THE UNIVERSITY *of* EDINBURGH

The roles of the transcriptional regulator Id1 in pluripotency and differentiation

Mattias Malaguti

Thesis presented for the degree of Doctor of Philosophy

MRC Centre for Regenerative Medicine

The University of Edinburgh

2014

DECLARATION

I declare that this thesis is of my own composition. The work presented in this thesis is my own, unless otherwise stated, and has not been presented for any other degree or professional qualification.

Mattias Malaguti

“Meno ignorante sei, più ignorante diventi.”

“The less ignorant you are, the more ignorant you will become.”

Elvio Malaguti

ACKNOWLEDGEMENTS

I owe an immense debt of gratitude to Sally for entrusting a young inexperienced biology enthusiast with the tools and the resources to carry out research in what is objectively and impartially speaking the most interesting of all scientific topics: stem cell biology. Her gentle manners, optimistic attitude and uncanny ability to make an unexpected observation sprout from a chaotic mess of data helped me cheer up on the long, cold and rainy winter days; her willingness to let me undertake side experiments and cloning follies, whilst making sure my primary aims did not lose focus, helped me energise on the long, cold and rainy summer days; her accessibility and availability to openly discuss results and plans in an earnest and relaxed way made my PhD studies more pleasant and enjoyable than I could have hoped for.

I have had the good fortune to be mentored by two other scientists who in sporting terms would undoubtedly be described as "world class". Ian has a tremendous analytical mind and has over the years helped me focus on those parts of my project that needed refining whilst being incredibly supportive of my work. Val's superhuman embryology skills and knowledge of mouse development opened up avenues of research that would have been unthinkable for a student with my molecular background, and have helped me place a lot of *in vitro* observations in perspective.

Every single person in the Lowell lab has influenced my project as well as me as a person, and my thanks go to everyone. Xinzhi was like a mother in the lab, and I am extremely envious of her daughters for taking her away from us. Paul was a great colleague and we both came to rely heavily on our scientific discussions during our two years as post-doc-less students. His peculiar yet sharp sense of humour never failed to keep me entertained. Alex was a freakishly skilled honours student who guided me in my first cloning experiments and who has since achieved tremendous success as a PhD student in Cambridge. Chia-Yi's arrival provided me with a great conversation companion and with a terrific example of tireless work ethics, as well as with someone who could fully understand and share my frustration at failed gene targeting experiments. She is also by far the best-informed member of the lab when it comes to the all-important institute gossip. Guillaume radically transformed the potential of the Lowell lab. It is unclear to me how a man who has two young daughters and who spends his nights looking through a telescope and reading programming books rather than sleeping could have any mental energy left to feed himself. Not only has he mastered this basic survival task, he has an in-depth understanding of the field of early development, and an ability to approach this field from different angles than those that have been thus far considered. His quantitative

approach to imaging and interest in fate specifiers other than transcription factors have influenced my work and my scientific thinking greatly. Amy is a great student whose embryological skills I greatly admire and who has always been of great company. Seeing Juraj join the lab was a pleasure almost as great as his beard. His incomprehensible conversations with Guillaume provided golden comedy moments and his wilfulness to help out with experiments was of great aid in periods of cell culture overload. Julia's enthusiasm and chirpiness are extremely contagious and she is a great person to be around. I'm sure she has a great future ahead of her despite her day-to-day supervisor's attempts to sabotage her with dodgy protocols and her allergy to maths. Tulin only rarely ventures out of the confines of the Chambers lab but it's always a pleasure to see her with a smile on her face.

A number of other people in the institute have provided me with invaluable assistance and advice over the years, and space constraints mean I can only mention those whom I have bothered the most. The Developmental Biology group members have always given great feedback on my work in a relaxed and cosy atmosphere. Nicola provided me with the Nanog-RFP targeting construct and has always been available for discussion, Yali has given me plenty of embryo advice and reagents, Kumiko was key in helping me to set up single-cell qRT-PCR experiments, and Freddy and Suresh were always at hand for cloning discussions. Lynsey performed all the morula aggregations and transfers, Ron sectioned embryos, Olivia, Simon and Claire performed sorts for me and taught me how to use the analysers, and without Audrey, Euan and Robbie the Institute could not function. A lot of my work has been carried out in cell culture and it is the cell culture staff that I am deeply indebted to. Jon and Theresa have done great jobs at the helm but it is Helen and Marylin who are the unsung heroes. They put in an unbelievable amount of work in a chronically understaffed department, and without them the amount of research that we carry out in cell culture would be unthinkable.

I also wish to thank Dr. Robert Benezra and Dr. Hyung-Song Nam for their kind gift of the Id1-Venus targeting construct which proved crucial to my studies.

My biggest thanks go to my family. My parents and sisters provided me with a loving and supportive environment whilst feeding and stimulating my curiosity and thirst for knowledge from when I was a small child.

I don't think any of this would have been possible without Angela. She has put up with my stress, frustrations, late hours and weekends in the lab, unwillingness to go on holiday, resistance to do any form of housework, interminable hours of writing, and has kept me clean, fed and loved throughout. I look forward to paying my debt to her when she starts writing her thesis in less than a year's time.

Apologies to my aunt Jasna for failing to create a stem cell cream to make her skin look younger. You will be missed.

ABSTRACT

The transition from pluripotency to differentiation is a key event in the life of all complex multicellular organisms. In the development of the mouse, the pluripotent epiblast undergoes gastrulation and gives rise to three multipotent germ layers, which will in turn form the tissues of the adult body. The events leading up to gastrulation have been extensively studied *in vivo* in developing embryos, and modelled *in vitro* making use of embryonic stem (ES) cells.

Bone morphogenic protein (BMP) signalling plays a key role in these processes. BMP can in fact maintain ES cells in a self-renewing state by inhibiting their differentiation into neural ectoderm, whilst at the same time being required for the specification of mesoderm in the developing embryo (Winnier *et al.* 1995, Ying *et al.* 2003a). A key intracellular target of BMP is the transcriptional regulator Id1, which can recapitulate the effects of BMP in the preservation of ES cell pluripotency and in the inhibition of neural specification from pluripotent cells (Ying *et al.* 2003a).

This thesis will focus on understanding the roles of this molecule in the early decisions affecting the transition from pluripotency to differentiation. In particular, I aim to study the expression pattern of Id1 in cultures of pluripotent cells, and to clarify which extracellular and intracellular molecules regulate the expression of the factor; I aim to understand how forced Id1 expression inhibits the differentiation of pluripotent cells, and whether Id1 may play a similar role in the regulation of the asynchronous exit from pluripotency observed in differentiating wild-type cells; finally, I aim to characterise the expression pattern of Id1 in the early stages of post-implantation development at the single-cell resolution, and to understand how the expression of the molecule correlates with the previously characterised expression patterns of key signalling molecules and transcription factors.

The generation of a reporter ES cell line expressing the yellow fluorescent protein Venus fused to the C-terminus of Id1 allowed me to assess the expression of the factor in culture on a single-cell basis, making use of immunofluorescence and flow

cytometry. I observed that expression of Id1 is reliant on active BMP signalling and low Activin/Nodal signalling, and I characterised the combinatory effects of the two pathways on Id1 expression. Furthermore, I demonstrated that high Nanog expression is incompatible with high Id1 expression in ES cell cultured in the presence of LIF and serum, which raises the possibility that Nanog may be affecting the expression of Id1 *in vivo*, both in pre-implantation and in post-implantation embryos.

I generated ES cell lines overexpressing Id1 and observed that the factor inhibits differentiation of pluripotent cells into neural ectoderm by delaying their exit from a post-implantation epiblast-like pluripotent state, and ultimately favouring mesodermal specification. This suggests that Id1 is acting at a specific stage of differentiation and that the differentiation process itself is following a similar developmental pathway to what is observed in the peri-gastrulation stage embryo.

I performed single-cell transcriptional analysis on differentiating wild-type ES cells and observed that Id1 is not expressed at an appropriate point in time to affect the asynchronous the exit from pluripotency observed in neural adherent monolayer differentiation, which suggests that other factors must be responsible for this phenomenon.

Finally, I addressed the expression pattern of Id1 protein in the embryonic tissue of gastrulating mouse embryos by imaging chimaeric embryos generated using the Id1-Venus reporter ES cells. I observed that Id1 is expressed in the proximal regions of streak stage embryos; in the epiblast and migrating mesendoderm of bud stage embryos; in cardiac, lateral and allantoic mesoderm and in foregut endoderm in headfold stage embryos. These expression patterns fit with the reported expression of BMP molecules at these stages of development, and suggest that Id1 expression is primarily dependent on BMP expression in early post-implantation embryos. However, I also observed Id1 expression in a ring of cells surrounding the node in headfold stage embryos, a previously uncharacterised expression pattern not directly attributable to BMP expression. This raises the intriguing question of what is regulating Id1 expression and what roles Id1 may be playing in this key embryonic structure.

LAY SUMMARY

Complex organisms such as mammals develop from a single egg cell fertilised by a single sperm cell. Over the course of the pregnancy, this fertilised egg cell divides multiple times and gives rise to all cells which form the foetus at birth. This transition from a single cell with the capability to form all cells of the body to multiple specialised cells such as neurones and muscle cells is governed by many genes and proteins, and the roles of many of these are poorly understood.

In order to study the role of one such protein called Id1, I made use of mouse embryonic stem (ES) cells. These cells are derived from a mouse embryo at a stage in which all of its cells can give rise to all tissues of the adult mouse, and they can be used to study some of the events that happen in the mouse embryo without having to sacrifice mice. I generated ES cells that produce a yellow fluorescent protein whenever Id1 is normally active, and monitored their fluorescence to study how various molecules affect the presence or absence of Id1 in these cells. I also generated ES cells that produce more Id1 protein than normal cells, and discovered that they tend to produce more mesoderm (the tissue that will give rise to muscle and bone) than neural tissue. I also studied where Id1 is present in mouse embryos around the time the cells that form neural tissue, mesoderm and endoderm (the tissue that will give rise to gut and other internal organs) arise, and observed that it is found in the future heart and gut.

This work helps us to understand the roles of a protein at the time when cells lose their ability to make all tissues of adult organisms. This study is therefore useful to paint a better picture of developmental events that, if malfunctioning, could result in foetal abnormalities. Furthermore, it can help improve culture methods to make specific tissue types from ES cells for therapeutic purposes.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFU	Arbitrary fluorescent unit
AVE	Anterior visceral endoderm
(b)HLH	(Basic) helix-loop-helix
BMP	Bone morphogenic protein
bp	Base pair
BSA	Bovine serum albumin
cDNA	Copy DNA
ChIP-seq	Chromatin immunoprecipitation sequencing
DAPI	4’,6-diamidino-2-phenylindole
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DVE	Distal visceral endoderm
EB	Embryoid body
EC	Embryonal carcinoma
EMT	Epithelial-to-mesenchymal transition
EpiSC	Epiblast stem cell
ES	Embryonic stem
EV	Empty vector
FCS	Foetal calf serum
Fgf	Fibroblast growth factor
GFP	Green fluorescent protein
ICM	Inner cell mass

Igf	Insulin-like growth factor
iPS	Induced pluripotent stem
IRES	Internal ribosome entry site
kb	Kilobase
kDa	Kilodalton
KSR	Knockout serum replacement
LB	Lysogeny broth
LIF	Leukaemia inhibitory factor
LUT	Lookup table
mRNA	Messenger RNA
NES	Nuclear export signal
NLS	Nuclear localisation signal
pA / polyA	Polyadenylation signal sequence
PBS	Phosphate buffered saline
PFA	Paraformaldehyde
(qRT-) PCR	(Quantitative real-time) polymerase chain reaction
RFP	Red fluorescent protein
RNA	Ribonucleic acid
SFEB	Serum-free embryoid body
TGFβ	Transforming growth factor β
UPL	Universal ProbeLibrary
UTR	Untranslated region
wt	Wild-type

CHAPTER 1

Introduction

1.1 Embryonic stem cells

1.1.1 The wonders and terrors of embryonic stem cell biology

Mouse embryonic stem (ES) cells are a type of cultured cell that has captured the public imagination for many years because of their two defining properties: the capability to generate at least one identical copy of themselves every time they divide ("self-renewal") and the ability to differentiate into all cell types present in the mouse embryo and adult ("pluripotency"). Their capacity of integrating with a live embryo, and of contributing to the tissues of the adult mouse developed from that embryo, could have been scripted by a science fiction writer, and has brought about both scientific elation and widespread prejudice, both of the cell type and of the researchers using them. Their amenity for genetic manipulation, which resulted in the 2007 Nobel Prize in Physiology or Medicine, has provided a platform for the study of the functions of molecules, complexes and pathways, for the modelling of diseases and the screening of novel drugs, yet the image of the mad scientist playing God is still a frighteningly common one amongst the public. These fears, caused by a lack of comprehension partially attributable to the media's sensationalistic approach to science, and partially attributable to the inability or disinterest of the scientific community to divulge their discoveries in a manner comprehensible to the general public, have important implications not only in the perception of research by the public, but in terms of scientific investment by governments. The Sherley v. Sebelius US Supreme Court case could have had potentially devastating effects on the American embryonic stem cell community had the judges not declined to hear the case (Wadman 2013), and in Italy projects involving the use of human ES cells have been specifically excluded from a public funding call (Cattaneo *et al.* 2010).

Embryonic stem cells have been and still are an incredibly powerful tool for the study of embryonic development and for the discovery of novel approaches to regenerative medicine. The recent drive for public engagement promoted by both research councils and universities alike will hopefully allow the newer generations to approach embryonic stem cell biology without pre-conceived prejudices.

1.1.2 Historical perspective

Scientific discoveries are rarely the result of an isolated event of momentary genius, but are rather the consequence of a spark of intuition rooted in the body of knowledge generated by years of research.

The derivation of embryonic stem cells followed this path, and stemmed from the fields of embryology, developmental biology and cancer biology. A particularly relevant set of studies was that performed on teratomas and embryonal carcinoma cells.

1.1.2.1 Embryonal carcinoma cells

Teratoma is a type of germ cell tumour which spontaneously arises in some mouse strains, such as 129, and has the property of containing cells from one or more of all three germ layers (ectoderm, mesoderm and endoderm) (Stevens & Little 1954). In a classic experiment, Kleinsmith and Pierce (1964) demonstrated the potential of a subset of individual cells from these tumours, named embryonal carcinoma (EC) cells, to give rise to a new tumour containing derivatives of all germ layers, as well as extraembryonic tissues, when grafted into recipient mice. This experiment illustrated some key concepts that would later come to the fore in ES cell biology: the clonogenicity of single cells, their pluripotency, and the heterogeneity and

potential lineage bias of single pluripotent EC cells, which gave rise to different proportions of cells from each germ layer in different tumours.

EC cell lines derived from different tumours were adapted to grow in culture as adherent monolayers in basal medium supplemented with serum, in the presence or absence of a layer of feeder cells (Kahan & Ephrussi 1970, Rosenthal *et al.* 1970). Their morphology is strikingly resemblant to that of ES cells (Jacob 1978, Rosenthal *et al.* 1970). Importantly, injection of tumour-derived EC cells into the blastocyst resulted in their contribution to normal mouse development, as opposed to tumour formation (Brinster 1974, Illmensee & Mintz 1976, Mintz & Illmensee 1975, Papaioannu *et al.* 1975), suggesting that these cancer cells can, under the appropriate conditions, behave as pluripotent cells of the early embryo. Mice with tissues derived from cells of separate genetic identity (in this case their own embryonic cells and EC cells) are known as chimaeras, from the ancient Greek mythological creature Chimaera, the body of which was a mixture of those of a lion, a goat and a snake. Martin Evans and colleagues also demonstrated that cultured EC cell lines could result in the generation of chimaeras upon blastocyst injection (Papaioannu *et al.* 1978), providing a crucial link between an *in vitro* established cell line and *in vivo* development.

1.1.2.2 The isolation of embryonic stem cells

The experience gained by two researchers in the field of EC cells would provide invaluable for the isolation of embryonic stem cells. Gail Martin worked in the laboratory of Martin Evans during the 1970s, where the two researchers experimented on EC cell biology (Martin & Evans 1974, Martin & Evans 1975). Martin moved to San Francisco to set up her own laboratory, whilst in Cambridge a talented embryologist named Matthew Kaufman started collaborating with Evans. The common interests of Martin and Evans and their experimental prowess allowed

both research units to simultaneously isolate pluripotent cells from mouse blastocysts. Martin named these cells embryonic stem cells and demonstrated their ability to differentiate into cells of all germ layers *in vitro* (Evans & Kaufman 1981, Martin 1981). The ability of ES cells to generate chimaeric embryos was addressed soon after, and ES cells were shown to contribute both to chimaeric embryos and to their offspring (Bradley *et al.* 1984, Kaufman *et al.* 1984).

1.1.2.3 Transgenesis

The potential of ES cells for the generation of transgenic mice was of extreme promise, as noted by Roberston and colleagues upon their confirmation of germline contribution of ES cells (Bradley *et al.* 1984). The following year the first report of transgenic ES cells was published. The researchers introduced a cosmid containing the human type II-collagen gene into ES cells by random integration, and assessed its expression in chimaeric mice (Lovell-Badge *et al.* 1985). In doing so, they demonstrated the ease of genetic modification of ES cells and their potential use for humanised mouse models. Thomas & Capecchi (1986) performed the first gene targeting experiment in ES cells, whereby a sequence of DNA in the genome of the ES cells was replaced by another sequence by means of homologous recombination between flanking regions.

ES cells genetically modified with random or targeted transgene integrations were then shown to successfully contribute to live animals and their offspring following blastocyst injection (Hooper *et al.* 1987, Robertson *et al.* 1986), proving the usefulness of *in vitro* transgenesis in ES cells for the *in vivo* study of the phenotype of transgenic mice.

1.1.3 Pre-implantation development of the mouse

Embryonic stem cells were isolated from the inner cell mass of the mouse blastocyst, and share the same capability of forming all tissue types of the adult mouse. It is therefore important to understand how the mouse embryo develops before implantation if we are to understand how ES cells are regulated.

1.1.3.1 Fertilisation and development to the blastocyst stage

When the sperm fertilises the oocyte the zygote is formed. Unlike ES cells, this cell is totipotent, which means it can form all cells of the adult mouse as well as all the conceptus-derived extraembryonic structures required for supporting the development of the embryo. The fertilised zygote undergoes a series of symmetric cell divisions during which the whole embryo does not increase in volume and the volume of the cells halves at every division event (Nichols & Smith 2012, Rossant & Tam 2009, Wennekamp *et al.* 2013). The individual cells, termed blastomeres, of these early embryos are still totipotent and can give rise to adult mice under the appropriate conditions (Tarkowski 1959, Tarkowski & Wróblewska 1967, Tarkowski *et al.* 2010). Once the embryo reaches the 8-16 cell stage, its cells form tight junctions and the structure undergoes a process known as compaction. This embryonic stage is known as morula (named after the Latin for mulberry, due to its resemblance to the fruit) and occurs approximately 2.5 days after fertilisation (embryonic day 2.5, noted as E2.5). At E3.5 the embryo reaches the 32-64 cell stage and forms the blastocyst, a structure characterised by an outer layer of epithelial cells, the trophectoderm, an inner cell mass (ICM), and a fluid filled cavity, known as the blastocoel cavity. During the course of a further day of development, the blastocyst expands and the cells in the ICM segregate into two cell types, the epiblast (sometimes referred to as primitive ectoderm), which will form the embryo proper, and the primitive endoderm (sometimes referred to as hypoblast), an epithelial layer

of cells that underlies the epiblast and will provide patterning signals to the embryo during development (Figure 1.1) (Nichols & Smith 2012, Rossant & Tam 2009, Wennekamp *et al.* 2013).

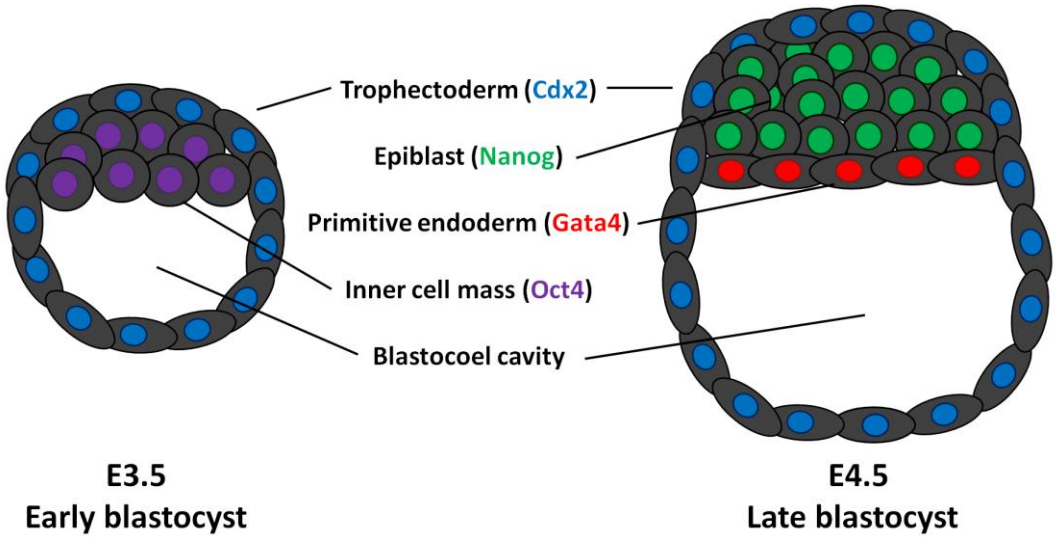


Figure 1.1 – Structure of the mouse blastocyst

The blastocyst forms approximately 3.5 days after fertilisation, and contains two cell types, an external layer of trophectoderm, expressing the transcription factor Cdx2 (blue), and an inner cell mass, expressing the pluripotency factor Oct4 (purple). The blastocyst undergoes further cell division and expansion, and the inner cell mass gives rise to two structures: the epiblast, which will form the embryo proper and expresses Nanog (green), and the primitive endoderm, which underlies the epiblast and expresses Gata4 (red).

1.1.3.2 Emergence of the pluripotent epiblast

When the blastocyst is formed, the cells composing it have already made a first crucial decision in their development: to differentiate into trophectoderm or to remain in the pluripotent ICM. Trophectodermal cells are now committed to

extraembryonic lineages, but ICM cells are still capable of giving rise to both embryonic lineages and extraembryonic primitive endoderm (Nichols & Smith 2012, Rossant & Tam 2009, Wennekamp *et al.* 2013). All cells in the ICM express Oct4, a transcription factor also known as Pou5f1 and Oct3/4, and its expression is essential for the development of the ICM: embryos that lack it can only form trophectoderm and cannot develop further (Nichols *et al.* 1998, Palmieri *et al.* 1994). Its negative relationship with the trophoblast determinant Cdx2 appears to be critical for the segregation of trophectoderm and ICM (Niwa *et al.* 2005). Whilst required for the establishment of the pluripotent ICM, Oct4 expression is not limited to the pluripotent epiblast at E4.5, and is also detected in the primitive endoderm layer (Dietrich & Hiiragi 2007, Grabarek *et al.* 2011, Nichols *et al.* 2009, Silva *et al.* 2009).

The epiblast and primitive endoderm can be distinguished by the respective expression of Nanog and Gata4 (Nichols *et al.*2009, Silva *et al.* 2009, Plusa *et al.* 2008). Nanog is a critical factor for the formation of the ICM, as *Nanog*-null embryos generate a non-viable unspecified ICM-like structure (Mitsui *et al.* 2003, Frankenberg *et al.* 2011, Silva *et al.* 2009). Cells in the ICM express Nanog and Gata4 in a mutually exclusive salt-and-pepper pattern, which allows to predict which cells will contribute to the epiblast and which to the primitive endoderm. Other factors expressed in the primitive endoderm, such as Gata6 and Pdgfra, are expressed in the embryo prior to Gata4, and can be co-expressed with Nanog in the early blastocyst. During development from the early to the late blastocyst stage, through a combination of cell sorting and apoptosis, these Gata6- and Pdgfra-expressing cells localise to the outside of the ICM and form the primitive endoderm (Plusa *et al.* 2008).

The segregation of the epiblast and primitive endoderm is mediated by Fgf signalling. Blastocysts deficient for Fgf signalling components or cultured in Fgf signalling inhibitors display an expanded epiblast and a lack of primitive endoderm, whereas culture of blastocysts in the presence of exogenous Fgf4 causes the

downregulation of Nanog and homogeneous expression of Gata6 (Chazaud *et al.* 2006, Nichols *et al.* 2009, Yamanaka *et al.* 2010). Interestingly, it is the factors that promote pluripotency that drive the Fgf-dependent induction of primitive endoderm. *Nanog*-null embryos fail to generate primitive endoderm due to their lack of *Fgf4* expression, and addition of exogenous Fgf4 to the cultures rescues the formation of primitive endoderm (Frankenberg *et al.* 2011). Sox2 is another key pluripotency transcription factor. It is expressed throughout early development and localises to the ICM and later to the epiblast region of the blastocyst. *Sox2*-null embryos die around implantation due to defects in the epiblast (Avilion *et al.* 2003, Bedzhov *et al.* 2012, Zhang *et al.* 2013). Oct4 and Sox2 have been found to co-operate in the direct activation of Fgf4 transcription in the ICM, and Oct4 is required both upstream and downstream of Fgf4 activation for the transcription of primitive endoderm markers (Fraidenraich *et al.* 1998, Frum *et al.* 2013). Its ability to direct primitive endoderm gene expression is achieved by switching its transcriptional activation partner from Sox2 to the endoderm-specific factor Sox17 (Aksoy *et al.* 2013).

1.1.3.3 Embryonic stem cells as a model for pre-implantation development

The development of the pluripotent ICM and epiblast, and the activity of pluripotency and pro-differentiation factors within these structures, are complex processes which are not yet fully understood. Embryonic stem cells represent a directly relevant *in vitro* counterpart to these *in vivo* pluripotent structures, and can be used as a model to study these events. They provide a number of advantages compared to direct experimentation on pre-implantation embryos: they can be maintained in a self-renewing state indefinitely, allowing the study of molecular relationships that may be hard to uncover in the rapidly developing mouse embryo; they can be cultured in bulk, providing a wealth of biological material; they can be maintained in chemically defined culture medium, allowing to dissect the roles of extracellular signalling molecules in their maintenance and differentiation; they

allow rapid and straightforward transgenesis experiments to address the roles of specific genes in pluripotency and differentiation; they can be visualised during experimentation with ease in their "natural" environment for a prolonged period of time.

They therefore represent a tremendously powerful tool for the study of developmental biology. It is however imperative that we relate ES-cell based findings to our knowledge of *in vivo* development, in order to formulate hypotheses that are plausible in the light of both our observations and of the events that regulate mouse development and that constrain developmental decisions in time and space.

1.1.4 Extracellular signalling molecules involved in the preservation of embryonic stem cell pluripotency

1.1.4.1 Leukaemia inhibitory factor

ES cells were isolated by culturing them in a basal medium supplemented with serum on a mitotically-inactivated fibroblast feeder layer (Evans & Kaufman 1981, Martin 1981). The feeder cells were required not only for adhesion, but because of their ability to secrete an uncharacterised differentiation-inhibiting activity entity (Smith & Hooper 1987), which could be found in Buffalo rat liver cell-conditioned medium, a culture medium which allowed for the dispensation of the feeder layer (Hooper *et al.* 1987). This differentiation-inhibiting activity entity was eventually identified as leukaemia inhibitory factor (LIF) (Smith *et al.* 1988, Williams *et al.* 1988). This molecule is a cytokine that binds to a heteromeric receptor formed by Lifr and gp130 (Davis *et al.* 1993, Gearing *et al.* 1992). Upon receptor binding, LIF stimulates interaction of the receptors with Janus kinases, which phosphorylate the receptors (Stahl *et al.*1994), allowing their interaction with a number of molecules. Amongst these molecules are Stat proteins, which bind the phosphorylated receptors and are in turn phosphorylated and activated, and SHP-2 (Stahl *et al.* 1995, Wen *et al.* 1995).

SHP-2 binding results in the activation of the mitogen-activated protein kinase intracellular cascade, mediated by the interaction of SHP-2 with Grb2 and Gab1, proteins which can both induce mitogen-activated protein kinase activation (Li *et al* 1994, Takahashi-Tezuka *et al.* 1998).

LIF is required for the implantation of the blastocyst to the uterine wall. The blastocysts of female mice lacking a *Lif* gene cannot implant, and enter diapause (Stewart *et al.* 1992). Diapause is a state of delayed implantation in which the blastocysts arrest development, whilst retaining the potential to reprise normal development under the appropriate conditions (Lopes *et al.* 2004). The activity of LIF is also essential following diapause: embryos lacking the *Gp130* gene, encoding the heteromerisation partner of the LIF receptor, survive diapause but are unable to reprise development, due to a loss of epiblast cells (Nichols *et al.* 2001).

LIF signalling is thought to maintain ES cell pluripotency primarily by the activation of the transcription factor Stat3. Lack of Stat3 activity results in ES cell differentiation (Burdon *et al.* 1999a, Niwa *et al.* 1998, Ying *et al.* 2008). Targets of this factor include the pluripotency factor *Klf4* and *Tcfp2l1*, a recently identified target which can sustain pluripotency in *Stat3*-null ES cells in LIF and serum culture (Martello *et al.* 2013). Activation of the mitogen-activated protein kinase signalling cascade negatively affects ES cell self-renewal and is required for ES cell differentiation (Burdon *et al.* 1999b, Kunath *et al.* 2007, Stavridis *et al.* 2007, see section 1.1.4.4). It therefore appears that LIF has a dual effect on ES cells: to maintain them in an undifferentiated state while at the same time generating intracellular signals that will allow them to differentiate.

1.1.4.2 Serum and bone morphogenic protein

The second component of ES cell culture medium required for the maintenance of pluripotency is foetal calf serum (FCS). ES cells cultured in the presence of LIF but

absence of FCS are incapable of remaining undifferentiated, and form neural ectoderm, albeit at a lower efficiency than cells cultured in the absence of both LIF and FCS (Ying *et al.* 2003a,b).

The components of FCS are undefined, and may vary from one batch to another. The differentiation-inhibiting activity entity present in FCS was identified in 2003 as bone morphogenic protein 4 (BMP4) (Ying *et al.* 2003a). ES cells cultured in the presence of LIF and BMP4 can self-renew and be propagated in culture, retain multilineage differentiation potency and can contribute to chimaeras. BMP4 is not sufficient to preserve pluripotency in the absence of LIF: ES cells cultured in BMP4 alone differentiate into non-neural fates (Ying *et al.* 2003a).

BMP's are a subfamily of proteins that is part of the transforming growth factor beta (TGFβ) superfamily of signalling molecules. These are secreted extracellular proteins that bind as dimers to heterotetrameric transmembrane serine/threonine protein kinase receptors. The heterotetrameric receptors are formed by two Type I and two Type II receptors, which recognise specific subsets of TGFβ superfamily ligands. The receptors associated with BMP4 recognition are Bmpr1a, Bmpr1b (Type I), Bmpr2, Acvr2a, Acvr2b (Type II) (Allendorph *et al.* 2006, Greenwald *et al.* 2003, Griffith *et al.* 1996, Kirsch *et al.* 2000, Koenig *et al.* 1994, Nagaso *et al.* 1999, Paralkar *et al.* 1991, Piek *et al.* 1999, Rosenzweig *et al.* 1995, ten Dijke *et al.* 1994, Weber *et al.* 2007). Upon binding of BMP (or other TGFβ ligands), the kinase domain of the Type II receptor phosphorylates the Type I receptor, activating its kinase activity (Chen & Weinberg 1995, Wrana *et al.* 1994). Activated receptors can then mediate the phosphorylation of specific Smad proteins, recognised via the interaction of two short amino acid sequences: the L45 loop on the Type I receptor and the L3 loop on the Smad protein (Feng & Derynck 1997, Lo *et al.* 1998). BMP4 drives the phosphorylation of Smad1, Smad5 and Smad8. Phosphorylated Smad's can form a complex with Smad4 and drive the transcription of target genes (Chen *et al.* 1997b, Kretzschmar *et al.* 1997, Lagna *et al.* 1996, Nishimura *et al.* 1998, Suzuki *et al.* 1997).

Id genes are direct targets of these active BMP-induced Smad complexes (Hollnagel *et al.* 1999, Katagiri *et al.* 2002, Korchynskyi & ten Dijke 2002, López-Rovira *et al.* 2002, Ogata *et al.* 1993), and Id protein overexpression can substitute for BMP4 in the maintenance of ES cell pluripotency (Ying *et al.* 2003a). Id proteins can therefore preserve the capability of ES cells to self-renew in a Smad-independent manner, and are thus the key mediators of pluripotency in LIF and serum culture in association with Stat3.

1.1.4.3 Activin and Nodal

Activin and Nodal are members of a separate subfamily of TGFβ signalling molecules. They also signal through specific TGFβ receptors: Acvr1b, Acvr1c (Type I), Acvr2a, Acvr2b (Type II); a co-receptor, Cripto, is also involved in Nodal signal transduction (Cheng *et al.* 2003, Reissman *et al.* 2001, Willis *et al.* 1996). Activin/Nodal signals lead to the phosphorylation of Smad2 and Smad3, which upon binding to Smad4 form transcriptionally active complexes and modulate the transcription of target genes (Kretzschmar *et al.* 1997, Lagna *et al.* 1996, Nakao *et al.* 1997, Wu *et al.* 1997).

Nodal is an important signal in embryo patterning and germ layer specification, and *Nodal*-null embryos die shortly after implantation (Mesnard *et al.* 2006). ES cells secrete Activin/Nodal agonists and these signals positively affect cell proliferation (Ogawa *et al.* 2007). However, Activin/Nodal signalling negatively regulates the transcription of the pluripotency-inducing BMP targets in ES cells (Galvin *et al.* 2010), and inhibition of Activin/Nodal signalling with the small molecule inhibitor SB431542 improves the efficiency of reprogramming mouse embryonic fibroblasts to induced pluripotent stem cells (Maherali & Hochedlinger 2009). Furthermore, Activin/Nodal-driven phosphorylation of Smad2/3 have been shown to induce activation of the promoter of pro-differentiation genes *Gsc* and *Mixl1* via the

formation of active Smad2/3/4 transcriptional complexes and through Smad4-independent displacement of a chromatin compaction factor (Xi *et al.* 2011).

The mechanisms by which Activin/Nodal signalling may inhibit BMP target transcription are various. Smad2/3 can directly bind the promoters of some BMP targets and repress their transcription (Kang *et al.* 2003, Lee *et al.* 2011). There may be competition for Smad4 binding between Smad2/3 and Smad1/5/8, although Smad4 overexpression does not rescue Activin/Nodal-driven BMP inhibition (Galvin *et al.* 2010). Smad2/3 can also drive the transcription of Smad6 and Smad7 (Afrakhte *et al.* 1998, Brodin *et al.* 2000, Hua *et al.* 2000). These are inhibitory Smad proteins, which can inhibit transcription driven by Smad1/5/8 (Smad6, Smad7) and by Smad2/3 (Smad7) (Hata *et al.* 1998, Imamura *et al.* 1997, Ogawa *et al.* 2007, Souchelnytskyi *et al.* 1998). Thus, Activin/Nodal-driven transcription of Smad6, an inhibitory Smad specific to BMP-responsive Smad proteins, could dampen BMP signalling. However, BMP can also induce transcription of Smad6 (Afrakhte *et al.* 1998, Takase *et al.* 1998), suggesting the factor may be expressed in ES cells regardless of Activin/Nodal activity. To complicate matters further, Smad2 and Smad3 can both co-operate and antagonise each other in the promotion of the transcription of target genes (Míguez *et al.* 2013).

The crosstalk between BMP and Activin/Nodal signals is likely to be required for the generation of a fine-tuned balance between different transcriptional programmes, critical for both the maintenance of pluripotency and the proliferation of ES cells.

1.1.4.4 Fibroblast growth factor and the mitogen-activated protein kinase pathway

Fibroblast growth factor (Fgf) molecules are key signalling components throughout the development of the mouse embryo. They act primarily by stimulating the mitogen-activated protein kinase pathway, in which sequential activation of Fgf receptors, Grb2, Ras, Raf, Mek and Erk lead to the transcription of target genes (Villegas *et al.* 2010).

The knockout of Fgf receptors leads to developmental abnormalities (Colvin *et al.* 1996, Yamaguchi *et al.* 1994), and knockout of *Fgf4* leads to lethality shortly after implantation (Feldman *et al.* 1995). *Fgf4* is first expressed at the 4-cell stage and is expressed in the inner cell mass of the blastocyst (Niswander & Martin 1992, Rappolee *et al.* 1994). It is also secreted by ES cells (Ma *et al.* 1992). Because of its *in vivo* requirement, and because of the roles of Fgf's in promoting cell proliferation in other systems (Thisse & Thisse 2005), it could be expected for Fgf4 to promote ES cell self-renewal and proliferation. However, Fgf signalling appears to negatively affect ES cell self-renewal, and its inhibition promotes and homogenises the expression of transcription factors associated with pluripotency (Burdon *et al.* 1999b, Hamazaki *et al.* 2006, Kunath *et al.* 2007, Wilder *et al.* 1997, Ying *et al.* 2008). The loss of Fgf/Erk activity also has a profound effect on the blastocyst: cells of the inner cell mass become uniformly positive for the epiblast determinant Nanog and fail to correctly segregate primitive endoderm (Chazaud *et al.* 2006, Nichols *et al.* 2009). This is the likely cause of the lethal phenotype seen in *Fgf4*-null embryos.

The pro-differentiation effect of Fgf/Erk holds true to the extent that *Fgf4*-null ES cells retain expression of the pluripotency factor Oct4 after 6 days of culture in basal medium devoid of LIF and BMP, and inhibition of Mek prevents the loss of Oct4 from ES cells undergoing embryoid body differentiation (Burdon *et al.* 1999b, Kunath *et al.* 2007). Fgf signalling is therefore required to exit the pluripotent state, as confirmed by the rescue of differentiation seen following the addition of Fgf4 to *Fgf4*-null ES cells or following the removal of a Mek inhibitor from *Sox1-GFP* reporter ES cells for the first 24 hours of differentiation (Kunath *et al.* 2007, Stavridis *et al.* 2007). This effect of Fgf4 on differentiation was thought to be mediated by the mitogen-activated protein kinases Erk1 and Erk2. *Erk2*-null ES cells were in fact shown to be resistant to neural and mesodermal differentiation (Kunath *et al.* 2007). Rederivation of *Erk2*-null ES cells on a different genetic background, however, allowed Kunath and colleagues to demonstrate that the lack of *Erk2* resulted in a more homogeneous expression of pluripotency factors under self-

renewing conditions, but that it did not affect the differentiation competence of the cells (Hamilton *et al.* 2013). This suggests that either Fgf4 is only required to exit pluripotency on certain genetic backgrounds, or that Erk2 is not the main mediator of Fgf signalling in ES cell differentiation.

1.1.4.5 Wnt

The Wnt pathway is a complex signalling cascade involving many Wnt proteins (19 for most mammals) and diverse intracellular responses. In canonical Wnt signalling, the effects of Wnt are mediated by modulating the transcriptional activity of β-catenin. In the absence of Wnt, cytoplasmic β-catenin is associated with a destruction complex consisting of Axin, Dvl, APC, CK1 and GSK3β. CK1 and GSK3β are kinases that phosphorylate β-catenin, causing it to be recognised by the E3 ubiquitin ligase β-TrCP, ubiquitinated and degraded by the proteasome. Upon Wnt binding to a heteromeric receptor formed by Frizzled and Lrp5/6, the degradation of β-catenin is inhibited and the protein can translocate to the nucleus and drive transcription of target genes. According to the established model, Wnt association with the receptor results in Axin, Dvl, CK1 and GSK3β association with the Wnt/receptor complex, resulting in the inability of the destruction complex to phosphorylate cytoplasmic β-catenin which is no longer targeted for degradation (Clevers & Nusse 2012, Kim *et al.* 2013, Wu & Pan 2009). A recent publication suggests that the entire destruction complex, including APC and phosphorylated β-catenin, associates with Frizzled/Lrp5/6 upon Wnt stimulation, but that phosphorylated β-catenin can no longer be ubiquitinated by β-TrCP and is no longer released from the destruction complex. It therefore saturates the destruction complex preventing the degradation of the remaining cytoplasmic β-catenin (Li *et al.* 2012).

Regulation of β-catenin function plays an important role in ES cell pluripotency. In 2002, Fodde and colleagues observed that lack of β-catenin phosphorylation in ES cells, obtained by truncation of APC proteins or with the use of mutant alleles of β-

catenin lacking the normally phosphorylated residues, resulted in severe differentiation impairment (Kielman *et al.* 2002). Inhibition of GSK3β with a small molecule inhibitor was then demonstrated to allow nuclear relocalisation of β-catenin and promote expression of the pluripotency factor *Rex1* (Sato *et al.* 2004). Association of β-catenin with the pluripotency factor Klf4 regulates the expression of telomerase, ensuring long-term self-renewal capability (Hoffmeyer *et al.* 2012). Stimulation of ES cells with Wnt3, Wnt3a, Wnt5a and Wnt6 was also shown to promote pluripotency (Hao *et al.* 2006, Ogawa *et al.* 2006), and this positive effect on self-renewal could be stimulated further with a combination of Wnt and LIF or Fgf inhibition (Singla *et al.* 2006).

1.1.4.6 Maintenance of ES cell pluripotency by combinatorial stimulation or inhibition of signalling pathways

All of the signalling pathways described above are capable of positively or negatively affecting the self-renewal ability of embryonic stem cells, but stimulation or inhibition of no single pathway can maintain ES cells in a pluripotent state. Culture of ES cells in LIF alone generates neural ectoderm (Ying *et al.* 2003a), culture in BMP4 alone generates mesodermal and epithelial cells (Johansson & Wiles 1995, Park *et al.* 2004, Ying *et al.* 2003a), Activin/Nodal stimulation alone drives mesendoderm formation (Fei *et al.* 2010, Hansson *et al.* 2009, Pfendler *et al.* 2005), inhibition of Activin/Nodal promotes neural differentiation (Watanabe *et al.* 2005), stimulation with Fgf accelerates neural differentiation (Chen *et al.* 2010). Inhibition of Fgf signalling alone can maintain self-renewal briefly, but the cells exhibit proliferation defects and eventually differentiate (Meek *et al.* 2013, Ying *et al.* 2008). Stimulation of Wnt signalling delays the exit from pluripotency but does not ultimately prevent it (Wray *et al.* 2011, Yi *et al.* 2011); Wnt also promotes mesodermal and inhibits neural differentiation, and Wnt inhibition favours neural specification (Arnold *et al.* 2000, Hansson *et al.* 2009, Watanabe *et al.* 2005).

It appears therefore that ES cell self-renewal can be promoted by multiple signalling events, but that a combination of these is required for long-term maintenance of pluripotency. The sufficiency of LIF and BMP4 for this purpose was the first to be characterised (Ying *et al.* 2003a), but other combinations of factors have successfully been used since.

The observations that the inhibition of Fgf signalling and of β-catenin phosphorylation promoted pluripotency led Ying *et al.* (2008) to develop a self-renewal medium containing only two small molecule inhibitors of Mek (PD0352901) and GSK3β (CHIR99021), a medium they termed "2i". Culture of ES cells in this medium resulted in homogenous gene expression of pluripotency factors and allowed derivation of ES cells from refractory mouse strains; these effects were potentiated by the addition of LIF to the medium. 2i+LIF medium was also used to successfully derive ES cells from rat embryos, a procedure which had previously been impossible as rat ES cells differentiate in LIF and serum (Buehr *et al.* 2008, Li *et al.* 2008, Nichols *et al.* 2009, Ying *et al.* 2008).

Smith and colleagues therefore formulated the hypothesis that these culture conditions allowed ES cells to remain in a "ground state", and suggested self-renewal was the default option of ES cells, providing these were shielded from autocrine pro-differentiation signals such as Fgf (Ying *et al.* 2008). This constitutes an attractive hypothesis, for it introduces the idea of ES cell culture conditions for the promotion of a default pluripotent state. The observation that ES cells will differentiate into neural ectoderm in the absence of any stimulus or the presence of LIF alone had previously led to the use of the anti-neural factor BMP4 as a differentiation-inhibiting entity (Ying *et al.* 2003a,b), and had resulted in the formulation of the concept that ES cells will differentiate by default, and that the maintenance of pluripotency must be achieved via the inhibition of differentiation.

The "default self-renewal" hypothesis cannot however hold true. The inhibition of GSK3β mimics the effects of Wnt stimulation, and therefore suggests that exogenous stimulation of this pathway or of its downstream effectors is required for the

maintenance of pluripotency. The shielding of cells from autocrine signals with Fgf signalling inhibitors is not sufficient to preserve long-term self-renewal capability (Meek *et al.* 2013, Ying *et al.* 2008). It is intriguing that rat ES cells do exhibit long-term self renewal in basal medium supplemented solely with a Mek inhibitor ("1i"), and suggestive of the fact that pluripotency states can vary between rodent species (Meek *et al.* 2013).

Other combinations of factors have been adopted since 2008 for the long-term maintenance of mouse ES cells. LIF and PD0352901, as well as LIF and Wnt3a/CHIR99021, have been shown to sustain pluripotency in the absence of exogenous BMP4 (ten Berge *et al.* 2011, Wray *et al.* 2009). Two recent publications have demonstrated that a homogenous pluripotent state reminiscent of 2i culture can be generated via small molecule-mediated inhibition of Activin/Nodal signalling receptors (SB431542, inhibitor of Acvr1b, Acvr1c, Tgfbr1, Inman *et al.* 2002) and of Mek (PD0352901) (Hassani *et al.* 2012, Hassani *et al.* 2013). This medium, termed "R2i", also allowed derivation of ES cells from recalcitrant mouse strains. The achievement of ground state pluripotency via inhibition of Fgf and Activin/Nodal signalling would appear to lend further weight to the "default self-renewal" hypothesis. However, it appears that R2i medium maintains ES cell pluripotency primarily by derepressing phospho-Smad2-inhibited BMP targets, and addition of BMP receptor inhibitors to R2i medium causes a collapse of pluripotency (Hassani *et al.* 2012, Hassani *et al.* 2013). This suggests that under these culture conditions BMP4 autocrine signalling is paramount to the maintenance of pluripotency, and that complete shielding of ES cells from autocrine signals would not result in the maintenance of the pluripotent state.

In summary, it appears the maintenance of ES cell pluripotency can be achieved by means of activating or inhibiting a variety of signalling pathways, but that no signalling pathway on its own is sufficient for long-term ES cell self-renewal. The adaptability of ES cells to various culture conditions is likely to reflect the necessity

for cells in the inner cell mass of the blastocyst to maintain pluripotency in a three-dimensional environment whilst being exposed to a variety of signals, which may differ both in type and intensity between neighbouring cells.

1.1.5 Intracellular mediators of ES cell pluripotency

1.1.5.1 Oct4 and Sox2

Oct4 is a transcription factor that is crucial for the establishment of pluripotency in the early embryo (Nichols *et al.* 1998, Palmieri *et al.* 1994, see section 1.1.3.2). It is also expressed in embryonic stem cells (Rosner *et al.* 1990), and it is essential for their self-renewal. Loss of Oct4 expression results in differentiation of ES cells towards the trophectodermal lineage (Niwa *et al.* 2000), an event reminiscent of its antagonism of the trophoblast specifier Cdx2 *in vivo* (Niwa *et al.* 2005).

Sox2 is a transcription factor essential for early mouse development (Avilion *et al.* 2003, see section 1.1.3.2). It is also required for embryonic stem cell pluripotency; its loss leads to the downregulation of *Oct4* and the upregulation of markers of the trophectodermal lineage (Ivanova *et al.* 2006, Misui *et al.* 2007). Oct4 and Sox2 co-operate to drive the transcription of target genes. These targets include the aforementioned *Fgf4*, the pluripotency factors *Nanog* and *Utf1*, as well as *Oct4* and *Sox2* themselves (Chew *et al.* 2005, Nishimoto *et al.* 1999, Rodda *et al.* 2005, Tomioka *et al.* 2002, Yuan *et al.* 1995). It would therefore appear likely that the upregulation of trophectodermal markers in Sox2- or Oct4-depleted ES cells is due to a loss of expression of Oct4/Sox2 target genes required for the maintenance of pluripotency. However, the loss of Sox2 protein is not accompanied by a rapid downregulation of Oct4/Sox2 target genes, the expression of which occurs in synchrony with *Oct4* downregulation. Furthermore, restoring Oct4 levels in Sox2-depleted cells by introduction of a transgene can rescue, at least in part, the expression of Oct4/Sox2 targets, the differentiation phenotype of the cells and the

capability of the cells to contribute to chimaeric embryos (Misui *et al.* 2007). This suggests that Sox2 acts primarily by promoting the expression of Oct4, and that the expression of shared Oct4/Sox2 targets is either dispensable or can be maintained by other Sox family members, as is the case for Sox15, which can bind Oct4/Sox2 target sites with weaker affinity than Sox2 (Maruyama *et al.* 2005).

Whilst Oct4 expression is essential for ES cell pluripotency, and the factor can substitute for the loss of the pluripotency mediator Sox2, overexpression of Oct4 also drives differentiation of ES cells, resulting in the upregulation of markers of endoderm and mesoderm (Niwa *et al.* 2000). Why this is the case is unclear. One possibility is that Oct4 can bind the promoters of both pluripotency factors and pro-differentiation genes, but that at wild-type levels it associates with DNA binding partners that bind preferentially or exclusively at the promoters of genes that stimulate self-renewal. Upon elevation of Oct4 levels, the excess Oct4 molecules unbound by the saturated pro-self-renewal partners could interact with DNA binding proteins capable of binding the promoters of pro-differentiation genes, driving the exit from pluripotency. It is intriguing to think that one such partner could be the Oct4 *in vivo* primitive endoderm-specifying partner Sox17 (Aksoy *et al.* 2013). The possibility that excess Oct4 may promote the exit from pluripotency was given further credit by the observation that *Oct4* heterozygous ES cells exhibit homogeneous expression of the pluripotency factor Nanog and display delayed differentiation kinetics compared to wild-type ES cells. These cells are also capable of self-renewal in the presence of LIF alone, possibly due to increased autocrine Wnt signalling (Karwacki-Neisius *et al.* 2013). Similar observations were also made in induced pluripotent stem cells (Radzisheuskaya *et al.* 2013). Oct4 has also been shown to interact with β-catenin and to localise at the cell membrane in complex with this protein and Cdh1 (commonly referred to as E-cadherin) (Faunes *et al.* 2013, Livigni *et al*. 2013). This suggests a mechanism by which excess Oct4 protein may be buffered to prevent it from driving the expression of pro-differentiation factors. The levels of Oct4 in 2i and 2i+LIF display a striking correlation to the levels of

Nanog, whereas this correlation is less evident in LIF and serum culture (Muñoz Descalzo *et al.* 2012, Muñoz Descalzo *et al.* 2013, see also Figure 3.13). It is therefore tempting to hypothesise that a strict regulation of Oct4 levels in 2i culture, potentially resulting in exclusive interactions with pluripotency determinants, may be responsible for the more homogeneous expression of pluripotency factors in these culture conditions.

