The Synchronisation of Pluripotency Transitions with Morphogenesis during Mouse Peri-implantation Development

Irene Escudero Selma

The Synchronisation of Pluripotency Transitions with Morphogenesis during Mouse Peri-implantation Development

Irene Escudero-Selma

Uitgegeven in eigen beheer - Publication privately printed – Irene Escudero-Selma

Cover-design: Irene Galera

The work presented in this joint doctorate thesis was performed in the Department of Cell Biology at Erasmus University Medical Center (Erasmus MC), Rotterdam, The Netherlands. The PhD training at the Erasmus University Rotterdam (EUR) was coordinated through the Erasmus MC Graduate School. The Department of Cell Biology is also member of the Erasmus MC – Leiden University joint research school *Medisch Genetisch Centrum (MGC) Zuid-West Nederland*. This research was funded by ZonMw (grant number 116006104). This booklet is the dissertation presented in partial fulfilment of the requirements for the doctorate degree at the MUL and EUR, respectively.

Lay-out and printing by ProefschriftMaken, proefschriftmaken.nl

The Synchronisation of Pluripotency Transitions with Morphogenesis during Mouse Peri-implantation Development

Synchronisatie van Pluripotentie Transities met Morfologie in Peri-Implantatie Muis Ontwikkeling

Thesis

to obtain the doctorate degree from the Erasmus University Rotterdam
by command of the
rector magnificus Prof. Dr. A.L. Bredenoord
and in accordance with the decision of the Doctorate Board
The public defense shall be held on
Tuesday 5th September, 2023, at 15.30 pm.

by
Irene Escudero-Selma
born in Barcelona (Spain)

ERASMUS UNIVERSITEIT ROTTERDAM

Doctoral promotion committee

Promoter:	Prof. Dr. D.F.E. Huylebroeck
Co-promoter:	Assoc. Prof. Dr. D. ten Berge
Other members:	Dr. W.M. Baarends Prof.dr. M.E. Drukker Prof. R. Fodde

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

+	Positive
2D	2 dimensional
3D	3 dimensional
AMIS	Apical Membrane Initiation Site
AZD	AZD4547, FGFR inhibitor
Chip-Seq	Chromatin Immunoprecipitation Sequencing
CHIR	Chir99021, GSK3 inhibitor
E	Embryonic day
ESCs	Embryonic stem cells
EpiLCs	Epiblast-like stem cells
EpiSCs	Epiblast stem cells
FGF	Fibroblast Growth Factor
h	Human
H3	Histone 3
H3K4me3	Histone 3 lysine 4 tri-methylation
H3K27me3	Histone 3 lysine 27 tri-methylation
ICM	Inner Cell Mass
KO	Knock-out
LIF	Leukaemia Inhibitory Factor
LIM	LIF, IWP2, MEK-inhibitor
m	Mouse
Otx2	Orthodenticle Homeobox 2
P	Pluripotency
PD	PD-0325901, MEK inhibitor
PODXL	Podocalyxin
PrAP	Pre-Apical Patch
PSCs	Pluripotent stem cells
RNA	Ribonucleic Acid
RNA-seq	RNA Sequencing
RSCs	Rosette-like stem cells
TF	Transcription factor
WT	Wild-type

Chapter 1

Introduction

Embryonic development is a complex and highly organized multi-step process. The single cell formed after the fertilization of the oocyte needs first to divide and differentiate in a timely fashion to generate specialized cell types, such as contracting cardiomyocytes and electrically excitable neurons. This means that one of the key requirements for development is the existence of an initial population of cells with the potential to expand and differentiate into all the cell types in the embryo. The identification of this cell type was made gradually, with evidence showing that individual cells as early as 2-cell stage were capable of generating a new embryo (for a review see Mummery, 2014). Later, Paul and Edwards identified a population of tightly attached embryonic cells that when transplanted were able to form aggregates containing multiple cell types (Mummery, 2014; Edwards, 2002).

Their discoveries opened up promising avenues not only to understand early embryogenesis and subsequent development, but also for clinical and translational research, such as reproducibly studying cell differentiation and developing *in vitro* fertilization techniques which would serve to deal with sub- and infertility in the clinic. In order to maximize this potential, it was paramount to describe these cells, their characteristics and how and when they appear in the embryo, also in order to culture them for *in vitro* applications.

THE MAKING OF A NEW INDIVIDUAL

The first step in mouse development is fertilization, which brings together the maternal oocyte and the paternal sperm to form a single cell, the fertilized egg or zygote (**Figure 1**). The zygote can give rise to all the tissues in the developing embryo and also the extra-embryonic tissues required to support its development, such as the yolk sac, the amnion and the chorion, and contribution of the allantois tissue to the later placenta. This potential of the fertilized egg or zygote is referred to as **totipotency**.

After fertilization, the mammalian zygote undergoes rounds of slow, but ordered mitotic division to increase the number of cells, named blastomeres (Boroviak, 2014). This is called cleavage, which comprises the 2-cell until the 8-16 cell embryo. At this stage, blastomeres can contribute to both embryonic and extra-embryonic tissues when transplanted to an acceptor embryo, and therefore retain totipotency (Kelly, 1977). At embryonic day (E) 2.5 (E2.5), the mouse embryo reaches the 8-cell stage. In order to construct the new embryo, blastomeres then need to make lineage commitment decisions to generate the different embryonic and extra-embryonic tissues. In the next few paragraphs, we will discuss the process by which the blastomeres commit or become instructed to acquire distinct fates.

First choice: inside or outside?

The first commitment decision is to differentiate a number of the embryonic cells towards the **trophectoderm (TE)**, which will serve implantation and contribute to the majority of the placenta, and at the same time establish the **inner cell mass (ICM)** cells that will generate the embryo, as well as specific extra-embryonic tissue cells derived from it. The cues for such a decision between cells that are apparently equal arise thanks to **compaction** around E2.75. Compaction creates spatial cues via an increase of intercellular interactions, mediated by E-cadherin: doing so, outer cells become differentiated from inner cells. This allows the outer cells in the embryo to undergo **polarization**, creating distinct apical and basal domains. This is key, because perturbation of polarization (e.g., through disruption of the levels of polarization factor Par3 or expression of a dominant-negative form of the apical determinant aPKC) increases the contribution of cells towards the ICM (Plusa, 2005).

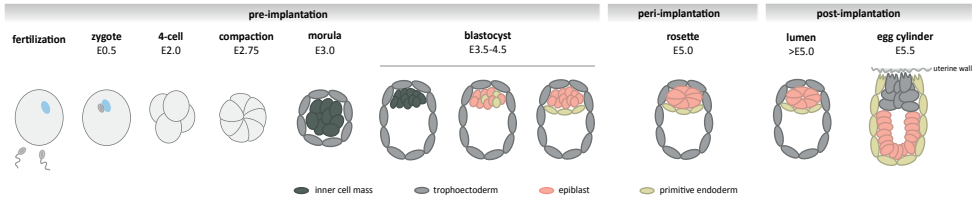


Figure 1. Early mouse development timeline, from fertilization to implantation into the maternal uterus.

Up until this point, cells retain some plasticity, as outer cells can still become inner cells if transplanted to an earlier-stage mouse embryo (Rossant, 1980). The cellular differences established by compaction and polarization are needed for these cells to become committed. Commitment is mediated by the key TE transcription factor (TF) CDX2 (a member of the caudal-related homeobox TF gene family) and the ICM TF OCT4 (also called OCT3/4 or POU5F1, a POU homeodomain, octamer-binding TF). Consequently, *Cdx2* knock-out (KO) embryos are able to undergo compaction and polarization, but do not form a mature TE, and collapse (Strumpf, 2005). *Cdx2* is expressed from the 8-cell stage, with CDX2 protein detectable in the outer cells of 16-cell stage embryos until its restriction to the TE in blastocysts. Initially, *Cdx2* is co-expressed with *Oct4*, whose mRNA expression is later limited to the inside cells (Strumpf, 2005). OCT4 and CDX2 interact with each other and are able as TFs to repress each other's targets, which can explain their exclusionary expression pattern (Niwa, 2005).

The pathway upstream of *Cdx2* and *Oct4* relaying positional information to initiate the separation of TE and ICM was not known until the observation that *Tead4*-KO embryos expressed *Oct4* and *Nanog* in all the cells, and failed to form TE (Nishioka, 2007). Research turned to the study of the Hippo pathway, which includes the TF TEAD4. In this pathway,

MTS1/2 kinases form heterodimers with SAV1, which leads to the phosphorylation of SAV1, MOB1 and the LATS1/2 kinases. LATS1/2 phosphorylate YAP (yes-associated protein) and TAZ (WW domain–containing transcription regulator protein-1), which are key regulators of the Hippo pathway, and causing the latter proteins to be sequestered in the cytoplasm. Neither YAP nor TAZ have DNA-binding domains, but rather act as binding partners of the TEAD family of TFs when translocated into the nucleus, inducing transcription (Ma, 2019). On the outside cells of pre-blastocyst stage embryos, YAP is in the nucleus where it can interact with TEAD4, leading to an upregulation of *Cdx2*. On the contrary, in inside cells YAP becomes phosphorylated and sequestered in the cytoplasm, so *Cdx2* expression is very low. Cell-cell contacts are key to maintain adequate YAP phosphorylation and cellular localization, and disruption of E-cadherin causes increased YAP in the nuclei of the inside cells of the embryo (Nishioka, 2009). In reverse, experimental decrease of expression of apical domain proteins increases retention of YAP in the cytoplasm and decreases *Cdx2*, leading to more cells acquiring ICM fate (Posfai, 2014).

Second choice: to be (an embryo) or not to be?

This first lineage specification period finishes with the formation of a blastocyst around the 32-cell stage. At around the ±64-cell stage such blastocyst has a fluid-filled cavity (the blastocoel) that contains about 20 cells of ICM on one side (the embryonic pole) and is entirely surrounded by an epithelial layer of TE cells. Totipotency is no longer experimentally detectable around this stage, as the ability of inner and outer blastomeres to generate a normal blastocyst that can implant and thereby generate a new individual is lost around the 5th cleavage division (±32-cell stage) (Suwinska, 2008). In fact, the ICM cells then go through another lineage decision between **primitive endoderm** (**PrE**, also named hypoblast), which will later form the visceral and parietal yolk sacs, and the **epiblast**, which will generate the three germ layers of the embryo proper and the entire allantois.

Initially, at the 32-cell stage, cells co-produce TFs associated with PrE fate (i.e., GATA6, SOX17) and epiblast fate (NANOG, SOX2, OCT4), respectively, but at the 64-cell stage their presence becomes mutually exclusive (Saiz, 2013; Posfai, 2014). The main effectors behind this change seem to be polypeptide growth factor FGF4, acting through its signalling receptor FGFR2, together with the TFs NANOG (in the epiblast cells) and GATA6 (in the hypoblast cells). FGF4 had been shown to be crucial for post-morula and hence also post-implantation embryo development, as *Fgf4*-KO embryos cannot properly grow an ICM (Feldman, 1995). At the 32-cell stage, *Fgf4* is expressed in cells fated to become epiblast, whereas *Fgfr2* is expressed in cells that will become PrE. The importance of these effectors is clear: treating embryos with kinase inhibitors of the tyrosine-kinase activity of FGF receptors causes all the ICM cells to become epiblast, whereas addition of FGF4 causes them to become PrE (Yamanaka, 2010). Further, commitment to either PrE or epiblast fate seems to be consolidated around E4.0-4.5, because mouse embryos treated

with inhibitors or Fgf4, and then switched to control media before, are able to restore PrE and epiblast cell populations (Yamanaka, 2010). Although NANOG is the key ICM factor, genetic inactivation of *Nanog* in mouse embryos not only abrogates epiblast formation, but also causes defects in the PrE lineage. This is due to lack of FGF4 production from *Nanog*-positive (+) cells, as indicated by the restoration of the PrE in homozygous *Nanog*-KO embryos by exogenous FGF4 (Frankenberg, 2011). NANOG is able to repress *Gata6* and binds to the *Gata6* gene promoter, supporting a functional antagonism between these two TFs (Singh, 2007). FGF4 might alleviate NANOG-dependent suppression of *Gata6* in PrE precursors, thus promoting increased expression of *Gata6*, and subsequent activation of *Gata4*, *Sox17* and *Pdgfra*.

At the end of this PrE and epiblast specification, these two cell types are still intermingled in a "salt-and-pepper" pattern. This is resolved around the peri-implantation period at E4.5, when PrE cells form an epithelium that separates the epiblast from the blastocoel at the embryonic pole of the blastocyst, where the ICM cells are located. This is achieved by both migration of the cells, mostly by passive mechanisms, and by apoptosis of PrE cells on the inside of the ICM (Plusa, 2008).

The initially totipotent zygote has now evolved into an organized structure that contains a specific ICM-derived cell subpopulation, the epiblast, that can generate a new individual. These cells are named **pluripotent** to acknowledge their developmental potential, although it is more limited as compared to the totipotent zygote. Here, I will use the term pluripotency as the ability of these early embryonic cells to differentiate into cell types of any embryonic (germ layer) lineage (endoderm, mesoderm and ectoderm) of the embryo proper, and thereby to contribute to the creation of a full new individual in the end, and also to specific, but not all, extra-embryonic tissues.

Importantly for some of the experimental work (see **Chapter 3**), the cells of the epiblast become organized in a transient, morphologically visible rosette, before the blastocyst implants in the uterine wall at around E5.0. In order to gain this implantation competency, the embryos need to leave their outer shells or *zona pellucida*, a translucent, protective glycoprotein layer already surrounding the plasma membrane of mammalian oocytes (Wang, 2006). Implantation is an organized process that depends on competence from both the blastocyst and the hormonally-prepared uterus (Paria, 1993), but optimal vs sub-optimal implantation sites likely exist.

Three main stages can be delineated in implantation: apposition, attachment and penetration (Wang, 2006). During apposition, the TE gets close to the luminal epithelium of the uterus, which adapts to allow interdigitation of microvilli from the TE and luminal epithelium (Wang, 2006). This initial contact is strengthened during attachment through adhesion-

serving integrin and selectin proteins, so that embryos cannot be dislodged from the uterus through flushing (Wang, 2006). The embryo then invades the uterus through the luminal epithelium and basal lamina into the uterine stroma, where it can establish vascular contacts with the mother (Wang, 2006). With implantation comes a series of intense morphological changes in the epiblast, with the formation of the mouse egg cylinder shortly after implantation, surrounded by visceral endoderm. The final ending of the pluripotent state of the mouse epiblast happens at around E8.0 (Huang, 2012; Osorno, 2012). Gastrulation, the process through which all three embryonic germ layers are specified, and which ultimately re-allocates cells drastically in the embryo and establishes the body plan, starts around E6.0 in mouse embryos (Arnold, 2009). Within the mouse egg cylinder, the future anterior-posterior (head-tail) and dorsal-ventral axes (back-belly) of the embryo are determined before gastrulation starts. How gastrulation occurs, and how it can morphologically vary between different animal species, is discussed in depth in reviews (e.g., Arnold, 2009), but is outside of the scope of this PhD thesis.

The transient nature of pluripotency and the difficulties of studying very tightly timely regulated stages in the embryo are a challenge to understanding the molecular underpinnings of these processes. Therefore, isolation and culture of pluripotent stem cell populations from several stages of the embryo in a scalable and reliable manner are a bottleneck for research investigating the underlying mechanisms of early-embryonic pluripotency.

BRINGING PLURIPOTENCY TO THE DISH: *IN VITRO* MODELS

Interest in elucidating the mechanisms underlying the developmental potential of the ICM of mammalian embryos initially dealt with significant technical barriers in culturing pluripotent stem cells in the dish, because only transient pluripotent cells, which became differentiated over time in culture, could be obtained (Atienza-Samols, 1978). That is why alternatives, like the use of embryonal carcinoma cells (EC cells), which are the pluripotent stem cells of teratocarcinomas, were initially more often used in the field. The first report of self-renewing pluripotent stem cells established from the mouse embryo comes from late-blastocyst pre-implantation ICM cells cultured in feeder layers with conditioned medium from PSA-1 EC cell lines (Martin, 1981). These cells were called **embryonic stem cells (ESCs)** to signal their embryonic origin.

Originally, a limitation of ESCs was their availability, as not all mammalian species and more notably also not all mouse strains were conducive of stable ESC line generation (the so-called non-permissive strains) (Brons, 2007). A notable exception to this rule were human ESCs (hESCs), although their signalling requirements were radically different than those of mouse ESCs (mESCs), for they required addition of FGF2 and Activin-A to

the culture medium (Thomson, 1998). The underlying signalling requirements were later shown to be conserved in the mouse. However, mESCs could not be maintained in FGF2/Activin-A, but rather non-ESC-like cell lines could be derived from the late epiblast from pre-gastrulation stages (i.e., E5.75) (Brons, 2007). These cell lines expressed pluripotency marker genes such as *Oct4* and *Nanog*, but did show stark differences compared to mESCs (Brons, 2007). To account for their origin and to indicate their differences with mESCs and analogy with later mouse cells, they were called **epiblast stem cells (EpiSCs)** (Brons, 2007) (**Figure 2**). EpiSCs can be derived from pre-somite post-implantation embryos up to E8.0 (Osorno, 2012).

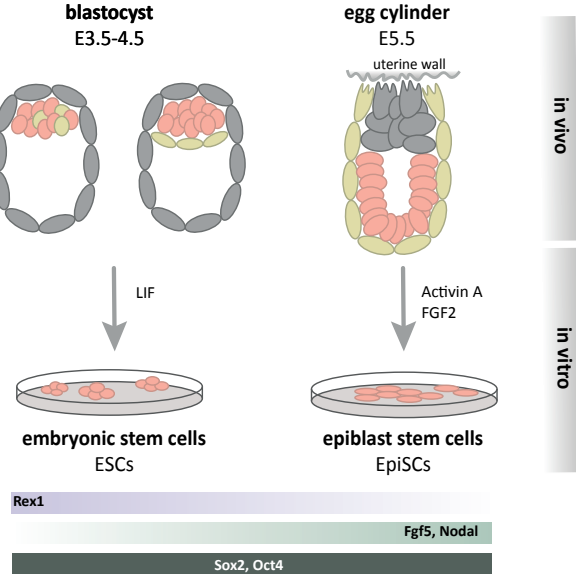


Figure 2. Pluripotent stem cell derivation from mouse embryos.

Both mESCs and EpiSCs are of course derived from the pluripotent epiblast, so if they are *bona-fide* representations of their embryonic counterparts, they should remain pluripotent. There are several tests and stringency requirements for pluripotency in the mouse. Functional assays include differentiation into the three germ layers *in vitro*, teratoma formation *in vivo*, capacity of chimaera formation, germline transmission after blastocyst injection and, the most stringent, tetraploid complementation and single-cell chimaera formation (Li, 2017; De Los Angeles, 2015). ESCs are able to differentiate *in vitro*, generate teratomas, and form germline competent chimaeras or produce tetraploid complementation when injected into blastocysts (Bradley, 1984; De Los Angeles, 2015). In contrast, EpiSCs do not efficiently form chimaeras when transplanted into pre-implantation blastocysts, and germline transmission is not observed (Brons, 2007; De Los Angeles, 2015). However, when grafted into late-

stage embryos at E7.5, but not later, EpiSCs are capable to contribute to a large proportion of tissues: at the same stage, ESCs cannot be grafted (Huang, 2012). These observations indicate that EpiSCs are incompatible with the embryonic environment in pre-implantation embryos, and ESCs are incompatible with the environment after implantation, so EpiSCs represent and occupy a later developmental stage than ESCs. Consequently, cultured ESCs need to transit to a more advanced stem cell population in the petri dish named epiblast-like stem cells (EpiLCs), which are more similar to E5.75 epiblasts, before they can differentiate (Hayashi, 2011). Owning to the aforementioned distinct developmental potential, ESCs and EpiSCs are also sometimes said to represent the "naive" and "primed" pluripotent state, respectively (see below, **Table 1**) (Nichols, 2009).

Table 1. Differences between naive and primed stem cells. Adapted from Nichols et al. (2009)

Property	Naive stem cells	Primed stem cells
Tissue of origin	ICM of the blastocyst	Epiblast post-implantation
Stem cell type	Embryonic stem cells	Epiblast stem cells
Blastocyst chimaeras	Yes	No
Teratomas	Yes	Yes
Pluripotency factors	OCT4, NANOG, SOX2, KLF2, KLF4	OCT4, SOX2, NANOG
Specific markers	REX1, NR0B1, FGF4	FGF5, T, OTX2
Response to LIF/STAT3	Self-renewal	None
Response to FGF/Erk	Differentiation	Self-renewal
Clonogenicity	High	Low
XX status	XaXa	XaXi
Metabolic profile	Bivalent	Primarily glycolysis

NAIVE VS PRIMED PLURIPOTENCY: DISTINCT STATES OF THE PLURIPOTENT CONTINUUM

Transcriptome analysis of ESCs and EpiSCs showed that their gene expression profiles are highly similar, but ESCs have a much higher degree of heterogeneity (Factor, 2014). Further studies confirmed that distinct cell populations existed within ESC cultures, which are sometimes said to be in a *metastable* state, although this concept remains vaguely defined (Choi, 2016; Hanna, 2009). Generally speaking, EpiSCs are transcriptionally closer to the post-implantation epiblast, whereas ESCs are more similar to E4.5 pre-implantation mouse epiblasts, as would be expected from the embryonic stage at which they can graft and contribute to embryo formation (Boroviak, 2014). Transcriptome analysis of pre-implantation and post-implantation epiblast also determined two sets of distinct pluripotency marker genes: *Fbxo15*, *Esrrb*, *Dppa3* (also called *Stella*), *Klf2*, *Klf4*,

Klf5, *Tbx3*, *Zfp42* (also called *Rex1*), *Tfcp2l1* and *Nr0b1* in the pre-implantation epiblast, and *Fgf5*, *Lef1* and *Nodal* in the post-implantation epiblast (Boroviak, 2014). In line with their developmental origin, ESCs express higher steady-state levels of *Esrrb*, *Zfp42*, *Tfcp2l1*, *Nr0b1* and *Dppa3*, whereas EpiSCs have higher levels of *Fgf5*, a gene that remains expressed in the primitive streak as well (Ghimire, 2018).

Even though some reports of spontaneous reversal of EpiSCs to naive pluripotency in culture exist, EpiSCs generally cannot be reverted efficiently to ESCs unless genetic or chemical manipulation is used, especially in non-permissive backgrounds (mouse strains from which ESCs could previously not be obtained, see above) (Hanna, 2009; Bao, 2009). This also points to epigenetic barriers being established in the ESC-to-EpiSC transition, which is also apparent by the distinct state of X-chromosome inactivation. X-chromosome inactivation is a mechanism of epigenetic inactivation of one of the X-chromosomes in mammalian (XX) females, to allow gene dosage compensation of X-linked genes compared to (XY) males (Fan, 2011). Female epiblast cells in the ICM and their derived mESCs have two active X-chromosomes (XaXa), but upon differentiation to EpiSCs, one X-chromosome of the two becomes randomly inactivated (Nichols, 2009). In line with the ability of the X-chromosome to indicate the naive or primed pluripotent state, EpiSCs that are reverted back to naive state were found to re-activate their X-chromosome (Bao, 2009).

The search for the key epigenetic barrier in the naive to primed transition has found no differences in the main epigenetic marks, namely histone (mainly H3) modifications and DNA-methylation (Takahashi, 2017). DNA-methylation levels are globally similar between ESCs and EpiSCs, although if divided in functional categories, promoter regions have higher methylation in primed cells than naive cells (Tosolini, 2018; Veillard, 2014). The functional meaning of these observations is questioned because *in vivo* embryos show lower levels of methylation, and increased methylation has been observed as a function of time in culture (Veillard, 2014; Meissner, 2008). In general, ESCs have a more open and accessible chromatin configuration (Azuara, 2006), which may partially explain how they are able to withstand the loss of multiple key epigenetic regulatory activities, such as those of the "maintenance" (upon cell division) DNA methyltransferase DNMT1, and EED and SUZ12, whereas EpiSCs cannot withstand this (Geula, 2015).

Naive cells further display unique epigenetic features. Genomic sites containing the active histone mark H3K4me3 and the repressive mark H3K27me3 simultaneously, therefore termed bivalent domains, have been described in mESCs (Hackett, 2014; Bernstein, 2006). Bivalent promoters are thought to be in a "poised" state, ready to be activated or silenced later on in the developmental program, as bivalent-associated genes show low expression but are resolved in differentiated cell types, either being upregulated and maintaining H3K4me3-only or being downregulated and maintaining H3K27me3-only (Bernstein,

2006). This is in line with the observation that half of methylated promoters in EpiSCs were bivalent in ESCs, where hypermethylation may be related to silencing (Veillard, 2014).

Enhancers also show a unique feature in naive pluripotent cells. Enhancers were classified depending on their activated status, and it was observed that active enhancers in mESC-enriched genes have active enhancer marks in mESCs, but these active marks are erased in EpiSCs. On the other hand, active enhancers in EpiSCs have an active signature both in EpiSCs and, albeit less enriched, in mESCs (Factor, 2014). The latter were called *seed enhancers* and it was shown that almost double of seed enhancers became active in analysis of differentiated tissues, compared with naive enhancers, indicating that seed enhancers may be primed for downstream activation (Factor, 2014). Another interesting observation was the differential use of enhancers in genes that are similarly expressed between ESCs and EpiSCs (Factor, 2014). A well-studied case happens with the *Oct4* gene during naive

to primed transition. *Oct4* regulation switches from a distal enhancer to a proximal enhancer in naive and primed cells, respectively, and this fine-tuning of the *Oct4* gene regulation has also been observed *in vivo* (Yeom, 1996; Choi, 2016). Finer mapping of histone modification changes between naive and primed cells are needed to understand how changes in the enhancer landscape contribute to the transition.

Epigenetic mechanisms are dependent on another change during the naive-to-primed transition, i.e., on metabolism. Metabolic activity produces reactants needed for epigenetic regulation. ESCs have a bivalent metabolism that uses both anaerobic glycolysis and oxidative phosphorylation, whereas EpiSCs primarily use glycolysis (Tsogtbaatar, 2020). This difference is not only relevant, for it establishes a difference between these pluripotent states that possibly reflects the different environments where they are present (blastocyst to the post-implantation epiblast in the hypoxic uterine wall), but also because metabolic activity does impact the naive-to-primed transition (Tsogtbaatar, 2020). For example, the key pluripotency gene *Lin28a/b* is involved in the control of glycolysis, and deletion of *Mtch2*, a key gene that regulates mitochondria fusion and elongation, also leads to a delayed ESC-to-EpiSC transition (Tsogtbaatar, 2020).

The establishment of *in vitro* cultures of mESCs and EpiSCs has allowed the characterization of pluripotent stem cell states in the early embryo. However, there are still limitations: 2D cultures do not reproduce the morphological changes that happen during development, and they rarely include interactions with additional cell types. These limitations have been slowly overcome by the use of 3D systems, such as growing mESCs embedded in Matrigel (Bedzhov, 2014). ESCs in this system form a polarized rosette and, later on, a cavity or lumen, similarly to Matrigel-embedded ICMs (Bedzhov, 2014). This approach can be used to investigate the relationship between pluripotency and morphogenetic change.

Previous research shows that the naive pluripotency network is disassembled at the onset of lumenogenesis and that maintenance of naive gene expression leads to incomplete lumen formation (Shahbazi, 2017). Additionally, 3D cultures of trophoblast stem cells and ESCs have given rise to blastocyst-like structures, called blastoids, which can be used to interrogate the relationship between different cell populations in the embryo (Rivron, 2018). The diversity and functional relevance of 3D systems for early embryonic development has been reviewed elsewhere (Posfai, 2021).

GENE REGULATORY NETWORKS MAINTAINING PLURIPOTENCY

Even though mESCs and EpiSCs differ in genetic, epigenetic and functional characteristics, they share their pluripotent hallmark. Pluripotency (P) is orchestrated and maintained by gene regulatory networks (pluripotent GRNs, hence PGRNs). The consensus is that OCT4, SOX2 and NANOG are the core pluripotency TFs from this network, although NANOG is not a general pluripotency factor, as will be discussed below.

