused broadly heterogeneous expression of a downstream gene in the gene regulatory network, leading to a threshold dependent allocation of cell fates (Raj et al., 2010). In mouse, heterogeneity in expression levels of a gene in hematopoietic progenitor cells could account for stochastic priming for cell fate decisions (Chang et al., 2008). These examples demonstrate how cellular noise is able to create lineage biases. These biases alone, however, cannot create bistable states. Instead, the fate decision has to be stabilized once the bias is created (reviewed in (Ferrell, 2002)). EPI and PrE fates could thus be biased in a stochastic fashion through transcriptional noise, with their fates subsequently being stabilized by feedback loops, possibly involving FGF signaling. Further considerations of how this could be achieved are discussed below (see Discussion p. 138, "Possible mechanisms for GATA6/NANOG/FGF interplay").

RELATIVE PROTEIN LEVELS AND LINEAGE CHOICE

If, as discussed above, a threshold has to be reached to ensure lineage bias, relative protein levels might be critical for lineage allocation. We inferred the nuclear concentration of NANOG and GATA6 from fluorescence intensity levels in ICMs of *Gata6* wildtype and heterozygous embryos. We found that heterozygous embryos, displaying lower levels of GATA6, had overall fewer PrE cells. This result was independently confirmed by Bessonard et al. (Bessonnard et al., 2014). Interestingly we were also able to demostrate that lineage specification of the PrE was retarded in *Gata6* heterozygous embryos in comparison with wildtype embryos (Publication I, Figure 4). These results demonstrate the importance of protein levels for the second cell fate decision. Lower levels of GATA6 could conceivably require more time to reach the threshold needed to overcome pluripotency and fewer cells would succeed in this process. In this scenario, GATA6 and NANOG are stably coexpressed in the early ICM until stochastic fluctuations lead to one of the two proteins reaching a required threshold level, required to establish a lineage bias. As discussed above, this bias could then be reinforced through FGF signaling or other mechanisms. It would be interesting to test to what extent NANOG levels play a role, and if lower levels of NANOG would concomitantly lower the threshold requirement for GATA6. Analysis of the relative versus absolute protein levels would allow the determination of relevance of the relative concentration of these two proteins and could be carried out in a compound allelic *Gata6;Nanog* series.

It is important to note that both, the analysis of cell fate correlation, as well as protein levels were carried out from still images of fixed embryos. To confirm the inferred results and gain more insight into these processes time lapse imaging of live embryos with subsequent correlation and level analysis would be necessary.

PARALLELS BETWEEN SOMATIC CELL REPROGRAMMING AND EPI/PRE CELL FATE CHOICE

A possible way to generate expression heterogeneity in order to create the phenotypic differences described above is through epigenetic regulation (Wong, 2005). ES cells are thought to maintain a metastable state, ready for differentiation cues, through epigenetic regulation. Similarly, such epigenetic states could play a role in breaking symmetry in the ICM or in reinforcement of the fate decision.

An important epigenetic mechanism to silence genes during differentiation is DNA methylation. During preimplantation development, this is one possible mechanism to establish lineage boundaries. For instance *Elf5*, which together with *Cdx2* and *Eomes* promotes TE development, is silenced through DNA methylation in ES cells but hypomethylated and thereby active in the TS cells. Similarly, based on cell culture models, the pluripotency marker *Stella* is speculated to be hypomethylated in ICM cells but gets methylated and silenced in the EPI upon implantation (Hayashi et al., 2008).

In some instances such repressive DNA methylation marks have to be removed. The highly methylated zygote undergoes massive demethylation before the first cell fate decision to open chromatin and thereby increase plasticity (see Introduction p. 24). Similarly, during reprogramming of somatic cells to iPS cells the epigenetic landscape of the somatic cell needs to be reprogrammed as well to create an open, pluripotent state (Huangfu et al., 2008; Mikkelsen et al., 2008; Shi et al., 2008). In general, active DNA demethylation might allow for dynamic regulation of the epigenetic state of regulatory genes in response to differentiation cues.

We tested the involvement of the DNA demethylation promoting deaminase AID during fate reversion by reprogramming *Aid* null mutant fibroblasts to iPS cells (Publication IV). Interestingly we found that mutant fibroblasts displayed enhanced response to the process of reprogramming, through faster change in morphology, activation of pluripotency markers and numbers of colonies. AID could be permissive for the expression of fate

specific genes and in its absence the differentiated state becomes destabilized. However, the so acquired pluripotent state proved unstable and AID null mutant iPS cells were unable to retain pluripotency over time and instead differentiated. We also found that secondary pluripotency genes were hypermethylated in these cells. It could be speculated that AID is important during a later phase of reprogramming to activate secondary pluripotency genes. AID null cells also failed to upregulate *Dnmt1,* which is responsible for methylation of newly synthesized DNA strands and its exclusion from the nucleus causes DNA demethylation in the early embryo (Bestor, 2000; Howell et al., 2001). Interestingly, when passaging AID null iPS cells a small population was able to retain pluripotency. This observation together with the lack of DNMT1 in these cells could suggest that DNA replication through passaging caused sufficient passive demethylation in some cells to compensate for the absence of AID.

Our results support reports of AID involvement in iPS reprogramming, possibly through demethylation of pluripotency gene promoters (Bhutani et al., 2013; 2010). However, a recent study, attempting reprogramming of *Aid* null mutant fibroblasts, found no differences in efficiency or maintenance of pluripotency in these cells compared to wildtype (Shimamoto et al., 2014). This discrepancy could be explained by strain differences of the parental mice (Shimamoto: C57BL6, Kumar: BALB/c). Mouse strains differ in their genetic background, which can lead to phenotypical variations and can have dramatic effects for ES cell derivation (Baharvand and Matthaei, 2004; Erickson, 1996; Kawase et al., 1994; Simpson et al., 1997; Suzuki et al., 1999). It is possible that the establishment of pluripotency as well as epigenetic regulation during iPS cell reprogramming is similarly strain dependent. iPS reprogramming in wildtype as well as *Aid* null mutant fibroblasts from different mouse strains should be performed to gain a clear understanding of the role of AID in the maintenance of the pluripotent state.

ES cells derived from *Aid* null mutant blastocysts displayed normal morphology and marker expression. However, derivation efficiency was

reduced in *Aid* null compared to wildtype embryos. One could speculate that AID plays a role during a phase of ICM to EPI development, analogous to the late phase in iPS reprogramming. Either way, the varying effects of *Aid* deletion on the maintenance of pluripotency suggest that another mechanism must be in place, potentially cooperating with AID in active demethylation.

Interestingly, computational modeling when used to explain the mechanisms of fate reversion through iPS reprograming reveals interesting parallels between fate reversion and bimodal cell fate choice as described for PrE/EPI (Chang et al., 2008; MacArthur et al., 2008). For both, cells seem inherently robust to reprogramming/differentiation respectively, but amplification of transcriptional noise from several factors and transcriptional feedback loops may be sufficient to trigger activation of the pluripotency/differentiation program. This explains the stochastic behavior of both, reprogramming and PrE/EPI cell fate choice. It also seems that both require higher expression levels of the relevant genes to overcome a necessary threshold, similar to what could be the case for the second cell fate decision.

SEGREGATING PRIMITIVE ENDODERM AND EPIBLAST

GATA6 GOVERNS PRIMITIVE ENDODERM FATE

Once gene expression homogeneity in the ICM is broken, cells progressively commit to their respective lineages. For the PrE this means significant changes and entails the activation of genetic programs, including the activation of transcription factor cascades and signaling pathways, which will eventually lead to the cells' transition to an epithelial extraembryonic endoderm identity (see Introduction p. 10, 15). A good candidate for master regulator of this process seemed to be *Gata6*, since its ectopic expression in ES cells is sufficient for their trans-differentiation to a PrE like state (Fujikura et al., 2002; Shimosato et al., 2007). However, the *Gata6* null mutant was reported to be postimplantation lethal with defects in PrE derivatives, which

suggested a less major role for *Gata6* in the regulation of PrE formation ((Cai et al., 2008; Koutsourakis et al., 1999; Morrisey et al., 1998b), see Introduction p. 13). However, genotyping in these embryos was performed using *in situ* hybridization, which lacks the accuracy of PCR approaches.

To gain a definite understanding of the role of GAT6 in preimplantation development we performed a more extended analysis on a *Gata6* wildtype, heterozygote, and null mutant allelic series to determine the role of GATA6 during the second cell fate decision. In order to gain a comprehensive picture of the phases that make up the second cell fate decision, embryos of various stages were obtained from *Gata6* heterozygous (source: (Sodhi et al., 2006)) intercrosses at sequential time points after mating. No *Gata6* null mutant embryos could be recovered after implantation (E5.5, see also Publication I: Table S1), which pointed to *Gata6* null mutants being peri-implantation rather than postimplantation lethal as previously suggested (see Introduction, p. 13). Subsequently, we analyzed large numbers of embryos during the three phases of the second cell fate decision using immunohistochemistry. GATA6 protein was indeed absent and instead all ICM cells ectopically expressed NANOG. Additionally, markers for the maturation of the PrE, i.e. SOX17, PDGFRA and GATA4, as well as markers for the lineage's polarization and epithelialization, DAB2 and aPKC, could not be detected in *Gata6* null mutant embryos (Publication I, Figures 2A and 3A, Movie S1). Rather, the known pluripotency and EPI markers SOX2 and OCT4, like NANOG, were expressed throughout the ICM (PUBLICATION I, Figure 3B). These results were recently confirmed by Bessonnard et al. (Bessonnard et al., 2014) and put GATA6 at the top of the hierarchy that regulates PrE formation and thereby the second cell fate decision.

It has indeed been suggested that *Gata6* might be upstream of some of the consecutive PrE markers *Sox17*, *Gata4* and *Pdgfra*. Accordingly, GATA binding motifs can be found in enhancer elements of *Pdgfrα* and *Gata4* (Niakan et al., 2010; Wang and Song, 1996). Moreover, in a recent study performing ChIP-Seq for GATA6 in XEN cells and induced XEN cells,

Wamaitha and colleagues showed enrichment of GATA6 near PrE specific genes, like *Gata4*, *Sox17* and *Pdgfra* (Wamaitha et al., 2015). They also demonstrated that extraembryonic endoderm specific genes (like *Gata4* and *Sox17*) displayed more than ten fold upregulation shortly after activation of *Gata6* in ES cells.

It seems, however, that GATA6 alone is not sufficient to activate this cascade. Interestingly, in *Nanog* null mutant embryos, which display an ICM solely comprising GATA6 positive PrE biased cells, it has been shown that these cells stall in PrE differentiation and cannot activate subsequent PrE markers like *Sox17* or *Gata4* (Frankenberg et al., 2011). Thus, *Nanog* is non-cell-autonomously required for PrE formation. Indeed, in the absence of *Nanog*, ICM cells do not express FGF4, which normally fulfills this non-cell-autonomous role by activating the FGF/ERK signaling pathway in neighboring cells and reinforcing their PrE identity (Frankenberg et al., 2011). It is therefore likely that GATA6 and FGF signaling are simultaneously required to activate secondary PrE genes. To understand if this is indeed the case, it would be interesting to ectopically express *Gata6* in *Fgf4* null mutant embryos. This would determine if maintained expression of *Gata6* is sufficient for PrE formation. Interestingly, this was recently demonstrated in *Fgf4* null mutant ES cells, which successfully convert to XEN like cells upon overexpression of *Gata6*, including upregulation of *Sox17* and *Gata4*, suggesting that GATA6 is indeed sufficient to activate these PrE genes (Wamaitha et al., 2015). However, regulation of this process might differ in XEN cells compared to the PrE. Furthermore, Wamaitha et al. showed that *Gata6* overexpression is sufficient to also reprogram neural cells to XEN cells. This underlines the role of GATA6 and points to it being upstream of all essential PrE specifying factors.

Regulation of Gata6 expression

Additionally to activation of PrE markers and repression of *Nanog* (see Introduction p. 19, 21), FGF signaling has also been suggested to activate *Gata6*. However, in the *Fgf4* null mutant *Gata6* expression was observed

initially, but its expression was not maintained beyond the 32-cell stage (Kang et al., 2013). FGF signaling therefore seems to play an additional role in the maintenance of *Gata6* rather than its activation.