1.1.5.2 Nanog

Nanog is a transcription factor indentified by two research groups in 2003. They demonstrated that its overexpression could maintain ES cells in a self-renewing state in basal medium without LIF supplementation, and named the gene after the mythological Celtic Land of Youth, Tír na nÓg (Chambers *et al.* 2003, Mitsui *et al.* 2003). This factor displays a stronger pluripotency-inducing phenotype than Oct4 and Sox2, and does not induce differentiation when overexpressed, but cannot rescue trophectodermal specification in Oct4-null or Sox2-depleted cells (Chambers *et al.* 2003, Ivanova *et al.* 2006). It is required for the specification of the pluripotent epiblast *in vivo* (Mitsui *et al.* 2003, Frankenberg *et al.* 2011, Silva *et al.* 2009, see section 1.1.3.2), but it is dispensable for ES cell self-renewal. *Nanog*-null ES cells can be maintained in culture, retain multilineage differentiation capacity and can contribute to chimaeras. They display lower clonogenicity than wild-type cells and express markers of embryonic and extraembryonic endoderm (Chambers *et al.* 2007, Mitsui *et al.* 2003). These dualities indicate that Nanog is an extremely potent pluripotency determinant, but that it is neither sufficient nor required for ES cell self-renewal and differentiation.

Nanog is expressed in a heterogeneous pattern in ES cells. This heterogeneity was detected both at the protein level (within the Oct4-expressing pluripotent compartment) and at the transcriptional level, making use of transcriptional reporters

replacing the endogenous *Nanog* open reading frame with a marker protein, GFP or βgeo. The cells expressing the highest levels of *Nanog* reporters proved to be less prone to differentiation than the cells expressing lower levels of *Nanog* reporters (Chambers *et al.* 2007, Singh *et al.* 2007). Sorting of the cells expressing high levels of the GFP reporter did not allow the long-term purification of a naïve pluripotent population, but rather led to the restoration of the original distribution of GFP high and low cells after 6 days of culture (Chambers *et al.* 2007). This implies that Nanog heterogeneity is intrinsic to LIF and serum culture, and led Chambers *et al.* (2007) to suggest that fluctuations in the levels of Nanog may offer a time window for the cells to escape the pluripotent state and commit to differentiation when the levels of the transcription factor are low.

Nanog binds DNA as a homodimer (Mullin *et al.* 2008, Wang *et al.* 2008), and directly interacts with a number of pluripotency factors, with which it also shares common targets and neighbouring DNA binding sites. These interactors include Sox2 and may include Oct4, although the data available for a direct Nanog-Oct4 interaction are discordant. Regardless of the veracity of a direct interaction, Nanog, Oct4 and Sox2 share a substantial number of targets, including *Nanog*, *Oct4* and *Sox2* themselves (Chen *et al.* 2008, Costa *et al.* 2013, Gagliardi *et al.* 2013, Kim *et al.* 2008, Liang *et al.* 2008, Wang *et al.* 2006). The roles of these factors in the promotion of self-renewal, their physical interactions and the existence of common transcriptional targets led to the idea that these three proteins act as master regulators of pluripotency, and boost the expression of each other to maintain cells in a pluripotent state. This, however, does not fit with the heterogeneous expression of Nanog observed in cells with similar expression levels of Oct4 and Sox2. Two recent publications have demonstrated an unexpected role for Nanog in the autorepression of its own transcription. Whilst Oct4 and Sox2 promote the transcription of *Nanog*, Nanog binds its enhancer region independently of Oct4 and Sox2 and drives the downregulation of its message. This is not only the case in LIF and serum culture, but also in 2i+LIF medium (Fidalgo *et al.* 2012, Navarro *et al.* 2012).

1.1.5.3 Other transcription factors regulating pluripotency

A number of other transcription factors have been implicated in the regulation of ES cell pluripotency. What follows is a brief description of some of the best characterised pluripotency factors.

Krüppel-like factors (Klf's) are a family of zinc finger proteins involved in many cellular processes. Klf2, Klf4 and Klf5 are expressed in ES cells and have been shown to promote pluripotency. Loss of all three factors drives differentiation of ES cells (Jiang *et al.* 2008). The data obtained for single-Klf loss-of-function experiments are discordant. *Klf4*-null mice survive until birth (Segre *et al.* 1999). *Klf4* knockdown has been reported to drive differentiation (Zhang *et al.* 2010), reduce clonogenicity (Bourillot *et al.* 2009), or have no effect (Jiang *et al.* 2008). Loss of *Klf5* has been shown to generate implantation defects and prevent ES cell derivation (Ema *et al.* 2008), drive ES cell differentiation (Parisi *et al.* 2008), reduce ES cell clonogenicity (Bourillot *et al.* 2009), or have no effect (Jiang *et al.* 2008). *Klf2* knockout ES cells can be derived, and *Klf2*-null mice survive until E12.5 (Kuo *et al.* 1997). *Klf2* knockdown has been reported to have no effect on self-renewal (Jiang *et al.* 2008). This variety of phenotypes may reflect differences in genetic background or in residual protein expression in knockdown experiments, but it is clear that expression of at the very least one out of the three factors is essential for self-renewal. In keeping with the positive role of the factors in the promotion of pluripotency, overexpression of Klf2, Klf4 or Klf5 can sustain LIF-independent self-renewal (to a variable extent) (Ema *et al.* 2008, Hall *et al.* 2009, Li *et al.* 2005, Niwa *et al.* 2009, Parisi *et al.* 2008, Zhang *et al.* 2010). Klf factor expression is induced by pluripotency factors: *Klf2* transcription is driven by Oct4, whereas *Klf4* and *Klf5* expression is driven by LIF/Stat3 (Hall *et al.* 2009, Li *et al.* 2005). Klf factors can in turn promote the transcription of Oct4 and Nanog, share a number of common targets with these factors, and are in some cases indispensable for the expression of shared

target genes, as is the case for the Oct4/Sox2/Klf4 target *Lefty1* (Jiang *et al.* 2008, Nakatake *et al.* 2006, Parisi *et al.* 2008).

Rex1 (also referred to as Zfp42) is a zinc-finger protein, identified as a factor rapidly downregulated at the onset of EC cell differentiation (Hosler *et al.* 1989). Its expression was later confirmed in ES cells (Rogers *et al.* 1991), and it is commonly employed as a marker of the pluripotent state (Okita *et al.* 2007, Toyooka *et al.* 2008, Wray *et al.* 2011). It is expressed heterogeneously in LIF and serum culture, and cells expressing high levels of Rex1 display enriched expression of other pluripotency factors, such as Klf factors, Nanog, Sox2, Tbx3 and Esrrb, suggesting it marks ES cells in a naïve pluripotent state (Toyooka *et al.* 2008). It is therefore surprising that neither its knockout nor its overexpression ostensibly affect ES cell pluripotency and differentiation nor embryonic development until after E17.5. *Rex1*-null mice are viable, although they are found in litters in numbers lower than the expected Mendelian rates (Masui *et al.* 2007, Scotland *et al.* 2009). It appears that its primary role is that of promoting the reactivation of the inactive X chromosome by promoting the transcription of *Tsix* and inhibiting that of *Xist* (Gontan *et al.* 2012, Navarro *et al.* 2010). Rex1 is also involved in the suppression of endogenous retroviral expression, and may thus aid in the preservation of the transcriptional identity of pluripotent cells (Guallar *et al.* 2012). It is therefore possible that a combination of inappropriate X inactivation and unsuppressed retroviral expression may lead to developmental defects resulting in a lower number of births of *Rex1*-null mice than expected.

Esrrb is an orphan nuclear receptor protein. It is required for placental development, and its knockout results in embryonic lethality between E9.5-E10.5 (Luo *et al.* 1997). It was first implicated in the regulation of ES cell pluripotency in 2006, when it was demonstrated that its knockdown compromised self-renewal capacity, that its expression was reduced upon *Nanog* and *Oct4* knockdown and induced by *Nanog* overexpression (Ivanova *et al.* 2006, Loh *et al.* 2006). Esrrb can interact with Nanog

to promote *Oct4* expression, it can interact with Oct4 to promote *Nanog* expression, and it can also interact with Sox2 (Hutchins *et al.* 2013, van den Berg *et al.* 2008, Zhang *et al.* 2008). It can bind the enhancer/promoter regions of a number of other pluripotency factors, including *Klf4*, *Rex1* and *Sox2*, and can directly interact with RNA Polymerase II (Feng *et al.* 2009, van den Berg *et al.* 2008, van den Berg *et al.* 2010). Interestingly, the *Esrrb* locus appears to be a convergence point of two processes that promote ES cell pluripotency: Wnt signalling and Nanog transcriptional activity. Inhibition of GSK3β or Wnt3a stimulation can induce *Esrrb* expression by inhibiting the transcriptional repression exerted by Tcf7l1 (commonly referred to as Tcf3. *Tcf3* is the official gene symbol for the Id1 binding partner E2A) (Martello *et al.* 2012). *Esrrb* is the transcript which is upregulated quickest and to the highest levels following tamoxifen-mediated nuclear translocation of Nanog-ERT2 in *Nanog*-null ES cells. Esrrb overexpression can sustain LIF-independent self-renewal even in the absence of *Nanog*, and *Esrrb*-null cells cannot sustain LIF-independent self-renewal upon Nanog overexpression, suggesting Esrrb is a crucial Nanog target (Festuccia *et al.* 2012). In keeping with its *in vivo* phenotype and with its role as a mediator of Nanog, *Esrrb*-null ES cells can be maintained in culture and contribute to chimaeric embryos, despite displaying impaired clonogenicity (Festuccia *et al.* 2012, Martello *et al.* 2012).

1.1.5.4 E-cadherin

E-cadherin (or Cdh1, for Cadherin 1) is a transmembrane protein that is involved in cell-cell adhesion. Its N-terminal extracellular domain mediates the formation of adherens junctions through homotypic interactions with the N-terminal domains of E-cadherin molecules on neighbouring cells. Its C-terminal domain interacts with β-catenin and, through this protein, with α-catenin and the actin cytoskeleton (Stemmler 2008). E-cadherin is expressed throughout pre-implantation development, and is required for embryo compaction and correct generation and segregation of

trophectoderm and inner cell mass (Bedzhov *et al.* 2012, de Vries *et al.* 2004, Fierro-González *et al.* 2013, Kan *et al.* 2007, Larue *et al.* 1994, Riethmacher *et al.* 1995, Stephenson *et al.* 2010). It is expressed in ES cells, where it is required for their adhesive phenotype but not for their maintenance in culture (Larue *et al.* 1996). *Cdh1*-null ES cells, however, display a gene expression profile reminiscent of post-implantation epiblast and do not respond to LIF signalling, but rather appear to acquire dependency on Activin/Nodal signalling (although *Cdh1*-null ES cells can also be maintained in N2B27 medium supplemented with LIF and BMP4, suggesting they can maintain pluripotency under the control of different extracellular signals) (Hawkins *et al.* 2012, Soncin *et al.* 2009, Soncin *et al.* 2011). E-cadherin interacts with the LIF receptor and gp130, and its loss results in the downregulation of the two proteins, a phenotype that can be rescued by overexpressing the extracellular domain of E-cadherin. Forced expression of a constitutively active form of Stat3 can rescue the expression of naïve pluripotency factors in *Cdh1*-null ES cells (del Valle *et al.* 2013). The adhesive phenotype of *Cdh1*-null ES cells can be resuced by expressing N-cadherin (or Cdh2, for Cadherin 2) or chimaeric constructs comprising the extracellular domain of E-cadherin and the intracellular domain of N-cadherin (and vice versa), under the control of the *Cdh1* promoter. The re-acquisition of an adhesive phenotype results in the restoration of a pre-implantation epiblast-like gene expression pattern (Bedzhov *et al.* 2013).

E-cadherin appears to mediate pluripotency in more ways than just through the transduction of LIF signalling. As mentioned previously, it can interact with the transcriptional regulator β-catenin, and through this with Oct4. The interaction of E-cadherin with β-catenin tethers the latter to the membrane and prevents its transcriptional activity, and it is likely that the transcriptional activity of Oct4 is regulated similarly (Faunes *et al.* 2013, Livigni *et al*. 2013, Orsulic *et al.* 1999). Surprisingly, E-cadherin can also replace the requirement for Oct4 in the generation of induced pluripotent stem cells (Redmer *et al.* 2011). E-cadherin is also involved in the mediation of BMP signalling: we have recently shown that culture of ES cells in

the absence of LIF and the presence of BMP results in the maintenance of E-cadherin and Oct4 expression, and that forced loss of E-cadherin rescues the BMP-driven block of differentiation (Malaguti *et al.* 2013, experiments carried out by Dr. Sally Lowell and Dr. Paul Nistor). E-cadherin appears to regulate BMP signalling itself, as *Cdh1*-null ES cells express approximately 80-fold lower levels of *Bmp4* transcripts than wild-type ES cells (Soncin *et al.* 2011).

1.1.6 Other types of pluripotent stem cells

1.1.6.1 Epiblast stem cells

Epiblast stem cells (EpiSCs) represent the *in vitro* equivalent of post-implantation pluripotent epiblast. They were simultaneously derived by two research groups in 2007 from the epiblast of E5.5-E6.5 mouse embryos and from the epiblast of E7.5 rat embryos (Brons *et al.* 2007, Tesar *et al.* 2007). It was later shown that EpiSCs can also be derived from blastocysts and later stage embryos, until the onset of somitogenesis (Najm *et al.* 2011, Osorno *et al.* 2012). It is puzzling that the cultured equivalent of post-implantation pluripotent epiblast can be generated from a pre-implantation stage embryo, but the derivation process takes place over an extended period of time, so it is likely that the ES cell-like ICM cells have progressed in their development during this time. EpiSCs can also be derived from ES cells by extensive passaging in epiblast stem cell culture medium (Guo *et al.* 2009).

Epiblast stem cell pluripotency is reliant on different extracellular signals than ES cell pluripotency, and is maintained by culturing the cells in N2B27 medium supplemented with Fgf2 and Activin A. These factors have a negative impact on ES cell self-renewal, illustrating how diversely pluripotency is protected in the two cell types (Brons *et al.* 2007, Greber *et al.* 2010, Tesar *et al.* 2007).

EpiSCs and ES cells display several differences in gene expression. Both cell types express Oct4 to a similar extent, but EpiSCs express lower levels of Nanog and

Sox2, and express very little if any mRNA of naïve pluripotency factors such as *Esrrb*, *Klf*'s and *Rex1*. They display relatively high expression of post-implantation markers such as *Fgf5* and *T*, which are expressed at low levels in ES cells (Brons *et al.* 2007, Guo *et al.* 2009, Najm *et al.* 2011, Tesar *et al.* 2007).

EpiSCs also display a different capability for contribution to chimaeric embryos. Blastocyst injection of EpiSCs is generally accompanied by a failure of the cells to contribute to the developing embryo (Brons *et al.* 2007, Tesar *et al.* 2007). EpiSC contribution to development after blastocyst injection has been achieved, but only after selecting for an ES cell-like subpopulation of cells by means of a fluorescent reporter (Han *et al.* 2010), or by overexpressing E-cadherin, and thus possibly increasing the adhesive propensity of the cells (Ohtsuka *et al.* 2012). It is not surprising that the *in vitro* equivalent of a cell derived from an apical-basal polarised epithelial sheet struggles to integrate and interact with the three-dimensional inner cell mass, and raises the question of whether EpiSCs would contribute to development more readily if integrated in post-implantation stage embryos. This was proved to be the case by Huang *et al.* (2012), who grafted EpiSCs into gastrulation stage embryos, and demonstrated that EpiSCs can contribute to all germ layers and primordial germ cells. This result further confirms that EpiSCs represent the *in vitro* equivalent of pluripotent post-implantation epiblast.

1.1.6.2 Induced pluripotent stem cells

A groundbreaking set of experiments performed by Kazutoshi Takahashi and Shinya Yamanaka in 2006 demonstrated that a terminally differentiated cell could be reprogrammed to a pluripotent state by means of overexpression of four transcription factors: Klf4, Myc, Oct4 and Sox2. The type of cell generated by this method was termed induced pluripotent stem (iPS) cell. iPS cells are morphologically indistinguishable from ES cells, display a similar gene expression pattern, can contribute to chimaeras upon blastocyst injection and can be transmitted to the

germline of the chimaeric embryos (Takahashi & Yamanaka 2006, Okita *et al.* 2007, Wernig *et al.* 2007). Various combination of transcription factors, signalling molecules, small molecules, RNAs have since been used to generate iPS cells from a variety of tissues (Anokye-Danso *et al.* 2011, Aoi *et al.* 2008, Chen *et al.* 2011, Esteban *et al.* 2010, Heng *et al.* 2010, Hou *et al.* 2013, Ichida *et al.* 2009, Li *et al.* 2011, Miyoshi *et al.* 2011, Redmer *et al.* 2011). iPS cells have also been derived from human cells (Park *et al.* 2008, Takahashi *et al.* 2007), and represent a tool of incredible potential both for the study of patient-specific disease phenotypes in culture, and for the generation of patient-specific cell types for transplantation. Furthermore, they lack the ethical problems associated with embryonic stem cell generation, but present problems of their own. The use of transgenesis for induction of a pluripotent phenotype using oncogenes such as Myc is a prospect that should be avoided if the cells are to be used for transplantation, and it is unclear how derivation of iPS cells from different types of tissues may influence the epigenetic and karyotypic properties of the cells (Kim *et al.* 2010, Ohi *et al.* 2011). Various strategies are being developed to overcome these problems, making use of technologies such as RNA-based or protein transduction-based reprogramming to avoid genomic integration of the reprogramming factors, and of alternative Myc-free transcription factor cocktails to reduce the potential for long-term oncogenicity of iPS cells (reviewed in Bayart & Cohen-Haguenauer 2013).

1.1.6.3 Human embryonic stem cells

Human embryonic stem cells were derived in 1998 by plating inner cell masses on irradiated fibroblasts in serum-containing medium (Thomson *et al.* 1998). The extracellular signalling molecule requirement of human and mouse ES cells varies greatly: LIF signalling does not affect human ES cell self-renewal (Dahéron *et al.* 2004, Reubinoff *et al.* 2000, Thomson *et al.* 1998), BMP4 promotes trophectodermal and endodermal differentiation (Pera *et al.* 2004, Xu *et al.* 2002),

Activin/Nodal/TGFβ signals promote pluripotency (James *et al.* 2005, Vallier *et al.* 2005, Xu *et al.* 2008), as does Fgf (Dvorak *et al.* 2005, Greber *et al.* 2007, Greber *et al.* 2010, Xu *et al.* 2005, Xu *et al.* 2008, Yu *et al.* 2011). The growth factor requirements of human ES cells thus resemble those of mouse EpiSCs, as does the cell morphology, the inability to retain pluripotency in 2i medium, the X chromosome inactivation status and the expression of markers of differentiated lineages (Hanna *et al.* 2010, Pera *et al.* 2004, Silva *et al.* 2008, Tesar *et al.* 2007, Thomson *et al.* 2008). Human pluripotent cells reminiscent of mouse ES cells were derived by means of reprogramming and culture in 2i+LIF with continuous expression of the reprogramming factors (Hanna *et al.* 2010). Recently, two groups identified culture conditions capable of sustaining a mouse ES cell-like naïve pluripotent state for human ES cells without the need for transgenesis, using combinations of growth factors and small molecule inhibitors. In both cases the culture medium contained Fgf2, Tgfβ1, LIF, PD0352901 and a GSK3β inhibitor (Chan *et al.* 2013, Gafni *et al.* 2013). These discoveries represent a great stepping stone for the study and the understanding of pluripotency states and of inter-specific differences in pre-implantation development. They also have an important cell signalling-specific implication: the simultaneous addition of Fgf2 and the Mek inhibitor PD0352901 to the culture media suggests that Fgf signalling exerts a pro-pluripotency effect through effectors alternative to mitogen-activated protein kinase pathway components, an observation compounded by the loss of pluripotency observed when either Fgf2 or PD0352901 are removed from the culture (Gafni *et al.* 2013).

1.2 Lineage specification

1.2.1 Post-implantation development of the mouse

1.2.1.1 Implantation and egg cylinder formation

As the inner cell mass of the blastocyst develops and segregates epiblast and primitive endoderm, the trophectodermal cells also specialise: the polar trophectoderm overlying the ICM remains diploid and proliferative, whereas the mural trophectoderm at the abembryonic pole differentiates into non-mitotic polyploid giant cells (Kunath *et al.* 2004, Rossant 2004). As the blastocyst expands, trophectodermal cells derived from the polar trophectoderm flow into the non-proliferative mural regions. The flow occurs in an asymmetrical manner preferentially on one side of the blastocyst, and may result in the tilt of the polar trophectoderm/epiblast relative to the embryonic/abembryonic axis which can be observed around the time of implantation (Copp 1979, Gardner 2000, Gardner & Davies 2002, Smith 1980, Smith 1985).

Upon hatching from the zona pellucida, a membrane which surrounds the pre-implantation embryo, the blastocysts implants to the uterus through attachment and invasion of the uterine wall by the mural trophectoderm (Cross *et al.* 1994, Kirby *et al*. 1967, Smith 1980, Smith 1985). The polar trophectoderm, the epiblast and the primitive endoderm underlying it expand into the blastocoel cavity (Beddington & Robertson 1999). Apoptosis is induced in the cells at the centre of the epiblast, which leads to the formation of an epithelial monolayer arranged in a hemispheroidal conformation (Coucouvanis & Martin 1985). This structure, which arises between 5 and 5.5 days post fertilisation, is commonly referred to as "egg cup" or "egg cylinder", due to its shape. The region of the epiblast in contact with the derivatives of the polar trophectoderm is referred to as proximal, whereas the region furthest away from the polar trophectoderm is referred to as distal. The proximal/distal axis

of the epiblast therefore coincides with the embryonic/abembryonic axis of the blastocyst (Figure 1.2).

The polar trophectoderm generates the extraembryonic ectoderm, in contact with the epiblast, and the ectoplacental cone, in contact with the uterus (Rossant & Tam 2009). The ectoplacental cone displays a tilt that coincides with the future anterior/posterior embryonic axis, but does not predict the polarity of the axis (Gardner *et al.* 1992, Rossant & Tam 2004, Smith 1985). Smith (1985) suggested that the tilt of the polar trophectoderm/epiblast at implantation may coincide with the tilt of the ectoplacental cone, and thus with the future anterior/posterior axis, but this has not been unequivocally demonstrated (Figure 1.2).

The primitive endoderm specifies two cell types: the parietal endoderm, that underlies the cells of the mural trophectoderm, and the visceral endoderm, which underlies the epiblast and the extraembryonic ectoderm (Figure 1.2) (Rossant 2004).

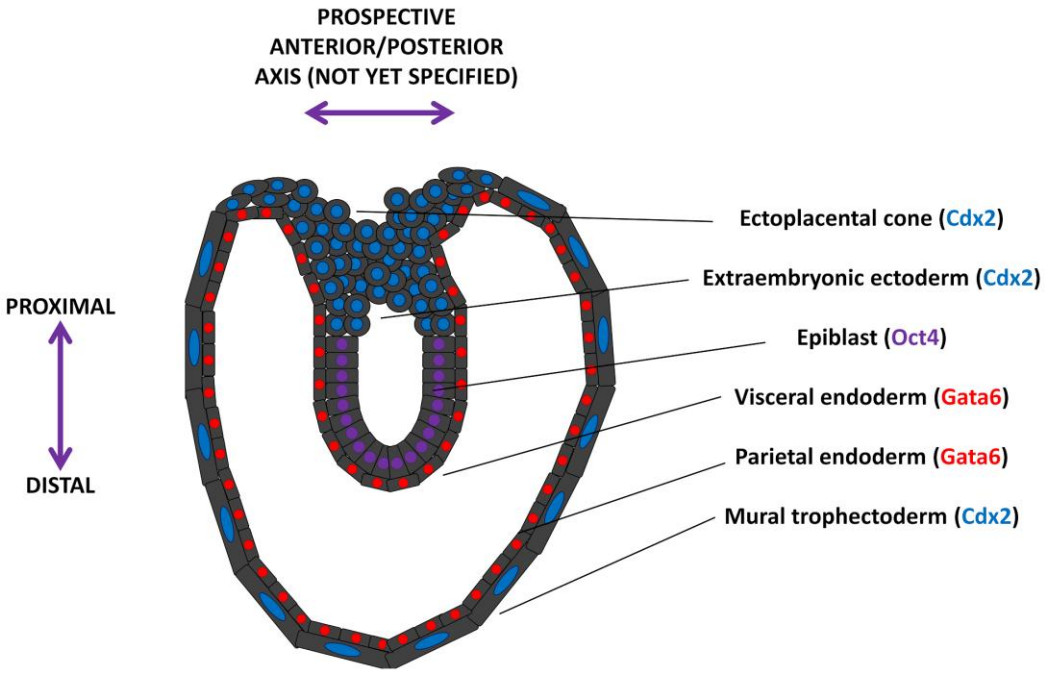


Figure 1.2 – Structure of the mouse embryo at E5.5

Diagrammatic structure of the mouse embryo shortly after implantation. The polar trophectoderm gives rise to the ectoplacental cone, in contact with the uterus, and the extraembryonic ectoderm, in contact with the epiblast. Giant cells of the mural trophectoderm mediate implantation. The epiblast cavitates to form the egg cylinder. The primitive endoderm gives rise to the visceral endoderm, underlying the epiblast and extraembryonic ectoderm, and to the parietal endoderm, underlying the mural trophectoderm. Characteristic markers of these cell types are indicated in the figure. The proximal/distal embryonic axis is depicted, as well as the prospective anterior/posterior axis, as predicted by the tilt in the ectoplacental cone.

1.2.1.2 Establishment of the anterior/posterior axis

The formation of the proximal/distal axis represents the first step in the generation of anterior and posterior regions of the epiblast. *Nodal* is initially expressed throughout

the epiblast (Varlet *et al.* 1997), but the expression of the pro-protein convertases *Furin* and *Pace4*, responsible for converting the inactive pro-Nodal into the active form of Nodal, becomes restricted to the extraembryonic regions (Mesnard *et al.* 2006). Furthermore, Nodal induces the transcriptional activity of Smad2 at the distal tip of the embryo, and this leads to the expression of the Nodal, BMP and Wnt inhibitors *Cer1*, *Dkk1* and *Lefty1* (Cer1 inhibits Nodal and BMP signalling, Dkk1 inhibits Wnt signalling and Lefty1 inhibits Nodal signalling). The expression of these factors, regulated in space by Smad2 activity and by the absence of BMP-induced Smad1 activity (which is antagonistic to Smad2 activity), specifies the distal visceral endoderm (DVE) (Arnold & Robertson 2009, Belo *et al.* 2000, Brennan *et al.* 2001, Chen & Shen 2004, Perea-Gomez *et al.* 2002, Rodriguez *et al.* 2005, Soares *et al.* 2008, Waldrip *et al.* 1998, Wu *et al.* 2000, Yamamoto *et al.* 2009).

DVE cells migrate proximally and establish the anterior side of the epiblast. It appears that *Cer1* and *Lefty1* are asymmetrically expressed and dictate the direction of DVE migration (Yamamoto *et al.* 2004), but it is unclear whether this asymmetric expression may be induced by localised repressive signals at the prospective proximal posterior side of the embryo, such as Wnt3 (induced by BMP4), or whether the inhibition of Nodal, BMP and Wnt signals by Cer1, Dkk1 and Lefty1 are responsible for the localised expression of the DVE-repressive signals (Arnold & Robertson 2009, Kimura-Yoshida *et al.* 2005, Rivea-Pérez & Magnuson 2005, Tam & Loebel 2007).

It was long believed that the anterior migration of the DVE lead to its repositioning at the prospective embryonic anterior and the formation of the anterior visceral endoderm (AVE) signalling centre, responsible for continued inhibition of BMP, Nodal and Wnt. However, live imaging and clonal analysis revealed that as the DVE migrates anteriorly, it is replaced at the distal tip by a new set of cells that start expressing *Cer1* and *Lefty1* after E5.5. These cells follow the DVE in its migration and at E6.5 they localise to the anterior of the embryo, where they form the AVE,

while the cells that formed the DVE localise at the proximal anterior and lateral regions of the embryo (Takaoka *et al.* 2011).

1.2.1.3 Gastrulation

The most important time in the life of the embryo is initiated upon formation of the primitive streak (sometimes referred to as streak for brevity). This structure arises in the epiblast at the embryonic/abembryonic boundary and unequivocally defines the posterior of the embryo. It is an elongated region of epiblast characterised by a furrow through which cells migrate as they lose epithelial identity. The exit from the epiblast and migration through the primitive streak gives rise to mesoderm and endoderm, whilst cells that fail to migrate through the streak will generate ectodermal derivatives (Beddington & Robertson 1999, Nowotschin & Hadjantonakis 2010).

The induction of the primitive streak is dependent upon BMP4, Nodal and Wnt3, which reinforce each other's expression in a positive regulatory loop (Ben-Haim *et al.* 2006, Brennan *et al.* 2001, Conlon *et al.* 1994, Huelsken *et al.* 2000, Liu *et al.* 1999, Mishina *et al.* 1995, Tam & Loebel 2007, Winnier *et al.* 1995). *Wnt3* transcripts are detectable on one side of the proximal epiblast at E5.5 prior to streak formation, which is consistent with the inductive role of Wnt3 in streak determination (Liu *et al.* 1999, Rivea-Pérez & Magnuson 2005). The combination of BMP4, Nodal and Wnt signalling induce the expression of transcriptional determinants of the primitive streak identity, such as *Eomes* and *T*.

T (Brachyury) is a transcription factor essential for gastrulation that is expressed throughout the primitive streak (Beddington *et al.* 1992, Inman & Downs 2006, Wilkinson *et al.* 1990, Willison 1990, Wilson & Beddington 1997, Wilson *et al.* 1995). It is a target of Wnt signalling, and amongst its reported targets are Fgf and Wnt signalling components. It is thus likely that it is crucial in the stabilisation of the

primitive streak signalling environment (Casey *et al.* 1998, Evans *et al.* 2012, Tada & Smith 2000, Yamaguchi *et al.* 1999).

Eomes (Eomesodermin) is a transcription factor expressed in extraembryonic ectoderm throughout early pre-implantation development, and in the proximal posterior epiblast at the time of primitive streak specification (Ciruna & Rossant 1999). It is required for the correct specification of the primitive streak and cell migration through it; *Eomes*-null cells fail to downregulate the epithelial marker E-cadherin and are thus incapable of delaminating from the epiblast through the primitive streak (Arnold *et al.* 2008, Russ *et al.* 2000).

Migration through the primitive streak is achieved by means of an epithelial-to-mesenchymal transition (EMT). Cells in the epiblast, an E-cadherin-expressing epithelium, activate the expression of *Snail*, a transcription factor responsible for promoting the acquisition of mesenchymal migratory characteristics through the repression of *Cdh1* expression and consequent loosening of cell-cell adhesions. *Snail* expression and EMT are dependent on the activity of Fgf8, which is expressed in the proximal posterior regions of the embryo and in the primitive streak (Carver *et al.* 2001, Ciruna & Rossant 2001, Crossley & Martin 1995, Mahmood *et al.* 1995, Sun *et al.* 1999).

As development progresses, the primitive streak elongates distally and reaches the distal tip of the embryo. Cells juxtaposed anteriorly to the primitive streak form a structure knows as the node, the mouse equivalent of Spemann's organiser in *Xenopus* and Hensen's node in the chick, which acts as a source of Nodal (Arnold & Robertson 2009, Beddington & Robertson 1999). The primitive streak is subject to a gradient of BMP4 from proximal to distal regions, a gradient of Nodal from distal to proximal regions, and Fgf8 expression throughout its length. The varying stimulation from these molecules and Wnt's drives the specification of different mesodermal subtypes in cells that migrate at different positions along the streak. From proximal to distal, these are: extraembryonic mesoderm, lateral plate mesoderm, paraxial

mesoderm, axial mesoderm. The cardiac mesoderm is specified at a similar level of the primitive streak to paraxial and lateral plate mesoderm, but migrates before these other subtypes. The cells that migrate through the anterior-most region of the primitive streak give rise to definitive endoderm, which intercalates and displaces the visceral endoderm. Not all visceral endoderm is displaced, as some has been suggested to contribute to the endodermal tissues of the embryo proper. High BMP signalling in the proximal regions of the embryo also gives rise to primordial germ cells (Arnold & Robertson 2009, Kinder *et al.* 1999, Kwon *et al.* 2008, Lawson 1999, Lawson *et al.* 1999, Smith *et al.* 1994, Robb & Tam 2004, Tam & Loebel 2007).

As the embryo continues its expansion and somitogenesis is initiated, the node and the primitive streak become localised to the posterior-most regions of the embryo, extending the node-derived notochord in the process, and supplying mesodermal and endodermal cells to the extending caudal regions. The primitive streak is eventually referred to as the tail bud from E9.0-E9.5 (Arnold & Robertson 2009, Cambray & Wilson 2002, Ramkumar & Anderson 2011, Wilson & Beddington 1996). Intriguingly, it appears that the tail bud harbours a progenitor cell type capable of contributing to both mesoderm and neural lineages (Tzouanacou *et al.* 2009).

1.2.1.4 Neural specification

The induction of a neural fate in the anterior epiblast is achieved via AVE-mediated repression of BMP signals. BMP was identified as a repressor of neural fate in *Xenopus laevis* animal cap explants, which differentiate into the neural lineage when dissociated and cultured in the absence of signalling factors (Grunz & Tacke 1989, Wilson & Hemmati-Brivanlou 1995). This led to the formulation of a widely debated and disagreed upon hypothesis: the default model of neural specification. According to the hypothesis, in the absence of instructive or repressive signals, unspecified

ectodermal cells will give rise to neural ectoderm. Support for the model is drawn by animal cap dissociation experiments in *Xenopus* and early experiments in axolotl, whereas criticism of the model stems primarily from work carried out in the chick embryo, which suggests a key role for Fgf signalling and an insufficient role for BMP inhibition in neural induction (De Robertis & Kuroda 2004, Hemmati-Brivanlou & Melton 1997, Muñoz-Sanjuán & Brivanlou 2002, Stern 2006, Streit *et al.* 1998, Streit *et al.* 2000). *In vitro* experiments on differentiating mouse embryonic stem cells do not ultimately resolve this issue (Kunath *et al.* 2007, Stavridis *et al.* 2007, Ying *et al.* 2003b, see Section 1.2.2 for a detailed discussion of *in vitro* data).

Mouse is not as amenable a model organism for the study of lineage specification as the frog and the chick, and thus there are fewer data available which may shed light on this topic. What is certain is the involvement of BMP repression in neural induction. Mice lacking the *Bmpr1* gene display premature induction of neural markers, not only anteriorly, but throughout the epiblast, and fail to upregulate markers of the primitive streak. Furthermore, the inhibition of Fgf signalling between E5.5 and E7.5 appears to promote the expression of the neural marker *Hesx1* in cultured embryos, suggesting Fgf is not required, but rather plays a negative role on neural induction at the egg cylinder stage (Di-Gregorio *et al*. 2007).

Further evidence implicates BMP signalling in the repression of neural fate *in vivo*. A similar set of experiments to those performed by Wilson & Hemmati-Brivanlou (1995) was recently carried out for mouse epiblasts. Explants of anterior and posterior epiblasts at E6.5, followed by 5 days of culture in serum-free N2B27 medium in the presence or absence of BMP4, revealed that the untreated explants generated neural tissue, whereas the BMP4-treated explants gave rise to surface ectoderm, mesoderm and endoderm. At E7.0, untreated explants generated neural ectoderm, with posterior explants expressing lower levels of neural markers than anterior explants; BMP4 treatment of anterior explants gave rise to surface ectoderm and BMP4 treatment of posterior explants generated mesoderm and endoderm. At E7.5 BMP4 no longer inhibited neural marker expression in anterior explants, but it

was capable of inducing mesodermal and endododermal markers in posterior explants (Li *et al.* 2013). These data have a number of implications: the neural differentiation of untreated explants in unsupplemented medium mimics the behaviour of *Xenopus* animal cap explants, and provides support for the default model; the anti-neural effect of BMP4 directly implicates it in the repression of neural induction until the stage at which cells have committed to the neural lineage (E7.5); the differential behaviour of explants following BMP4 treatment suggests that regional and temporal differences within the embryo affect the response of epiblast cells to BMP4.

Whilst these observations lend support to the default model, it is impossible to ascertain whether instructive signals which may be necessary for neural induction, other than Fgf, are expressed in the *Bmpr1* mutants or in the embryo explant cultures.

One such signal may be represented by Igf (insulin-like growth factor), which is a neural inducer in *Xenopus*, and is expressed in gastrulation-stage embryos (Lee *et al.* 1990, Pera *et al.* 2001, Pera *et al.* 2003, Telford *et al.* 1990). The N2B27 culture medium widely adopted for ES cell and EpiSC differentiation, and used by Li *et al.* (2013) for their explants cultures, contains insulin (Ying *et al*. 2003b).

Nodal is also involved in the repression of neural fates, as demonstrated by the expansion of neural domains in *Nodal*-null embryos (Camus *et al.* 2006). *Nodal* expression is absent from *Bmpr1*-null embryos (Di-Gregorio *et al.* 2007), as may be expected due to the disruption of the positive regulatory loop involving Nodal, BMP and Wnt signals (see section 1.2.1.3). This may also suggest a requirement for low or absent Nodal signalling for neural induction. In keeping with this hypothesis, whilst the knockout of either of the Nodal antagonists *Cer1* or *Lefty1* does not result in a disruption in development, the knockout of both factors can lead to an anterior expansion of the domain of expression of primitive streak determinants (as well as the generation of multiple primitive streaks in a subset of mutants), and this

phenotype can be partially rescued upon deletion of one *Nodal* allele (Perea-Gomez *et al.* 2002).

1.2.2 Embryonic stem cell differentiation as a model for lineage specification

The study of differentiating embryonic stem cells represents an extremely convenient experimental tool for the analysis of the molecular events that accompany lineage specification, for the reasons described in paragraph 1.1.3.3. Their usefulness, however, depends on their capability to replicate the exact molecular decisions that occur *in vivo*, and do so in an appropriate timescale.

Ever since their isolation, it has been known that suspension culture of ES cells results in the formation of cystic embryoid bodies (EBs), hollow spherical cell aggregates containing cells from different germ layers (Doetschman *et al.* 1985, Martin 1981). In particular, the outside of the EBs is lined by endoderm, and the inside cells contain variable mixtures of mesoderm and ectoderm depending on the specific culture conditions. This structure is reminiscent of the endoderm lining the ectoderm and mesoderm in the post-implantation embryo and suggests that differentiating ES cells are capable of three-dimensional self-organisation. Furthermore, the sequence of gene activation over the timescale of differentiation experiments is the same as that seen *in vivo*, which implies that EB differentiation is a developmentally relevant ES cell differentiation system (Keller 2005). EBs present disadvantages however: they contain a mixture of different cell types, making the specification of one cell type impossible to study in isolation; they are three-dimensional, making them less practical than two-dimensional cultures for imaging purposes; they are cultured in suspension, which is troublesome for live imaging.

The differentiation of ES cells in adherent monolayers offers the solution to all of these problems. The culture of ES cells on feeder cells has been used to induce specification of different lineages (Kawasaki *et al.* 2000, Nakano *et al.* 1994),

presumably due to the factors secreted by the different types of feeder cells used in the various protocols. This present a problem in itself, since it does not allow for the assessment of ES cell behaviour in a chemically defined context. This in turn means that drawing conclusions on the effects of soluble molecules of interest in the differentiation process is an imprecise process, as the molecules themselves or their inhibitors may already be secreted by the feeder layer.

ES cells can however be differentiated on extracellular matrix in feeder-free culture (Coraux *et al.* 2003, Nishikawa *et al.* 1998). A widely adopted medium for ES cell differentiation was devised in 2003 by Smith and colleagues by combining N2 and B27-supplemented basal media, which had previously been used for serum-free culture of neural cells. They named the new culture medium N2B27 (Bottenstein & Sato 1979, Brewster *et al.* 1993, Ying *et al.* 2003a, Ying *et al.* 2003b). This medium was devised to support cell survival and proliferation without providing lineage-specifying signals, although, as discussed in paragraph 1.2.1.4, it contains insulin, a protein related to Igf, which can promote neural fates in *Xenopus*.

Culture of ES cells as an adherent monolayer plated on gelatine in N2B27 medium results in their conversion to the neural lineage (Pollard *et al.* 2006, Ying *et al.* 2003a, Ying *et al.* 2003b). This suggests that in the absence of exogenous factors ES cells will specify neural ectoderm, thus providing apparent support to the default model of neural specification in the mouse. However, as discussed in paragraph 1.1.4.4, ES cells secrete Fgf4, and Fgf activity is required during the early stages of neural induction, presumably to specify a post-implantation epiblast-like gene expression pattern (Burdon *et al.* 1999b, Kunath *et al.* 2007, Ma *et al.* 1992, Stavridis *et al.* 2007, Sterneckert *et al.* 2010). Consequently, whilst lack of exogenous signals will result in "default" neural specification, paracrine signalling appears to be essential in the first step of neural induction from ES cells, namely the exit from naïve pluripotency, as also noted by Smith and colleagues in the formulation of their "ground state" of ES cell self-renewal hypothesis (Ying *et al.* 2008). Importantly, ES cells differentiating in N2B27 appear to progress through the molecular transitions

that occur *in vivo*, with initial expression of transcripts characteristic of pluripotent post-implantation epiblast followed by conversion to a neural ectoderm gene expression pattern, suggesting that N2B27 culture represents a developmentally relevant *in vitro* differentiation system (Aiba *et al.* 2006, Aiba *et al.* 2009, Sterneckert *et al.* 2010, Trott & Martinez Arias 2013). Furthermore, adherent monolayer differentiation of ES cells in N2B27 medium supplemented with specific molecules also allows to generate mesodermal, endodermal and epithelial cell types (Hansson *et al.* 2009, Malaguti *et al.* 2013, Ying *et al.* 2003a. The experiments carried out in Malaguti *et al.* 2013 were performed by Dr. Sally Lowell), and supplementation of the medium with factors that promote pluripotency allows for long-term maintenance of ES cells and EpiSCs (Brons *et al.* 2007, ten Berge *et al.* 2011, Tesar *et al.* 2007, Wray *et al.* 2009, Ying *et al.* 2003a, Ying *et al.* 2008). This culture system is therefore an appropriate, powerful and convenient tool for the study of the molecular events that accompany lineage specification decisions.

1.3 The biology of Id1

1.3.1 The helix-loop-helix family of transcriptional regulators

1.3.1.1 Protein domains of basic helix-loop-helix factors

The basic helix-loop-helix (bHLH) protein family was characterised as a set of transcriptional regulators sharing a structurally similar protein domain comprising a basic region adjacent to two α-helices separated by a loop region (Ellenberger *et al.* 1994, Ferré-D'Amaré *et al.* 1993, Murre *et al.* 1989a, Villares & Cabrera 1987). This family is found in a variety of eukaryotic organisms, ranging from the yeast *Saccharomyces cerevisiae* to humans, and in the mouse it comprises over 100 genes, which are involved in a multiplicity of biological processes (Massari & Murre 2000, Skinner *et al.* 2010). There have been various phylogenetic classifications of bHLH proteins, based on sequence analysis, DNA binding properties, function and tissue specificity, and the proteins have been subdivided into 5, 6 or 7 subfamilies (Atchley & Fitch 1997, Ledent *et al.* 2002, Massari & Murre 2000, Stevens *et al.* 2008, Skinner *et al.* 2010). Some bHLH proteins contain domains shared with members of their subfamily: a leucine zipper DNA binding domain (shared by factors such as Myc, Max and Mxd1), a PAS dimerisation motif (shared by factors such as Ahr, Hif1a, Sim1), an Orange protein interaction domain (shared by factors such as Hes proteins, Hey proteins and Stra13), a WRPW motif involved in the repression of transcription (present in Hes proteins) (Blackwood & Eisenman 1991, Dawson *et al.* 1995, Hooker & Hurlin 2006, Huang *et al.* 1993, Kewley *et al.* 2004, Prendergast *et al.* 1991, Sun *et al.* 2007).

1.3.1.2 Interaction of bHLH dimers with DNA

bHLH transcription factors contact DNA as dimers. The HLH domain mediates protein dimerisation, whereas the basic region is responsible for contacting the DNA

at specific recognition sequences termed E-boxes, of sequence CANNTG (Figure 1.3). Specific E-box sequences can be recognised by particular subsets of bHLH proteins, providing high target specificity for bHLH homo- or heterodimers (Ellenberger *et al.* 1994, Ephrussi *et al.* 1985, Jones 2004, Ruzinova & Benezra 2003). In classical models of bHLH activity, tissue-specific bHLH factors heterodimerise with the ubiquitously expressed E proteins in order to contact the DNA (Murre *et al.* 1989b, Norton 2000). E proteins comprise E12, E47 (splice variants of the gene *Tcf3*, also referred to as *E2A*), E2-2 (encoded by *Tcf4*) and HEB (encoded by *Tcf12*) (Massari & Murre 2000).

1.3.2 Id1: discovery and interactors

1.3.2.1 Identification of Id proteins

Id1 was identified in a screen for nucleotide sequences homologous to those encoding the second α-helix of the bHLH factors *Myc*, *Myod* and *Myog*. Weintraub and colleagues noticed that whilst the factor contained a helix-loop-helix domain, it lacked the basic region adjacent to it. They proposed that the protein may act as a dominant-negative heterodimerisation partner, and confirmed this hypothesis by demonstrating it could interact with E12, E47 and Myod, and inhibit their association with DNA. They therefore termed the protein Id, for "Inhibitor of DNA binding" (Benezra *et al.* 1990).

Three more Id proteins have been discovered in the mouse: Id2, Id3 and Id4 (Christy *et al.* 1991, Riechmann *et al.* 1994, Sun *et al.* 1991).

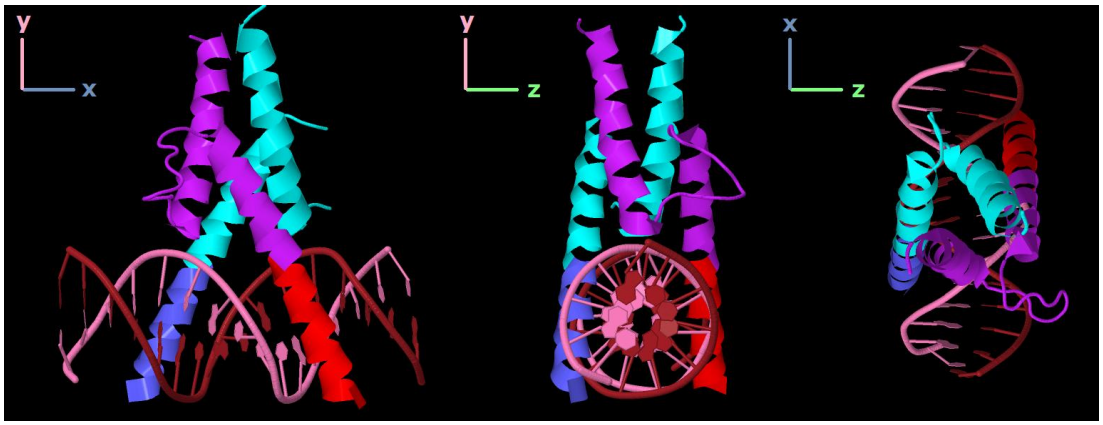


Figure 1.3 – Structure of a basic helix-loop-helix dimer in contact with DNA

Crystal structure of the bHLH regions of a heterodimer of E47 (in cyan and blue) and Neurod1 (in purple and red) in contact with DNA (in pink and dark red). Three alternative views are shown; the orientation of arbitrary x, y and z axes is depicted to clarify the position of each view with respect to the others. The basic region of helix 1 (in red for Neurod1 and in blue for E47) contacts the major groove of DNA, with a glutamate and its neighbouring arginine residue recognising the cytosine and adenine nucleotides of the E-box. The remainder of helix 1, the loop and helix 2 (in purple for Neurod1 and in cyan for E47) mediate the dimerisation of the two proteins (Ellenberger *et al.* 1994, Longo *et al.* 2008). The image was generated using the Jmol Protein Explorer software with the crystal structure generated by Longo *et al.* (2008) (Protein Data Bank accession code: 2QL2). The loop region of E47 is not displayed as it is missing from the crystal structure.

1.3.2.2 Id1 protein properties

Id1 is encoded by a small locus containing 1160 nucleotides of exonic DNA. The locus encodes two splice isoforms, which encode proteins of 148 (Id1-001) and 168 (Id1-002) amino acids, 41 of which constitute the HLH domain (Benezra *et al.* 1990, Hernandez *et al.* 1996, see Figure 3.1 for a diagrammatic view of the locus). The HLH domain of Id1 was shown to be necessary and almost sufficient for the repression of the DNA binding activity of E2A and Myod (Pesce & Benezra 1993). The remainder of the proteins contains no conserved protein domains but was found to contain a nuclear export signal (NES) and a nuclear localisation signal (NLS) (Makita *et al.* 2006, Trausch-Azar *et al.* 2004). Id1 is approximately 15kDa in size, and thus is able to freely diffuse through the nuclear pore (Mattaj & Englmeier 1998). It is therefore likely that the intracellular localisation of Id1 is tightly regulated and of importance for its activity. Id1 can also be imported into the nucleus in complex with E12 and E47, and can mediate nuclear import of NLS-deficient E12 and E47 proteins (Deed *et al.* 1996, Lingbeck *et al.* 2005). Co-transfection of Id1 with E12 or E47 in HeLa cells increases the half-life of the protein from 0.9 hours (transfection of Id1 alone) to 10.3 hours (Id1+E12) or 9.5 hours (Id1+E47), suggesting E proteins can protect Id1 from the proteasome, which mediates its degradation (Bounpheng *et al.* 1999, Lingbeck *et al.* 2005).

Id1 can be subject to post-translational modification. It has been shown to be phosphorylated *in vitro* by protein kinase A and protein kinase C. Id1 phosphorylation did not impair its ability to inhibit the DNA binding of E47 (Nagata *et al.* 1995).

1.3.2.3 Regulation of bHLH transcriptional activity

Despite the ability of its HLH domain to mediate dimerisation, Id1 appears to be unable to homodimerise (Sun *et al.* 1991). It has been proposed this is the result of

repulsion between the positively-charged arginine 68 residue on one Id1 monomer and arginine 96 on a second monomer, which would find themselves in close proximity in a homodimer (Wibley *et al.* 1996). Id2 shares its inability to homodimerise, whereas Id3 is capable of doing so (Sun *et al.* 1991, Wibley *et al.* 1996). Id1 can interact with tissue-specific bHLH factors such as Myod, but appears to do so with low efficiency, whereas it has high propensity for interaction with E proteins (Benezra *et al.* 1990, Sun *et al.* 1991). It therefore appears that Id1 inhibits DNA binding primarily by sequestering E proteins and preventing their homo- or hetero-dimerisation with tissue-specific bHLH factors. A yeast two-hybrid screen performed in our laboratory to identify Id1 interactors in mouse ES cells identified E12 and E47 exclusively, providing support to this model of Id1 action (Davies *et al.* 2013). bHLH factors that preferentially dimerise with factors other than E proteins, and that have weak affinity for Id's, such as Myc (Loveys *et al.* 1996, Murre *et al.* 1989b), would therefore be able to escape Id-mediated transcriptional regulation.