Genetic inactivation of *Oct4*, *Nanog* and *Sox2* each leads to defects in the embryo, which display abnormal or inexistent epiblasts (Nichols, 1998; Mitsui, 2003; Avilion, 2003). Of these encoded factors, OCT4 and SOX2 are often thought as the main players for pluripotency, as their KO ESCs cannot be generated (Masui, 2007; Nichols, 1998). In contrast, *Nanog*-KO ESC clones and lines can be generated, even though their pluripotent status is affected, and display reduced expression of pluripotency marker genes and an upregulation of lineage marker genes (Mitsui, 2003). As an attestation to their key role in pluripotency, a cocktail containing cDNA-based production of OCT4 and SOX2, together with KLF4 and c-MYC (often referred to as OSKM), has been successfully used to reprogram terminally differentiated fibroblasts into cells with pluripotency features, enabling to obtain (self-renewable) induced pluripotent stem cells (iPSCs) (Takahashi, 2006). Although this is the paradigm-shifting seminal study, later on the same group reported efficient generation of iPSCs without c-MYC by adjusting the timing of the selection of reprogrammed colonies (Nakagawa, 2008). Although NANOG is not included in the reprogramming cocktail, studies have shown that *Nanog*-null cells are only able to generate partially reprogrammed colonies, and that NANOG co-occupies multiple binding sites of the TFs (often co-operating) OCT4 and SOX2, therefore being considered a core factor of the PGRN (Silva, 2009; Chen, 2008).

How these TFs regulate pluripotency is not totally clear. An analysis of the core pluripotency factors together with additional TFs implicated in pluripotency has shown that multiple such pluripotency factors bind the same loci more frequently than expected by chance

(Chen, 2008). In addition, such factors often interact with each other at the protein level, indicating a degree of co-operativity (van den Berg, 2010). There are multiple genomic loci co-occupied by OCT4 and SOX2 binding, and a large fraction of these sites are also bound by NANOG (Li, 2017). In particular the co-operative binding of OCT4 and SOX2 has been studied intensely. OCT4 is a member of the Oct family of proteins (for octamer binding) which comprise 8 different TFs that also are part of the larger POU family of proteins, and they all bind the octamer sequence ATGCAAAT. They do so via a bipartite DNA-binding domain with two domains: an N-terminal POU-type domain and a C-terminal homeodomain (HD), and these two connected through a linker region. Additionally, they also have two transactivation domains. Most Oct proteins bind as homodimers on the so-called MORE motif, which is enriched on somatic cells. On the contrary, OCT4 binds to the SoxOct motif, which is most prevalent in mESCs. In order to do so, it forms a heterodimer with SOX2 (Kim, 2021; Jerabek, 2016). The relevance of this Sox:Oct heterodimer binding has been experimentally confirmed. For example, a mutant OCT4 protein that cannot bind to SOX2 is accompanied by complete loss of its ability to generate iPSCs; OCT6 can become a reprogramming factor when the key residues from OCT4 that allow SOX2-binding are mutated (Jerabek, 2016). Interestingly, the disruption of SOX2 and OCT4 binding to their cognate DNA target sequences not only affects reprogramming, but also affects clonogenicity in ESC lines, pointing to a role in pluripotency maintenance (Tapia, 2015).

However, the binding of the core pluripotency factors does not correlate well with differential gene expression, so there may be additional TFs required, certainly in this cell model (Li, 2017). Such efforts in expanding the PGRN showed that there is a high degree of connectivity between pluripotent factors, and co-operative binding of NANOG, SOX2, DAX1, NAC1, OCT4 and KLF4 correlates with higher gene expression, whereas single-TF binding corresponds with mostly repressed genes (Kim, 2008).

How do the acknowledged core pluripotency factors promote pluripotency? Initial models postulated that pluripotent cells were protected from differentiation by the expression of pluripotent genes. However, it has been pointed out that this would make the system unlikely to be able to differentiate appropriately (Loh, 2011; Shu, 2013). Further, overexpression experiments (using cDNA) have shown that pluripotency factors act as lineage specifiers, and not blocking cell differentiation. *Oct4* overproduction causes stem cells to differentiate towards an endoderm-like lineage, whereas *Sox2* overproduction triggers an increase in ectoderm and mesoderm markers and a reduction in pluripotent markers (Kopp, 2008; Niwa, 2000). Remarkably, OCT4 can also be substituted by (the otherwise differentiation factor) GATA3 for reprogramming, seemingly via prevention of differentiation into ectodermal or mesodermal lineages (Shu, 2013). *Sox2* overproduction causes differentiation towards neuroectodermal lineages, and can also be substituted by other lineage specifiers such as SOX1 and SOX3 (also studied in other embryos, including *Xenopus* embryos, as relevant to

neurogenesis), and NCOR2 (Shu, 2013). Reprogramming of fibroblasts via substitution of core pluripotency genes with known lineage specifiers has also been proven in human cells (Montserrat, 2013). These results are more consistent with a model in which pluripotency factors are lineage specifiers that maintain pluripotency when properly balanced, rather than a model in which they prevent differentiation.

Although the studies of somatic cell reprogramming have produced relevant data on the GRNs that are key for establishment of pluripotency, they are also limited in their scope. Reprogramming and embryogenesis, although similar in some ways, are still distinct processes, so how relevant the insights gathered from reprogramming are for the genesis of pluripotency during development needs to be carefully considered and experimentally assessed. Further, reprogramming studies focus on the reversal of somatic cells to naive pluripotency, but do not include the distinct pathways required for maintenance of naive and primed pluripotency, and transition to primed pluripotency. As a key example, NANOG is considered a core pluripotency factor, but *Nanog*-null EpiSCs can be generated from ESCs, indicating that NANOG is not required for primed pluripotency and therefore should not be considered that strictly as a core pluripotency factor (Osorno, 2012).

MAINTENANCE OF NAIVE PLURIPOTENCY BY EXTRINSIC SIGNALS

The existence of naive and primed stem cell pluripotent states highlights the relevance of research not only on pluripotency acquisition and maintenance, but also into the signalling pathways and downstream molecular players that regulate these transitions, in order to better define these cultures *in vitro* and to understand pluripotent transitions in the embryo.

Leukemia inhibitory factor (LIF) signalling

Initially, mESCs were obtained in a feeder layer of mitotically inactivated fibroblasts (MEFs) in the presence of serum (Martin, 1981; Evans, 1981). The molecule secreted by MEFs that supported mESC culture was later identified as LIF (Moreau, 1988; Smith, 1988; Williams, 1988). LIF is a member of the interleukin-6 (IL-6) family of cytokines, which includes LIF, IL-6, IL-11, IL-27, CNTF, CT-1 and OSM. LIF binds to its membrane receptor LIF-R (also known as CD118, or LIF-R alpha) and induces its heterodimerization with the high-affinity transmembrane glycoprotein 130 (gp130). Both of these receptors do not have intrinsic catalytic tyrosine (Tyr) or serine/threonine (Ser/Thr) kinase activity, but are associated with janus kinases (JAKs), and LIF binding leads to the phosphorylation and activation of JAKs. These in turn phosphorylate Tyr residues on gp130 and LIF-R. These phospho-sites are then able to recruit and phosphorylate proteins that contain SH2 domains, such as STAT3 (a member of the STAT protein family). STAT3 forms a dimer and translocates into

the nucleus, where it can directly bind to DNA and act as a TF (Onishi, 2015; Ohtsuka, 2015) (**Figure 3**). LIF signalling is also controlled by negative feedback via repressors of this pathway, such as SOCS3 and PIAS3 (Ohtsuka, 2015). Mutation of the Tyr residues that mediate STAT3 interaction has an impact in ESC self-renewal, providing a direct link between the JAK/STAT pathway and stem cell self-renewal (Niwa, 1998).

Signalling via LIF thus activates the JAK/STAT pathway, but it also activates the MAPK and PI3K-Akt pathways. LIF activates MAPK via Grb2/SOS, and the PI3K-Akt pathway via JAKs. Consequently, it has been reported that inactivation of these pathways, independently of the JAK/STAT pathway, also allows ESC self-renewal in the absence of LIF (Ohtsuka, 2015). However, other groups have found that MAPK signalling through LIF is not relevant for self-renewal of ESCs, and in contrast, that STAT3 activation alone is sufficient to maintain the undifferentiated phenotype (Matsuda, 1999).

Although the role of LIF signalling in maintenance of ESCs *in vitro* had been established, apparently normal embryos could be derived until E6.5 with inactivated *Stat3* (Do, 2013). A role of LIF signalling in the embryo was first described during diapause. Diapause is a developmental delay that occurs in some species as a result of adverse environmental conditions, causing the arrest of the embryo in the pre-implantation blastocyst stage. Embryos harbouring a knock-out of *Lifr* and/or *gp130* could not be recovered, or only at very low efficiencies, from females in which diapause was induced (Nichols, 2001). In addition, the ICM of these blastocysts was defective, with overall less cells and loss of pluripotency markers (Nichols, 2001).

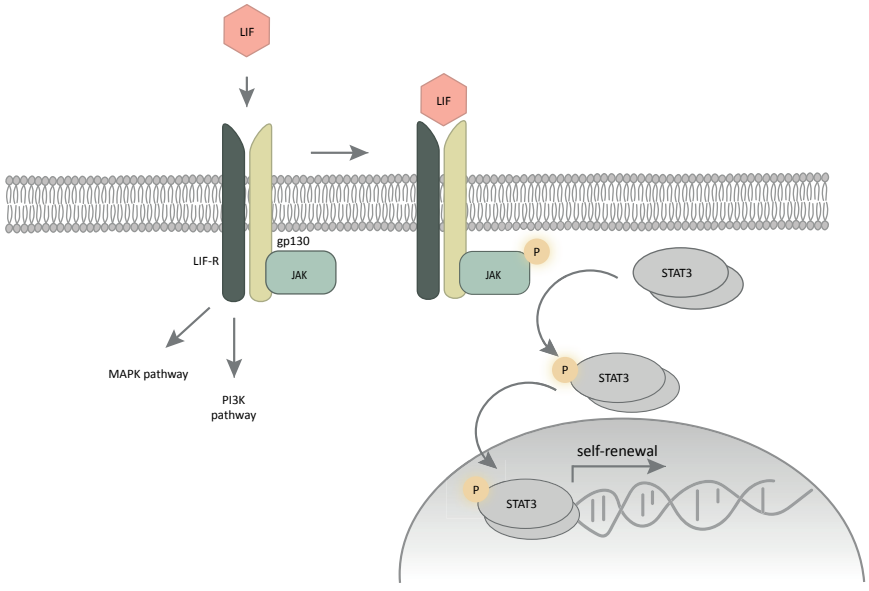


Figure 3. A schematic overview of the LIF pathway activation and function.

A more general role of LIF signalling *in vivo* could be observed in *Stat3*-KO embryos that were obtained from mothers that also had mutant *Stat3*, thereby eliminating any influence of maternal STAT3 in early embryogenesis (Do, 2013). These embryos appear normal if recovered at E3.5, but if recovered at later stages the markers for ICM and TE are abnormal, and no ESC lines can be derived from them (Do, 2013). It was postulated that STAT3 may orchestrate its effect in pluripotency maintenance through direct regulation of *Oct4* gene transcription (Do, 2013).

There are multiple key TFs that mediate LIF-controlled cell self-renewal, which is in line with its activity upstream of multiple signalling pathways. The main TF associated with STAT3 LIF-mediated self-renewal is TFCP2L1 (Martello, 2013). Consequently, exogenous expression of *Tfcp2l1* from cDNA can rescue the defects in *Stat3*-null ESCs, which can then be propagated in an undifferentiated state, whereas knock-down of TFCP2L1 leads to a reduction of pluripotency markers and clonogenicity (Martello, 2013). The JAK/STAT pathway is also associated with regulation of key pluripotency factors *Klf4* and *Sox2*, whereas PI3K-Akt signalling may regulate *Tbx3* and *Nanog* mRNA levels. Therefore, LIF signalling may contribute to pluripotency maintenance with different gene networks using parallel pathways (Niwa, 2009).

A confounding factor in ESC culture up until this point was the presence of serum in the media, the composition of which is not fully known and can vary. This could not rule out other signalling requirements for pluripotency maintenance, as LIF alone was only able to support stem cell culture at moderate to high densities. In order to address that, early studies used a combination of bone morphogenetic protein (BMP) in combination with LIF in defined N2B27 media (Ying, 2003). The combination of LIF and BMP2/4 was soon replaced by a new defined culture media that combined a GSK3 inhibitor that acts as WNT pathway agonist (i.e., CHIR99021), and a MEK pathway inhibitor (i.e., PD0325901), allowing the culture of ESCs even in the absence of LIF (Ying, 2008). This culture media was called "2i" referring to the simultaneous use of the two inhibitors and led to what was defined as a (*in vitro*-only) "ground" state of pluripotency, since the authors claimed that the cells were in a basal state that was "intrinsically self-maintaining" (Ying, 2008).

WNT signalling

Activation of the WNT pathway through CHIR99021 affects the canonical part of the pathway, which is mediated by b-catenin (b-cat), encoded by *Ctnnb1* (**Figure 4**). In absence of WNT ligands, cytoplasmic b-cat is phosphorylated by the destruction complex, which comprises APC, AXIN and GSK3B (Yang, 2016). Phosphorylated b-cat is targeted for proteasomal degradation, so its cytoplasmic levels can be kept low (Yang, 2016; Wray, 2012). When WNT ligands are present, they interact with a membrane receptor complex formed by Frizzled and the low-density lipoprotein receptor-related protein (LRP) family members

LRP5/6 (Wray, 2012). Autocrine secretion of WNT ligands is possible and it involves Porcupine (PORCN), which allows acylation and secretion of WNT proteins. The direct effect of WNT interaction with its membrane receptors is the inhibition of the destruction complex as well as the accumulation of non-phosphorylated b-cat in the cytoplasm. This b-cat can then accumulate in the nucleus, where it interacts with its target TFs from the TCF/LEF family (Wray, 2012). The traditional model views b-cat as a co-activator, as TCF/LEF TFs are in complex with co-repressors until they interact with b-cat in the nucleus at the target DNA-binding sites.

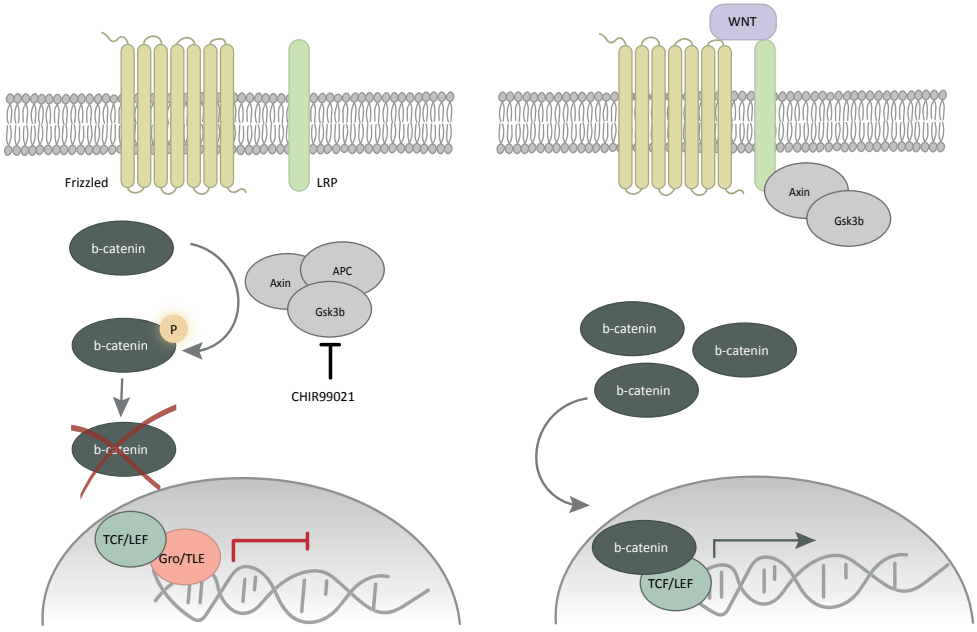


Figure 4. Scheme of the WNT pathway in the absence *(left panel)* **and presence** *(right panel)* **of WNT.**

A lot of research efforts have been invested in establishing the relationship between the WNT pathway and the maintenance of pluripotency. Previous studies showed no defects of *Ctnnb1* mutant embryos previous to gastrulation, casting some doubt on whether the effects of the WNT pathway were due to off-targets of, for example, CHIR99021 (Rudloff, 2012; Huelsken, 2000). In addition, ESCs lacking CTNNB1 appeared normal under self-renewing conditions, and the cells continued to express the genes encoding core pluripotency markers, even when the whole *Ctnnb1* gene body was excised (Rudlof, 2012; Wray, 2012; Aulicino, 2020).

Newer insights start to provide a finer description of the role of b-cat, and therefore, the WNT pathway, on pluripotency. On the one hand, b-cat appears to be essential for maintenance

of the pluripotent epiblast during embryonic diapause, and WNT (canonical) reporter activity has been detected in pre-implantation epiblast (Fan, 2020; ten Berge, 2011). On the other hand, defects in ESC self-renewal have definitely been attributed to WNT through the canonical pathway. When ESCs containing a *Ctnbb1* deletion were compared to wild-type cells cultured in either serum/LIF or CHIR99021/LIF, wild-type cells in serum/LIF clustered together with b-cat-null cells either in serum/LIF or CHIR99021/LIF, pointing to a requirement for b-cat to mediate the effects of CHIR99021 (Aulicino, 2020). Expansion of ESCs was observed with a panel of Gsk3 inhibitors, and deletions in *Ctnnb1* were associated with differentiation and not self-renewal, in response to CHIR99021/LIF or in the presence of 2i (Wray, 2011). In this line, endogenous WNT signalling is key to maintain ESCs, as ESCs differentiate into EpiSCs when treated with an inhibitor of Porcupine (ten Berge, 2011).

What are the transcriptional targets of nuclear b-cat and how does it regulate transcription? There are 4 *Tcf/Lef* genes in mouse: *Tcf3* (or *Tcf7l1*), *Tcf1* (or *Tcf7*), *Tcf4* (or *Tcf7l2*; not to be confused with the ubiquitous bHLH-TF E2-2, also named *Tcf4*) and *Lef1*. TCF7L1 is the most abundant TCF/LEF protein in mouse (60%), followed by TCF1(23%), LEF1(12%) and TCF4(5%) (Pereira, 2006). In contrast with the classical model, the current model views TCF7L1 as a repressor of the pluripotency network whose repressive function is suppressed by b-cat (Merrill, 2012). TCF7L1 binds to the enhancer regions of *Nanog*, and repression of *Nanog* by TCF7L1 is dose-dependent (Pereira, 2006). This function of TCF7L1 depends on its DNA-binding and its Groucho-binding domains, as *Nanog* cannot be repressed if those domains are deleted from the *Tcf7l1* gene (Pereira, 2006). Cells that contain a critical-exon *Tcf7l1* deletion are insensitive to CHIR99021 and form alkaline phosphatase+ colonies in absence of CHIR99021, whereas TCF7L1 overproduction suppresses formation of such colonies, which is restored upon stimulation with Wnt3a (Wray, 2011; Yi, 2011). Although this points to a key role of TCF7L1 in mediating the WNT response, a putative role for other TCF family members is emerging. Some authors suggested that TCF7 may be able to substitute for TCF7L1 in differentiating conditions (Moreira, 2017), while others noted that TCF1 is required for full activation of WNT target genes (Yi, 2011).

MAPK/ERK signalling, including FGF signalling

For the maintenance of ESCs in culture, another determining pathway is the MAPK/ERK pathway, which needs to be inhibited to prevent differentiation (Ying, 2008). As mentioned before, this pathway can receive input from LIF signalling, but the most relevant upstream activation of the MAPK pathway in pluripotent stem cells is initiated by FGF family signals. FGF signalling further activates three main pathways: MAPK/ERK, PI3K and PLC-y. We will focus on response through MAPK/ERK, as this has been shown to be the key transduction pathway (Ying, 2008).

In mammals, there are four FGF receptor (FGFR) encoding genes (*Fgfr1-4*) and most of them can generate multiple receptor isoforms through alternative splicing, determining the affinity of each of the >20 FGF ligands. FGFRs are receptor Tyr-kinases that contain an extracellular ligand-binding domain, a single transmembrane domain, and an intracellular cytoplasmic domain that has regulatory regions and a characteristic split, catalytic tyrosine kinase core. Binding of FGFs to their receptors causes dimerization of FGFRs. X-ray structures of this interaction show that it is stabilized both by FGF and receptor-receptor mediated interactions, but also by interaction with heparan sulfate proteoglycans (HSPG), which bind to a charged region of the extracellular domain on the FGFRs. Upon receptor dimerization, the tyrosine kinase activity of the receptor is activated, leading to autophosphorylation of the tyrosine residues in the receptor, which recruits the docking proteins FRS2α and FRS2β through their PTB (phosphotyrosine binding) domain. FRS2 proteins become then phosphorylated and lead to recruitment of SHP2, GRB2 and GAB1, which are key proteins for activation of downstream effector pathways RAS-MAPK and PI3K-Akt. After assembly of this Frs2 signalling complex, the guanine nucleotide exchange factor (GEF) SOS is recruited through binding with GRB2, which then leads to the activation of RAS through exchange of GTP for GDP (Dailey, 2005; Eswarakumar, 2005; Schlessinger, 2000) (**Figure 5**).

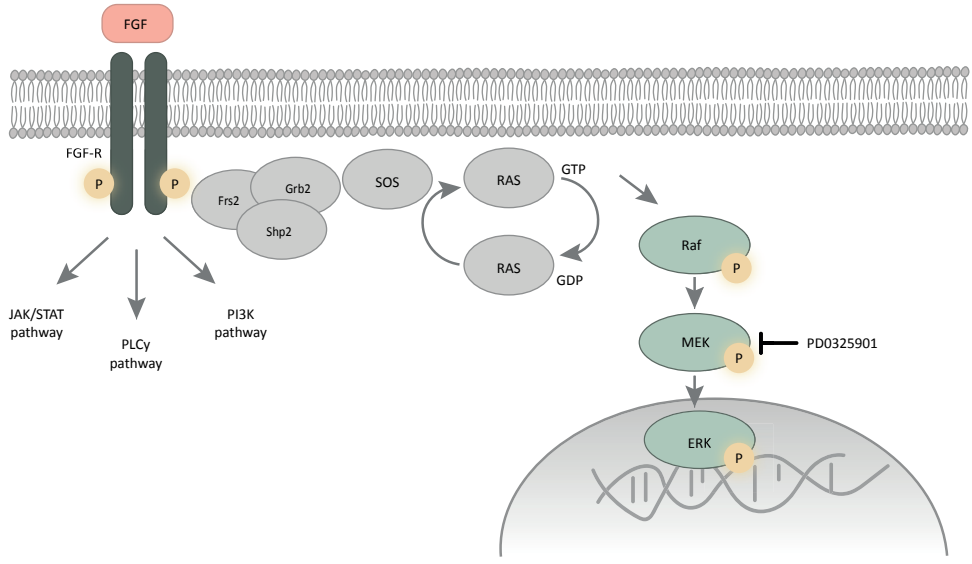


Figure 5. Simplified representation of the FGF signalling pathway.

There are 3 levels within the MAPK activation cascade, which are evolutionarily conserved: the MAPK, the MAPK kinase (MAPKK or MEK) and the MAPKK kinase or MEK kinase (MAPKKK or MEKK). Mammals have 4 groups of MAPKs: extracellular signal-related

kinases (ERK)-1/2, Jun amino-terminal kinases (JNK1/2/3), p38 proteins (p38a/b/g/d) and ERK5. Active RAS interacts with and activates RAF (a MAPKKK), which then interacts with the MAPKK MEK1/2 and the latter phosphorylates Tyr and Thr residues on the MAPKs ERK1/2 (Chang, 2001; Schlessinger, 2000). Phosphorylation of ERK1/2 through this kinase cascade leads to translocation of the latter to the nucleus and phosphorylation of its targets (e.g., ELK1, c-FOS, c-JUN) (Ma, 2016; Schlessinger, 2000) (**Figure 5**). The relevance of this pathway for *in vivo* pluripotency is not yet defined, as *Erk1/2* double KO (dKO) ESCs cannot be generated and *Erk1* or *Erk2* single deletions allow embryogenesis to progress through implantation stages (Ma, 2016).

Data coming from pharmacological inhibition, such as the 2i system discussed above, supported a model in which ERK is not necessary for ESC maintenance, as these cells can be maintained in MEK inhibitor. However, genetic manipulation data shows that may not be the case. As previously mentioned, *Erk1/2*-dKO ESCs cannot be generated, and simultaneous Raf-A/B/C knockdown impacts ESC proliferation and survival (Ma, 2016; Chen, 2015). When inducible systems that allowed transient culture of ESCs without ERK1 or ERK2 were created, cultures crashed within a few passages, and apoptosis and telomere shortening were detected in the cells (Chen, 2015).

The MAPK/ERK pathway is seemingly not only relevant for cell integrity, but also for exit from pluripotency. FGF4 acting through the MAPK/ERK pathway is essential to allow mESCs to exit from their pluripotent state and undergo neuronal and general (including mesodermal) differentiation (Kunath, 2007). A period of ERK signalling seems also key to allow ESCs to differentiate, an effect that was not rescued through signalling by alternate FGF-activated pathways PI3K-Akt or PLC-y (Stavridis, 2007). A role for ERK2 in establishing a poised state on developmental gene promoters, with accumulation of H3K27me3 and poised RNA Polymerase-II, allowing fast activation during the transition to a primed state, has also been described (Tee, 2014). However, studies following ESCs exiting from naive pluripotency in real time were not able to correlate p-ERK dynamics and NANOG (Deathridge, 2019). These findings argue that suppression of the MEK/ERK pathway may maintain a naive pluripotent cell state and prevent differentiation.

NAIVE AND GROUND PLURIPOTENCY

Although "ground" pluripotency is commonly used together with or to refer to naive pluripotency, it is important to distinguish between ground (cultured in 2i) and conventional (cultured in serum) naive pluripotency (**Table 2**). Since naive pluripotency is defined on the basis of the ability to generate all embryonic lineages following blastocyst injection, both serum and 2i-cultured cells have the potential to be considered naive (Hackett, 2014).

However, studies have shown that conventional naive stem cells switched to 2i media undergo profound changes in their transcriptome and epigenome, showing that there are differences between these two conditions (Marks, 2012; Ghimire, 2012; Galonska, 2015).

Table 2. Comparison between ground and naive pluripotency.

Property	Conventional naive pluripotency	Ground state pluripotency
Culture media	LIF + serum or MEFs + serum	2i (WNT agonist and MEK inhibitor)
Chimaera formation	Yes	Abrogated at high passage
Metastability	Yes	No
Methylation	High	Low
Bivalency	High	Low

Ground stem cells were thought to represent more faithfully the ICM of the blastocyst, as methylation levels were reduced compared to conventional ESCs without aberrant increased expression of satellite regions, and their transcriptome clustered more closely to that of the pre-implantation epiblast (Tosolini, 2018: Leitch, 2013; Boroviak, 2014). However, it has been later shown that culture of mESCs in L2i (LIF and 2i) leads to widespread demethylation and imprint loss due to exposure to MEK1/2 inhibition, which is associated with a lower contribution to chimaera formation or tetraploid complementation, and therefore loss of pluripotency in these cells (Choi, 2017; Yagi, 2017).

Ground state stem cells also display transcriptional and epigenetic differences with conventional naive stem cells. Although most stem cell maintenance genes are similarly expressed between cells cultured in serum and 2i, some pluripotency genes are differentially expressed; serum cells also show increased expression of lineage genes (Marks, 2012; Ghimire, 2018). Bivalency is also reduced when cells are grown in 2i, compared to serum (Marks, 2012; Galonska, 2015). H3K27me3 levels are reportedly higher in serum (Marks, 2012). In line with the differences seen in epigenetic characteristics, *Oct4* regulation shows a change from the proximal element in serum to the distal element in 2i (Galonska, 2015). As mentioned before, methylation levels are higher in serum+LIF cells (both REX1-high and REX1-low, therefore this effect is not explained by metastability of the culture) compared to L2i cells (Habibi, 2013). Both transcriptome and epigenome differences seem to be interconvertible when serum cells are switched to 2i media or vice versa (Marks, 2012).

EXIT FROM NAIVE PLURIPOTENCY AND TRANSITION TO PRIMED

The establishment of defined culture conditions for naive stem cells also generated interest on the key molecular players responsible for dismantling the pluripotency network to initiate differentiation and lineage commitment. A series of studies looked into key genes for this exit in a high-throughput unbiased manner, using reporter cell lines and siRNA libraries, transposon gene traps or, more recently, CRISPR-Cas9. These approaches were able to, on one hand, support the role for the WNT and MEK pathways on the maintenance of pluripotency, since multiple mutations affecting the exit from pluripotency were detected in components of those pathways (Leeb, 2014). It is particularly noteworthy that *Tcf7l1* was a hit in multiple screens (Leeb, 2014; Betschinger, 2013; Yang, 2020). Multiple genes have been found to affect pluripotency exit, such as those encoding epigenetic modifiers PTEN and RBPJ, PUM1, ZFP706, Folliculin (FLCN) and TFE3 (Betschinger, 2013; Leeb, 2014).