Alternatively, it has been speculated that GATA6 could maintain its own expression since canonical GATA binding motifs can be found in the *Gata6* sequence (Frankenberg et al., 2011). Preliminary data using a $Gata6^{H2B\text{-}Venus}$ knock-in reporter allele (generated by Dr. Christian Schröter in Prof. Alfonso Martinez-Arias' lab at the University of Cambridge, UK) shows Venus expression throughout the ICM, even in the absence of GATA6 protein. These data imply that the regulatory machinery necessary for *Gata6* maintenance is not dependent on GATA6 protein (Additional results related to Publication I, Figure 7, p. 55). However, since NANOG is expressed in these cells, this also suggests that NANOG does not transcriptionally repress *Gata6*, and that instead *Gata6* repression is dependent on GATA6 or one of its downstream targets. Nevertheless, this does not exclude the possibility of post-transcriptional or post-translational regulation of GATA6 through NANOG.

A potential GATA factor reinforcing *Gata6* expression could potentially also be GATA4, as *Gata4* overexpression in ES cells is sufficient to stimulate *Gata6* upregulation (Fujikura et al., 2002). It was shown that ERK phosphorylates and activates GATA4 directly in cardiomyocytes (Liang et al., 2001). This could be a mechanism for FGF signaling to indirectly regulate *Gata6* expression through GATA4. Furthermore, SOX17 binds to regulatory regions of *Gata6* and *Gata4* in XEN cells (Niakan et al., 2010). The activation of certain successive PrE markers after *Gata6* could set positive feedback loops, thereby progressively securing commitment to a PrE fate.

Strong evidence also suggests that *Gata6* expression is regulated by NANOG. Inactivation of *Nanog* in cells of the ICM induces *Gata6* expression (Frankenberg et al., 2011) and in ES cells NANOG directly binds the *Gata6* promoter and represses gene expression (Frankenberg et al., 2011; Mitsui et al., 2003; Singh et al., 2007). A study using sodium vanadate to activate

FGFR/ERK signaling in ES cells further showed that while treated cells repressed *Nanog* and upregulated *Gata6*, additional ectopic *Nanog* expression in these cells was sufficient to inhibit *Gata6* upregulation (Hamazaki et al., 2006). These findings highly suggest a repressive role of NANOG for *Gata6*. This suggests that in the absence of NANOG, *Gata6* would be upregulated. In the future it would be interesting to examine if GATA6 is increased in *Nanog* null mutant blastocysts. The absence of FGF4 in these embryos however, could mask a potential effect of lack of NANOG, since FGF/ERK signaling has also been suggested to upregulate *Gata6* (Chazaud et al., 2006; Frankenberg et al., 2011).

This multilayered interdependence of factors regulating gene expression in the ICM makes it extremely difficult to find clear answers. A solution to finally elucidate if FGF4 or NANOG or potentially both are the main regulators of *Gata6* in the ICM, could accrue from single cell protein level analysis of a *Nanog*;*Fgf4* allelic series.

REGULATION OF NANOG EXPRESSION

It is still popular belief in the field that *Gata6* and *Nanog* repress each other and that this mutual repression governs the initiation of the second cell fate decision. While there is strong evidence supporting *Gata6* repression through NANOG, as discussed above, not much has been actually shown to prove direct repression of *Nanog* through GATA6 other than their mutually exclusive expression in the ICM. Indeed, while all ICM cells in the *Gata6* null mutant express NANOG (Publication I, Figure 2A), immunohistochemistry and subsequent measurement of fluorescence intensity levels showed no upregulation of NANOG levels in *Gata6* null embryos in comparison with wildtype embryos (Publication I, Figure 2B). These results show that *Nanog* expression is indeed maintained in the absence of *Gata6* but in contrast to popular belief they also indicate that *Nanog* might not be directly repressed by GATA6. This means another factor has to be in place to downregulate *Nanog* in the ICM even in the absence of *Gata6*.

Nanog, similarly to *Gata6*, is thought to regulate its own expression, as observed in ES cells (Boyer et al., 2005; Fidalgo et al., 2012; Navarro et al., 2012). Imaging of a transcriptional BAC *Nanog:H2B-GFP* reporter (Xenopoulos et al., 2015) on a *Nanog* null mutant (Mitsui et al., 2003) background, shows unperturbed GFP expression. This observation suggests *Nanog* expression does not require NANOG protein. However, this also shows that *Nanog* repression is dependent on NANOG protein or one of its downstream targets but independent of GATA6 protein, which is expressed in all cells of *Nanog* null mutant ICMs.

Indeed, an ideal candidate for *Nanog* regulation has already been identified. Work in ES cells revealed that FGF signaling transcriptionally represses Nanog (Santostefano et al., 2012). This repression seems to occur specifically through the FGF/ERK pathway in these cells and likely in the embryo. Members of the FGF/ERK signaling pathway are reciprocally expressed in the ICM, with EPI cells mainly expressing *Fgf4* and PrE cells expressing its putative receptor *Fgfr2* (Guo et al., 2010; Ohnishi et al., 2014). Furthermore, null mutant embryos for FGF/ERK signaling components display all-EPI ICMs with ectopic expression of *Nanog* (Arman et al., 1998; Chazaud et al., 2006; Feldman et al., 1995; Kang et al., 2013). This evidence points to direct inhibition of *Nanog* in PrE cells through ERK signaling.

Integration of Gata6, Nanog and FGF signaling

Taken together, the FGF/ERK pathway has been implicated in repression of *Nanog,* as well as the activation of not only *Gata6,* but also subsequent factors of the PrE program in the ICM. Indeed, treatment of embryos with exogenous FGF4 leads to the downregulation of *Nanog* in all ICM cells, as well as the upregulation of PrE markers (Nichols et al., 2009; Yamanaka et al., 2010). To assess the effect of FGF signaling on *Gata6* null mutant embryos, we treated them with saturated levels of FGF4 and performed immunohistochemistry for NANOG and PrE markers. As expected, wildtype embryos displayed complete absence of NANOG and ectopic expression of PrE markers in all ICM cells. In contrast, forced activation of

FGF signaling was not sufficient to activate any PrE markers in *Gata6* null mutant embryos. Surprisingly however, NANOG was still present in all cells of the ICM in these embryos (Publication I, Figure 5A-D). This means that in the absence of *Gata6*, FGF signaling is not sufficient to downregulate *Nanog* and therefore reveals that the FGF signaling mediated repression of Nanog is *Gata6* dependent. This unexpected finding sheds more light on the mechanism in which the master regulators GATA6, NANOG and FGF4 cooperate to govern the segregation of PrE and EPI. This newly discovered dependency between GATA6 and FGF signaling can potentially explain the lack of evidence for direct *Nanog* repression through GATA6 in spite of observations pointing to an apparent mutual repression of the two factors. In this new light it is likely that *Nanog* repression through *Gata6* is indeed occurring, but indirectly via FGF signaling rather than directly and in parallel to FGF signaling as assumed before. To conclusively determine if *Gata6* can or cannot directly repress Nanog it would be interesting to ectopically express *Gata6* in *Fgf4* null mutant embryos and establish if *Nanog* is being downregulated in this context or not.

Very recently *Nanog* was shown to be downregulated upon ectopic expression of *Gata6* in *Fgf4* null mutant ES cells (Wamaitha et al., 2015). This result points to GATA6 being sufficient to downregulate *Nanog* in an ES cell context. However, the possibility remains that this regulation takes place differently *in vivo* than *in vitro*. XEN cells mostly represent a derivative of the PrE, the parietal endoderm (Brown et al., 2010; Kunath, 2005), in which additional genes are activated and gene regulation might differ substantially from that of immature PrE cells. The same experiments have to be performed in the embryo to interpret these results in an *in vivo* context. In one possible scenario ERK could be lowering the relative threshold needed for *Gata6* to overcome *Nanog in vivo*. Ectopic expression of *Gata6* could exceed this threshold on its own *in vitro*.

POSSIBLE MECHANISMS FOR GATA6/NANOG/FGF INTERPLAY

To further investigate the position of *Gata6* in the FGF signaling pathway, we treated embryos with a potent ERK inhibitor, which as expected led to all ICM cells expressing dramatically elevated NANOG levels and complete downregulation of PrE markers in wildtype embryos. Interestingly, the same observation was made in *Gata6* null mutant embryos (Publication I, Figure 5A-D). Therefore, since ERK can still directly affect *Nanog* expression, GATA6 has to act upstream of ERK. It is likely that GATA6 acts on the level of either the FGF receptor or *Grb2*, since those are the most specific to the PrE and show correlating expression with *Gata6* (Guo et al., 2010; Ohnishi et al., 2014). A recently published mathematical model supports a scenario where GATA6 acts upstream of *Fgfr2* (Bessonnard et al., 2014). The ideal approach to further test this hypothesis would be chromatin immunoprecipitation (ChIP) of GATA6 in PrE cells, as well as luciferase assays using promoter regions of the candidate FGF signaling components. Unfortunately, it is nearly impossible to harvest sufficient material from blastocyst stage embryos for biochemical assays of this kind. Alternatively, ChIP has recently been performed in XEN cells, where GATA6 enrichment was indeed found upstream of *Fgfr2* (Wamaitha et al., 2015). Unfortunately, this only provides partial evidence since XEN cells may not entirely recapitulate the *in vivo* state. Yet another approach to narrow down possible targets would be treatment of *Gata6* null mutant embryos with inhibitors for different components of the FGF signaling pathway, especially with inhibitors targeting FGF receptors or GRB2. These experiments are however further complicated by the diversity of FGF receptors (see Introduction p. 14) in combination with the lack of specificity of many of the available inhibitors.

If FGFR2 is indeed the predominant receptor relaying FGF/ERK signaling in the ICM then *Gata6* cannot be genetically upstream of *Fgfr2.* If this were the case, it would completely abolish FGF signaling in the absence of *Gata6* and cause upregulation of *Nanog* in the ICM. Instead, NANOG levels were not elevated in *Gata6* null mutant ICMs (Publication I, Figure 2B), which

means marginal ERK signaling levels have to be active even in the EPI and in *Gata6* null mutants, which both display absence of GATA6. This suggests that FGF/ERK signaling plays a cell-autonomous role in EPI cells to maintain NANOG at physiological levels, presumably to ensure the ability of these cells to readily respond to differentiation cues after implantation (Chambers et al., 2003; Hatano et al., 2005; Mitsui et al., 2003; Shin et al., 2011; Suzuki et al., 2006; Vallier et al., 2009). Regulative FGF signaling in ES and EPI cells would also explain why *Nanog* levels are very heterogeneous in these cells and become dramatically upregulated upon treatment with ERK inhibitors. Active FGF signaling in the absence of Gata6 could instead point to GATA6 acting on FGFR2 in a post-transcriptional or post-translational manner, for example by stabilizing the protein.

Alternatively, another receptor or receptor isoform could be in play. It is generally assumed that FGFR2 is the predominant receptor that plays an active role during EPI and PrE segregation since its knockout leads to peri-implantation lethality (Arman et al., 1998), and resembles the *Fgf4* or *Grb2* null mutants. *Fgfr2* deletion has been attempted by several groups and while Arman et al. report peri-implantation lethality (Arman et al., 1998), Xu et al. and Yu et al. both report lethality between E10.5 and E11.5 (Yu 2003, Xu 1998) , which would attribute a far lesser role to *Fgfr2* than previously thought. Indeed, as mentioned before (see Introduction p. 15), *Fgfr1* is also expressed in the ICM, but does not acquire reciprocal expression in PrE and EPI cells as *Fgfr2* does but is expressed similarly in both cell types (Ohnishi et al., 2014). Unfortunately, it is yet unknown which isoforms of the two receptors are expressed. Both FGFR1 and FGFR2 have several isoforms, which show differential specificity for FGF4 binding (Ornitz et al., 1996). FGFR1b and 2b both display low FGF4 binding, while FGFR1c and 2c show high specificity (Ornitz et al., 1996). The continuous expression of *Fgfr1* in EPI cells could account for the continued marginally active FGF signaling in these cells if the 1b isoform is expressed.