Id proteins preferentially interact with specific E proteins: Id1 has a similar propensity for interaction with E12, E47 and E2-2 but lower propensity for interaction with HEB, Id2 favours interactions with E47 over the other factors, whereas Id3 displays reduced differences in its interaction preferences (Langlands *et al.* 1997). These results are supported by the observations that neonatal lethality in *Tcf3*-null mice (i.e. mice that lack E12 and E47) is partially rescued in *Tcf3*$^{-/-}$*;Id1*$^{-/-}$ mice, whereas the perinatal lethality in *Tcf12*-null (HEB-null) mice is partially rescued in *Tcf12*$^{-/-}$*;Id3*$^{-/-}$ mice, but not in *Tcf12*$^{-/-}$*;Id1*$^{-/-}$ mice (Barndt & Zhuang 1999, Yan *et al.* 1997).

1.3.2.4 Interactions with non-bHLH proteins

Id1 has been found to interact with proteins that lack a helix-loop-helix domain. It can interact with Psmd4, a component of the proteasome (also referred to as S5a).

Psmd4 interacts with Id1 in a region N-terminal to the HLH domain and prevents Id1 from heterodimerising with bHLH factors. It also diminishes its half-life (Anand *et al.* 1997). Id1 can interact with Pax2, Pax5 and Pax8, and inhibit their DNA binding and transcriptional activity. These proteins regulate the development of B-cells and kidney, so their inhibition by Id1 may play a role in the specification of these cell types (Roberts *et al.* 2001). Id1 can also interact with Ets family transcription factors. This protein family comprises many members involved in many biological processes, and it includes Fgf signalling mediators (Jedlicka & Gutierrez-Hartmann 2008). Id1 has been shown to interact with Elk1 and Ets2 and to repress their transcriptional activity (Ohtani *et al.* 2001, Yates *et al.* 1999). Transcriptional activation of *Cdkn2a* (often referred to as *INK4a* or *p16-INK4a*) by Ets2 promotes cellular senescence, and Id1 can inhibit this process (Ohtani *et al.* 2001). Id1 can also interact and inhibit the activity of another factor involved in the activation of *Cdkn2a*, Dnajc2 (also referred to as MIDA1 and ZRF1) (Ribeiro *et al.* 2013, Shoji *et al.* 1995). The involvement of Id1 in the repression of *Cdkn2a* transcription through alternative pathways suggests the HLH factor may be crucial for the inhibition of premature cellular senescence.

Other Id proteins interact with cell cycle components. Human ID2 can directly interact with the cell cycle regulator and tumour suppressor RB, and human ID2 and ID3 are phosphorylated by CDK2 (Deed *et al.* 1997, Iavarone *et al.* 1994). Human ID1 is not phosphorylated by CDK2 nor by CDK4 and does not interact with RB (Deed *et al.* 1997, Hara *et al.* 1996), suggesting that it might not be regulated by cell-cycle components.

1.3.2.5 Regulation of Id1 expression

As discussed in paragraphs 1.1.4.2 and 1.1.4.3, *Id1* is a positive target of BMP signalling and a negative target of Activin/Nodal/TGFβ signalling (Galvin *et al.* 2010, Hollnagel *et al.* 1999, Katagiri *et al.* 2002, Korchynskyi & ten Dijke 2002,

López-Rovira *et al.* 2002, Ogata *et al.* 1993). BMP activation of Smad1 and Smad5 results in their interaction with Smad4 and direct binding of the *Id1* promoter at two Smad Binding Elements and a GC-rich region, resulting in the induction of *Id1* transcription (Katagiri *et al.* 2002, Korchynskyi & ten Dijke 2002). Nanog can bind to Smad1 and repress the transcriptional activation of *Id1* (Suzuki *et al.* 2006), and *Id1* message is enriched in cells expressing low levels of a Nanog-GFP reporter (Galvin *et al.* 2010).

In addition to the potential mechanisms for Activin/Nodal-mediated inhibition of BMP signalling described in paragraph 1.1.4.3, SMAD3, but not SMAD2, can directly bind the *ID1* promoter in human cell lines in a region overlapping that bound by SMAD1. The repression of *ID1* transcription is dependent on the interaction of SMAD3 with ATF3 (Kang *et al.* 2003). The interaction of both Smad2 and Smad3 with the *Id1* locus has been observed in mouse ES cells (Lee *et al.* 2011). Interestingly, TGFβ signalling can induce the expression of *ID1* in human breast cancer cells, suggesting that SMAD3 may interact with cell-type specific factors promoting transcriptional activation of *ID1* (Padua *et al.* 2008, Stankic *et al.* 2013).

Id1 transcription is induced by serum as well as BMP4 (Barone *et al.* 1994, Benezra *et al.* 1990, Ying *et al.* 2003a). This induction was shown to be carried out by a protein complex including the zinc-finger protein Egr1, and an Egr1 binding site was identified on the *Id1* promoter (Tournay & Benezra 1996). A recent publication has shown that *Id1* induction in response to serum is dependent on active BMP signalling, Smad4 expression and a BMP response element on the *Id1* promoter (Lewis & Prywes 2013). This suggests that Egr1 may be in complex with Smad proteins during the serum-induced activation of *Id1* transcription.

The *Id1* promoter also contains a binding site for C/EBPβ, which drives *Id1* expression in pro-B-cells (Saisanit & Sun 1995, Saisanit & Sun 1997).

1.3.3 The expression and function of Id1 in development

1.3.3.1 Expression of Id genes in early development

The expression of Id factors in pre-implantation development has not been assessed by means of *in situ* hybridisation or immunohistochemistry. The data available were generated in a single-cell quantitative real-time (qRT) PCR experiment on individual cells taken from the E3.5 and E4.5 inner cell masses and on single ES cells. *Id1* was detected in 6 out of 14 cells from the E3.5 ICM, in 6 out of 10 cells from the E4.5 ICM and in 10 out of 14 ES cells. In the E4.5 ICM, it is expressed in 3 out of 4 cells co-expressing *Gata4* and *Gata6*, and in 3 out of 6 cells expressing *Nanog*, suggesting it is expressed both in the primitive endoderm and in the epiblast, although firm conclusions cannot be drawn from the analysis of so few cells (Tang *et al.* 2010). Phosphorylated Smad1 is heterogeneously distributed in the blastocyst, in line with this expression pattern for *Id1* (James *et al.* 2005). *Id2* was not detected at E3.5, and detected in 1 cell (expressing *Gata4* and *Gata6*) at E4.5. It displays heterogeneous expression in ES cells. *Id3* is expressed at similar levels in all cells of the E3.5 ICM, and is expressed in 8 out of 10 cells at E4.5. Its expression pattern in ES cells is more heterogeneous. *Id4* was only detected in one E4.5 cell (expressing *Nanog*), and in 1 ES cell (Tang *et al.* 2010).

The data available for pre-gastrula and gastrula-stage embryos were generated in 1992 and 1997 by *in situ* hybridisation of radioactively labelled probes (Jen *et al.* 1997, Wang *et al.* 1992). The images are at low resolution, and it is somewhat hard to specifically identify which tissues have been hybridised and whether the hybridisation signal is above background levels. A recent publication has investigated the expression pattern of *Id1* and *Id2* at E6.5, E7.0 and E7.5, helping to shed some light on the expression pattern of these two genes (Li *et al.* 2013).

At E5.5, *Id1* is expressed in the proximal regions of the embryo and in the extraembryonic endoderm. It is unclear whether *Id1* expression is detected in the epiblast or in the underlying visceral endoderm and overlying extraembryonic

ectoderm. At E6.5 the expression pattern remains virtually identical. At E7.5 it is still expressed proximally, and is also detected in the allantois and in the chorion (Jen *et al.* 1997, Wang *et al.* 1992). The data by Li *et al.* (2013) show that at E6.5 *Id1* appears to be expressed both in the visceral endoderm and in the epiblast, but it is unclear whether this expression also extends to the extraembryonic ectoderm. At the early bud stage (using the embryo staging guidelines of Downs & Davies 1993) the expression of *Id1* remains proximal and extends to the allantoic bud. At the headfold stage, *Id1* is expressed in the cardiac crescent, in lateral mesoderm and in the chorion.

At E6.5, *Id2* is expressed in the extraembryonic ectoderm. At E7.5, it is expressed in the chorion. It is unclear whether it is also expressed in the epiblast (Jen *et al.* 1997). The data by Li *et al.* (2013) confirm the expression of *Id2* in the extraembryonic ectoderm at E6.5. At the late bud stage, *Id2* is expressed strongly in the chorion, but it also displays weaker expression in the proximal regions of the embryo. At the headfold stage, *Id2* is expressed in the chorion and in the cardiac crescent. It can also be detected in the posterior of the embryo, but the published image does not clarify whether it is expressed along the primitive streak or whether the signal results from probe trapping.

The *in situ* hybridisation images for *Id3* and *Id4* are too unclear to draw firm conclusions upon. The authors suggest *Id3* is expressed throughout the epiblast and *Id4* is not expressed (Jen *et al.* 1997).

All four *Id* genes can be detected at later stages in a variety of cell types and tissues, derived from all three germ layers. The expression pattern of the genes is not identical, suggesting that their expression is differentially regulated, and that they may be involved in the regulation of different developmental processes. *Id1* and *Id3* display the most overlapping expression patterns, whereas *Id4* is primarily expressed in the neural lineage (Ellmeier & Weith 1995, Jen *et al.* 1996, Jen *et al.* 1997, Wang *et al.* 1992).

1.3.3.2 Knockout of Id genes in vivo

The knockout of *Id1* has no effect on the viability of mice and results in no obvious phenotypic defect (Yan *et al.* 1997). *Id2*-null mice are indistinguishable from their littermates at birth, but a quarter of pups dies neonatally. The surviving mice exhibit retarded growth, lack lymph nodes and Peyer's patches, and have a reduced number of natural killer cells (Yokota *et al.* 1999). Pregnant females also display defects in their mammary glands and are unable to lactate (Mori *et al.* 2000). The knockout of *Id3* does not affect the viability of mice. *Id3*$^{-/-}$ mice have a reduced number of CD4^{+} and CD8^{+} thymocytes, lower amounts of IgG1 and IgG2a in serum and lower B-cell proliferation than wild-type mice in response to treatment with an anti-mouse IgM antibody. The B-cell proliferation defect can be rescued by B-cell-specific overexpression of *Id1* in *Id3*$^{-/-}$ mice (Pan *et al.* 1999). Only 50% of *Id4*-null embryos survive until birth, and only 20% survive until aldulthood. The mutants display smaller brain sizes apparent from E11.5 due to premature terminal differentiation of progenitor cells (Bedford *et al.* 2005, Yun *et al.* 2004).

Compound mutants for more than one *Id* gene have also been generated. Because of their overlapping expression pattern and of the ability of Id1 to rescue some of the phenotypic defects of *Id3*-null mice, *Id1*$^{-/-}$*;Id3*$^{-/-}$ embryos were generated. The embryos exhibited angiogenic defects and smaller brain sizes from E10.5 and invariably died at E13.5, presumably because of cranial haemorrhaging from E12.5. A single intact allele of either *Id1* or *Id3* is sufficient for seemingly normal development and postnatal survival (Lyden *et al.* 1999). Knockout of any 4 out of 6 alleles of *Id1*, *Id2* and *Id3* results in embryonic lethality, including genotypes such as *Id1*$^{+/-}$*;Id2*$^{+/-}$*;Id3*$^{-/-}$, in which only one of the three genes is knocked out on both alleles. Lethality was due to defects in cardiogenesis from as early as E9.5. The cardiac defects in *Id1*$^{-/-}$*;Id3*$^{-/-}$ embryos appear to be due to impaired Igf1 and Wnt5a secretion, and intraperitoneal injection of Igf1 allowed to obtain live born *Id1*$^{-/-}$*;Id3*$^{-/-}$

pups. This study did not report on whether any defects were observed in other tissues (Fraidenraich *et al.* 2004).

1.3.3.3 Knockout of Id1 in ES cells

The phenotype of *Id1*-null ES cells has also been investigated. These cells express lower levels of *Nanog* and *Rex1* and higher levels of *T*. They have a reduced clonogenic potential, possibly as a result of the lower levels of *Nanog*. It is unclear whether Id1 has a direct effect on *T* expression, or whether the cells are simply adopting a gene expression pattern more similar to that of epiblast stem cells. Analysis of the microarray data comparing $Id1^{-/-}$ ES cells to wild-type ES cells confirms that other post-implantation epiblast markers, such as *Fgf5* and *Otx2*, are upregulated in the knockout cells (Romero-Lanman *et al.* 2011).

1.3.3.4 Helix-loop-helix networks

The development of many tissue types relies on a correct balance between the self-renewal of stem and progenitor cells and the differentiation of these cells into the specific cell types required for the development and the homeostasis of the tissue in question. A commonly adopted mechanism for achieving this balance is the establishment of a cross-regulatory network of helix-loop-helix factors. This type of network is involved in a variety of developmental events, including myogenesis and neurogenesis (Massari & Murre 2000). I shall describe how an HLH network regulates neurogenesis in order to present the mechanisms underlying this type of transcriptional network.

The undifferentiated status of neural stem cells and neural progenitor cells is maintained by two families that negatively regulate bHLH factor transcription: the Id protein family and the Hes protein family.

As previously discussed, Id proteins inhibit the DNA binding and transcriptional activity of E proteins and of their heterodimerisation partners, which can induce cell differentiation. Id proteins play a crucial role in the maintenance of neural stem and progenitor cell multipotency, as exemplified by the reduced brain size and premature differentiation observed in *Id1*$^{-/-}$*;Id3*$^{-/-}$ embryos and in *Id4*$^{-/-}$ embryos (Bedford *et al.* 2005, Lyden *et al.* 1999, Yun *et al.* 2004).

Hes proteins derive their name from their homology to the *hairy* and *enhancer of split* genes of *Drosophila melanogaster*. They are bHLH factors which also contain an Orange protein interaction domain and a WRPW C-terminal motif. They can achieve transcriptional repression in a number of ways: they can prevent DNA binding by E47 in a manner reminiscent of Id proteins; they can bind target DNA sequences (E-boxes or N-boxes for some Hes family members) and repress the transcription of pro-differentiation bHLH factors such as *Ascl1* (also known as *Mash1*) through interaction with TLE corepressor factors; they can interact with pro-differentiation factors already bound to DNA and inactivate transcription through association with TLE's (Chen *et al.* 1997a, Dai *et al.* 2010, Giagtzoglou *et al.* 2003, Ross *et al.* 2003, Sasai *et al.* 1992). Furthermore, Hes family members can cross-repress each other's activity, and Hes factors can directly bind Id proteins (Bai *et al.* 2007, Jhas *et al.* 2006, Jögi *et al.* 2002). Many Hes family members are Notch target genes, and can thus integrate cell-cell signalling with transcriptional activity (Ohtsuka *et al.* 1999, Sun *et al.* 2007).

Neuronal specification occurs after the asymmetric division of an undifferentiated neural stem cell. One daughter cells remains undifferentiated and expresses Id and Hes proteins, and the other daughter cell loses expression of these factors and upregulates pro-differentiation bHLH factors, such as *Ascl1* or *Neurog2*. These factors are capable of initiating neuronal differentiation and drive the formation of different neuronal subtypes (GABAergic and glutamatergic respectively). Neurog2 activates the expression of *Neurod1*, a bHLH factor capable of inducing terminal neuronal differentiation (Ma *et al.* 1999). The asymmetricity of the division of the

stem cell is reinforced through the activity of pro-neural genes, which induce the activity of the Notch ligand Delta, and drive the expression of Hes factors in the neighbouring sister cell (Bertrand *et al.* 2002). Self-renewal and differentiation decisions are sometimes enforced by stabilisation of oscillatory gene expression patterns. Hes1 and Ascl1 protein levels display oscillatory anti-phase behaviour in neural progenitor cells. If the oscillations in a cell terminate when Ascl1 levels are high and Hes1 levels are low, the cell can differentiate, if they terminate when Hes1 levels are high and Ascl1 levels are low, the cell will retain multipotency (Imayoshi *et al.* 2013, Kageyama *et al.* 2007, Kageyama *et al.* 2008). bHLH factors also regulate astrocyte differentiation, through Hes activity, and oligodendrocyte differentiation, through Olig activity. Interestingly, Olig2 also displays oscillatory patterns in neural progenitor cells (Imayoshi *et al.* 2013).

The oscillations in Hes protein expression are not limited to neural cells, and can be observed in ES cells. Differences in Hes1 expression affect the behaviour of ES cells during differentiation, with Hes1-high cells more prone to mesodermal fates and Hes1-low cells more prone to neural fates (Kobayashi *et al.* 2009, Zhou *et al.* 2013). Other HLH factors are also expressed in ES cells: Id proteins mediate the BMP-induced maintenance of pluripotency, E proteins are expressed, and three pro-differentiation factors expressed in LIF+FCS culture (Neurod1, Tcf15 and Twist1) have been found to interact with E12 and E47 and to promote the exit from a pluripotent state (Davies *et al.* 2013, Malaguti *et al.* 2013, Marchand *et al.* 2009, Ying *et al.* 2003a. The experiments on Twist1 in Malaguti *et al.* 2013 were performed by Dr. Paul Nistor). This suggests that a helix-loop-helix network may be involved in the regulation of ES cell pluripotency and differentiation. Notch may be involved in the regulation of this network, as it is known to affect the differentiation of ES cells, but it does not influence ES cell self-renewal, and it is not upstream of Hes1 activity or expression in this cell type (Kobayashi & Kageyama 2010, Lowell *et al.* 2006, Nemir *et al.* 2006, Schroeder *et al.* 2006).

1.3.3.5 The roles of Id1 in ES cell self-renewal and differentiation

Upon discovery of BMP signalling as the key component of serum for the maintenance of ES cell pluripotency, Id1 was identified as a key effector of this signalling pathway, and Id1 overexpression was found to be sufficient to sustain ES cell self-renewal in the presence of LIF (Ying *et al.* 2003a). Id1 appears to be dispensable for the maintenance of pluripotency in 2i culture, as it is expressed at extremely low levels (Marks *et al.* 2012), and its knockout does not result in overt differentiation in ES cell culture, although this could be due to the redundant expression of other Id family members (Romero-Lanman *et al.* 2011, see paragraph 1.3.3.3). Its expression in LIF+FCS culture has only been investigated in the single-cell qRT-PCR experiments described in paragraph 1.3.3.1, and appears to be heterogeneous (Tang *et al.* 2010).

Id1 is rapidly downregulated at the onset of differentiation, both in N2B27 and in embryoid bodies (Aiba *et al.* 2006, Aiba *et al.* 2009, Romero-Lanman *et al.* 2011), an observation consistent with its role in the maintenance of ES cell self-renewal. Forced expression of Id1 inhibits cell differentiation towards the neural lineage and generates a mixture of morphologically diverse non-neural cells (Ying *et al.* 2003a).

Despite its potent effect in the maintenance of pluripotency and in the inhibition of neural specification, very little is known of the expression pattern and of the roles of Id1 in pluripotent and differentiating cells. The work presented in this thesis aims to fill this void at least in part, in order to do justice to a small protein with enormous potential for developmental regulation.

CHAPTER 2

Materials and Methods

2.1 Materials

2.1.1 General reagents

Product	**Vendor**	**Catalogue number**
0.25% Trypsin EDTA, phenol red	Gibco	25200-056
1 kb DNA Ladder	New England Biolabs	N3232L
1,4-Dithiothreitol (DTT)	Roche	10197785103
10 mM dNTP Mix	Invitrogen	18427-013
100 bp DNA Ladder	New England Biolabs	N3231L
100mm TC-Treated Culture Dish	Corning	430167
2-Mercaptoethanol	BDH	44143-3A
75cm^2 Rectangular Canted Neck Cell Culture Flask with Plug Seal Cap	Corning	430720
λ DNA-HindIII digest	New England Biolabs	N3012S
Absolutely RNA Miniprep Kit	Agilent Technologies	400800
Accutase® solution	Sigma	A6964
Acetic acid glacial AnalR Normapur®	VWR	20104.334
Amersham ECL Blocking Agent	GE Healthcare	RPN2125
Amersham Hybond ECL Nitrocellulose Membrane	GE Healthcare	RPN203D
Amersham Hyperfilm ECL (24x30cm)	GE Healthcare	28-9068-40
Amersham Rediprime II DNA Labeling System	GE Healthcare	RPN1633
Ammonium chloride	Acros Organics	423285000

Ampicillin	Calbiochem	171254
Bacto™ Agar	BD	214010
BamHI-HF®	New England Biolabs	R3136T
Blasticidine S hydrochloride	Sigma	15205-100MG
Bovine Albumin Fraction V (7.5% solution)	Gibco	15260-037
Bovine Serum Albumin 100X (10mg/ml)	NEB	B9001S
CellsDirect™ One-Step RT-PCR Kit	Invitrogen	11753-100
CHIR 99021	Axon	1386
cOmplete ULTRA Tablets, EDTA-Free, Glass Vials	Roche	05892953001
Containers, Polystyrene, 30ml Universal, Flow Seal Cap	Sterilin	128A/FS
Costar® 12 well clear TC-treated multiple well plates, individually wrapped, sterile	Corning	3513
Costar® 24 well clear TC-treated multiple well plates, individually wrapped, sterile	Corning	3524
Costar® 48 well clear TC-treated multiple well plates, individually wrapped, sterile	Corning	3584
Costar® 6 well clear TC-treated multiple well plates, individually wrapped, sterile	Corning	3516
Costar® 96 well clear flat bottom TC-treated microplate	Corning	3595

CryoTube™ Vials	Thermo Scientific	377224
Deoxyribonucleic acid from herring sperm	Sigma	D7290-1ML
DAPI in water at 10mg/mL	Biotium	40043
DMEM/F-12	Gibco	21331-020
DNeasy® Blood and Tissue Kit	Qiagen	69504
Donkey serum	Sigma	D9663-10ML
Dulbecco's Phosphate Buffered Saline	Sigma	D8537
Ethanol absolute AnalaR NORMAPUR®	VWR	101074F
Fibronectin, from bovine plasma	Sigma	F4759
Foetal bovine serum, qualified, E.U.-approved, South America origin	Gibco	10270, batch 40F0240K
G418 sulphate	PAA	P27-011
Gelatine from porcine skin	Sigma	G1890
GenePulser®/MicroPulser™ electroporation cuvettes, 0.4cm gap	Bio-Rad	165-2088
Glasgow Minimum Essential Medium	Sigma	G5154
Glycerol	Fisher	56-81-5
Glycine	Sigma	G8898-500G
Illustra ProbeQuant G-50 Micro Columns	GE Healthcare	28-9034-08
Insulin from bovine pancreas	Sigma	I1882
Isopropanol, 99.5%, for molecular biology, DNAse RNAse and Protease free	Acros Organics	327272500
Kanamycin	Calbiochem	420311
KasI	New England Biolabs	R0544S

KnockOut™ Serum Replacement	Gibco	10828-028
KpnI	New England Biolabs	R0142S
L-Glutamine 200mM (100X)	Gibco	25030-024
Laminin, 1mg/ml, from Engelbreth-Holm-Swann murine sarcoma basement membrane	Sigma	L2020
LB Broth, Miller (Luria-Bertani)	BD	244620
LDN 193189	Axon	1509
Lightcycler® 480 Multiwell Plate 384, white	Roche	04729749001
Lightcycler® 480 Probes Master	Roche	04887301001
Lightcycler® 480 SYBR Green I Master	Roche	04707516001
Lipofectamine™ 2000 Reagent	Life Technologies	11668-027
M-MLV Reverse Transcriptase	Invitrogen	28025-013
M2 medium	Sigma	M7167-100ML
MEM NEAA	Gibco	11140-035
Methanol AnalaR Normapur®	VWR	20847.240
Millex®-GP syringe filter unit, 0.22 µm, polyethersulfone, 33 mm, gamma sterilized	Millipore	SLGP033RS
Neurobasal® medium	Gibco	21103-049
NotI-HF®	New England Biolabs	R3189L
NuPAGE® 4-12% Bis-Tris Gels, 1.0 mm, 10 well	Novex	NP0321BOX
NuPAGE® LDS Sample Buffer (4X)	Novex	NP0007
NuPAGE® MOPS SDS Running Buffer (20X)	Novex	NP0001

NuPAGE® Transfer Buffer (20X)	Novex	NP0006
Nylon membranes, positively charged	Roche	11417240001
Orange G	Sigma	O3756
Orthoboric acid	Fisher	10043-35-3
Paraformaldehyde	Sigma	158127-500G
PD 0325901	Axon	1408
Pefabloc® SC	Sigma	76307-100MG
Penicillin-Streptomycin (10,000 U/mL)	Gibco	15140-122
PerfectHyb™ Plus	Sigma	H7033
Petri dish, 140mm	Sterilin	501V
pGem®-T Easy Vector System I	Promega	A1360
Phosphate buffered saline tablet	Sigma	P4417-100TAB
Plasmid Maxi Kit	Qiagen	12162
Platinum® *Pfx* DNA Polymerase	Invitrogen	11708-013
Poly-L-ornithine solution	Sigma	P4957
Progesterone	Sigma	P8783
ProLong® Gold Antifade Reagent	Molecular Probes	P36930
Propidium Iodide, 1mg/mL solution in water	Biotium	40017
Puromycin dihydrochloride from *Streptomyces alboniger*	Sigma	P8833
Putrescine dihydrochloride	Sigma	P5780
QIAprep Spin Miniprep Kit	Qiagen	27104
QIAquick Gel Extraction Kit	Qiagen	28704
Random Primers	Invitrogen	48190-011
rAPid Alkaline Phosphatase	Roche	04898133001
Recombinant human BMP-4	R&D	314-BP-010

Recombinant human Fgf basic (146aa)	R&D	233-FB-025
Recombinant human/rat/mouse Activin A	R&D	338-AC-050
Restriction Endonuclease HindIII	Roche	11274040001
Restriction Endonuclease ScaI	Roche	10775266001
RIPA Buffer	Sigma	R0278-50ML
RNase H	New England Biolabls	M0297L
RNaseOUT™ Recombinant Ribonuclease Inhibitor	Invitrogen	10777-019
SB431542	Abcam	ab120163
SeeBlue® Pre-Stained Standard	Novex	LC5625
Sodium chloride	Fisher	S/3105/63
Sodium dodecyl sulphate	Sigma	L3771-25G
Sodium hydroxide	Fisher	S/4880/53
Sodium selenite	Sigma	S5261
Spermidine	Sigma	S2626
Standard Petri Dish, 90mm, Single Vent, AS	Sterilin	101R20
Subcloning Efficiency™ DH5α™ Competent Cells	Invitrogen	18265-017
SuperSignal™ West Pico Chemiluminescent Substrate	Thermo Scientific	34079
SYBR® Safe DNA Gel Stain	Invitrogen	S33102
T4 DNA Ligase	New England Biolabs	M0202L
Taq DNA Polymerase	Qiagen	201203
Triton™ X-100	Sigma	X100-100ML
Tween® 20	Promega	H5151

Tris Base	Fisher	BP152-1
UltraPure™ Agarose	Invitrogen	16500100
UltraPure™ DNase/RNase-Free Distilled Water	Gibco	10977
XhoI	New England Biolabs	R0146S
Zero Blunt® TOPO® PCR Cloning Kit, without competent cells	Invitrogen	450245

2.1.2 Instruments

FACSAria (BD)

FACSCalibur (BD)

G:BOX F3 (Syngene)

Gel tanks (Engineering & Design Plastics Ltd)

Gene Pulser (Bio-Rad)

IX51 inverted microscope (Olympus) with Retiga-2000R Fast 1394 camera (QImaging)

LightCycler® 480 Instrument II (Roche)

LSRFortessa (BD)

NanoDrop ND-1000 Spectrophotometer (Thermo Scientific)

NanoVue Plus Spectrophotometer (GE Healthcare)

PowerPac™ Basic Power Supply (Bio-Rad)

StereoZoom SMZ-U dissection microscope (Nikon)

TCS SPE inverted confocal microscope (Leica)

TProfessional Standard Thermocycler (Biometra)

Xcell SureLock® Mini-cell with Xcell II™ Blot Module Kit (Novex)

2.1.3 PCR primer sequences

2.1.3.2 qRT-PCR primer sequences and UPL probe number

Target	Forward primer	Reverse primer	UPL probe
Ascl1	TCTCCTGGGAATGGACTTTG	CGTTGGCGAGAAACACTAAAG	74
Cdh1	ATCCTCGCCCTGCTGATT	ACCACCGTTCTCCTCCGTA	18
Cdh2	GCCATCATCGCTATCCTTCT	CCGTTTCATCCATACCACAAA	18
Dlx5	AGCCCCTACCACCAGTACG	GCTCCGCCACTTCTTTCTC	SYBR Green
Eomes	ACCGGCACCAAACTGAGA	AAGCTCAAGAAAGGAAACATGC	9
Esrrb	CGATTCATGAAATGCCTCAA	CCTCCTCGAACTCGGTCA	89
Fgf5	AAAACCTGGTGCACCCTAGA	CATCACATTCCCGAATTAAGC	29
Gapdh	CCCACTAACATCAAATGGGG	CCTTCCACAATGCCAAAGTT	SYBR Green
Gata2	CACAAGATGAATGGACAGAACC	ACAGGTGCCCGCTCTTCT	75
Gata6	GGTCTCTACAGCAAGATGAATGG	TGGCACAGGACAGTCCAAG	40
Gsc	GAGACGAAGTACCCAGACGTG	GGCGGTTCTTAAACCAGACC	32
Hex	CTACACGCACGCCCTACTC	CAGAGGTCGCTGGAGGAA	50
Id1	TCCTGCAGCATGTAATCGAC	GGTCCCGACTTCAGACTCC	78
Kdr	CCCCAAATTCCATTATGACAA	CGGCTCTTTCGCTTACTGTT	18
Klf4	CGGGAAGGGAGAAGACACT	GAGTTCCTCACGCCAACG	62
Krt8	AGTTCGCCTCCTTCATTGAC	GCTGCAACAGGCTCCACT	67

Nanog	CCTCCAGCAGATGCAAGAA	GCTTGCACTTCATCCTTTGG	25
Oct4	GTTGGAGAAGGTGGAACCAA	CTCCTTCTGCAGGGCTTTC	95
Otx2	GACCCGGTACCCAGACATC	GCTCTTCGATTCTTAAACCATACC	103
Pdgfra	GTCGTTGACCTGCAGTGGA	CCAGCATGGTGATACCTTTGT	80
Rex1	CAGCTCCTGCACACAGAAGA	ACTGATCCGCAAACACCTG	16
Snail	GTCTGCACGACCTGTGGAA	CAGGAGAATGGCTTCTCACC	71
Sox1	GTGACATCTGCCCCCATC	GAGGCCAGTCTGGTGTCAG	60
T	ACTGGTCTAGCCTCGGAGTG	TTGCTCACAGACCAGAGACTG	27
Tbp	GGGGAGCTGTGATGTGAAGT	CCAGGAAATAATTCTGGCTCA	97
Tcf15	GTGTAAGGACCGGAGGACAA	GATGGCTAGATGGGTCCTTG	104
Vimentin	CCAACCTTTTCTTCCCTGAA	TGAGTGGGTGTCAACCAGAG	109
Wnt3a	AATGGTCTCTCGGGAGTTTG	CTTGAGGTGCATGTGACTGG	53
Zeb1	GCCAGCAGTCATGATGAAAA	TATCACAATACGGGCAGGTG	48
Zeb2	CAAGAGGCGCAAACAAGC	TGCGTCCACTACGTTGTCAT	79
Zfp521	AAGCAAGCGAAACCGAGAT	TTCTGGCCTCTTCTTGCAGT	16

2.1.3.3 Other PCR primer sequences

Target	Forward primer	Reverse primer	Use in thesis
Flag-Id1	TGCTGGTTGTTGTGCTGT		Figure 4.3
Id1	TCCTGCAGCATGTAATCGAC	GGTCCCGACTTCAGACTCC	Figure 3.1
Id1-002	GAACGGCGAGATCAGTGC	TCGGTTTCCCACTGAAGAAT	Figure 3.1
Id1 multiple integration probe	TTGGGTGGTGTGACTTTGAG	AACAGAACCCAAACTCCACC	Figure 3.5
Id1 targeting probe	CTGGCTGCCGTAGTGAAAGG	AGTCTGTCCCTAGTGCATGG	Figures 3.5, 3.7
mKate2	CTAGCTCGAGGCCGCCACCATGGTGAGCGAGCTGATTAA	CATGGGCGCCTCTGTGCCCCAGTTTGCTAG	Section 2.2.4.1.7
Nanog multiple integration probe	CATGGGCGCCTCTGTGCCCCAGTTTGCTAG	GAGTCCCTGGATTTGATCCC	Figure 3.14
Nanog targeting probe	AAATTTATTTCTGGGCATGTGGTGAC	GGATCCAGATTTCAGGATTGGAGG	Figure 3.14

2.1.4 Antibodies

2.1.4.1 Primary Antibodies

Target	Supplier	Antibody name	Raised in	Dilution	Catalogue number
β-tubulin	Sigma	AA2	Mouse	1:4000	T8328
Cdh1	Invitrogen	ECCD-2	Rat	1:200	13-1900
Cdh2	BD	32/N-cadherin	Mouse	1:200	610920
Esrrb	Perseus Proteomics	H6705	Mouse	1:200	PP-H6705-00
Flag (HRP conjugated)	Sigma	M2-Peroxidase	Mouse	1:2000	A8592
GFP	Molecular Probes	anti-GFP, chicken IgY fraction	Chicken	1:1000	A10262
GFP	Molecular Probes	anti-GFP, rabbit IgG fraction	Rabbit	1:1000	A11122
Id1	Chemicon	7D4.2	Mouse	1:200	MAB4372
Id1	Santa Cruz	C-20	Rabbit	1:200	sc-488
Klf4	R&D	klf4aa11-483	Goat	1:200	AF3158
Nanog	eBiosciences	eBIOMLC-51	Rat	1:200	14-5761-80
Nkx2.5	Abcam	Anti-Nkx2.5	Rabbit	1:200	ab35842
Oct4	Santa Cruz	C-10	Mouse	1:200	sc-5279
T	Santa Cruz	C-19	Goat	1:200	sc-17745
tagRFP	Evrogen	anti-tRFP	Rabbit	1:1000	AB233
Tcfap2a	DSHB	AP-2 alpha	Mouse	1:10	3B5
Tubb3	Covance	Tuj1	Mouse	1:1500	MMS-435P

2.1.4.2 Secondary antibodies

Alexa Fluor® antibodies (Molecular Probes) conjugated to appropriate fluorophores were used as secondary antibodies for immunostaining experiments.

Target species	Raised in	Fluorophore	Catalogue number
Chicken	Goat	FITC	A16055
Goat	Donkey	568	A11057
Goat	Donkey	647	A21447
Mouse	Donkey	488	A21202
Mouse	Donkey	568	A10037
Mouse	Donkey	647	A31571
Rabbit	Donkey	488	A21206
Rabbit	Donkey	568	A10042
Rat	Donkey	488	A21208
Rat	Goat	568	A11077
Rat	Donkey	594	A21209

For detection of primary antibodies in Western blots the following two secondary antibodies were used: ECL Mouse IgG, HRP-linked whole Ab (from sheep) (GE Healthcare, catalogue number NA931-1ML); ECL Rabbit IgG, HRP-linked whole Ab (from donkey) (GE Healthcare, catalogue number NA934-100UL). Both antibodies were used at a 1:4000 dilution.

2.1.5 Formulation of solutions

Denaturing solution: 87.66g/l sodium chloride, 20g/l sodium hydroxide in water.

DNA loading dye: Orange G (Sigma) 0.25%, glycerol (Fisher) 30%, in water.

Genomic DNA Lysis Buffer composition: Tris Base 100mM pH8.5, EDTA 5mM, NaCl 200mM, 0.2% SDS, in water.

Lysogeny broth (LB): 2.5% LB powder (BD) in water. This formulation of LB contains 10g/l NaCl. This solution was prepared and autoclaved by the Institute Wash Staff.

Neutralising solution: 116.8g/l sodium chloride, 121.1g/l Tris base, pH8.0 in water.

N2 (homemade): 27.5ml DMEM/F12 (Gibco), 4ml sterile bovine serum albumin (75mg/ml in PBS, Gibco), 4ml insulin (25mg/ml in 0.01M HCl, Sigma), 4ml apo-transferrin (100mg/ml in H_2O, Sigma), 40µl sodium selenite (3mM in H_2O, Sigma), 400µl putrescine (160mg/ml in H_2O, Sigma), 132µl progesterone (0.6mg//ml in ethanol, Sigma). This medium supplement was prepared in our Institute by the Transgenics Service.

PBS (not for cell culture): 5 PBS tablets/l (Sigma) in water.

SSC 20X solution: 175.2g/l NaCl, 88.2g/l Tris Base, pH7.4 in water.

TAE: Tris Base (Fisher) 40mM, acetic acid 20mM, EDTA 1mM, pH 8.4, in water.

TBE: Tris Base (Fisher) 22.5mM, orthoboric acid (Fisher) 22.5mM, EDTA 500µM pH 8.0, in water.

TE: 10mM Tris Base, 1mM EDTA, pH8.0, in water.

2.2 Methods

2.2.1 DNA Methods

2.2.1.1 DNA cloning methods

2.2.1.1.1 Restriction enzyme digestion

DNA digestions were performed by digesting 1-100 µg DNA with 5-100 units of the appropriate restriction enzymes, using the buffers and digestion conditions recommended by the enzyme manufacturers (New England Biolabs, Roche). Digestions were typically performed for 1-3 hours at 37ºC.

2.2.1.1.2 Preparation of electrophoretic gels

To generate electrophoretic gels, agarose was mixed with TBE and heated with a microwave until the agarose powder was fully melted. The solution was allowed to cool but not to solidify, and SYBR® Safe DNA Gel Stain (Invitrogen) was added at a 1:10000 dilution. The solution was mixed, poured into a gel mould and allowed to cool to room temperature and solidify. The concentration of agarose mixed with TBE varied from 0.8%-1.5% (weight/volume), depending on the size of the DNA fragments to be visualised.

2.2.1.1.3 Gel electrophoresis

The samples to be analysed were mixed with DNA loading dye (dye added at a 1:6 dilution), loaded into a well of the prepared electrophoresis gel submerged in TBE, and subject to electrophoresis at 80-100V for 30-90 minutes, depending on the size of the DNA to be visualised. A DNA ladder was loaded next to the samples for size comparison (1 kb DNA Ladder or 100 bp DNA Ladder, New England Biolabs).

After completion of the run, the gel was imaged under an ultraviolet transilluminator (G:BOX F3, Syngene) to verify the correct size of the DNA fragment.

2.2.1.1.4 Purification of DNA from agarose gels

If a DNA fragment was to be purified for further use, only one tenth of the restriction digestion reaction was run on an agarose gel for verification of the correct size under an ultraviolet transilluminator, as described above. Following this step, the remainder of the reaction was run on a separate agarose gel and briefly illuminated under an ultraviolet transilluminator. The DNA band of interest was excised from the gel using a sterile blade, and the gel fragment was transferred to a microcentrifuge tube. The length of the ultraviolet transillumination was kept to under 3 seconds to avoid the induction of mutations in the DNA sequence to be purified. The excised band was then purified using the QIAquick Gel Extraction Kit (Qiagen) following the manufacturer's instructions.

2.2.1.1.5 Quantification of DNA concentration

The concentration of DNA in solution was measured using a NanoVue Plus Spectrophotometer (GE Healthcare) according to the manufacturer's instructions.

2.2.1.1.6 Dephosphorylation of linear DNA ends

The 5' phosphate groups of linear DNA were removed using rAPid Alkaline Phosphatase (Roche) according to the manufacturer's instructions. Briefly, linear DNA was incubated in 1X rAPid Alkaline Phosphatase Buffer and 1U rAPid Alkaline Phosphatase for 10-30 minutes at 37ºC. The enzyme was then heat inactivated at 75ºC for 2 minutes.

2.2.1.1.7 DNA fragment ligation

Linear DNA fragments were ligated using T4 DNA Ligase (New England Biolabs). Ligation reactions were performed in a 10µl reaction volume containing 1X T4 DNA Ligase Reaction Buffer and 400U T4 DNA Ligase. A 3:1 molar ratio of insert:vector was used. Reactions were performed at 37ºC for 1 hour or at 4ºC overnight.

2.2.1.1.8 Preparation of selective bacterial culture plates

LB-agar was prepared by the Institute Wash Staff, by dissolving 1.5% Bacto™ Agar (BD) in LB. Ampicillin 1000X stocks were prepared by dissolving 100mg ampicillin powder (Calbiochem) in 1ml DNase/RNase-free water (Gibco) and filtering the solution through a 0.22µm filter (Millipore). Kanamycin 1000X stocks were prepared by dissolving 50mg kanamycin powder (Calbiochem) in 1ml DNase/RNase-free water (Gibco) and filtering the solution through a 0.22µm filter (Millipore). Antibiotic stock aliquots were kept at -20ºC, and upon thawing an aliquot was kept at 4ºC for no longer than 4 weeks.

LB-agar was melted by microwaving the mixture and allowed to cool but not to solidify. The appropriate antibiotic was added at a 1:1000 dilution, and 15ml of LB-agar/antibiotic mixture were pipetted into 90mm Standard Petri Dishes (Sterilin). The LB-agar was allowed to solidify at room temperature prior to use. Unused plates were conserved at 4ºC for no longer than 4 weeks.

2.2.1.1.9 Plasmid transformation into DH5α bacteria

Plasmids were transformed into chemically competent DH5α *Escherichia coli* (Invitrogen) according to the distributor's instructions. Briefly, fewer than 200ng of DNA, in a solution of 10µl or less, were added to a tube containing 50µl of DH5α bacteria on ice. The mixture was incubated on ice for 30 minutes, heated to 42ºC for

20 seconds in a water bath, then placed on ice for a further 2 minutes. 950µl of LB were added to the mixture, and the tube was incubated at 37ºC for 1 hour, shaking at 225rpm. 100µl of the mixture were plated onto selective LB plates overnight at 37ºC.

2.2.1.1.10 Plasmid purification from bacteria

To purify 10-25µg of plasmid DNA, I picked one colony from the transformation plates into 5 ml LB, supplemented with the appropriate antibiotics, and incubated the liquid culture overnight at 37ºC, shaking at 225rpm.
I then purified DNA using the QIAprep Spin Miniprep Kit (Qiagen), according to the manufacturer's instructions.

To purify 100-500µg of plasmid DNA, I picked one colony from the transformation plates into 1 ml LB, supplemented with the appropriate antibiotics, and incubated the liquid culture for 8 hours at 37ºC, shaking at 225rpm. I added 100-500µl of the culture to 100-500ml LB, supplemented with the appropriate antibiotics, and incubated the liquid culture overnight at 37ºC, shaking at 225rpm.
I then purified DNA using the Qiagen Plasmid Maxi Kit, according to the manufacturer's instructions.

2.2.1.1.11 Plasmid sequencing

Plasmid sequencing was performed by the GenePool service at The University of Edinburgh, making use of the BigDye® Terminator Cycle Sequencing technology (Invitrogen). I supplied the GenePool service with 6µl solutions containing 200-500ng plasmid DNA and 1µl of a 3.2µM primer solution. Sequencing reads were analysed using the ApE software.

2.2.1.2 PCR methods

2.2.1.2.1 Taq PCR

Taq DNA Polymerase was used to efficiently amplify DNA in those instances in which high fidelity of amplification was not required.

PCR's was performed according to the manufacturer's instructions in a reaction volume of 50µl. The components of the reactions were as follows:

Reagent	**Volume/reaction**	**Final concentration**
10X PCR Buffer	5µl	1X
25mM $MgCl_2$	0-2µl	1.5-2.5mM
dNTP mix (10mM each)	1µl	200µM each
Forward primer (10µM)	2µl	400nM
Reverse primer (10µM)	2µl	400nM
Taq DNA polymerase	0.25-0.5µl	1.25-2.5 units/reaction
Template DNA	Variable	Variable
Water	To 50µl	-

Thermal cycling was performed using a TProfessional Standard Thermocycler (Biometra), with the following cycle conditions:

Step	**Temperature**	**Duration (minutes:seconds)**
Denaturation	94ºC	03:00
35 cycles of		
Denaturation	94ºC	01:00
Annealing	55-58ºC	00:30
Extension	72ºC	01:00 (for amplicons longer than 1kb, 1 minute/kb)
Following the 35 cycles		
Extension	72ºC	10:00

Annealing temperatures were optimised for each primer pair; for the experiments reported in this thesis they lied within the range described above.

2.2.1.2.2 Platinum® Pfx PCR

Platinum® Pfx DNA Polymerase was used to amplify DNA in those instances in which high fidelity of amplification was required.

PCR's was performed according to the manufacturer's instructions in a reaction volume of 50µl. The components of the reactions were as follows:

Reagent	**Volume/reaction**	**Final concentration**
10X *Pfx* Amplification Buffer	5µl	1X
50mM $MgSO_4$	1-2µl	1-2mM
dNTP mix (10mM each)	1.5µl	300µM each
Forward primer (10µM)	1.5µl	300nM
Reverse primer (10µM)	1.5µl	300nM
Platinum® Pfx DNA polymerase	0.4µl	1 unit/reaction
Template DNA	Variable	Variable
Water	To 50µl	-

Thermal cycling was performed using a TProfessional Standard Thermocycler (Biometra), with the following cycle conditions:

Step	Temperature	Duration (minutes:seconds)
Denaturation	94ºC	05:00
35 cycles of		
Denaturation	94ºC	00:15
Annealing	55-58ºC	00:30
Extension	68ºC	01:00 (for amplicons longer than 1kb, 1 minute/kb)
Following the 35 cycles		
Extension	68ºC	05:00

Annealing temperatures were optimised for each primer pair; for the experiments reported in this thesis they lied within the range described above.

2.2.1.2.3 Quantitative real-time PCR

Quantitative real-time PCR was performed using the Roche LightCycler® 480 Real-Time PCR System. This system makes use of *Taq* DNA polymerase to amplify short DNA amplicons, and uses fluorescent molecules to measure the number of DNA molecules present after each cycle of amplification. Whenever possible, I made use of the Universal ProbeLibrary (UPL) system, in which a short nucleic acid probe coupled to a fluorophore and a quencher specifically binds the DNA region to be amplified. Upon amplification of the DNA, the fluorophore is separated from the quencher and fluorescence can be detected by the LightCycler® 480 instrument. In alternative, I made use of the SYBR® Green I system, in which all DNA fluoresces upon binding the SYBR® Green I dye. Melting curves were analysed to ensure amplification of a single species of DNA.

Primers used for qRT-PCR were designed using the online Roche Universal ProbeLibrary Assay Design Centre. The primer binding sequences were verified on the Ensembl online database to ensure the absence of single nucleotide polymorphisms. The primers were tested by performing *Taq* PCR on a cDNA sample and verifying the presence of a single amplicon on an electrophoretic gel. The amplicon generated by each primer pair was cloned into a pGem®-T Easy backbone (Promega), to be used as a serial dilution in qRT-PCR experiments.

qRT-PCR experiments were carried out in 384-well plates, using a 10µl reaction volume.

UPL reaction components:

Reagent	**Volume/reaction**	**Final concentration**
LightCycler® 480 Probes Master	5µl	1X
Water	1.5µl	-
Forward primer (10µM)	0.45µl	450nM
Reverse primer (10µM)	0.45µl	450nM
Universal ProbeLibrary Probe (10µM)	0.1µl	100nM
Template cDNA	2.5µl	-

UPL cycling conditions:

Step	**Temperature**	**Duration (minutes:seconds)**
Denaturation	95ºC	05:00
45 cycles of		
Denaturation	95ºC	00:05
Annealing	60ºC	00:10
Extension	72ºC	00:01
Following the 45 cycles		
Cooling	40ºC	00:10

Fluorescence readings were taken after each extension step.

SYBR® Green I reaction components:

Reagent	Volume/reaction	Final concentration
SYBR® Green I Master	5µl	1X
Forward primer (10µM)	1µl	1µM
Reverse primer (10µM)	1µl	1µM
Template cDNA	3µl	-

SYBR® Green I cycling conditions:

Step	Temperature	Duration (minutes:seconds)
Denaturation	95ºC	05:00
45 cycles of		
Denaturation	95ºC	00:05
Annealing	58ºC	00:10
Extension	72ºC	00:20
Acquisition	81ºC	00:01

Following the 45 cycles, a melting curve was generated in order to observe whether amplification of multiple DNA sequences had occurred. The machine settings were as follows.

Temperature	Ramp rate	Time at indicated temperature (minutes:seconds)	Acquisition
95ºC	4.8ºC/s	00:01	None
65ºC	2.5ºC/s	00:10	None
95ºC	0.11ºC/s		Continuous (5 acquisitions/ ºC)

The number of molecules of the amplicon of interest in each reaction was calculated by the LightCycler® 480 Software, referring to a serial dilution of the pGem®-T Easy/amplicon plasmids described above, loaded in the same qRT-PCR plate at known concentrations (6 10-fold dilutions of plasmid at an initial copy number of 10^8/reaction). 3 replicate reactions were loaded on the qRT-PCR plates for each cDNA sample. To calculate the normalised expression value for a gene of interest for each biological sample, the average expression value of the gene of interest in the 3 replicate reactions was divided by the average expression value of the housekeeping gene *Tbp* in 3 replicate reactions:

If X_1, X_2, X_3 are the expression values of 3 replicate reactions for gene *X*, and T_1, T_2, T_3 are the expression values of 3 replicate reactions for *Tbp*,

Gene expression value in sample = $[(X_1+X_2+X_3) / 3] / [(T_1+T_2+T_{3)} / 3]$.

To calculate the biological average expression value for a gene of interest in biological replicate samples, the gene expression values in each sample, calculated as above, were averaged across biological replicates.

2.2.1.2.4 Single-cell quantitative real-time PCR

Single-cell quantitative real-time PCR was performed using the UPL system as described above, with some modifications. Experiments were carried out in 384-well plates, using an 8µl reaction volume.

Reaction components:

Reagent	Volume/reaction	Final concentration
LightCycler® 480 Probes Master	4µl	1X
Water	1.2µl	-
Forward primer (10µM)	0.36µl	450nM
Reverse primer (10µM)	0.36µl	450nM
Universal ProbeLibrary Probe (10µM)	0.08µl	100nM
Template cDNA	2µl	-

Cycling conditions:

Step	Temperature	Duration (minutes:seconds)
Heating	50ºC	02:00
Denaturation	95ºC	10:00
45 cycles of		
Denaturation	95ºC	00:15
Annealing & Extension	60ºC	01:00
Following the 45 cycles		
Cooling	40ºC	00:10

Fluorescence readings were taken after each extension step.
These cycling conditions were optimised for the LightCycler® 480 by Dr. Kumiko Iwabuchi, an expert in single-cell transcriptional analysis from Dr. Keisuke Kaji's laboratory.

The LightCycler® 480 Software was used to calculate the gene expression levels for each gene of interest based on a serial dilution of cDNA extracted from 32 cells and subject to pre-amplification alongside the analysed samples (10 2-fold serial dilutions). The *Tbp* level of the 1-cell equivalent of the serial dilution (i.e. the 1:32

dilution of the 32-cell cDNA sample) was used to select single-cell samples to be analysed. Only cells with a *Tbp* expression value falling within a 0.5-1.5-fold range of this *Tbp* expression value were subject to qRT-PCR for a panel of genes of interest.
2 replicate reactions were loaded on the qRT-PCR plates for each cDNA sample, in order to minimise the use of single-cell cDNA. *Tbp*-normalised gene expression values were calculated as described in the previous section (2.2.1.2.3).