These studies also detected Orthodenticle Homeobox 2 (*Otx2*) as a key gene for pluripotency progression. OTX2 protein levels increase upon exit from naive pluripotency and in the embryo shortly after implantation, which corresponds with the loss of naive pluripotency in the embryo (Acampora, 2013; Acampora, 2016). Increased activation of target genes for OTX2 has been observed upon exit from naive pluripotency, arguing towards a role for OTX2 directing this process (Yang, 2014). Cells that contain a homozygous deletion in the *Otx2* gene are impaired in entering the EpiSC state and become unstable upon differentiating (Acampora, 2013). OTX2 seems to have multiple roles during this transition. On one hand, OTX2 binds upstream of *Nanog*, an interaction that seems to be key for the maintenance of the naive pluripotency state (Acampora, 2016). On the other, OTX2 binds to regions upon exit from naive pluripotency and causes them to gain H3K27Ac, a mark of active enhancers, or binds to regions previously bound by OCT4, maintaining their active state (Yang, 2014). OTX2 not only co-binds regions where OCT4 is found, but is also able to re-organize OCT4 occupancy in the genome and can partially pioneer the epigenetic landscape of the EpiLC transition when overproduced in 2i conditions (Buecker, 2014). This data together points to the crucial role of OTX2 in orchestrating the changes in the naive to primed transition.

Although most of the efforts to understand the molecular underpinnings of the naive to primed transition have focused on signalling pathways and key TF networks, new studies are highlighting the need for a more comprehensive approach. Recently, two groups have described a reduction in cell membrane tension during this transition and have proved a direct effect of membrane tension in the dissolution of naive pluripotency. Maintenance of high membrane tension via different mechanisms both increased the level of naive factors (e.g., NANOG, TFCP2L1) and slowed the dynamics of naive pluripotency exit (Bergert, 2021; DeBelly, 2021).

A direct consequence of the naive-to-primed transition studies is the need for an intermediate state. Two main reasons explain this requirement: naive cells are refractory to transcription factor determinants and inductive cues and primed cells have already initiated a response to differentiation cues (Smith, 2017). Therefore, there is a requirement for a "formative phase" in which extensive chromatin remodelling and increase in bivalent promoters leave the cells sensitized to differentiation cues (Smith, 2017).

Multiple protocols claim to have uncovered intermediate pluripotent states during the naive to primed transition. Cells with characteristics consistent of those of a formative phase have been detected 25 hours after release from 2i, showing changes in metabolic and epigenetic profile (Kalkan, 2017). An intermediate "poised" state controlled by key miRNA regulator ISY1 has also been described, and deletion of *Isy1* consequently impairs differentiation to primed state (Du, 2018). These cells express naive marker genes such as *Nanog* and *Esrrb*, but upregulate some other genes and a large subset of miRNAs (Du, 2018). In human pluripotent stem cells, an intermediate state has also been described when cells are cultured in a low-lipid media (Cornacchia, 2019). The main flaw of these studies is the lack of *in vivo* representation, as no counterpart to these states has been defined in the embryo and most rely only on comparisons of RNA-seq data to datasets from E4.5-E5.5 embryos.

CONCLUSIONS AND FUTURE DIRECTIONS

The development of pluripotency is tightly regulated and timed, accompanied by multiple relevant processes for successful embryogenesis, such as changes on the epigenome and transcriptome and a morphogenetic transformation from an unorganized mass of cells to an open lumen. Pluripotency is therefore not a static state, but is an umbrella term for multiple states, of which naive and primed pluripotency have been captured *in vitro* as ESCs and EpiSCs. There is a solid rationale to believe that additional intermediate states between naive and primed pluripotency exist and are needed for pluripotent stem cells to adopt the epigenetic and genetic makeover required for lineage decisions. However, although efforts are being made to describe intermediate pluripotent states, their representability of states *in vivo* is unclear.

Identification of intermediate pluripotent states is dependent on the knowledge available on signalling pathways relevant for pluripotent stem cells. The transition from previously metastable, heterogeneous ESC cultures into defined media allows in-depth dissection of the signalling pathways and key effectors that allow the transition between naive and primed pluripotency and of the key characteristics of each state. The discovery that pluripotency maintenance, as determined by the ability of stem cells to contribute to embryo development after blastocyst injection, may be compromised following extended culture, suggests that

there are still gaps in our understanding of pluripotency in the dish. Advances in sequencing techniques, such as ChIP-seq and single-cell (sc) RNA-seq, in order to be able to reliably apply them to the *in vivo* embryo, will be needed to fully understand how pluripotency develops *in vivo* and how accurate *in vitro* cultures are of the *in vivo* states they represent.

Finally, pluripotency in the embryo does not progress in a 2D, isolated environment. Pluripotent stem cells undergo a morphogenetic transformation, and receive inputs from additional cell types. Improvement in 3D models *in vitro* allows to start resolving pending open questions in the field, such as the relationship between pluripotency and morphogenesis progression, and the requirement for extra-embryonic signals for the progression of pluripotency and/or morphogenesis in the embryo.

REFERENCES

Acampora D, Di Giovannantonio LG, Simeone A. Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition. Dev. 2013;140(1):43-55.

Acampora D, Omodei D, Petrosino G, et al. Loss of the Otx2-Binding Site in the Nanog Promoter Affects the Integrity of Embryonic Stem Cell Subtypes and Specification of Inner Cell Mass-Derived Epiblast. Cell Rep. 2016;15(12):2651-2664.

Arnold SJ, Robertson EJ. Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. Nat Rev Mol Cell Biol. 2009 Feb;10(2):91-103.

Atienza-Samols SB, Sherman MI. Outgrowth promoting factor for the inner cell mass of the mouse blastocyst. Dev Biol. 1978;66(1):220-231.

Aulicino F, Pedone E, Sottile F, Lluis F, Marucci L, Cosma MP. Canonical Wnt Pathway Controls mESC Self-Renewal Through Inhibition of Spontaneous Differentiation via β-Catenin/TCF/LEF Functions. Stem Cell Reports. 2020;15(3):646-661.

Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R. Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev. 2003 Jan 1;17(1):126-40.

Azuara V, Perry P, Sauer S, et al. Chromatin signatures of pluripotent cell lines. Nat Cell Biol. 2006;8(5):532-538.

Bao S, Tang F, Li X, et al. Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. Nature. 2009;461(7268):1292-1295.

Bedzhov I, Zernicka-Goetz M. Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. Cell. 2014;156(5):1032-1044.

De Belly H, Stubb A, Yanagida A, Labouesse C, Jones PH, Paluch EK, Chalut KJ. Membrane Tension Gates ERK-Mediated Regulation of Pluripotent Cell Fate. Cell Stem Cell. 2021 Feb 4;28(2):273-284.e6.

Berge D Ten, Kurek D, Blauwkamp T, et al. Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. Nat Cell Biol. 2011;13(9):1070-1077.

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell. 2006 Apr 21;125(2):315-26.

Bergert M, Lembo S, Sharma S, Russo L, Milovanović D, Gretarsson KH, Börmel M, Neveu PA, Hackett JA, Petsalaki E, Diz-Muñoz A. Cell Surface Mechanics Gate Embryonic Stem Cell Differentiation. Cell Stem Cell. 2021 Feb 4;28(2):209-216.e4.

Betschinger J, Nichols J, Dietmann S, Corrin PD, Paddison PJ, Smith A. Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. Cell. 2013;153(2):335-347.

Boroviak T, Nichols J. The birth of embryonic pluripotency. Philos Trans R Soc B Biol Sci. 2014;369(1657).

Boroviak T, Loos R, Bertone P, Smith A, Nichols J. The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. Nat Cell Biol. 2014 Jun;16(6):516-28.

Bradley A, Evans M, Kaufman MH, Robertson E. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature. 1984 May 17-23;309(5965):255-6.

Brons IG, Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA, Vallier L. Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature. 2007 Jul 12;448(7150):191-5.

Buecker C, Srinivasan R, Wu Z, et al. Reorganization of enhancer patterns in transition from naive to primed pluripotency. Cell Stem Cell. 2014;14(6):838-853.

Chang L, Karin M. Mammalian MAP kinase signalling cascades. Nature. 2001 Mar 1;410(6824):37-40.

Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, Loh YH, Yeo HC, Yeo ZX, Narang V, Govindarajan KR, Leong B, Shahab A, Ruan Y, Bourque G, Sung WK, Clarke ND, Wei CL, Ng HH. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell. 2008 Jun 13;133(6):1106-17.

Chen H, Guo R, Zhang Q, et al. Erk signaling is indispensable for genomic stability and self-renewal of mouse embryonic stem cells. Proc Natl Acad Sci U S A. 2015;112(44):E5936-E5943.

Choi J, Huebner aaron J, Clement K, et al. Prolonged Mek1/2 suppression impairs the developmental potential of embryonic stem cells. Nat Publ Gr. 2017;548.

Choi HW, Joo JY, Hong YJ, et al. Distinct Enhancer Activity of Oct4 in Naive and Primed Mouse Pluripotency. Stem Cell Reports. 2016;7(5):911-926.

Cornacchia D, Zhang C, Zimmer B, et al. Lipid Deprivation Induces a Stable, Naive-to-Primed Intermediate State of Pluripotency in Human PSCs. Cell Stem Cell. 2019;25(1):120-136.e10.

Dailey L, Ambrosetti D, Mansukhani A, Basilico C. Mechanisms underlying differential responses to FGF signaling. Cytokine Growth Factor Rev. 2005 Apr;16(2):233-47.

Deathridge J, Antolović V, Parsons M, Chubb JR. Live imaging of erk signalling dynamics in differentiating mouse embryonic stem cells. Dev. 2019;146(12).

De Belly H, Stubb A, Yanagida A, Labouesse C, Jones PH, Paluch EK, Chalut KJ. Membrane Tension Gates ERK-Mediated Regulation of Pluripotent Cell Fate. Cell Stem Cell. 2021 Feb 4;28(2):273-284.e6.

De Los Angeles A, Ferrari F, Xi R, Fujiwara Y, Benvenisty N, Deng H, Hochedlinger K, Jaenisch R, Lee S, Leitch HG, Lensch MW, Lujan E, Pei D, Rossant J, Wernig M, Park PJ, Daley GQ. Hallmarks of pluripotency. Nature. 2015 Sep 24;525(7570):469-78.

Do DV, Ueda J, Messerschmidt DM, et al. A genetic and developmental pathway from STAT3 to the OCT4-NANOG circuit is essential for maintenance of ICM lineages in vivo. Genes Dev. 2013;27(12):1378-1390.

Du P, Pirouz M, Choi J, et al. An Intermediate Pluripotent State Controlled by MicroRNAs Is Required for the Naive-to-Primed Stem Cell Transition. Cell Stem Cell. 2018;22(6):851-864.e5.

Edwards RG. Personal pathways to embryonic stem cells. Reprod Biomed Online. 2002;4(3):263-278.

Eswarakumar VP, Lax I, Schlessinger J. Cellular signaling by fibroblast growth factor receptors. Cytokine Growth Factor Rev. 2005 Apr;16(2):139-49.

Evans MJ. The isolation and properties of a clonal tissue culture strain of pluripotent mouse teratoma cells. J Embryol Exp Morphol. 1972 Aug;28(1):163-76.

Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. Nature. 1981;292(5819):154-156.

Factor DC, Corradin O, Zentner GE, Saiakhova A, Song L, Chenoweth JG, McKay RD, Crawford GE, Scacheri PC, Tesar PJ. Epigenomic comparison reveals activation of "seed" enhancers during transition from naive to primed pluripotency. Cell Stem Cell. 2014 Jun 5;14(6):854-63.

Fan G, Tran J. X chromosome inactivation in human and mouse pluripotent stem cells. Hum Genet. 2011 Aug;130(2):217-22.

Fan R, Kim YS, Wu J, et al. Wnt/Beta-catenin/Esrrb signalling controls the tissue-scale reorganization and maintenance of the pluripotent lineage during murine embryonic diapause. Nat Commun. 2020;11(1).

Feldman B, Poueymirou W, Papaioannou VE, DeChiara TM, Goldfarb M. Requirement of FGF-4 for postimplantation mouse development. Science. 1995 Jan 13;267(5195):246-9.

Galonska C, Ziller MJ, Karnik R, Meissner A. Ground State Conditions Induce Rapid Reorganization of Core Pluripotency Factor Binding before Global Epigenetic Reprogramming. Cell Stem Cell. 2015 Oct 1;17(4):462-70.

Geula S, Moshitch-Moshkovitz S, Dominissini D, et al. m6A mRNA methylation facilitates resolution of naive pluripotency toward differentiation. Science. 2015 Feb 27;347(6225):1002-6.

Ghimire, S., Van der Jeught, M., Neupane, J. et al. Comparative analysis of naive, primed and ground state pluripotency in mouse embryonic stem cells originating from the same genetic background. Sci Rep 8, 5884 (2018).

Habibi E, Brinkman AB, Arand J, Kroeze LI, Kerstens HH, Matarese F, Lepikhov K, Gut M, Brun-Heath I, Hubner NC, Benedetti R, Altucci L, Jansen JH, Walter J, Gut IG, Marks H, Stunnenberg HG. Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. Cell Stem Cell. 2013 Sep 5;13(3):360-9.

Hackett JA, Azim Surani M. Regulatory principles of pluripotency: From the ground state up. Cell Stem Cell. 2014;15(4):416-430.

Hanna J, Markoulaki S, Mitalipova M, et al. Metastable Pluripotent States in NOD-Mouse-Derived ESCs. Cell Stem Cell. 2009;4(6):513-524.

Huang Y, Osorno R, Tsakiridis A, Wilson V. In Vivo differentiation potential of epiblast stem cells revealed by chimeric embryo formation. Cell Rep. 2012 Dec 27;2(6):1571-8.

Huelsken J, Vogel R, Brinkmann V, Erdmann B, Birchmeier C, Birchmeier W. Requirement for β-catenin in anterior-posterior axis formation in mice. J Cell Biol. 2000;148(3):567-578.

Jerabek S, Ng CK, Wu G, Arauzo-Bravo MJ, Kim KP, Esch D, Malik V, Chen Y, Velychko S, MacCarthy CM, Yang X, Cojocaru V, Schöler HR, Jauch R. Changing POU dimerization preferences converts Oct6 into a pluripotency inducer. EMBO Rep. 2017 Feb;18(2):319-333.

Kalkan T, Olova N, Roode M, et al. Tracking the embryonic stem cell transition from ground state pluripotency. Dev. 2017;144(7):1221-1234.

Kelly SJ. Studies of the developmental potential of 4- and 8-cell stage mouse blastomeres. J Exp Zool. 1977 Jun;200(3):365-76.

Kim J, Chu J, Shen X, Wang J, Orkin SH. An extended transcriptional network for pluripotency of embryonic stem cells. Cell. 2008 Mar 21;132(6):1049-61.

Kim KP, Han DW, Kim J, Schöler HR. Biological importance of OCT transcription factors in reprogramming and development. Exp Mol Med. 2021 Jun;53(6):1018-1028.

Kopp JL, Ormsbee BD, Desler M, Rizzino A. Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. Stem Cells. 2008 Apr;26(4):903-11.

Kunath T, Saba-El-Leil MK, Almousailleakh M, Wray J, Meloche S, Smith A. FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. Development. 2007;134(16):2895-2902.

Leeb M, Dietmann S, Paramor M, Niwa H, Smith A. Genetic exploration of the exit from self-renewal using haploid embryonic stem cells. Cell Stem Cell. 2014;14(3):385-393.

Leitch HG, Mcewen KR, Turp A, et al. Naive pluripotency is associated with global DNA hypomethylation. Nat Struct Mol Biol. 2013;20(3):311-316.

Li M, Belmonte JCI. Ground rules of the pluripotency gene regulatory network. Nat Rev Genet. 2017;18(3):180-191.

Loh KM, Lim B. A precarious balance: Pluripotency factors as lineage specifiers. Cell Stem Cell. 2011;8(4):363-369.

Ma X, Chen H, Chen L. A dual role of Erk signaling in embryonic stem cells. Exp Hematol. 2016;44(3):151-156.

Ma S, Meng Z, Chen R, Guan KL. The Hippo Pathway: Biology and Pathophysiology. Annu Rev Biochem. 2019 Jun 20;88:577-604.

Marks H, Kalkan T, Menafra R, Denissov S, Jones K, Hofemeister H, Nichols J, Kranz A, Stewart AF, Smith A, Stunnenberg HG. The transcriptional and epigenomic foundations of ground state pluripotency. Cell. 2012 Apr 27;149(3):590-604.

Martello G, Bertone P, Smith A. Identification of the missing pluripotency mediator downstream of leukaemia inhibitory factor. EMBO J. 2013 Oct 2;32(19):2561-74.

Martin GR. Isolation of a Pluripotent Cell Line from Early Mouse Embryos Cultured in Medium Conditioned by Teratocarcinoma Stem Cells (Embryonic Stem Cells/Inner Cell Masses/ Differentiation in Vitro/Embryonal Carcinoma Cells/Growth Factors). Vol 78.; 1981.

Masui S, Nakatake Y, Toyooka Y, Shimosato D, Yagi R, Takahashi K, Okochi H, Okuda A, Matoba R, Sharov AA, Ko MS, Niwa H. Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nat Cell Biol. 2007 Jun;9(6):625-35.

Matsuda T, Nakamura T, Nakao K, Arai T, Katsuki M, Heike T, Yokota T. STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. EMBO J. 1999 Aug 2;18(15):4261-9.

Meissner A, Mikkelsen TS, Gu H, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature. 2008;454(7205):766-770.

Merrill BJ. Wnt pathway regulation of embryonic stem cell self-renewal. Cold Spring Harb Perspect Biol. 2012;4(9):a007971.

Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell. 2003 May 30;113(5):631-42.

Montserrat N, Nivet E, Sancho-Martinez I, Hishida T, Kumar S, Miquel L, Cortina C, Hishida Y, Xia Y, Esteban CR, Izpisua Belmonte JC. Reprogramming of human fibroblasts to pluripotency with lineage specifiers. Cell Stem Cell. 2013 Sep 5;13(3):341-50.

Moreau JF, Donaldson DD, Bennett F, Witek-Giannotti J, Clark SC, Wong GG. Leukaemia inhibitory factor is identical to the myeloid growth factor human interleukin for DA cells. Nature. 1988 Dec 15;336(6200):690-2.

Moreira S, Polena E, Gordon V, et al. A Single TCF Transcription Factor, Regardless of Its Activation Capacity, Is Sufficient for Effective Trilineage Differentiation of ESCs. Cell Rep. 2017;20(10):2424-2438.

Mummery C, van de Stolpe A, Roelen BAJ, Clevers H, eds. Chapter 4 - Of Mice and Men: The History of Embryonic Stem Cells. In: Stem Cells (Second Edition). Second Edition. Academic Press; 2014:69-100.

Nakagawa M, Koyanagi M, Tanabe K, Takahashi K, Ichisaka T, Aoi T, Okita K, Mochiduki Y, Takizawa N, Yamanaka S. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat Biotechnol. 2008 Jan;26(1):101-6.

Nichols J, Zevnik B, Anastassiadis K, Niwa H, Klewe-Nebenius D, Chambers I, Schöler H, Smith A. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell. 1998 Oct 30;95(3):379-91.

Nichols J, Smith A. Naive and primed pluripotent states. Cell Stem Cell. 2009 Jun 5;4(6):487-92.

Nichols J, Chambers I, Taga T, Smith A. Physiological Rationale for Responsiveness of Mouse Embryonic Stem Cells to Gp130 Cytokines. Vol 128. The Company of Biologists; 2001.

Nishioka N, Yamamoto S, Kiyonari H, Sato H, Sawada A, Ota M, Nakao K, Sasaki H. Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos. Mech Dev. 2008 Mar-Apr;125(3-4):270-83

Nishioka N, Inoue K, Adachi K, Kiyonari H, Ota M, Ralston A, Yabuta N, Hirahara S, Stephenson RO, Ogonuki N, Makita R, Kurihara H, Morin-Kensicki EM, Nojima H, Rossant J, Nakao K, Niwa H, Sasaki H. The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. Dev Cell. 2009 Mar;16(3):398-410.

Niwa H, Burdon T, Chambers I, Smith A. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes Dev. 1998;12(13):2048-2060.

Niwa H, Miyazaki J, Smith AG. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet. 2000 Apr;24(4):372-6.

Niwa H, Ogawa K, Shimosato D, Adachi K. A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. Nature. 2009 Jul 2;460(7251):118-22.

Ohtsuka S, Nakai-Futatsugi Y, Niwa H. LIF signal in mouse embryonic stem cells. JAKSTAT. 2015 Sep 11;4(2):e1086520.

Onishi K, Zandstra PW. LIF signaling in stem cells and development. Development. 2015 Jul 1;142(13):2230-6.

Osorno R, Tsakiridis A, Wong F, Cambray N, Economou C, Wilkie R, Blin G, Scotting PJ, Chambers I, Wilson V. The developmental dismantling of pluripotency is reversed by ectopic Oct4 expression. Development. 2012 Jul;139(13):2288-98.

Paria BC, Huet-Hudson YM, Dey SK. Blastocyst's state of activity determines the "window" of implantation in the receptive mouse uterus. Proc Natl Acad Sci U S A. 1993;90(21):10159-10162.

Pereira L, Yi F, Merrill BJ. Repression of Nanog Gene Transcription by Tcf3 Limits Embryonic Stem Cell Self-Renewal. Mol Cell Biol. 2006;26(20):7479-7491.

Plusa B, Frankenberg S, Chalmers A, Hadjantonakis AK, Moore CA, Papalopulu N, Papaioannou VE, Glover DM, Zernicka-Goetz M. Downregulation of Par3 and aPKC function directs cells towards the ICM in the preimplantation mouse embryo. J Cell Sci. 2005 Feb 1;118(Pt 3):505-15.

Plusa B, Piliszek A, Frankenberg S, Artus J, Hadjantonakis AK. Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. Development. 2008;135(18):3081-3091.

Posfai E, Tam OH, Rossant J. Mechanisms of Pluripotency In Vivo and In Vitro. In: Current Topics in Developmental Biology. Vol 107. Academic Press Inc.; 2014:1-37.

Posfai E, Lanner F, Mulas C, Leitch HG. All models are wrong, but some are useful: Establishing standards for stem cell-based embryo models. Stem Cell Reports. 2021 May 11;16(5):1117-1141.

Rivron NC, Frias-Aldeguer J, Vrij EJ, Boisset JC, Korving J, Vivié J, Truckenmüller RK, van Oudenaarden A, van Blitterswijk CA, Geijsen N. Blastocyst-like structures generated solely from stem cells. Nature. 2018 May;557(7703):106-111.

Rossant J, Vijh KM. Ability of outside cells from preimplantation mouse embryos to form inner cell mass derivatives. Dev Biol. 1980;76(2):475-482.

Rudloff S, Kemler R. Differential requirements for β-catenin during mouse development. Dev. 2012;139(20):3711-3721.

Saiz N, Plusa B. Early cell fate decisions in the mouse embryo. Reproduction. 2013;145(3):145-165.

Schlessinger J. Cell signaling by receptor tyrosine kinases. Cell. 2000 Oct 13;103(2):211-25.

Shahbazi M, Scialdone A, Skorupska N, et al. Pluripotent state transitions coordinate morphogenesis in mouse and human embryos. Nature. 2017;552(7684):239-243.

Shu J, Wu C, Wu Y, et al. Induction of pluripotency in mouse somatic cells with lineage specifiers. Cell. 2013;153(5):963.

Silva J, Nichols J, Theunissen TW, Guo G, van Oosten AL, Barrandon O, Wray J, Yamanaka S, Chambers I, Smith A. Nanog is the gateway to the pluripotent ground state. Cell. 2009 Aug 21;138(4):722-37.

Smith AG, Heath JK, Donaldson DD, Wong GG, Moreau J, Stahl M, Rogers D. Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature. 1988 Dec 15;336(6200):688-90.

Smith A. Formative pluripotency: The executive phase in a developmental continuum. Dev. 2017;144(3):365-373.

Stavridis MP, Lunn JS, Collins BJ, Storey KG. A discrete period of FGF-induced Erk1/2 signalling is required for vertebrate neural specification. Development. 2007 Aug;134(16):2889-94.

Strumpf D, Mao CA, Yamanaka Y, Ralston A, Chawengsaksophak K, Beck F, Rossant J. Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. Development. 2005 May;132(9):2093-102.

Suwińska A, Czołowska R, Ozdzeński W, Tarkowski AK. 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. Dev Biol. 2008;322(1):133-144.

Takahashi S, Kobayashi S, Hiratani I. Epigenetic differences between naive and primed pluripotent stem cells. Cell Mol Life Sci. 2018;75(7):1191-1203.

Takahashi K, Yamanaka S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell. 2006;126(4):663-676.

Tapia N, MacCarthy C, Esch D, Gabriele Marthaler A, Tiemann U, Araúzo-Bravo MJ, Jauch R, Cojocaru V, Schöler HR. Dissecting the role of distinct OCT4-SOX2 heterodimer configurations in pluripotency. Sci Rep. 2015 Aug 28;5:13533.

Tee WW, Shen SS, Oksuz O, Narendra V, Reinberg D. Erk1/2 activity promotes chromatin features and RNAPII phosphorylation at developmental promoters in mouse ESCs. Cell. 2014;156(4):678-690.

Tosolini M, Brochard V, Adenot P, et al. Contrasting epigenetic states of heterochromatin in the different types of mouse pluripotent stem cells. Sci Rep. 2018;8(1):5776.

Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. Embryonic stem cell lines derived from human blastocysts. Science. 1998 Nov 6;282(5391):1145-7.

Tsogtbaatar E, Landin C, Minter-Dykhouse K, Folmes CDL. Energy Metabolism Regulates Stem Cell Pluripotency. Front Cell Dev Biol. 2020;8.

van den Berg DL, Snoek T, Mullin NP, Yates A, Bezstarosti K, Demmers J, Chambers I, Poot RA. An Oct4-centered protein interaction network in embryonic stem cells. Cell Stem Cell. 2010 Apr 2;6(4):369-81.

Veillard AC, Marks H, Bernardo AS, et al. Stable methylation at promoters distinguishes epiblast stem cells from embryonic stem cells and the in vivo epiblasts. Stem Cells Dev. 2014;23(17):2014-2029.

Wang H, Dey SK. Roadmap to embryo implantation: clues from mouse models. Nat Rev Genet. 2006 Mar;7(3):185-99.

Williams RL, Hilton DJ, Pease S, et al. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature. 1988;336(6200):684-687.

Wray J, Hartmann C. WNTing embryonic stem cells. Trends Cell Biol. 2012;22(3):159-168.

Wray J, Kalkan T, Gomez-Lopez S, et al. Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. Nat Cell Biol. 2011;13(7):838-845.

Yagi M, Kishigami S, Tanaka A, Semi K, Mizutani E, Wakayama S, Wakayama T, Yamamoto T, Yamada Y. Derivation of ground-state female ES cells maintaining gamete-derived DNA methylation. Nature. 2017 Aug 10;548(7666):224-227.

Yamanaka Y, Lanner F, Rossant J. FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. Development. 2010 Mar;137(5):715-24.

Yang S-H, Kalkan T, Morrisroe C, Smith A, Sharrocks AD. A Genome-Wide RNAi Screen Reveals MAP Kinase Phosphatases as Key ERK Pathway Regulators during Embryonic Stem Cell Differentiation. Schübeler D, ed. PLoS Genet. 2012;8(12):e1003112.

Yang K, Wang X, Zhang H, et al. The evolving roles of canonical WNT signaling in stem cells and tumorigenesis: Implications in targeted cancer therapies. Lab Investig. 2016;96(2):116-136.

Yeom YI, Fuhrmann G, Ovitt CE, et al. Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. Development. 1996;122(3):881-894.

Yi F, Pereira L, Hoffman JA, et al. Opposing effects of Tcf3 and Tcf1 control Wnt stimulation of embryonic stem cell self-renewal. Published online 2011.

Ying QL, Nichols J, Chambers I, Smith A. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell. 2003;115(3):281-292.

Ying QL, Wray J, Nichols J, et al. The ground state of embryonic stem cell self-renewal. Nature. 2008;453(7194):519-523.

Chapter 3

In vitro capture and characterization of embryonic rosette-stage pluripotency between naive and primed states

Neagu A*, van Genderen E*, **Escudero I*** et al.