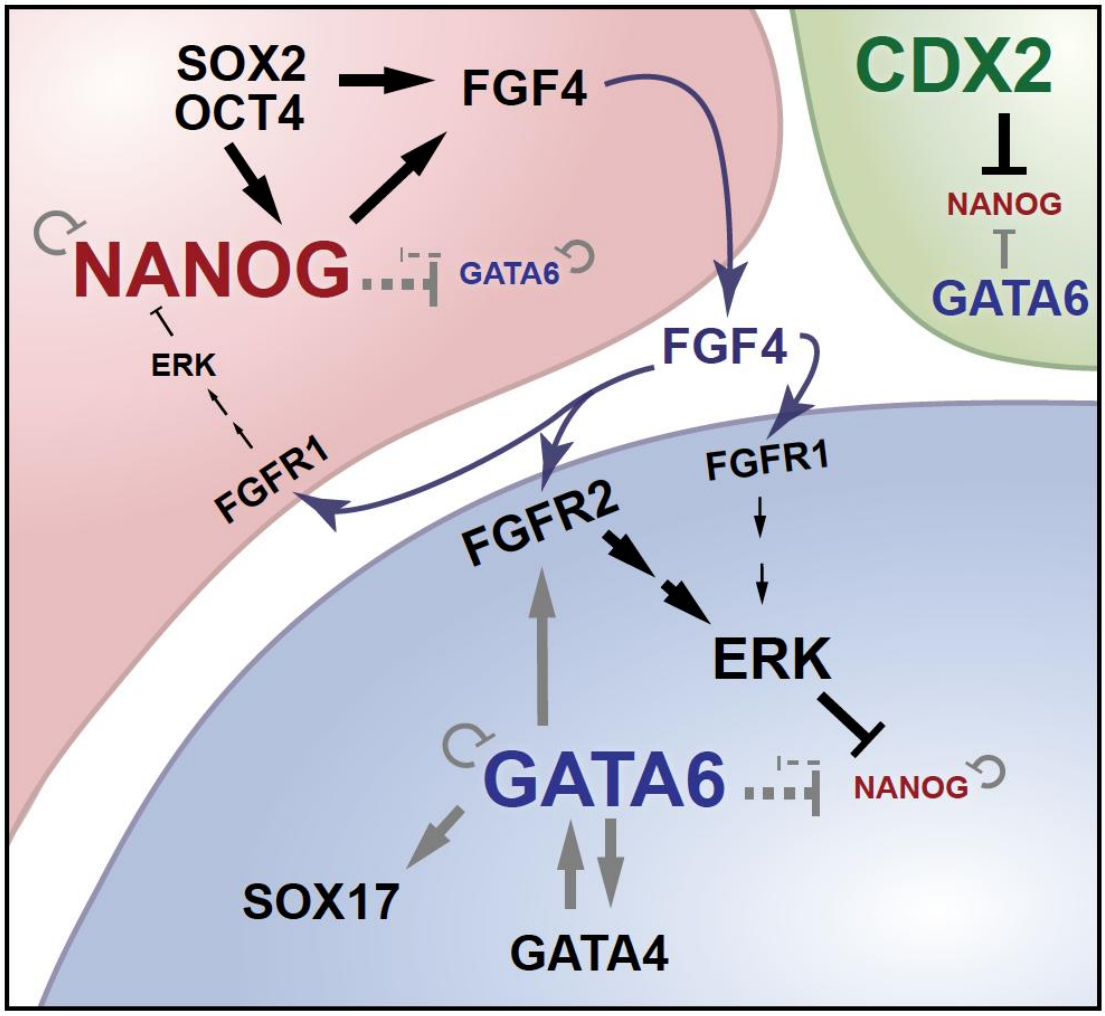


Figure 8: Model for the segregation of EPI and PrE. GATA6 acts upstream of *Fgfr2*, enhancing FGF/ERK signaling **in PrE biased cells**. Signaling through FGFR2 in concert with FGFR1 is sufficient to repress *Nanog*, which in turn releases *Gata6* repression. This feedback loop reinforces the PrE fate in addition to activation of the PrE specific transcription factors *Gata4* and *Sox17* through GATA6. GATA4 may reinforce *Gata6* expression. **In EPI biased cells** weaker signaling through only FGFR1 allows for moderation of NANOG levels rather than repression. Both NANOG and GATA6 are able to repress their own expression. Mutual repression of GATA6 and NANOG likely takes place indirectly. **In TE cells** GATA6 and CDX2 repress *Nanog*. Refer to text for further information. Green: TE. Blue: PrE. Red: EPI. Arrows: positive regulation. Blocked arrows: repression. Grey arrows: speculated interactions. Dashed arrows: indirect interaction.

Alternatively or in addition, the receptors could rely on cooperative activation. Preliminary data suggests that while both *Fgfr1* and *Fgfr2* null mutant embryos show postimplantation lethality, only embryos depleted for both genes phenocopy the *Fgf4* null mutant (Ms. Minjung Kang – Hadjantonakis Lab, personal communication of unpublished data). This suggests that FGFR1 and FGFR2 cooperate to relay FGF/ERK signaling in

the ICM. In that case FGFR1 would be active in both cell types, possibly only moderately downregulating *Nanog* because both receptors are needed for full activity. Since their expression is highly correlated, *Gata6* is likely upstream of only *Fgfr2*, which would lead to additional downregulation of *Nanog* in *Gata6* expressing cells, i.e. PrE cells. Indeed, ChIP in XEN cells has shown enrichment of GATA6 upstream the *Fgfr2* gene, but not *Fgfr1* (Wamaitha et al., 2015). Testing this hypothesis *in vivo* is similarly problematic as before (see p. 134) and either specific inhibitors for the different receptors have to be found or novel genetic approaches utilized. Double null mutants for either receptor and *Gata6* would give additional insights. Alternatively, ectopic expression of either receptor on a *Gata6* null mutant background could give answers in a more straightforward fashion. If GATA6 indeed acts upstream of one of the two receptors, forced expression of the receptor would result in a rescue of the *Gata6* mutant cells' response to FGF signaling.

Interestingly, Wamaitha and colleagues recently found that FGF signaling does not seem necessary for reprogramming of ES cells to XEN cells through overexpression of *Gata6*, which involves downregulation of *Nanog* (Wamaitha et al., 2015). This is yet another discrepancy between *in vitro* and *in vivo* results and underlines the scrutiny that should be applied when interpreting such data.

FURTHER ROLES OF GATA6

We have shown that deletion of *Gata6* results in embryos lacking the PrE lineage and that downregulation of *Nanog* through FGF signaling in the ICM is dependent on GATA6 (Publication I)). We concluded that *Gata6* is the first single transcription factor deletion resulting in a complete failure of PrE formation and maturation. However, another transcription factor deletion has recently been shown to result in ablation of the PrE. The pluripotency marker *Oct4*, which becomes restricted to EPI cells in the late blastocyst, also seems to be necessary for PrE formation, as its deletion in embryos leads to a failure in PrE specification (Frum et al., 2013; Le Bin et al., 2014). Moreover, *Nanog* was not repressed upon treatment with FGF4 in these embryos (Frum et al.,

2013), similarly to *Gata6* null mutant embryos. The authors concluded a function for *Oct4* in PrE formation. However, their work also showed that *Oct4* null mutant embryos initially express *Gata6* but fail to maintain it beyond the mid blastocyst stage. This means GATA6 is lost during the window of FGF treatment and its absence could readily explain the cell's lack of response to FGF signaling, as we have shown that the FGF signaling mediated downregulation of *Nanog* is *Gata6* dependent. Ectopic expression of *Gata6* in *Oct4* null mutant embryos, with the expectation to rescue the cell's response to FGF treatment, would be necessary to confirm this hypothesis. Similarly, *Oct4* was shown to be dispensable for reprogramming of ES cells to a XEN like state through ectopic expression of *Gata6*, which involved upregulation of PrE markers like *Sox17* and *Gata4* (Wamaitha et al., 2015).

In other respects, *Gata6* also seems involved in TE specification. We found ectopic *Nanog* expression in the TE of *Gata6* null mutant embryos (Publication I, Figure 2C), suggesting that GATA6 is involved in repression of *Nanog* in the TE under wildtype conditions. In this case, *Nanog* repression through GATA6 functions differently in the TE than in the ICM. As discussed above, in the absence of *Gata6*, ICM cells fail to respond to FGF signaling which is responsible for *Nanog* repression in these cells. In contrast, in the TE FGF signaling fulfills a mitogenic role (see Introduction p. 9) and *Fgf4* null mutant embryos have fewer TE cells (Kang et al., 2013). Taking into account that TE size was not affected in *Gata6* null mutant embryos, it can be concluded that FGF signaling in the TE lineage is intact even in the absence of GATA6. Ectopic expression of *Nanog* in these cells therefore points to GATA6 regulating *Nanog* independently of FGF signaling in the TE. This presumably takes place in cooperation with CDX2, which also has been shown to repress *Nanog* in TE cells (see Introduction p. 9).

Regulation of ICM cell plasticity through the Gata6/Nanog/FGF network

By the time of implantation the ratio of EPI and PrE inside the ICM is surprisingly consistent. As can be seen in Figure 8, for CD1 mice this ratio lays at around 40% EPI and 60% PrE (Publication I, Figure 4B; Additional results related to Publication I, Figure 8, p. 55) and therefore seems tightly regulated. Several mechanisms could contribute to this regulation. It has been shown that a subpopulation of ICM cells remains interconvertible between EPI and PrE for an extended period. These cells stay plastic, not merely during the process of initial cell fate specification, but up until the late blastocyst stage when the two cell types are seemingly committed (Grabarek et al., 2012). This becomes especially evident when cells are pushed towards either lineage by treatment with FGF4 or ERK inhibitor (Yamanaka et al., 2010). These studies showed that a change in an ICM cell's environment is able to direct it towards either lineage until the late blastocyst stage.

Plusa and colleagues also demonstrated that this extended plasticity is necessary during blastocyst development. By following the expression of the PrE specific reporter *Pdgfra*$^{H2B\text{-}GFP}$ they showed that cells that do not reach their appropriate position through sorting in the late blastocyst could change their fate to match their neighbors' (Plusa et al., 2008). It is understood that FGF signaling is involved in this plasticity (Yamanaka et al., 2010) but how cells retain or loose plasticity remains unknown. It seems likely that the cells' inherent ability to respond to FGF signaling, rather than a change in its environment, is responsible for this feature. When treating embryos with exogenous FGF4 during different time frames, we also noted that cells reacted variably, with later treatments causing less cells to downregulate *Nanog* and adopt a PrE fate (Publication I, Figure 6). These observations confirmed the gradual loss of plasticity inside the ICM, which seems to be accompanied by reduced ability to respond to FGF signaling.

As we also demonstrated that the presence of GATA6 was necessary for cells in the ICM to respond to FGF signaling. It could therefore be

speculated that cells retain their ability to respond to FGF signaling until they are fully committed to an EPI fate and GATA6 is absent. Subsequently those cells would be protected from a forced fate switch inside the FGF4 rich, sorted EPI compartment of the implanting blastocyst. Work by Grabarek and colleagues, which showed that PrE cells remain plastic longer than EPI cells, which seemed more sheltered, supports this hypothesis (Grabarek et al., 2012).

CONCLUSION AND PERSPECTIVES

The present work has given many new insights into the mechanisms that govern the second cell fate decision and the establishment as well as maintenance of pluripotency. Foremost, GATA6 has been shown to be a major regulator of the second cell fate decision and it was integrated into the regulatory network governing this process, by revealing that the FGF-mediated downregulation of *Nanog* is GATA6 dependent. This discovery elucidates one step of the complex interactions in the interdependent network of GATA6, NANOG and FGF. In the future it will be interesting to find how GATA6 acts on the FGF signaling pathway in this context, which might give insight into the same kind of interaction in other fields, as in gastric cancer for instance, where both FGFR2 as well as GATA6 have been suggested as targets (Deng et al., 2012).

It was further shown that GATA6 acts in a dosage dependent manner and a study using a *Gata6;Nanog* allelic series, currently taking place, will reveal if the relative dosages of GATA6 and NANOG are imperative for the lineage decision between PrE and EPI. A dosage dependent lineage decision would also support the hypothesis that the emergence of the two lineages is stochastic, since expression noise could lead to one factor stochastically reaching a dosage threshold and initiating differentiation. More experiments, especially live imaging of reporters for both markers will be necessary to understand this process.

This work has also provided efficient protocols for the derivation of ES and XEN cells, which can serve as models for EPI and PrE respectively. For

instance, it was shown that ES cells could be derived from AID null mutant embryos, though with reduced efficiency. Furthermore iPS cells generated from fibroblasts lacking AID exhibited unstable pluripotency states. Both results suggest the involvement of DNA demethylation in the establishment and maintenance of pluripotency, but also point to AID possibly cooperating with other factors, e.g. TET proteins to fulfill this function. It will be interesting to gain further insight into this level of regulation of lineage decisions in the embryo.

This work has provided answers as well as newly opened questions. Future studies will elucidate how transcription factors, signaling pathways and epigenetic regulation cooperate to allow the proper allocation of PrE and EPI.