Spearman's rank correlation coefficients for *Tbp*-normalised gene expression values were calculated with the R software, using the rcorr() function in the Hmisc package. This function also calculates the approximated asymptotic p-values for the correlations using the *t*-distribution.

2.2.1.3 DNA isolation from mouse embryonic stem cells

2.2.1.3.1 DNA isolation from 24-well plates

Mouse ES cells were grown to confluency in 24-well plates. They were washed twice in PBS, then lysed in 500µl Genomic DNA Lysis Buffer supplemented with 200µg/ml Proteinase K. The lysates were transferred to microcentrifuge tubes and incubated overnight at 55ºC shaking at 300rpm. The DNA was then precipitated by adding 1 volume of isopropanol (Acros Organics) and centrifuged. Excess salts were removed with two washes in 70% ethanol (VWR). The pellets were air-dried and resuspended in 500µl of water.

2.2.1.3.2 DNA isolation from 6-well plates

Mouse ES cells were grown to confluency in 6-well plates. Genomic DNA was then extracted using the Qiagen DNeasy® Blood and Tissue Kit, following the

manufacturer's instructions for animal blood or cultured cells. Following a first elution step, the eluate was passed through the column again to maximise DNA concentration and yield.

2.2.1.4 Southern Blotting

2.2.1.4.1 Genomic DNA digestion

5µg of genomic DNA extracted from mouse ES cells were digested overnight at 37ºC, in 100U of high-concentration restriction enzymes, in the appropriate digestion buffer, in the presence of 100µg/ml bovine serum albumin (New England Biolabs) and 2.5mM spermidine (Sigma). The DNA was then precipitated by adding 1 volume of isopropanol and centrifuged. Excess salts were removed with two 70% ethanol washes. The pellet was air-dried and resuspended in 20µl of water.

2.2.1.4.2 Genomic DNA electrophoresis and denaturation

The digested DNA was loaded on a 0.8% (w/v) agarose/TAE gel. λ DNA-HindIII digest (New England Biolabs) was loaded as a size marker. Electrophoresis was performed overnight at 10-25V, depending on the size of the electrophoretic gel and of the DNA fragments to be detected. The gel was stained with SYBR® Safe (Invitrogen) at a 1:10000 dilution in TAE for 15 minutes, then imaged under an ultraviolet transilluminator to verify the complete digestion of the DNA samples. The gel was then incubated in Denaturing Solution for 40 minutes on a rocking platform, rinsed with distilled water, and incubated in Neutralising Solution for 40 minutes on a rocking platform.

2.2.1.4.3 Transfer to nylon membrane

DNA was then transferred onto a positively charged nylon membrane (Roche) making use of the transfer strategy depicted in Figure 2.1.

The transfer was carried out for 48 hours, after which the nylon membrane was rinsed in 2X SSC and baked for 1 hour at 120ºC.

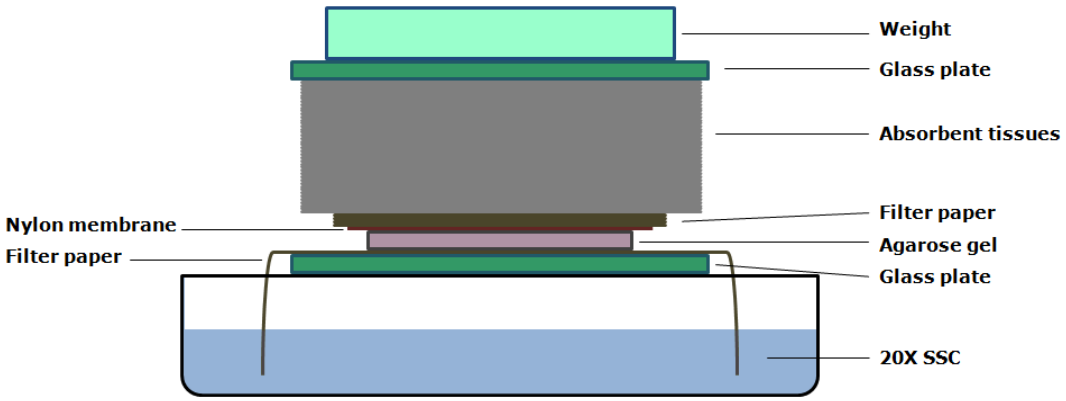


Figure 2.1 – Southern blot transfer strategy

Diagram of the strategy used to transfer DNA from an agarose gel to a nylon membrane for Southern blotting applications. A plastic tank was filled with 20X SSC. A glass plate (green) was placed upon the plastic tank. Filter paper (brown) was soaked in 20X SSC and wrapped around the glass pate, leaving the extremities of the paper bathed in 20X SSC, to enable the solution to maintain the paper wet by capillary action. The agarose gel containing the digested genomic DNA (pink) was placed on top of the filter paper, the nylon membrane (red) was pre-soaked in 20X SSC and rested on top of the gel. Four sheets of filter paper (brown) were soaked in 20X SSC and placed on top of the nylon membrane. Absorbent tissue (grey) was placed on top of the filter paper. A second glass plate (green) was placed on top of the filter paper, and a weight (light blue) was balanced on top of the glass plate.

2.2.1.4.4 Generation of probes for Southern blot

The probes for Southern blotting were designed to be between 400-600 nucleotides in size and to be present in a single genomic location. The probes were generated by *Taq* PCR using wild-type E14tg2a ES cell genomic DNA as a template, purified by gel extraction, subcloned into a pGem®-T Easy backbone and sequenced. Probe fragments were generated by EcoRI digestion-mediated excision of the insert from the pGem®-T Easy backbone, followed by gel extraction.

Probe labelling was carried out using the Amersham Rediprime II DNA Labeling System (GE Healthcare) according to the manufacturer's instructions. Briefly, 25ng probe DNA were diluted in a total volume of 45µl of TE buffer in a screw cap tube. The tube was incubated at 100ºC for 5 minutes to denature the DNA, then placed on ice for 5 minutes. The DNA was then added to a reaction tube (containing dried dATP, dGTP, dTTP, random primers and exonuclease free Klenow enzyme), alongside 5µl [α-^{32}P]dCTP (3000Ci/mmol). The reaction was incubated at 37ºC for 10 minutes. In order to remove unincorporated nucleotides, the reaction contents were then added sequentially to two Illustra ProbeQuant G-50 Micro Columns (GE Healthcare) and centrifuged. The eluate was incubated at 100ºC for 5 minutes and chilled on ice for 5 minutes prior to hybridisation to the membrane.
The same protocol was used to label a probe to detect the λ DNA-HindIII digest, using only 2µl [α-^{32}P]dCTP (3000Ci/mmol), and 3µl of non-radioactive dCTP.

2.2.1.4.5 Probe hybridisation to the membrane

The baked membrane was rinsed with 2X SSC and placed into a glass hybridisation bottle. It was incubated with 20ml PerfectHyb™ Plus hybridisation buffer (Sigma) at 65ºC for a minimum of 1 hour. 500µl sonicated herring sperm DNA (Sigma) were incubated at 100ºC for 5 minutes and cooled on ice for 5 minutes. They were added

to the hybridisation bottle alongside the labelled probes of interest. The hybridisation was carried out overnight at 65ºC.

2.2.1.4.6 Membrane washing and film exposure

The hybridisation solution was removed and the membrane was washed twice for 15 minutes in 2X SCC+0.1% SDS, and once for 30 minutes in 0.5X SSC+0.1% SDS. The membrane was wrapped in cling film and placed in a film developing cassette, where it was used to expose Amersham Hyperfilm ECL (GE Healthcare). Exposure time varied between 1-7 days at -80ºC. The film was developed in the dark using a film developer.

2.2.2 RNA Methods

2.2.2.1 Total RNA isolation

Total RNA was extracted from cultured cells using the Absolutely RNA Miniprep Kit (Agilent Technologies), following the manufacturer's instructions.
The concentration of the isolated RNA was measured using a NanoVue Plus Spectrophotometer (GE Healthcare) according to the manufacturer's instructions.

2.2.2.2 cDNA synthesis

First-strand cDNA synthesis was performed using M-MLV Reverse Transcriptase (Invitrogen), following the manufacturer's instructions. Briefly, 1µg RNA was mixed with 1µl 10mM dNTP Mix (Invitrogen), 50ng Random Primers (Invitrogen) in a 12µl reaction volume. The mixture was heated to 65ºC for 5 minutes and cooled on ice for 1 minute. 4µl 5X First-Strand Buffer, 2µl 0.1M DTT and 1µl RNaseOUT™

Recombinant Ribonuclease Inhibitor (40U/µl) (Invitrogen) were added. The mixture was incubated at 37ºC for 2 minutes. 1µl (200U) M-MLV Reverse Transcriptase was added, and the reaction was incubated at 25ºC for 10 minutes. Reverse transcription was performed at 37ºC for 50 minutes, and was followed by heat inactivation of the enzymes at 70ºC for 15 minutes. To digest the RNA template, 1µl (5U) RNase H (New England Biolabs) was added to the mixture, which was incubated at 37ºC for 20 minutes. The cDNA was diluted 1:5 in DNase/RNase-free water (Gibco), and stored at -20ºC.

2.2.2.3 cDNA generation from single cells

RNA was extracted from single cells and reverse transcribed into cDNA using the CellsDirect™ One-Step RT-PCR Kit (Invitrogen), making use of the protocol described by Dalerba *et al.* (2011) with minor modifications. Each well of a 96-well PCR plate was preloaded on ice with 5µl 2X Reaction Mix, 0.2µl Superscript® III RT/Platinum® *Taq* Mix (with RNaseOUT™ Ribonuclease Inhibitor), 2.5µl primer mix (containing 200nM of each gene-specific primer to be used in later qRT-PCR experiments), 1.3µl DNase/RNase-free water (Gibco). Live single cells were sorted into separate wells of the PCR plate using a FACSAria cell sorter (BD Biosciences). 32 cells were sorted into one well, to be used for the generation of qRT-PCR standard curves by serial dilution of the generated cDNA (see paragraph 2.2.1.2.4).
The plate was placed in a TProfessional Standard Thermocycler (Biometra). RNA was reverse transcribed to cDNA and the cDNA was subject to 22 cycles of pre-amplification of target amplicons (to ensure sufficient abundance of the amplicons of interest for qRT-PCR), as described by Dalerba *et al.* (2011). Briefly, the reactions were incubated at 50ºC for 15 minutes to lyse the cells and reverse transcribe the cDNA, the Superscript® III Reverse Transcriptase was inactivated and the Platinum® *Taq* was activated by heating the reactions to 95ºC, and pre-amplification

was performed with cycles of 95ºC for 15 seconds and 60ºC for 4 minutes. The cDNA was diluted 1:5 with DNase/RNase-free water and stored at -20ºC

2.2.3 Protein Methods

2.2.3.1 Western Blotting

2.2.3.1.1 Protein isolation from cultured cells

Cultured cells were lysed in RIPA buffer (Sigma) supplemented with 0.5mM Pefabloc® SC (Sigma), 1mM DTT (Roche) and 1X Protease Inhibitor Cocktail (Roche cOmplete ULTRA Tablets). Cells were allowed to lyse on ice for 30 minutes before transferring the lysate to a microcentrifuge tube. 1µl (250U) Benzonase Nuclease (Sigma) was added to the tube and the solution was incubated at room temperature for 10 minutes.
Protein concentration was estimated by measuring the absorbance of the sample at 280nm with a NanoDrop ND-1000 spectrophotometer (Thermo Scientific).

2.2.3.1.2 Sample preparation and gel loading

The samples to be loaded on a Western blot gel were diluted in RIPA buffer to a uniform concentration. 25µl of protein sample were mixed with 12.5µl 4X NuPAGE LDS Sample Buffer (Novex) and with 12.5µl of a 1M DTT aqueous solution. The samples were boiled for 10 minutes and chilled on ice.

A 10-well NuPAGE® 4-12% Bis-Tris Gel (Novex) was loaded in an XCell SureLock™ Mini-Cell electrophoresis tank (Novex), and submerged with 1X NuPAGE® MOPS SDS Running Buffer (Novex).
10µl of each sample were loaded on to the gel alongside 10µl SeeBlue® Pre-Stained Standard (Novex). The gel was subject to electrophoresis at 200V for 50 minutes.

2.2.3.1.3 Protein transfer to nitrocellulose membrane

Transfer buffer was generated as an aqueous solution of 1X NuPAGE® Transfer Buffer (Novex) and 10% methanol (VWR). Amersham Hybond ECL Nitrocellulose Membrane (GE Healthcare) was cut to the size of the gel and soaked in transfer buffer for 10 minutes. Filter paper was cut to the size of the gel and soaked in transfer buffer for 10 minutes, alongside 4 blotting pads of the same size as the gel.
The gel was removed from the plastic case it was lodged in and the protein transfer was set up in an XCell II™ Blot Module (Novex), following the manufacturer's instructions. Briefly, the components for the transfer were assembled in the blotting module in the following order: 2 blotting pads, filter paper, gel, membrane, filter paper, 2 blotting pads. The transfer was carried out in the XCell SureLock™ Mini-Cell tank in transfer buffer for 1 hour at 25V.

2.2.3.1.4 Antibody staining

The membrane was transferred to blocking solution, containing 5% (w/v) Amersham ECL Blocking Agent (GE Healthcare) and 0.1% Tween® 20 (Promega) in PBS. Blocking was carried out for 1 hour at room temperature on a rocking platform. The membrane was then incubated in primary antibody diluted in blocking solution for 1 hour at room temperature on a rocking platform. The membrane was subject to three washes of 15 minutes in PBS+-0.1% Tween® 20 at room temperature on a rocking platform. Binding of the appropriate HRP-conjugated secondary antibody and subsequent membrane washing were carried out in the same manner.

2.2.3.1.5 Chemiluminescent film impression

The membrane was placed on cling film. 1ml SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific) was added onto the membrane for 5 minutes. The membrane was wrapped in cling film taking care to avoid the formation of air bubbles, and placed in a Western blot developing cassette. Amersham Hyperfilm ECL (GE Healthcare) was exposed to the chemiluminescent signals generated by the membrane for varying amounts of time, prior to film development.

2.2.3.1.6 Membrane stripping

If required, the primary and secondary antibodies were stripped from the membrane using a stripping buffer composed of 1.5% glycine (Sigma), 0.1% SDS (Sigma), 1% Tween® 20 (Promega) in water at pH 2.2. Stripping was performed at room temperature on a rocking platform for 10 minutes, and was followed by 2 washes in PBS and 2 washes in PBS+0.1% Tween® 20. Each wash was performed at room temperature on a rocking platform for 10 minutes. Further membrane labelling was performed as described in section 2.2.3.1.4.

2.2.3.2 Immunofluorescence

2.2.3.2.1 Sample fixation

Cells cultured with the intent of imaging nuclear proteins at high magnification using the Leica TCS SPE inverted confocal microscope were fixed in PBS+4% PFA (Sigma)+0.1% Triton™ X-100 (Sigma) for 20 minutes at room temperature. The fixative was then neutralised with 50mM NH_4Cl (Acros Organics) in PBS for 5 minutes at room temperature, then subject to three 5-minute washes with PBS.

Cells cultured for imaging with the widefield Olympus IX51 microscope were fixed in PBS+4% PFA for 20 minutes at room temperature, then subject to three 5-minute washes with PBS.

Dissected embryos were fixed in PBS+4% PFA for 2 hours at room temperature. The fixative was then neutralised with 50mM NH_4Cl in PBS for 20 minutes at room temperature. The embryos were then transferred to PBS.

2.2.3.2.2 Antibody stain of cells

Fixed cell samples were incubated in blocking solution for 30 minutes. Blocking solution consisted of PBS+3% donkey serum (Sigma)+0.1% Triton™ X-100. Primary antibodies were diluted in blocking solution and added to the cells for 1-3 hours at room temperature or overnight at 4ºC. The samples were then subject to three 5-minute washes in PBS. Secondary antibodies were diluted 1:1000 in blocking solution and added to the cells for 1 hour at room temperature. The samples were subject to three 5-minute washes in PBS. 100ng/ml DAPI in PBS was added to the samples for 5 minutes. The samples were then subject to three 5-minute washes in PBS.

When the staining was performed on cells cultured in wells, the samples were maintained in PBS prior to imaging. When the staining was performed on cells cultured on coverslips, the coverslips were mounted onto glass slides using a droplet of ProLong® Gold Antifade Reagent (Molecular Probes). The antifade reagent was allowed to cure for 48 hours at 4ºC, then the edges of the coverslips were sealed with transparent nail varnish.

2.2.3.2.3 Whole mount staining of mouse embryos

The embryos were permeabilised in PBS+0.5% Triton™ X-100 for 1 hour at room temperature, after which they were transferred to blocking solution (same recipe as in paragraph 2.2.3.2.2) overnight at 4ºC. Primary antibodies were diluted in blocking solution and added to the embryos for 48 hours at 4ºC. The embryos were subject to five 10-minute washes in PBS+0.1% Triton™ X-100. Secondary antibodies were diluted 1:1000 in blocking solution and added to the embryos for 48 hours at 4ºC. When a DAPI stain was performed, DAPI was added to the secondary antibody stain solution at 100ng/ml. The embryos were subject to five 10-minute washes in PBS+0.1% Triton™ X-100.

2.2.3.2.4 Staining of mouse embryo sections

The slides containing the gelatine-mounted sections were immersed in PBS at 55ºC for 15 minutes to melt the gelatine. The slides were carefully dried with medical wipes, and circles were drawn around the sections with PAP pen. The sections were permeabilised in PBS+0.1% Triton™ X-100 for 10 minutes at room temperature, after which they were stained as described for cells in paragraph 2.2.3.2.2. Upon completion of staining, they were mounted using ProLong® Gold Antifade Reagent. The antifade reagent was allowed to cure for 48 hours at 4ºC, then the edges of the coverslips were sealed with transparent nail varnish.

2.2.3.2.5 Imaging stained samples

Cells and embryo sections were imaged using a widefield Olympus IX51 microscope with a QImaging Retiga-2000R Fast 1394 camera, or using a Leica TCS SPE inverted confocal microscope. The microscope used for each experiment is indicated in the relevant figure legend.

Whole mount embryo stains were imaged using a Leica TCS SPE inverted confocal microscope. The embryos were imaged in a cell culture imaging chamber, submerged in PBS.

Optimal exposure, gain and offset settings were empirically determined for each experiment. Pixel intensity in 8-bit images ranges between 0 and 255 (2^8=256 possible values). In order to adopt the appropriate settings, the sample with the highest expression of a specific factor of interest for each experiment was used to find the combination of gain and exposure that would result in pixel saturation at the highest intensity value (255). The negative control sample was then used with those gain and exposure settings to identify the offset setting that would result in pixel saturation at the lowest intensity value (0). This system ensures that only very few pixels in any given image are saturated at the lowest and highest fluorescent intensity values.

When multiple fluorescent images were acquired for one field of view, the order of fluorophore acquisition was as follows: far red, red, green, blue, phase contrast. This was done to ensure minimal photobleaching of the sample.

All images were acquired in 8-bit grayscale format. The coloured images presented in this thesis were generated by modifying the lookup table (LUT) associated with the pixel intensity values in Fiji (a distribution of ImageJ and Java enriched with multiple plugins) (Schindelin *et al.* 2012). For grayscale, the red, green and blue values associated with the 256 available pixel intensities are always equal (i.e. if a pixel has intensity 100, the red, green and blue values in the lookup table for that pixel will all be equal to 100). If, for example, an 8-bit grayscale image obtained from a green secondary antibody needs to be displayed in green, the red and blue values for each pixel intensity in the lookup table are set by Fiji to 0, so the displayed pixels range from 0 to 255 only in the green channel.

2.2.3.2.6 Three-dimensional rendering of image stacks

Images of embryos obtained as stacks of various focal plains were rendered in three dimensions using the 3D Viewer plugin in Fiji (Schmid *et al.* 2010).

2.2.3.2.7 Nuclear segmentation for immunostaining quantification

In order to perform nuclear immunostaining quantification, an automated software called Nucleus Editor (part of the Farsight Toolkit, available at http://www.farsight-toolkit.org/) was used for nuclear segmentation. Nuclear segmentation is the name of the process by which the nuclei of all cells in an image are identified and given specific coordinates within the image (Lin *et al.* 2003). Nucleus Editor allows to adjust a variety of parameters to empirically determine the best segmentation values for nuclei in different images. For example, the nuclei of cells cultured in 2i are more spherical than the nuclei of cells within the post-implantation epiblast (three-dimensional data not shown, for two-dimensional comparison see Figures 3.12 and 5.2). It also provides the user with the possibility of manually editing segmented images to ensure the highest achievable accuracy of segmentation. This also gives the user the possibility to exclude dividing cells and cells not fully included in the field of view from further analysis. The result of a segmentation on the image of DAPI-stained nuclei of mouse embryonic stem cells is shown in Figure 2.2.

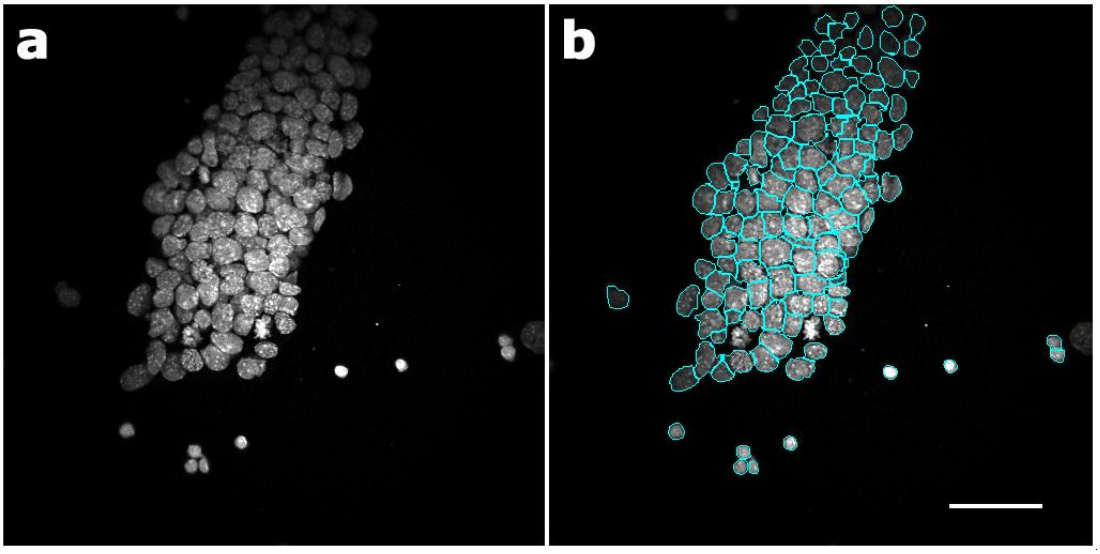


Figure 2.2 – Nuclear segmentation of mouse embryonic stem cells

a. E14tg2a mouse embryonic stem cells cultured in LIF+FCS, fixed and stained with DAPI. The image was acquired on a Leica TCS SPE confocal microscope.

b. Segmentation result for the image in (a) using Nucleus Editor. The automated software result was manually edited to exclude dividing nuclei and nuclei not fully included in the field of view from the final segmentation result. Scale bar: 30µm.

2.2.3.2.8 Nuclear immunostaining quantification using Multicell3D

Having defined the spatial coordinates of each nucleus in a set of pictures of interest, I made use of a software to measure the average intensity of all pixels within each nucleus for every channel. This software, named Multicell3D, was developed by Dr. Guillaume Blin, a member of our laboratory, and has been successfully used for nuclear immunostaining quantification in cultured cells and in mouse embryos (Davies *et al.* 2013, Osorno *et al.* 2012). It confers every nucleus a unique identifier

number and calculates a number of parameters for each nucleus. I made use of the average fluorescence intensity in each analysed channel, and of the volume of the nucleus.

I will present two types of nuclear immunostaining quantification graph within this thesis:

- A direct comparison of the average fluorescence intensity values (0-255) of two channels within single cells, by means of dot plots.
- The analysis of the fluorescence distribution for a single channel across all analysed cells, by means of flow cytometry-like distribution plots.

2.2.3.2.9 Average fluorescence intensity and integrated fluorescence intensity

For flow cytometry-like distribution plots, I will present the intensity values for cells imaged in widefield or as a single confocal plan, and the integrated intensity values for cells imaged in 3D. The integrated intensity equals the fluorescence intensity multiplied by the nuclear volume, and provides an indication of the total amount of protein present within a nucleus. It therefore adjusts for nucleus size:

Let nucleus *x* have a volume $V_x = 100$ pixels, and a number of molecules of a protein of interest $P_x = 100$ molecules. Let each protein molecule have an average fluorescence intensity $I_p = 1$ AFU (arbitrary fluorescence unit)/molecule. The total fluorescence in nucleus *x* is:

$$F_x = P_x * I_p = 100 \text{ molecules} * 1 \text{ AFU/molecule} = 100 \text{ AFU}.$$

The average pixel fluorescence intensity in nucleus *x* is:

$$F_{px} = F_x / V_x = 100 \text{ AFU} / 100 \text{ pixels} = 1 \text{ AFU/pixel}.$$

The integrated fluorescence intensity in nucleus *x* is:

$$IF_x = F_{px} * V_x = F_x / V_x * V_x = F_x = 100 \text{ AFU}.$$

The integrated fluorescence intensity is equal to the total fluorescence in nucleus *x*. It is therefore a representation of the total number of protein molecules in a nucleus, whereas the average fluorescence intensity is an indication of the local concentration of protein molecules in a nucleus.

For comparison, let nucleus *y* have $V_y = 50$ pixels and $P_y = 100$ molecules. The total intensity (and integrated fluorescence intensity) in nucleus *y* is:

$$F_y = IF_y = P_y * I_p = 100 \text{ molecules} * 1 \text{ AFU/molecule} = 100 \text{ AFU}.$$

The average pixel fluorescence intensity in nucleus *y* is:

$$F_{py} = F_y / V_y = 100 \text{ AFU} / 50 \text{ pixels} = 2 \text{ AFU/pixel}.$$

Thus, nuclei *x* and *y* have the same integrated fluorescence intensity (number of protein molecules), but different average fluorescence intensity (concentration of protein molecules).

As mentioned at the start of the paragraph, whenever nuclear volume measurements are available, by means of three-dimensional imaging, I will make use of the integrated intensity value for plotting distributions of a single factor. This will provide a better description of the amount of protein present in each analysed cell within the culture represented in the plot.

2.2.4 Cell culture

2.2.4.1 Cell lines

2.2.4.1.1 Wild-type embryonic stem cell lines

Two wild-type embryonic stem cell lines were used in this study.

E14tg2a ES cells are a clonal cell line derived from ES cells generated from 129/Ola blastocysts. They are hypoxanthine phosphoribosyltransferase deficient (Hooper *et al.* 1987).

E14Ju09 ES cells are a clonal cell line derived by the Transgenics Unit in our Institute from ES cells obtained from chimaeric embryos generated with E14tg2a ES cells. They have a 129/Ola genetic background and are hypoxanthine phosphoribosyltransferase deficient. They have a high propensity for germline colonisation (observations made by the Transgenics Unit staff).

2.2.4.1.2 Generation of transgenic Id1-overexpressing ES cells

Id1 overexpressing ES cells were generated using an Id1 overexpression plasmid used by Ying *et al.* (2003a). The salient features of the plasmid are depicted in Figure 2.3. The plasmid was linearised by digestion with ScaI and ethanol precipitated. Its concentration was quantified and 1µg was lipofected into 10^5 E14tg2a ES cells. The cells were transferred to three 9cm dishes after 24 hours at 1/100, 1/10 and 8.9/10 dilutions. 2µg/ml puromycin were added to the cells after a further 24 hours, and resistant clones were picked and expanded after 1 week of selection.

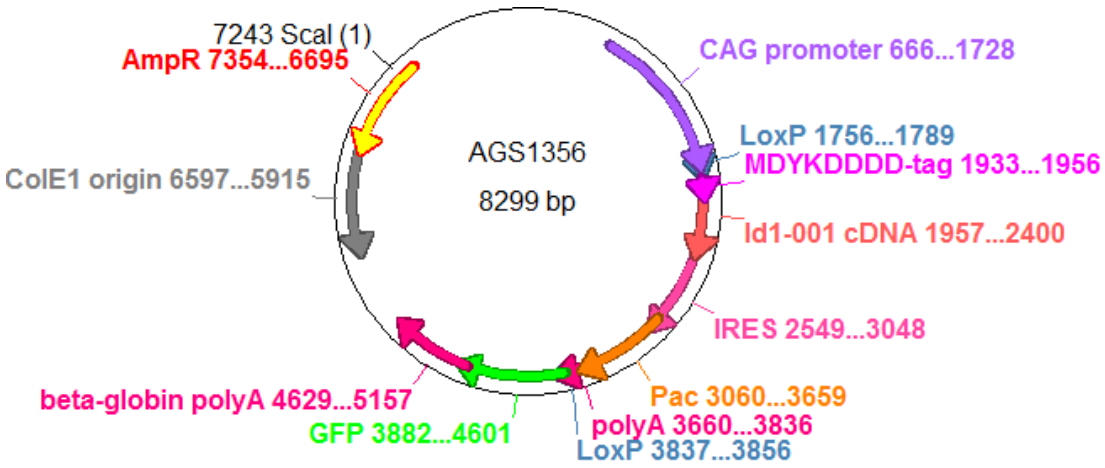


Figure 2.3 – Id1 overexpression plasmid

Structure of the plasmid used for the generation of Id1 overexpressing cells (plasmid code: AGS1356). The strong *CAG* promoter (Niwa *et al.* 1991) drives the expression of the Flag (MDYKDDDD)-tagged cDNA of *Id1*, which is followed by an internal ribosome entry site (*IRES)*, a puromycin resistance gene (*Pac*), and a polyadenylation signal sequence (*polyA*). The *Flag-Id1-IRES-Pac-polyA* sequence is flanked by *loxP* sites in the same orientation, and followed by the cDNA of *GFP* and a second polyadenylation signal. The position of the ScaI restriction site used for the linearisation of the plasmid is indicated.

2.2.4.1.3 Generation of Id1 reverted ES cells

In order to excise the *Flag-Id1-IRES-Pac-polyA* sequence from the genome of clonal Id1 overexpressing ES cell lines, I lipofected 10^5 cells with 3µg of a plasmid containing a *CAG* promoter driving expression of the *Cre* recombinase gene followed by an *IRES* and a puromycin resistance gene (*CAG-Cre-IRES-Pac*, plasmid code: AGS844). The next day I plated the cells at clonal density (10 cells/cm^2) in the absence of selection. After 1 week I observed the clones under the fluorescent microscope and I picked and expanded clones that were uniformly expressing GFP.

2.2.4.1.4 Generation of empty vector control ES cells

ES cells to be used as a control line for experiments involving the Id1 overexpressing cells were generated in parallel to these cell line to minimise variability. The cells were generated using an "empty vector" plasmid, the salient features of the plasmid are depicted in Figure 2.4. The plasmid was linearised by digestion with ScaI and ethanol precipitated. Its concentration was quantified and 1µg was lipofected into 10^5 E14tg2a ES cells. The cells were transferred to three 9cm dishes after 24 hours at 1/100, 1/10 and 8.9/10 dilutions. 2µg/ml puromycin were added to the cells after a further 24 hours, and resistant clones were picked and expanded after 1 week of selection.

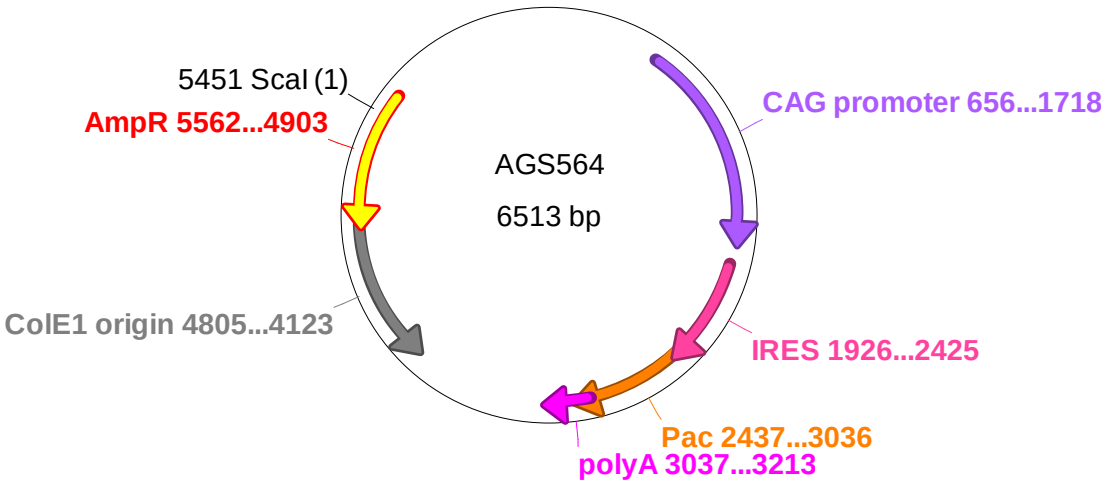


Figure 2.4 – Empty vector control plasmid

Structure of the plasmid used for the generation of Id1 overexpressing cells (plasmid code: AGS564). The *CAG* promoter is directly followed by an internal ribosome entry site (*IRES)*, a puromycin resistance gene (*Pac*), and a polyadenylation signal sequence (*polyA*). The position of the ScaI restriction site used for the linearisation of the plasmid is indicated.

2.2.4.1.5 Generation of Id1-Venus reporter ES cells

Id1-Venus reporter ES cells (named IdV) were generated using an updated version of the plasmid described by Nam & Benezra (2009), kindly gifted to our laboratory by Dr. Robert Benezra and Dr. Hyung-song Nam (Memorial Sloan-Kettering Cancer Center, New York, USA). The salient features of the plasmid are depicted in Figure 2.5. 100µg of the plasmid were linearised by digestion with KpnI, ethanol precipitated and electroporated into 10^7 E14Ju09 ES cells. 10^6 cells were transferred to ten 9cm dishes immediately following electroporation. 300µg/ml G418 were added to the cells after 24 hours, and resistant clones were picked and expanded after 10 days of selection.

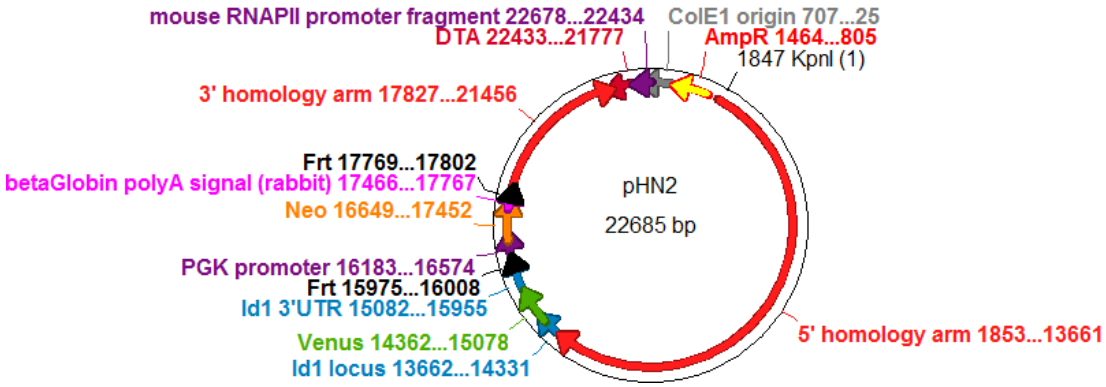


Figure 2.5 – Id1-Venus targeting plasmid

Structure of the plasmid used for the generation of Id1-Venus reporter ES cells (plasmid code: pHN2). The *Id1* open reading frame is followed by sequence coding for a flexible linker and Venus. This sequence is placed before the stop codon of *Id1-001*. *Venus* is followed by the 3'UTR of *Id1* and an excisable selection cassette. The cassette contains a *Pgk* promoter driving expression of *Neo*, a gene conferring resistance to G418, which is followed by a polyadenylation signal sequence. The *Pgk-Neo-polyA* sequence is flanked by *Frt* sites in the same orientation. The 5' and 3' homology arms to the endogenous locus are 12479bp and 3630bp long. Outwith

the regions of homology, an *RNA polymerase II* promoter fragment drives the expression of *DTA*, which encodes diphtheria toxin A, a toxic molecule that should kill cells which express this transgene following non-targeted integration events. The position of the KpnI restriction site used for the linearisation of the plasmid is indicated.

2.2.4.1.6 Excision of the selection cassette from IdV ES cells

In order to excise the *Pgk-Neo-polyA* sequence from the genome of IdV ES cells, I lipofected 10^5 cells with 3µg of a plasmid containing a *Pgk* promoter driving expression of the *FlpO* recombinase gene. The next day I plated the cells at clonal density (10 cells/cm^2) in the absence of selection. After 10 days I picked 96 colonies into a 96-well plate, replica plated them, and cultured one plate in the absence of G418 and the other plate in the presence of 300µg/ml G418. Only 1 colony displayed subsceptibility to G418, suggestive of excision of the selection cassette. I expanded that clonal ES cell line and named it IdV-SC.

2.2.4.1.7 Generation of lineage labelled Id1-Venus reporter ES cells

In order to genetically label IdV ES cells, I generated a plasmid encoding the red fluorescent protein mKate2 (Evrogen) fused to three SV40 nuclear localisation signal (NLS) sequences. mKate2 has minimal spectral overlap with Venus: the excitation and emission peaks of mKate2 are 589nm and 634nm, the excitation and emission peaks of Venus are 515nm and 528nm. The SV40 NLS sequences were used to ensure nuclear localisation of the protein to aid with cell detection in chimaeric embryos, whilst avoiding potential problems to due overexpression of mKate2-tagged histone proteins as an alternative strategy for nuclear localisation of the fluorescent signal.

In order to generate the overexpression plasmids, I amplified the open reading frame of *mKate2* by overlapping extension PCR. A XhoI restriction site followed by a Kozak sequence were added at the 5' end of *mKate2*, and a KasI restriction site was added at the 3' end in place of the stop codon. The amplicon was cloned into a pCR®-Blunt II-TOPO backbone (Invitrogen) and sequence verified. The XhoI-mKate2-KasI insert was purified by digestion of this intermediate plasmid with XhoI and KasI, followed by gel purification of this fragment. The opposite strands of three SV40 NLS sequences were synthesised as oligonucleotides, and were designed to leave overhangs on either end following their annealing. The overhang upstream of the NLS sequences corresponded to the overhang generated by the KasI restriction enzyme (5'-GCGC-3'), the overhang on the minus strand corresponded to that generated by the NotI restriction enzyme (5'-GGCC-3'). The final acceptor vector was digested with XhoI and NotI and the backbone fragment was gel extracted and purified. The backbone, *mKate2* fragment and annealed NLS oligonucleotides were mixed at a 1:1:1 molar ratio and ligated with T4 DNA ligase. The assembled plasmid is depicted in Figure 2.6.

The plasmid was linearised by digestion with ScaI and ethanol precipitated. Its concentration was quantified and 1µg was lipofected into 10^5 IdV ES cells. The cells were transferred to three 9cm dishes after 24 hours at 1/100, 1/10 and 8.9/10 dilutions. 2µg/ml puromycin were added to the cells after a further 24 hours, and 20 resistant clones were picked and expanded after 1 week of selection. I screened the clones for maintenance nuclear mKate2 expression following 5 days of differentiation in culture medium containing FCS without LIF in the absence of selection. I generated frozen stocks of three clones with high transgene expression in all cells after differentiation, and used one of these for the experiments presented in this thesis. This clonal cell line was named IdVnK1, for "IdV nuclear mKate2 clone 1".

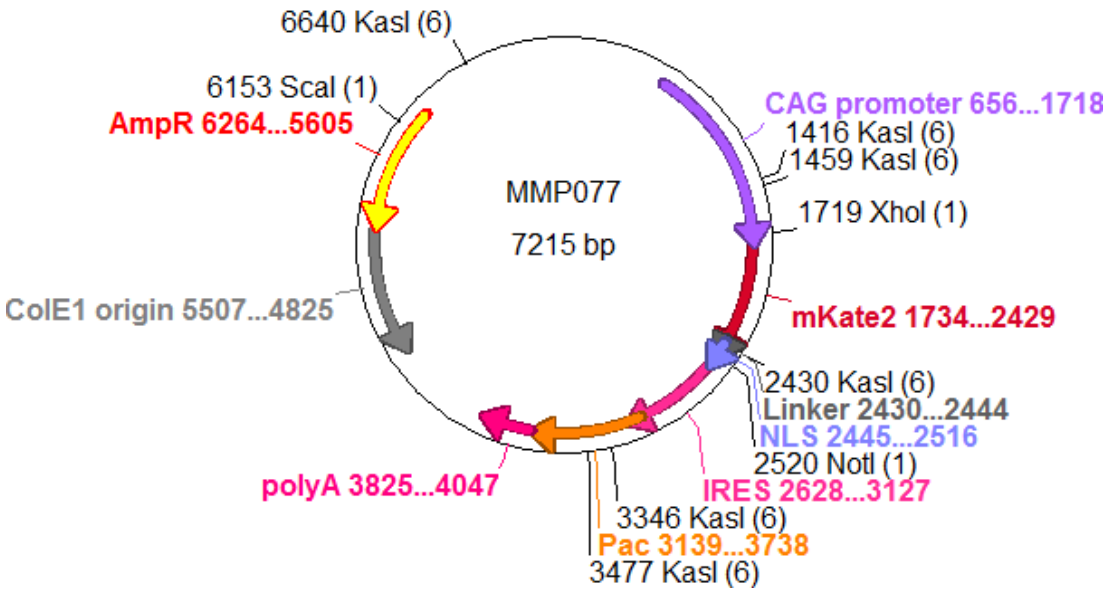


Figure 2.6 – mKate2-NLS lineage labelling plasmid

Structure of the plasmid used for the generation of lineage-labelled IdV ES cells (plasmid code: MMP077). The *CAG* promoter drives the expression of *mKate2-linker-NLS*, which is followed by an internal ribosome entry site (*IRES*), a puromycin resistance gene (*Pac*), and a polyadenylation signal sequence (*polyA*). The position of the ScaI restriction site used for the linearisation of the plasmid and the positions of the restriction sites used for the generation of the plasmid are indicated (see main text).

2.2.4.1.8 Generation of Id1-Venus Nanog-tagRFP double reporter ES cells

In order to generate Id1-Venus Nanog-tagRFP ES cell lines, I made use of a construct generated and kindly gifted to me by Dr. Nicola Festuccia in Dr. Ian Chambers's laboratory. The salient features of the plasmid are depicted in Figure 2.7. 100µg of the plasmid were linearised by digestion with NotI, ethanol precipitated and electroporated into 10^7 IdV ES cells. 10^6 cells were transferred to ten 9cm dishes immediately following electroporation. 10µg/ml blasticidin S were added to the cells

after 24 hours, and resistant clones were picked and expanded after 10 days of selection.

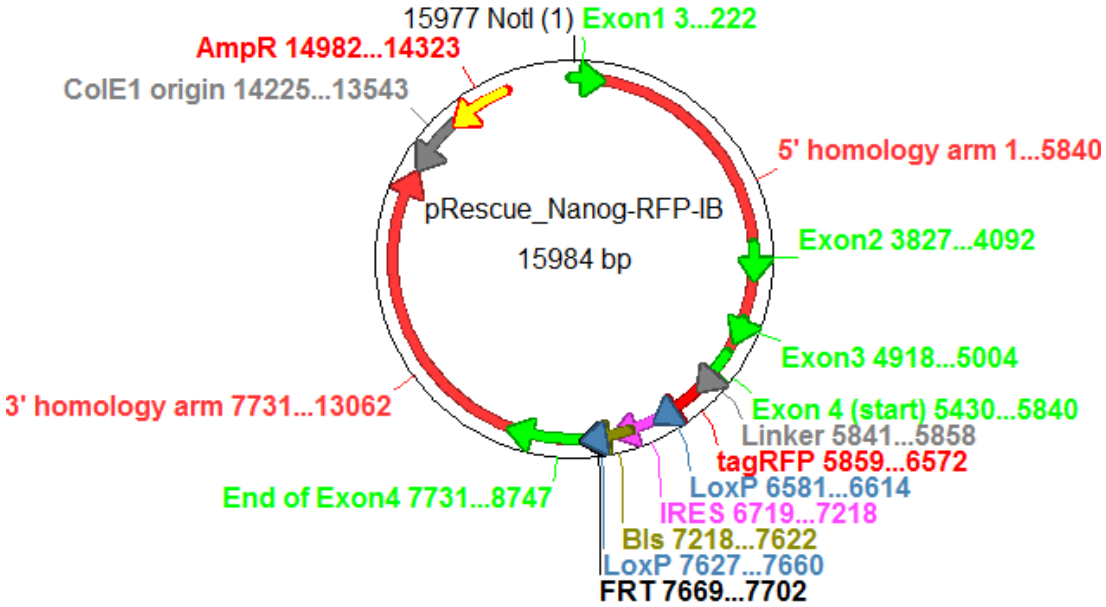


Figure 2.7 – Nanog-tagRFP targeting construct

Structure of the plasmid used for the generation of Nanog-tagRFP reporter ES cells (plasmid name: pRescue_Nanog-RFP-IB). The 5' homology arm (5840bp) is constituted by the four exons and 3 introns of the *Nanog* locus (as indicated in figure). The stop codon is replaced by the DNA encoding for a flexible linker, fused in frame to *tagRFP*. This sequence is followed by an *IRES* and *Bls*, a gene encoding for blasticidin S resistance. The *IRES-Bls* cassette is flanked by *loxP* sites. The 3' homology arm (5331bp) contains the 3' UTR of *Nanog*, including the endogenous polyadenylation signal sequence. The position of the NotI restriction site used for the linearisation of the plasmid is indicated.

2.2.4.2 Culture conditions

All cell culture was performed on Corning cell culture plastics.

2.2.4.2.1 LIF+FCS culture

ES cells were propagated in LIF+FCS culture medium unless otherwise specified. This culture medium is similar to that used by Smith *et al.* (1988) upon characterisation of LIF as a crucial factor for the maintenance of pluripotency. Its composition is as follows:

Component	**Volume**	**Final concentration**
Glasgow Minimum Essential Medium (GMEM, Sigma)	500ml	
Foetal Calf Serum (FCS, Gibco)	51ml	10%
L-Glutamine/Sodium pyruvate solution	11ml	L-Glutamine 2mM Sodium pyruvate 1mM
Non-Essential Amino Acids (NEAA, Gibco)	5.5ml	1X
0.1M 2-mercaptoethanol	570µl	100nM
LIF	570µl	100U/ml

To passage ES cells cultured in LIF+FCS, Cell culture dishes were coated with 0.1% gelatine in PBS for 30 minutes prior to cell passaging. The medium was aspirated from the cells, sterile PBS (Sigma) was used to wash the cells once, and a 0.05% solution of trypsin EDTA (1:5 dilution of 0.25% trypsin EDTA, Gibco, in PBS) was added to the cultures. The cells were incubated at 37°C until detachment (typically 1-3 minutes). 5-10 volumes of LIF+FCS medium were added to the cells to quench the action of trypsin EDTA. The cells were pelleted in a universal tube (Sterilin) at 300g for 3 minutes. The medium was aspirated, the cells were resuspended in fresh

LIF+FCS medium and plated at the desired density on the prepared gelatinised dishes, after aspiration of the 0.1% gelatine solution.
A standard haemocytometer was used to count cell numbers when required.

2.2.4.2.2 N2B27 culture

N2B27 culture for ES cell differentiation was performed as previously described (Pollard *et al.* 2006, Ying *et al.* 2003a, Ying *et al.* 2003b).

Briefly, ES cells cultured in LIF+FCS were trypsinised, quenched, pelleted, resuspended in N2B27 medium, pelleted again, resuspended in N2B27 medium and plated at the desired concentration for differentiation. The composition of N2B27 medium is as follows:

Component	**Volume**	**Final concentration**
DMEM/F12 (Gibco)	50ml	
Neurobasal (Gibco)	50ml	
L-Glutamine	1ml	2mM
0.1M 2-mercaptoethanol	100µl	100nM
N2 (homemade)	500µl	
B27 (Gibco)	1ml	0.5X

2.2.4.2.3 2i and LIF+BMP4 culture

2i and LIF+BMP4 culture was performed as previously described (Ying *et al.* 2003a, Ying *et al.* 2008). 2i and LIF+BMP4 are N2B27-based media. The composition of the N2B27 used for these cultures is identical to that used for differentiation, with the exception that the homemade N2 supplement is replaced by the N2 supplement produced by Gibco (500µl N2/100ml medium, final concentration 0.5X).

To generate 2i medium, N2B27 was supplemented with 1µM PD0325901 and 3µM CHIR99021. 100U/ml LIF were added to make 2i+LIF medium.
For the transfer of ES cells cultured in LIF+FCS to 2i culture medium, culture dishes were coated with poly-L-ornithine (Sigma) for 2 hours at room temperature. The poly-L-ornithine coating was removed and the dishes were then coated with laminin (Sigma) for 2 hours at room temperature. LIF+FCS cultures were trypsinised, quenched, pelleted, resuspended in N2B27, pelleted again, resuspended in 2i and counted. The cells were plated in 2i at a density of 10^4 cells/cm^2. 2i cultures were passaged every three days. Cells were washed with DMEM/F12 and detached from the dish by incubating them in accutase (Sigma) at 37°C until detachment (typically 3 minutes). They were quenched in N2B27, pelleted, resuspended in 2i and replated at a density of 10^4 cells/cm^2.

To generate LIF+BMP4 medium, N2B27 was supplemented with 100u/ml LIF and 10ng/ml BMP4.
For the transfer of ES cells cultured in LIF+FCS to LIF+BMP4 culture medium, culture dishes were coated with 0.1% gelatine for 30 minutes. LIF+FCS cultures were trypsinised, quenched, pelleted, resuspended in N2B27, pelleted again, resuspended in LIF+BMP4, counted and plated in LIF+BMP4 at the desired density. To passage LIF+BMP4 cultures, cells were washed with DMEM/F12 and detached from the dish by incubating them in accutase (Sigma) at 37°C until detachment (typically 3 minutes). They were quenched in N2B27, pelleted, resuspended in LIF+BMP4 and replated at the desired density.

2.2.4.2.4 Epiblast stem cell culture

Epiblast stem cell cultures were cultured as previously reported (Brons *et al.* 2007, Tesar *et al.* 2007) with minor modifications. The N2B27-based EpiSC culture medium composition was optimised in our Institute for EpiSC viability and

proliferative capacity by Dr. Rodrigo Osorno and Dr. Anestis Tsakiridis in Dr. Ian Chambers and Dr. Val Wilson's laboratories. Its composition is as follows.

Component	Volume	Final concentration
DMEM/F12 (Gibco)	50ml	
Neurobasal (Gibco)	50ml	
0.1M 2-mercaptoethanol	100µl	100nM
L-Glutamine	1ml	2mM
Non-Essential Amino Acids (Gibco)	1ml	1X
N2 (Gibco)	500µl	0.5X
B27 (Gibco)	2ml	1X
Activin A, 10µg/ml in 0.1% BSA (R&D)	200µl	20ng/ml
Fgf2, 10µg/ml in 0.1% BSA (R&D)	100µl	10ng/ml

The EpiSC basal medium (modified N2B27) was prepared without addition of Activin A and Fgf2. These molecules were added to EpiSC basal medium at the concentrations described above immediately prior to use.