* These authors contributed equally

Nat Cell Biol. 22, 534–545 (2020)

ABSTRACT

Following implantation, the naive pluripotent epiblast of the mouse blastocyst generates a rosette, undergoes lumenogenesis and forms the primed pluripotent egg cylinder, which is able to generate the embryonic tissues. How pluripotency progression and morphogenesis are linked and whether intermediate pluripotent states exist remain controversial. We identify here a rosette pluripotent state defined by the co-expression of naive factors with the transcription factor OTX2. Downregulation of blastocyst WNT signals drives the transition into rosette pluripotency by inducing OTX2. The rosette then activates MEK signals that induce lumenogenesis and drive progression to primed pluripotency. Consequently, combined WNT and MEK inhibition supports rosette-like stem cells, a self-renewing naive-primed intermediate. Rosette-like stem cells erase constitutive heterochromatin marks and display a primed chromatin landscape, with bivalently marked primed pluripotency genes. Nonetheless, WNT induces reversion to naive pluripotency. The rosette is therefore a reversible pluripotent intermediate whereby control over both pluripotency progression and morphogenesis pivots from WNT to MEK signals.

ARTICLES
https://doi.org/10.1038/s41556-020-0508-x

nature cell biology

In vitro capture and characterization of embryonic rosette-stage pluripotency between naive and primed states

Alex Neagu[1,11], Emiel van Genderen[1,11], Irene Escudero[1,11], Lucas Verwegen[1], Dorota Kurek[1], Johannes Lehmann[1], Jente Stel[1], René A. M. Dirks[2], Guido van Mierlo[2], Alex Maas[1], Cindy Eleveld[3], Yang Ge[1], Alexander. T. den Dekker[4], Rutger W. W. Brouwer[4], Wilfred F. J. van IJcken[4], Miha Modic[5,6], Micha Drukker[7], Joop H. Jansen[8], Nicolas C. Rivron[9], Esther B. Baart[3,10], Hendrik Marks[2] and Derk ten Berge[1]✉

Following implantation, the naive pluripotent epiblast of the mouse blastocyst generates a rosette, undergoes lumenogenesis and forms the primed pluripotent egg cylinder, which is able to generate the embryonic tissues. How pluripotency progression and morphogenesis are linked and whether intermediate pluripotent states exist remain controversial. We identify here a rosette pluripotent state defined by the co-expression of naive factors with the transcription factor OTX2. Downregulation of blastocyst WNT signals drives the transition into rosette pluripotency by inducing OTX2. The rosette then activates MEK signals that induce lumenogenesis and drive progression to primed pluripotency. Consequently, combined WNT and MEK inhibition supports rosette-like stem cells, a self-renewing naive-primed intermediate. Rosette-like stem cells erase constitutive heterochromatin marks and display a primed chromatin landscape, with bivalently marked primed pluripotency genes. Nonetheless, WNT induces reversion to naive pluripotency. The rosette is therefore a reversible pluripotent intermediate whereby control over both pluripotency progression and morphogenesis pivots from WNT to MEK signals.

Pluripotent cells generate all embryonic tissues and arise in the blastocyst in an immature naive state. These cells become responsive to lineage-inductive signals only after maturation into the transcriptionally and epigenetically different primed pluripotent state of the post-implantation epiblast[1]. Following implantation, the naive cells polarize and arrange into a rosette[2]. Negatively charged sialomucins, including podocalyxin (PODXL), are exposed and repulse the apical cell membranes, which form a central lumen[2,3]. Pluripotency therefore progresses through several developmental intermediates before acquiring its true potential[1,4,5]. How these transitions are regulated is unclear, and whether intermediate pluripotent states occur during this transition remains controversial. WNT and MEK signals mediate the naive-to-primed pluripotency transition in vitro[6,7], and the combination of a WNT agonist, the GSK3 inhibitor CHIR99021 (CHIR), and the MEK inhibitor PD325901 (PD; also known as 2i)[7], together with the cytokine LIF maintains naive pluripotency in embryonic stem cells (ESCs). Here, we show that WNT and MEK signals control successive pluripotency and morphogenetic transitions in the embryo that are centred on the rosette. Simultaneous WNT and MEK inhibition stabilizes rosette pluripotency in vitro. Our study demonstrates a genuine intermediate pluripotent state with its markers and the signals that control its entrance and exit (Fig. 1a).

Results

We analysed peri-implantation embryos to pinpoint the naive-primed transition in vivo. By checking copulation plugs every 3 h, we obtained blastocysts attached to the uterine wall around embryonic day 4.9 (E4.9) to E5.0. Rosettes were buried in implantation crypts around E4.9–5.2, and embryos displayed a lumen around E5.1–5.25. Several attached embryos also displayed a rosette. We identified rosettes by their central aggregation of actin fibres (Fig. 1c,f, arrows (not always visible due to the sectioning plane)), while lumens were surrounded by PODXL (Fig. 1d,g, arrows). The naive markers KLF4, NANOG and ESRRB were expressed in the blastocyst and in most rosettes, but were absent from lumen-stage embryos (Fig. 1b–e,g–j; Extended Data Fig. 1a–d, ESRRB levels slightly lower in rosettes). Importantly, rosettes expressed the primed marker OTX2, which indicates a more advanced pluripotent state (Fig. 1f; Extended Data Fig. 1b). Half of the rosettes displayed some cells lacking NANOG, some of which expressed OCT6, another primed marker (Fig. 1j, $n=12$). Lumen-stage epiblasts coexpressed OCT6 with OTX2 in all cells (Fig. 1g). Analysis

[1]Department of Cell Biology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands. [2]Department of Molecular Biology, Faculty of Science, Radboud University, Radboud Institute for Molecular Life Sciences (RIMLS), Nijmegen, The Netherlands. [3]Division of Reproductive Medicine, Department of Obstetrics and Gynaecology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands. [4]Center for Biomics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands. [5]The Francis Crick Institute, London, UK. [6]Department for Neuromuscular Diseases, UCL Queen Square Institute of Neurology, London, UK. [7]Institute of Stem Cell Research, Helmholtz Zentrum München, Neuherberg, Germany. [8]Department of Laboratory Medicine, Laboratory of Hematology, Radboud University, Nijmegen Medical Centre and Radboud Institute for Molecular Life Sciences (RIMLS), Nijmegen, The Netherlands. [9]Institute of Molecular Biotechnology of the Austrian Academy of Sciences, Vienna Biocenter (VBC), Vienna, Austria. [10]Department of Developmental Biology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands. [11]These authors contributed equally: Alex Neagu, Emiel van Genderen, Irene Escudero. ✉e-mail: d.tenberge@erasmusmc.nl

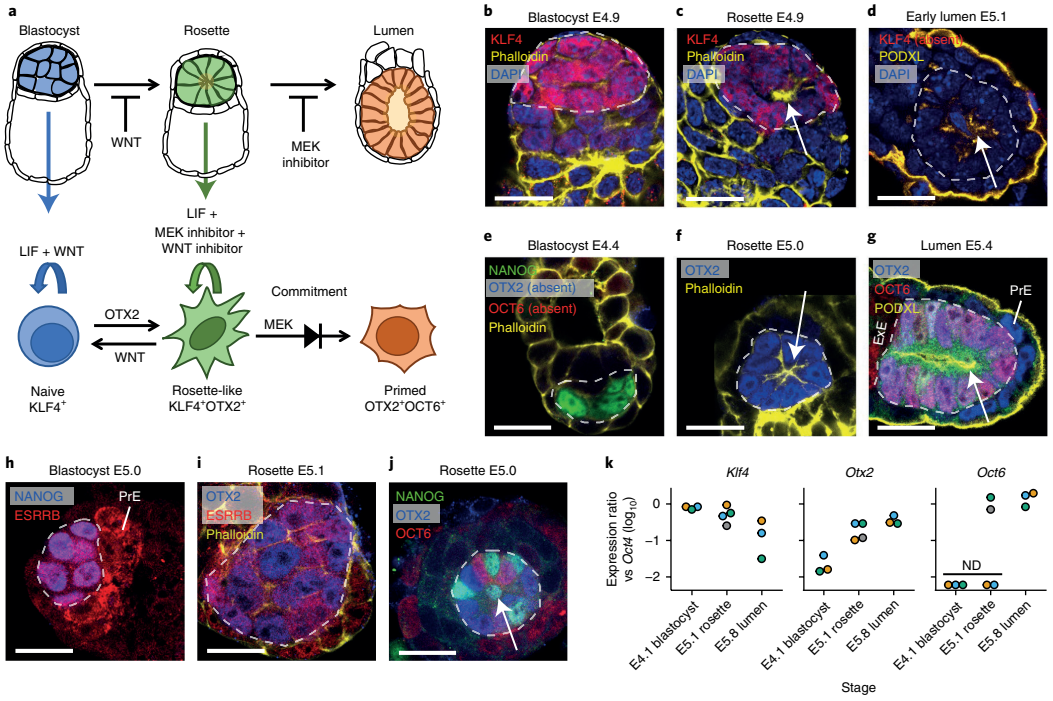


Fig. 1 | Naive and primed markers in the embryonic rosette. a, Summary of the main findings indicating the logic of signalling and embryonic and in vitro relationships. Downregulation of WNT allows the induction of OTX2, the transition from naive to rosette pluripotency and the acquisition of MEK responsivity. MEK signals drive commitment, lumenogenesis and further progression to primed pluripotency. Combined WNT and MEK inhibition captures rosette-stage pluripotency in vitro. WNT reverts rosette cells to naive pluripotency. **b–j**, Embryos of the indicated stages (blastocysts (**b**,**e**,**h**), rosettes (**c**,**f**,**i**,**j**) and early lumen (**d**,**g**)) stained for the indicated markers. The epiblast is encircled by a dashed white line. Arrows indicate the rosette centre (**c**, **f** and **j** visualized by phalloidin staining of the apical actin accumulation (in **c** and **f**)) or the lumen surrounded by PODXL (**d** and **g**). KLF4, NANOG, ESRRB, OTX2 and OCT6 were detected in 17 out of 17, 12 out of 12, 5 out of 5, 3 out of 17 and 0 out of 6 blastocysts, respectively, and in 14 out of 14, 18 out of 20, 7 out of 8, 38 out of 38 and 6 out of 12 rosettes, respectively. KLF4, NANOG, OTX2 and OCT6 were detected in 0 out of 15, 0 out of 4, 9 out of 9 and 22 out of 22 lumen-stage embryos, respectively. OCT6 was also detected in extraembryonic ectoderm (ExE), ESRRB in primitive endoderm (PrE) and OTX2 in lumen-stage PrE. **k**, Expression of *Klf4*, *Otx2* and *Oct6* relative to *Oct4* (as an epiblast-specific marker) in embryos of the indicated stages, determined by RT–PCR. Each colour-coded dot represents a group of four embryos. ND, not detected. Adjustments of individual colour channels were performed equally for all related images (**b–j**). Scale bars, 20 µm (**b–j**).

using PCR with reverse transcription (RT–PCR) confirmed that *Klf4* was downregulated during the progression from blastocyst to lumen. Meanwhile *Otx2* was induced in the rosette and *Oct6* in the lumen stage and in some rosettes (Fig. 1k). The epiblast therefore progresses to a rosette-specific state expressing the naive markers KLF4, NANOG and ESRRB with OTX2 before converting to the OTX2$^+$OCT6$^+$ primed state.

WNT and MEK signals couple successive steps in morphogenesis to pluripotency progression. We modelled rosette formation and lumenogenesis using ESCs seeded in basement membrane extract (BME)[2] to study the roles of WNT and MEK signals. After 48 h, 75% of ESC aggregates formed rosettes displaying apical–basal polarity of the actin cytoskeleton (Fig. 2a,b). About one-quarter of the aggregates contained KLF4$^+$ cells, more than 90% contained OTX2$^+$ cells and 60% contained OCT6$^+$ cells (Fig. 2c–e; Extended Data Fig. 2a,b). We monitored lumenogenesis by live staining for exposed PODXL (Extended Data Fig. 2c,d). About 70% of aggregates displayed a lumen after 65 h, similar to the percentage that formed rosettes (Fig. 2f,g).

The presence of 2i maintained KLF4 while blocking rosette formation, OTX2 and OCT6 induction and lumenogenesis. The GSK3 inhibitor CHIR (which acts as a WNT agonist) and WNT3A protein induced similar effects. Conversely, inhibiting endogenous WNT signals via the WNT inhibitor IWP2 (inhibitor of WNT production 2) promoted rosette formation, OCT6 induction and lumenogenesis (Fig. 2a–g; Extended Data Fig. 2a,b). WNT3A protein was less potent than CHIR, which is probably due to its low stability[8]. WNT/β-catenin acts by relieving transcriptional repression by TCF7L1 (refs. [9,10]). Accordingly, *Tcf7l1*$^{-/-}$ ESC aggregates, which lack TCF7L1 repression, did not form rosettes, even after WNT inhibition by IWP2 (Fig. 2h). WNT inhibition induced OTX2, which was dependent on TCF7L1 (Fig. 2i). OTX2 is a critical driver of rosette formation[3]. Indeed, ESCs in which *Otx2* was knocked out (*Otx2*KO) formed fewer rosettes, while *Otx2* overexpression (*Otx2*OE) induced rosettes even in the presence of the WNT agonist

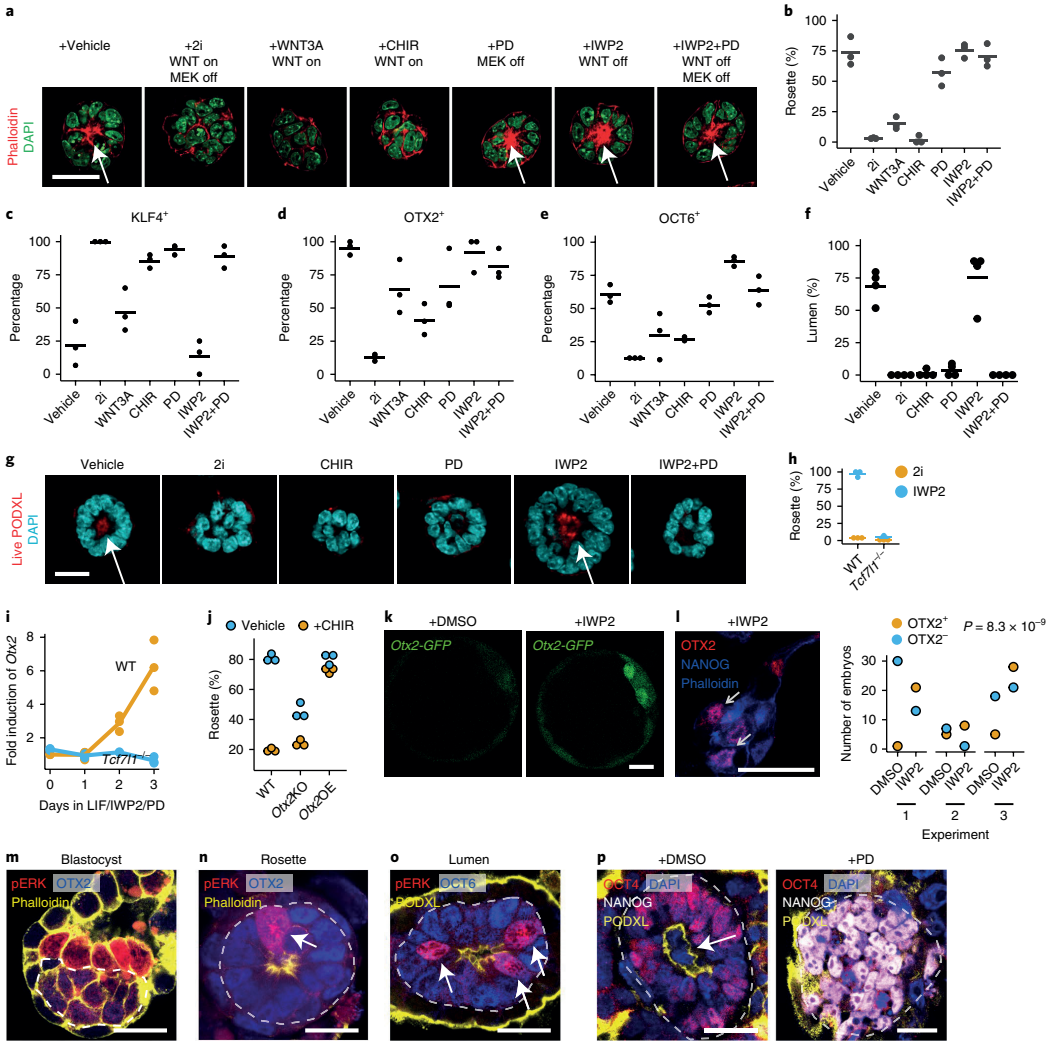


Fig. 2 | WNT and MEK signals couple successive steps in morphogenesis to pluripotency progression. a, Rosettes and aggregates 48 h after seeding ESCs in BME in the indicated conditions, stained for markers. The status of the WNT and MEK pathways as controlled by the treatments is indicated. Arrows indicate the rosette centres. **b**, The percentage of ESCs generating rosettes 48 h after seeding. $n = 3$ biological replicates. **c**–**e**, The percentage of ESC aggregates containing KLF4$^+$ (**c**), OTX2$^+$ (**d**) or OCT6$^+$ (**e**) cells 48 h after seeding. $n = 3$ biological replicates. **f**, The percentage of ESCs generating lumens 65 h after seeding. $n = 4$ biological replicates. **g**, ESC aggregates 65 h after seeding and live stained for PODXL, followed by DAPI. Arrows indicate lumens surrounded by PODXL. **h**, The percentage of wild-type (WT) or *Tcf7l1*$^{-/-}$ ESCs generating rosettes 48 h after seeding in the indicated conditions. $n = 3$ independent experiments. **i**, Induction of *Otx2* in *Tcf7l1*$^{-/-}$ ESCs compared to WT after transfer to LIF/IWP2/PD (RT–PCR, three technical replicates). **j**, Percentage of *Otx2*KO or *Otx2*OE ESCs generating rosettes 48 h after seeding. $n = 3$ independent experiments. **k**, Four-cell-stage *Otx2*–GFP embryos cultured for 75 h and imaged for GFP. A GFP$^+$ DMSO-treated (3 from 11) and representative IWP2-treated (10 from 11) embryos are shown. **l**, Four-cell-stage embryos cultured to the expanded blastocyst stage. Arrows indicate OTX2 in NANOG$^+$ epiblast of an IWP2-treated embryo. Three independent experiments with $n = 158$ embryos (66 DMSO, 92 IWP2) are summarized in the chart on the right. Two-sided Fisher's exact test. **m**–**o**, Embryos of the indicated stages stained for markers. Arrows indicate pERK$^+$ epiblast cells. Images representative of three blastocysts (**m**), none positive; nine rosettes (**n**), three displaying a pERK$^+$ cell; four lumen stages (**o**), with six, four, three and four pERK$^+$ cells. pERK was also observed in primitive endoderm (**m**). **p**, Blastocysts cultured for 77 h with DMSO or PD and stained for markers; 8 out of 16 and 1 out of 46 OCT4$^+$ embryos displayed a lumen (arrow). Dotted white line indicates the epiblast (**m**–**p**). Images represent three (**a**) or four (**g**) independent experiments. Scale bars, 20 µm (**a**,**g**,**k**–**p**). For **b**–**f** and **h**, bars represent the mean.

NATURE CELL BIOLOGY ARTICLES

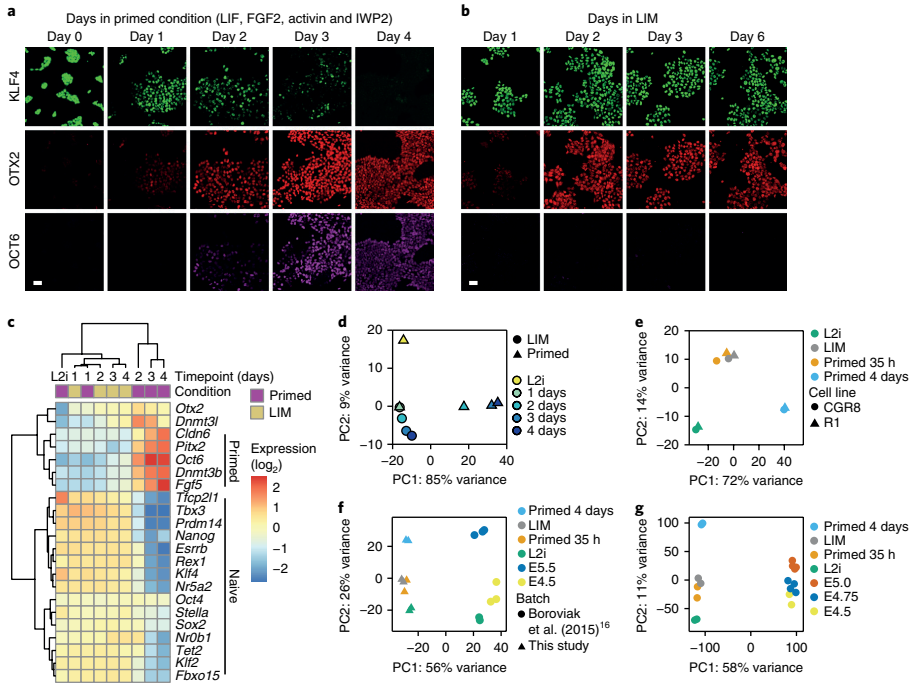


Fig. 3 | WNT controls the initial phase of the naive-primed transition. a,**b**, ESCs plated in primed (**a**) or LIM (**b**) conditions for the indicated number of days and immunostained for the indicated markers (3 independent experiments). **c**, RNA-seq heatmap from R1 ESCs plated in primed or LIM conditions for the indicated number of days. **d**, PCA of RNA-seq data from R1 ESCs plated in primed or LIM conditions for the indicated number of days. At day 1 the samples for both conditions overlap. One sample for each condition. **e**–**g**, PCA of RNA-seq data from R1 and CGR8 ESC lines maintained in the indicated conditions compared with L2i ESCs (**e**), E4.5 and E5.5 epiblast (**f**, from ref. [16]), or with E4.5, E4.75 and E5.5 epiblast (**g**, from ref. [3]). Each dot indicates a biological replicate. Scale bars, 40 µm (**a**,**b**).

CHIR (Fig. 2j). OTX2 is therefore a key effector repressed by the WNT pathway. The MEK inhibitor PD permitted rosette formation and OTX2 induction, especially in combination with IWP2, but most rosettes maintained KLF4 expression and few induced OCT6 and progressed to lumenogenesis, arresting in the KLF4$^+$OTX2$^+$ state (Fig. 2a–g; Extended Data Fig. 2a,b). Downregulation of the WNT/β-catenin–TCF7L1 pathway therefore induces OTX2 and rosette formation, whereas MEK signals drive the subsequent lumenogenesis and progression to primed pluripotency.

WNT signals are active in the blastocyst[6,11–13]. To investigate whether these prevented the inner cell mass (ICM) from inducing OTX2, we cultured four-cell-stage *Otx2*–GFP (green fluorescent protein) reporter embryos in the presence of IWP2. After 56 h, blastocysts formed in which the ICM displayed faint GFP signals that rapidly increased in the subsequent 6 h. In contrast, in the absence of IWP2, the ICM induced faint GFP in few embryos after 75 h (Fig. 2k). Immunostaining of four-cell-stage embryos cultured until the expanded blastocyst stage confirmed this result, whereas phalloidin staining did not reveal evidence of rosette morphogenesis (Fig. 2l). Thus, blastocyst WNT signals prevent premature progression of the ICM to OTX2$^+$ rosette pluripotency, while additional factors, for example, basal lamina deposition[2], are required for rosette morphogenesis.

MEK signals induce the phosphorylation of ERK, typically as a series of stochastic pulses at a frequency reflecting the strength of the signal[14]. The epiblast of late blastocysts and most rosettes did not display phosphorylated ERK (pERK) (Fig. 2m). However, some rosettes contained a single pERK$^+$ cell (Fig. 2n). These rosettes displayed OTX2 but not OCT6 expression (Extended Data Fig. 2e). In lumen-stage epiblasts, all cells were OCT6$^+$ and many displayed ERK phosphorylation (Fig. 2o; Extended Data Fig. 2f). Thus, pERK pulses increase in frequency during progression to the lumen stage. To investigate their function, we cultured blastocysts to the lumen stage. The embryos lost KLF4 expression 41 h after attachment to the substrate and displayed lumenogenesis after 77 h. However, MEK inhibition maintained the expression of KLF4 and NANOG in the epiblast, which generated a ball-shaped or disorganized mass lacking a lumen (Fig. 2p). These data suggest that pluripotency progression is coupled to embryonic morphogenesis first by WNT signals and, subsequently, in the rosette stage, by MEK signals (Fig. 1a).

WNT controls the initial phase of the naive-primed transition. We observed a transient appearance of rosette-like KLF4$^+$OTX2$^+$ OCT6$^-$ cells 0.5–1 day after ESCs that were grown in N2B27 medium supplemented with LIF and 2i (hereafter called L2i) were differentiated in N2B27 medium without supplements, or differentiated towards epiblast-like cells (EpiLCs), which are primed cells resembling egg cylinder epiblast[15] (Extended Data Fig. 3a,b). To focus on the roles of WNT and MEK, we maintained the

presence of LIF while inhibiting endogenous WNT signals using IWP2. We also activated MEK using FGF2 and supported the primed cells with activin. LIF slows but does not inhibit the naive-primed transition in this condition (hereafter called 'primed condition')[6]. OTX2 was induced after 1 day and coexpressed with KLF4 until conversion to the KLF4$^-$OTX2$^+$OCT6$^+$ primed state after 3 days (Fig. 3a). Importantly, ESCs transferred from L2i into conditions in which both WNT and MEK were inhibited (hereafter called LIM for LIF, IWP2 and MEK inhibitor) also induced OTX2, but arrested in the KLF4$^+$OTX2$^+$OCT6$^-$ rosette-like state (Fig. 3b). Since L2i and LIM differ only in WNT pathway activity (active versus repressed, respectively), this result indicates that WNT signals are required to prevent OTX2 induction and pluripotency progression.

RNA-sequencing (RNA-seq) showed that repression of naive and induction of primed genes mostly occurred after 2 days in primed conditions and appeared complete after 3–4 days, while after 4 days in LIM, pluripotency gene expression resembled that after 1–2 days in primed conditions (Fig. 3c; Supplementary Table 1). Principal component analysis (PCA) separated primed from other conditions alongside principal component 1 (PC1), and L2i from LIM alongside PC2 (Fig. 3d). During the first day, however, the transcriptomes for both LIM and primed conditions changed only alongside PC2, despite the opposite status of the MEK pathway. Extending this time frame, the transcriptomes of two ESC lines after 8 (R1) or 15 (CGR8) days in LIM grouped together with those after 35 h in primed conditions (Fig. 3e; Supplementary Table 2). Loss of WNT signals therefore drives part of the naive-primed transition, independent from MEK activity. Moreover, combined WNT and MEK inhibition arrests pluripotent cells in an intermediate naive-primed, rosette-like state.

A previous transcriptome analysis showed that naive pluripotency appears sometime within E3.75–4.5 and is absent from E5.5 epiblasts[16]. PCA of these data with ours captured the developmental progression in PC2 (Fig. 3f). ESCs from both studies mapped before E4.5 ICM alongside this developmental axis, while the 4-day-primed samples mapped alongside E5.5 epiblasts. Importantly, cells in LIM mapped together with 35-h primed cells, confirming their partway transition towards primed pluripotency. A comparative transcriptome analysis with peri-implantation embryos[3] again mapped the L2i ESCs before E4.5 ICM, while cells in LIM showed most similarity with E5.0 epiblasts (Fig. 3g). These data support our findings that loss of WNT signalling induces the naive-primed transition, while MEK inhibition arrests the cells only at a more advanced rosette-like state.

MEK and WNT inhibition maintains rosette characteristics. We derived stable cell lines from blastocysts in LIM. Outgrowths generated dispersed colonies that expanded long term (Extended Data Fig. 3c). After blastocyst injection, all three tested cell lines contributed to live chimeras, induced strong sex conversion and showed germline transmission (Extended Data Fig. 3d; Supplementary Table 3). Although characteristic for ESCs, contribution to blastocyst chimeras is found for a spectrum of pluripotent states, including E5.0 peri-implantation epiblast cells[17] and EpiSCs maintained in differentiation-suppressing conditions[18,19].