4. REFERENCES

Abe, Y., Minegishi, T., and Leung, P.C.K. (2004). Activin receptor signaling. Growth Factors *22*, 105–110.

Ambrosetti, D.C., Scholer, H.R., Dailey, L., and 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-4Enhancer. Journal of Biological Chemistry *275*, 23387–23397.

Anani, S., Bhat, S., Honma-Yamanaka, N., Krawchuk, D., and Yamanaka, Y. (2014). Initiation of Hippo signaling is linked to polarity rather than to cell position in the pre-implantation mouse embryo. Development *141*, 2813–2824.

Arceci, R.J., King, A.A., Simon, M.C., Orkin, S.H., and Wilson, D.B. (1993). Mouse GATA-4: a retinoic acid-inducible GATA-binding transcription factor expressed in endodermally derived tissues and heart. Molecular and Cellular Biology *13*, 2235–2246.

Arman, E., Haffner-Krausz, R., Chen, Y., Heath, J.K., and Lonai, P. (1998). Targeted disruption of fibroblast growth factor (FGF) receptor 2 suggests a role for FGF signaling in pregastrulation mammalian development. Proceedings of the National Academy of Sciences *95*, 5082–5087.

Artus, J., Panthier, J.J., and Hadjantonakis, A.K. (2010). A role for PDGF signaling in expansion of the extra-embryonic endoderm lineage of the mouse blastocyst. Development *137*, 3361–3372.

Artus, J., Piliszek, A., and Hadjantonakis, A.-K. (2011). The primitive endoderm lineage of the mouse blastocyst: Sequential transcription factor activation and regulation of differentiation by Sox17. Developmental Biology *350*, 393–404.

Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. Genes & Development *17*, 126–140.

Baharvand, H., and Matthaei, K.I. (2004). Culture condition difference for establishment of new embryonic stem cell lines from the C57BL/6 and BALB/c mouse strains. In Vitro Cell. Dev. Biol. Anim. *40*, 76–81.

Balakier, H., and Pedersen, R.A. (2003). Allocation of Cells to Inner Cell Mass and Trophectoderm Lineages in Preimplantation Mouse Embryos'. Developmental Biology *90*, 352–362.

Bangs, F.K., Schrode, N., Hadjantonakis, A.-K., and Anderson, K.V. (2015). Lineage specificity of primary cilia in the mouse embryo. Nat Cell Biol *17*, 113–122.

Basilico, C., Ambrosetti, D., Fraidenraich, D., and Dailey, L. (1997). Regulatory mechanisms governing FGF-4 gene expression during mouse development. J. Cell. Physiol. *173*, 227–232.

Bessonnard, S., De Mot, L., Gonze, D., Barriol, M., Dennis, C., Goldbeter, A., Dupont, G., and Chazaud, C. (2014). Gata6, Nanog and Erk signaling control cell fate in the inner cell mass through a tristable regulatory network. Development *141*, 3637–3648.

Bestor, T.H. (2000). The DNA methyltransferases of mammals. Human Molecular Genetics *9*, 2395–2402.

Bhutani, N., Decker, M.N., Brady, J.J., Bussat, R.T., Burns, D.M., Corbel, S.Y., and Blau, H.M. (2013). A critical role for AID in the initiation of reprogramming to induced pluripotent stem cells. The FASEB Journal *27*, 1107–1113.

Bhutani, N., Brady, J.J., Damian, M., Sacco, A., Corbel, S.Y., and Blau, H.M. (2010). Reprogramming towards pluripotencyrequires AID-dependent DNAdemethylation. Nature *463*, 1042–1047.

Bird, A.P., and Wolffe, A.P. (1999). Methylation-induced repression--belts, braces, and chromatin. Cell *99*, 451–454.

Blake, W.J., KAErn, M., Cantor, C.R., and Collins, J.J. (2003). Noise in eukaryotic gene expression. Nature *422*, 633–637.

Blij, S., Parenti, A., Tabatabai-Yazdi, N., and Ralston, A. (2015). Cdx2 efficiently induces trophoblast stem-like cells in naïve, but not primed, pluripotent stem cells. Stem Cells and Development *24*, 1352–1365.

Boroviak, T., Loos, R., Bertone, P., Smith, A., and 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.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. Cell *122*, 947–956.

Bradley, A., and Robertson, E. (1986). Embryo-derived stem cells: a tool for elucidating the developmental genetics of the mouse. Curr. Top. Dev. Biol. *20*, 357–371.

Brook, F.A., and Gardner, R.L. (1997). The origin and efficient derivation of embryonic stem cells in the mouse. Proceedings of the National Academy of Sciences *94*, 5709–5712.

Brown, K., Legros, S., Artus, J., Doss, M.X., Khanin, R., Hadjantonakis, A.-K., and Foley, A. (2010). A Comparative Analysis of Extra-Embryonic Endoderm Cell Lines. PLoS ONE *5*, e12016.

Buehr, M., Meek, S., Blair, K., Yang, J., Ure, J., Silva, J., McLay, R., Hall, J., Ying, Q.-L., and Smith, A. (2008). Capture of authentic embryonic stem cells from rat blastocysts. Cell *135*, 1287–1298.

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

Cai, K.Q., Capo-Chichi, C.D., Rula, M.E., Yang, D.-H., and Xu, X.-X. (2008). Dynamic GATA6 expression in primitive endoderm formation and maturation in early mouse embryogenesis. Dev. Dyn. *237*, 2820–2829.

Capo-Chichi, C.D., Rula, M.E., Smedberg, J.L., Vanderveer, L., Parmacek, M.S., Morrisey, E.E., Godwin, A.K., and Xu, X.-X. (2005). Perception of differentiation cues by GATA factors in primitive endoderm lineage determination of mouse embryonic stem cells. Developmental Biology *286*, 574–586.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and 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., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. Nature *450*, 1230–1234.

Chang, H.H., Hemberg, M., Barahona, M., Ingber, D.E., and Huang, S. (2008). Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. Nature *453*, 544–547.

Chazaud, C., Yamanaka, Y., Pawson, T., and Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Devcel *10*, 615–624.

Cheng, A.M., Saxton, T.M., Sakai, R., Kulkarni, S., Mbamalu, G., Vogel, W., Tortorice, C.G., Cardiff, R.D., Cross, J.C., Muller, W.J., et al. (1998). Mammalian Grb2 regulates multiple steps in embryonic development and malignant transformation. Cell *95*, 793–803.

Chew, J.L., Loh, Y.H., Zhang, W., Chen, X., Tam, W.L., Yeap, L.S., Li, P., Ang, Y.S., Lim, B., Robson, P., et al. (2005). Reciprocal Transcriptional Regulation of Pou5f1 and Sox2 via the Oct4/Sox2 Complex in Embryonic Stem Cells. Molecular and Cellular Biology *25*, 6031–6046.

Cho, L.T.Y., Wamaitha, S.E., Tsai, I.J., Artus, J., Sherwood, R.I., Pedersen, R.A., Hadjantonakis, A.-K., and Niakan, K.K. (2012). Conversion from mouse embryonic to extra-embryonic endoderm stem cells reveals distinct differentiation capacities of pluripotent stem cell states. Development *139*, 2866–2877.

Ciruna, B.G., and Rossant, J. (1999). Expression of the T-box gene Eomesodermin during early mouse development. Mod *81*, 199–203.

Cockburn, K., Biechele, S., Garner, J., and Rossant, J. (2013). The Hippo Pathway Member Nf2 Is Required for Inner Cell Mass Specification. Current Biology 23, 1195–1201.

Cortázar, D., Kunz, C., Saito, Y., Steinacher, R., and Schär, P. (2007). The enigmatic thymine DNA glycosylase. DNA Repair (Amst.) *6*, 489–504.

Coucouvanis, E., and Martin, G.R. (1999). BMP signaling plays a role in visceral endoderm differentiation and cavitation in the early mouse embryo. Development *126*, 535–546.

Cross, S.H., and Bird, A.P. (1995). CpG islands and genes. Current Opinion in Genetics & Development *5*, 309–314.

Dailey, L., Ambrosetti, D., Mansukhani, A., and Basilico, C. (2005). Mechanisms underlying differential responses to FGF signaling. Cytokine & Growth Factor Reviews *16*, 233–247.

Dard, N., Le, T., Maro, B., and Louvet-Vallee, S. (2009). Inactivation of aPKCλ Reveals a Context Dependent Allocation of Cell Lineages in Preimplantation Mouse Embryos. PLoS ONE *4*, e7117.

Davies, T.J., and Fairchild, P.J. (2012). Optimization of protocols for derivation of mouse embryonic stem cell lines from refractory strains, including the non obese diabetic mouse. Stem Cells and Development *21*, 1688–1700.

Dean, W., Santos, F., Stojkovic, M., Zakhartchenko, V., Walter, J., Wolf, E., and Reik, W. (2001). Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. Proceedings of the National Academy of Sciences *98*, 13734–13738.

Deng, N., Goh, L. K., Wang, H., Das, K., Tao, J., Tan, I. B., et al. (2012). A comprehensive survey of genomic alterations in gastric cancer reveals systematic patterns of molecular exclusivity and co-occurrence among distinct therapeutic targets. Gut, 61, 673–684.

Deng, X.-Y., Wang, H., Wang, T., Fang, X.-T., Zou, L.-L., Li, Z.-Y., and Liu, C.-B. (2015). Non-viral methods for generating integration-free, induced pluripotent stem cells. Curr Stem Cell Res Ther *10*, 153–158.

Dietrich, J.-E., and Hiiragi, T. (2007). Stochastic patterning in the mouse pre-implantation embryo. Development *134*, 4219–4231.

Ducibella, T., and Anderson, E. (2003). Cell shape and membrane changes in the eight-cell mouse embryo: Prerequisites for morphogenesis of the blastocyst. Developmental Biology *47*, 45–58.

Eiselleova, L., Peterkova, I., Neradil, J., Slaninova, I., Hampl, A., and Dvorak, P. (2008). Comparative study of mouse and human feeder cells for human embryonic stem cells. Int. J. Dev. Biol. *52*, 353–363.

Erickson, R.P. (1996). Mouse models of human genetic disease: which mouse is more like a man? Bioessays *18*, 993–998.

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

Feldman, B., Poueymirou, W., Papaioannou, V.E., DeChiara, T.M., and Goldfarb, M. (1995). Requirement of FGF-4 for postimplantation mouse development. Science *267*, 246–249.

Ferrell, J.E., Jr (2002). Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. Current Opinion in Cell Biology *14*, 140–148.

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

Fierro-González, J.C., White, M.D., Silva, J.C., and Plachta, N. (2013). Cadherin-dependent filopodia control preimplantation embryo compaction. Nat Cell Biol *15*, 1424–1433.

Francois, M., Koopman, P., and Beltrame, M. (2010). SoxF genes: Key players in the development of the cardio-vascular system. The International Journal of Biochemistry & Cell Biology *42*, 445–448.

Frankenberg, S., Gerbe, F., Bessonnard, S., Belville, C., Pouchin, P., Bardot, O., and Chazaud, C. (2011). Primitive Endoderm Differentiates via a Three-Step Mechanism Involving Nanog and RTK Signaling. Developmental Cell *21*, 1005–1013.

Freyer, L., Schroeter, C., Saiz, N., Schrode, N., Nowotschin, S., Martinez-Arias, A., Hadjantonakis, A.-K. (submitted). A loss-of-function and H2B-Venus transcriptional reporter allele for *Gata6* in mice. BMC Developmental Biology.

Frum, T., Halbisen, M.A., Wang, C., Amiri, H., Robson, P., and Ralston, A. (2013). Oct4 Cell-Autonomously Promotes Primitive Endoderm Development in the Mouse Blastocyst. Developmental Cell *25*, 610–622.

Fujikura, J., Yamato, E., Yonemura, S., Hosoda, K., Masui, S., Nakao, K., Miyazaki Ji, J.-I., and Niwa, H. (2002). Differentiation of embryonic stem cells is induced by GATA factors. Genes & Development *16*, 784–789.

Gerbe, F., Cox, B., Rossant, J., and Chazaud, C. (2008). Dynamic expression of Lrp2 pathway members reveals progressive epithelial differentiation of primitive endoderm in mouse blastocyst. Developmental Biology *313*, 594–602.

Gertsenstein, M., Nutter, L.M.J., Reid, T., Pereira, M., Stanford, W.L., Rossant, J., and Nagy, A. (2010). Efficient Generation of Germ Line

Transmitting Chimeras from C57BL/6N ES Cells by Aggregation with Outbred Host Embryos. PLoS ONE *5*, e11260.