EpiSCs were derived from ES cells following the protocol of Guo *et al.* (2009). A well of a 6-well plate was coated with 7.5ng/ml fibronectin in PBS for 1 hour. The fibronectin solution was removed and $3x10^4$ ES cells were plated onto the well in LIF+FCS. After 24 hours, the well was washed twice with PBS and EpiSC culture medium was added to the cells. The cells were allowed to reach high density before passaging. To passage the cells, they were washed in DMEM/F12 and incubated in accutase for 1-2 minutes at 37°C. The dish was forcefully slapped to induce cell detachment from the basal membrane without generating a single-cell suspension. N2B27 was used to quench the accutase, the cells were pelleted, resuspended in EpiSC culture medium and replated at high density in 1:2-1:4 dilutions. The EpiSCs were used for experimentation after 10 passages, to minimise the chances of leftover

undifferentiated ES cell in the cultures, and before 20 passages, to decrease the likelihood of genetic or epigenetic abnormalities induced by extended passaging.

2.2.4.2.5 Cell culture on coverslips

Cells to be subject to confocal microscopy were cultured on confocal microscopy-compatible plastic coverslips custom-made for the use of our laboratory by Ibidi. For each sample to be prepared, a coverslip was placed in a well of a 6-well plate and coated with the appropriate coating medium for a minimum of 1 hour prior to cell plating.

2.2.4.2.6 Cell stimulation with exogenous molecules

Cells stimulated with exogenous factors for flow cytometry or immunostaining analysis were treated as follows.

For ES cells cultured in LIF+FCS, cells were plated at a density of $2x10^4$ cells/cm^2 in the presence of the exogenous factors and cultured for 48 hours.

For ES cells cultured in 2i, cells were plated at a density of 10^4 cells/cm^2. After 24 hours, the culture medium was replaced with 2i supplemented with the exogenous factors, and the cells were cultured for a further 48 hours. For the R2i sample described in section 3.2.5.3, the medium was switched from 2i to R2i after 24 hours of plating, and the cells were cultured for a further 48 hours.

For EpiSCs, the cells were plated in the presence of the exogenous factors at an appropriate dilution for them to reach confluency after 48 hours of culture (1:4-1:6).

2.2.4.3 ES cell transgenesis

2.2.4.3.1 Lipofection

10^5 ES cells were plated in LIF+FCS and allowed to attach to a well of a 6-well plate (minimum 1 hour, generally 5-6 hours). 1-3µg of linear or circular plasmids were diluted in 100µl GMEM. 1-2µl Lipofectamine™ 2000 were diluted in 100µl GMEM and incubated for 5 minutes at room temperature. The diluted plasmid solution was gently added to the Lipofectamine 2000 solution and incubated for 20 minutes at room temperature. The Lipofectamine 2000/plasmid solution was then added to the ES cells in a drop-wise manner. The plate was gently rocked to spread the lipofection solution and incubated overnight at 37°C. Mock lipofections were performed by diluting 1-2µl Lipofectamine 2000 in 200µl GMEM and adding it to cells. Positive control lipofections were performed by lipofecting a fluorescent plasmid and screening for fluorescence under a fluorescent microscope 24 hours post-lipofection.

2.2.4.3.2 Electroporation

Prior to electroporation, thirteen 9cm dishes were gelatinised for 30 minutes at room temperature. The gelatine was removed, LIF+FCS medium was added to the dishes and they were transferred to a 37°C incubator for 1 hour. 9.2ml LIF+FCS medium were added to two universal tubes, which were transferred to a 37°C waterbath for 1 hour. 100µg of the plasmid to be electroporated were linearised with the appropriate restriction enzyme, precipitated in 100% ethanol and washed once with 70% ethanol. For the second 70% ethanol wash, the microfuge tube in which the DNA was being precipitated was filled entirely with 70% ethanol to minimise the chance of bacterial carryover and spun again. The tube was then transferred to a cell culture hood, the ethanol removed and the pellet allowed to dry before resuspension in 100µl DNase/RNase-free distilled water (Gibco).

ES cells to be electroporated were grown to confluency in two 75cm^2 flasks. The cells were trypsinised, pelleted, resuspended in GMEM, pelleted and resuspended in 700µl GMEM. The cells were counted and 10^7 cells were added to a 4mm electroporation cuvette. The linearised plasmid in 100µl water was added to the cells and GMEM was added to a total volume of 800µl. The electroporation mixture was incubated at room temperature for 3 minutes, following which the cells were electroporated at 3µF (0.8kV). The 800µl containing electroporated cells were immediately transferred to a universal containing 9.2ml of pre-warmed LIF+FCS medium, and 1ml of this solution (10^6 cells) was added to each of ten 9cm dishes. For the mock electroporation, 10^7 cells were subject to the same electroporation procedure in the absence of plasmid DNA, transferred to the other universal containing 9.2ml of pre-warmed LIF+FCS, and 1ml of this solution was added to each of three 9cm dishes. Selective medium was added 24 hours after cell plating.

2.2.4.4 Cryopreservation of cells

2.2.4.4.1 Cell freezing

To cryopreserve cells, they were subject to trypinisation/accutase treatment and pelleted following standard passaging routines. The supernatant was aspirated off and the cells were resuspended in FCS+10% dimethylsulphoxide (DMSO) (LIF+FCS cultures) or knockout serum replacement (KSR)+10% DMSO (N2B27-based cultures). 1ml of FCS/KSR+DMSO was used for each aliquot to be frozen (usually 10^6 cells were frozen per aliquot). Upon resuspension, each aliquot was transferred to a cryovial (Thermo Scientific) and immediately placed on dry ice. They were then transferred to -80°C freezers for temporary storage and liquid nitrogen tanks (-178°C) for long-term storage.

2.2.4.4.2 Cell defrosting

To defrost cells, 9ml LIF+FCS (for cells frozen from LIF+FCS) or N2B27 (for cells frozen from N2B27-based cultures) were added to a universal and incubated at 37°C. The aliquot containing the frozen cells was thawed at room temperature and immediately added to the pre-warmed medium. The cells were pelleted and the supernatant aspirated. The cells were then resuspended in the culture medium of choice and plated onto pre-coated plates or dishes.

2.2.4.5 ES cell differentiation protocols

2.2.4.5.1 N2B27 differentiation

The preparation of ES cell samples for N2B27 neural adherent monolayer differentiation was performed as described in paragraph 2.2.4.2.2 and as previously reported (Pollard *et al.* 2006, Ying *et al.* 2003a, Ying *et al.* 2003b). The optimal plating density was empirically determined for each cell line to maximise cell survival and efficiency of differentiation. For the differentiation experiments using Id1 overexpressing and empty vector control ES cells, the densities were optimised to obtain similar cell numbers at the final timepoint of the differentiation experiments (day 5). Due to the continuous rapid proliferation rate of Id1 overexpressing ES cells during the first three days of differentiation, presumably due to their delayed exit from pluripotency, the plating densities were as follows: empty vector control, Id1 reverted ES cells: $1.8x10^4/cm^2$ (10^6 cells/9cm dish); Id1 overexpressing ES cells: $5.5x10^3/cm^2$ ($3x10^5$ cells/9cm dish). Plating of the Id1 overexpressing ES cells at the density used for control cell lines resulted in complete confluency on day 3 of differentiation.

2.2.4.5.2 Embryoid body differentiation

ES cells were prepared for embryoid body (EB) differentiation by trypsinising, quenching and pelleting them as for routing passaging. They were then resuspended twice in LIF+FCS ES cell culture medium to which LIF had not been added ("-LIF medium") and counted. Their density was adjusted to 25000 cells/ml –LIF medium. 10ml PBS were added to a non-cell culture treated 140mm Petri dish (Sterilin). The lid of the dish was positioned so that its outside was in contact with the base plates of the cell culture hood. 100 droplets of 25µl of the cell suspension (625 cell/droplet) were pipetted onto the inside of the lid. The lid was turned over and closed over the PBS-filled bottom of the Petri dish, and carefully moved to a 37°C incubator. After three days, the PBS was removed from the Petri dish, the droplets were dropped onto the bottom of the Petri dish and 30ml –LIF medium were added to maintain the expanding EB's in culture. To collect the EB's for RNA extraction, the 30ml medium was transferred to a universal tube, the EB's were pelleted by centrifugation at 300g for 3 minutes, the supernatant was carefully aspirated off, and the EB's were lysed following the same protocol adopted for cell pellets.

For the setup of EB differentiation from sorted cells, the cells in the sorted samples were pelleted, resuspended twice in –LIF medium supplemented with penicillin and streptomycin (Gibco) to prevent bacterial contamination deriving from the fluorescence activated cell sorter. The differentiation was then set up as described for trypsinised cells above, maintaining penicillin and streptomycin in the culture medium throughout the differentiation experiment.

2.2.4.6 Flow cytometry and fluorescence activated cell sorting

2.2.4.6.1 Sample preparation

Cell samples were prepared for flow cytometry and fluorescence activated cell sorting by generating a single-cell suspension with standard passaging methods.

Following trypsin/accutase treatment, quenching and pelleting, the cells were resuspended in PBS+10% FCS and immediately placed on ice in the dark. To remove dead cells from the flow cytometry analyses and the sorted samples, these resuspensions were supplemented with 100ng/ml DAPI or 1µg/ml propidium iodide. DAPI was used for samples analysed on the LSRFortessa (IVNR cells) or samples sorted on the FACSAria, propidium iodide was used for samples analysed on the FACSCalibur (flow cytometry experiments on IdV and IdV-SC cells), since this analyser lacks the laser/detector combination to detect DAPI. These fluorescent molecules cannot enter intact cells, but can penetrate dead cells with disrupted cell membranes and bind DNA, thus allowing to discriminate between live DAPI/propodium iodide-negative and dead DAPI/propidium iodide-positive cells during flow cytometry experiments.

2.2.4.6.2 Sample analysis

Flow cytometry experiments were performed as recommended by the manufacters of the analysers. The softwares used on the machines were CellQuest for the FACSCalibur and FACSDiva for the LSRFortessa. 10000-50000 cells were analysed for each sample. The FACSAria cell sorter was operated by the Flow Cytometry Staff in my presence, using the FACSDiva software. E14Ju09 ES cells were used as negative controls, and single-fluorophore controls were used whenever possible.

2.2.4.6.3 Data analysis

The data generated on the analysers and the sorter was exported as .fcs files. Analysis of the data was performed with the FlowJo software. Debris was excluded by gating using forward scatter and side scatter. Comparison of the amplitude and height of the side scatter signal was used to exclude doublets (two cells stuck

together and simultaneously analysed). Dead cells were then excluded by gating for DAPI-negative or propidium iodide-negative cells. This data was then plotted to observe the expression pattern of the fluorescent proteins of interest. When generating overlay histograms, the same number of events were inputed for each separate sample. Gating to calculate percentages of positive and negative cells was performed using negative control cells (E14Ju09 for Venus and tagRFP, IdV for tagRFP).

2.2.5 Embryology

2.2.5.1 Maintenance of mice

Mice were housed and bred in the Animal Unit of the Centre for Regenerative Medicine, in accordance to the provisions of the Animals (Scientific Procedures) Act 1986 and the 2010/63/EU Directive.

2.2.5.2 Morula aggregation

Morula aggregations were performed by the Transgenics Unit staff. ES cells used for morula aggregations were cultured in LIF+FCS prior to dissociation and aggregation into C57/BL6 host embryos. Embryos were cultured until the blastocyst stage and injected into the uterus of pseudopregnant females.

2.2.5.3 Embryo dissections

Female mice that were injected with chimaeric blastocysts were culled by cervical dislocation by the staff of the Animal Unit. The uterus was dissected from the mice, and embryos were dissected from the uterus in M2 culture medium (Sigma) at room

temperature using forceps. The decidua and Reichert's membrane were removed before the embryos were fixed as described in section 2.2.3.2.1. The dissections were performed under a StereoZoom SMZ-U dissection microscope (Nikon).

2.2.5.4 Sectioning fixed embryos

Fixed embryos were sectioned by the Histology Service of the institute. Embryos were embedded in gelatine, 7µm-thick sections were generated with a cryostat and mounted on glass microscopy slides.

CHAPTER 3

The expression pattern of Id1 in pluripotent cultures

3.1 Introduction

The helix-loop-helix factor Id1 carries out a crucial role in the maintenance of embryonic stem cell pluripotency. ES cells are routinely cultured in medium containing LIF and foetal calf serum, and the key component of FCS for ES cell self-renewal is BMP4, an extracellular signalling molecule which drives the transcription of *Id1*. Id1 overexpression can functionally substitute for BMP4/FCS to preserve the self-renewal capability of ES cells (Ying *et al.* 2003a).

The expression of *Id1* in pluripotent cultures has been studied at the population level. *Id1* is expressed in ES cells cultured in LIF+FCS and rapidly downregulated at the onset of differentiation, both in neural adherent N2B27 differentiation and undirected embryoid body differentiation protocols; it is expressed at low levels in 2i culture (Aiba *et al.* 2009, Marks *et al.* 2012, Romero-Lanman *et al.* 2011, Ying *et al.* 2003a). The expression pattern of *Id1* within pluripotent cultures at a higher resolution has been addressed as part of a single-cell qRT-PCR study, which revealed a heterogeneous expression pattern for the HLH factor both in ES cell cultures and in the inner cell mass of the blastocyst (Tang *et al.* 2010).

The absence of comprehensive information on the expression pattern of Id1 at the protein level within single cells in culture led me to generate Id1-Venus fluorescent reporter ES cells. These cells were used to address the pattern of Id1 expression within pluripotent cultures at the single cell level by flow cytometry and imaging, to study its response to activation and inhibition of specific signalling pathways, and to investigate its relationship with other transcription factors involved in the maintenance of ES cell pluripotency.

3.2 Results

3.2.1 Id1-001 is the only expressed Id1 isoform in pluripotent and differentiated ES cells

The *Id1* locus encodes two splice variants of the gene, *Id1-001* and *Id1-002* (Hernandez *et al.* 1996). These alternative transcripts are highly similar, and code for proteins which share 135 amino acids out of a total of 148 (Id1-001) or 168 (Id1-002). The transcripts differ at their 3' ends, *Id1-001* contains an intron which is not spliced in *Id1-002* (Figure 3.1a).

I investigated which of these transcripts were expressed in ES cells and cells differentiated for 5 days in N2B27 medium in a neural adherent monolayer differentiation protocol (Ying *et al.* 2003a,b). I designed two sets of primers to amplify sequences contained in both transcripts or in *Id1-002* only (Figure 3.1a). I could not design a set of primers specific to *Id1-001* since its entire sequence is present in *Id1-002* mRNA. I performed PCR's on cDNA from ES cells cultured in LIF+FCS and ES cells cultured in N2B27 for 5 days, and observed that only *Id1-001* is expressed in these cultures (Figure 3.1b).

3.2.2 The expression pattern of Id1 mRNA in pluripotent and differentiating cultures

I set out to recapitulate published qRT-PCR and microarray results on the expression of *Id1* mRNA in pluripotent cultures and differentiating cells. I performed qRT-PCR experiments on cDNA obtained from E14tg2a ES cells cultured in 2i, LIF+FCS or differentiated into EpiSCs (Figure 3.2a). *Id1* is expressed at very low levels in 2i culture, at slightly higher levels in EpiSCs and at much higher levels in LIF+FCS. I then performed an adherent neural monolayer differentiation experiment by plating E14tg2a ES cells (cultured in LIF+FCS) in N2B27 for 5 days, extracting RNA at

daily timepoints. I performed qRT-PCR experiments on these samples and observed that *Id1* is sharply downregulated at the onset of differentiation, and its expression is progressively re-acquired during the course of the differentiation experiment (Figure 3.2b), as had been previously observed in microarray experiments (Aiba *et al.* 2009).

3.2.3 The quality of commercially available anti-Id1 antibodies is insufficient for reliable Id1 protein expression analysis

In order to address the pattern of Id1 expression in culture at the protein level, I tested two commercially available anti-Id1 antibodies. The first antibody (Santa Cruz SC-488) had been used by Ying *et al.* (2003a) with promising results, the second (Chemicon 7D4.2) was tested for comparison. I performed Western blots on the lysate of E14tg2a ES cells cultured in LIF+FCS, and probed the membranes with the two antibodies (Figure 3.3). Neither of the antibodies resulted in the presence on the membranes of a single band of approximately 15 kDa, corresponding to Id1. The Santa Cruz antibody detected Id1, but also detected a further 9 non-specific bands. The Chemicon antibody resulted in a smear on the membrane. This confirmed a previous report which suggested the Santa Cruz SC-488 antibody was unsuitable for the specific detection of Id1 in mouse tissues (Perk *et al.* 2006).

3.2.4 Generation of an Id1-Venus reporter ES cell line

The lack of a commercially available high quality specific anti-Id1 antibody raised the problem of how to address the expression of Id1 protein in pluripotent cultures at single-cell resolution. When faced with the same issue, Perk *et al.* (2006) opted for the generation of anti-Id1 antibodies. Whilst this strategy offers the advantage of being able to study the expression of Id1 in all cells and tissues of interest, the successful generation of a specific antibody is by no means certain. I selected an

alternative strategy, with a higher likelihood of success and other benefits: the generation of an Id1 fluorescent reporter ES cell line. The advantages of a fluorescent reporter line are manifold: not only can the expression of Id1 be studied by imaging and Western blotting, it can also be studied by flow cytometry, and live cells can be sorted into separate populations based on expression of the fluorescent reporter protein. This cannot be done with antibody staining, since Id1 is an intracellular protein and permeabilisation of live cells for immunostaining prior to flow cytometry would result in their death.

3.2.4.1 Selection of a reporter system

I decided to generate a fusion of Id1 with Venus, a yellow fluorescent protein with the appealing properties of being bright and fast-folding (Nagai *et al.* 2002, Shaner *et al.* 2005), and thus coupling ease of detection with minimal likelihood of disruption of the properties of wild-type Id1. The use of a fusion protein offers advantages over systems that rely on gene replacement, on insertion of an internal ribosomal entry site followed by the fluorescent protein sequence downstream of the gene of interest, or on 2A peptide-fluorescent protein fusions. The replacement of one of the alleles of a gene of interest with a fluorescent protein may result in the loss of potential regulatory sequences within the gene body, resulting in a misregulation of the transcription of the reporter allele. It can also lead to haploinsufficiency phenotypes. Whilst this is unlikely to be the case for Id1, since its knockout has no phenotype *in vivo* (Fraidenraich *et al.* 2004, Lyden *et al.* 1999), the observation that its knockout does have a phenotype in ES cells (Romero-Lanman *et al.* 2011) means subtle haploinsufficiency phenotypes have the potential to arise under specific culture conditions. Internal ribosome entry sites cannot guarantee equimolar expression of the mRNAs of the fluorescent protein and of the gene of interest. Conventional IRES's can lead to lower expression of the transcript downstream of the IRES (Chan *et al.* 2011), which could result in undetectable fluorescent protein expression in a

cell that is expressing the gene of interest. The *gtx-IRES* transcriptional amplification system results in higher expression of the transcript downstream of the IRES (Canham *et al.* 2010, Chappell *et al.* 2000). If the transcriptional amplification is not constant amongst all cells in culture, differences in fluorescent protein expression levels could arise in cells expressing the same amount of gene of interest mRNA. Fusion of the protein of interest to a fluorescent protein via a 2A peptide allows equimolar expression of the two proteins with minimal disruption to the function of the protein of interest. 2A peptides, in fact, lead to the co-translational separation of the two linked proteins, so that expression of each molecule of a protein of interest leads to expression of a molecule of fluorescent protein (Donnelly *et al.* 2001, Palmenberg *et al.* 1992, Ryan *et al.* 1991). This system should allow for accurate reporting of the expression of the gene of interest, but does not allow to study the subcellular localisation of the protein of interest, and, if the stability of the fluorescent protein and the protein of interest vary greatly, could result in persistence of fluorescence after loss of expression of the protein of interest. The use of fusion proteins is also subject to disadvantages: the fusion protein may be defective in cell localisation or in protein activity, and the fluorescent protein may affect protein stability, but it guarantees the presence of the protein of interest whenever fluorescence is detected. Importantly, in the case of Id1, an Id1-Venus fusion protein has already been generated and shown to accurately report Id1 expression within the nervous system, without causing any phenotype in the reporter mouse harbouring two copies of the reporter construct targeted into both endogenous *Id1* loci (Nam & Benezra 2009).

3.2.4.2 Id1-Venus targeting strategy

To generate Id1-Venus reporter ES cells, I made use of an updated version of the targeting construct used by Nam & Benezra (2009), kindly gifted to our laboratory by Dr. Robert Benezra and Dr. Hyung-song Nam (Memorial Sloan-Kettering Cancer

Center, New York, USA). This targeting vector contains a 5' targeting arm with 12479bp of homology to the wild-type *Id1* locus (including *Id1* itself), the sequence encoding a flexible linker and Venus fused to the 3' end of *Id1-001* upstream of the stop codon, a further 874bp of homologous DNA containing the polyadenylation site for the transcript, an *Frt*-flanked Flp-excisable selection cassette containing a *Pgk* promoter driving transcription of the selectable marker *Neo*, a 3630bp-long 3' homology arm, and an *RNAPII* promoter fragment driving expression of the toxic molecule *DTA* outwith the regions of homology, to promote the death of cells harbouring random integrations of the construct (Figure 3.4b). Correct targeting of this construct to the *Id1* locus would result in the replacement of the endogenous *Id1* sequence (Figure 3.4a) with the *Id1-Venus* fusion followed by the selection cassette (Figure 3.4c). Episomal transfection of a plasmid containing a *Pgk* promoter driving expression of *FlpO* recombinase would then result in recombination of the two *Frt* sites and loss of the selection cassette (Figure 3.4d). For more information refer to section 2.2.4.1.5 of the Materials and Methods chapter.

3.2.4.3 Generation of Id1-Venus reporter ES cells

I electroporated the linearised targeting construct into E14Ju09 ES cells. These ES cells were derived in our Institute from E14tg2a chimaeric embryos and share the same genotype of E14tg2a ES cells. They were selected for use due to their high rate of germline transmission and chimaeric contribution, as determined by our Institute's Transgenics Service staff. Electroporated cells were cultured in the presence of 300µg/ml G418 for 10 days before clones were picked into 96-well plates. Genomic DNA was extracted from each of the clones undergoing expansion. I performed Southern blots on the genomic DNA samples digested with HindIII using the probe designed by Nam & Benezra (2009) for identification of correctly targeted clones (Figure 3.5a). I identified three clonal ES cell lines with the correct targeting event (clones 47, 72 and 77). The ~6kb band corresponding to the targeted allele band in

the blots appeared to be weaker than the ~12kb wild-type band for clones 72 and 77. I re-extracted genomic DNA from the three clones and the parental E14Ju09 ES cells and repeated the Southern blots, probing the membrane with the same "targeting" probe as in Figure 3.5a as well as with a probe to detect spurious integrations of the construct (Figure 3.5b, probe binding location depicted in Figure 3.4). All three clones had appropriately sized wild-type and targeted bands, but the targeted band for clones 72 and 77 appeared to be weaker than the wild-type band again. Clone 77 also had an extra construct integration event. I therefore elected to use clone 47 for future experiments, and I re-named this clonal ES cell line IdV, for Id1-Venus.

3.2.4.4 Verification of Venus expression in the reporter cells

I proceeded to test whether the Venus fluorescence was detectable from the fusion protein by performing flow cytometry and immunostaining of ES cells cultured in LIF+FCS. Flow cytometry revealed that Venus fluorescence can be detected in this cell line, and that it is distributed normally, with the majority of cells expressing the fluorescent protein (Figure 3.5c). I then stained IdV ES cells using a polyclonal anti-GFP antibody (A111222, Molecular Probes) capable of detecting the Venus epitope, and counterstained cell nuclei using DAPI. This revealed a heterogeneous expression pattern for Id1-Venus within ES cell colonies (Figure 3.5d). I made use of a nuclear immunostaining quantification software named MultiCell3D, developed by Dr. Guillaume Blin, a member of our laboratory, to measure the Venus fluorescence intensity within the each nucleus of over 900 IdV and over 1000 E14Ju09 ES cells (more information on MultiCell3D can be found in section 2.2.3.2.8 of the Materials and Methods chapter). A logarithm-scale plot of the intensity distribution of the nuclear Venus signals revealed a fluorescence distribution nearly identical to that observed in the flow cytometry experiment (Figure 3.5d), giving further credit to the flow cytometry result and confirming the validity of the nuclear immunostaining quantification technique for analysis of protein expression at the single cell level.

3.2.4.5 Venus expression reports Id1 expression

In order to verify whether Venus fluorescence correctly reported expression of *Id1*, I performed fluorescence-activated cell sorting of the 20% of cells expressing the highest and the 20% of cells expressing the lowest Venus fluorescence in a LIF+FCS ES cell culture. RNA was collected from the sorted samples and the starting culture and reverse transcribed to cDNA. qRT-PCR experiments were performed on the three samples to measure the expression level of *Id1*, *Venus* and the housekeeping gene *Tbp*, used to normalise the levels of *Id1* and *Venus* transcripts (Figure 3.6). Both *Id1* and *Venus* were expressed at the highest level in the Venus-high population, at an intermediate level in unsorted cells, and at the lowest level in the Venus-low sample. This suggests the expression of Venus in the reporter cells is accurately reporting the expression of *Id1*.

3.2.4.6 Excision of the selection cassette from IdV ES cells

I then set out to excise the selection cassette from the IdV ES cells. I lipofected the cells with a plasmid encoding FlpO recombinase, plated the cells at clonal density and allowed them to form colonies. I picked 96 colonies into a 96-well plate, replica plated them, and cultured one plate in the absence of G418 and the other plate in the presence of 300µg/ml G418. Only 1 colony displayed subsceptibility to G418, suggestive of excision of the selection cassette. Genomic DNA was extracted and digested with HindIII, and a Southern blot was performed with this genomic DNA alongside that of E14Ju09 and IdV ES cells (Figure 3.7a). The Southern blot confirmed the removal of the selection cassette. I named this cell line IdV-SC for IdV minus Selection Cassette. For more information refer to sections 2.2.4.1.6 of the Materials and Methods chapter.

3.2.4.7 Comparison of Venus fluorescence in IdV and IdV-SC ES cells

I next set out to compare the expression level of Venus between IdV and IdV-SC ES cells. I decided not to perform this comparison between cells cultured in LIF+FCS because these culture conditions can contain different ratios of cells in different self-renewal states (naïve pluripotency, primed pluripotency and differentiation) between parallel cultures, and this property could in turn influence the number of Id1-Venus expressing cells and the intensity of Id1-Venus expression in those cells. I therefore decided to culture the two ES cell lines in 2i, a defined culture medium which homogenises cell potency and gene expression (Ying *et al.* 2008, Nichols *et al.* 2009), and supplemented this medium with 10ng/ml of the *Id1* inducer BMP4 and with 10μM SB431542, a small-molecule inhibitor of Activin/Nodal signalling, shown to induce *Id1* expression in ES cells cultured in LIF+FCS (Galvin *et al.* 2010). I performed flow cytometry analysis of the two cell lines alongside E14Ju09 non-fluorescent control ES cells, and observed that Id1-Venus was expressed in 2i+BMP4+SB431542 at similar levels between IdV and IdV-SC cells, suggesting either cell line is fit for Id1-Venus expression analysis in ES cells (Figure 3.7b). The IdV ES cells offer the possibility of applying G418 selection during routine cell passaging, thus ensuring that the transgene is not lost by mitotic recombination (Stern 1936) or that the locus is not epigenetically silenced. The IdV-SC ES cells have the advantage of lacking exogenous DNA downstream of the *Id1* locus (with the exception of an Frt site), thus ensuring that no regulatory DNA sequence is disrupted. Serum-free culture results in greater susceptibility of cells to antibiotic selection, and to prevent excessive cell death I avoided applying the selective medium in these conditions. I consequently performed all experiments in serum-free medium with IdV-SC ES cells, and I will present flow cytometry data for LIF+FCS culture performed with the same cell line for coherence. Most of these data were generated in parallel with IdV ES cells, and the cell lines displayed the same behaviour. Their identical response to BMP4 stimulation and inhibition is illustrated

as an example in Figure 3.7c (see the next paragraph for further details on the BMP4 response of Id1-Venus reporter cells).

3.2.5 Analysis of Id1-Venus expression in pluripotent cultures by flow cytometry

3.2.5.1 Id1 is induced by BMP4 and repressed by Activin A in LIF+FCS culture

Id1 has been shown to respond to treatment with BMP4 and the Activin/Nodal inhibitor SB431542 in LIF+FCS culture of ES cells at the transcript level (Galvin *et al.* 2010). I set out to investigate the response of Id1 protein to BMP4 and Activin/Nodal stimulation and inhibition at the single cell level by performing flow cytometry analyses of IdV ES cells cultured in LIF+FCS and treated with agonists and antagonists of the BMP and Activin A signalling pathways for 48 hours. I selected this experimental strategy to understand how ES cells in LIF+FCS culture conditions would respond to alterations in these signalling pathways over a time window that would allow them to comprehensively alter their gene expression, without conceding a long enough time frame for the cells to overtly differentiate. The observed changes in Id1-Venus expression should therefore represent a mid-term reaction to exogenous stimuli at steady state, rather than the intrinsic capability of an ES cell culture to respond to said stimuli after starvation in the absence of any other instructive signal.

Firstly, I assessed the Id1-Venus response to BMP signalling stimulation and inhibition. To stimulate BMP signalling, I made use of recombinant BMP4 at a concentration of 10ng/ml, the concentration used by Ying *et al.* (2003) to maintain ES cell self-renewal in LIF+BMP4 culture. To inhibit BMP signalling, I cultured cells in the presence of 100nM LDN-193189, a small-molecule compound which at this concentration selectively inhibits phosphorylation of Smad1/5/8 without affecting phosphorylation of Smad2 in mouse pulmonary artery smooth muscle cells cultured in the presence of FCS (Yu *et al.* 2008). The compound is also routinely

used at this concentration to inhibit BMP signalling in the neural induction of human pluripotent stem cells (Chambers *et al.* 2012).

I performed flow cytometry on untreated, BMP4-stimulated and LDN-193189-treated LIF+FCS cultures of IdV-SC ES cells, alongside E14Ju09 parental non-fluorescent control ES cells (Figure 3.8a. This is the same experimental setup as in Figure 3.7c; the data are displayed for ease of comparison with the rest of the figure). I observed that, as expected, Id1-Venus fluorescence was increased upon BMP4 treatment and was decreased upon LDN-193189 treatment. This suggests that whilst there is BMP in FCS, it is not at a concentration which would saturate BMP receptors, or BMP4 addition to LIF+FCS culture would have no effect. It is interesting to note that LDN-193189 treatment decreased Id1-Venus expression but did not abolish it, suggesting that either the LDN-193189 dosage was not sufficient to fully suppress BMP signalling, or that factors other than Smad1/5/8 were promoting Id1-Venus expression (see paragraph 1.3.2.5 for a description of non-Smad *Id1* inducers). Increasing the concentration of LDN-193189 would likely result in non-specific inhibition of Activin/Nodal receptors (Yu *et al.* 2008), thus the question remains unanswered.

I then proceeded to investigate the effects of Activin A stimulation and inhibition on expression of Id1-Venus in LIF+FCS culture. Activin A stimulates the phosphorylation of Smad2/3, which can act antagonistically to the BMP-induced phosphorylation of Smad1/5/8 (Candia *et al.* 1997, Yamamoto *et al.* 2009), and it is therefore perhaps unsurprising that Activin/Nodal inhibition can induce *Id1* at the transcript level (Galvin *et al.* 2010). To stimulate Activin/Nodal signalling, I made use of recombinant Activin A at a concentration of 20ng/ml, the concentration used to culture Activin/Nodal-dependent epiblast stem cells (Brons *et al.* 2007, Tesar *et al.* 2007). To inhibit Activin/Nodal signalling, I used 10μM SB431542, a small molecule compound that at this concentration selectively inhibits phosphorylation of Smad2 without affecting the phosphorylation of Smad1 in a variety of cell types cultured in FCS (Inman *et al.* 2002). The compound is also routinely used at this

concentration to inhibit Activin/Nodal signalling in the neural induction of human pluripotent stem cells (Chambers *et al.* 2009, Chambers *et al.* 2012).

I performed flow cytometry on untreated, Activin A-stimulated and SB431542-treated LIF+FCS cultures of IdV-SC ES cells, alongside E14Ju09 parental non-fluorescent control ES cells (Figure 3.8b). I observed that, as expected, Id1-Venus fluorescence was increased upon SB431542 treatment. Activin A stimulation of ES cells decreased Id1-Venus fluorescence, but not by a large amount. This suggests that FCS contains a high amount of Activin/Nodal signalling agonists, which almost saturate Activin/Nodal receptors. It would appear that Activin/Nodal inhibition is the key limiting step for Id1-Venus expression in LIF+FCS culture: when gating the flow cytometry plot so that fewer than 1% of E14Ju09 ES cells result Venus-positive, approximately 90% of SB431542-treated IdV-SC ES cells express Id1-Venus, compared to approximately 65% of untreated ES cells (gate not shown, average of three experiments).

I then tested the effects of administering combinations of agonists and antagonists of Activin/Nodal and BMP signalling (Figure 3.8c,d). Activin A reduced Id1-Venus expression in BMP4-treated cells, but did not further decrease Id1-Venus expression in LDN-193189-treated cells, confirming the near-saturating levels of Activin/Nodal agonists in FCS. SB431542 treatment potentiated the BMP4-driven Id1-Venus induction, and overcame LDN-193189 inhibition to drive similar levels of Id1-Venus expression to LIF+FCS culture.

It would therefore appear that there is a high level of Activin/Nodal signalling in LIF+FCS culture. Inhibition of this with SB431542 results in a strong increase in fluorescence that cannot be suppressed by BMP inhibition. It is also interesting to note that LDN-193189-mediated BMP inhibition leads to a reduction in Id1-Venus fluorescence that is not affected by further administration of Activin A, which suggests a basal level of Id1-Venus expression may be intrinsic to LIF+FCS culture,

or that other Activin A inhibition-resistant factors contribute to Id1-Venus expression.

In summary, the expression of Id1-Venus at the protein level in LIF+FCS culture is positively regulated by BMP4 and negatively regulated by Activin/Nodal signalling.

3.2.5.2 The expression of Id1-Venus in different ES cell culture systems

ES cells can be maintained in a self-renewing state using different combinations of molecules supplemented to basal culture media. I set out to investigate the changes in expression of Id1-Venus in four ES cell culture media. Alongside the canonical LIF+FCS culture, I tested Id1-Venus expression in three N2B27-based media: LIF+BMP4, in which the self-renewal-stimulating activity of serum is directly supplanted by that of BMP4 (concentration: 10ng/ml) (Ying *et al.* 2003a); 2i, in which two small molecule inhibitors of Mek (PD0325901, concentration: 1μM) and of GSK3β (CHIR99021, concentration: 3μM) shield cells from autocrine signals that would drive their differentiation, without the need for supplementation with further extracellular signals (Ying *et al.* 2008); and 2i+LIF, in which LIF stimulation strengthens the maintenance of self-renewal induced by 2i.

I performed flow cytometry on IdV-SC ES cells cultured in each of these four conditions for a minimum of three passages, alongside non-fluorescent control parental E14Ju09 ES cells cultured in LIF+FCS (Figure 3.9a). Id1-Venus was not expressed in 2i nor in 2i+LIF, and was expressed at similar levels in LIF+FCS and LIF+BMP4. The lack of fluorescence in 2i and 2i+LIF confirms the transcript level data in Figure 3.2a. The observation that replacement of FCS with BMP4 results in similar Id1-Venus expression strengthens the argument that BMP is the crucial component of FCS for Id1 expression in ES cell culture (Ying *et al.* 2003a).

3.2.5.3 Activin/Nodal inhibition and BMP4 stimulation are essential for Id1-Venus expression in 2i culture

In order to understand whether the signals that regulate Id1 expression in LIF+FCS play a similar role in 2i culture, I performed flow cytometry on IdV-SC cells cultured in 2i and supplemented with LIF, 20ng/ml Activin A, 10ng/ml Fgf2 or 100nM LDN-193189 for 48 hours (Figure 3.9b). I included Fgf2 amongst the test conditions to address whether it could induce Id1-Venus expression via Mek-independent pathways. The adopted concentration of Fgf2 is the concentration used to maintain pluripotency in EpiSC culture (Osorno *et al.* 2012). Supplementation of 2i medium with either of these molecules had no effect on Id1-Venus expression, which was unsurprising since the reporter is not expressed in 2i. This also suggests that Fgf2 cannot induce Id1-Venus expression in the presence of PD0325901.

I then tested the effects of adding 10ng/ml BMP4 and 10μM SB431542 to 2i cultures for 48 hours (Figure 3.9c,e). BMP4 stimulation drastically increased the expression of Id1-Venus, but the expression of the reporter was fully silenced upon simultaneous addition of BMP4 and Activin A (the histogram for the 2i+BMP4+Activin A sample overlaps that of 2i alone and is hidden by it). SB431542 promoted a mild increase in fluorescence, that was suppressed upon simultaneous addition of SB431542 and LDN-193189 (the histogram for the 2i+SB431542+LDN-193189 sample overlaps that of 2i alone and is hidden by it). The combination of BMP4 and SB431542 resulted in the expression of Id1-Venus to levels similar to those of LIF+FCS culture.

I also compared the expression levels of the reporter cells cultured in 2i+SB431542 or in R2i, a culture medium alternative to 2i that also promotes ground state pluripotency, the composition of which is N2B27+1μM PD0352901+10μM SB431542 (Figure 3.9d) (Hassani *et al.* 2012, Hassani *et al.* 2013). The expression levels of Id1-Venus in the two media are similar, suggesting that inhibition of GSK3β by CHIR99021 does not affect the levels of Id1-Venus.

The combination of these data suggests that the expression of Id1-Venus in 2i culture is dependent on simultaneous activation of BMP and inactivation of Activin/Nodal signalling, as exemplified by the lack of Venus fluorescence in 2i+BMP+Activin A and 2i+LDN-193189+SB431542. The levels of BMP in 2i culture appear to be lower than those of Activin/Nodal agonists, since Id1-Venus displays a stronger induction upon BMP4 addition than SB431542 addition (assuming Id1-Venus is induced and repressed equally by equimolar amounts of BMP and Activin/Nodal agonists).

3.2.5.4 Id1 expression in LIF+FCS is inhibited by Mek inhibition independently of BMP and Activin/Nodal

I proceeded to test the effects of the inhibitors used in 2i culture on Id1-Venus expression in LIF+FCS. I cultured IdV-SC ES cells in LIF+FCS supplemented for 48 hours with 1μM of the Mek inhibitor PD0325901, 3μM of the GSK3β inhibitor CHIR99021, a combination of both inhibitors or with 10ng/ml Fgf2, to address whether Fgf stimulation had an opposite effect to Mek inhibition. This concentration of Fgf2 is the concentration used to maintain pluripotency in EpiSC culture (Osorno *et al.* 2012). I performed flow cytometry on the cells alongside an unsupplemented LIF+FCS culture and a non-fluorescent parental E14Ju09 ES cell control (Figure 3.10a). Inhibition of Mek resulted in a marked reduction of Id1-Venus expression, to a level lower than that seen with LDN-193189. Addition of Fgf2 to the culture did not increase Venus fluorescence (the histogram for the LIF+FCS+Fgf2 sample overlaps that of LIF+FCS alone and is hidden by it), suggesting that either Fgf's (or other Mek agonists) are present in FCS to a high amount, or that Mek inhibition downregulates Id1-Venus expression via an indirect mechanism. Inhibition of GSK3β resulted in a slight downregulation of Id1-Venus. Simultaneous inhibition of GSK3β and Mek reduced the expression level of Id1-Venus to a similar level to that seen for Mek inhibition alone.

I then tested whether the Mek inhibition-driven downregulation of Id1-Venus could be rescued by addition of BMP4 (Figure 3.10b). Addition of BMP4 to PD0325901-treated cells promoted expression of Id1-Venus to levels lower than LIF+FCS, a similar result to that seen for 2i+BMP4 (Figure 3.9c). This suggests that the mechanism of inhibition of Id1-Venus expression by PD0325901 has a more pronounced effect on Id1-Venus expression than BMP4 stimulation.

Next, I asked whether the combination of Mek inhibition, BMP inhibition and Activin/Nodal stimulation could abolish Id1-Venus expression in LIF+FCS culture (Figure 3.10c). The combination of PD0325901 with Activin A, LDN-193189 or both further downregulated the expression of Id1-Venus compared to either of the molecules alone (refer to Figure 3.8 for the effects of Activin A and LDN-193189 in isolation or in combination), but could not fully suppress Id1-Venus fluorescence. Fewer than 5% of the cells expressed the reporter, and in these cells the levels of Id1-Venus were very low, suggesting that the combination of these factors has a profound inhibitory effect on Id1-Venus expression.

These observations suggests that the process of Mek inhibition is a major regulator of Id1 expression in addition to BMP and Activin/Nodal. The observation that the combination of Mek inhibition with either BMP inhibition or Activin A stimulation resulted in the downregulation of Id1-Venus compared to either of these treatments in isolation implies that Mek inhibition negatively regulates Id1 expression through a mechanism that is likely to be independent of BMP and Activin/Nodal. This conclusion is compounded by the observation that BMP stimulation of PD0325901-treated cells can rescue Id1-Venus expression in part both in LIF+FCS and in 2i. Furthermore, it appears that Id1 expression can be drastically reduced but not fully abolished in LIF+FCS. This is suggestive of the presence of one or more factors that can positively regulate Id1 expression independently of the effects of BMP, Activin A and Mek.

3.2.5.5 Id1 is induced by BMP4 and repressed by Activin A in epiblast stem cell culture

Having characterised the expression pattern of Id1-Venus in ES cells, I wished to study its expression and its response to exogenous signals in EpiSCs. I therefore derived EpiSCs from IdV-SC ES cells by plating ES cells in EpiSC culture medium (N2B27+10ng/ml Fgf2+20ng/ml Activin A) and passaging them at high density for 10 passages to ensure that the great majority (if not the totality) of cells in the culture have exited naïve pluripotency. This derivation method was originally described by Guo *et al.* (2009) and is routinely adopted by various research groups in our Institute (Osorno & Chambers 2011, Osorno *et al.* 2012).

In first place I assessed the effect of withdrawing Fgf2 or Activin A from EpiSC culture medium on the expression of Id1-Venus. To ensure suppression of autocrine signals, I added 10μM SB431542 to the cells which I cultured in the absence of Activin A, and 1μM PD0325901 to the cells which I cultured in the absence of Fgf2. I cultured EpiSCs in these conditions for 48 hours and assessed their expression of Id1-Venus by flow cytometry (Figure 3.11a). EpiSCs express reduced levels of Id1-Venus compared to ES cells, but the expression of the reporter is not abolished, in accordance with the transcript-level data presented in Figure 3.2a. Immunofluorescence analysis of EpiSCs reveals that Id1-Venus is expressed in the Oct4-expressing pluripotent compartment, as well as in Oct4-negative cells (Figure 3.11d). Culture of EpiSCs in the absence of Fgf2 and the presence of PD0325901 results in a slight downregulation of Id1-Venus. Removal of Activin A and addition of SB431542, on the other hand, drives the upregulation of Id1-Venus. This suggests that in EpiSC culture the expression of Id1 is dampened by Activin A, and that Fgf2 only plays a minor role in the promotion of Id1 expression.

I next investigated the response of EpiSCs to BMP stimulation and inhibition, by culturing EpiSCs in 10ng/ml BMP4 or 100nM LDN-193189 for 48 hours (Figure 3.11b,d). Treatment of EpiSCs with LDN-193189 did not alter the expression of Id1-

Venus, suggesting that BMP signalling is low in EpiSC cultures or that its effectors are inhibited. Culture of EpiSCs in the presence of BMP4 resulted in a surprising bimodal expression pattern: a proportion of the cells responded by upregulating Id1-Venus to levels similar to those seen in LIF+FCS culture, whilst the remainder of the cells appeared refractory to BMP4 stimulation and retained an expression pattern similar to that of non-stimulated EpiSCs. The proportion of the cells that respond or do not respond to BMP varies between experiments, potentially as a result of slightly different cell densities. This suggests that one or more factors present in EpiSC cultures can desensitise cells from BMP stimulation.

I proceeded to test whether one of such factors may be Activin A by removing it from the EpiSC culture medium and by suppressing paracrine Activin/Nodal signalling with 10μM SB431542, whilst adding 10ng/ml BMP4 to the culture medium (Figure 3.11c). This resulted in the loss of bimodal Id1-Venus expression, with cells now expressing the reporter in a normal distribution, suggesting that Activin A may be responsible for the inhibition of the BMP4 response. Surprisingly, however, cultures treated with BMP4+SB431542 expressed lower average levels of Id1-Venus than cultures treated with SB431542 alone. This may be due to the fact that EpiSCs rely on Activin A for self-renewal, and are thus likely to have initiated differentiation after 48 hours in its absence. The cell type generated by BMP4+SB431542 stimulation may express other factors that may negatively regulate the expression of Id1-Venus. Stimulation of EpiSC cultures for a shorter time may aid in the study of the effects of these molecules within cells which have not lost their pluripotent identity.

In summary, epiblast stem cells express low levels of Id1 and have inactive BMP signalling. BMP stimulation results in upregulation of Id1-Venus in a subset of cells only, and inhibition of Activin/Nodal allows a higher proportion of cells to respond to BMP.

3.2.6 Analysis of Id1-Venus expression in pluripotent cultures by nuclear immunostaining quantification

Having studied the expression pattern of Id1-Venus in ES cell cultures by flow cytometry, I set out to investigate its expression pattern and its relationship with pluripotency factors by immunostaining. This technique allows one to observe a freeze-frame image of the state of cells within a culture, and to study the expression of multiple proteins of interest simultaneously within each individual cell. By measuring the fluorescence intensity for each protein of interest within each imaged nucleus with the MultiCell3D software (Chapter 2 section 2.2.3.2.8), it is also possible to obtain quantitative expression data at the single cell level.

I therefore decided to recapitulate some of the expression data I generated by flow cytometry making use of immunostaining to study the expression pattern of Id1-Venus and of the pluripotency factors Nanog and Oct4. Given that Id1 is expressed heterogeneously, I wished to investigate whether this factor is specifically associated with either naïve (Nanog-high) or primed (Nanog-low) subpopulations of ES cells, or with spontaneously differentiated cells (Oct4-negative). I performed immunostains on cells cultured in 2i, LIF+FCS, LIF+FCS+BMP4 (10ng/ml), LIF+FCS+LDN-193189 (100nM) or LIF+FCS+PD0325901 (1μM) (Figure 3.12a). I observed that, in accordance with the flow cytometry data, Id1-Venus was not expressed in 2i, was expressed at low levels in the presence of LDN-193189 or PD0325901, and was expressed at higher levels in LIF+FCS and LIF+FCS+BMP4. Nanog was expressed uniformly at high levels in 2i, and slightly less homogeneously but still at high levels in the presence of PD0325901. It was expressed at a lower level and more heterogeneously in the other culture conditions. Oct4 was expressed at similar levels in a rather uniform manner in all culture conditions (Figure 3.12). This acts as a control confirming that differences in expression of Id1-Venus and Nanog are not artefacts resulting from differences in size or shape of cells in the different conditions.

The agreement of these data with the findings of the flow cytometry experiments validate this quantitative immunostaining approach.

I performed pairwise comparisons of gene expression within single cells by generating dot plots plotting the expression level of one factor over another within each nucleus (Figure 3.13). I observed that in all of the conditions, only very few (if any) cells were co-expressing high levels of Id1-Venus and high levels of Nanog. It appears that cells that are expressing high levels of one of the two factors will express intermediate or low levels of the other. This observation is in accordance with a previous report that Nanog negatively regulates the expression of *Id1* (Suzuki *et al.* 2006). The negative relationship between the two factors had not previously been described at the single-cell level. The relationship between Id1-Venus and Oct4 appears to be different. High levels of Id1-Venus are often associated with high or intermediate levels of Oct4, with very few cells expressing high levels of Id1-Venus and low levels of Oct4. A similar, yet even more striking, distribution is observed when plotting the levels of Nanog and Oct4. No cells appear to express high levels of Nanog and low levels of Oct4, with a visually evident maximum Nanog:Oct4 ratio being present in each of the plots. In 2i, the variability of Oct4 for each specific Nanog intensity is decreased, so that all cells lie in a line close to the maximum Nanog:Oct4 ratio described above. This effect has previously been described by Muñoz Descalzo *et al.* (2012) using a different nuclear immunostaining quantification software. This confirms the replicability of the result and the validity of MultiCell3D as an analysis platform. It is interesting to note that cells cultured in LIF+FCS+PD0325901 appear to tend to the maximum Nanog:Oct4 ratio more than cells cultured in the other LIF+FCS-based conditions, suggesting this culture condition may contain a higher number of cells in a naïve pluripotent state, as may be expected from the high levels of Nanog seen in the presence of PD0325901.

In conclusion, Id1-Venus and Nanog appear to be in a negative relationship, as high expression of both factors within the same cell was rarely observed in any of the analysed conditions.

3.2.7 Investigation of the relationship between Id1 and Nanog through the generation of an Id1-Venus Nanog-tagRFP double reporter ES cell line

3.2.7.1 Generation of Id1-Venus Nanog-tagRFP reporter ES cells

In order to further investigate the relationship between Id1 and Nanog, I decided to target one *Nanog* allele in IdV ES cells to generate a double fluorescent reporter ES cell line. I made use of a construct generated and kindly gifted to me by Dr. Nicola Festuccia in Dr. Ian Chambers's laboratory. Targeting of this construct to the *Nanog* locus results in the insertion of DNA encoding a flexible linker and the red fluorescent protein tagRFP at the 3' end of the *Nanog* open reading frame, prior to the stop codon, resulting in the production of a Nanog-tagRFP fusion protein. This sequence is followed by an *IRES* and *Bls*, a gene conferring resistance to blasticidin S (Figure 3.14a-c). The targeting construct is promoterless, the transcribed region of the *Nanog* locus constitutes the entire 5' homology arm, the 3' untranslated region and the 4315 nucleotides of DNA downstream of that constitute the 3' homology arm. Random integration of the construct should therefore be unlikely to confer resistance to blasticidin S, as the expression of *Bls* is dependent on integration downstream of a promoter. For more information refer to section 2.2.4.1.8 of the Materials and Methods chapter.

I electroporated 100µg of the linearised targeting construct into 10^7 IdV ES cells. After 10 days of selection with 10µg/ml blasticidin S, only 9 colonies had grown. I picked and expanded the clones, and performed Southern blots on BamHI-digested genomic DNA extracted from the 9 cell lines, using the targeting probe depicted in Figure 3.14. All 9 samples displayed the correct banding pattern (data not shown). I

repeated the DNA isolation, digestion and Southern blotting for 3 of the clones. I confirmed the correct targeting of the construct to one allele of *Nanog*, and verified that the construct had not randomly integrated in other regions of the genome for any of the clones (Figure 3.14d). These double reporter ES cell lines were named IVNR, for Id1-Venus Nanog-tagRFP.