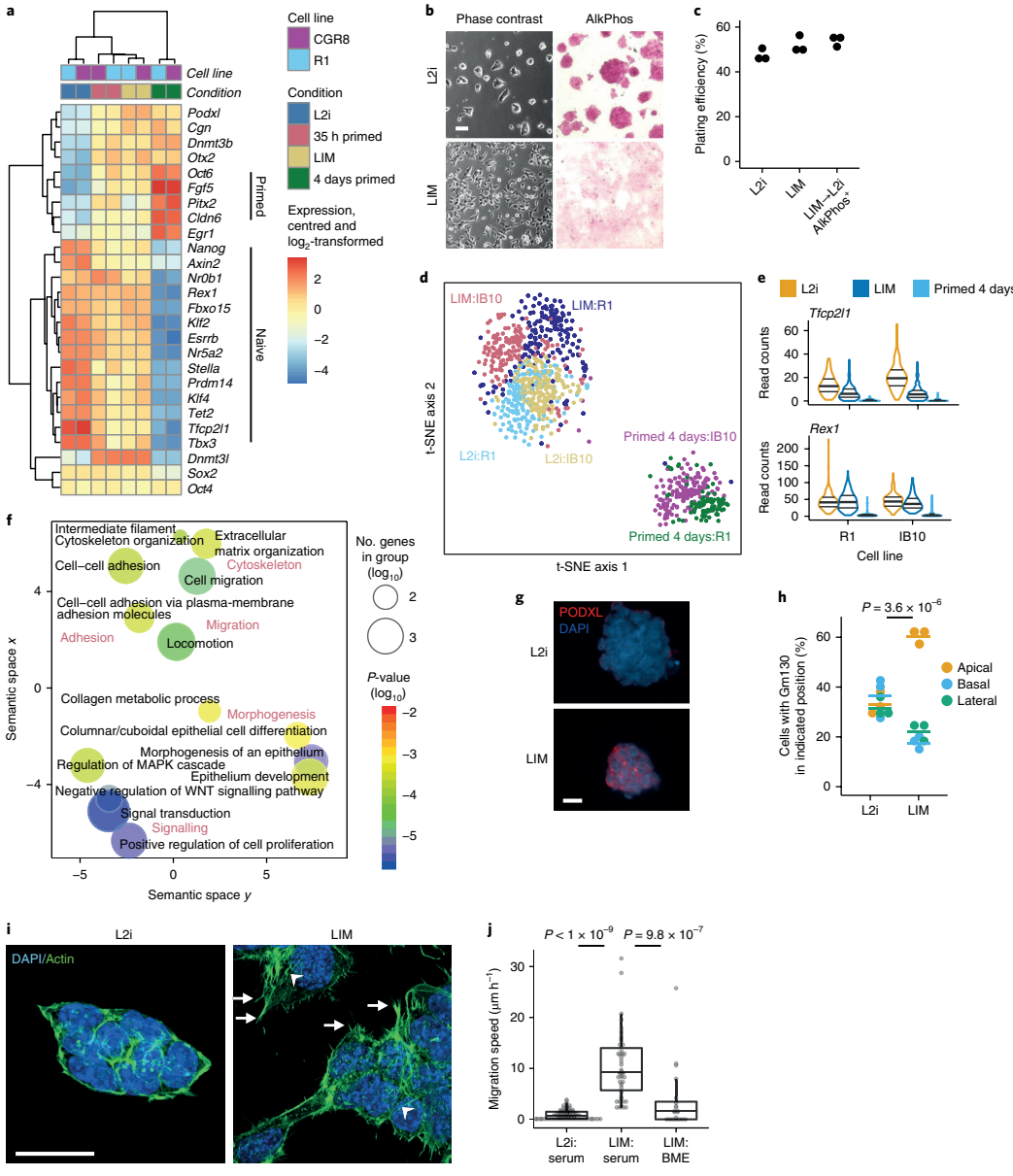
LIM supported the expansion of ESCs and blastocyst-derived cell lines for more than 28 passages on serum or BME coatings. Repression of the WNT target *Axin2* and the MEK target *Egr1* verified pathway inhibition (Extended Data Fig. 3e). The lines maintained expression of OCT4, SOX2, the naive markers KLF4, NANOG and ESRRB (at a slightly reduced level), and OTX2 while repressing OCT6 (Extended Data Fig. 4a). The cells displayed dispersed monolayer morphology, low alkaline phosphatase activity, tolerated clonal passaging, retained naive gene expression, LIF dependency and active X chromosomes, and proliferated slightly slower than ESCs (Fig. 4a–c; Extended Data Figs. 3e and 4b–d). Single-cell sequencing and t-distributed stochastic neighbour embedding (t-SNE) analysis revealed a KLF4$^+$OTX2$^+$OCT6$^-$ LIM population distinct from naive and primed cells (Fig. 4d; Extended Data Fig. 4e). While expression of the pluripotency marker TFCP2L1 was reduced, most pluripotency markers, such as REX1, showed similar expression level distributions in cells grown under LIM and L2i conditions (Fig. 4e). Of note, LIM cells repressed OTX2, regenerated alkaline-phosphatase-positive colonies and acquired ESC-like proliferation rates following transfer to L2i (Fig. 4c; Extended Data Fig. 4d,f), which indicates that WNT induced reversion to the ESC state.

Based on bulk transcriptome data, 531 and 196 genes were more than twofold differentially expressed between LIM and L2i cells (Supplementary Table 2). Gene Ontology (GO) analysis revealed processes involving cell adhesion, cell migration, epithelial cell differentiation and morphogenesis, and MEK signalling, which is consistent with the phenotypic changes observed in vitro and in the rosette (Fig. 4f; Supplementary Table 4). Motility and adhesion-related genes are induced early in the naive-primed transition[20,21], which possibly reflects morphogenetic processes such as polarization and lumenogenesis[2]. Indeed, two factors instrumental for rosette lumenogenesis were induced: PODXL, which induces cell membrane separation, and cingulin (CGN), which mediates its cell surface exposure[3] (Fig. 4a,g). Furthermore, LIM cells gained apical–basal polarity and displayed filopodia, stress fibres and a dispersed appearance, which all indicate motility (Fig. 4b,h,i; Extended Data Figs. 3c and 4g). Live imaging showed the cells migrating at approximately 10 μm h^{-1} on serum coatings (Fig. 4j; Extended Data Fig. 4h). E5.0–6.0 epiblast cells disperse after detaching from the basal lamina during mitosis[22], which suggests that the absence of lamina contacts induces motility. Accordingly, surfaces coated with the basal lamina extract BME reduced motility (Fig. 4j). Together,

Fig. 4 | MEK and WNT inhibition maintains rosette characteristics. a, RNA-seq heatmap from R1 and CGR8 ESCs cultured in the indicated conditions. LIM 8 days (R1) or 15 days (CGR8). **b**, Phase-contrast and alkaline-phosphatase (AlkPhos)-stained images of R1 ESCs maintained in L2i or LIM for three passages. Three independent experiments were performed, with similar results obtained. **c**, The plating efficiency of ESCs and LIM cells, and the formation of alkaline-phosphatase-positive colonies by LIM cells after plating in L2i. 3 independent experiments. **d**, t-SNE plot of single-cell RNA-seq data of the indicated cell lines and conditions. **e**, Distribution of the expression levels of the indicated genes in the single-cell RNA-seq data. Violin plots show the median (centre bold bar), the first and third quartiles on top of the expression level distribution. $n = 159$, 178, 102 (R1) and 184, 149 and 164 (IB10) cells for L2i, LIM and primed, respectively (**d** and **e**). **f**, GO analysis of genes differentially expressed between L2i and LIM cells (two samples each). The *P* values were calculated from minimum hypergeometric scores from a ranked gene list using GOrilla[58]. GO terms were aggregated and plotted according to their semantic relationships using REVIGO[59]. **g**, Immunostaining for PODXL of R1 ESCs maintained in L2i or LIM for six passages (three biological replicates). **h**, Apical–basal polarity in L2i and LIM cells was determined by Gm130 immunostaining for polarized location of the Golgi complex. Two-way analysis of variance for effect of the condition. $n = 3$ independent experiments; bar represents the mean. **i**, L2i and LIM cells stained with phalloidin to visualize the actin cytoskeleton. Filopodia (arrows) and stress fibres (arrowheads) are indicated (three independent experiments). **j**, Migration speed of L2i and LIM cells on serum-coated or BME-coated slides. The centre line indicates the median, the boxes the first and third quartiles, while the whiskers extend to the most extreme value no further than 1.5-times the inter-quartile range. $n = 80$, 47, 20 cells for L2i, serum and BME, respectively. Two-tailed Mann–Whitney test. Scale bars, 20 μm (**g**,**i**) or 40 μm (**b**).

our data argue that WNT inhibition induces pluripotency progression to the rosette-like state, while MEK inhibition arrests further progression and maintains this state. Therefore, we propose that WNT/MEK-inhibited cells are called rosette-like stem cells (RSCs).

OTX2 enables MEK-driven transition to primed pluripotency. We monitored the naive-primed transition using the *Rex1*::destabilized-GFP (RGd2) reporter, which is thought to track exit from naive pluripotency[21]. While RGd2 ESCs and RSCs displayed similar levels of the reporter, consistent with our single-cell RNA-seq data, RSCs downregulated the reporter within 24 h in the primed condition, more than a day earlier than ESCs (Fig. 5a; Extended Data Fig. 5a). Likewise, RSCs carrying *Tbx3*–RFP naive and *Oct6*–GFP reporters[23] switched more quickly to primed



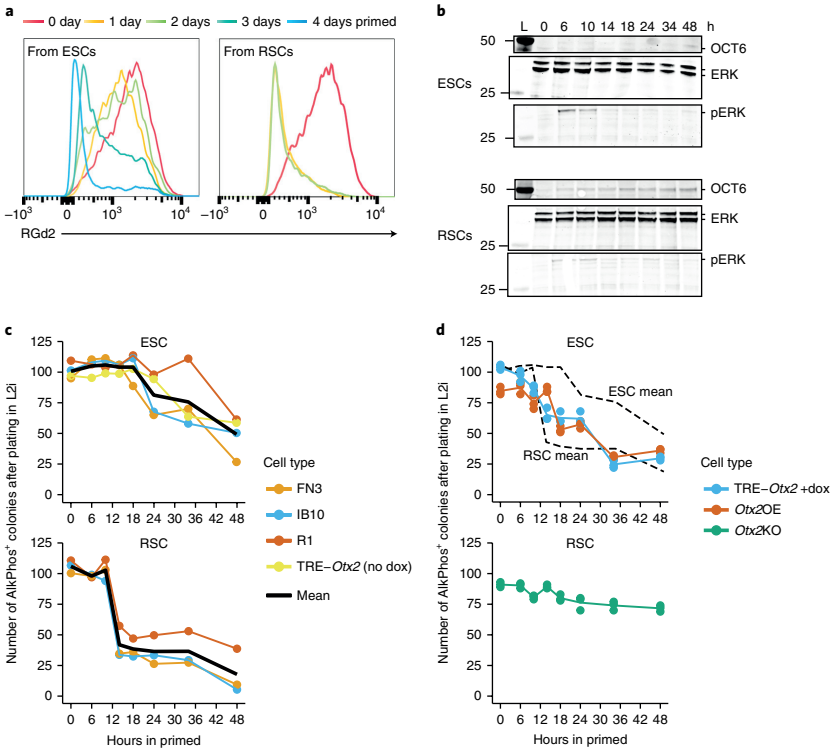


Fig. 5 | OTX2 enables MEK-driven transition to primed pluripotency. a, RGd2 ESCs and RSCs were plated in primed conditions and analysed daily for GFP by flow cytometry (three independent experiments). **b**, Western blots showing OCT6, ERK (same gel) and pERK (on different gel) in FN3 ESCs and RSCs after transfer to primed medium for the indicated durations (three biological replicates; full blots are available as source data). L, ladder. **c**,**d**, Number of alkaline-phosphatase-positive colonies generated in L2i by 200 ESCs (upper) or RSCs (lower) after incubation in primed medium for the indicated times. Cell lines used: FN3, IB10, R1 and TRE–*Otx2* (no doxycycline (dox) (**c**), and TRE–*Otx2* (+ dox), *Otx2*OE and *Otx2*KO (**d**). $n=4$ (ESC) or $n=3$ (RSC) biological (**c**) or $n=4$ technical (**d**) replicates.

reporter expression than ESCs (Extended Data Fig. 5b). Moreover, RSCs induced OCT6 protein expression after 6 h in the primed condition, followed by a sudden drop in their ability to establish alkaline-phosphatase-positive colonies in L2i, which indicates their commitment to primed pluripotency. In contrast, ESCs induced OCT6 expression after 34 h, and their reversibility declined only slowly after 24 h (Fig. 5b,c; Extended Data Fig. 5c). Cell lines of diverse backgrounds, including non-permissive FVB/N (FN3), showed this difference. Therefore, RSCs are poised for conversion to primed pluripotency after release of MEK inhibition, whereas ESCs require a prior transition to the rosette-like state. Nonetheless, both ESCs and RSCs responded with similar pERK oscillations following transfer to primed conditions (Fig. 5b; Extended Data Fig. 5c). The WNT pathway—the only difference between L2i and LIM—must therefore modify the pERK response. OTX2 is induced following the loss of WNT activity (Fig. 2i) and acts early in the naive-primed transition[23–25]. Importantly, RSCs lacking OTX2 failed to commit to primed pluripotency and maintained their reversal efficiency even after 48 h in primed condition (Fig. 5d). Moreover, ESCs that over-expressed OTX2, either constitutively or by TRE-based doxycycline induction, showed a similarly rapid drop in reversibility as that of RSCs in the primed condition (Fig. 5d; Extended Data Fig. 5d). OTX2 therefore enables the progression to primed pluripotency in response to MEK signals.

A primed chromatin landscape in RSCs. L2i ESCs display DNA hypomethylation, which is associated with a low expression of DNA methyltransferases[26–30]. After transfer to primed conditions, 5-methylcytosine (5mC) levels rapidly increased from around 1% to around 4% after 3 days, while 5-hydroxymethylcytosine (5hmC) was low in naive and primed cells but elevated during the transition, which suggests that there is 5mC turnover (Fig. 6a). RSCs expressed *Dnmt3b* and *Dnmt3l*, and the protein DNMT3L was strongly induced in RSCs and embryonic rosettes, which suggests that there is a departure from the hypomethylated ground state (Figs. 4a and 6b; Extended Data Fig. 6a,b). Indeed, 5mC and 5hmC levels were similarly elevated in RSCs to that of cells halfway through the naive-primed transition (Fig. 6a).

L2i ESCs are further characterized by the presence of bivalent gene promoters[27,31], carrying both activating H3K4me3 and repressive H3K27me3 chromatin marks[32–34]. Comparative chromatin immunoprecipitation (ChIP) with sequencing (ChIP-seq) analysis of two RSC lines with L2i ESCs[27] and EpiLCs[35] showed that RSCs clustered with EpiLCs regarding H3K4me3, thus suggesting

NATURE CELL BIOLOGY ARTICLES

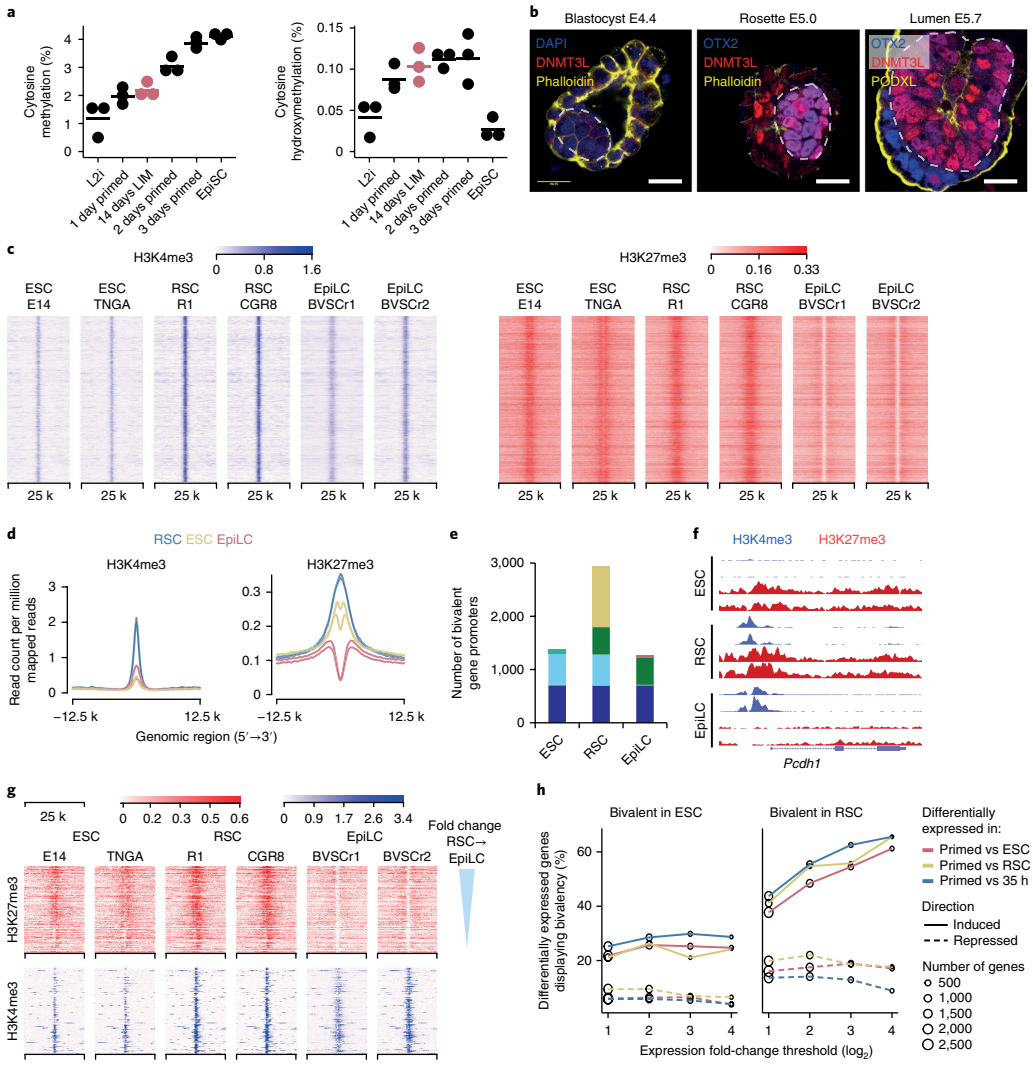


Fig. 6 | A primed chromatin landscape in RSCs. a, Levels of cytosine methylation and cytosine hydroxymethylation in ESCs (black) or RSCs (red) cultured in the indicated conditions ($n=3$ biological replicates, bars represent the mean). **b**, Blastocyst-stage ($n=4$), rosette-stage ($n=7$) and lumen-stage ($n=5$) embryos stained for DNMT3L, OTX2 and phalloidin or PODXL. The epiblast is encircled by a dotted white line. DNMT3L is also visible in the extraembryonic ectoderm (rosette-stage). **c**, H3K4me3 and H3K27me3 heatmaps of the bivalent cluster. **d**, Average intensity plots of H3K4me3 and H3K27me3 in bivalent regions. **e**, The number of promoters carrying bivalent marks in naive, rosette-like or primed cells. Groups of promoters are coloured depending on the cell type(s) in which they display bivalency. **f**, Screenshot from the UCSC genome browser showing H3K4me3 (blue) and H3K27me3 (red) peaks on the *Pcdh1* gene. **g**, H3K27me3 and H3K4me3 heatmaps centred over promoters of genes induced during the RSC-to-primed transition, ranked according to the fold-change in induction. **h**, The fraction of gene promoters carrying bivalent marks in ESCs (in L2i) or RSCs, plotted against expression level changes. Scale bar, 20 μm (**b**).

a shift towards the primed state, but revealed a distinct pattern of H3K27me3 (Extended Data Fig. 6c–e). *K*-means clustering of the ChIP-seq peaks revealed bivalent (Fig. 6c,d), H3K4me3-enriched and H3K27me3-enriched clusters (Extended Data Fig. 6f,g). RSCs gained H3K4me3 and H3K27me3 marks in the bivalent cluster compared to ESCs, while EpiLCs lost H3K27me3. Consequently, RSCs contained more than twice as many bivalent genes as ESCs or EpiLCs (Fig. 6c–e; Supplementary Table 5). Many primed genes, such as *Pcdh1*, *Oct6* and *Fgf5*, gained bivalency (Fig. 6f; Extended Data Fig. 6h). Moreover, genes that showed a stronger expression

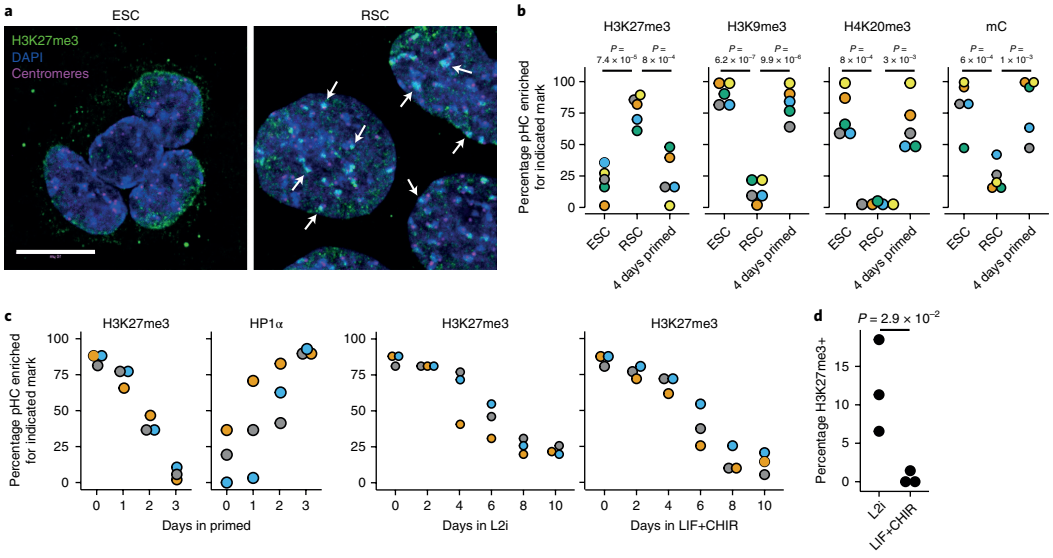


Fig. 7 | RSCs remodel pHC. a, Immunostaining for H3K27me3 (green) and the centromeres (magenta) in R1 ESCs and RSCs. Arrows indicate examples of pericentric H3K27me3 foci (light blue, overlap of H3K27me3 and DAPI). Images are representative of three experiments. **b**, The percentage of metaphases in indicated conditions in which the pericentric regions are enriched for the indicated mark. Each dot represents the average of 100 metaphases. $n = 5$ biological replicates. **c**, RSCs were cultured in the indicated conditions and the percentage of cells displaying pHC foci positive for the indicated marks was determined. $n = 3$ biological replicates. **d**, ESCs were cultured for 4 weeks in L2i or LIF+CHIR, and the fraction of cells displaying H3K27me3$^+$ pHC foci was determined. $n = 3$ biological replicates. (**b** and **d**) Unpaired two-tailed *t*-test. Scale bars, 10 µm (**a**).

difference between primed cells and RSCs were more likely to display bivalency in RSCs, a correlation that was not observed with bivalent genes in ESCs or with downregulated or naive genes (Fig. 6g,h). Thus, the increase in bivalent genes in RSCs foreshadowed their induction following the transition to primed pluripotency, information that is absent from ESCs. The naive-to-rosette transition therefore sets up a chromatin landscape that is permissive for primed pluripotency, which is consistent with the rapid response of RSCs to MEK activation.

Rosette-specific formatting of pericentric heterochromatin. H3K27me3 is a feature of facultative heterochromatin, which evolved to regulate gene expression. However, immunostaining of RSCs revealed its accumulation on pericentric heterochromatin (pHC), which is 4,6-diamidino-2-phenylindole (DAPI)-dense heterochromatin surrounding the centromeres that is normally characterized by constitutive heterochromatin marks, including H3K9me3, H4K20me3, DNA methylation and HP1α[34] (Fig. 7a). Metaphase chromosome spreads showed enrichment of pericentric H3K27me3 in 60–90% of RSCs and only in a minority of ESCs or primed cells, while pericentric H3K9me3, H4K20me3 and 5-methylcytosine were depleted in most RSCs (Fig. 7b; Extended Data Fig. 7a–d), which indicates that RSCs erased constitutive heterochromatin marks and replaced them with facultative marks.

After transfer to primed conditions, RSCs lost KLF4 and erased most pericentric H3K27me3 within 2 days while simultaneously gaining pericentric HP1α, which indicates that there is a reversion to canonical constitutive heterochromatin following the transition to primed pluripotency (Fig. 7c; Extended Data Fig. 7e). Reverting the cells to ESCs by L2i also induced erasure of pericentric H3K27me3, but at a slower rate, and a subset of cells retained the mark, even after extensive culture (Fig. 7c,d; Extended Data Fig. 7f). Indeed, pericentric H3K27me3 was recently observed in a subset of L2i ESCs[36]. However, transfer to LIF + CHIR erased virtually all pericentric H3K27me3 (Fig. 7c,d), which suggests that MEK inhibition induces epigenetic heterogeneity in ESCs.

Importantly, while peri-implantation blastocysts displayed little H3K27me3 immunostaining, OTX2$^+$ embryos displayed nuclear accumulation of H3K27me3 in the epiblast, including staining of the pericentric regions. Post-implantation KLF4$^+$OTX2$^+$ embryos displayed further accumulation of H3K27me3 and intense pericentric staining in many epiblast cells, which was lost after lumenogenesis, and loss of KLF4 (Fig. 8a–c; Extended Data Fig. 8a,b). Conversely, rosettes lost pericentric H3K9me3 and displayed mosaic 5-methylcytosine staining (Fig. 8d,e). Thus, the rosette exchanges constitutive with facultative heterochromatin marks, which is reversed after lumenogenesis. Moreover, pericentric H3K27me3 appeared after the premature induction of rosette-stage pluripotency in blastocysts by IWP2, and persisted when arresting embryos in rosette pluripotency by MEK inhibition (Fig. 8f,g). The striking accumulation of pericentric H3K27me3 in the embryonic rosette and in RSCs supports the shared nature of their pluripotent state. Finally, *Otx2*KO ESCs repressed pericentric H3K27me3 accumulation in LIM, while *Otx2*OE ESCs increased it (Fig. 8h), thus confirming the central role of OTX2 in controlling rosette pluripotency.