Ghatpande, S., Brand, T., Zile, M., and Evans, T. (2006). Bmp2 and Gata4 function additively to rescue heart tube development in the absence of retinoids. Dev. Dyn. *235*, 2030–2039.

Goldin, S.N., and Papaioannou, V.E. (2003). Paracrine action of FGF4 during periimplantation development maintains trophectoderm and primitive endoderm. Genesis *36*, 40–47.

Grabarek, J.B., Zyzyńska, K., Saiz, N., Piliszek, A., Frankenberg, S., Nichols, J., Hadjantonakis, A.-K., and Plusa, B. (2012). Differential plasticity of epiblast and primitive endoderm precursors within the ICM of the early mouse embryo. Development *139*, 129–139.

Graham, S.J.L., Wicher, K.B., Jedrusik, A., Guo, G., Herath, W., Robson, P., and Zernicka-Goetz, M. (2014). BMP signalling regulates the pre-implantation development of extra-embryonic cell lineages in the mouse embryo. Nature Communications *5*, 5667.

Guo, G., Huss, M., Tong, G.Q., Wang, C., Sun, L.L., Clarke, N.D., and Robson, P. (2010). Resolution of Cell Fate Decisions Revealed by Single-Cell Gene Expression Analysis from Zygote to Blastocyst. Developmental Cell *18*, 675–685.

Guo, J.U., Su, Y., Zhong, C., Ming, G.-L., and Song, H. (2011). Emerging roles of TET proteins and 5-hydroxymethylcytosines in active DNA demethylation and beyond. Cc *10*, 2662–2668.

Hamazaki, T., Kehoe, S.M., Nakano, T., and Terada, N. (2006). The Grb2/Mek Pathway Represses Nanog in Murine Embryonic Stem Cells. Molecular and Cellular Biology *26*, 7539–7549.

Hamilton, T.G., Klinghoffer, R.A., Corrin, P.D., and Soriano, P. (2003). Evolutionary Divergence of Platelet-Derived Growth Factor Alpha Receptor Signaling Mechanisms. Molecular and Cellular Biology *23*, 4013–4025.

Hancock, S.N., Agulnik, S.I., Silver, L.M., and Papaioannou, V.E. (1999). Mapping and expression analysis of the mouse ortholog of Xenopus Eomesodermin. Mod *81*, 205–208.

Handyside, A.H. (1978). Time of commitment of inside cells isolated from preimplantation mouse embryos. J Embryol Exp Morphol *45*, 37–53.

Hanna, J.H., Saha, K., and Jaenisch, R. (2010). Pluripotency and Cellular Reprogramming: Facts, Hypotheses, Unresolved Issues. Cell *143*, 508–525.

Hart, A.H., Hartley, L., Ibrahim, M., and Robb, L. (2004). Identification, cloning and expression analysis of the pluripotency promoting Nanog genes in mouse and human. Dev. Dyn. *230*, 187–198.

Hasegawa, Y., Takahashi, N., Forrest, A.R.R., Shin, J.W., Kinoshita, Y., Suzuki, H., and Hayashizaki, Y. (2011). CC Chemokine Ligand 2 and Leukemia Inhibitory Factor Cooperatively Promote Pluripotency in Mouse Induced Pluripotent Cells. Stem Cells *29*, 1196–1205.

Hatano, S.-Y., Tada, M., Kimura, H., Yamaguchi, S., Kono, T., Nakano, T., Suemori, H., Nakatsuji, N., and Tada, T. (2005). Pluripotential competence of cells associated with Nanog activity. Mech. Dev. *122*, 67–79.

Hayashi, K., Lopes, S.M.C. de S., Tang, F., and Surani, M.A. (2008). Dynamic Equilibrium and Heterogeneity of Mouse Pluripotent Stem Cells with Distinct Functional and Epigenetic States. Cell Stem Cell *3*, 391–401.

Heller, L.C., Li, Y., Abrams, K.L., and Rogers, M.B. (1999). Transcriptional regulation of the Bmp2 gene. Retinoic acid induction in F9 embryonal carcinoma cells and Saccharomyces cerevisiae. J. Biol. Chem. *274*, 1394–1400.

Hillman, N., Sherman, M.I., and Graham, C. (1972). The effect of spatial arrangement on cell determination during mouse development. J Embryol Exp Morphol *28*, 263–278.

Hirai, H., Karian, P., and Kikyo, N. (2011). Regulation of embryonic stem cell self-renewal and pluripotency by leukaemia inhibitory factor. Biochem. J. *438*, 11–23.

Hirate, Y., Hirahara, S., Inoue, K.-I., Suzuki, A., Alarcon, V.B., Akimoto, K., Hirai, T., Hara, T., Adachi, M., Chida, K., et al. (2013). Polarity-Dependent Distribution of Angiomotin Localizes Hippo Signaling in Preimplantation Embryos. Current Biology 1–14.

Hochedlinger, K., and Jaenisch, R. (2003). Nuclear transplantation, embryonic stem cells, and the potential for cell therapy. N. Engl. J. Med. *349*, 275–286.

Howell, C.Y., Bestor, T.H., Ding, F., Latham, K.E., Mertineit, C., Trasler, J.M., and Chaillet, J.R. (2001). Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. Cell *104*, 829–838.

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

Hyafil, F., Morello, D., Babinet, C., and Jacob, F. (1980). A cell surface glycoprotein involved in the compaction of embryonal carcinoma cells and cleavage stage embryos. Cell *21*, 927–934.

Inoue, H., Nagata, N., Kurokawa, H., and Yamanaka, S. (2014). iPS cells: a game changer for future medicine. The EMBO Journal *33*, 409–417.

Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-

renewal and inner cell mass specification. Nature *466*, 1129–1133.

Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. A., Chuan, H., and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science. *333*, 1300–1303.

Jaenisch, R., and Young, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell *132*, 567–582.

Johnson, M.H., and Ziomek, C.A. (1981a). Induction of polarity in mouse 8-cell blastomeres: specificity, geometry, and stability. The Journal of Cell Biology *91*, 303–308.

Johnson, M.H., and Ziomek, C.A. (1981b). The foundation of two distinct cell lineages within the mouse morula. Cell *24*, 71–80.

Johnson, M.H., Maro, B., and Takeichi, M. (1986). The role of cell adhesion in the synchronization and orientation of polarization in 8-cell mouse blastomeres. J Embryol Exp Morphol *93*, 239–255.

Jones, P.A., and Takai, D. (2001). The role of DNA methylation in mammalian epigenetics. Science *293*, 1068–1070.

Julio, M.K.D., Alvarez, M.J., Galli, A., Chu, J., Price, S.M., Califano, A., and Shen, M.M. (2011). Regulation of extra-embryonic endoderm stem cell differentiation by Nodal and Cripto signaling. Development *138*, 3885–3895.

Kadokawa, Y., Kato, Y., and Eguchi, G. (1987). Cell lineage analysis of the primitive and visceral endoderm of mouse embryos cultured in vitro. Cell Differ. *21*, 69–76.

Kalmar, T., Lim, C., Hayward, P., Muñoz-Descalzo, S., Nichols, J., Garcia-Ojalvo, J., and Martinez-Arias, A. (2009). Regulated Fluctuations in Nanog Expression Mediate Cell Fate Decisions in Embryonic Stem Cells. PLoS Biol *7*, e1000149.

Kanai-Azuma, M., Kanai, Y., Gad, J.M., Tajima, Y., Taya, C., Kurohmaru, M., Sanai, Y., Yonekawa, H., Yazaki, K., Tam, P.P.L., et al. (2002). Depletion of definitive gut endoderm in Sox17-null mutant mice. Development *129*, 2367–2379.

Kang, M., Piliszek, A., Artus, J., and Hadjantonakis, A.-K. (2013). FGF4 is required for lineage restriction and salt-and-pepper distribution of primitive endoderm factors but not their initial expression in the mouse. Development *140*, 267–279.

Kawase, E., Suemori, H., Takahashi, N., Okazaki, K., Hashimoto, K., and Nakatsuji, N. (1994). Strain difference in establishment of mouse embryonic stem (ES) cell lines. Int. J. Dev. Biol. *38*, 385–390.

Keramari, M., Razavi, J., Ingman, K.A., Patsch, C., Edenhofer, F., Ward, C.M., and Kimber, S.J. (2010). Sox2 Is Essential for Formation of

Trophectoderm in the Preimplantation Embryo. PLoS ONE *5*, e13952.

Kim, I., Saunders, T.L., and Morrison, S.J. (2007). Sox17 Dependence Distinguishes the Transcriptional Regulation of Fetal from Adult Hematopoietic Stem Cells. Cell *130*, 470–483.

Kobayashi, H., Sakurai, T., Imai, M., Takahashi, N., Fukuda, A., Yayoi, O., Sato, S., Nakabayashi, K., Hata, K., Sotomaru, Y., et al. (2012). Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. PLoS Genetics *8*, e1002440.

Kodaira, K., Imada, M., Goto, M., Tomoyasu, A., Fukuda, T., Kamijo, R., Suda, T., Higashio, K., and Katagiri, T. (2006). Purification and identification of a BMP-like factor from bovine serum. Biochemical and Biophysical Research Communications *345*, 1224–1231.

Kostetskii, I., Jiang, Y., Kostetskaia, E., Yuan, S., Evans, T., and Zile, M. (1999). Retinoid signaling required for normal heart development regulates GATA-4 in a pathway distinct from cardiomyocyte differentiation. Developmental Biology *206*, 206–218.

Koutsourakis, M., Langeveld, A., Patient, R., Beddington, R., and Grosveld, F. (1999). The transcription factor GATA6 is essential for early extraembryonic development. Development *126*, 723–732.

Krawchuk, D., Honma-Yamanaka, N., Anani, S., and Yamanaka, Y. (2013). Developmental Biology. Developmental Biology *384*, 65–71.

Kriaucionis, S. and Heintz, N. (2009). The nuclear DNA base, 5-hydroxymethylcytosine is present in brain and enriched in Purkinje neurons. Science 324, 929–930.

Kropp, E.M., Bhattacharya, S., Waas, M., Chuppa, S.L., Hadjantonakis, A.-K., Boheler, K.R., and Gundry, R.L. (2014). N-glycoprotein surfaceomes of four developmentally distinct mouse cell types. Proteomics Clin Appl *8*, 603–609.

Kunath, T. (2005). Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. Development *132*, 1649–1661.

Kuo, C.T., Morrisey, E.E., Anandappa, R., Sigrist, K., Lu, M.M., Parmacek, M.S., Soudais, C., and Leiden, J.M. (1997). GATA4 transcription factor is required for ventral morphogenesis and heart tube formation. Genes & Development *11*, 1048–1060.

Kurimoto, K. (2006). An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis. Nucleic Acids Res. *34*, e42–e42.

Kuroda, T., Tada, M., Kubota, H., Kimura, H., Hatano, S.Y., Suemori, H., Nakatsuji, N., and Tada, T. (2005). Octamer and Sox Elements Are Required for Transcriptional cis Regulation of Nanog Gene Expression. Molecular and Cellular Biology *25*, 2475–2485.

Kwon, G.S., Viotti, M., and Hadjantonakis, A.-K. (2008). The Endoderm of the Mouse Embryo Arises by Dynamic Widespread Intercalationof Embryonic and Extraembryonic Lineages. Developmental Cell *15*, 509–520.

Larsen, F., Gundersen, G., Lopez, R., and Prydz, H. (1992). CpG islands as gene markers in the human genome. Genomics *13*, 1095–1107.

Larue, L., Ohsugi, M., Hirchenhain, J., and Kemler, R. (1994). E-cadherin null mutant embryos fail to form a trophectoderm epithelium. Proceedings of the National Academy of Sciences *91*, 8263–8267.

Le Bin, G.C., Munoz-Descalzo, S., Kurowski, A., Leitch, H., Lou, X., Mansfield, W., Etienne-Dumeau, C., Grabole, N., Mulas, C., Niwa, H., et al. (2014). Oct4 is required for lineage priming in the developing inner cell mass of the mouse blastocyst. Development *141*, 1001–1010.

Lerou, P.H., and Daley, G.Q. (2005). Therapeutic potential of embryonic stem cells. Blood Rev. *19*, 321–331.

Leung, C.Y., and Zernicka-Goetz, M. (2013). Angiomotin prevents pluripotent lineage differentiation in mouse embryos via Hippo pathway-dependent and -independent mechanisms. Nature Communications *4*, 1–11.