3.2.7.2 Verification of tagRFP expression in the reporter cells

I proceeded to test whether the tagRFP and Venus fluorescence were detectable in the three IVNR clones, and if the fluorescence levels were consistent with the expected behaviour of Nanog and Id1 proteins in various culture conditions. I cultured the three IVNR clones in LIF+FCS, LIF+FCS+10µg/ml blasticidin S (in order to select for *Nanog-tagRFP-IRES-Bls* expression from the targeted *Nanog* locus), 2i or N2B27 for 48 hours, then analysed the expression of Venus and tagRFP by flow cytometry (Figure 3.15). The 2i samples had been cultured in 2i medium for 4 passages prior to the setup of the experiment. The fluorescence distribution of Id1-Venus met expectations: expression of the factor was at its maximum in LIF+FCS, decreased in N2B27 and was detected at low levels in very few cells in 2i, for all of the three clones. Addition of blasticidin S to LIF+FCS culture also decreased Id1-Venus expression, an observation which fits with blasticidin S-mediated selection of *Nanog-tagRFP*-expressing cells, and with the negative input of Nanog on *Id1* expression (Suzuki *et al.* 2006).

Nanog-tagRFP expression was extremely low. Less than 10% of IVNR ES cells cultured in LIF+FCS expressed tagRFP, with the percentage increasing only marginally upon addition of blasticidin S. 2i culture, which results in uniform Nanog expression (Figure 3.12a, Ying *et al.* 2008), led to a slight increase in tagRFP expression, but did not result in uniform expression of the fluorescent reporter, with 40-70% of cells still within the tagRFP-negative gate in the three IVNR clonal lines. This suggests that whilst the three cell lines express Nanog-tagRFP in accordance

with their culture conditions, the fluorescence generated by the fusion protein is too low to be detected with the BD LSRFortessa analyser used for this experiment (the only analyser in the Institute capable of detecting tagRFP). The same observations were made by Dr. Nicola Festuccia upon generation of Nanog-RFP reporter cell lines from wild-type ES cells.

In order to verify that the weak expression of tagRFP truly recapitulated the expression of Nanog, I randomly selected IVNR clone 2 ES cells and immunostained them using anti-tagRFP and anti-Nanog antibodies in LIF+FCS culture (Figure 3.16a). I performed nuclear immunostaining quantification on the samples, and observed a positive correlation between Nanog and tagRFP expression (Pearson product-moment correlation coefficient: 0.8629) (Figure 3.16b). This confirms that tagRFP fluorescence accurately reports the expression of Nanog. Having verified this for IVNR clone 2, I used this clone for further experiments.

3.2.7.3 Attempted purification of populations based on fluorescent protein expression

I then tested whether fluorescence activated cell sorting of ES cells based on expression of Venus and tagRFP could result in the isolation of enriched populations expressing a selected combination of the two proteins. I cultured IVNR ES cells in LIF+FCS and sorted them into four populations based on Venus and tagRFP expression (Figure 3.17a). I immediately analysed the sorted samples on the BD LSRFortessa analyser (Figure 3.17b). Whilst separation based on Venus expression worked with considerable success, separation based on tagRFP expression did not. The vast majority of cells from the tagRFP-negative sorted samples lacked expression of tagRFP, but only a minor fraction of cells from the tagRFP-positive sorted samples were expressing the red fluorescent protein, although the proportion of tagRFP-expressing cells was increased in these populations compared to unsorted cells (data not shown, Figure 3.15). Thus, it would appear that sorting based on

tagRFP expression succeeds in isolating a Nanog-tagRFP-negative population (with more than 98% purity), but does not succeed in purifying a Nanog-tagRFP-positive population. It can however generate a Nanog-tagRFP enriched population (24% tagRFP-positive cells after sorting, compared to 4.5% tag-RFP positive cells in the unsorted sample. These numbers and the previous one for the tagRFP-negative population represent the average of the Venus-high and Venus-low samples).

3.2.7.4 Venus and tagRFP report Id1 and Nanog expression

I repeated the experiment described in the previous paragraph and extracted RNA from the four sorted populations as well as an unsorted control. I reverse transcribed the RNA and performed qRT-PCR experiments on the samples (Figure 3.18). *Id1* expression was enriched in the Venus-positive tagRFP-negative (V+R-) and Venus-positive tagRFP-positive (V+R+) samples. *Nanog* expression was enriched in the Venus-negative tagRFP-positive (V-R+) and the V+R+ samples. The expression of both factors was low in the double negative samples (V-R-) (Figure 3.18a). This suggests that the two fluorescent proteins report the expression of *Id1* and *Nanog*, and confirms tagRFP-based sorting is capable of enriching the proportion of Nanog-tagRFP-expressing cells within a sample. The expression of the Nanog target *Esrrb* reflected the expression of *Nanog*, whereas that of the general pluripotency marker *Oct4* was similar across all sorted populations (Figure 3.18b). The expression of the primed pluripotency markers *Fgf5* and *T* (but not that of *Otx2*) was enriched in the V-R- sample, suggestive of a possible differentiation-primed state within cells not expressing neither Id1 nor Nanog in LIF+FCS culture (Figure 3.18c,d). The pro-differentiation factor *Tcf15* was lower in the R+ samples, consistent with its negative relationship with Nanog (Figure 3.18c,d) (Davies *et al.* 2013). Neither *Hex* nor *Sox1* displayed enrichment in any of the samples, suggesting that neither of these four subpopulations is primed for primitive endoderm or neural commitment (Figure 3.18d). In conclusion, sorting of IVNR cells based on expression of Id1-Venus and

Nanog-tagRFP results in the isolation of cell populations displaying differential expression of pluripotency and differentiation markers.

3.2.7.5 Levels of Nanog, but not of Id1, in ES cell cultures influence the kinetics of EB differentiation

Having established that the four populations sorted from IVNR ES cells were characterised by different gene expression patterns, I tested whether they showed functional differences in their ability to undergo differentiation.

I repeated the sorting experiments and generated embryoid bodies (in the presence of FCS and in the absence of LIF) from each of the four sorted samples as well as from an unsorted control culture. I extracted RNA after 3 and 5 days of differentiation, reverse transcribed it, and performed RT-qPCR experiments to measure the expression levels of markers of various lineages (Figure 3.19). I confirmed that the sorted populations were correctly reporting expression levels of *Id1* and *Nanog* in the starting populations (Figure 3.19a). Both the tagRFP-positive samples (Venus-positive and Venus-negative) displayed a delayed exit from pluripotency, as indicated by higher expression of *Cdh1*, *Fgf5*, *Nanog*, *Oct4*, and the lower expression of the differentiation marker *Cdh2* after 3 days of differentiation, compared to the two tagRFP-negative populations (Figure 3.19a,b,f). The EBs contained little or no neural ectoderm, as indicated by the low levels of *Ascl1* (commonly referred to as *Mash1*) and *Sox1* (Figure 3.19c). The four populations upregulated the endodermal marker *Gata6* and the pro-EMT factors *Zeb1* and *Zeb2* with the same kinetics (Figure 3.19d,f). The delayed exit from pluripotency was evident when observing differentiation towards the mesodermal lineage. The tagRFP-positive populations expressed higher levels of the primitive streak marker *T* and lower levels of the differentiated mesoderm marker *Kdr* (also commonly referred to as *Flk1*) on the third

day of differentiation (Figure 3.19e), but both markers were expressed at similar levels between the samples after 5 days of differentiation.

Therefore, it appears that different expression levels of Id1 in LIF+FCS culture are not associated with differences in the kinetics of ES cell differentiation towards mesendodermal lineages in FCS-containing medium. High expression of Nanog, on the other hand, is associated with a slower exit from the pluripotent state.

3.3 Discussion

3.3.1 Flow cytometric analysis of Id1 expression in ES cell cultures

The generation of an Id1-Venus reporter ES cell line provided me with a useful tool to address the expression pattern of Id1 in pluripotent cultures at the protein level and at single-cell resolution.

I made use of this cell line to confirm published RNA expression data which suggested a heterogeneous expression pattern in LIF+FCS cultures (Tang *et al.* 2010), and observed that Id1-Venus fluorescence exhibits what appears to be a normal distribution, with some cells displaying high fluorescence and some not expressing the reporter allele (Figure 3.5c,d). The distribution of Id1-Venus fluorescence therefore suggests a continuous spectrum of Id1 protein levels within ES cells, as opposed to tightly regulated specific levels of expression.

3.3.1.1 The response of Id1-Venus to BMP stimulation and inhibition in LIF+FCS

I investigated the effect of agonists and antagonists of common signalling pathways on the expression of Id1-Venus. Id1 is a well-established target of BMP signalling in a variety of organisms and cell types, including mouse ES cells (Hollnagel *et al.* 1999, Katagiri *et al.* 1994, Ogata *et al.* 1993, Ying *et al.* 2003a). It is therefore

unsurprising that Id1-Venus fluorescence was increased upon BMP4 stimulation and decreased upon BMP inhibition of LIF+FCS cultures (Figures 3.7c, 3.8a). The latter observation confirms that BMPs are present in FCS: if this were not the case, then Id1-Venus fluorescence would be unaffected by BMP inhibition. The former observation suggests that the BMP levels in FCS are not sufficient to saturate BMP receptors, because if this were not the case then BMP4 stimulation would not increase Id1-Venus fluorescence. The approximate combined concentration of BMP4, BMP6 and BMP9 in FCS was reported to range between 11 and 25 ng/ml (equating to 1.1-2.5 ng/ml in culture medium containing 10% FCS), depending on the supplier and batch of FCS (Herrera & Inman 2009). It is therefore conceivable that supplementation with 10ng/ml BMP4 could result in receptor saturation, but addition of BMP4 at higher concentrations is required to test this hypothesis. BMP inhibition by treatment of ES cells with 100nM LDN-193189 did not result in abolition of Id1-Venus expression. It is unclear whether this is due to incomplete inhibition of BMP signalling or whether other factors are responsible for Id1 expression in LIF+FCS culture. Whilst increased LDN-193189 concentrations result in unspecific inhibition of Smad2 phosphorylation, implying the question above cannot be answered by increasing the concentration of this inhibitor, a Western blot against phosphorylated Smad's may clarify whether 100nM LDN-193189 can effectively inhibit Smad1/5/8 phosphorylation. The possibility that other factors are inducing Id1 expression will be discussed later in this chapter.

3.3.1.2 The response of Id1-Venus to Activin/Nodal stimulation and inhibition in LIF+FCS

Id1 has been reported to be a negative target of Activin/Nodal signalling in mouse embryonic stem cells (Galvin *et al.* 2010). I therefore set out to investigate how the supplementation of LIF+FCS culture medium with recombinant Activin A or with the Activin/Nodal signalling inhibitor SB431542 would affect Id1-Venus expression.

I observed that addition of Activin A to LIF+FCS had little effect on Id1-Venus fluorescence, but that administration of SB431542 dramatically increased Id1-Venus expression (Figure 3.8b). This suggests that Activin/Nodal signalling agonists are present at near-saturating levels in LIF+FCS culture, confirming previously reported observations (Ogawa *et al.* 2007). It also implies that Activin/Nodal signalling has a profound inhibitory effect on Id1-Venus expression.

3.3.1.3 The integration of multiple signals regulates Id1 expression in ES cell cultures

I analysed the expression of Id1-Venus in various ES cell culture conditions (Figure 3.9a). I observed that Id1-Venus expression is similar in LIF+FCS and N2B27+LIF+BMP4, which suggests the balance of BMP's and Activin/Nodal agonists is similar in the two culture conditions. ES cells cultured in LIF+FCS and LIF+KSR secrete Activin/Nodal agonists (Ogawa *et al.* 2007), so it is possible that the same is true for ES cells cultured in LIF+BMP4. I also observed that Id1-Venus is not expressed in 2i nor in 2i+LIF, in keeping with published RNA-Seq data suggesting low expression in 2i (Marks *et al.* 2012). This is an interesting observation in the light of the persistence of Id1-Venus fluorescence in LIF+FCS cultures treated with the BMP inhibitor LDN-193189 and with Activin A (Figure 3.8). There are various possible explanations to this difference in basal expression. The first is that 2i cultures could be devoid of BMP signals, whereas 100nM LDN-193189 may not be sufficient to inhibit all of the BMP in LIF+FCS. 2i cultures do appear to have low amounts of BMP, as judged by the strong induction of fluorescence which followed BMP4 administration to 2i cultures (Figure 3.9c,e).

A second possibility is that Mek and/or GSK3β inhibition could be responsible for a decrease in Id1-Venus fluorescence. This also appears to be the case, since both PD0325901 and CHIR99021 administration to LIF+FCS culture resulted in Id1-Venus downregulation, and simultaneous administration of PD0352901, Activin A

and LDN-193189 to LIF+FCS cultures caused a further downregulation of Id1-Venus expression (Figure 3.10a,c).

A third explanation could be that factors expressed in 2i culture may be responsible for the repression of *Id1* expression. Nanog is uniformly expressed in 2i (Figure 3.12a, Ying *et al.* 2008), and it can repress *Id1* transcription (Suzuki *et al.* 2006), suggesting this explanation may also be true.

Finally, it is also conceivable that factors expressed in LIF+FCS but not in 2i could be responsible for the transcription of *Id1*.

The combination of these data and considerations leads to the formulation of a model whereby Id1 expression in LIF+FCS is promoted by BMP agonists, present at levels below receptor saturation, and inhibited by Activin/Nodal agonists, present at levels close to receptor saturation, resulting in heterogeneous Id1-Venus expression. Treatment of cells with LDN-193189 reduces Id1-Venus expression, but does not abolish it, due to the positive input of one or more unidentified factors. Treatment of cells with Activin A does not affect the expression of Id1-Venus greatly, since Activin/Nodal agonists are present in culture at near-saturating levels. Treatment of cells with a Mek inhibitor drives the downregulation of Id1-Venus, possibly through the action of Nanog, and simultaneous inhibition of Mek and BMP signalling drives most, but not all, cells to lose expression of Id1-Venus. The residual expression may be due to incomplete Mek and BMP inhibition or to the positive input of one or more unidentified factors. Fgf2 does not induce expression of Id1-Venus, suggesting it is either found in serum at saturating levels, or that it does not promote *Id1* transcription (Figure 3.20a).

In 2i culture, levels of neither BMP nor Activin/Nodal agonists are saturating, whilst Mek and GSK3β inhibition are fully repressing the expression of Id1-Venus, possibly through the action of Nanog. The induction of Id1-Venus expression depends on simultaneous BMP stimulation and Activin/Nodal inhibition, as the repressive action of Mek inhibition will overcome the positive effect of BMP4 if Activin A is added to the culture at saturating levels, and it will overcome the

positive effect of SB431542-mediated Activin/Nodal signalling inhibition if BMP signalling is inhibited with LDN-193189. The complete silencing of Id1-Venus expression in 2i may be due to the lack of the unidentified factors promoting Id1 expression mentioned above for LIF+FCS culture, or to the uniform and high expression of Nanog, which can exert a negative effect on *Id1* transcription (Figure 3.20b).

3.3.1.4 The expression of Id1-Venus in epiblast stem cells

Id1-Venus is expressed at low levels in epiblast stem cells, due to the repressive action of the Activin A present in the culture medium and to the presence of low levels of BMP agonists, as confirmed by the lack of Id1-Venus downregulation in the cultures upon LDN-193189 treatment (Figure 3.11b). The expression of Id1-Venus in EpiSC cultures must therefore result from the positive influence of other factors. Removal of Fgf2 from EpiSC culture medium does not result in a marked downregulation of Id1-Venus (Figure 3.11a), suggesting that if it does induce the expression of Id1-Venus, it is not the sole factor to do so (Figure 3.20c).

BMP4 stimulation of EpiSCs results in a bimodal expression pattern, with a proportion of the cells unable to respond to BMP. The removal of Activin A from the culture medium coupled to Activin/Nodal inhibition resulted in the restoration of a normally-distributed expression pattern, but not all the cells express Id1-Venus, implying that other negative regulators of *Id1* transcription exist in these cultures (Figure 3.11b-d). It is also possible that EpiSCs express variable levels of BMP receptors, and that some cells fail to respond to BMP stimulation by expressing Id1-Venus for this reason.

3.3.2 Negative relationship between Id1-Venus and Nanog in ES cell cultures

3.3.2.1 Nuclear immunostaining quantification of ES cell cultures

In order to address the relationship between Id1-Venus and the pluripotency factors Nanog and Oct4, I performed immunostains of ES cells cultured in various conditions and quantified the immunofluorescent signal for each of the factors in the nuclei of each of the imaged cells (Figure 3.12). The expression pattern of Id1-Venus reflected that seen in flow cytometry analyses, displaying a response to BMP stimulation and inhibition, and low expression in 2i and LIF+FCS+PD0325901 (Figure 3.12b). This confirms the validity of immunostaining quantification as a technique for analysing protein expression at the single cell level.

Comparison of the expression level of pairs of factors within single cells indicated that cells expressing high levels of Id1-Venus tend to express Oct4, suggesting they are not differentiated cells within cultures, but also tend to express low or intermediate levels of Nanog. There appears to be a negative relationship between the two factors, with very few cells expressing high levels of both. This could be explained by the observation that *Id1* is a negative transcriptional target of Nanog (Suzuki *et al.* 2006), although the negative relationship between the two factors had not been previously characterised at the single cell level. The negative effect of Nanog on Id1 may explain why the HLH factor displays a certain heterogeneity of expression within monolayer cultures uniformly exposed to FCS.

I was also able to confirm a previous report on the relationship between Nanog and Oct4 in LIF+FCS and 2i cultures (Muñoz Descalzo *et al.* 2012), and to extend its findings to BMP4-, LDN-193189- and PD0325901-treated LIF+FCS cultures. ES cells appear to be incapable of expressing high levels of Nanog without expressing high levels of Oct4. There is a maximum Nanog:Oct4 ratio of fluorescent intensity which cannot be surpassed. This is clearly visible on the dot plots in Figure 3.13a as the limit before the empty triangle of space in the bottom right regions of the plots. In 2i culture, the Nanog:Oct4 ratio in all cells tends to this maximum ratio, so that while

levels of the factors in individual cells will vary, the expression ratio of the two proteins will remain relatively constant. I observed that in LIF+FCS cultures treated with PD0325901 the cells appeared to have less variability in Oct4 expression and to tend to the maximum Nanog:Oct4 ratio, suggesting this culture system is more similar to 2i than the other tested LIF+FCS-based cultures.

3.3.2.2 Issues with the generation of an Id1-Venus Nanog-tagRFP double reporter ES cell line

To further characterise the relationship between Id1 and Nanog, I generated Id1-Venus Nanog-tagRFP double reporter ES cells. Flow cytometry analysis revealed that the tagRFP fluorescence was too low for accurate detection with a BD LSRFortessa analyser, the only analyser in our Institute with the correct laser-detector combination for detection of tagRFP fluorescence. Even in 2i culture, where Nanog is expressed uniformly (Figure 3.12, Ying *et al.* 2008), tagRFP was undetectable in approximately half of the cells in the culture (Figure 3.15). This observation prevented an in-depth flow cytometric study of the expression pattern of Id1-Venus and Nanog-tagRFP in response to perturbation of signalling pathways in culture, since the data collected for Nanog-tagRFP would not be reflective of the true expression of Nanog. In addition, purification of Nanog-tagRFP-positive populations proved impossible (Figure 3.17).

3.3.2.3 Differences in gene expression in Id1-Venus Nanog-tagRFP sorted populations

Whilst not fluorescing at a sufficient intensity to capture the full range of Nanog expression, tagRFP did at least report high levels of Nanog, as seen by the shift towards positivity observed in 2i culture compared to N2B27 culture (Figure 3.15),

by the correlation of Nanog and tagRFP immunostaining signals (Figure 3.16), and by the enrichment for *Nanog* message in tagRFP-positive sorted cells (Figure 3.18a, Figure 3.19a). Interestingly, Id1-Venus-negative Nanog-tagRFP-negative cells displayed higher levels of the primed pluripotency markers *Fgf5* and *T*, lower levels of the naïve pluripotency marker *Esrrb* and similar levels of *Oct4* compared to the other sorted populations (Figure 3.18). This may suggest that these cells, unshielded from neural commitment due to the lack of Id1, and not constrained in a pluripotent state due to the lack of Nanog, may represent a cell type in a differentiation-favourable state. Undirected embryoid body differentiation of the sorted populations did not reveal accelerated kinetics of differentiation for this population compared to Id1-Venus-positive Nanog-tagRFP-negative cells (Figure 3.19). Perhaps this is not surprising, given that the EB differentiation protocol used in these experiments primarily induced mesendodermal differentiation, which is not strongly suppressed by Id1. Furthermore, it is possible that the serum present in the differentiation medium rapidly upregulated Id1 in the Nanog-tagRFP-negative Id1-Venus-negative subpopulation. A serum-free neural induction experiment may reveal differences in differentiation kinetics between these populations, given that Id1 gain-of function has a powerful inhibitory effect on neural differentiation (Ying *et al.* 2003a). Analysis of earlier timepoints may also be required to observe differences in the rate of exit from pluripotency.

3.3.2.4 Significance of the negative relationship between Id1 and Nanog

Nanog negatively regulates the transcription of *Id1* (Suzuki *et al.* 2006), and I have shown that the two factors display a negative relationship at the protein level. It is however unclear why the pluripotency factor Nanog would negatively regulate the inhibitor of differentiation Id1. The data I have generated is not sufficient to move us past the realm of speculation.

One possibility is that Id1 may be capable of inducing primitive endoderm differentiation of ES cells, and that Nanog is repressing this effect. Nanog is in fact expressed in an exclusive pattern with the primitive endoderm specifier Gata6 in the inner cell mass (Chazaud *et al.* 2006) and displays a negative relationship with primitive endoderm inducers in ES cell cultures (Singh *et al.* 2007). *Id1* expression was found to be associated with a primitive endoderm-like population of ES cells characterised by expression of *Hex* (Canham *et al.* 2010), although *Hex* is not enriched in Id1-Venus-positive cells (Figure 3.18d).

Alternatively, Nanog may be acting as a guardian of pluripotency by repressing the transcription of the neural inhibitor *Id1*, thus allowing ES cells to differentiate into any germ layer without the anti-neural bias which may be introduced by *Id1* expression.

A third possibility is that the negative relationship between the factors is required later in development, and its existence in pluripotent cultures is merely a by-product of later developmental events. Nanog and Id1 expression overlap in the proximal posterior regions of the gastrulating mouse embryo (Jen *et al.* 1997, Morkel *et al.* 2003, Osorno *et al.* 2012). Id1 repression by Nanog may be required at this time and location for the correct sequence of events to take place for the correct migration of mesendoderm through the primitive streak.

FIGURE 3.1

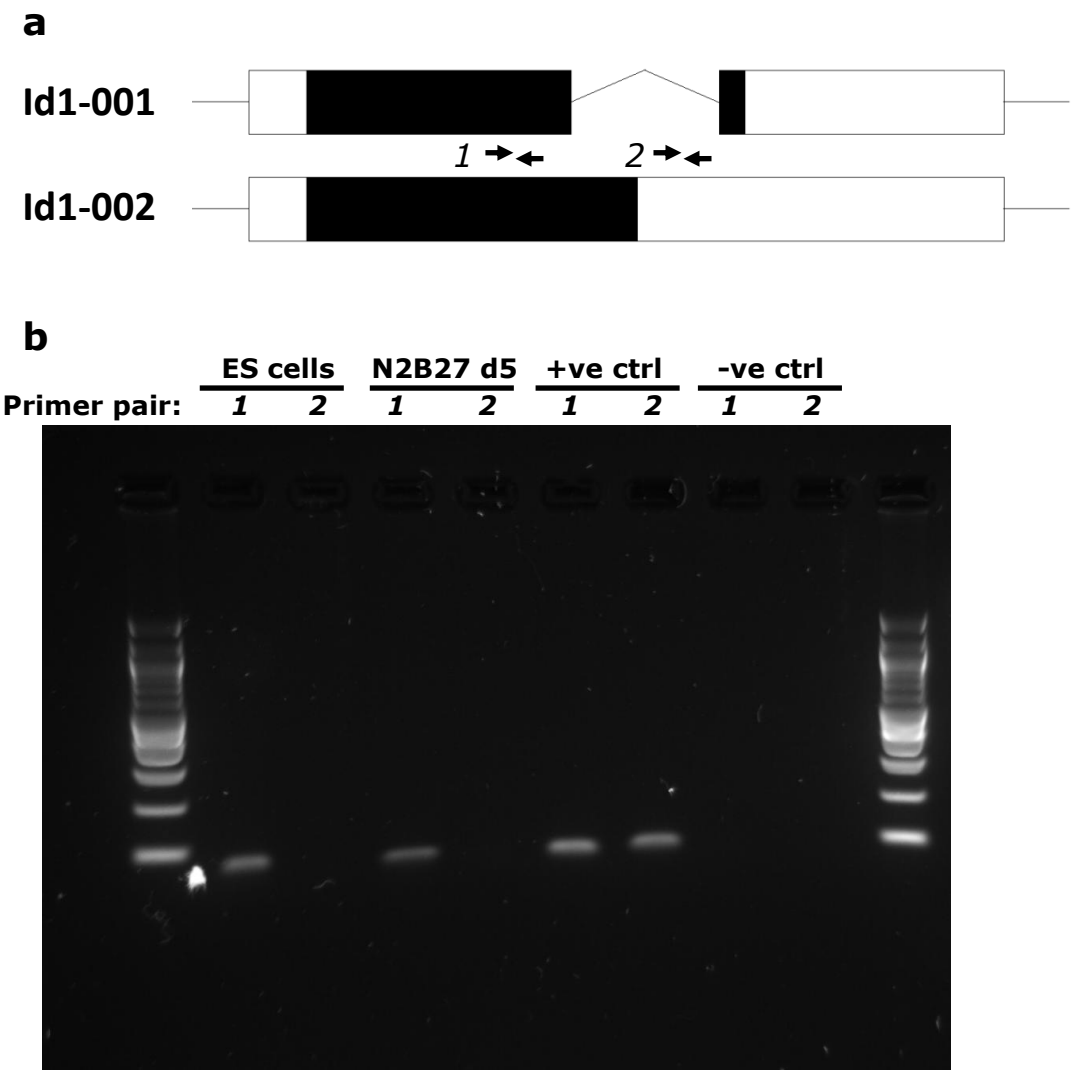

a
Id1-001
1
2
Id1-002
b
ES cells
N2B27 d5
+ve ctrl
-ve ctrl
Primer pair:
1
2
1
2
1
2
1
2

Figure 3.1 – Id1 isoforms and their expression

a. Structure of the *Id1* locus illustrating the DNA regions encoding two alternative mRNA isoforms.

Black boxes: coding region, white boxes: untranslated regions, diagonal lines: intronic DNA, straight lines: untranscribed intergenic DNA, arrows: primer pairs (numbered 1 and 2) used to detect both transcripts (1) or *Id1-002* only (2).

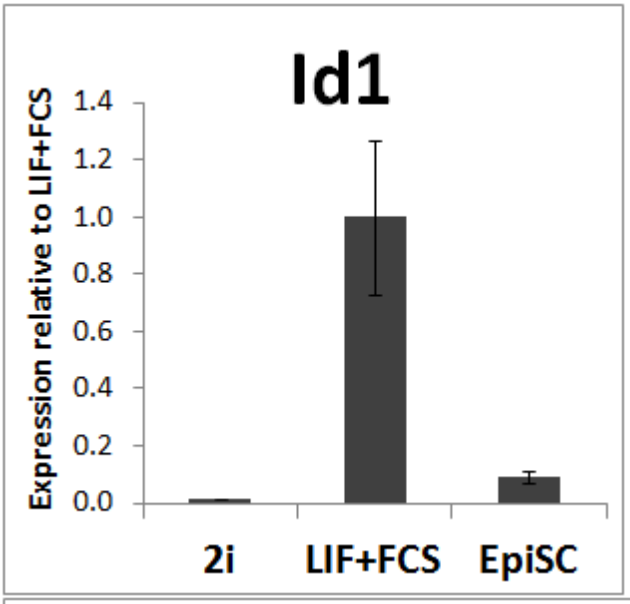
b. PCR performed using the primer pairs depicted in (a) to detect total *Id1* transcript (1) or *Id1-002* specifically (2).

Template DNA was cDNA from ES cells cultured in LIF+FCS ("ES cells"), cDNA from ES cells cultured in N2B27 for 5 days ("N2B27 d5"), a plasmid containing the entire *Id1* locus ("+ve ctrl") or water ("-ve ctrl").

The DNA ladder loaded either side of the samples is the NEB 100 bp DNA ladder.

FIGURE 3.2

a


Id1
Expression relative to LIF+FCS
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
2i
LIF+FCS
EpiSC

b

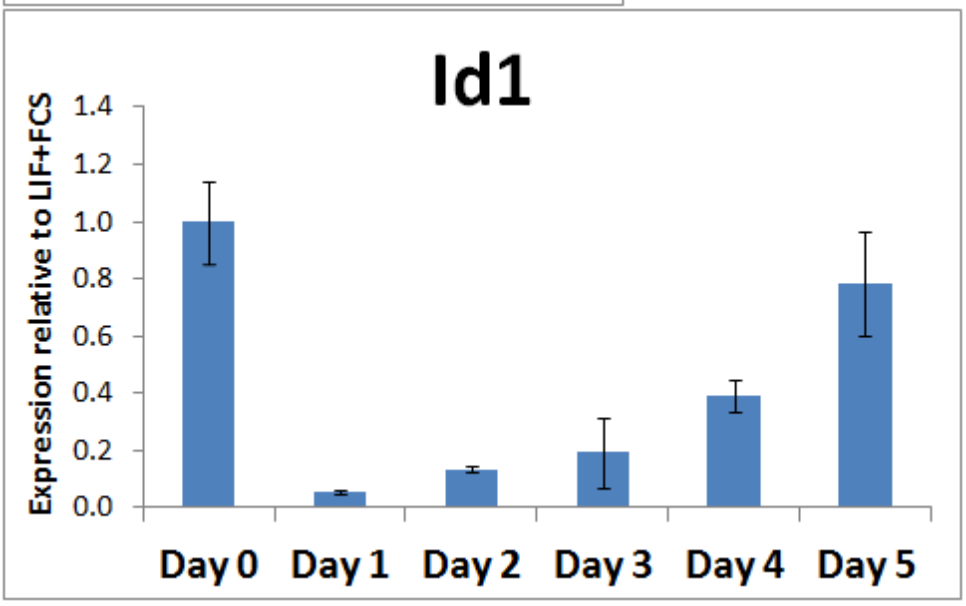

Id1
Expression relative to LIF+FCS
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Day 0
Day 1
Day 2
Day 3
Day 4
Day 5

Figure 3.2 – *Id1* expression in pluripotency and differentiation

a. *Id1* mRNA levels in 2i+LIF, LIF+FCS and EpiSCs. Error bars: standard deviation of two biological replicates.

b. *Id1* mRNA levels at daily timepoints of neural adherent monolayer differentiation of ES cells. ES cells cultured in LIF+FCS (Day 0) were replated in N2B27 and RNA extracted daily. Error bars: standard deviation of two biological replicates.

FIGURE 3.3

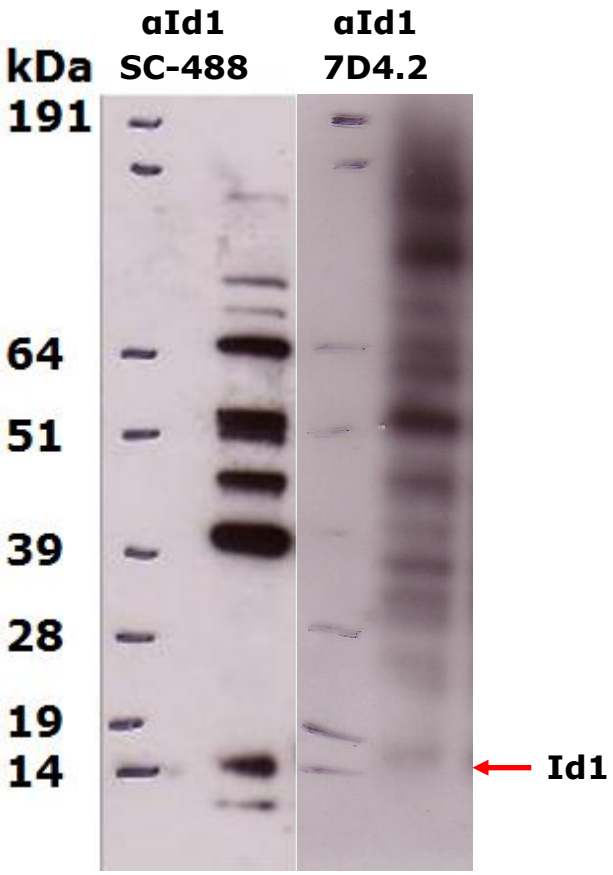
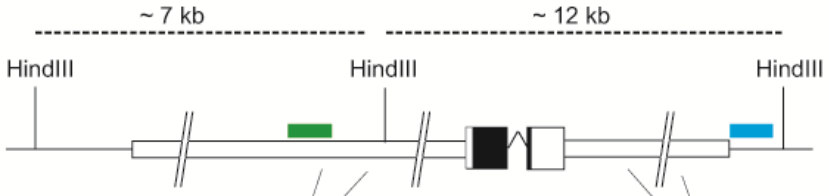


Figure 3.3 – Poor quality of anti-Id1 antibodies

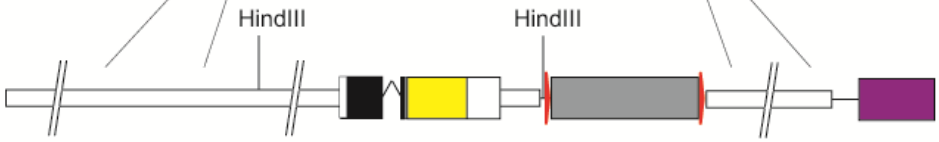
Western Blots on protein lysate from ES cells cultured in LIF+FCS, probing with two different commercially available anti-Id1 antibodies: SC-488 (Santa Cruz) and 7D4.2 (Chemicon). The expected molecular weight of Id1 (approximately 15 kDa) is indicated by the red arrow.

FIGURE 3.4

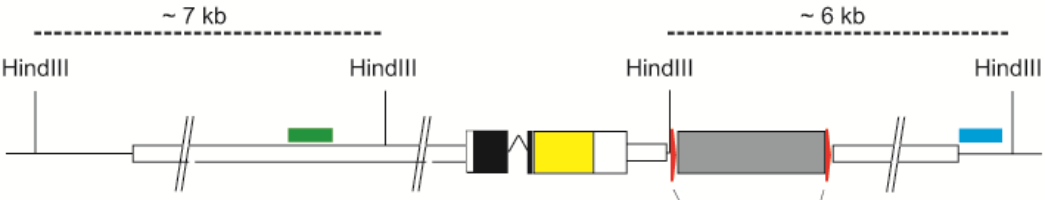
a – Wild-type locus



b – Targeting construct



c – Targeted locus



d – Targeted locus, *Pgk-Neo* removed

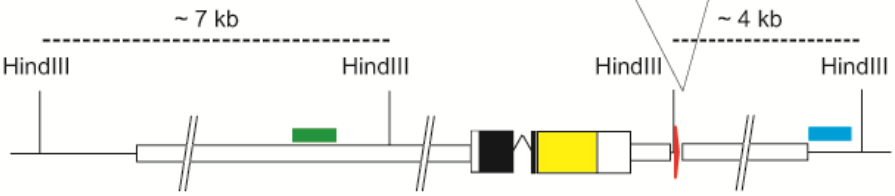


Figure 3.4 – Id1 reporter targeting strategy

a. Wild-type *Id1* locus.
Single horizontal lines: intergenic DNA, medium-sized white boxes: homology regions with targeting vector, big white boxes: untranslated regions, black boxes: coding regions, parallel diagonal lines: sequence of homology region not displayed, green box: Southern blot multiple integration probe, blue box: Southern blot targeting probe.

b. Targeting vector containing a flexible linker followed by the *Venus* cDNA (yellow box) fused in frame with the 3' end of *Id1*.
Red triangles: *Frt* sites, grey box: *Pgk-Neo* selection cassette, purple box: *DTA* negative selection marker outwith homology regions.

c. *Id1* locus following a correct targeting event.

d. Removal of the *Pgk-Neo* selection cassette following transfection with a plasmid encoding FlpO.

The sizes of bands detected on Southern blot membranes following genomic DNA digestion with HindIII are indicated for the various alleles.

FIGURE 3.5

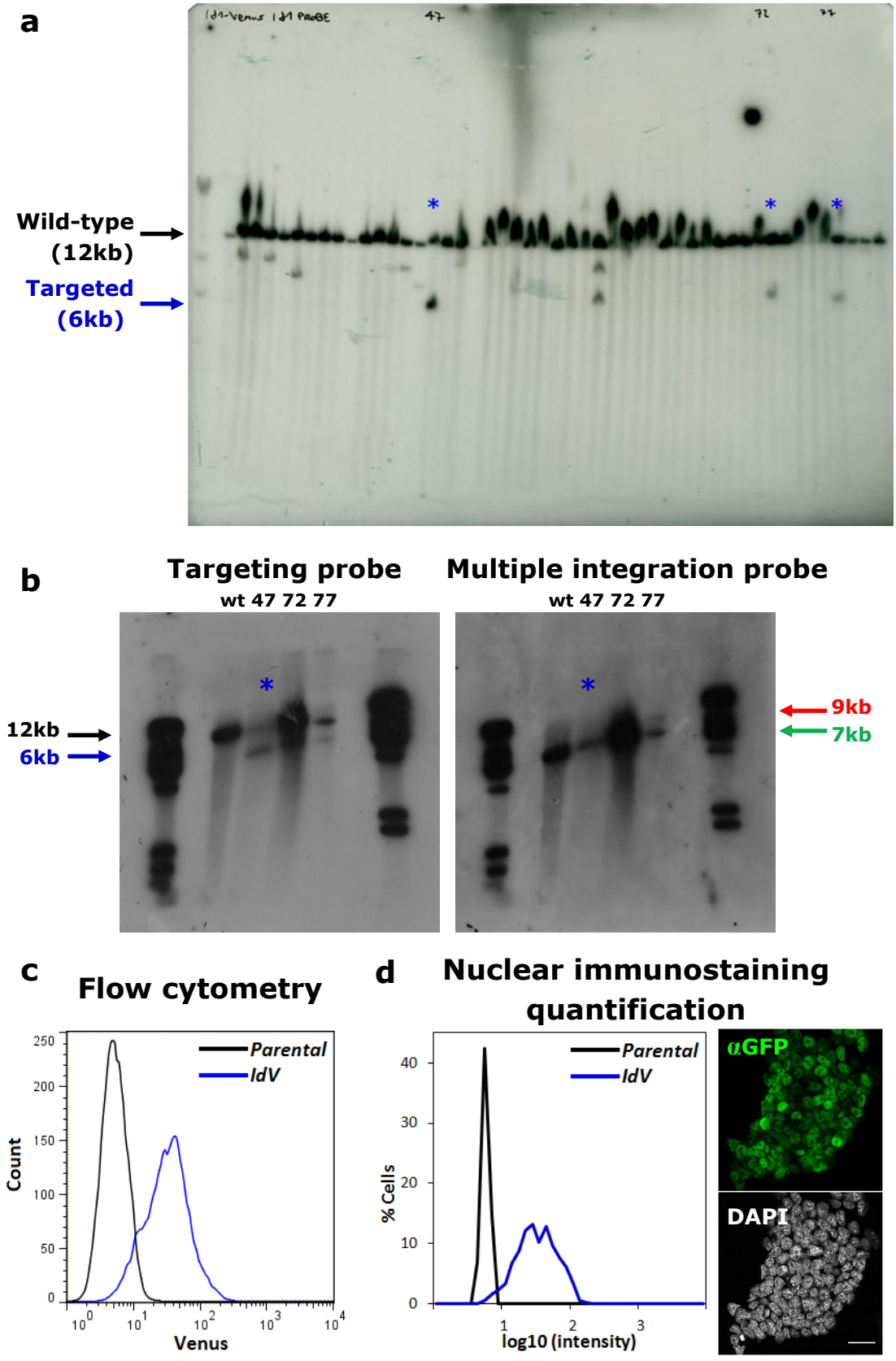

a
Wild-type (12kb)
Targeted (6kb)
b
Targeting probe
wt 47 72 77
Multiple integration probe
wt 47 72 77
12kb
6kb
9kb
7kb
c
Flow cytometry
Parental
IdV
Count
Venus
d
Nuclear immunostaining quantification
Parental
IdV
% Cells
log10 (intensity)
αGFP
DAPI

Figure 3.5 – Screening and identification of an Id1-Venus reporter ES cell clone

a. Southern blot using the targeting probe depicted in Fig.3.4a. The λ DNA HindIII digest (NEB) was loaded to the left of the samples as a size marker. Three clones with the correct banding pattern were identified (clones 47, 72, 77). Black arrow: wild-type band (~12kb), blue arrow: targeted band (~6kb).

b. Repeat of the Southern blot on the parental cell line and the three targeted clones using the targeting probe and a multiple integration probe. The λ DNA HindIII digest (NEB) was loaded to the left of the samples as a size marker.
Green arrow: correct targeting event (~7kb), red arrow: aberrant targeting event (~9kb).
Clone 72 has a more intense wild-type band than targeted band, clone 77 has a second integration event, clone 47 has the correct banding pattern.

c. Flow cytometric analysis of the parental ES cell line and the Id1-Venus reporter ES cell line (IdV clone 47) cultured in LIF+FCS. This experiment was repeated 3 times; the median Venus fluorescence intensities for the biological replicates of the IdV sample have a standard deviation of approximately 15%.

d. Quantification of nuclear Venus signal of parental and IdV ES cells cultured in LIF+FCS and stained with a polyclonal anti-GFP antibody capable of recognising the Venus epitope. The signal from over 900 nuclei was scored for each cell line using the Multicell3D software. An image of a representative IdV ES cell colony stained with the anti-GFP antibody and imaged on a Leica TCS SPE inverted confocal microscope is included to display the expression pattern of Id1-Venus. Scale bar: 30μm.

FIGURE 3.6

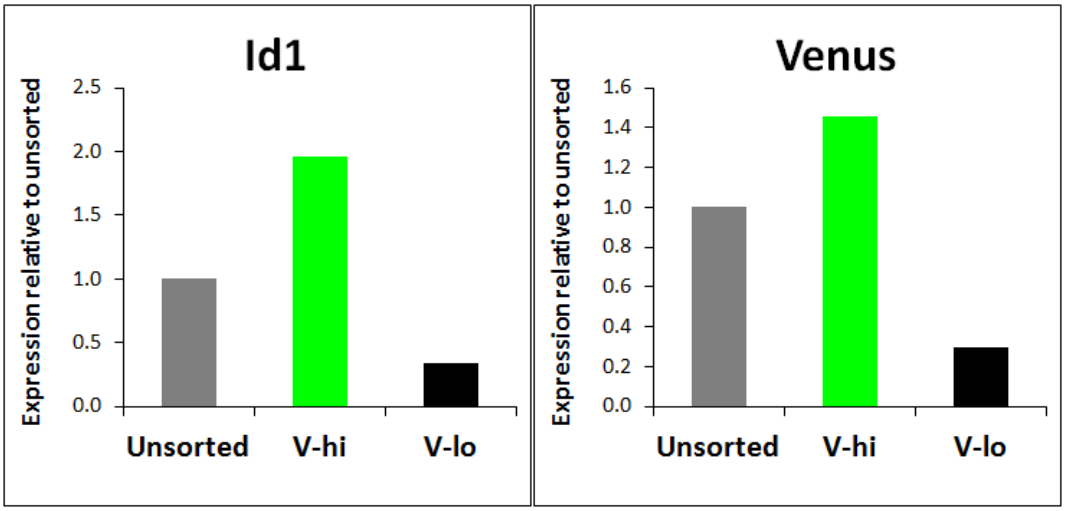

Id1
Expression relative to unsorted
2.5
2.0
1.5
1.0
0.5
0.0
Unsorted
V-hi
V-lo
Venus
Expression relative to unsorted
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Unsorted
V-hi
V-lo

Figure 3.6 – Venus expression recapitulates *Id1* expression

Levels of *Id1* and *Venus* transcript in IdV ES cells cultured in LIF+FCS and sorted by fluorescence-activated cell sorting into Venus-high ("V-hi") and Venus-low ("V-lo") populations. The 20% of live cells displaying the highest or lowest Venus fluorescence were sorted to generate the two populations. RNA was extracted from these samples as well as the starting unsorted population ("Unsorted") and qRT-PCR was performed on reverse transcribed cDNA. Transcript levels were normalised to the levels of *Tbp*. No error bars are displayed as there was only one biological replicate for each sample.

FIGURE 3.7

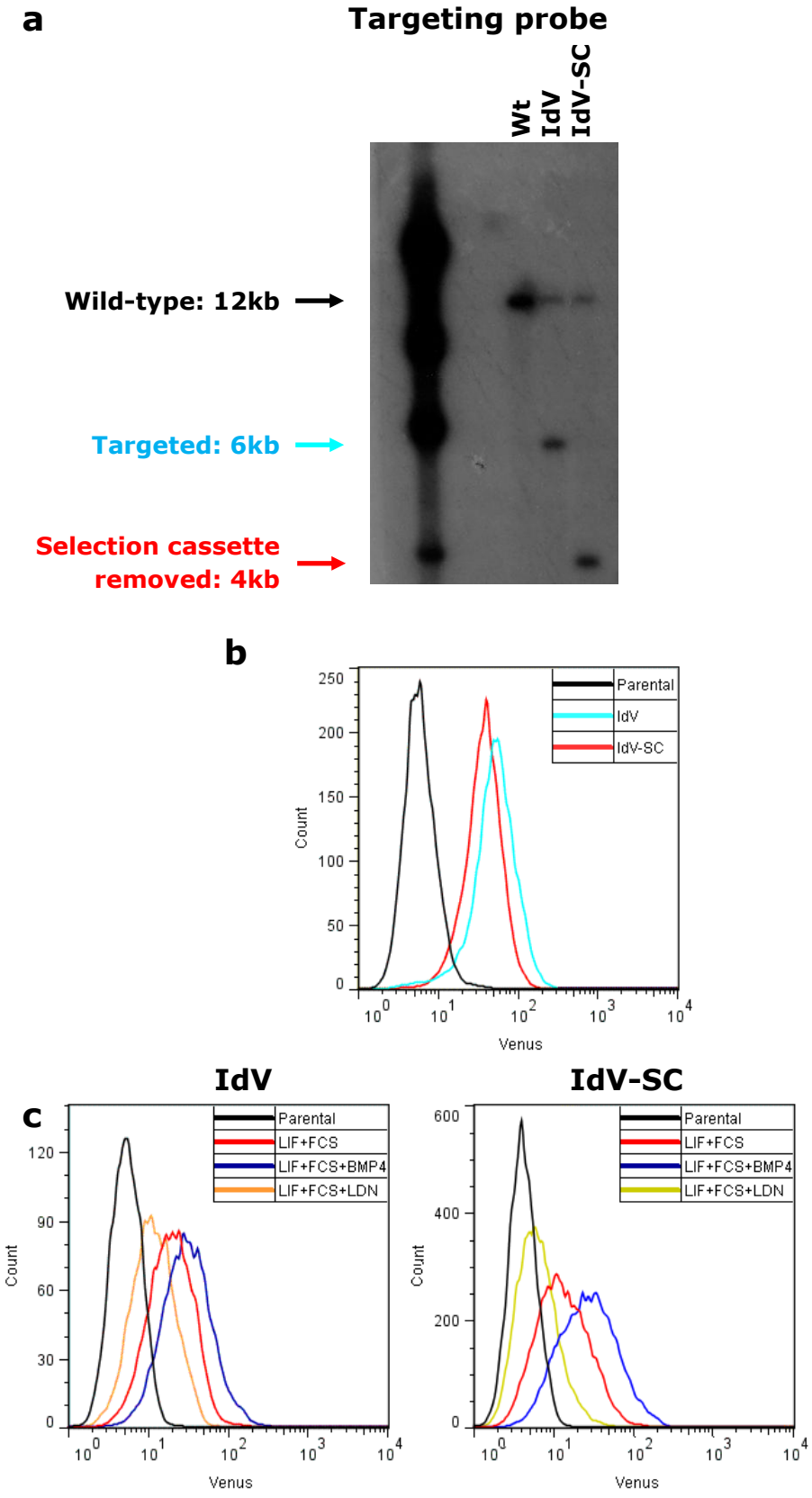

a
Targeting probe
Wt
IdV
IdV-SC
Wild-type: 12kb
Targeted: 6kb
Selection cassette removed: 4kb
b
Parental
IdV
IdV-SC
Count
Venus
c
IdV
IdV-SC
Parental
LIF+FCS
LIF+FCS+BMP4
LIF+FCS+LDN
Count
Venus

Figure 3.7 – Id1 expression is unaffected by the Pgk-Neo cassette

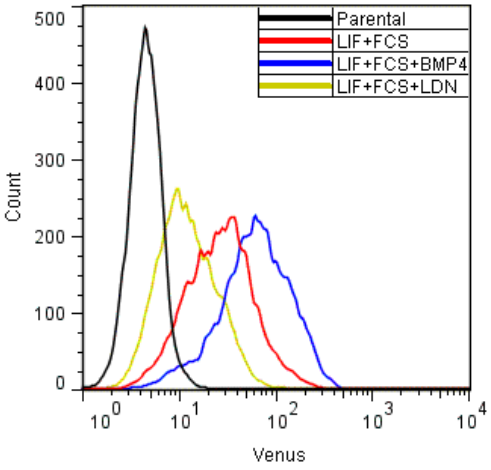
a. Southern blot using the targeting probe depicted in Fig.3.4a on genomic DNA from wild-type ES cells, IdV cells and an IdV clone transfected with a plasmid encoding FlpO which had become subsceptible to G418, suggestive of excision of the *Pgk-Neo* cassette.
Black arrow: wild-type band (~12kb), light blue arrow: targeted band (~6kb), red arrow: targeted band without the selection cassette (~4kb).
The G418-subsceptible clone appears to have lost the selection cassette. It was renamed IdV-SC.

b. Flow cytometric analysis of wild-type, IdV and IdV-SC cells.
The cells were cultured in 2i+10μM SB431542+10ng/ml BMP4 to minimise variations in Id1 expression due to different levels of differentiation in LIF+FCS parallel cultures. Parental non-fluorescent cells are included as a control. This experiment was repeated 3 times; the median Venus fluorescence intensities for the biological replicates of the IdV and IdV-SC samples have a standard deviation of less than 5%.

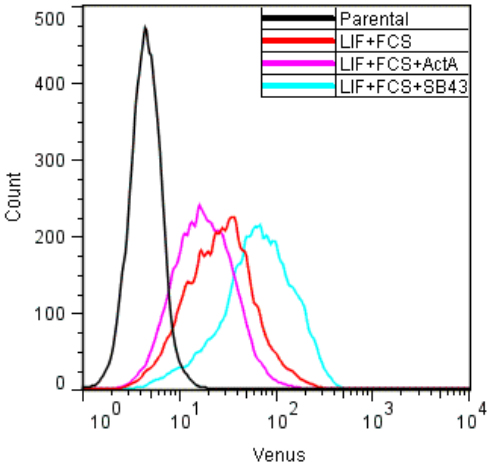
c. Flow cytometric analysis of IdV and IdV-SC ES cells cultured in LIF+FCS and stimulated for 48 hours with 10ng/ml BMP4 or 100nM LDN-193189 (LDN, a small-molecule BMP signalling inhibitor). Parental non-fluorescent cells are included as a control. The experiment was repeated 3 times. The median Venus fluorescence intensity values for the biological replicates of each of the IdV samples (LIF+FCS, LIF+FCS+BMP4, LIF+FCS+LDN-193189) have a standard deviation of less than 15%; the median Venus fluorescence intensity values for the biological replicates of each of the IdV-SC samples have a standard deviation of less than 5%. In all replicate experiments the data follow the same trend as those displayed in figure.