Discussion

We identified an intermediate pluripotent state in the embryonic rosette and demonstrated that WNT controls the naive-rosette transition and MEK the rosette-primed transition. This sequential action allows the capture of embryonic (ESC)[37,38], rosette-like (RSC) and epiblast (EpiSC)[39,40] stem cells. RSCs differ from ESCs

NATURE CELL BIOLOGY ARTICLES

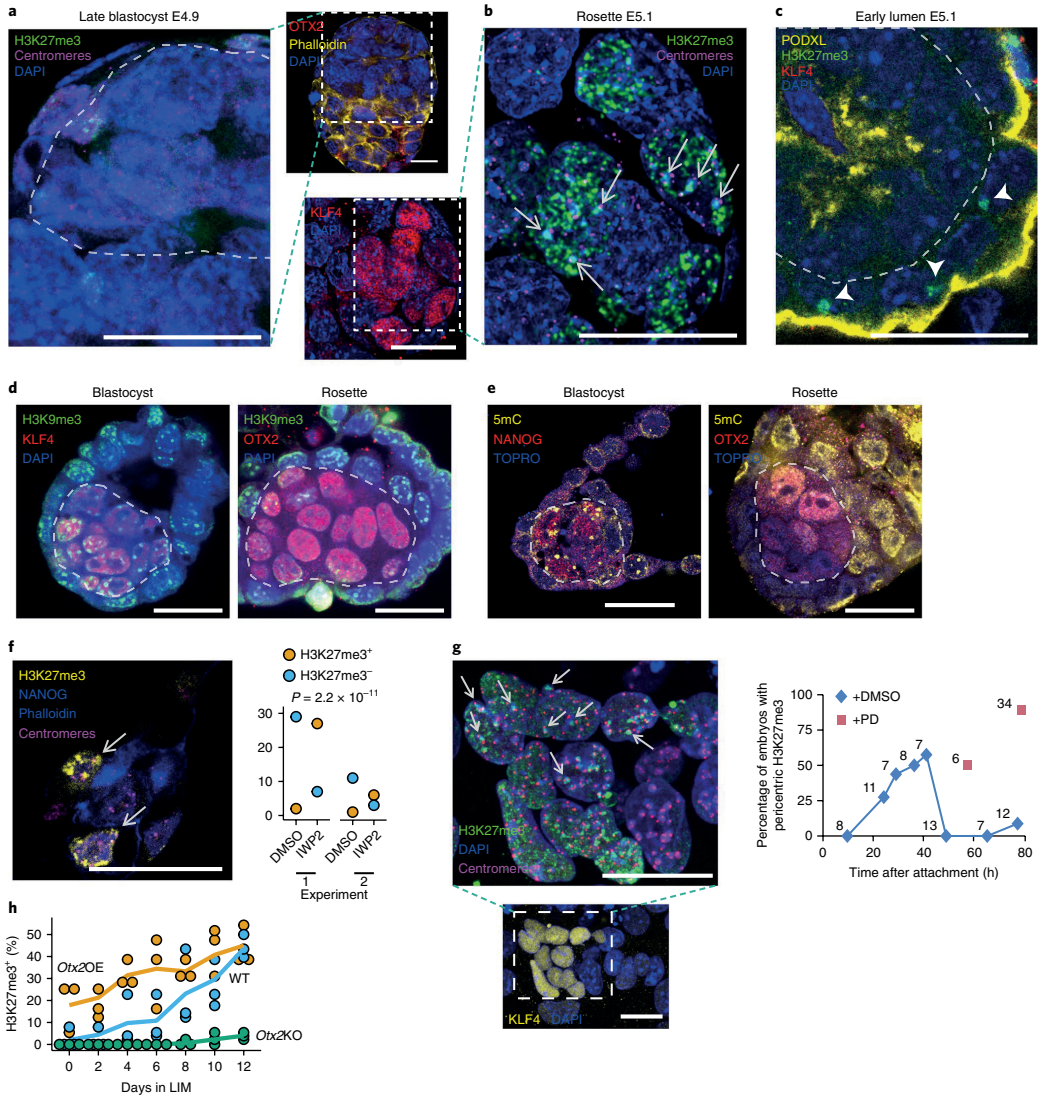


Fig. 8 | Rosette-specific formatting of pHC. a, Blastocyst (16 samples) stained for H3K27me3 (green), centromeres (magenta), OTX2 (red), phalloidin (yellow) and DAPI (blue) shows absence of OTX2 and H3K27me3 foci in the epiblast. **b**, Rosette (23 samples) stained for H3K27me3 (green), centromeres (magenta), KLF4 (red) and DAPI (blue). White arrows indicate some H3K27me3^{+} pHC foci (light blue, overlap of H3K27me3 and DAPI). **c**, Early lumen-stage embryo (15 samples) stained for H3K27me3 (green), PODXL (yellow) and DAPI (blue). No enrichment of H3K27me3 at the pHC is observed. The outer primitive endoderm layer displays enrichment of H3K27me3 on the inactive X chromosomes (arrowheads). **d**, Blastocyst (8 samples) and rosette (7 samples) embryos stained for KLF4 or OTX2 (red), H3K9me3 (green) and DNA (blue). **e**, Blastocyst (15 samples) and rosette (18 samples) embryos stained for NANOG or OTX2 (red), 5mC (yellow) and DNA (blue). The epiblast is encircled by dotted white line (**a**,**c**–**e**). **f**, Four-cell-stage embryos were cultured to the expanded blastocyst stage in the indicated conditions and the presence of pericentric H3K27me3 (arrows) indicated in a representative IWP2-treated embryo. Two independent experiments, $n = 86$ embryos (43 for each condition); two-sided Fisher's exact test. **g**, Left: embryos cultured in the presence of a MEK inhibitor (PD) for 57 h after attachment and stained for KLF4 (yellow), H3K27me3 (green), centromeres (magenta) and DAPI (blue). Right: the number of embryos cultured for the indicated time, and the fraction that displayed H3K27me3-enriched pHC foci. Each time point indicates an independent experiment. **h**, *Otx2*OE, WT or *Otx2*KO ESCs were cultured in LIM for the indicated duration and the percentage of cells displaying H3K27me3^{+} pHC foci quantified ($n = 3$ technical replicates). Scale bars, 20 μm (**a**–**g**).

Table 1 | Characteristics of mouse pluripotent stem cells

Characteristic	ESCs	RSCs	EpiSCs
Embryonic equivalent	E3.75–4.5 ICM[16]	Rosette	Pre-streak to early-streak epiblast[18]
Contribution to blastocyst chimeras	High efficiency	High efficiency	Low efficiency
Growth factor requirements	LIF + WNT[6]	LIF; inhibition of WNT and MEK signals	FGF2 + activin A[39,40]
Morphology	Dome-shaped colonies	Flat dispersed colonies	Flat colonies
Polarity	Unpolarized	Apical–basal polarity	Apical–basal polarity
Alkaline phosphatase	Strong	Faint	Faint
Clonogenicity	High	High	Low
Migratory	No	Yes	No
Transcriptome	Naive transcription factors	Naive transcription factors + OTX2	Primed transcription factors
X-inactivation status	XaXa	XaXa	XaXi
DNA methylation	Low	Intermediate	High
DNA hydroxymethylation	Low	High	Low
Promoter bivalency	Low	High in primed genes	Low
Main pHC mark	H3K9me3/5meC	H3K27me3	H3K9me3/5meC
Ability to efficiently revert to earlier state	Low	High	Low

XaXa, two active X chromosomes; XaXi, one inactive X chromosome.

by their bivalent marking of primed genes, facultative formatting of pHC, primed pluripotency progression in response to MEK activation and a similar transcriptome as cells halfway through the naive-primed transition, among other features (Table 1). We further showed that OTX2 acts downstream of WNT to establish rosette-like pluripotency, including the MEK responsiveness and pHC formatting. Finally, WNT and MEK signals control rosette formation and lumenogenesis, respectively, thereby coordinating pluripotency progression with embryonic morphogenesis (Fig. 1a).

Several studies argue that ESCs do not require WNT signals since they maintain naive gene expression and chimera-forming ability in their absence[10,41–43]. Our identification of rosette pluripotency—which shares those characteristics with ESCs—resolves this controversy: in the absence of WNT, ESCs convert into RSCs that share defining features with the embryonic rosette. Furthermore, our findings that WNT activation and MEK inhibition block sequential embryonic transitions—naive-rosette and rosette-primed, respectively—strongly argue that ESCs and RSCs represent distinct embryonic stages. Of note, the proposed attributes of the hypothetical 'formative' naive-primed intermediate, such as downregulation of naive gene expression and random X chromosome inactivation[5], are not observed in the rosette or RSCs.

Despite their primed chromatin landscape, WNT induces RSC-to-ESC reversion. This contrasts with the lack of reversion found in primed cells[21,44,45]. Epigenetic barriers such as DNA methylation and the repressive H3K9me3 histone mark enforce cell fate commitment[46–49]. The irreversibility of the primed state argues that MEK signals establish epigenetic barrier function in preparation for lineage specification. The erasure of H3K9me3 and 5meC and the accumulation of H3K27me3 in the rosette point towards unsuspected genome remodelling events, possibly clearing constitutive heterochromatin that would interfere with embryogenesis. In this light, it is suggestive that pericentric H3K27me3 maintains repression of the major satellites in ESCs when the H3K9me3 heterochromatin pathway is compromised[36,50]. H3K27me3 may act as a compensatory repressive mechanism in rosette-stage pluripotency until conversion to primed pluripotency. Further understanding of the ESC–RSC reversibility and RSC-primed transition may provide insight into the mechanisms of cell fate commitment and epigenetic barriers to cell reprogramming.

Recently, it was shown that L2i inhibits lumenogenesis in mouse embryos and maintains naive pluripotency, outlining a mechanism that coordinates the exit from naive pluripotency with amniotic cavity formation[3]. Our findings extend this seminal work by demonstrating that WNT signals link pluripotency progression to embryogenesis, not only by preventing rosette formation but also by preventing MEK signals from driving primed pluripotency progression. The rosette forms because of polarization of the ICM, which is induced by the deposition of a basal lamina by the primitive endoderm and the trophoblast[2]. The primitive endoderm differentiates from the ICM between E3.0 and E3.75 (refs. [51,52]). Interestingly, the mature primitive endoderm expresses the WNT antagonist DKK1 (refs. [53,54]) and may thereby signal its completion by repressing WNT signals, allowing *Otx2* induction, rosette formation and MEK-dependent pluripotency progression to proceed. Indeed, the pro-amniotic cavity is severely compromised in *Otx2* mutants[55], while mutants with constitutively activate β-catenin fail both to induce *Otx2* and form a pro-amniotic cavity[56,57]. Such a mechanism would constitute a checkpoint linking epiblast progression to maturation of the primitive endoderm by WNT signals.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41556-020-0508-x.

Received: 25 October 2018; Accepted: 20 March 2020;
Published online: 4 May 2020

References

1. Weinberger, L., Ayyash, M., Novershtern, N. & Hanna, J. H. Dynamic stem cell states: naive to primed pluripotency in rodents and humans. *Nat. Rev. Mol. Cell Biol.* **17**, 155–169 (2016).
2. Bedzhov, I. & Zernicka-Goetz, M. Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. *Cell* **156**, 1032–1044 (2014).
3. Shahbazi, M. N. et al. Pluripotent state transitions coordinate morphogenesis in mouse and human embryos. *Nature* **552**, 239–243 (2017).

4. Morgani, S., Nichols, J. & Hadjantonakis, A. K. The many faces of pluripotency: in vitro adaptations of a continuum of in vivo states. *BMC Dev. Biol.* **17**, 7 (2017).
5. Smith, A. Formative pluripotency: the executive phase in a developmental continuum. *Development* **144**, 365–373 (2017).
6. ten Berge, D. et al. Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. *Nat. Cell Biol.* **13**, 1070–1075 (2011).
7. Ying, Q. L. et al. The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519–523 (2008).
8. Tuysuz, N. et al. Lipid-mediated Wnt protein stabilization enables serum-free culture of human organ stem cells. *Nat. Commun.* **8**, 14578 (2017).
9. Yi, F. et al. Opposing effects of Tcf3 and Tcf1 control Wnt stimulation of embryonic stem cell self-renewal. *Nat. Cell Biol.* **13**, 762–770 (2011).
10. Wray, J. et al. Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nat. Cell Biol.* **13**, 838–U246 (2011).
11. Kemp, C., Willems, E., Abdo, S., Lambiv, L. & Leyns, L. Expression of all Wnt genes and their secreted antagonists during mouse blastocyst and postimplantation development. *Dev. Dyn.* **233**, 1064–1075 (2005).
12. ten Berge, D. et al. Wnt signaling mediates self-organization and axis formation in embryoid bodies. *Cell Stem Cell* **3**, 508–518 (2008).
13. Wang, Q. T. et al. A genome-wide study of gene activity reveals developmental signaling pathways in the preimplantation mouse embryo. *Dev. Cell* **6**, 133–144 (2004).
14. Albeck, J. G., Mills, G. B. & Brugge, J. S. Frequency-modulated pulses of ERK activity transmit quantitative proliferation signals. *Mol. Cell* **49**, 249–261 (2013).
15. Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S. & Saitou, M. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146**, 519–532 (2011).
16. Boroviak, T. et al. Lineage-specific profiling delineates the emergence and progression of naive pluripotency in mammalian embryogenesis. *Dev. Cell* **35**, 366–382 (2015).
17. Gardner, R. L., Lyon, M. F., Evans, E. P. & Burtenshaw, M. D. Clonal analysis of X-chromosome inactivation and the origin of the germ line in the mouse embryo. *J. Embryol. Exp. Morphol.* **88**, 349–363 (1985).
18. Kurek, D. et al. Endogenous WNT signals mediate BMP-induced and spontaneous differentiation of epiblast stem cells and human embryonic stem cells. *Stem Cell Reports* **4**, 114–128 (2015).
19. Masaki, H. et al. Interspecific in vitro assay for the chimera-forming ability of human pluripotent stem cells. *Development* **142**, 3222–3230 (2015).
20. Veluscek, G., Li, Y., Yang, S. H. & Sharrocks, A. D. Jun-mediated changes in cell adhesion contribute to mouse embryonic stem cell exit from ground state pluripotency. *Stem Cells* **34**, 1213–1224 (2016).
21. Kalkan, T. et al. Tracking the embryonic stem cell transition from ground state pluripotency. *Development* **144**, 1221–1234 (2017).
22. Gardner, R. L. & Cockroft, D. L. Complete dissipation of coherent clonal growth occurs before gastrulation in mouse epiblast. *Development* **125**, 2397–2402 (1998).
23. Buecker, C. et al. Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell Stem Cell* **14**, 838–853 (2014).
24. Yang, S. H. et al. Otx2 and Oct4 drive early enhancer activation during embryonic stem cell transition from naive pluripotency. *Cell Rep.* **7**, 1968–1981 (2014).
25. Acampora, D., Di Giovannantonio, L. G. & Simeone, A. Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition. *Development* **140**, 43–55 (2013).
26. Habibi, E. et al. Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. *Cell Stem Cell* **13**, 360–369 (2013).
27. Marks, H. et al. The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* **149**, 590–604 (2012).
28. Leitch, H. G. et al. Naive pluripotency is associated with global DNA hypomethylation. *Nat. Struct. Mol. Biol.* **20**, 311–316 (2013).
29. Ficz, G. et al. FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. *Cell Stem Cell* **13**, 351–359 (2013).
30. von Meyenn, F. et al. Impairment of DNA methylation maintenance is the main cause of global demethylation in naive embryonic stem cells. *Mol. Cell* **62**, 848–861 (2016).
31. Galonska, C., Ziller, M. J., Karnik, R. & Meissner, A. Ground state conditions induce rapid reorganization of core pluripotency factor binding before global epigenetic reprogramming. *Cell Stem Cell* **17**, 462–470 (2015).
32. Bernstein, B. E. et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006).
33. Azuara, V. et al. Chromatin signatures of pluripotent cell lines. *Nat. Cell Biol.* **8**, 532–538 (2006).
34. Mikkelsen, T. S. et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).
35. Kurimoto, K. et al. Quantitative dynamics of chromatin remodeling during germ cell specification from mouse embryonic stem cells. *Cell Stem Cell* **16**, 517–532 (2015).
36. Tosolini, M. et al. Contrasting epigenetic states of heterochromatin in the different types of mouse pluripotent stem cells. *Sci. Rep.* **8**, 5776 (2018).
37. 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).
38. Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154–156 (1981).
39. Brons, I. G. et al. Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191–195 (2007).
40. Tesar, P. J. et al. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, 196–199 (2007).
41. Lyashenko, N. et al. Differential requirement for the dual functions of β-catenin in embryonic stem cell self-renewal and germ layer formation. *Nat. Cell Biol.* **13**, 753–U365 (2011).
42. Biechele, S., Cockburn, K., Lanner, F., Cox, B. J. & Rossant, J. Porcn-dependent Wnt signaling is not required prior to mouse gastrulation. *Development* **140**, 2961–2971 (2013).
43. Augustin, I. et al. Autocrine Wnt regulates the survival and genomic stability of embryonic stem cells. *Sci. Signal.* **10**, eaah6829 (2017).
44. Bao, S. et al. Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. *Nature* **461**, 1292–1295 (2009).
45. Guo, G. et al. Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* **136**, 1063–1069 (2009).
46. Watanabe, A., Yamada, Y. & Yamanaka, S. Epigenetic regulation in pluripotent stem cells: a key to breaking the epigenetic barrier. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **368**, 20120292 (2013).
47. Messerschmidt, D. M., Knowles, B. B. & Solter, D. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes Dev.* **28**, 812–828 (2014).
48. Becker, J. S., Nicetto, D. & Zaret, K. S. H3K9me3-dependent heterochromatin: barrier to cell fate changes. *Trends Genet.* **32**, 29–41 (2016).
49. Bogdanovic, O. & Lister, R. DNA methylation and the preservation of cell identity. *Curr. Opin. Genet. Dev.* **46**, 9–14 (2017).
50. Saksouk, N. et al. Redundant mechanisms to form silent chromatin at pericentromeric regions rely on BEND3 and DNA methylation. *Mol. Cell* **56**, 580–594 (2014).
51. Chazaud, C., Yamanaka, Y., Pawson, T. & Rossant, J. Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2–MAPK pathway. *Dev. Cell* **10**, 615–624 (2006).
52. Chazaud, C. & Yamanaka, Y. Lineage specification in the mouse preimplantation embryo. *Development* **143**, 1063–1074 (2016).
53. Boroviak, T., Loos, R., Bertone, P., Smith, A. & Nichols, J. 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 (2014).
54. Hoshino, H., Shioi, G. & Aizawa, S. AVE protein expression and visceral endoderm cell behavior during anterior–posterior axis formation in mouse embryos: asymmetry in OTX2 and DKK1 expression. *Dev. Biol.* **402**, 175–191 (2015).
55. Ang, S. L. et al. A targeted mouse *Otx2* mutation leads to severe defects in gastrulation and formation of axial mesoderm and to deletion of rostral brain. *Development* **122**, 243–252 (1996).
56. Chazaud, C. & Rossant, J. Disruption of early proximodistal patterning and AVE formation in *Apc* mutants. *Development* **133**, 3379–3387 (2006).
57. Moser, A. R. et al. Homozygosity for the *Min* allele of *Apc* results in disruption of mouse development prior to gastrulation. *Dev. Dyn.* **203**, 422–433 (1995).
58. Eden, E., Navon, R., Steinfeld, I., Lipson, D. & Yakhini, Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* **10**, 48 (2009).
59. Supek, F., Bosnjak, M., Skunca, N. & Smuc, T. REVIGO summarizes and visualizes long lists of Gene Ontology terms. *PLoS ONE* **6**, e21800 (2011).

Methods

Isolation of timed embryos. All animal experiments were conducted after approval by the Centrale Commissie Dierproeven (the Hague, the Netherlands) and by the Erasmus MC Animal Ethical Committee. Embryos were obtained from crosses between actin–GFP male[60] and superovulated C57Bl/6JolaHsd or FVB/N female mice. The male mice and newly weaned female mice were maintained under a shifted day–night rhythm (lights on at 1:00, lights off at 13:00). Superovulation was performed using an adjusted protocol: 5-week-old females received 0.1 ml of CARD HyperOva (Cosmo Bio) intraperitoneally at 8:00, and 48 h later, 5 units of 50 U ml^{-1} human chorionic gonadotropin. The superovulated females were housed with actin-GFP males after the human chorionic gonadotropin injection and checked for copulation plugs 3 and 6 h later. For isolation of peri-implantation and post-implantation embryos, isolated uteri were dissected under the lens of a fluorescence stereomicroscope to facilitate the detection of implanted GFP$^+$ embryos. The embryos were then used in immunostaining, RNA isolation or culture protocols as described below.

The mouse *Otx2*–GFP strain, STOCK Tg(*Otx2*-EGFP)BD199Gsat/Mmnc, RRID:MMRRC_000239-UNC, was obtained from the Mutant Mouse Resource and Research Center (MMRRC) at Taconic/University at Albany, a NIH-funded strain repository, and was donated to the MMRRC by N. Heintz, The Rockefeller University. Embryos were obtained from mating *Otx2*–GFP males with superovulated females and isolation of the embryos at E1.5. Embryo culture was performed as described below.

Immunostaining and imaging of embryos. Embryos were fixed for 20 min in 4% paraformaldehyde (PFA), washed twice in PBS containing 0.05% Tween-20 (PBT) and permeabilized for 20 min in 0.3% Triton X-100 plus 0.1 M glycine in PBT. Samples were then washed three times in PBT and blocked for 1 h in blocking buffer consisting of 5% normal donkey serum plus 1% BSA in PBT. For 5mC and 5hmC staining, embryos were treated after permeabilization with 4 M HCl plus 0.1% Triton X-100 in water for 10 min, rinsed in PBT plus 20% blocking buffer and blocked as described above. The embryos were then incubated overnight at room temperature with primary antibodies in blocking buffer. The next day, embryos were washed three times in PBT, incubated for 2 h with secondary antibodies and again washed three times in PBT. If applicable, embryos were then incubated with 5 units per ml of phalloidin (labelled with Alexa-546 or Alexa-488) (ThermoFisher) in PBT for 20 min and washed in PBT. Embryos were then incubated with DAPI or, for 5mC or 5hmC staining, TO-PRO-3 in PBT for 10 min, rinsed in PBT and mounted on coverslips in Prolong Gold (ThermoFisher). The Prolong Gold was allowed to set for at least 1 day, after which embryos were imaged using a Leica SP5 confocal laser scanning microscope with 405-, 488-, 561-, 594- and 633-nm laser lines and ×40 or ×63 oil objectives. Images depicting pHC in embryos were deconvolved using Huygens Deconvolution software (Scientific Volume Imaging). Contrast settings were adjusted if necessary and equally over all relevant images using ImageJ. The antibodies used are listed in Supplementary Table 6.

Embryo culture. For culture of four-cell-stage embryos, embryos were isolated at E1.5 from superovulated females and cultured in KSOM medium (Millipore) until the four-cell stage, or in vitro fertilizations were performed as previously described[61] with modifications. In vitro fertilization culture medium, Mouse Vitro Fert (MVF; Cook Medical), was used for sperm incubation, in vitro fertilization and zygote culture. Straws with frozen sperm were thawed in a 37 °C water bath for 30 s. The sperm was pushed out of the straw into 500 µl of MVF and incubated for 11–15 min at 37 °C, 5% CO_2, 5% O_2. Meanwhile four to five superovulated 4-week-old female mice per fertilization drop were used as oocyte donors. After 4 h of co-incubation, the presumptive zygotes were washed three times in 50-µl MVF drops and those appearing normal were cultured overnight in a 50-µl MVF drop at 37 °C, 5% CO_2, 5% O_2. The next day, two-cell embryos were transferred into KSOM medium and cultured at 37 °C, 5% CO_2. Next, four-cell embryos were collected during a 10-h window and transferred into KSOM containing 2 µM IWP2 (Merck, 681671) or the same concentration of dimethylsulfoxide (DMSO) and cultured for 60 h. They were then fixed, immunostained and imaged as described above.

To culture blastocysts to the egg cylinder stage, morula-stage embryos (C57BL/6NRj, Janvier Quickblasto, or freshly isolated) or E4.0 blastocysts (see "Isolation of timed embryos") were cultured in KSOM medium until the expanded blastocyst stage. They were then transferred to 8-well iBidi plates containing IVC1 medium[62]. E4.0 blastocysts were cultured directly into iBidi plates containing IVC1. After attachment of the embryos, the medium was switched to IVC2 (ref. [62]), containing 1 µM PD as indicated. Embryos were then fixed, immunostained and imaged as described above.

Cell culture. Mouse ESCs were cultured on gelatine and FCS-coated plates in N2B27 medium supplemented with 1,000 U ml^{-1} LIF, 3 µM CHIR and 1 µM PD (L2i ESCs). Coating was performed by first briefly incubating the wells with 0.1% gelatin in PBS, then incubating them with FCS for 1 h at 37 °C followed by a rinse with PBS. N2B27 medium consisted of one volume DMEM/F12 combined with one volume Neurobasal medium supplemented with 0.5% N2 supplement, 1% B27 supplement, 0.033% BSA 7.5% solution, 50 µM β-mercaptoethanol, 2 mM Glutamax, 100 U ml^{-1} penicillin and 100 µg ml^{-1} streptomycin (all from Invitrogen). Cells were cultured at 37 °C and 5% CO_2 and passaged 1:3–1:10 every 2–3 days by triturating the colonies into a single cell suspension using 0.05% trypsin–EDTA. Trypsin was quenched using soybean trypsin inhibitor (Sigma). Epiblast stem cells (EpiSCs) were cultured on gelatine and FCS-coated plates in N2B27 medium supplemented with 20 ng ml^{-1} activin A, 12 ng ml^{-1} bFGF (both Peprotech) and 2 µM IWP2 to suppress spontaneous differentiation[18]. EpiSCs were passaged 1:4–1:10 every 3 days by triturating the colonies into small clumps using 0.5 mg ml^{-1} collagenase IV (Sigma). All cell lines in culture were tested for mycoplasma contamination every 2–3 months.

To transit from embryonic stem to MEK/WNT-inhibited or primed cells, trypsinized ESCs were seeded onto gelatine and FCS-coated plates in N2B27 medium supplemented with 1,000 U ml^{-1} LIF, 2 µM IWP2 and either 1 µM PD for MEK/WNT-inhibited cells or 20 ng ml^{-1} activin A and 12 ng ml^{-1} bFGF for primed cells. After 2 days, MEK/WNT-inhibited cells were passaged as ESCs, whereas primed cells were passaged as EpiSCs.

The following ESC lines were used (see Supplementary Table 6): R1 and CGR8 (both obtained from Stanford Transgenic Facility), IB10, SV7, FN3 and SV8 (all described in ref. [6]; SV8 is female and FN3 is from the non-permissive FVB/N background), RGd2 (ref. [21]), *Tcf7l1*$^{-/-}$ (ref. [9]), *Tbx3*–RFP;*Oct6*–GFP[23], *Otx2*KO and *Otx2*OE[25]. The following RSC lines were derived from blastocysts in this study (see below): 129LIMA2, 129LIMB1, 129LIMB3 (all male) and 129LIMA6 (female). For the generation of TRE–*Otx2* ESCs, a mouse *Otx2* complementary DNA (accession number BC017609) was amplified using primers incorporating an amino-terminal Kozak sequence and SalI and NotI restriction sites. The vector p2Lox.GFP[63] was digested with SalI and NotI to remove the GFP sequence, and the SalI/NotI-digested *Otx2* fragment ligated into its place to create p2Lox.*Otx2*. A total of 25 µg of p2Lox.*Otx2* was electroporated into A2Lox.Cre ESCs, followed by induction of the cells with 1 µg ml^{-1} doxycycline overnight to induce recombination, and the cells selected on 250 µg ml^{-1} of G418, as previously described[64].

IWP2 (Merck, 681671), CHIR (Axon Medchem), PD (Merck) were diluted from 2 mM (IWP2) or 10 mM stocks in DMSO and used at 2, 3 and 1 µM, respectively. WNT3A protein was used at 400 ng ml^{-1}. Recombinant mouse WNT3A protein was produced in *Drosophila* S2 cells grown in suspension culture (a gift from R. Nusse, Stanford University) and purified using Blue Sepharose affinity and gel filtration chromatography[65]. For this, the S2 cells were expanded in Schneider's *Drosophila* medium (Lonza) containing 10% fetal bovine serum and antibiotics, and media were collected when cell expansion reached a plateau. Up to 12 litres of conditioned medium was 0.45-µm filtered, adjusted to 1% Triton X-100 and then applied to a fast protein liquid chromatography column containing 200 ml Blue Sepharose 6 Fast Flow (GE Healthcare, 17094801). After washing with four volumes of washing buffer (150 mM KCl, 20 mM Tris-HCl, 1% CHAPS, pH 7.5), bound proteins were eluted with elution buffer (1.5 M KCl, 20 mM Tris-HCl, 1% CHAPS, pH 7.5). After analysis on a Coomassie gel, WNT3A-containing fractions were selected, combined and concentrated to 10 ml using Pierce Concentrators 20 K MWCO (Thermo Scientific, 89887A). The combined fractions were fractionated on a HiLoad 26/60 Superdex 200 gel filtration column (GE Healthcare) in PBS, 0.5 M NaCl, 1% CHAPS, pH 7.3. Fractions were analysed for purity and WNT activity by Coomassie gels and WNT activity reporter assays, respectively, and selected fractions were combined for use.

Derivation and blastocyst injection of RSCs. E3.5 blastocysts were obtained from crosses between 129S2/SvHsd males and C57Bl/6JolaHsd females (Envigo). Blastocysts were flushed from the uteri and placed intact in 96-wells plates coated with gelatine and FCS and filled with 100 µl N2B27 medium plus LIF. Late next day, when hatching and attachment of the embryos had occurred, an equal volume of medium containing 4 µM IWP2 and 2 µM PD was added. Embryo outgrowths were trypsinized 4 days later, the trypsin quenched with soybean trypsin inhibitor, and plated in 96-wells plates coated with gelatine and FCS in N2B27 containing LIF, 2 µM IWP2 and 1 µM PD (RSC medium). After 4 days, wells containing colonies were passaged and expanded further in RSC medium. RSCs were recovered by treatment with 0.05% trypsin–EDTA and 10–15 cells were injected into blastocysts collected from C57Bl/6 mice. Blastocysts (10–15 per mouse) were transferred into the uterus of a pseudo pregnant mouse. The percentage of coat colour chimerism of the chimeras was visually estimated. Chimeras were mated with C57Bl/6 mice to determine germ line transmission.

Xist fluorescent in situ hybridization (FISH). For Xist RNA FISH and immunofluorescence staining, cells were adsorbed to poly-lysinated coverslips. Cells were fixed for 10 min with 4% PFA in PBS at room temperature, washed with 70% ethanol, permeabilized for 4 min with 0.2% pepsin at 37 °C and post-fixed with 4% PFA in PBS for 5 min at room temperature. Coverslips were washed twice with PBS and dehydrated in a gradient of 70%, 90% and 100% ethanol. A nick-labelled DNA probe (digoxigenin) was dissolved in hybridization mixture (50% formamide, 2XSSC (1XSSC: 0.15 M NaCl, 0.015 M sodium citrate), 50 mM phosphate buffer (pH 7.0) and 10% dextran sulfate) and 100 ng µl^{-1} mouse Cot DNA to a final concentration of 1 ng µl^{-1}. Probe was denatured for 5 min, prehybridized for 45 min at 37 °C, and coverslips were incubated overnight in a humid chamber at 37 °C. After hybridization, coverslips were washed once

in 2XSSC, three times in 50% formamide–2XSSC, both at 37 °C, and twice in TST (0.1 M Tris, 0.15 M NaCl, 0.05% Tween 20) at room temperature. Blocking was done in BSA–TST for 30 min at room temperature. Detection was done via subsequent steps of incubation with anti-digoxigenin (Boehringer) and two fluorescein isothiocyanate (FITC)-labelled antibodies for Xist RNA detection in blocking buffer for 30 min at room temperature. Coverslips were washed twice with TST between detection steps and once finally with TS (0.1 M Tris, 0.15 M NaCl). Dehydrated coverslips were mounted with ProLong Gold Antifade with DAPI (Molecular Probes).