Li, D., Guo, B., Wu, H., Tan, L., and Lu, Q. (2015). TET Family of Dioxygenases: Crucial Roles and Underlying Mechanisms. Cytogenet. Genome Res. DOI:10.1159/000438853.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915–926.

Li, F., Cowley, D.O., Banner, D., Holle, E., Zhang, L., and Su, L. (2014). Efficient genetic manipulation of the NOD-Rag1-/-IL2RgammaC-null mouse by combining in vitro fertilization and CRISPR/Cas9 technology. Sci Rep *4*, 5290.

Li, M., Pevny, L., Lovell-Badge, R., and Smith, A. (1998). Generation of purified neural precursors from embryonic stem cells by lineage selection. Curr. Biol. *8*, 971–974.

Li, P., Tong, C., Mehrian-Shai, R., Jia, L., Wu, N., Yan, Y., Maxson, R.E., Schulze, E.N., Song, H., Hsieh, C.-L., et al. (2008). Germline competent embryonic stem cells derived from rat blastocysts. Cell *135*, 1299–1310.

Li, X., Talts, U., Talts, J.F., Arman, E., Ekblom, P., and Lonai, P. (2001). Akt/PKB regulates laminin and collagen IV isotypes of the basement membrane. Proceedings of the National Academy of Sciences *98*, 14416–14421.

Liang, J., Wan, M., Zhang, Y., Gu, P., Xin, H., Jung, S.Y., Qin, J., Wong, J., Cooney, A.J., Liu, D., et al. (2008). Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. Nature Publishing Group *10*, 731–739.

Liang, Q., Wiese, R.J., Bueno, O.F., Dai, Y.S., Markham, B.E., and Molkentin, J.D. (2001). The Transcription Factor GATA4 Is Activated by Extracellular Signal-Regulated Kinase 1- and 2-Mediated Phosphorylation of Serine 105 in Cardiomyocytes. Molecular and Cellular Biology *21*, 7460–7469.

Lim, J.W.E., and Bodnar, A. (2002). Proteome analysis of conditioned medium from mouse embryonic fibroblast feeder layers which support the growth of human embryonic stem cells. Proteomics *2*, 1187–1203.

Loh, K.M., Lim, B., and Ang, L.T. (2015). Ex uno plures: molecular designs for embryonic pluripotency. Physiol. Rev. *95*, 245–295.

Lou, X., Kang, M., Xenopoulos, P., Muñoz-Descalzo, S., and Hadjantonakis, A.-K. (2014). A rapid and efficient 2D/3D nuclear segmentation method for analysis of early mouse embryo and stem cell image data. Stem Cell Reports *2*, 382–397.

Louvet, S., Aghion, J., Santa-Maria, A., Mangeat, P., and Maro, B. (1996). Ezrin becomes restricted to outer cells following asymmetrical division in the preimplantation mouse embryo. Developmental Biology *177*, 568–579.

Lowry, W.E. (2012). Does transcription factor induced pluripotency accurately mimic embryo derived pluripotency? Current Opinion in Genetics & Development *22*, 429–434.

Lu, W., Fang, L., Ouyang, B., Zhang, X., Zhan, S., Feng, X., Bai, Y., Han, X., Kim, H., He, Q., et al. (2015). Actl6a protects embryonic stem cells from differentiating into primitive endoderm. Stem Cells *33*, 1782–1793.

Ma, Y., Gu, J., Li, C., Wei, X., Tang, F., Shi, G., Jiang, J., Kuang, Y., Li, J., Wang, Z., et al. (2012). Human foreskin fibroblast produces interleukin-6 to support derivation and self-renewal of mouse embryonic stem cells. Stem Cell Res Ther *3*, 29.

MacArthur, B.D., Please, C.P., and Oreffo, R.O.C. (2008). Stochasticity and the molecular mechanisms of induced pluripotency. PLoS ONE *3*, e3086.

Mannello, F., and Tonti, G.A. (2007). Concise review: no breakthroughs for human mesenchymal and embryonic stem cell culture: conditioned medium, feeder layer, or feeder-free; medium with fetal calf serum, human serum, or enriched plasma; serum-free, serum replacement nonconditioned medium, or ad hoc formula? All that glitters is not gold! Stem Cells *25*, 1603–1609.

Margot, J.B., Ehrenhofer-Murray, A.E., and Leonhardt, H. (2003). Interactions within the mammalian DNA methyltransferase family. BMC Mol. Biol. *4*, 7.

Martin, G.R. (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 *78*, 7634–7638.

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A.A., et al. (2007). Pluripotency

governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nature Publishing Group *9*, 625–635.

Matsuda, T., Nakamura, T., Nakao, K., Arai, T., Katsuki, M., Heike, T., and Yokota, T. (1999). STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. The EMBO Journal *18*, 4261–4269.

Matsui, T. (2006). Redundant roles of Sox17 and Sox18 in postnatal angiogenesis in mice. Journal of Cell Science *119*, 3513–3526.

Mayer, W., Niveleau, A., Walter, J., Fundele, R., and Haaf, T. (2000). Demethylation of the zygotic paternal genome. Nature *403*, 501–502.

McWhir, J., Schnieke, A.E., Ansell, R., Wallace, H., Colman, A., Scott, A.R., and Kind, A.J. (1996). Selective ablation of differentiated cells permits isolation of embryonic stem cell lines from murine embryos with a non-permissive genetic background. Nat. Genet. *14*, 223–226.

Medvedev, S.P., Shevchenko, A.I., and Mazurok, N.A. (2008). OCT4 and NANOG are the key genes in the system of pluripotency maintenance in mammalian cells. Russian Journal of ….

Meilhac, S.M., Adams, R.J., Morris, S.A., Danckaert, A., Le Garrec, J.-F., and Zernicka-Goetz, M. (2009). Active cell movements coupled to positional induction are involved in lineage segregation in the mouse blastocyst. Developmental Biology *331*, 210–221.

Meng, G.L., Nieden, zur, N.I., Liu, S.Y., Cormier, J.T., Kallos, M.S., and Rancourt, D.E. (2008). Properties of murine embryonic stem cells maintained on human foreskin fibroblasts without LIF. Mol. Reprod. Dev. *75*, 614–622.

Merzouk, S., Deuve, J.L., Dubois, A., Navarro, P., Avner, P., and Morey, C. (2014). Lineage-specific regulation of imprinted X inactivation in extraembryonic endoderm stem cells. Epigenetics & Chromatin *7*, 11.

Messerschmidt, D.M., and Kemler, R. (2010). Nanog is required for primitive endoderm formation through a non-cell autonomous mechanism. Developmental Biology *344*, 129–137.

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., and Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. Nature *454*, 49–55.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631–642.

Moerkamp, A.T., Paca, A., Goumans, M.-J., Kunath, T., Kruithof, B.P.T., and Kruithof-de Julio, M. (2013). Extraembryonic Endoderm cells as a model of

endoderm development. Develop. Growth Differ. *55*, 301–308.

Molkentin, J.D. (2000). The Zinc Finger-containing Transcription Factors GATA-4, -5, and -6: ubiquitously expressed regulators of tissue-specific gene expression. Journal of Biological Chemistry *275*, 38949–38952.

Molkentin, J.D., Lin, Q., Duncan, S.A., and Olson, E.N. (1997). Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis. Genes & Development *11*, 1061–1072.

Morgan, H.D., Dean, W., Coker, H.A., Reik, W., and Petersen-Mahrt, S.K. (2004). Activation-induced Cytidine Deaminase Deaminates 5-Methylcytosine in DNA and Is Expressed in Pluripotent Tissues: implications for epigenetic reprogramming. Journal of Biological Chemistry *279*, 52353–52360.

Morris, S.A., Graham, S.J.L., Jedrusik, A., and Zernicka-Goetz, M. (2013). The differential response to Fgf signalling in cells internalized at different times influences lineage segregation in preimplantation mouse embryos. Open Biology *3*, 130104–130104.

Morris, S.A., Teo, R.T.Y., Li, H., Robson, P., Glover, D.M., and Zernicka-Goetz, M. (2010). Origin and formation of the first two distinct cell types of the inner cell mass in the mouse embryo. Proceedings of the National Academy of Sciences *107*, 6364–6369.

Morrisey, E.E., Ip, H.S., Lu, M.M., and Parmacek, M.S. (1996). GATA-6: a zinc finger transcription factor that is expressed in multiple cell lineages derived from lateral mesoderm. Developmental Biology *177*, 309–322.

Morrisey, E.E., Tang, Z., Sigrist, K., Lu, M.M., Jiang, F., Ip, H.S., and Parmacek, M.S. (1998a). GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. Genes & Development *12*, 3579–3590.

Morrisey, E.E., Tang, Z., Sigrist, K., Lu, M.M., Jiang, F., Ip, H.S., and Parmacek, M.S. (1998b). GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. Genes & Development *12*, 3579–3590.

Motosugi, N., Bauer, T., Polanski, Z., Solter, D., and Hiiragi, T. (2005). Polarity of the mouse embryo is established at blastocyst and is not prepatterned. Genes & Development *19*, 1081–1092.

Mummery, C.L., Feyen, A., Freund, E., and Shen, S. (1990). Characteristics of embryonic stem cell differentiation: a comparison with two embryonal carcinoma cell lines. Cell Differ. Dev. *30*, 195–206.

Murohashi, M., Nakamura, T., Tanaka, S., Ichise, T., Yoshida, N., Yamamoto, T., Shibuya, M., Schlessinger, J., and Gotoh, N. (2009). An FGF4-FRS2α-Cdx2 Axis in Trophoblast Stem Cells Induces BMP4 to Regulate Proper Growth of Early Mouse Embryos. Stem Cells N/A–N/A.

Narita, N., Bielinska, M., and Wilson, D.B. (1997). Wild-type endoderm abrogates the ventral developmental defects associated with GATA-4 deficiency in the mouse. Developmental Biology *189*, 270–274.

Navarro, P., Festuccia, N., Colby, D., Gagliardi, A., Mullin, N.P., Zhang, W., Karwacki-Neisius, V., Osorno, R., Kelly, D., Robertson, M., et al. (2012). OCT4/SOX2-independent Nanog autorepression modulates heterogeneous Nanog gene expression in mouse ES cells. The EMBO Journal 31, 4547–4562.

Niakan, K.K., Ji, H., Maehr, R., Vokes, S.A., Rodolfa, K.T., Sherwood, R.I., Yamaki, M., Dimos, J.T., Chen, A.E., Melton, D.A., et al. (2010). Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating extraembryonic gene expression and indirectly antagonizing self-renewal. Genes & Development *24*, 312–326.

Niakan, K.K., Schrode, N., Cho, L.T.Y., and Hadjantonakis, A.-K. (2013). Derivation of extraembryonic endoderm stem (XEN) cells from mouse embryos and embryonic stem cells. *8*, 1028–1041.

Nichols, J., Silva, J., Roode, M., and Smith, A. (2009). Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. Development *136*, 3215–3222.

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

Nishioka, N., Inoue, K.-I., Adachi, K., Kiyonari, H., Ota, M., Ralston, A., Yabuta, N., Hirahara, S., Stephenson, R.O., Ogonuki, N., et al. (2009). The Hippo Signaling Pathway Components Lats and Yap Pattern Tead4 Activity to Distinguish Mouse Trophectoderm from Inner Cell Mass. Developmental Cell *16*, 398–410.

Nishioka, N., Yamamoto, S., Kiyonari, H., Sato, H., Sawada, A., Ota, M., Nakao, K., and Sasaki, H. (2008). Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos. Mech. Dev. *125*, 270–283.

Niwa, H., Burdon, T., Chambers, I., and Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes & Development *12*, 2048–2060.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat. Genet. *24*, 372–376.

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

Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 Determines Trophectoderm Differentiation. Cell *123*, 917–929.

Ogawa, K., Matsui, H., Ohtsuka, S., and Niwa, H. (2004). A novel mechanism for regulating clonal propagation of mouse ES cells. Genes Cells *9*, 471–477.

Ohnishi, Y., Huber, W., Tsumura, A., Kang, M., Xenopoulos, P., Kurimoto, K., Oleś, A.K., Araúzo-Bravo, M.J., Saitou, M., Hadjantonakis, A.-K., et al. (2014). Cell-to-cell expression variability followed by signal reinforcement progressively segregates early mouse lineages. Nat Cell Biol *16*, 27–37.

Okamoto, I., Otte, A.P., Allis, C.D., Reinberg, D., and Heard, E. (2004). Epigenetic dynamics of imprinted X inactivation during early mouse development. Science *303*, 644–649.