FIGURE 3.8

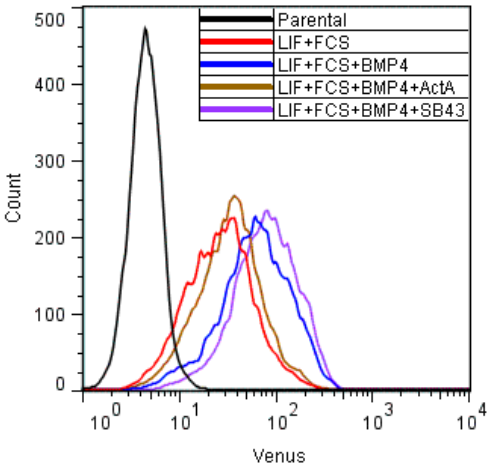
a – BMP response

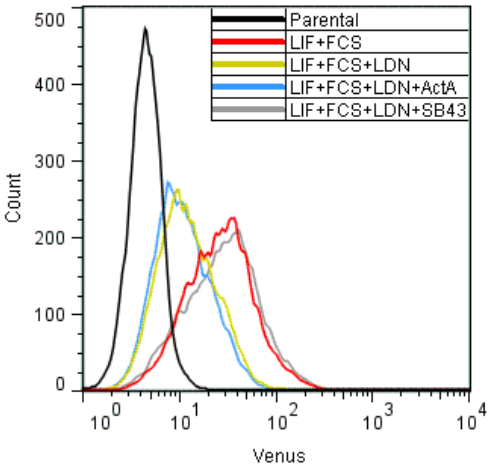


b – Activin response

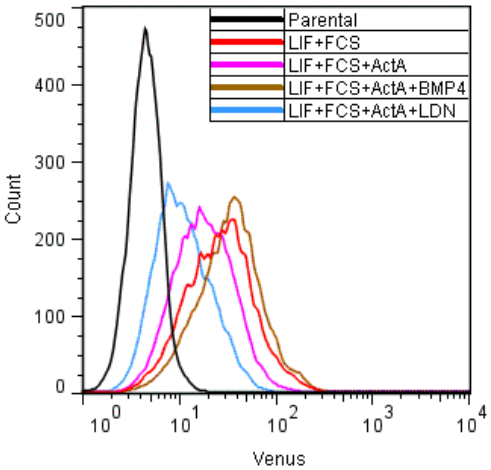


c – Effect of Activin on BMP response





d – Effect of BMP on Activin response



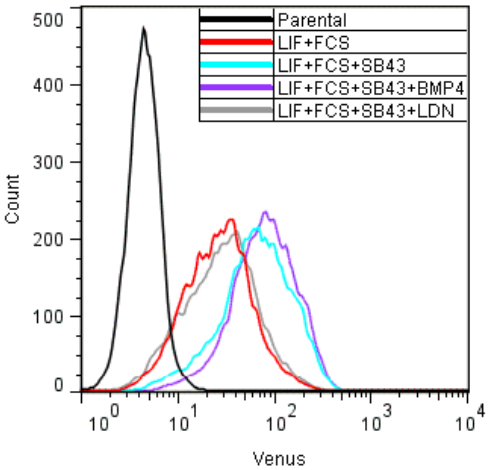


Figure 3.8 – Id1-Venus response to BMP and Activin A stimulation and inhibition in LIF+FCS culture

Flow cytometric analysis of IdV-SC ES cells cultured in LIF+FCS and stimulated with agonists and antagonists of the BMP4 and Activin/Nodal signalling pathways for 48 hours. Parental non-fluorescent cells are included as a control.

a. Id1-Venus response to addition of 10ng/ml BMP4 or 100nM LDN-193189 (LDN, a small-molecule BMP signalling inhibitor).

b. Id1-Venus response to addition of 20ng/ml Activin A (ActA) or 10μM SB431542 (SB43, a small molecule Activin/Nodal signalling inhibitor).

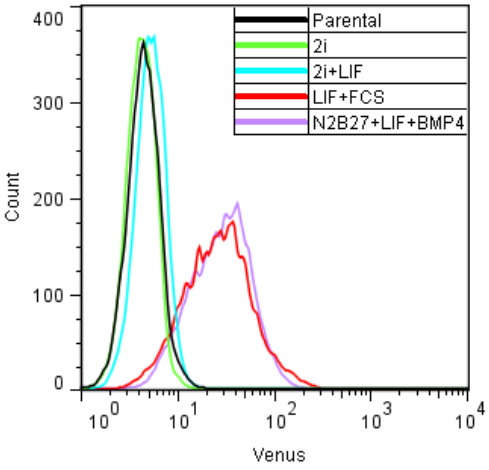
c. Effect of Activin A or SB431542 administration to BMP4- or LDN-193189-treated ES cells on Id1-Venus expression.

d. Effect of BMP4 or LDN-193189 administration to Activin A- or SB431542-treated ES cells on Id1-Venus expression.

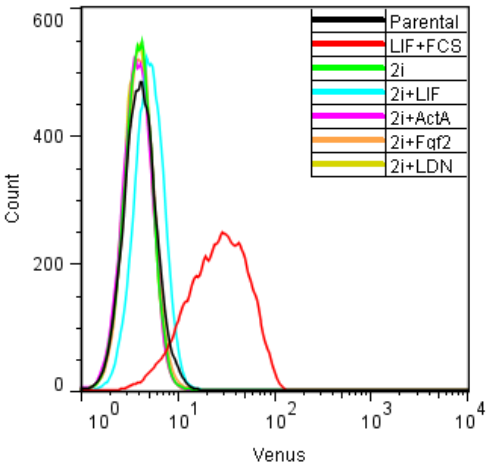
The experiments were repeated 3 times. The median Venus fluorescence intensity values for the biological replicates of each of the samples have a standard deviation of less than 20%. In all replicate experiments the data follow the same trend as those displayed in figure.

FIGURE 3.9

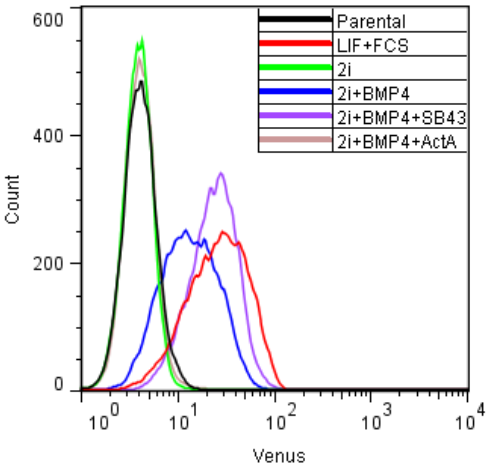
a – ES cell culture

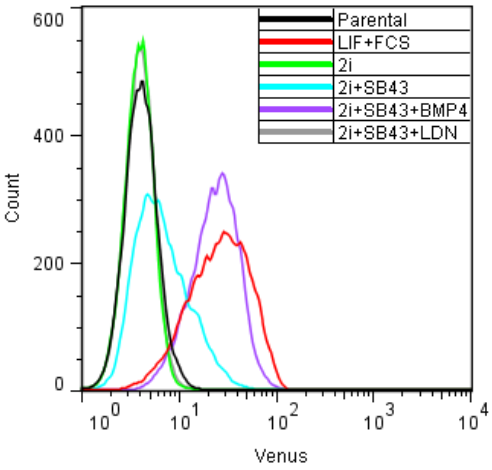


b – 2i culture

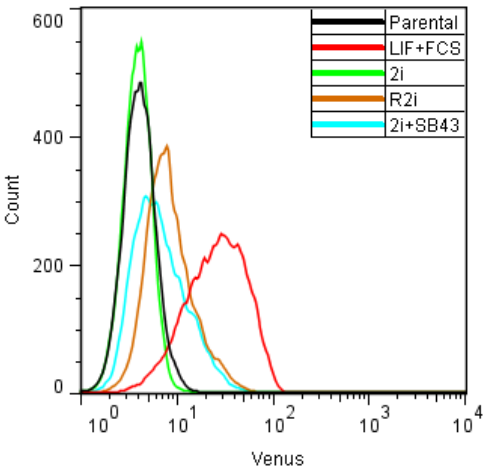


c – BMP/Activin crosstalk in 2i





d – R2i



e – Response to BMP in 2i

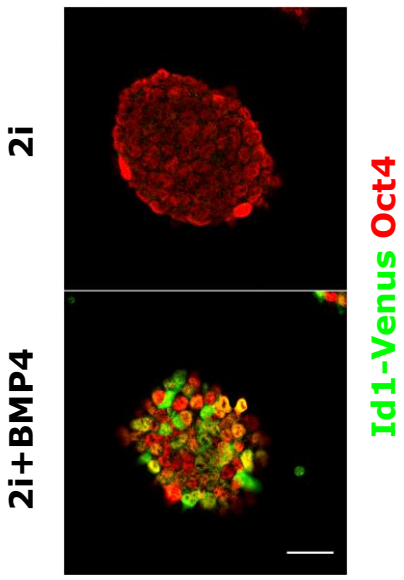


Figure 3.9 – Id1-Venus expression in ES cell cultures and 2i

a.-d. Flow cytometric analysis of IdV-SC ES cells cultured as described below. Parental non-fluorescent cells are included as a negative control.

a. Id1-Venus expression in four alternative ES cell culture conditions: LIF+FCS, N2B27+LIF+BMP4, 2i and 2i+LIF. Cells were cultured in N2B27+LIF+BMP4, in 2i and 2i+LIF for a minimum of 3 passages before performing the flow cytometry experiment.

b. Id1-Venus expression in 2i culture in response to the addition of LIF, 20ng/ml Activin A (ActA), 10ng/ml Fgf2 or 100nM LDN-193189 (LDN) for 48 hours. The LIF+FCS sample is included as a reference for Id1-Venus expression.

c. Id1-Venus expression in 2i culture in response to the addition of 20ng/ml Activin A (ActA), 10ng/ml BMP4, 100nM LDN-193189 (LDN), 10µM SB431542 (SB43) or a combination of these for 48 hours. The LIF+FCS sample is included as a reference for Id1-Venus expression.

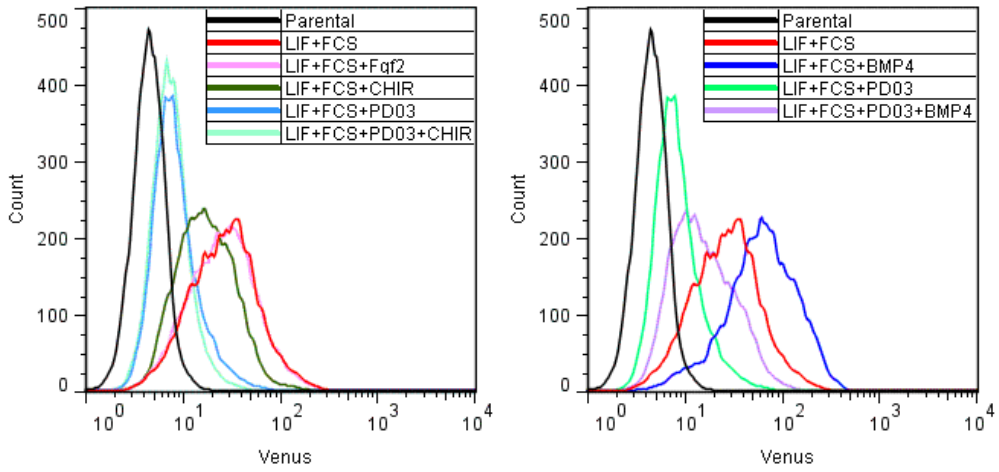
d. Id1-Venus expression in 2i culture following the addition of 10µM SB431542 for 48 hours or following a change of medium to R2i (N2B27+SB431542+PD0325901) for 48 hours. The LIF+FCS sample is included as a reference for Id1-Venus expression.

The experiments were repeated 3 times. The median Venus fluorescence intensity values for the biological replicates of each of the samples have a standard deviation of less than 15%. In all replicate experiments the data follow the same trend as those displayed in figure.

e. ES cells cultured in 2i in the presence or absence of 10ng/ml BMP4 (48 hour stimulation). The samples were fixed and immunostained with anti-GFP and anti-Oct4 antibodies. The cells were imaged on a Leica TCS SPE inverted confocal microscope. Scale bar: 30 µm.

FIGURE 3.10

a – Fgf/CHIR response b – BMP/PD03 crosstalk



c – Combination of antagonists of Id1 expression

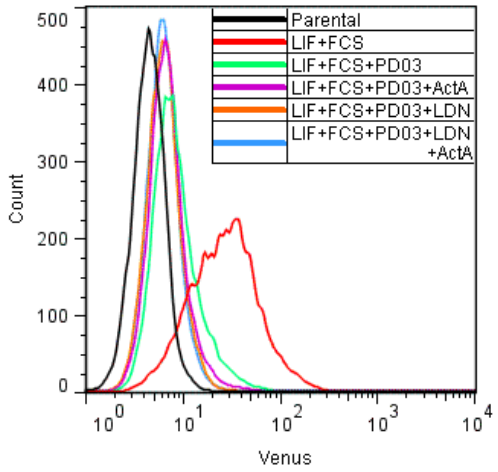


Figure 3.10 – Id1-Venus response to perturbations of GSK3β, Fgf/Mek, Activin/Nodal and BMP signalling in LIF+FCS culture

Flow cytometric analysis of IdV-SC ES cells cultured in LIF+FCS and stimulated with the factors described below for 48 hours. Parental non-fluorescent cells are included as a control.

a. Id1-Venus expression in response to treatment with the inhibitors used in 2i. LIF+FCS cultures were supplemented with 10ng/ml Fgf2, 3μM CHIR99021 (CHIR, a GSK3β inhibitor), 1μM PD0325901 (PD03, a Mek inhibitor), or a combination of these.

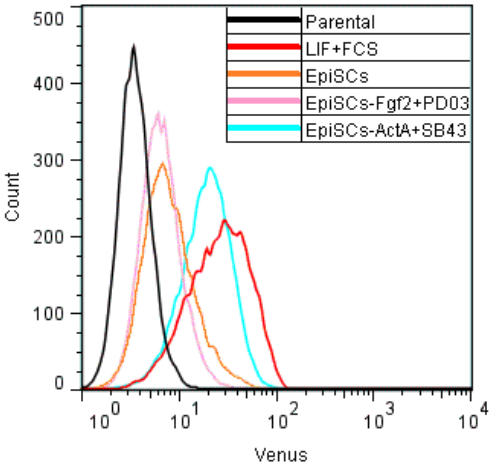
b. Id1-Venus expression following stimulation with 10ng/ml BMP4, 1μM PD0325901 (PD03), or both.

c. Id1-Venus expression following stimulation with 20ng/ml Activin A, 100nM LDN-193189 (LDN), 1μM PD0325901 (PD03), or combinations of these molecules.

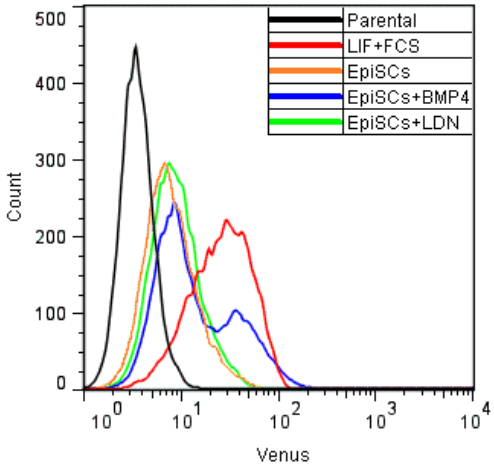
The experiments were repeated 3 times. The median Venus fluorescence intensity values for the biological replicates of each of the samples have a standard deviation of less than 20%. In all replicate experiments the data follow the same trend as those displayed in figure.

FIGURE 3.11

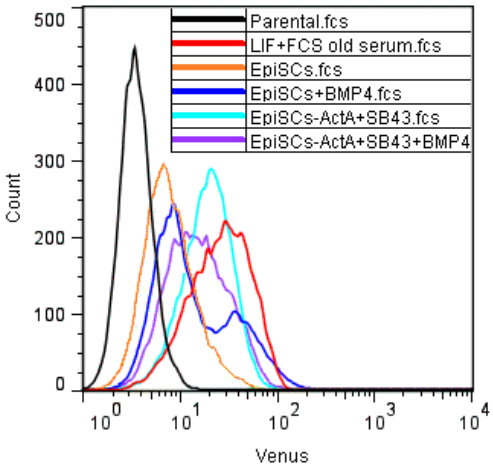
a – Removal of Fgf2/ActA



b – Response to BMP



c – Combination of BMP and SB43



d – Response to BMP

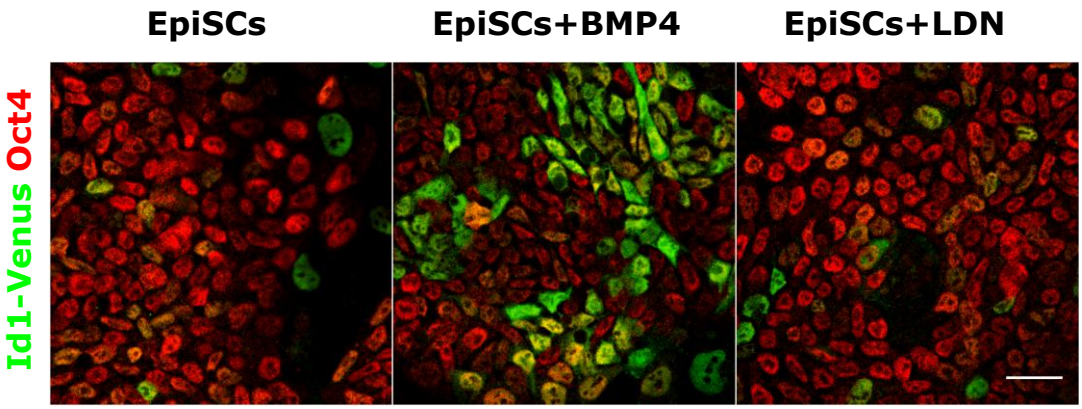


Figure 3.11 – Id1-Venus expression in epiblast stem cells

a.-c. Flow cytometric analysis of Id1-Venus expression in IdV-SC EpiSCs derived from ES cell cultures and cultured for 48 hours as described below. Parental non-fluorescent cells and IdV-SC ES cells cultured in LIF+FCS are included as a reference.

a. EpiSCs cultured in N2B27 supplemented with 20ng/ml Activin A and 10ng/ml Fgf2 (EpiSCs), in N2B27 supplemented with 20ng/ml Activin A and 1μM PD0325901 (EpiSCs-Fgf2+PD03), in N2B27 supplemented with 10ng/ml Fgf2 and 10μM SB431542 (EpiSCs-ActA+SB43).

b. Response of EpiSCs to addition of 10ng/ml BMP4 or 100nM LDN-193189 (LDN).

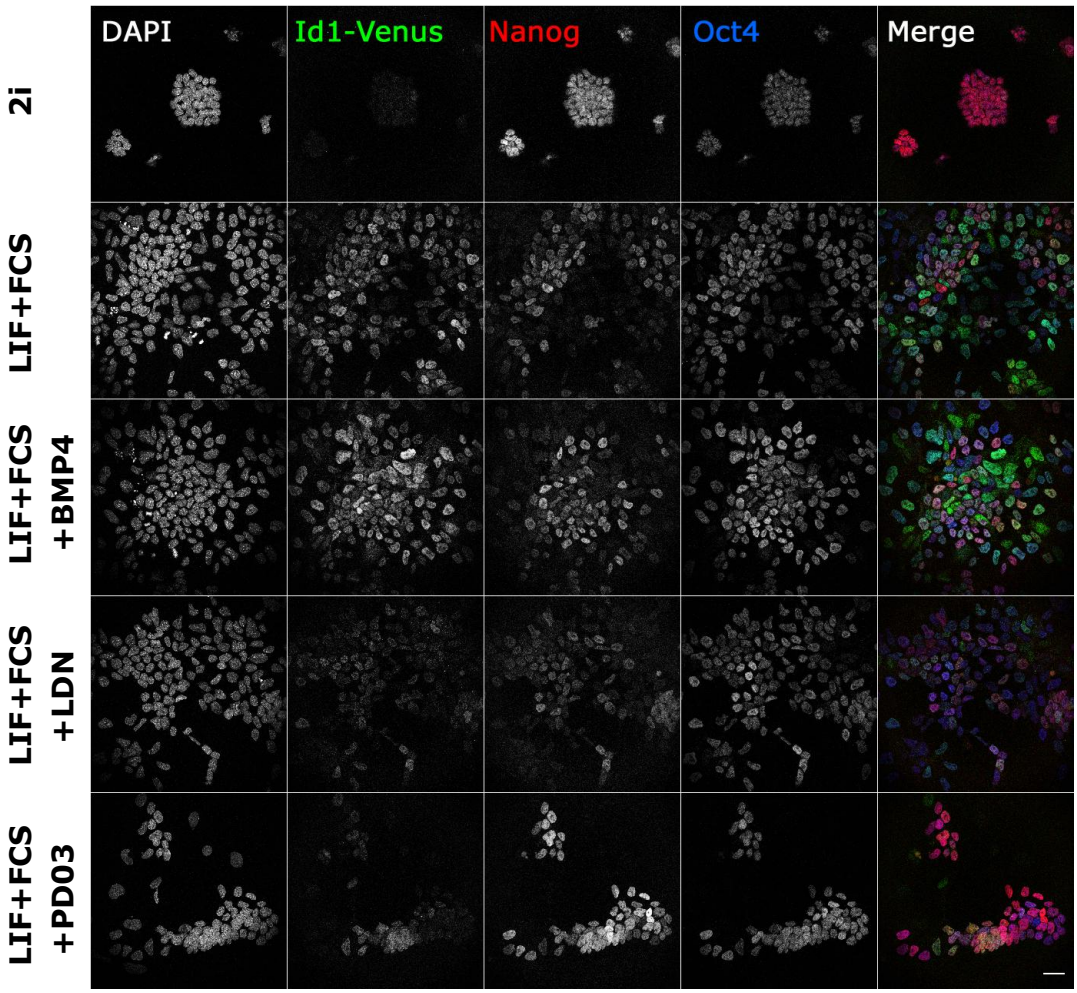
c. Response of EpiSCs to simultaneous stimulation of BMP4 and inhibition of Activin/Nodal signalling. EpiSCs were cultured as described above and in N2B27 supplemented with 10ng/ml Fgf2, 10ng/ml BMP4 and 10μM SB431542 (EpiSCs-ActA+SB43+BMP4).

The experiments were repeated 3 times. The median Venus fluorescence intensity values for the biological replicates of each of the samples have a standard deviation of less than 15%, with the exception of the samples supplemented with BMP4 and SB431542, for which the standard deviation is approximately 25%, potentially as a result of varying cell density and/or of the onset of differentiation (see main text). In all replicate experiments the data follow the same trend as those displayed in figure.

d. EpiSCs cultured in the presence of 10ng/ml BMP4, 100nM LDN-193189 (LDN) or in the absence of both molecules for 48 hours, fixed and immunostained with anti-GFP and anti-Oct4 antibodies. The cells were imaged on a Leica TCS SPE inverted confocal microscope. Scale bar: 30 μm.

FIGURE 3.12

a



b

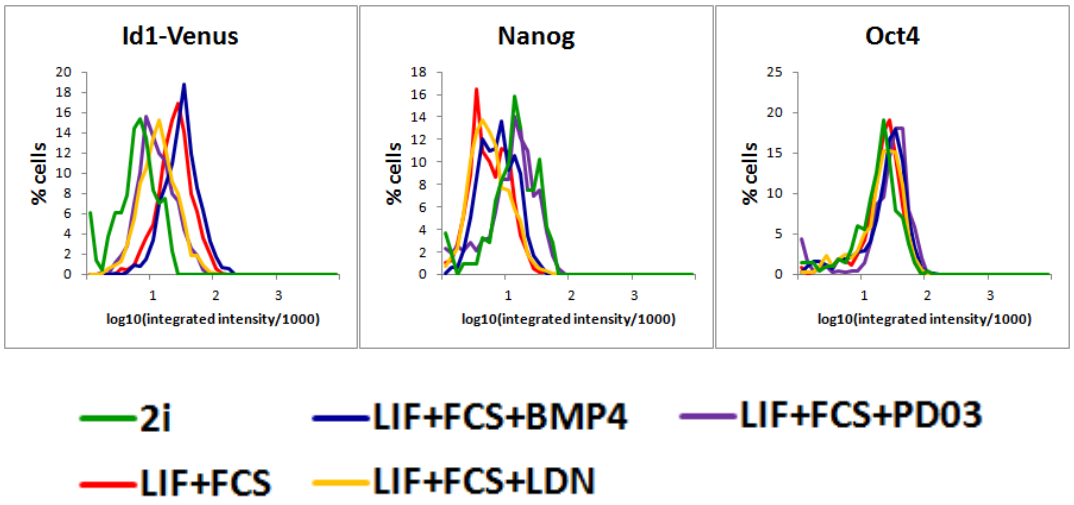


Figure 3.12 – Immunostaining quantification of Id1, Nanog and Oct4 in LIF+FCS

Response of IdV ES cells to BMP4 stimulation, BMP inhibition and Mek inhibition in LIF+FCS culture. IdV ES cells cultured in 2i were analysed as a negative control for Id1-Venus expression and as a positive control for Nanog expression.

a. IdV ES cells cultured in 2i, in LIF+FCS, in LIF+FCS with addition of 10ng/ml BMP4, 100nM LDN-193189 (LDN, BMP inhibitor) or 1µM PD0325901 (PD03, Mek inhibitor) as indicated in figure, immunostained for GFP, Nanog, Oct4 and counterstained with DAPI. The cells were imaged on a Leica TCS SPE inverted confocal microscope. Scale bar: 30 µm.

b. Immunostaining quantification for Id1-Venus, Nanog and Oct4 in all culture conditions in (a). A minimum of 400 nuclei were scored for each of the LIF+FCS-based culture conditions, and over 200 nuclei were scored for 2i. A minimum of five fields of view was used for each condition.

FIGURE 3.13

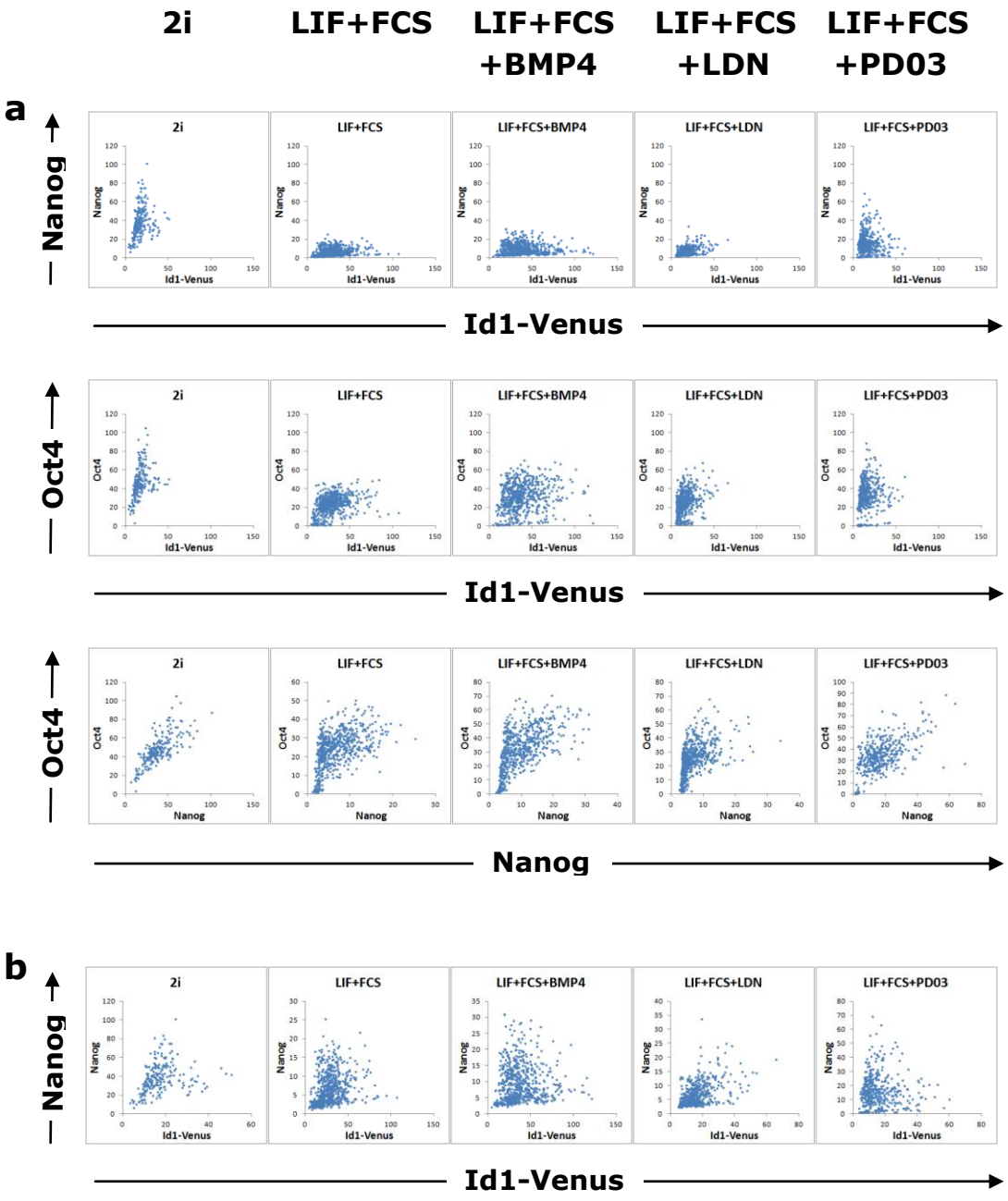

2i
LIF+FCS
LIF+FCS +BMP4
LIF+FCS +LDN
LIF+FCS +PD03
a
Nanog
Id1-Venus
Oct4
Id1-Venus
Oct4
Nanog
b
Nanog
Id1-Venus

Figure 3.13 – Immunostaining quantification of Id1, Nanog and Oct4 in LIF+FCS (2)

a. Dot plots indicating the relationship between Id1-Venus, Nanog and Oct4 in the samples shown in Fig 3.12. A minimum of 400 nuclei were scored for each of the LIF+FCS-based culture conditions, and over 200 nuclei were scored for 2i. A minimum of five fields of view was used for each condition.

b. Zoom in on the Id1-Venus over Nanog plots shown in (a). Note that the axes in these plots have different maximum fluorescent intensity values.

FIGURE 3.14

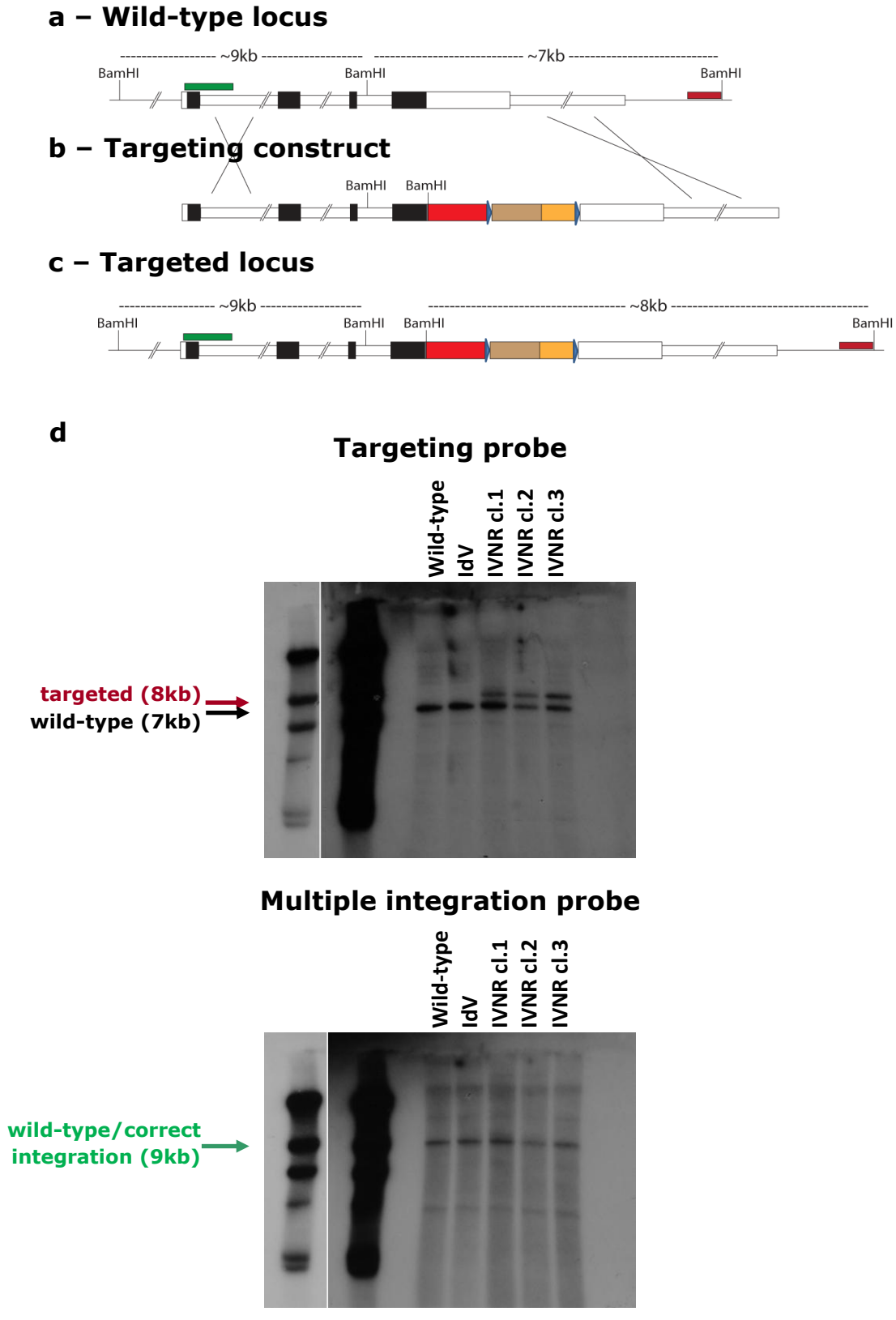

a – Wild-type locus
~9kb
~7kb
BamHI
BamHI
BamHI
b – Targeting construct
BamHI
BamHI
c – Targeted locus
~9kb
~8kb
BamHI
BamHI
BamHI
BamHI
d
Targeting probe
Wild-type
IdV
IVNR cl.1
IVNR cl.2
IVNR cl.3
targeted (8kb)
wild-type (7kb)
Multiple integration probe
Wild-type
IdV
IVNR cl.1
IVNR cl.2
IVNR cl.3
wild-type/correct integration (9kb)

Figure 3.14 – Generation of an Id1-Venus Nanog-tagRFP double reporter ES cell line

a-c Nanog-tagRFP targeting strategy. The sizes of bands detected on Southern blot membranes following genomic DNA digestion with BamHI are indicated for the various alleles.

a. Wild-type *Nanog* locus. Single horizontal lines: intergenic DNA, medium-sized white boxes: homology regions with targeting vector, big white boxes: untranslated regions, black boxes: coding regions, green box: Southern blot multiple integration probe, dark red box: Southern blot targeting probe.

b. Targeting vector containing a flexible linker followed by the *tagRFP* cDNA (red box) fused in frame with the 3' end of *Nanog*.
Blue triangles: *loxP* sites, brown box: internal ribosome entry site, yellow box: *Bls* (blasticidin S resistance gene).

c. *Nanog* locus following a correct targeting event.

d. Southern blots on wild-type, IdV and three blasticidin-resistant clones, using the targeting probe and the multiple integration probe depicted in (a). The λ DNA HindIII digest (New England Biolabs) was loaded to the left of the samples as a size marker.
Black arrow: wild-type band (~7kb), dark red arrow: targeted band (~8kb), green arrow: wild-type locus and correct integration event (~9kb).

FIGURE 3.15

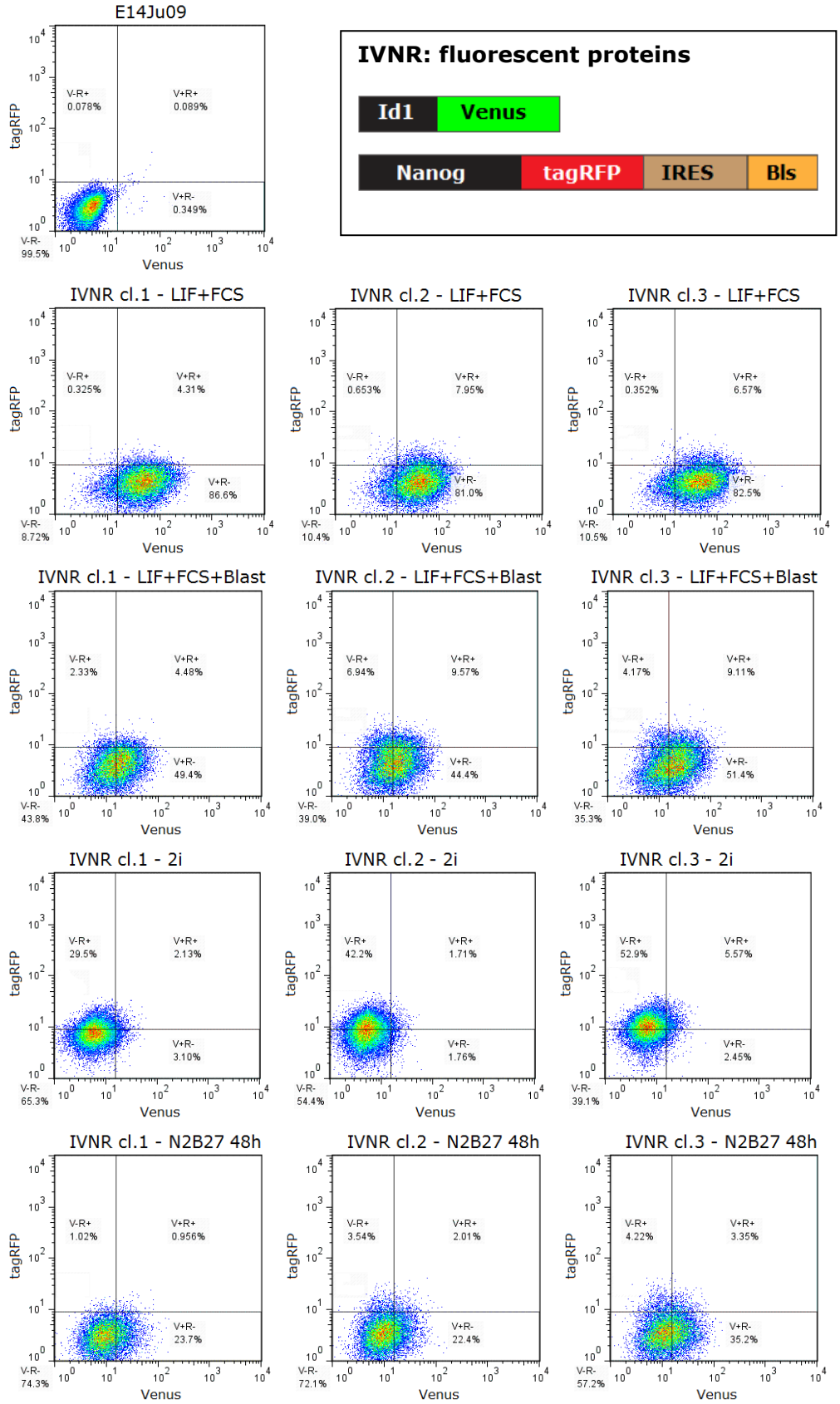

E14Ju09
V-R+ 0.078%
V+R+ 0.089%
V+R- 0.349%
V-R- 99.5%
tagRFP
Venus
IVNR: fluorescent proteins
Id1
Venus
Nanog
tagRFP
IRES
Bls
IVNR cl.1 - LIF+FCS
V-R+ 0.325%
V+R+ 4.31%
V+R- 86.6%
V-R- 8.72%
IVNR cl.2 - LIF+FCS
V-R+ 0.653%
V+R+ 7.95%
V+R- 81.0%
V-R- 10.4%
IVNR cl.3 - LIF+FCS
V-R+ 0.352%
V+R+ 6.57%
V+R- 82.5%
V-R- 10.5%
IVNR cl.1 - LIF+FCS+Blast
V-R+ 2.33%
V+R+ 4.48%
V+R- 49.4%
V-R- 43.8%
IVNR cl.2 - LIF+FCS+Blast
V-R+ 6.94%
V+R+ 9.57%
V+R- 44.4%
V-R- 39.0%
IVNR cl.3 - LIF+FCS+Blast
V-R+ 4.17%
V+R+ 9.11%
V+R- 51.4%
V-R- 35.3%
IVNR cl.1 - 2i
V-R+ 29.5%
V+R+ 2.13%
V+R- 3.10%
V-R- 65.3%
IVNR cl.2 - 2i
V-R+ 42.2%
V+R+ 1.71%
V+R- 1.76%
V-R- 54.4%
IVNR cl.3 - 2i
V-R+ 52.9%
V+R+ 5.57%
V+R- 2.45%
V-R- 39.1%
IVNR cl.1 - N2B27 48h
V-R+ 1.02%
V+R+ 0.956%
V+R- 23.7%
V-R- 74.3%
IVNR cl.2 - N2B27 48h
V-R+ 3.54%
V+R+ 2.01%
V+R- 22.4%
V-R- 72.1%
IVNR cl.3 - N2B27 48h
V-R+ 4.22%
V+R+ 3.35%
V+R- 35.2%
V-R- 57.2%

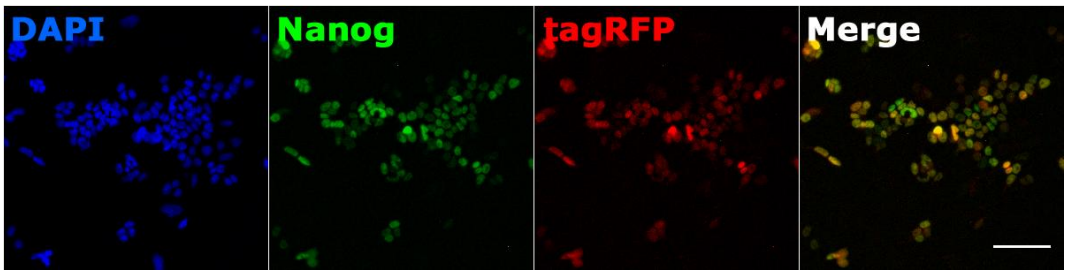
Figure 3.15 – Flow cytometric analysis of IVNR cells

Density plots generated by flow cytometric analysis of the levels of Venus and tagRFP in three IVNR clones cultured in LIF+FCS, LIF+FCS+10µg/ml blasticidin S (to select for expression from the *Nanog-tagRFP* locus), 2i or N2B27 for 48 hours, alongside a non-fluorescent control (E14Ju09). Density ranges from red (highest) to blue (lowest). Percentages of cells within each of the quadrants drawn in the figure are displayed within the density plots. The experiment was performed once for all clones and conditions.

A diagrammatic representation of the fluorescent protein transcripts present in IVNR ES cells is included.

FIGURE 3.16

a



b

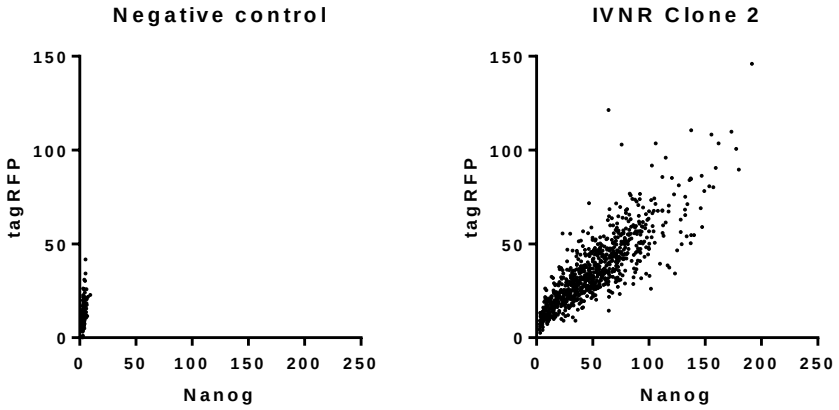


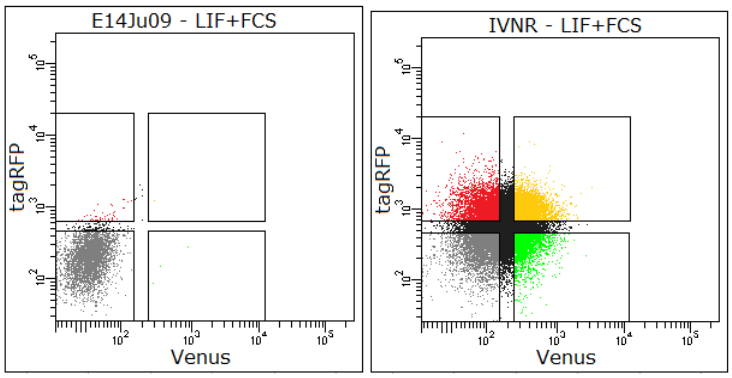
Figure 3.16 – tagRFP accurately reports Nanog expression in IVNR ES cells

a. IVNR clone 2 ES cells cultured in LIF+FCS, fixed and immunostained with anti-Nanog and anti-tagRFP antibodies. Nuclei were counterstained with DAPI. The cells were imaged on an Olympus IX51 microscope. Scale bar: 100µm.

b. Nuclear immunostaining quantification of the cells in (a) alongside a parental (E14Ju09) control stained with an anti-tagRFP primary antibody and with secondary antibodies against both anti-Nanog and anti-tagRFP ("Negative control"). A minimum of 950 nuclei were scored for each condition.

FIGURE 3.17

a



b

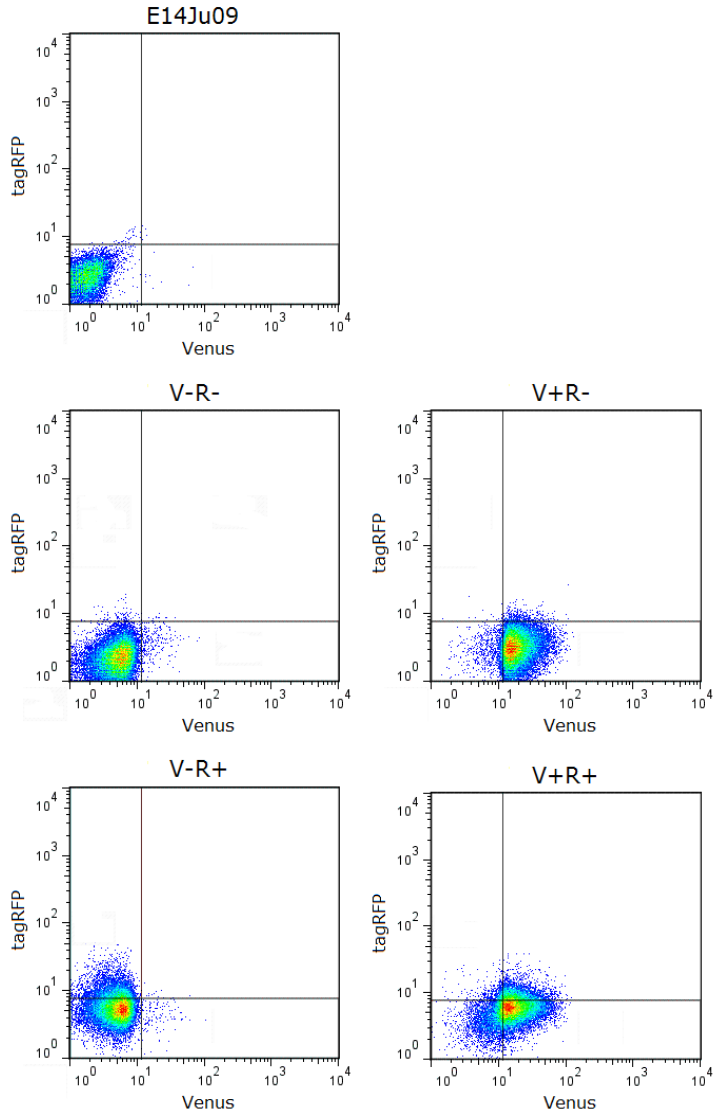


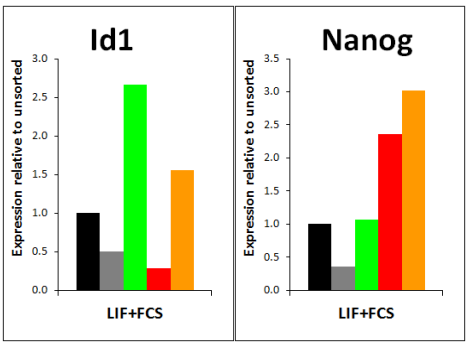
Figure 3.17 – tagRFP-based sorting does not purify the tagRFP-positive population

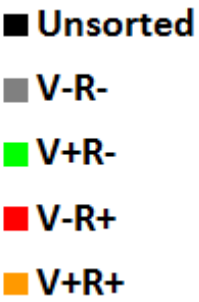
a. Fluorescence-activated cell sorting of IVNR clone 2 ES cells cultured in LIF+FCS, based on expression of Venus and tagRFP as illustrated in figure. E14Ju09 ES cells were used as a negative control.
The sorting was performed on a BD FACSAria sorter.

b. Flow cytometric analysis of the sorted cells immediately after sorting. V: Venus, R: tagRFP, -: negative, +: positive.
The analysis was performed on a BD LSRFortessa analyser, hence the gates are not in the same place on the graphs as the data generated by the BD FACSAria sorter in (a). The gates are positioned similarly to those in (a) with respect to the negative control sample.