Immunohistochemistry of cells. Cells were passaged onto gelatine- and FCS-coated glass slides 1 day before fixation. Cells were fixed with 4% PFA in PBS at room temperature for 10 min, washed with PBS containing 0.05% Tween 20 (PBST), permeabilized with 0.3% Triton X-100 plus 0.1 M glycine in PBST, washed with PBST, incubated in blocking buffer for 1 h (1% BSA, 5% donkey serum in PBST) and incubated overnight at room temperature in blocking buffer with the primary antibodies. The next day, slides were washed with PBST before incubation for 2 h with the secondary antibodies in blocking buffer. Nuclei were then stained with DAPI (Molecular Probes) in PBST for 5 min at room temperature, and slides mounted using Prolong Gold. Immunofluorescence confocal images were acquired using a Leica SP5 inverted laser-scanning confocal microscope. Contrast settings were adjusted if necessary and equally over all relevant images using ImageJ. The antibodies used are listed in Supplementary Table 6.

For alkaline phosphatase staining, cells were stained according to the manufacturer's protocol (Alkaline Phosphatase Detection kit SCR004, Millipore).

In vitro rosette and lumen assays. ESCs were triturated to a single cell suspension using 0.05% trypsin–EDTA (Life Technologies) and washed with PBS. Cells were resuspended in ice-cold BME (BME 2 RGF, AMSBIO, 3533-010-02) at a density of 5,000 cells per µl. Drops of 15 µl were plated on coverslips or in a prewarmed 8-chamber slide (Thermo Scientific, 154534) and incubated for 5 min at 37 °C until the BME solidified. The plate was then filled with prewarmed medium consisting of DMEM plus 15% FCS, 1× MEM nonessential amino acids, 50 µM β-mercaptoethanol, 100 U ml^{-1} penicillin and 100 µg ml^{-1} streptomycin (all from Life Technologies), containing supplements as indicated, and cultured at 37 °C and 5% CO_2. The cells and BME drops were then fixed with 4% PFA, prewarmed to 37 °C, for 30 min. For phalloidin staining, the fixed cells were rinsed with PBS, incubated with phalloidin–Alexa-488 (5 units per ml, Invitrogen) and DAPI in PBS for 1.5 h, and washed three times with PBS. For immunostaining, the fixed cells were permeabilized with 0.3% Triton/0.1 M glycine in PBS for 30 min at room temperature, washed with PBS containing 0.05% Tween 20 (PBST) for 20 min, incubated in blocking buffer for 1 h (1% BSA, 5% donkey serum in PBST) and incubated overnight at room temperature in blocking buffer with the primary antibodies. The next day, cells were washed three times for 30 min with PBST before incubation for 2 h with the secondary antibodies in blocking buffer. After three washes with PBST for 30 min, nuclei were stained with DAPI (Molecular Probes) for 30 min at room temperature and washed three times with PBS. For lumen assays, the live cultures were incubated for 1 h in a cell culture incubator with culture medium containing 1:500 anti-PODXL antibody (MAB1556, R&D Systems), followed by fixation and staining with secondary antibody as described above. The live staining ensured that only PODXL exposed at the cell exteriors was stained. The stained preparations were covered with Prolong Gold and imaged using a Zeiss Axioplan microscope or a Leica SP5 inverted laser-scanning confocal microscope. Structures were designated as rosette if they showed polarized, central localization of phalloidin or intracellular PODXL. Structures displaying central localization of PODXL after live staining were designated as lumen. The antibodies used are listed in Supplementary Table 6.

RNA-seq and data analysis. Total RNA was prepared using TriReagent (Sigma) according to the manufacturer's protocol. RNA-seq was performed at the Erasmus MC Center for Biomics (Rotterdam, the Netherlands). The RNA-seq library was prepared for analysis according to the Illumina TruSeq stranded mRNA protocol (www.illumina.com). Briefly, 200 ng of total RNA was purified using poly-T oligo-attached magnetic beads to result in poly-A containing mRNA. The poly-A tailed mRNA was fragmented, and cDNA was synthesized using SuperScript II and random primers in the presence of actinomycin D. cDNA fragments were end-repaired, purified with AMPure XP beads, and A-tailed using Klenow exo-enzyme in the presence of dATP. Paired-end adapters with dual index (Illumina) were ligated to the A-tailed cDNA fragments and purified using AMPure XP beads. The resulting adapter-modified cDNA fragments were enriched by PCR using Phusion polymerase as follows: 30 s at 98 °C, 15 cycles of (10 s at 98 °C, 30 s at 60 °C, 30 s at 72 °C), 5 min at 72 °C. PCR products were purified using AMPure XP beads and eluted in 30 µl of resuspension buffer. One microliter was loaded on an Agilent Technologies 2100 Bioanalyzer using a DNA 1000 assay to determine the library concentration and for quality checking. Cluster generation was then performed according to the Illumina TruSeq SR Rapid Cluster kit v2 (cBot) reagents preparation guide (www.illumina.com). Briefly, six RNA-seq libraries were pooled together to get a stock of 10 nM. One microlitre of the 10 nM stock was denatured with NaOH, diluted to 6 pM and hybridized onto the flow cell. The hybridized products were sequentially amplified, linearized and end-blocked according to the Illumina Single Read Multiplex Sequencing user guide. After hybridization of the sequencing primer, sequencing-by-synthesis was performed using the HiSeq 2500 with a single-read 50-cycle protocol followed by dual index sequencing.

Base calling and demultiplexing was performed using Illumina bcl2fastq v.2.19. Sequencing reads were aligned against GRCm38 (mm10) using HiSat2 v.2.0.4. Count data files were created using htseq-count v.0.6.0. Additional embryo and L2i ESC data[3,16] were downloaded from ArrayExpress E-MTAB-2958 and E-MTAB-5147. E-MTAB-2958 was aligned against GRCm38 and count data files created as described above, whereas count data files were already available from E-MTAB-5147.

Normalization and differential expression analysis were performed using R v.3.4.3 and the DESeq2 Bioconductor package v.1.18.1 with betaPrior set to TRUE, a False Data Rate cut-off alpha value of 0.1, and a log-fold change threshold lfcThreshold of 1. Heatmaps and PCA plots were generated from rlog-transformed data using the pheatmap package 1.0.8 and the DESeq2 function plotPCA, respectively, using the 500 genes displaying the highest row variance for PCA.

Significant differentially expressed genes were analysed for GO Biological Process terms enrichment using GOrilla[58], using a *P* value threshold of 0.001 and a background list of all genes that had an average read count of at least 1 over all samples. GO analysis was further processed using REVIGO[59].

Single-cell sequencing. Single-cell mRNA sequencing was performed according to an adapted version of the SORT-seq protocol[66]. In brief, single cells were FACS-sorted into 384-well plates containing 384 primers and mineral oil (Sigma). After sorting, plates were snap-frozen on dry ice and stored at −80 °C. For amplification, cells were heat-lysed at 65 °C followed by cDNA synthesis using the CEL-Seq2 protocol[67] and robotic liquid-handling platforms. After second-strand cDNA synthesis, the barcoded material was pooled into libraries of 384 cells and amplified using in vitro transcription. Following amplification, the rest of the CEL-seq2 protocol was followed for preparation of the amplified cDNA library using TruSeq small RNA primers (Illumina). The DNA library was paired-end sequenced on an Illumina Nextseq 500, high output, with a 1 × 75 bp Illumina kit (read 1: 26 cycles, index read: 6 cycles, read 2: 60 cycles).

During sequencing, read 1 was assigned 26 bp and was used for identification of the Illumina library barcode, cell barcode and unique molecular identifier. Read 2 was assigned 60 bp and used to map to the reference transcriptome of mm10 with Burrows–Wheeler Aligner. Data demultiplexing and statistical UMI poisson counting correction was done as previously described[68]. Mapping and generation of count tables was performed using the MapAndGo script (https://github.com/anna-alemany/transcriptomics/tree/master/mapandgo). Single-cell transcriptomics analysis was done using the RaceID3 algorithm, following an adapted version of the RaceID manual (https://cran.r-project.org/web/packages/RaceID/vignettes/RaceID.html). In total, 1,152 cells were sequenced. After removing cells with fewer than 8,000 UMIs and only keeping genes that were detected with at least 3 UMIs in at least 2 cells, 938 cells and 11,070 genes were left for further analysis. During clustering, genes contributing to batch effects (*Rn45s*, *Kcnq1ot1*, *Malat1*, A630089N07Rik, *Rpl13*, *Rpl39* and *Rps7*) were excluded from clustering, but were kept in the dataset for downstream visualization and analysis. Clustering was done using hclust. Differential gene expression analysis was done as previously described[66] using an adapted version of the DESeq2 algorithm[69].

Flow cytometry. For flow cytometry experiments, single-cell suspensions were made using trypsin–EDTA (5 min in 37 °C) or Enzyme-Free Cell Dissociation Buffer (Gibco) for 30 min at 37 °C and washed with 5% FCS in PBS. Hoechst 33258 (1 µg ml^{-1}) was used for live/dead cell assessment. Cells were analysed for GFP reporter gene expression using a FACS Fortessa (BD Biosciences), and data were analysed using FlowJo.

Live imaging. R1 L2i and MEK/WNT-inhibited cells were plated at clonal density on serum-coated or BME-coated plates and incubated overnight at 37 °C. The cells were tracked with CytoMate and pictures taken every minute. For the analysis, a stack of 12 pictures (taken at 5-min intervals, representing 1 h of tracking) was imported to ImageJ. Velocity was calculated, and the movement dotplots created using the 'Manual Tracking' plugin.

Real time RT–PCR analysis. Total RNA was prepared using a RNeasy mini or micro kit (Qiagen) according to the manufacturer's instructions. To extract RNA from embryos, four embryos from the same litter were combined, while other littermates were used for immunofluorescence to establish the presence of a rosette or lumen. Reverse transcription was performed using Superscript II (Invitrogen) or, for embryo RNA, a RevertAid RT Reverse Transcription kit (ThermoFisher, K1691). Quantitative PCR was carried out on a BioRad CFX machine using a SYBR-green-based master mix. Relative quantification was done using *Gapdh* as a reference gene or, for the analysis of embryo RNA, *Oct4* as a specific epiblast gene. All PCR assays were carried out in triplicate and the mean cycle threshold (CT) value was used for quantification. Primer sequences were designed such that they spanned splice junctions whenever possible, and sequences are provided in Supplementary Table 6.

Mass spectrometry of 5mC and 5hmC nucleosides. Mass spectrometry of 5mC and 5hmC was performed as previously described[70]. In short, for each sample, 2 µg of DNA (isolated from ~1×10^6 cells) was degraded into individual nucleosides using DNA Degradase Plus (Zymo Research). The individual nucleosides were measured using a high-performance liquid chromatography–tandem mass spectrometry (HPLC–MS/MS) system consisting of an Acquity UPLC (Waters) containing a Waters Atlantis Hilic column (2.1 mm × 100 mm, 3 µm) connected to a Micromass Quattro Premier XE (Waters). Quantification was performed using area-based linear regression curves derived from calibration standards containing internal standard solutions. The 5mC and 5hmC levels were calculated as a concentration percentage ratio of the percentage of 5-methyl-2′-deoxycytidine/2′-deoxyguanosine (%mdC/dG) and the percentage of 5-hydroxymethyl-2′-deoxycytidine/2′-deoxyguanosine (%hmdC/dG), respectively.

ChIP-seq analysis. R1 and CGR8 ESCs were cultured on gelatine and FCS-coated plates in N2B27 supplemented with LIF, 2 µM IWP2 and 1 µM PD for 8 days. Confluent cultures were fixed in PBS with 1% formaldehyde (Sigma) for 8 min and quenched using 100 mM fresh glycine for 15 min at room temperature. Fixed cell preparations were collected by scraping, then washed and sonicated in 1% SDS, 30 mM Tris pH 8.0, 150 mM NaCl using a Diagenode Bioruptor UCD-300 to a size range with a median of 200 bp. A total of 1 million cell equivalent chromatin was diluted seven times to 15 mM Tris pH 8.0, 1 mM EDTA, 150 mM NaCl, 1% Triton X-100 with fresh protease inhibitor cocktail (Roche) and 0.5 µg H3K27me3 or 1 µg H3K4me3 antibody (Diagenode) and incubated overnight at 4 °C with rotation. Protein A/G magnetic beads were washed in the same dilution buffer with 0.15% SDS and 0.1% BSA, added to the chromatin–antibody mix and rotated for 60 min at 4 °C. Beads were washed with 400 µl buffer for 5 min at 4 °C using (1) first a 150 mM NaCl wash buffer, (2) then twice a 500 mM NaCl buffer and (3) finally a 10 mM Tris buffer. DNA was eluted from the beads using 150 mM NaCl, 30 mM Tris, 1% SDS, 0.1 µg µl^{-1} Proteinase K buffer for 30 min at 55 °C followed by 2 h shaking at 65 °C. The beads were removed, after which the DNA was purified using a Qiaquick MinElute PCR purification kit (Qiagen) according to the manufacturer's protocol, and eluted in 20 µl of buffer EB.

Illumina library preparation was performed using a Kapa Hyper Prep kit using 5 ng of DNA isolated via ChIP. For end repair and A-tailing, double-stranded DNA was incubated with end repair and A-tailing buffer and enzyme and incubated first for 30 min at 20 °C and then for 30 min at 65 °C. Adapters were ligated by adding 30 µl ligation buffer, 10 µl Kapa DNA ligase, 5 µl 600 nM adaptor in a total volume of 110 µl and incubated for 15 min at 15 °C. Post-ligation clean-up was performed using Agencourt AMPure XP reagent, and products were eluted in 20 µl elution buffer. Libraries were amplified by adding 25 µl 2× KAPA HiFi Hotstart Ready-mix and 5 µl 10× Library Amplification Primer Mix followed by 10 PCR cycles. Samples were purified using a QIAquick MinElute PCR purification kit. Libraries were size- selected to ~300 bp using Agencourt Ampure XP beads. Correct size selection was confirmed using an Agilent BioAnalyzer. Sequencing was performed using an Illumina NextSeq500 with 42-bp read lengths. Sequencing of the input showed no specific signal.

Additional L2i[27] and EpiLC[35] H3K4me3 and H3K27me3 ChIP-seq data were downloaded from NCBI's Gene Expression Omnibus (GEO). The data were mapped against the mm9 reference genome using bowtie v.2.0.2. Duplicate and low-confidence mapping reads (mapq < 15) were removed. After sequencing and mapping of the ChIP-seq libraries, we normalized for the total number of mapped sequencing reads to allow quantitative comparisons between samples by generating reads per million mapped reads (RPM) values. The RPM values allow samples to be compared regardless of differences in sequencing depth. Peak calling was performed using SICER v.1.1 with window size 200, gap size 200 (for H3K4me3) or 600 (for H3K27me3) at an *E*-value of 0.1. Bivalency was demarcated using bedtools v.2.20.1 via the intersection of H3K4me3 and H3K27me3 replicate consistent peak-calls. Bedtools was used to collect read counts within peaks. DESeq2 v.1.16.1 was used for rlog transformation and normalization of the read counts. The data were clustered and visualized using R v.3.1.1, ngsplot v.2.61, bam2bw v.1.21 and the UCSC genome browser.

Preparation and immunostaining of chromosome spreads. Chromosome spreads were prepared as previously described[71]. Briefly, cells were treated with 0.1 µg ml^{-1} colcemid (Gibco) for 3 h. Cells arrested in mitotic prometaphase were collected by mitotic shake-off and spun down in culture medium containing colcemid for 10 min at 200 × *g*. Supernatant was removed to leave 1 ml of culture medium and cells were resuspended. An equal volume of a hypotonic buffer (pH 8.2) containing 30 mM Tris-HCl, 17 mM trisodium citrate dihydrate, 50 mM sucrose and 5 mM EDTA was added. After a 7 min of incubation, cells were spun down for 7 min at 200 × *g* and supernatant removed. Cells were resuspended in 100 mM sucrose (pH 8.2) to obtain a cell concentration of $5–15 \times 10^6$ cells per ml. From this suspension, 10 µl was applied to a slide dipped in a solution of 1% PFA and 0.15% Triton X-100 (pH 9.2). After drying for 1.5–2 h in a humid chamber, slides were washed with 0.08% Photo-Flo (Kodak Industrie), air dried and stored at −20 °C until further processing.

Immunofluorescence was performed as previously described[72]. The antibodies used are listed in Supplementary Table 6. Results from five independent experiments were evaluated by two observers, blinded to the culture condition and the type of modification detected. At least 100 metaphases were randomly selected and categorized per slide. The pHC region was identified by the DAPI-dense area between the centromeres and scored for absence or enrichment of the marker.

Statistics and reproducibility. No statistical methods were used to predetermine the sample sizes. The number and type of replicates and any statistical tests used are indicated in the figure legends, and all replicates successfully reproduced the presented findings. Most key experiments were repeated with cell lines and embryos from different genetic backgrounds (that is, from 129, C57Bl6 and FVB strains). Investigators were blinded to the experimental groups for analysis of in vitro rosette assays (Fig. 2j) and for analysis of chromosome spreads (Fig. 7b). Normality of data was tested using Kolmogorov–Smirnov test. Statistics were calculated using Graphpad Prism 8.3.0 or as indicated.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Figures 3 and 4 have associated RNA-seq and single-cell RNA-seq data deposited in the GEO and are accessible through GEO series accession number GSE145727. Figure 6 has associated ChIP-seq data deposited in GEO with the accession number GSE112234 . Previously published RNA-seq data that were re-analysed here are available from ArrayExpress under accession codes E-MTAB-2958 and E-MTAB-5147. Previously published ChIP-seq data that were re-analysed here are available from the GEO under accession codes GSE23943 and GSE60204. Source data for Figs. 1, 2 and 4–8 and Extended Data Figs. 1, 3–6 and 8 are presented with the paper. All other data supporting the findings of this study are available from the corresponding author upon reasonable request.

Code availability

No custom computer code was used in this study.

References

60. Okabe, M., Ikawa, M., Kominami, K., Nakanishi, T. & Nishimune, Y. 'Green mice' as a source of ubiquitous green cells. *FEBS Lett.* **407**, 313–319 (1997).
61. Ostermeier, G. C., Wiles, M. V., Farley, J. S. & Taft, R. A. Conserving, distributing and managing genetically modified mouse lines by sperm cryopreservation. *PLoS ONE* **3**, e2792 (2008).
62. Bedzhov, I., Leung, C. Y., Bialecka, M. & Zernicka-Goetz, M. In vitro culture of mouse blastocysts beyond the implantation stages. *Nat. Protoc.* **9**, 2732–2739 (2014).
63. Iacovino, M., Roth, M. E. & Kyba, M. Rapid genetic modification of mouse embryonic stem cells by Inducible Cassette Exchange recombination. *Methods Mol. Biol.* **1101**, 339–351 (2014).
64. Kyba, M., Perlingeiro, R. C. & Daley, G. Q. HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors. *Cell* **109**, 29–37 (2002).
65. Willert, K. et al. Wnt proteins are lipid-modified and can act as stem cell growth factors. *Nature* **423**, 448–452 (2003).
66. Muraro, M. J. et al. A single-cell transcriptome atlas of the human pancreas. *Cell Syst.* **3**, 385–394.e3 (2016).
67. Hashimshony, T. et al. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-seq. *Genome Biol.* **17**, 77 (2016).
68. Grun, D., Kester, L. & van Oudenaarden, A. Validation of noise models for single-cell transcriptomics. *Nat. Methods* **11**, 637–640 (2014).
69. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
70. Kroeze, L. I. et al. Characterization of acute myeloid leukemia based on levels of global hydroxymethylation. *Blood* **124**, 1110–1118 (2014).
71. van de Werken, C. et al. A universal method for sequential immunofluorescent analysis of chromatin and chromatin-associated proteins on chromosome spreads. *Chromosome Res.* **21**, 475–489 (2013).
72. Grewal, S. I. & Jia, S. Heterochromatin revisited. *Nat. Rev. Genet.* **8**, 35–46 (2007).

Acknowledgements

We are grateful to the following colleagues for providing cell lines used in this study: B. Merrill (*Tcf7l1*$^{-/-}$), A. Smith (RGd2), J. Wysocka (*Tbx3*-RFP;*Oct6*-GFP), A. Simeone (*Otx2*KO and *Otx2*OE) and M. Kyba (A2lox.Cre). We thank M. Muraro from Single Cell Discoveries for help with single-cell sequencing and data analysis, and R. van Albada and A. Korporaal for technical assistance. NWO ECHO.10.B1.064, TI Pharma D5-402, Marie Curie FP7-PEOPLE-2009-RG-256560, ZonMW 116006104 (all to D.t.B.), NWO-VIDI 864.12.007 (to H.M.) and FES NIRM (Dutch Innovation Award) supported this study.

Author contributions

A.N., I.E., E.v.G., L.V., D.K., J.L., J.S., M.M. and D.t.B. designed and performed experiments and analysed data. A.N., I.E., D.K., J.L., J.S., R.A.M.D., G.v.M., M.M., M.D., E.B.B., H.M. and D.t.B. prepared the manuscript. J.S., D.t.B., R.A.M.D., G.v.M. and H.M. designed, performed and analysed the ChIP-seq experiments. D.t.B., J.S., A.T.d.D., R.W.W.B. and W.F.J.v.I. performed and analysed the RNA-seq experiments. E.v.G., D.t.B. and N.C.R. performed and analysed the single-cell sequencing experiments. A.N., C.E., E.v.G. and E.B.B. prepared and analysed the chromosome spreads. A.M. assisted with the mouse experiments and Y.G. performed additional experiments. H.M. and J.H.J. designed, performed and analysed nucleoside mass spectrometry analyses. D.t.B. conceived the study and coordinated the work.

Competing interests

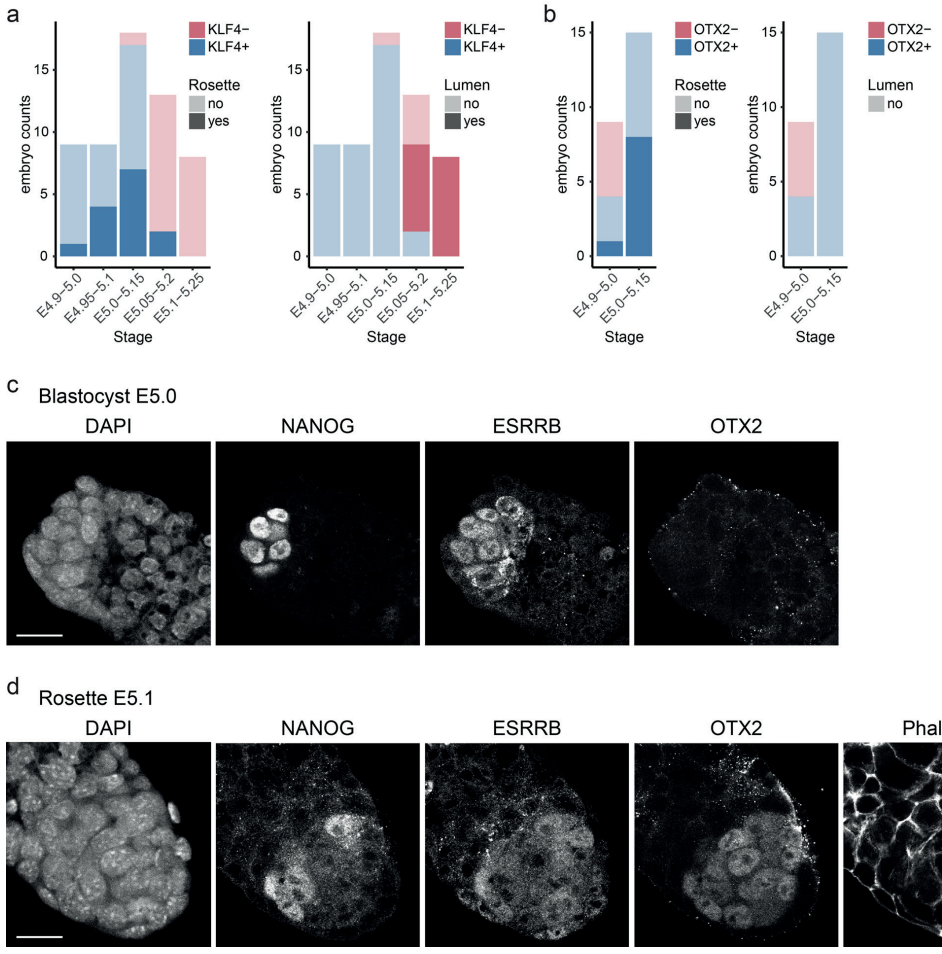
The authors declare no competing interests.

Additional information

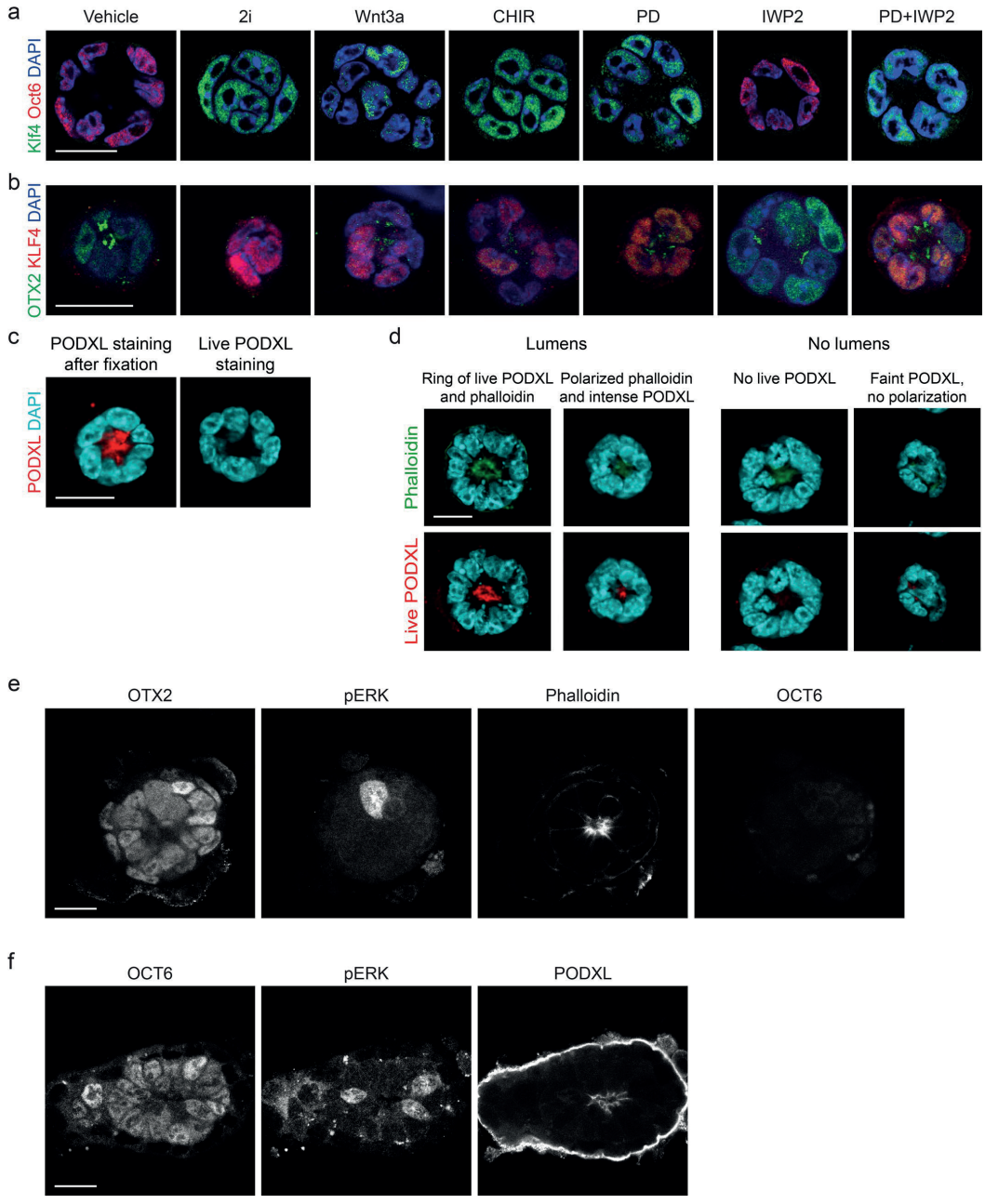
Extended data is available for this paper at https://doi.org/10.1038/s41556-020-0508-x.