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

Okumura-Nakanishi, S., Saito, M., Niwa, H., and Ishikawa, F. (2005). Oct-3/4 and Sox2 Regulate Oct-3/4 Gene in Embryonic Stem Cells. Journal of Biological Chemistry *280*, 5307–5317.

Ornitz, D.M., Xu, J., Colvin, J.S., McEwen, D.G., MacArthur, C.A., Coulier, F., Gao, G., and Goldfarb, M. (1996). Receptor specificity of the fibroblast growth factor family. J. Biol. Chem. *271*, 15292–15297.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. Curr. Biol. *10*, 475–478.

Ozbudak, E.M., Thattai, M., Kurtser, I., Grossman, A.D., and van Oudenaarden, A. (2002). Regulation of noise in the expression of a single gene. Nat. Genet. *31*, 69–73.

Palmieri, S.L., Peter, W., Hess, H., and Scholer, H.R. (1994). Oct-4 transcription factor is differentially expressed in the mouse embryo during establishment of the first two extraembryonic cell lineages involved in implantation. Developmental Biology.

Pauken, C.M., and Capco, D.G. (1999). Regulation of cell adhesion during embryonic compaction of mammalian embryos: roles for PKC and beta-catenin. Mol. Reprod. Dev. *54*, 135–144.

Pauken, C.M., and Capco, D.G. (2000). The Expression and Stage-Specific Localization of Protein Kinase C Isotypes during Mouse Preimplantation Development. Developmental Biology *223*, 411–421.

Pesce, M., Wang, X., Wolgemuth, D.J., and Schöler, H. (1998). Differential expression of the Oct-4 transcription factor during mouse germ cell differentiation. Mod *71*, 89–98.

Pfaffeneder, T., Hackner, B., Truss, M., Munzel, M., Muller, M., Deiml, C.A., Hagemeier, C., and Carell, T. (2011). The discovery of 5-formylcytosine in embryonic stem cell DNA. Angew. Chem. 50, 7008-7012.

Plusa, B. (2005). Downregulation of Par3 and aPKC function directs cells towards the ICM in the preimplantation mouse embryo. Journal of Cell Science *118*, 505–515.

Plusa, B., Piliszek, A., Frankenberg, S., Artus, J., and Hadjantonakis, A.K. (2008). Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. Development *135*, 3081–3091.

Popp, C., Dean, W., Feng, S., Cokus, S.J., Andrews, S., Pellegrini, M., Jacobsen, S.E., and Reik, W. (2010). Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. Nature *463*, 1101–1105.

Price, P.J., Goldsborough, M.D., and Tilkins, M.L. (1998). Embryonic stem cell serum replacement.

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

Rai, K., Huggins, I.J., James, S.R., Karpf, A.R., Jones, D.A., and Cairns, B.R. (2008). DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. Cell *135*, 1201–1212.

Raj, A., Rifkin, S.A., Andersen, E., and van Oudenaarden, A. (2010). Variability in gene expression underlies incomplete penetrance. Nature *463*, 913–918.

Ralston, A., Cox, B.J., Nishioka, N., Sasaki, H., Chea, E., Rugg-Gunn, P., Guo, G., Robson, P., Draper, J.S., and Rossant, J. (2010). Gata3 regulates trophoblast development downstream of Tead4 and in parallel to Cdx2. Development *137*, 395–403.

Reeve, W.J., and Ziomek, C.A. (1981). Distribution of microvilli on dissociated blastomeres from mouse embryos: evidence for surface polarization at compaction. J Embryol Exp Morphol *62*, 339–350.

Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. Nature *447*, 425–432.

Robertson, E.J. (1997). Derivation and maintenance of embryonic stem cell cultures. Methods Mol. Biol. *75*, 173–184.

Rodda, D.J. (2005). Transcriptional Regulation of Nanog by OCT4 and SOX2. Journal of Biological Chemistry *280*, 24731–24737.

Roelen, B.A., Goumans, M.J., Zwijsen, A., and Mummery, C.L. (1998).

Identification of two distinct functions for TGF-beta in early mouse development. Differentiation *64*, 19–31.

Rogers, M.B., Rosen, V., Wozney, J.M., and Gudas, L.J. (1992). Bone morphogenetic proteins-2 and -4 are involved in the retinoic acid-induced differentiation of embryonal carcinoma cells. Molecular Biology of the Cell *3*, 189–196.

Rohwedel, J., Guan, K., and Wobus, A.M. (1999). Induction of cellular differentiation by retinoic acid in vitro. Cells Tissues Organs (Print) *165*, 190–202.

Rosner, M.H., Vigano, M.A., Ozato, K., Timmons, P.M., Poirie, F., Rigby, P.W.J., and Staudt, L.M. (1990). A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. Nature *345*, 686–692.

Rossant, J. (1975). Investigation of the determinative state of the mouse inner cell mass. I. Aggregation of isolated inner cell masses with morulae. J Embryol Exp Morphol *33*, 979–990.

Rossant, J., and Lis, W.T. (2003). Potential of isolated mouse inner cell masses to form trophectoderm derivatives in vivo. Developmental Biology *70*, 255–261.

Rossant, J., and Vijh, K.M. (1980). Ability of outside cells from preimplantation mouse embryos to form inner cell mass derivatives. Developmental Biology *76*, 475–482.

Rossant, J., Chazaud, C., and Yamanaka, Y. (2003). Lineage allocation and asymmetries in the early mouse embryo. Philos. Trans. R. Soc. Lond., B, Biol. Sci. *358*, 1341–8–discussion1349.

Rougier, N., Bourc'his, D., Gomes, D.M., Niveleau, A., Plachot, M., Pàldi, A., and Viegas-Péquignot, E. (1998). Chromosome methylation patterns during mammalian preimplantation development. Genes & Development *12*, 2108–2113.

Russ, A.P., Wattler, S., Colledge, W.H., Aparicio, S.A., Carlton, M.B., Pearce, J.J., Barton, S.C., Surani, M.A., Ryan, K., Nehls, M.C., et al. (2000). Eomesodermin is required for mouse trophoblast development and mesoderm formation. Nature *404*, 95–99.

Saiz, N., Grabarek, J.B., Sabherwal, N., Papalopulu, N., and Plusa, B. (2013). Atypical protein kinase C couples cell sorting with primitive endoderm maturation in the mouse blastocyst. Development *140*, 4311–4322.

Saiz, N., Plusa, B., and Hadjantonakis, A.-K. (2015). Single cells get together: High-resolution approaches to study the dynamics of early mouse development. Seminars in Cell and Developmental Biology 1–9.

Santos, F., Hendrich, B., Reik, W., and Dean, W. (2002). Dynamic Reprogramming of DNA Methylation in the Early Mouse Embryo.

Developmental Biology *241*, 172–182.

Santostefano, K.E., Hamazaki, T., Pardo, C.E., Kladde, M.P., and Terada, N. (2012). Fibroblast Growth Factor Receptor 2 Homodimerization Rapidly Reduces Transcription of the Pluripotency Gene Nanog without Dissociation of Activating Transcription Factors. Journal of Biological Chemistry *287*, 30507–30517.

Schöler, H.R., Ruppert, S., Suzuki, N., Chowdhury, K., and Gruss, P. (1990). New type of POU domain in germ line-specific protein Oct-4. Nature *344*, 435–439.

Schrode, N., Saiz, N., Di Talia, S., and Hadjantonakis, A.-K. (2014). GATA6 Levels Modulate Primitive Endoderm Cell Fate Choice and Timing in the Mouse Blastocyst. Developmental Cell *29*, 454–467.

Schrode, N., Xenopoulos, P., Piliszek, A., Frankenberg, S., Plusa, B., and Hadjantonakis, A.-K. (2013). Anatomy of a blastocyst: cell behaviors driving cell fate choice and morphogenesis in the early mouse embryo. Genesis *51*, 219–233.

Shi, L.H., Ai, J.S., Ouyang, Y.C., Huang, J.C., Lei, Z.L., Wang, Q., Yin, S., Han, Z.M., Sun, Q.Y., and Chen, D.Y. (2008). Trichostatin A and nuclear reprogramming of cloned rabbit embryos. J. Anim. Sci. *86*, 1106–1113.

Shimamoto, R., Amano, N., Ichisaka, T., Watanabe, A., Yamanaka, S., and Okita, K. (2014). Generation and Characterization of Induced Pluripotent Stem Cells from Aid-Deficient Mice. PLoS ONE *9*, e94735.

Shimosato, D., Shiki, M., and Niwa, H. (2007). Extra-embryonic endoderm cells derived from ES cells induced by GATA Factors acquire the character of XEN cells. BMC Dev Biol *7*, 80.

Shin, M., Alev, C., Wu, Y., Nagai, H., and Sheng, G. (2011). Activin/TGF-beta signaling regulates Nanog expression in the epiblast during gastrulation. Mod *128*, 268–278.

Silva, J., Nichols, J., Theunissen, T.W., Guo, G., van Oosten, A.L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., and Smith, A. (2009). Nanog Is the Gateway to the Pluripotent Ground State. Cell *138*, 722–737.

Simmons, D.G., and Cross, J.C. (2005). Determinants of trophoblast lineage and cell subtype specification in the mouse placenta. Developmental Biology *284*, 12–24.

Simonson, O.E., Domogatskaya, A., Volchkov, P., and Rodin, S. (2015). The safety of human pluripotent stem cells in clinical treatment. Annals of Medicine 47, 370–380.

Simpson, E.M., Linder, C.C., Sargent, E.E., Davisson, M.T., Mobraaten, L.E., and Sharp, J.J. (1997). Genetic variation among 129 substrains and its importance for targeted mutagenesis in mice. Nat. Genet. *16*, 19–27.

Singh, A.M., Hamazaki, T., Hankowski, K.E., and Terada, N. (2007). A heterogeneous expression pattern for Nanog in embryonic stem cells. Stem Cells *25*, 2534–2542.

Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature *336*, 688–690.

Smith, R., and McLaren, A. (1977). Factors affecting the time of formation of the mouse blastocoele. J Embryol Exp Morphol *41*, 79–92.

Sodhi, C.P., Li, J., and Duncan, S.A. (2006). Generation of mice harbouring a conditional loss-of-function allele of Gata6. BMC Dev Biol *6*, 19.

Soprano, D.R., Teets, B.W., and Soprano, K.J. (2007). Role of retinoic acid in the differentiation of embryonal carcinoma and embryonic stem cells. Vitam. Horm. *75*, 69–95.

Soudais, C., Bielinska, M., Heikinheimo, M., MacArthur, C.A., Narita, N., Saffitz, J.E., Simon, M.C., Leiden, J.M., and Wilson, D.B. (1995). Targeted mutagenesis of the transcription factor GATA-4 gene in mouse embryonic stem cells disrupts visceral endoderm differentiation in vitro. Development *121*, 3877–3888.

Spindle, A.I. (2005). Trophoblast regeneration by inner cell masses isolated from cultured mouse embryos. J. Exp. Zool. *203*, 483–489.

Stewart, C.L., Kaspar, P., Brunet, L.J., Bhatt, H., Gadi, I., Köntgen, F., and Abbondanzo, S.J. (1992). Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. Nature *359*, 76–79.

Strumpf, D. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. Development *132*, 2093–2102.

Sun, Y., Li, H., Yang, H., Rao, M.S., and Zhan, M. (2006). Mechanisms Controlling Embryonic Stem Cell Self-Renewal and Differentiation. Critical Reviews in Eukaryotic Gene Expression *16*, 211–232.

Suwinska, A., Czołowska, R., Ożdżeński, W., and Tarkowski, A.K. (2008). Blastomeres of the mouse embryo lose totipotency after the fifth cleavage division: Expression of Cdx2 and Oct4 and developmental potential of inner and outer blastomeres of 16- and 32-cell embryos. Developmental Biology *322*, 133–144.

Suzuki, A., Raya, A., Kawakami, Y., Morita, M., Matsui, T., Nakashima, K., Gage, F.H., Rodríguez-Esteban, C., and Izpisúa Belmonte, J.C. (2006). Nanog binds to Smad1 and blocks bone morphogenetic protein-induced differentiation of embryonic stem cells. Proceedings of the National Academy of Sciences *103*, 10294–10299.

Suzuki, O., Matsuda, J., Takano, K., Yamamoto, Y., Asano, T., Naiki, M., and

Kusanagi, M. (1999). Effect of genetic background on establishment of mouse embryonic stem cells. Exp. Anim. *48*, 213–216.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R.,Aravind, L., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324, 930-935.

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

Tanaka, S., Kunath, T., Hadjantonakis, A.K., Nagy, A., and Rossant, J. (1998). Promotion of trophoblast stem cell proliferation by FGF4. Science *282*, 2072–2075.

Tarkowski, A.K., and Wróblewska, J. (1967). Development of blastomeres of mouse eggs isolated at the 4- and 8-cell stage. J Embryol Exp Morphol *18*, 155–180.

Torres-Padilla, M.E., and Chambers, I. (2014). Transcription factor heterogeneity in pluripotent stem cells: a stochastic advantage. Development *141*, 2173–2181.

Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S.-I., Sakaue, M., Matsuoka, C., Shimotohno, K., Ishikawa, F., Li, E., Ueda, H.R., et al. (2006). Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes Cells *11*, 805–814.

Umehara, H., Kimura, T., Ohtsuka, S., Nakamura, T., Kitajima, K., Ikawa, M., Okabe, M., Niwa, H., and Nakano, T. (2007). Efficient derivation of embryonic stem cells by inhibition of glycogen synthase kinase-3. Stem Cells *25*, 2705–2711.

Uy, G.D., Downs, K.M., and Gardner, R.L. (2002). Inhibition of trophoblast stem cell potential in chorionic ectoderm coincides with occlusion of the ectoplacental cavity in the mouse. Development *129*, 3913–3924.

Vallier, L., Mendjan, S., Brown, S., Chng, Z., Teo, A., Smithers, L.E., Trotter, M.W.B., Cho, C.H.H., Martinez, A., Rugg-Gunn, P., et al. (2009). Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. Philosophical Transactions of the Royal Society B: Biological Sciences *366*, 2230–2237.

Vanroose, G., Van Soom, A., and de Kruif, A. (2001). From co-culture to defined medium: state of the art and practical considerations. Reprod. Domest. Anim. *36*, 25–28.

Villa-Diaz, L.G., Pacut, C., Slawny, N.A., Ding, J., O'Shea, K.S., and Smith, G.D. (2009). Analysis of the factors that limit the ability of feeder cells to maintain the undifferentiated state of human embryonic stem cells. Stem Cells

and Development *18*, 641–651.

Vinot, S., Le, T., Ohno, S., Pawson, T., Maro, B., and Louvet-Vallee, S. (2005). Asymmetric distribution of PAR proteins in the mouse embryo begins at the 8-cell stage during compaction. Developmental Biology *282*, 307–319.

Viotti, M., Nowotschin, S., and Hadjantonakis, A.-K. (2014). SOX17 links gut endoderm morphogenesis and germ layer segregation. Nat Cell Biol *16*, 1146–1156.

Wamaitha, S.E., del Valle, I., Cho, L.T.Y., Wei, Y., Fogarty, N.M.E., Blakeley, P., Sherwood, R.I., Ji, H., and Niakan, K.K. (2015). Gata6 potently initiates reprograming of pluripotent and differentiated cells to extraembryonic endoderm stem cells. Genes & Development *29*, 1239–1255.

Wang, C., and Song, B. (1996). Cell-type-specific expression of the platelet-derived growth factor alpha receptor: a role for GATA-binding protein. Molecular and Cellular Biology *16*, 712–723.

Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D.N., Theunissen, T.W., and Orkin, S.H. (2006). A protein interaction network for pluripotency of embryonic stem cells. Nature *444*, 364–368.

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

Watanabe, T., Biggins, J.S., Tannan, N.B., and Srinivas, S. (2014). Limited predictive value of blastomere angle of division in trophectoderm and inner cell mass specification. Development *141*, 2279–2288.

Watson, A.J., Natale, D.R., and Barcroft, L.C. (2004). Molecular regulation of blastocyst formation. Animal Reproduction Science *82–83*, 583–592.

Wen, D., Saiz, N., Rosenwaks, Z., Hadjantonakis, A.-K., and Rafii, S. (2014). Completely ES cell-derived mice produced by tetraploid complementation using inner cell mass (ICM) deficient blastocysts. PLoS ONE *9*, e94730.

Wicklow, E., Blij, S., Frum, T., Hirate, Y., Lang, R.A., Sasaki, H., and Ralston, A. (2014). HIPPO Pathway Members Restrict SOX2 to the Inner Cell Mass Where It Promotes ICM Fates in the Mouse Blastocyst. PLoS Genetics *10*, e1004618.

Wilder, P.J., Kelly, D., Brigman, K., Peterson, C.L., Nowling, T., Gao, Q.S., McComb, R.D., Capecchi, M.R., and Rizzino, A. (1997). Inactivation of the FGF-4 gene in embryonic stem cells alters the growth and/or the survival of their early differentiated progeny. Developmental Biology *192*, 614–629.

Williams, R.L., Hilton, D.J., Pease, S., Willson, T.A., Stewart, C.L., Gearing, D.P., Wagner, E.F., Metcalf, D., Nicola, N.A., and Gough, N.M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of

embryonic stem cells. Nature *336*, 684–687.

Wong, A.H.C. (2005). Phenotypic differences in genetically identical organisms: the epigenetic perspective. Human Molecular Genetics *14*, R11–R18.

Wu, H., and Zhang, Y. (2014). Reversing DNA Methylation: Mechanisms, Genomics, and Biological Functions. Cell *156*, 45–68.

Wu, Q., Chen, X., Zhang, J., Loh, Y.-H., Low, T.-Y., Zhang, W., Zhang, W., Sze, S.-K., Lim, B., and Ng, H.-H. (2006). Sall4 interacts with Nanog and co-occupies Nanog genomic sites in embryonic stem cells. J. Biol. Chem. *281*, 24090–24094.

Xenopoulos, P., Kang, M., Puliafito, A., Di Talia, S., and Hadjantonakis, A.-K. (2015). Heterogeneities in Nanog Expression Drive Stable Commitment to Pluripotency in the Mouse Blastocyst. Cell Reports *10*, 1508–1520.

Xin, M., Davis, C.A., Molkentin, J.D., Lien, C.-L., Duncan, S.A., Richardson, J.A., and Olson, E.N. (2006). A threshold of GATA4 and GATA6 expression is required for cardiovascular development. Proceedings of the National Academy of Sciences *103*, 11189–11194.

Xuan, S., Borok, M.J., Decker, K.J., Battle, M.A., Duncan, S.A., Hale, M.A., Macdonald, R.J., and Sussel, L. (2012). Pancreas-specific deletion of mouse Gata4 and Gata6 causes pancreatic agenesis. J. Clin. Invest. *122*, 3516–28.

Xue, Z., Huang, K., Cai, C., Cai, L., Jiang, C.-Y., Feng, Y., Liu, Z., Zeng, Q., Cheng, L., Sun, Y.E., et al. (2013). Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. Nature *500*, 593–597.

Yagi, R., Kohn, M.J., Karavanova, I., Kaneko, K.J., Vullhorst, D., DePamphilis, M.L., and Buonanno, A. (2007). Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development. Development *134*, 3827–3836.

Yamanaka, Y., Lanner, F., and Rossant, J. (2010). FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. Development *137*, 715–724.

Yeom, Y.I., Fuhrmann, G., Ovitt, C.E., Brehm, A., Ohbo, K., Gross, M., Hübner, K., and Scholer, H.R. (1996). Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. Development *122*, 881–894.

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

Yuan, H., Corbi, N., Basilico, C., and Dailey, L. (1995). Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. Genes & Development *9*, 2635–2645.

Zappone, M.V., Galli, R., Catena, R., Meani, N., De Biasi, S., Mattei, E., Tiveron, C., Vescovi, A.L., Lovell-Badge, R., Ottolenghi, S., et al. (2000). Sox2 regulatory sequences direct expression of a (beta)-geo transgene to telencephalic neural stem cells and precursors of the mouse embryo, revealing regionalization of gene expression in CNS stem cells. Development *127*, 2367–2382.

Zhang, X., Ibrahimi, O.A., Olsen, S.K., Umemori, H., Mohammadi, M., and Ornitz, D.M. (2006). Receptor specificity of the fibroblast growth factor family. The complete mammalian FGF family. J. Biol. Chem. *281*, 15694–15700.

Zhao, R., Watt, A.J., Li, J., Luebke-Wheeler, J., Morrisey, E.E., and Duncan, S.A. (2005). GATA6 Is Essential for Embryonic Development of the Liver but Dispensable for Early Heart Formation. Molecular and Cellular Biology *25*, 2622–2631.

Zhao, R., Watt, A.J., Battle, M.A., Li, J., Bondow, B.J., and Duncan, S.A. (2008). Loss of both GATA4 and GATA6 blocks cardiac myocyte differentiation and results in acardia in mice. Developmental Biology *317*, 614–619.

DECLARATION OF CONTRIBUTION AS CO-AUTHOR

Publication I:

N.Sc. and A.-K.H. designed the study. N.Sc. and N.Sa. performed experiments. N.Sc., N.Sa. and S.D.T. performed computational analysis. N.Sc. and A.-K.H. wrote the manuscript.

Publication II:

K.K.N. and A.-K.H. outlined the protocol. K.K.N., N.S. and A.-K.H. wrote the protocol manuscript and L.T.Y.C. contributed to the manuscript.

Publication III:

A.C. and L.G.R. developed the protocol and wrote the manuscript. C.B., I.G. and L.R.D. supported the development of the protocol. N.S. and A.-K.H. contributed to and expanded the protocol and contributed to the manuscript.

Publication IV:

R.K., J.C., O.E. and T.E. conceived the study. R.K., L.D., N.S., T.-C.L., P.F. and S.M.-D. carried out experiments. R.K., J.C., O.E. and T.E. wrote the manuscript. A.A.Z. and A.-K.H. provided essential reagents and expertise. O.E. carried out computational and informatics analyses.

I hereby confirm the above statements:

Nadine Schrode Prof. Anna-Katerina Hadjantonakis

ACKNOWLEDGEMENTS

First, I'd like to greatly thank my adviser at Sloan-Kettering Institute Kat Hadjantonakis for giving me the opportunity to do my thesis work in her lab. It was a very educational time in many ways and much of my successful doctorate is owed to her guidance.

A special thanks also goes to Prof. Heinrich Leonhardt, who acted as official adviser at LMU Munich. He thereby made it possible for me to do an external doctorate in the first place. He was understanding of all arising circumstances and took interest in my future.

Further I'd like to thank Prof. Charles David for agreeing to write the Zweitgutachten and Prof. Barbara Conradt and Prof. Marc Bramkamp for the additional evaluation of my thesis and defense.

Many thanks also to all my collaborators throughout the years for the great team work!

I would also like to thank my coworkers at SKI, who often made both life and research easier. Panos Xenopoulos, Nestor Saiz, Anna Ferrer, Sophie Balmer and Vidur Garg were supportive parts of my life, not just scientifically but also personally. Thank you for the great time, and Sophie, Nestor and Vidur for reading and commenting on my thesis! A particular thanks goes here to Min Kang, my fellow grad student, with whom I had many scientific, philosophical as well as personal discussions; Thank you for your friendship Min!

One of the greatest thanks goes to Piero Sanfilippo, who started out as a coworker but became my closest friend in the far and foreign and sometimes overwhelming city of New York. We went through a myriad of ups and downs and only with his continued support did I make it to this point. Thanks for everything!

Further thanks goes to my non-coworker friends in New York as well as in Germany, who were great in getting my mind off work and making life more enjoyable every time I saw them. Thanks for being there, Jessica Johnson, Sandra Runkel, Andreas Jungfer, Steffi Thamm and Sarah Krumpa.

Writing my thesis and getting through the last stressful stages of my doctorate was lightened by J.T. Poirier. Thanks for enduring stressed-me, for lots of chocolates and sweets and motivational post-its to keep me going.

Finally I want to thank my family, who was always supportive in every way possible. I know it wasn't easy with me being so far away and often too busy to be in touch. Thank you for always being there! And a special thanks to my brother Marcel who visited every year and who got my thesis printed, bound and delivered. I know you're busy and I really appreciate all your support!

STATUTORY DECLARATION AND STATEMENT

(EIDESSTATTLICHE VERSICHERUNG UND ERKLÄRUNG)

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbständig und ohne unerlaubte Hilfe angefertigt ist.

Des Weiteren erkläre ich, dass die vorliegende Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist, dass ich mich nicht anderweitig einer Doktorprüfung ohne Erfolg unterzogen habe und dass ich nicht ohne Erfolg versucht habe, eine Dissertation einzureichen oder mich der Doktorprüfung zu unterziehen.