Supplementary information is available for this paper at https://doi.org/10.1038/s41556-020-0508-x.

Correspondence and requests for materials should be addressed to D.t.B.

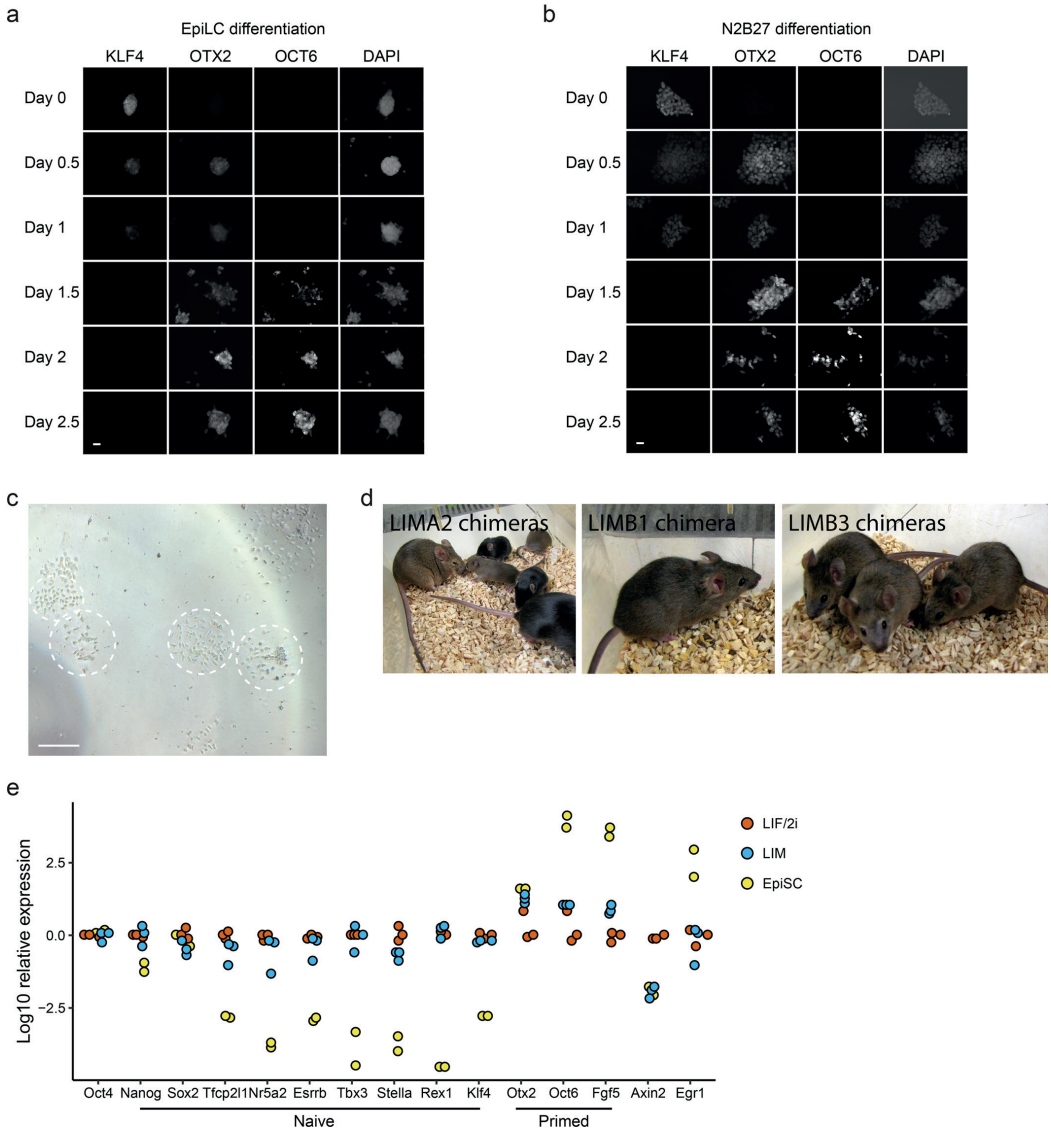


Extended Data Fig. 1 | Naive and primed markers in the embryonic rosette. a,**b,** Counts of embryos of the indicated stages displaying a rosette or lumen and KLF4 (**a**) or OTX2 (**b**) expression status. **a**, and (**b**) show different litters. **c**, Individual colour channels for Fig. 1h (representative of 5 experiments). **d**, Individual colour channels for Fig. 1i (representative of 8 experiments). Scale bars 20 µm.

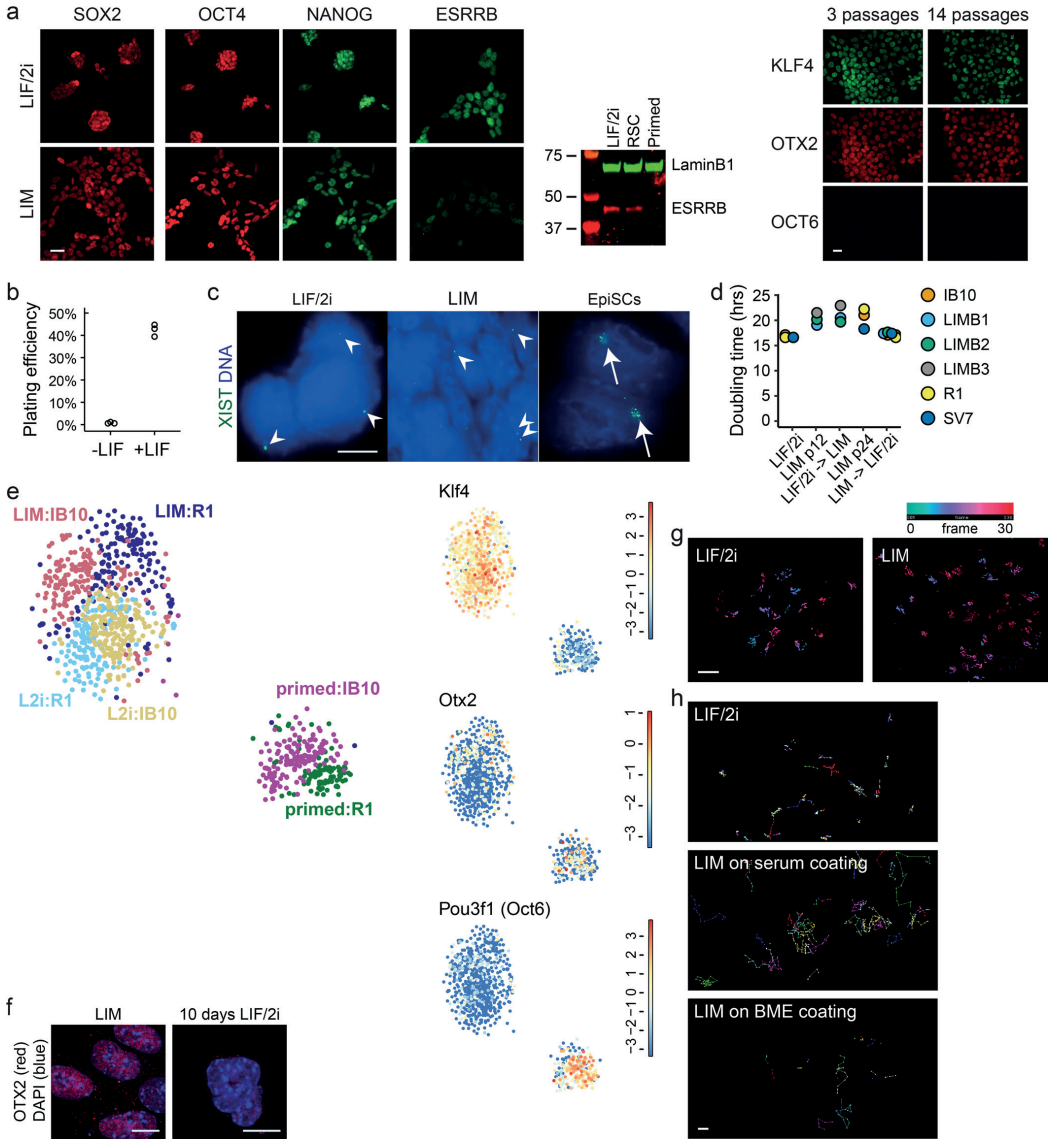


Extended Data Fig. 2 | See next page for caption.

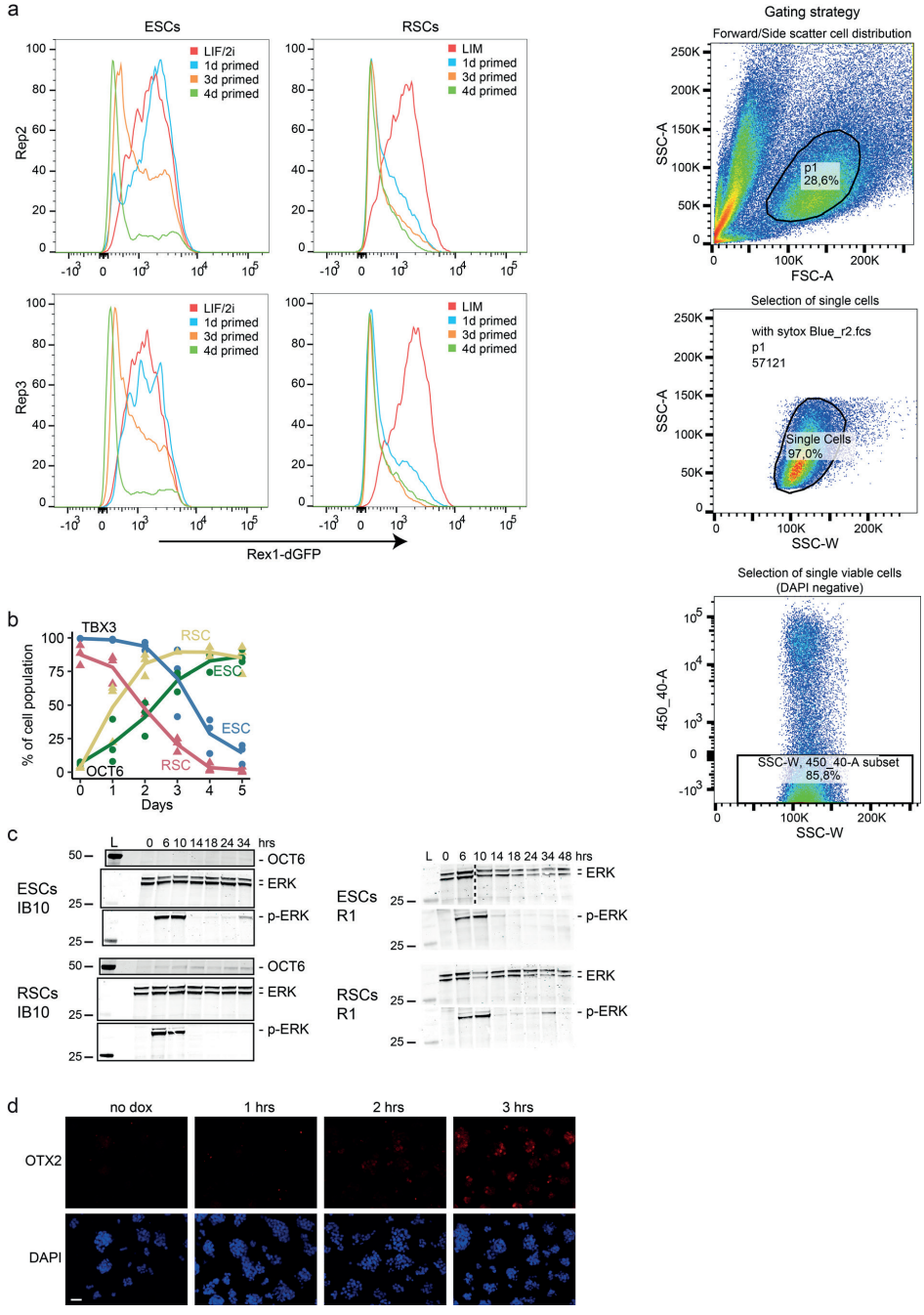
Extended Data Fig. 2 | WNT and MEK signals couple successive steps in morphogenesis to pluripotency progression. (**a**, **b**) Representative images of rosettes and aggregates generated by ESCs 48 hrs after seeding in BME in the indicated conditions and stained for the indicated markers. **a**, KLF4 (green) and OCT6 (red). **b**, OTX2 (green) and KLF4 (red). 3 independent experiments with similar results. **c**, Rosettes generated by ESCs 48 hrs after seeding in BME and stained for PODXL live or after fixation and permeabilization. Permeabilized samples showed PODXL staining, incorrectly suggesting presence of a lumen, while live staining revealed that the PODXL was not exposed at the cell surface and no lumen was actually present (illustrative examples, repeated > 5 times). **d**, Examples to illustrate criteria for calling ESC aggregates displaying a lumen, using live PODXL staining combined with phalloidin to reveal polarization (illustrative examples, repeated > 5 times). **e**, Individual colour channels for Fig. 2n (representative of 3 experiments). **f**, Individual colour channels for Fig. 2o (representative of 4 experiments). Scale bars 20 μm (a–f).



Extended Data Fig. 3 | WNT controls the initial phase of the naive-primed transition. (**a**, **b**) Immunostaining for the indicated markers of ESCs plated in EpiLC conditions (**a**) or N2B27 medium (**b**) for the indicated number of days (single experiments). **c**, Dispersed colonies (dashed white circles) obtained after passaging of a blastocyst grown out in the presence of IWP2, MEK inhibitor and LIF (LIM). 15 from 18 and 12 from 13 embryos yielded cell lines. **d**, Chimeras (brown), obtained after blastocyst injection of newly derived LIM cell lines (LIMA2, LIMB1 and LIMB3), with mate (black) and pups (only for LIMA2). Brown pups indicate that germline transmission occurred. **e**, RT-PCR analysis of ESCs maintained for 6 passages in L2i or MEK/WNT-inhibited conditions and of EpiSCs. n = 3 (LIF/2i and LIM) or n = 2 (EpiSC) biological replicates. Scale bars 20 μm (**a**,**b**), 100 μm (**c**).

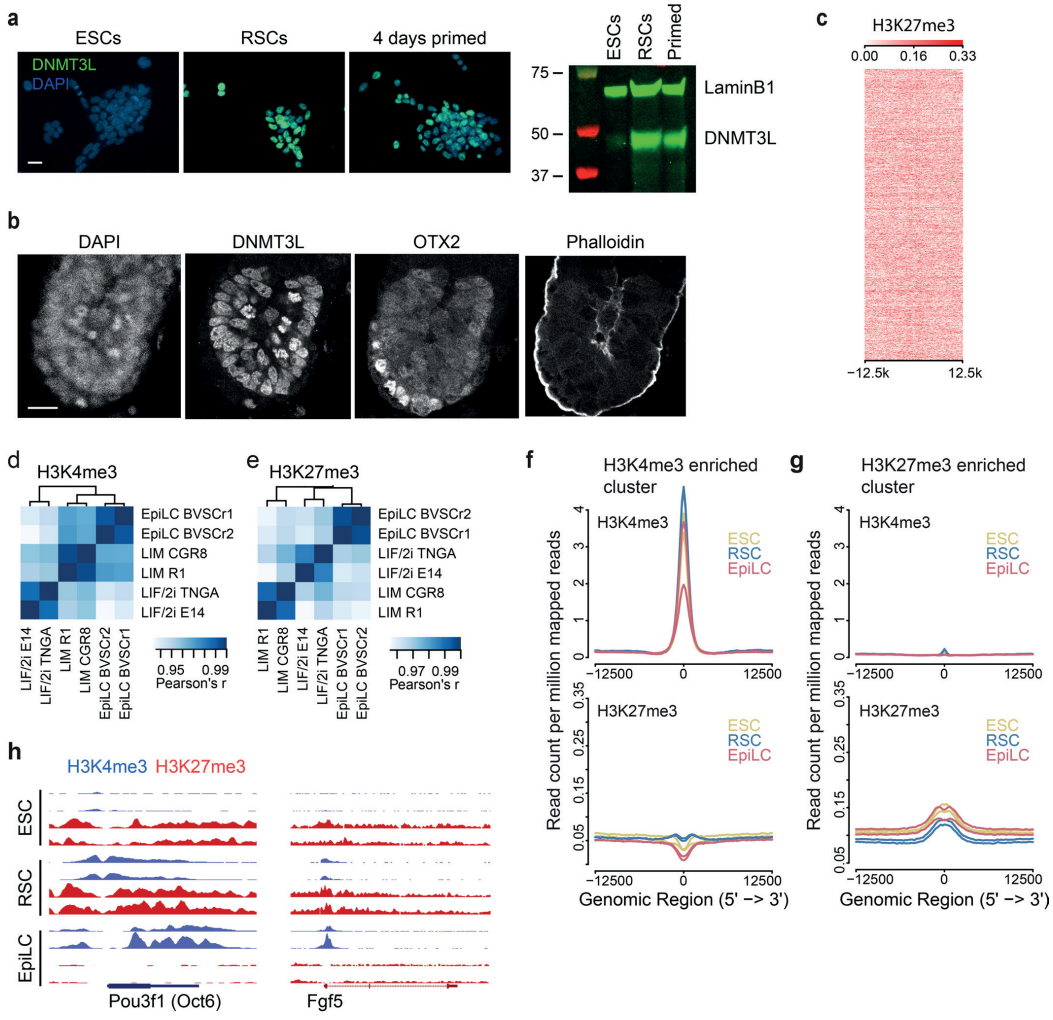


Extended Data Fig. 4 | MEK/WNT-inhibition maintains rosette characteristics. a, Immunostainings and western blot for the indicated markers of R1 ESCs maintained in L2i or LIM for 6 passages or as indicated (3 independent repeats with similar results; western blot 1 experiment, full blot in Source Data). **b**, Plating efficiencies of LIM cells in the presence and absence of LIF (n = 3 biological replicates, R1, FN3 and IB10 cells). **c**, Female SV8 L2i ESCs, LIM cells and EpiSCs stained for Xist by RNA-FISH. Puncta indicating Xist expression (arrowheads) are visible in both L2i and LIM cells, while only EpiSCs display Xist clouds indicating X-chromosome inactivation (arrows) (two independent experiments). **d**, Doubling times of the indicated cell lines in the indicated conditions. p12 and p24 indicate passage numbers of the cell lines. **e**, t-SNE plot of single cell RNA-Seq data, and log2 expression levels of the indicated genes plotted into the t-SNE plot. n = 159, 178, 102 (R1) and 184, 149 and 164 (IB10) cells for L2i, LIM and primed, respectively. **f**, RSCs were cultured for 10 days in the indicated conditions and immunostained for OTX2 (red). Representative images from 3 independent replicates. **g**, To determine apicobasal polarity, the Golgi apparatus was marked by Gm130 antibody staining, imaged by confocal microscopy and colorized according to z-stack level. Basal side of the cells correspond to low z-levels (green), and apical side to high levels (red). The data show apical Golgi localization in LIM cells, indicating their apicobasal polarity. Three experiments with similar results. **h**, Live tracking of individual cell locations during 24 hrs of L2i ESCs, LIM cells on serum-coating, and LIM cells on BME-coating. 80, 47 and 20 cells, respectively. Representative of 4 experiments. Scale bars 20 µm (a,g), 10 µm (c,f), 100 µm (h).

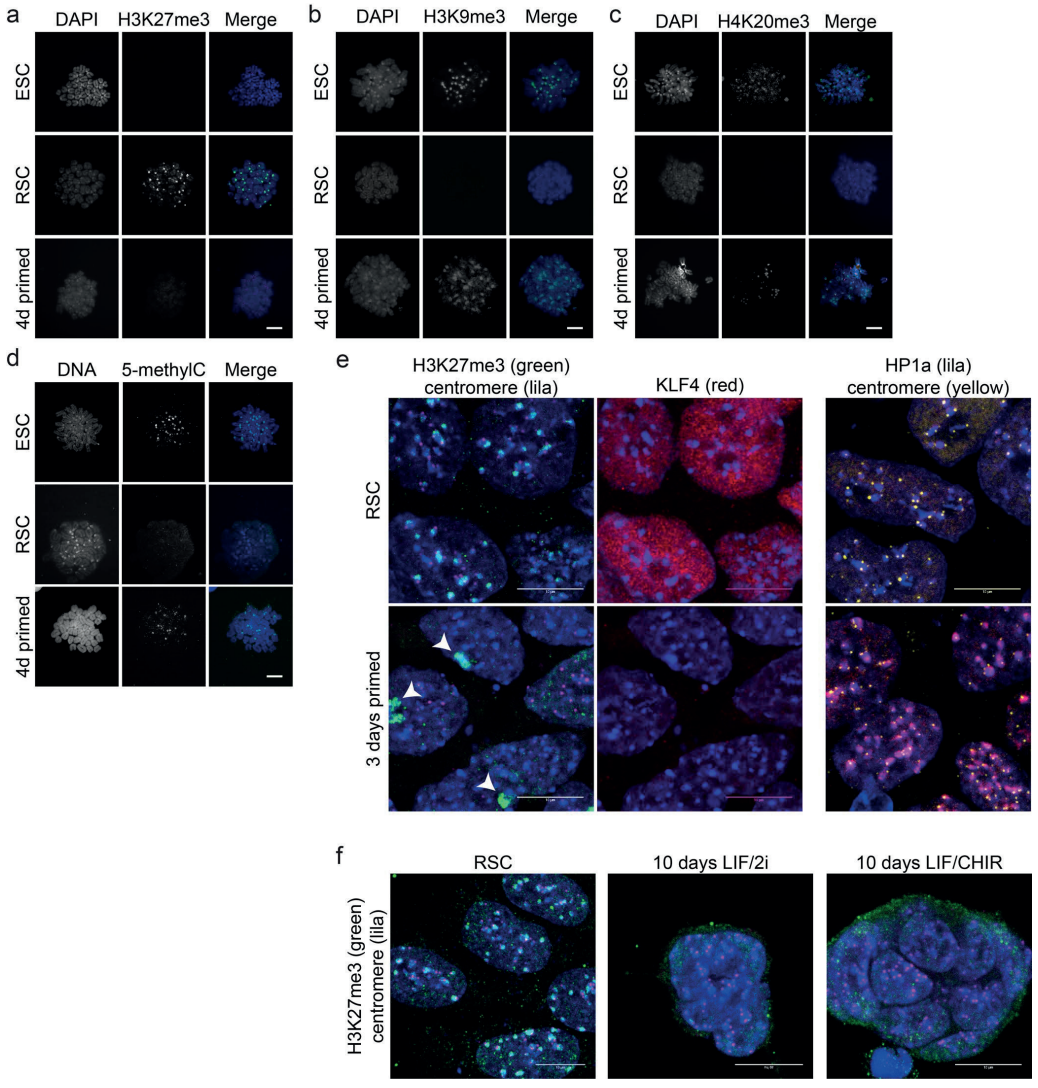


Extended Data Fig. 5 | See next page for caption.

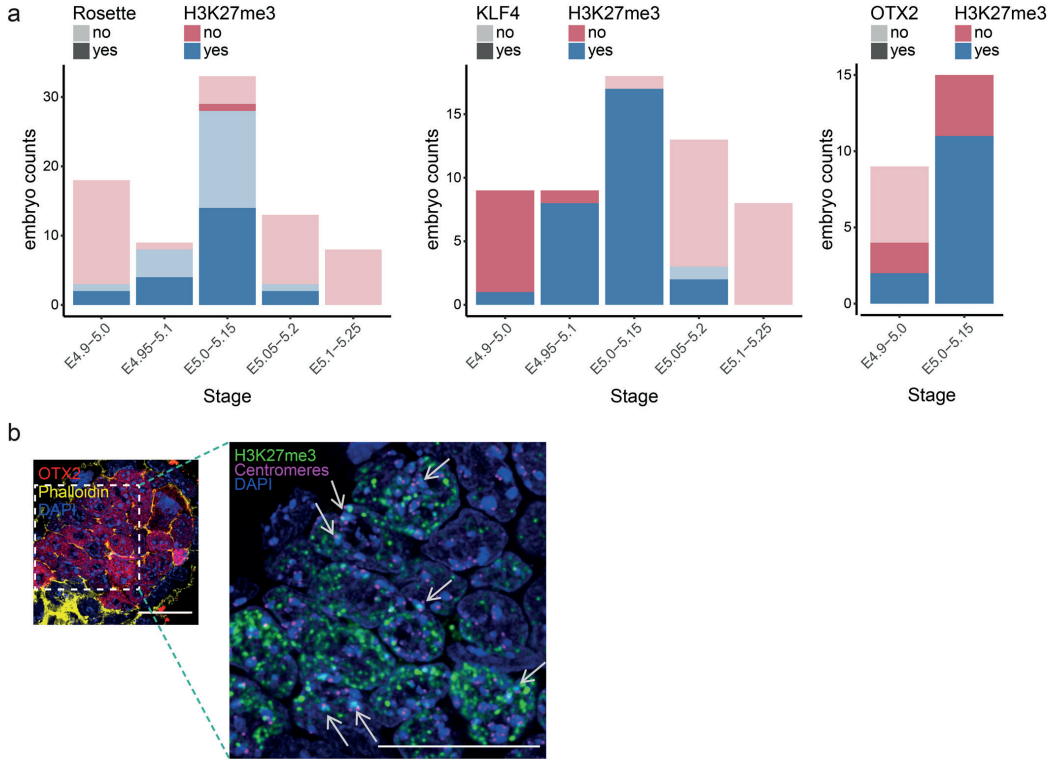
Extended Data Fig. 5 | Otx2 enables MEK-driven transition to primed pluripotency. a, RGd2 ESCs and RSCs were plated in primed conditions and analysed at several time points for GFP by flow cytometry. n = 3 independent replicates (3rd is shown in main Fig. 5a). **b**, ESCs and RSCs, carrying the Tbx3-RFP;Oct6-GFP reporters, were plated in primed conditions and analysed daily for expression of the reporters. Lines go through means, n = 3 independent experiments. **c**, Western blots showing OCT6 (only for IB10), ERK and phospho-ERK in IB10 and R1 ESCs and RSCs upon transfer to primed medium. Two (for OCT6) and three biological replicates. Full blots in Source Data. **d**, Immunostaining for OTX2 in TRE-Otx2 ESCs following induction by doxycycline for the indicated duration (single experiment). Scale bar 40 µm.



Extended Data Fig. 6 | A primed chromatin landscape in rosette-like stem cells. a, Immunofluorescence and western blot analysis for DNMT3L expression in the indicated cell types (3 independent repeats with similar results; western blot 1 experiment). **b**, Individual colour channels for Fig. 6b, right panel. Representative of 5 experiments. **c**, Representative input plot for H3K27me3 ChIP-Seq for the bivalent cluster. (**d**,**e**) H3K4me3 (d) or H3K27me3 (**e**) ChIP-Seq correlogram of ESCs, RSCs and EpiLCs. Two biological repeats for each condition. (**f**, **g**) Average intensity plots of H3K4me3 and H3K27me3 in the H3K4me3 (f) or H3K27me3 (g) enriched cluster. Two biological repeats for each condition. **h**, Screenshots from the UCSC genome browser showing H3K4me3 (blue) and H3K27me3 (red) peaks on the *Pou3f1* and *Fgf5* genes. Scale bars 20 μm (a,b).



Extended Data Fig. 7 | Rosette-like stem cells remodel pericentric heterochromatin. (**a–d**) Metaphase chromosome spreads of ESCs, RSCs and primed cells stained for the indicated markers. DAPI (blue), centromeres (red). **e**, Female RSCs (LIMA6) were cultured for 3 days in primed conditions and immunostained for the indicated markers. Arrowheads indicate H3K27me3-positive inactive X chromosomes demonstrating transition to primed pluripotency. **f**, RSCs were cultured for 10 days in the indicated conditions and immunostained for the indicated markers. Five (a–d) or three (e, f) biological replicates. Scale bars 20 μm (a-d), 10 μm (e, f).



Extended Data Fig. 8 | Rosette-specific formatting of pericentric heterochromatin. a, Counts of embryos of the indicated stages displaying pericentric H3K27me3 accumulation, and in which the presence of a rosette, KLF4 or OTX2 expression is detected. Embryos are the same as those of Extended Data Fig. 1a, b. **b**, Confocal microscopy images of E5.1 embryo stained for H3K27me3 (green), centromeres (magenta), OTX2 (red), and DAPI (blue). Some H3K27me3-positive pHC foci are indicated by white arrows (light blue, overlap of H3K27me3 and DAPI). Representative example from the embryos analysed in Extended Data Fig. 8a. Scale bars 20 μm.