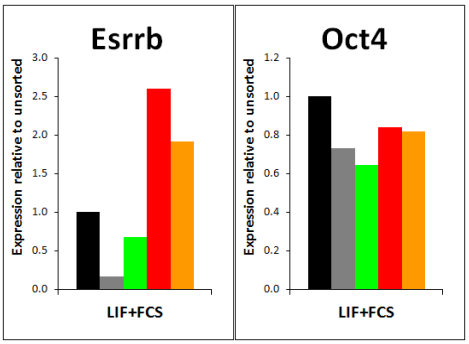
FIGURE 3.18

a – Id1, Nanog

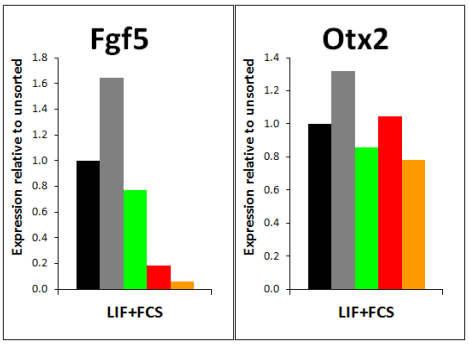




b – Pluripotency markers



c – Primed pluripotency



d – Priming/differentiation markers

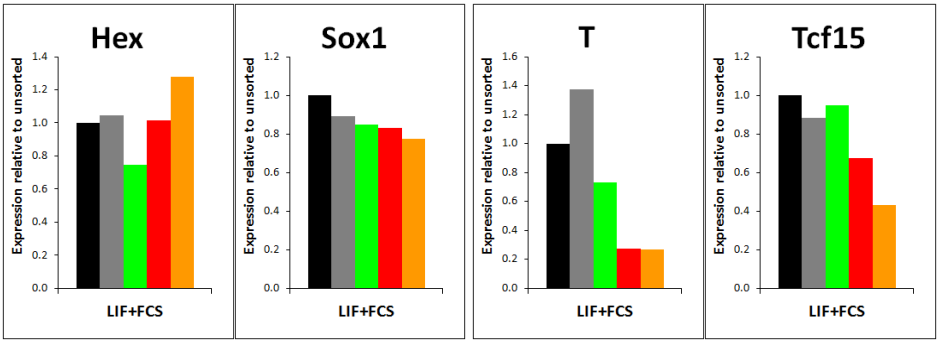
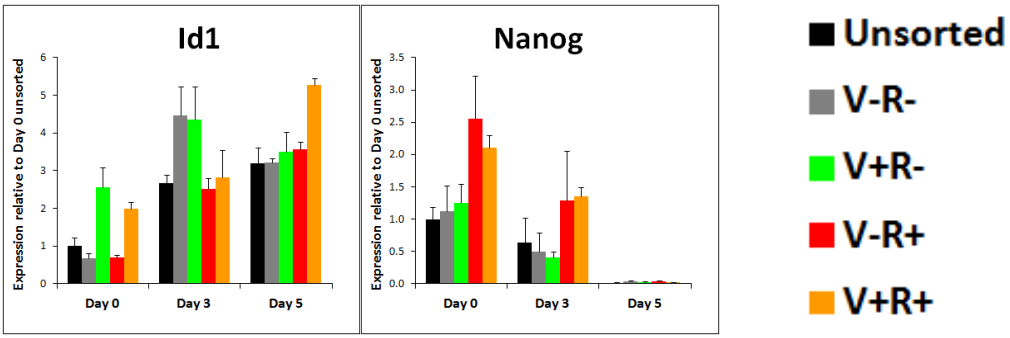


Figure 3.18 – tagRFP and Venus accurately report Nanog and Id1 expression

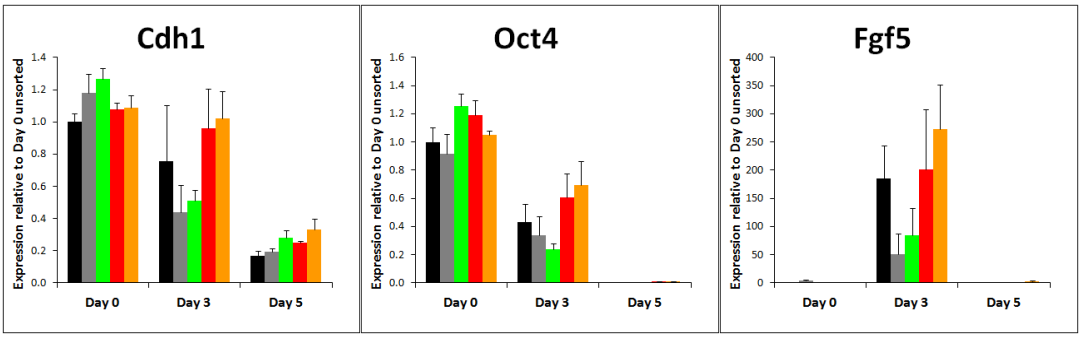
qRT-PCR on IVNR clone 2 ES cells cultured in LIF+FCS sorted based on Venus and tagRFP expression as illustrated in Fig.3.17a. V: Venus, R: tagRFP, -: negative, +: positive.

FIGURE 3.19

a – Id1, Nanog



b – Pluripotent epiblast



c – Neural ectoderm

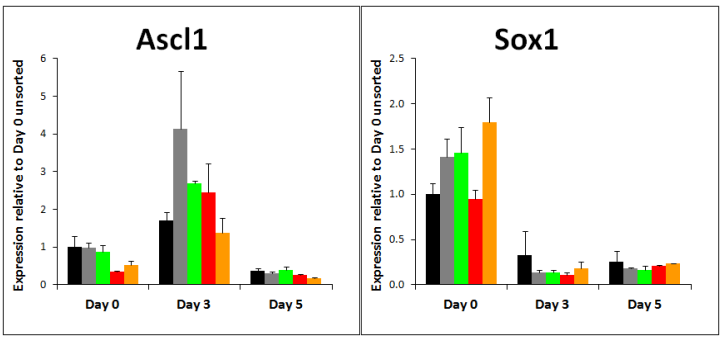
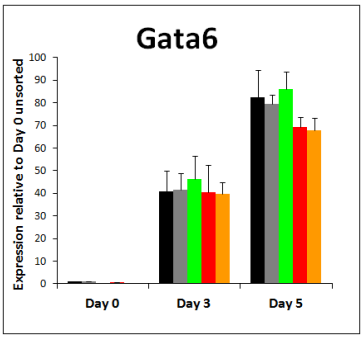


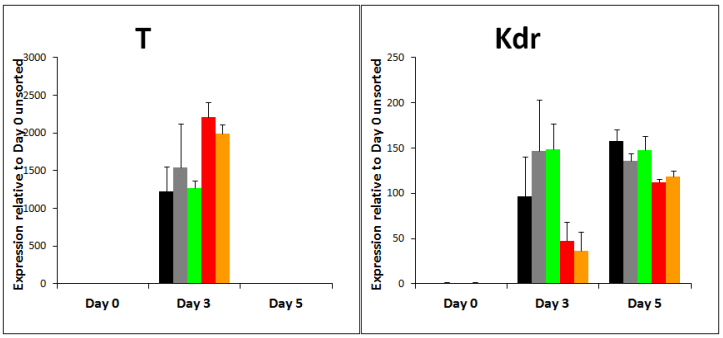
FIGURE CONTINUED ON THE NEXT PAGE

FIGURE CONTINUED FROM THE PREVIOUS PAGE

d – Endoderm



e – Primitive streak, mesoderm



f – Exit from epiblast (multilineage markers)

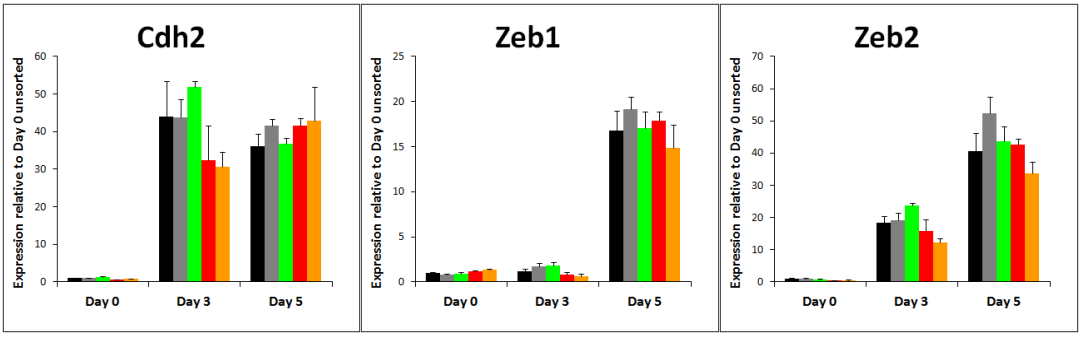


Figure 3.19 – Expression of Nanog, but not of Id1, affects kinetics of EB differentiation

qRT-PCR on IVNR clone 2 ES cells cultured in LIF+FCS, sorted based on Venus and tagRFP expression as illustrated in Fig.3.17a and differentiated into embryoid bodies in GMEM+10% FCS. V: Venus, R: tagRFP, -: negative, +: positive.
Error bars represent the standard error of the mean of three independent biological replicates.

FIGURE 3.20

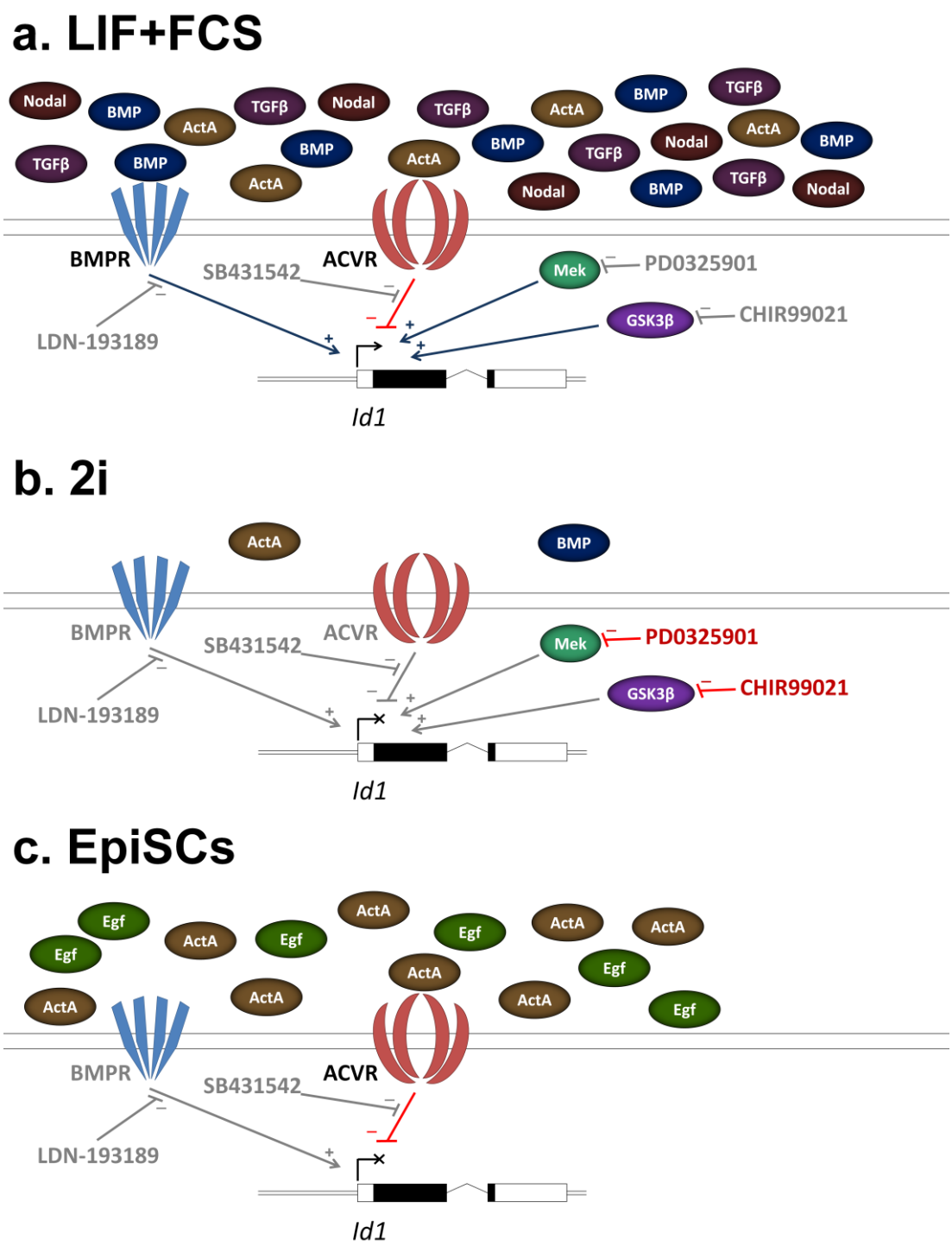

a. LIF+FCS
Nodal
BMP
ActA
TGFβ
BMPR
ACVR
SB431542
LDN-193189
Mek
PD0325901
GSK3β
CHIR99021
Id1
b. 2i
c. EpiSCs
Egf

Figure 3.20 – Regulation of Id1 expression in pluripotent cells

Diagramatic representation of the signalling pathways shown in this Chapter to regulate the expression of Id1 in ES cells and epiblast stem cells.

a. In LIF+FCS culture, BMP and Activin/Nodal signalling agonists are present in the culture medium. BMP promotes Id1 expression, Activin/Nodal agonists inhibit it. Mek and GSK3β also positively regulate Id1 expression. The combinatorial effect of these pathways results in heterogeneous Id1 expression within the culture.

b. In 2i culture, Mek and GSK3β inhibition have a profound inhibitory effect on Id1 expression, and BMP signals are low or absent from the culture medium. Id1 is therefore not expressed in these culture conditions.

c. In epiblast stem cell culture, exogenous Fgf2 and Activin A are present in the culture medium. Fgf2 does not affect Id1 expression, Activin A represses Id1. BMP signals are low or absent from the culture medium. The factor(s) that drive low levels of Id1 expression in a small subset of cells are currently unknown.

BMPR: BMP receptor; ACVR: Activin receptor; ActA: Activin A.

CHAPTER 4

The effects of Id1 overexpression in the differentiation of ES cell towards the neural lineage

4.1 Introduction

Id1 is a key downstream effector of BMP4 signalling in the maintenance of ES cell pluripotency and in the modulation of ES cell differentiation. The overexpression of the helix-loop-helix factor can sustain ES cell self-renewal in combination with LIF, and can recapitulate the block on the neural differentiation of ES cells imposed by BMP4 stimulation (Finley *et al.* 1999, Hollnagel *et al.* 1999, Ying *et al.* 2003a, Zhang *et al.* 2010).

BMP4 is required for mesoderm formation *in vivo*, with mutants either failing to express the primitive streak determinant *T* or exhibiting mesodermal defects until E9.5 when they invariably die (Winnier *et al.* 1995). It is therefore not surprising that serum and BMP4 have been employed in the differentiation of ES cells towards mesodermal lineages (Keller *et al.* 1993, Johansson & Wiles 1995, Nishikawa *et al.* 1998, Park *et al.* 2004).

However, BMP4 has also been used to generate endoderm and surface ectoderm from mouse ES cells (Coraux *et al.* 2003, Gouon-Evans *et al.* 2006), and to derive trophoblast from human ES cells (Sudheer *et al.* 2012), raising the question of what its precise role may be, and to which, if any, of these cell types is Id1 directing differentiation towards.

ES cells cultured in an adherent monolayer in the absence of LIF and the presence of BMP4 generate a mixture of two cell types: colonies of densely packed cells expressing Oct4, and flat epithelial sheets with “big flat cells” co-expressing Cdh1 and Tcfap2a (also known as AP-2α), a combination of markers that characterises surface ectoderm and trophoblast (Arkell & Beddington 1997, Mitchell *et al.* 1991)

(Figure 4.1). It would therefore appear that even in the simplest of BMP4 stimulation experiments a complex phenotype arises.

I set out to characterise the effect of Id1 overexpression on ES cell differentiation towards the neural lineage in greater detail. I aimed to discover whether the Id1-imposed inhibition of neural specification was due to a delay in the exit from pluripotency, a switch in differentiation trajectory, or both; to assess to what differentiated cell type seen in BMP4-treated cultures Id1 was directing differentiation towards, if any; to understand whether the presence of Id1 in wild-type differentiating cells may act to delay or block their commitment to the neural lineage.

4.2 Results

4.2.1 Generation and validation of Id1 overexpressing ES cell lines

4.2.1.1 Generation of Id1 overexpressing ES cell lines

In order to assess the effect of Id1 overexpression in differentiation, I transfected E14tg2a ES cells with the Id1 overexpression plasmid used by Ying *et al.* (2003a). This plasmid, diagrammatically depicted in Figure 4.2a, contains the strong *CAG* promoter (Niwa *et al.* 1991) driving the expression of the Flag-tagged cDNA of *Id1*, which is followed by an internal ribosome entry site (*IRES)*, a puromycin resistance gene (*Pac*), and a polyadenylation signal sequence (*pA*). The *Flag-Id1-IRES-Pac-pA* sequence is flanked by *loxP* sites in the same orientation, and followed by the cDNA of *GFP* and a second polyadenylation signal. This allows for the Cre recombinase-mediated excision of the *Id1* transgene from the genome of ES cells and expression of *GFP* in its place. I also transfected E14tg2a ES cells with a plasmid containing the *CAG* promoter driving expression of *Pac* to generate an "empty vector" control cell line (for more information refer to sections 2.2.4.1.2 and 2.2..4.1.4 of the Materials

and Methods chapter). I selected for stable transfectants by culturing the transfected cells in 2µg/ml puromycin and expanded clonal cell lines.

I performed a qRT-PCR analysis of RNA samples extracted from a clonal empty vector control cell line and three Id1 overexpressing clones cultured in LIF+FCS (Figure 4.2b). All three lines expressed higher levels of *Id1* than the control ES cells, ranging between 20- and 65-fold higher expression. I then assessed expression at the protein level by performing Western blots on the protein lysate of the same samples cultured in LIF+FCS and in N2B27 for 5 days (Figure 4.2c,d). Probing the Western blot membrane with anti-Id1 antibody (Santa Cruz SC-488) revealed that Id1 protein was overexpressed in both ES cells and cells that had undergone neural differentiation. The wild-type and Flag-tagged exogenous proteins are distinguishable due to the increased size of the exogenous protein. It appears that Id1 expression is lost from all cell lines upon differentiation: the endogenous protein can no longer be detected in any of the cell lines, and the transgenic protein is downregulated in both clones 7 and 11, and possibly in clone 4, which was already expressing low levels of the transgene in LIF+FCS culture.

4.2.1.2 The Id1 overexpressing ES cell lines lack a functional Flag epitope

I performed a Western blot using the lysates described above and probed the membrane with an anti-Flag antibody (Figure 4.2d). Surprisingly, none of the clones produced detectable bands, unlike a positive control lysate of wild-type ES cells transfected with a plasmid encoding *Flag-Neurod1*. Immunostaining of fixed ES cells from the same cell lines also resulted in no Flag signal being detected (data not shown).

I re-sequenced the plasmid used to make the cell lines (Figure 4.3a). Alignment of the sequencing read to the sequence of endogenous *Id1* revealed that the final lysine residue on the Flag epitope was also the second residue of wild-type Id1 (Figure 4.3b,c). It is therefore possible that the biochemical and structural properties of the

transgenic protein mask the Flag epitope making it impossible to detect by Western blot and immunofluorescence. It is worth noting that Ying *et al.* (2003a) observed a protein band of the same size to that detected in Figure 4.2c when performing Western blots with anti-Id1 antibody, but that they did not perform Flag immunoblots or immunostains. Hence, it is possible that the cell lines they used also lacked a detectable Flag epitope.

4.2.1.3 Excision of the Id1-IRES-Pac-pA *cassette by Cre-mediated recombination*

I generated a "rescued" cell line by transfecting one Id1-overexpressing clone (clone 7) with a circular plasmid containing a *CAG* promoter driving expression of a *Cre* recombinase gene. I plated the cells at clonal density, then screened the plate by fluorescence microscopy to pick and expand clones expressing GFP in place of Id1 (for more information refer to section 2.2.4.1.3 of the Materials and Methods chapter). I verified loss of *Id1* expression and overexpression of *GFP* by qRT-PCR (Figure 4.4a, data not shown). These clones will be referred to henceforth as "Id1 reverted". These lines serve as a control to ensure that any Id1 overexpression phenotype can be directly attributed to the *Id1* transgene and not to a secondary genetic change.

4.2.1.4 Functional validation of Id1 overexpressing cell lines

In order to verify the functional competence of the *Id1* transgene within these cell lines, I set up adherent monolayer neural differentiation experiments in N2B27 culture medium to verify the capability of Id1 to inhibit this process, as described in Ying *et al.* (2003a). I made use of the empty vector control clone, Id1 overexpressing clones 7 and 11, and one Id1 reverted clone.

I performed qRT-PCR on RNA samples collected daily from ES cells cultured in N2B27 for 5 days, and on the starting population sample of ES cells cultured in

LIF+FCS (Figure 4.4a, data not shown for clone 11 since it behaved as clone 7). The Id1 overexpressing clones expressed higher levels of *Id1* transcript than the control clones, and displayed severely dampened upregulation of the neural marker *Sox1*, which was highly upregulated in the empty vector control and the Id1 reverted cells. The amount of *Id1* transcript was reduced (approximately 4-fold) after just 24 hours in serum-free N2B27 medium. This transgene downregulation was common to the two tested clones and a third Id1 overexpressing clonal ES cell line generated using a different plasmid (plasmid structure: *CAG-Id1-IRES-Pac*, data not shown) and is therefore unlikely to be a clone- or plasmid-specific phenotype.

I then repeated the experiment and fixed the samples at day 5 of differentiation. I stained the cells with an antibody against the neuronal marker Tubb3 and imaged them to detect expression of Tubb3 and of GFP (Figure 4.4b, data not shown for clone 11 since it behaved as clone 7). The empty vector and Id1 reverted control cell lines displayed widespread expression of Tubb3, which was not detectable in Id1 overexpressing cells. The Id1 reverted control was uniformly GFP-positive, indicating successful excision of the *Id1* transgene.

Id1 overexpression therefore inhibits neural differentiation, as expected, confirming that the *Id1* transgene is functional in these cell lines.

4.2.2 Id1 overexpression does not prevent the loss of naïve pluripotency markers

Having illustrated the functional capability of the cell lines to reproduce the previously reported inhibition of neural differentiation, I set out to characterise the kinetics of the differentiation process.

I set up an adherent neural monolayer differentiation experiment by plating empty vector control or Id1 overexpressing ES cells cultured in LIF+FCS in N2B27 for 3 days. I extracted RNA from the starting population and at daily timepoints during differentiation, and performed qRT-PCR analysis of the samples for *Id1* and markers

of naïve pluripotency (Figure 4.5). I confirmed *Id1* was overexpressed in the test samples compared to control, and recapitulated the downregulation of *Id1* transcript after just 24 hours in serum-free N2B27 medium observed in Figure 4.4a (Figure 4.5a). *Esrrb*, *Klf4*, *Rex1* and *Nanog* were all downregulated to basal levels of expression within 24 hours of cell culture in N2B27, with comparable kinetics between control and Id1 overexpressing cells (Figure 4.5b). I stained empty vector control and Id1 overexpressing ES cells cultured in LIF+BMP4 or in N2B27 for 2 days with antibodies against Esrrb and Klf4. The expression of both markers was lost after 48 hours of culture in N2B27 (Figure 4.6a). Their downregulation was confirmed by quantification of the nuclear immunostaining signal intensity of a minimum of 130 nuclei for each of the four samples (Figure 4.6b).

I have confirmed the findings presented in this section using an independently-generated line of ES cells that I engineered to allow doxycycline-inducible expression of an *Id1* transgene. These data are not presented in this thesis, but can be found in our recent publication (Malaguti *et al.* 2013).

We can therefore conclude that Id1 overexpression in neural adherent monolayer differentiation does not affect the downregulation of naïve pluripotency markers on the timescale I have adopted for my experiments.

4.2.3 Id1 overexpression delays the exit from a pluripotent epiblast-like state whilst promoting a proximal posterior epiblast identity

Having established Id1 overexpression does not delay the exit from naïve pluripotency, I set out to investigate whether it affected gene expression during early timepoints of neural differentiation, or whether it played its role in the inhibition of neural specification at a later stage.

I used the RNA samples described in section 4.2.2 to investigate the expression of markers of various lineages during the first three days of neural adherent monolayer

differentiation (Figure 4.7). Whilst Id1 did not affect the kinetics of downregulation of naïve pluripotency markers (Figure 4.5b), it delayed the downregulation of *Oct4* and *Cdh1*, factors expressed in both pre-implantation and post-implantation pluripotent epiblast. It also delayed the downregulation of the post-implantation pluripotent epiblast markers *Fgf5* and *Otx2* (Figure 4.7a). This evidence points towards a delay in the exit from a post-implantation state of pluripotency. This was corroborated by a delay in the upregulation of *Zeb1* and *Zeb2*, transcription factors capable of driving the epithelial-to-mesenchymal transition and forcing cells to exit the pluripotent state, and of *Cdh2*, which encodes a transmembrane protein expressed in cells that have undergone EMT (Figure 4.7c). Interestingly, whilst delaying the loss of markers of pluripotent epiblast, Id1 overexpression also induced the expression of *T*, *Gsc* and *Eomes*, markers of proximal posterior epiblast/primitive streak, and inhibited the upregulation of neural markers *Sox1* and *Zfp521* (Figure 4.7b,c). This suggests a multifaceted role for Id1: on one hand it delays the exit from a post-implantation epiblast-like pluripotent state, on the other it provides the differentiating cells with a regionally specified identity, that of proximal posterior epiblast. In that respect it is interesting to note that *Nanog* is re-upregulated in Id1 overexpressing cells at day 3 of differentiation (Figure 4.5b). *Nanog* is a marker of the naïve pluripotency state, its overexpression can sustain LIF-independent ES cell self-renewal, and it is a key determinant of the epiblast in the inner cell mass of the blastocyst (Chambers *et al.* 2003, Chambers *et al.* 2007, Mitsui *et al.* 2003, Silva *et al.* 2009). However, it is also expressed in the posterior half of the E6.5 epiblast (Morkel *et al.* 2003, Osorno *et al.* 2012). Thus, its downregulation and consequent upregulation in Id1 overexpressing cells would add further weight to the conclusion that Id1 is promoting a pluripotent proximal posterior epiblast cell type.

I assessed the expression of some of these markers during this differentiation process at the protein level by performing immunofluorescence staining of empty vector control and Id1 overexpressing ES cells cultured in LIF+FCS or differentiated for 1, 2 or 3 days in N2B27. Control cells replace Cdh1 with Cdh2 at their cell membranes

starting at day 2 of differentiation. Id1 overexpressing cells, on the other hand, maintain expression of Cdh1 and do not upregulate Cdh2 within the timeframe of this experiment (Figure 4.8a). Control cells gradually downregulate Oct4, with cells expressing low levels of the protein at day 3. They do not visibly upregulate T protein. Id1 overexpressing cells, instead, retain Oct4 expression and upregulate T on the third day of differentiation (Figure 4.8b). I performed immunostaining quantification of the nuclear fluorescence signal of empty vector control and Id1 overexpressing cells cultured in LIF+FCS or in N2B27 for 3 days and stained with antibodies against Oct4 and T. Whilst the factors display a similar expression pattern in LIF+FCS culture, there is a very clear difference in their expression after 3 days of differentiation. Approximately 40% of control cells express Oct4, with 6% of these co-expressing T. Approximately 75% of Id1 overexpressing cells express Oct4, with 45% of these co-expressing T. For both cell lines, less than 2% of cells express T without co-expressing Oct4. Oct4 is downregulated as epiblast cells become mesodermal. This suggests that the T-positive cells that are being generated in this culture system represent the Oct4-positive T-positive cells present in the epiblast rather than an Oct4-negative early mesodermal cell type (Rivea-Pérez & Magnuson 2005).

I have also confirmed the findings presented in this section, for a subset of the markers described, in independently-generated ES cells that allow for doxycycline-inducible expression of an *Id1* transgene (Malaguti *et al.* 2013).

We can therefore conclude that, based on detailed marker analysis at the transcript and at the protein level, Id1 overexpression can inhibit neural differentiation by delaying the exit from a pluripotent post-implantation epiblast-like state and by imposing a proximal posterior epiblast-like regional identity on the cultured cells.

4.2.4 Id1 overexpression ultimately favours mesodermal differentiation

4.2.4.1 Id1 overexpression cannot prevent the exit from a pluripotent epiblast-like state

In order to understand into what lineage Id1 overexpressing cells eventually differentiate into, I set up a neural adherent monolayer differentiation of empty vector control and Id1 overexpressing ES cells in N2B27 medium. This experiment is similar to the one described above, but the gene expression of the cells was characterised at later points in time. I collected RNA from the starting populations of ES cells cultured in LIF+FCS and from cells cultured in N2B27 for 3, 4 or 5 days, and performed qRT-PCR analysis for a range of markers of different germ layers (Figure 4.9). I confirmed Id1 was overexpressed in the test samples, and noticed once more that the transgene was downregulated after differentiation (Figure 4.9a). The naïve pluripotency marker *Esrrb* was downregulated in both control and Id1 overexpressing cells after 3 days of differentiation (Figure 4.9b), as expected from the results presented in the previous section. Control cells exhibited very little detectable expression of the pluripotent post-implantation epiblast markers *Oct4*, *Cdh1* and *Fgf5* on the third day of differentiation, whereas the three factors were still readily detectable in Id1 overexpressing cells. These epiblast markers were ultimately downregulated, however, after 4 days of culture in N2B27 (Figure 4.9c). Immunostaining of empty vector control and Id1 overexpressing cells at day 3 and day 5 of N2B27 differentiation confirmed at the protein level that Id1 overexpression can maintain Cdh1 and Oct4 expression until day 3 but not until day 5 of differentiation (Figure 4.10). Id1 overexpressing cells upregulated proximal posterior epiblast markers *T*, *Eomes* and *Gsc* on day 3 of differentiation, and these were also downregulated subsequently (Figure 4.9d). Immunostaining of empty vector control and Id1 overexpressing cells at day 3 and day 5 of N2B27 differentiation confirmed at the protein level that Id1 overexpression can maintain Cdh1 and T expression until day 3 but not until day 5 of differentiation (Figure 4.11). The co-expression of Cdh1 and T at day 3 of differentiation confirms the previous observation that the T-

expressing cells generated by the Id1 overexpressing cells represent the Cdh1-positive T-positive cells expressed in the epiblast rather than an early mesodermal cell type.

These observations suggest that Id1 overexpression can delay, but not prevent, the exit from a pluripotent post-implantation epiblast-like state.

4.2.4.2 Id1 overexpression in neural differentiation generates mesodermal cells

I then went on to test the expression of markers of differentiated lineages. The general EMT markers *Cdh2*, *Vimentin* and *Zeb1* are expressed by cells differentiating into neural ectoderm or mesoderm. Interestingly the upregulation of these markers was delayed in Id1 overexpressing cells (Figure 4.9e). As expected, Id1 overexpression also inhibits the expression of the neural marker *Sox1*, and it slightly inhibits the expression of the neural/mesenchymal marker *Zfp521* (Figure 4.9f). It does not affect the expression of the endodermal marker *Gata6* (Figure 4.9g). Surface ectoderm is a subset of ectoderm that differentiates without losing an epithelial character and Cdh1 expression, so it is a viable hypothesis that the effect of Id1 overexpression on the inhibition of the upregulation of EMT markers may result in differentiation towards this particular lineage. This does not appear to be the case, however, as there is no significant difference between control and Id1 overexpressing cells in the expression of *Krt8* and *Gata2*, two markers of this lineage (Figure 4.9h). Finally, I tested the expression of a range of mesodermal markers. Id1 overexpression appears to drive the expression of the late primitive streak markers *Wnt3a* and *Snail* at day 4 of differentiation, 24 hours after the peak in expression of the primitive streak determinants *T* and *Eomes*. It also appears to induce the upregulation of the paraxial mesoderm markers *Pdgfra* and *Tcf15* (also commonly referred to as *Paraxis*) at days 4 and 5 of differentiation, and of the vascular, haematopoietic, lateral and extraembryonic mesoderm marker *Kdr* (also commonly referred to as *Flk1* and *Vegfr2*) at day 5 of differentiation (Figure 4.9i). Thus, Id1

overexpression appears to favour differentiation of ES cells cultured in serum-free unsupplemented N2B27 medium into cells of mesodermal lineages.

Cdh2 replaces Cdh1 at the membrane of cells undergoing EMT as they ingress through the primitive streak to form mesoderm, and is expressed by a large number of mesodermal subtypes thereafter (Hirano *et al.* 2008, Nakamura *et al.* 2013, Radice *et al.* 1997). It is therefore surprising that, despite upregulating a range of mesodermal markers, the Id1 overexpressing cells do not appear to upregulate *Cdh2* transcript to the same levels as the control cells expressing the factor in neural ectoderm. I examined the expression of Cdh2 at the protein level by performing immunostains on empty vector control and Id1 overexpressing cells at day 3 and day 5 of culture in N2B27 medium (Figure 4.12, also see Figure 4.8a for expression at day 3). Control cells express Cdh2 in a high proportion of cells at day 3 and are uniformly expressing the factor at day 5. Id1 overexpressing cells express Cdh2 in a few cells at day 3 and in a high proportion of cells at day 5. It appears that the cells at day 5 of differentiation are divided in two subsets, a first one with densely packed Cdh2-expressing cells and a second one with more loosely separated cells that appear to be stemming from the packed colony and downregulating Cdh2. This heterogeneity could explain the lower level of *Cdh2* transcript in Id1 overexpressing cells than in control cells at day 5 of differentiation.

I have also confirmed the findings presented in this section in independent experiments, for a subset of the markers described, using the Id1 reverted ES cells as a negative control (Malaguti *et al.* 2013).

In conclusion, Id1 overexpression cannot prevent the exit from a pluripotent state, but it reroutes cells destined for neural commitment to mesodermal specification.

4.2.5 Id1 expression in single wild-type differentiating cells is not responsible for the timing of the loss of pluripotency factors

ES cells differentiate asynchronously (Lowell *et al.* 2006). Having observed that Id1 overexpression is capable of delaying the loss of post-implantation pluripotency markers, I hypothesised that the heterogeneity in the rate of neural differentiation may be due to heterogeneity in Id1 expression during differentiation. I set out to test this by assaying whether there is a correlation between Id1 and pluripotency or neural marker expression at the single cell level.

I therefore decided to perform qRT-PCR on RNA extracted from single cells at various timepoints of N2B27 differentiation, and to compare the pattern of expression of *Id1* to that of a range of pluripotency and neural markers in the same single-cell samples. The expression pattern of the selected markers during the first four days of neural adherent monolayer differentiation at a whole population level is shown in Figure 4.13. These data confirm that the pluripotent epiblast markers *Cdh1*, *Fgf5*, *Nanog* and *Oct4* are progressively lost during the course of the differentiation experiment, whereas the neural/EMT markers *Cdh2*, *Sox1*, *Zeb1* and *Zeb2* are progressively upregulated.

I set up a first experiment with the aim of understanding how reliable this novel technique truly was, whilst at the same time obtaining some potentially valuable insight in the biological process I was studying. I made use of a fluorescence activated cell sorter to sort 90 E14tg2a ES cells cultured in LIF+FCS and 90 E14tg2a ES cells cultured in N2B27 for 4 days into single wells of a 96-well PCR plate. These cell numbers were chosen so that the maximum number of samples (180) could be loaded in duplicate on a 384-well qRT-PCR plate alongside positive and negative controls and a serial dilution of cDNA generated by sorting 32 cells into a well of a 96-well plate alongside the samples (see Materials and Methods paragraph 2.2.2.3). I performed reverse transcription of the RNA into cDNA, and 22 cycles of pre-amplification of the cDNA using primers to detect *Id1*, the pluripotency markers *Cdh1* and *Oct4*, the neural markers *Cdh2* and *Sox1*, and the housekeeping gene *Tbp*. I

performed qRT-PCR experiments on the samples to measure the expression level of the six factors listed above (Figure 4.14a). As expected, the majority of ES cells expressed *Cdh1* and *Oct4* at high levels and *Cdh2* and *Sox1* at basal levels, and the opposite was true for the differentiated cells. Two cells in the differentiated sample displayed an ES cell-like expression pattern, suggesting that they had yet to undergo neural commitment. *Id1* was expressed heterogeneously in ES cells and in a similar pattern, yet to a lower level, in differentiated cells. *Sox1* expression was very variable at day 4 of differentiation, in keeping with the flow cytometry profiles of differentiating Sox1-GFP reporter ES cells (Ying *et al.* 2003b). Expression of *Cdh2* expression was also heterogeneous in day 4 cultures.

In order to understand how the expression of one factor may reflect the expression status of the others, I calculated Spearman's rank correlation coefficients for the factors across all cells (Figure 4.14b). The coefficient is equal to +1 for a perfect positive correlation and to -1 for a perfect negative correlation. A value of 0 indicates a complete absence of correlation. *Oct4* and *Cdh1* displayed a statistically significant positive correlation with a coefficient of 0.82, as expected, given that both markers are expressed in ES cells and absent from neural cells. *Cdh2* and *Sox1* also displayed a statistically significant positive correlation with a coefficient of 0.5, as expected, given that both markers are expressed in neural cells and absent from ES cells. The ES cell markers were negatively correlated with the neural markers.

The combination of these observations confirms that the differentiation process is heterogeneous, as previously reported (Lowell *et al.* 2006, Ying *et al.* 2003b). The observation of the expected correlations between neural and pluripotency markers validates the reliability of single-cell qRT-PCR as an analytical technique, and suggests that it can detect meaningful correlations within heterogeneous populations of cells.

The analysis of the correlation of *Id1* with the other markers allows us to address the question of whether *Id1* expression is associated with cells that tend to resist differentiation. *Id1* displayed a weak positive correlation with *Cdh1* and *Oct4* and a

weak negative correlation with *Sox1*, but the rank correlation coefficients were low, and therefore unlikely to indicate true biologically meaningful correlations. It would appear that at the selected timepoints the shift of gene expression from a pluripotent signature to a neural signature has not yet started (ES cells) or has already been completed (day 4 of N2B27 culture). This in turn implies that the levels of pluripotency factors are low in the vast majority of differentiated cells, and thus that comparison of the levels of these with the levels of neural factors is not effectively giving an insightful indication into how the transition in expression from one to the other may operate.

I therefore set up a second experiment to attempt to capture a better description of the transition from pluripotency to the neural lineage. In order to maximise the chances of identifying timepoints providing high information, but still have a sufficient number of cells to perform statistically relevant analyses, I chose to analyse E14tg2a ES cells cultured in N2B27 for 1, 2 or 3 days. I sorted 60 cells from each timepoint into single wells of 96-well PCR plates and performed RNA retrotranscription and 22 cycles of cDNA pre-amplification with primers designed to amplify *Id1*, the pluripotency markers *Cdh1*, *Fgf5*, *Nanog*, *Oct4*, the neural and EMT markers *Cdh2*, *Sox1*, *Zeb1*, *Zeb2* and the normalising gene *Tbp*. I then performed qRT-PCR experiments to measure the expression levels of these factors across all samples and calculated Spearman's rank correlation coefficients for the factors separately for the three timepoints, in order to understand how the expression of the factors is correlated at specific times during the differentiation experiment (Figures 4.15, 4.16). The markers display heterogeneous expression both between different timepoints and within the same timepoint. As expected from the population-level qRT-PCR analysis presented (Figure 4.13), the levels of *Cdh1* and *Oct4* are downregulated over the course of the experiment, whereas the neural markers *Cdh2*, *Sox1*, *Zeb1* and *Zeb2* are upregulated as the differentiation progresses. The correlation analysis reveals that the markers behave as two separate sets of transcripts: a "pluripotent epiblast" set (*Cdh1*, *Fgf5*, *Nanog*, *Oct4*) and a "neural" set (*Cdh2*, *Sox1*, *Zeb1*, *Zeb2*). The markers within

the two sets of transcripts display weak statistically significant correlations at day 1 of differentiation. At days 2 and 3 of the differentiation experiment, the pluripotent epiblast markers are strongly correlated with one another, the neural markers are strongly correlated with one another, and the two sets are negatively correlated between each other.

These observations confirm that cells differentiate heterogeneously. The observation of the expected correlations between pluripotency markers and between neural markers once again confirms the validity of single-cell qRT-PCR as a tool for expression analysis. Single-cell qRT-PCR can therefore detect the heterogeneity in the differentiation process between individual cells, and can successfully identify correlations between the expression of different genes within heterogeneous cultures.

I proceeded to analyse the expression pattern of *Id1* to understand whether expression of the HLH factor is associated with a resistance to differentiation. *Id1* is expressed at extremely low levels in all cells on days 1 and 2 of differentiation, and is upregulated on day 3 by all cells. This suggests that it is unlikely to influence the rate of neural differentiation and suggests my initial hypothesis was wrong.

A few more interesting insights can be gained by the analysis of these data.

The first is the observation of the statistically significant positive association of *Cdh1* and *Oct4* at all timepoints of the differentiation process. Cdh1 and Oct4 are co-expressed from pre-implantation development until cells exit pluripotency in the post-implantation embryo (Fierro-González *et al.* 2013, Malaguti *et al.* 2013, Nichols *et al.* 1998, Palmieri *et al.* 1994), and the transcriptional association I observed may underlie a common mechanism promoting the expression of these factors.

The second is the observation that the pattern of upregulation of the neural transcription factors differs. *Sox1* is not expressed on day 1, is upregulated by few cells at low levels on day 2, and is then expressed at a higher level by a higher number of cells on day 3. *Zeb1* is expressed by many cells at a low level on day 1, and is then expressed at a higher level on days 2 and 3. *Zeb2* is expressed in some

cells at a high level on day 1, and the proportion of expressing cells increases on days 2 and 3. The three neural transcription factors therefore differ in the timing of their activation, the proportion of expressing cells and the level of expression upon activation.

We can conclude that ES cells differentiating in N2B27 in an adherent monolayer switch from a pluripotent epiblast gene expression signature to an early neural gene expression signature, and are therefore following a differentiation route that is compatible with the normal development of the mouse embryo. This suggests that neural adherent monolayer differentiation is a valid method for the study of neural specification and of the early lineage decisions in which Id1 is involved.

We can also conclude that in this differentiation system, *Id1* does not appear to be expressed at a level which would allow it to influence the differentiation direction or kinetics of the cells. The initial hypothesis of the HLH factor influencing the heterogeneity of the neural differentiation process is therefore unlikely to be correct.

Finally, single-cell qRT-PCR analysis revealed considerable variability between individual differentiating cells. This variability is not a simple artefact of the procedure (such as a lack of precision in the quantification of mRNA levels), since pluripotency and neural markers displayed the expected positive and negative correlations with each other.

4.3 Discussion

4.3.1 Id1 overexpression and the exit from naïve pluripotency

Id1 is a transcriptional regulator with a critical function in the maintenance of the self-renewal status of ES cell cultures. Whilst its knockout does not have a phenotype *in vivo*, and does not result in overt differentiation of ES cells, it does affect the clonogenic properties of ES cells and their expression of *Nanog* and *Rex1*

(Lyden *et al.* 1999, Romero-Lanman *et al.* 2011, Yan *et al.* 1997). The fact that this is the case despite ES cells expressing the structurally and functionally similar HLH factors Id2 and Id3 is testament to the importance that Id1 plays in cell fate decisions.

In a 2010 study, Jing and colleagues investigate the effects of BMP4 in two distinct stages of ES cell differentiation: the transition from a pre-implantation to a post-implantation type of pluripotency (referred to as ES cell to EpiSC transition), and the transition from post-implantation pluripotency to neural progenitor cells (Zhang *et al.* 2010). They adopted a serum-free embryoid body (SFEB) differentiation protocol (Watanabe *et al.* 2005) in which ES cells are allowed to aggregate in suspension in basal medium containing KSR and no LIF. They demonstrated that after two days of culture in this medium, cells treated with BMP4 expressed higher levels of *Klf4*, *Nanog* and *Rex1* and lower levels of *Fgf5* than untreated controls. They set up a similar experiment using Id2 overexpressing cells in place of BMP4-treated cells and showed that the Id2 overexpressing cells expressed higher levels of *Klf4*, similar levels of *Nanog* and *Rex1*, and lower levels of *Fgf5* to control cells. They also claimed that they obtained similar results with Id1 overexpressing ES cells, although they did not present these results. They therefore concluded that BMP4 and its targets Id1 and Id2 inhibit the ES cell to EpiSC transition.

In paragraph 4.2.2, I presented data which indicate that Id1 overexpression does not seem to delay the loss of *Klf4*, *Esrrb* and *Nanog* in the differentiation of ES cells towards the neural lineage. I have also confirmed these findings using a cell line that allows for doxycycline-inducible expression of an *Id1* transgene (Malaguti *et al.* 2013). How can these apparently contradicting datasets be reconciled?

The explanation may lie in timing differences between the two differentiation protocols used. The transition from pre-implantation to post-implantation pluripotency in the SFEB protocol takes between 48 and 72 hours, whereas the same transition in the N2B27 protocol occurs in approximately 24 hours. The delay in downregulation of *Klf4* and upregulation of *Fgf5* after 48 hours of SFEB differentiation reported by Zhang *et al.* (2010) for Id1 overexpressing cells may

therefore be difficult to capture in Id1 overexpressing cells cultured for 24 hours in N2B27. The use of shorter timepoints in N2B27 culture may have allowed to make similar observations. The differences in expression of *Klf4*, *Nanog* and *Rex1* between differentiated Id2-overexpressing and wild-type cells observed by Jing and colleagues were small in magnitude. Should their data for Id1-overexpressing cells reflect the expression levels of *Klf4*, *Nanog* and *Rex1* in the Id2-overexpressing cells, it is also possible that these gene expression differences may be too subtle to detect in a short time window in N2B27 culture.

4.3.2 The effects of Id1 overexpression on the loss of epiblast markers and the regional identity of differentiating cells

Despite not affecting the kinetics of the loss of expression of markers of naïve pluripotency, Id1 overexpression has a clear effect on the kinetics of the downregulation of makers of the early post-implantation pluripotent epiblast. Id1 overexpressing cells delayed the downregulation of *Cdh1*, *Fgf5*, *Oct4* and the upregulation of Cdh2 until day 4 of differentiation, unlike control cells, which were mostly negative for Oct4 protein at day 3 and started expressing Cdh2 at day 2 of differentiation. Concomitantly with this maintenance of Cdh1 and Oct4 expression, Id1 overexpressing cells displayed an inhibited upregulation of neural markers and in their place expressed primitive streak determinants such as T, within the Cdh1- and Oct4-expressing compartment. This suggests that Id1 overexpression is inducing the formation of a cell type with the gene expression pattern of proximal posterior epiblast.

4.3.2.1 Generation of a proximal posterior epiblast-like cell type

It is not surprising that Id1 overexpression results in the formation of cells with a proximal posterior epiblast regional identity. A proportion of *Bmp4* knockout mice

fail to express *T*, and BMP4 stimulation has been used to generate T-expressing cells and to promote mesoderm differentiation of ES cells (Johansson & Wiles 1995, Park *et al.* 2004, Winnier *et al.* 1995). *Bmp4* is expressed in the extraembryonic ectoderm adjacent to the proximal epiblast in E5.5, E6.5 and E7.5 embryos, with embryos at E7.5 displaying stronger expression of the morphogen in the posterior regions of the extraembryonic ectoderm adjacent to the proximal posterior epiblast expressing *T* and *Gsc* (Beddington *et al.* 1992, Coucouvanis & Martin 1999, Norris *et al.* 2002, Perea-Gómez *et al.* 1999, Perea-Gomez *et al.* 2001, Rivea-Pérez & Magnuson 2005, Tremblay *et al.* 2001, Wilkinson *et al.* 1990, Willison 1990). *Id1* is also expressed in the proximal regions of the embryo between E5.5 and E7.5 (Jen *et al.* 1997). Although *Id1* knockout embryos do not exhibit defects in the formation of the primitive streak (Fraidenraich *et al.* 2004, Lyden *et al.* 1999, Yan *et a.* 1997), it is conceivable one or more Id factors may be involved in instructing cells in the post-implantation epiblast about their location in the embryo and the structures they should be generating. Both *Id2* and *Id3* are expressed in the proximal posterior epiblast, in agreement with this hypothesis (Jen *et al.* 1997, Lin *et al.* 2013, unpublished observations by Amy Pegg, a PhD student in our laboratory).

4.3.2.2 Identity of the remaining cells in culture

The observation that approximately 45% of Oct4-positive Id1 overexpressing cells express T on the third day of differentiation in N2B27 raises an interesting question. Are the remaining 55% of Oct4-expressing cells yet to acquire a proximal posterior identity, are they a set of proximal posterior epiblast-like cells which are not expressing T, are they the equivalent of lateral and anterior proximal epiblast or of epiblast with more distal character? The question is an open one. The cells are not expressing high levels of *Sox1*, which is expressed in the anterior epiblast. Expression of Sox1 in N2B27 culture occurs after downregulation of Cdh1 (Kamiya *et al.* 2009, Malaguti *et al.* 2013, Watanabe *et al.* 2005), and the Id1 overexpressing cells appear to be uniformly Cdh1-positive at day 3 of differentiation. A qRT-PCR

for *Dlx5*, a gene expressed rostrally to *Sox1* in E6.5 embryos and in the proximal anterior of E7.5 embryos, did not detect any transcript in either control or Id1 overexpressing cells (data not shown). Markers of surface ectoderm, which has a proximal anterior component as E7.5, were not enriched in Id1 overexpressing cells. In the gastrulating embryo, T becomes restricted to cells at the posterior-most region of the embryo, rather than being widely expressed throughout the posterior half of the epiblast (Figures 5.2, 5.4), thus Oct4-positive T-negative cells exist in the proximal posterior epiblast.

We can therefore infer that, if the Oct4-positive T-negative cells have an *in vivo* counterpart, it is either lateral epiblast, medial/distal posterior epiblast, proximal posterior epiblast not expressing T, or epiblast of an earlier character.

4.3.2.3 Delay in the loss of post-implantation pluripotency factors

Another matter of interest is the significance of the Id1-induced delay in the downregulation of markers of the pluripotent post-implantation epiblast. Is this action of Id1 independent of its role in regionalising the culture or is the expression of *Cdh1*, *Fgf5* and *Oct4* a secondary property of the cell type that is being generated (i.e. does posterior epiblast inherently differentiate slower than anterior epiblast)? The former possibility is certainly an intriguing one. Embryos lacking the BMP4 receptor *Bmpr1a* exhibit premature loss of *Fgf5* and *Oct4* and premature differentiation into the neural lineage (Di-Gregorio *et al.* 2007). Furthermore, we have seen in Figure 4.1 that BMP4 treatment of cells cultured in N2B27 prevents the loss of Cdh1 expression and maintains colonies of Oct4-positive cells for at least 6 days of differentiation, suggesting the factor has a strong anti-EMT effect. Id1 overexpression seems to recapitulate this prolonged maintenance of Cdh1 until day 3, after which the cells lose the expression of Cdh1 and upregulate Cdh2. The *Id1* transgene is downregulated during the course of the differentiation experiments, and reduced levels of Id1 protein may be unable to sustain Cdh1 expression past day 3.

The second possibility is also very plausible. The expression of *Oct4* and *Cdh1* persists for a longer time in the posterior epiblast than in the anterior epiblast, with expression of both factors stronger in the posterior of headfold stage embryos before complete loss of *Oct4* transcript (Perea-Gómez *et al.* 1999, Stemmler *et al.* 2005). It is thus possible that the cells generated by overexpression of Id1 express Cdh1 and Oct4 because of the intrinsic properties of posterior proximal epiblast cells rather than because of an anti-differentiation action of Id1.

As with the proverbial case of the chicken and the egg, it is not fundamental for the purposes of this investigation to distinguish between these two possibilities by asking whether the ability of Id1 to delay differentiation is secondary to its effects on positional identity, or vice versa. Whether Id1 has an intrinsic anti-differentiation effect or whether its primary effect is to generate a posterior cell type that expresses pluripotency factors because of the structure of its gene expression network, the end result is that overexpression of the HLH factor delays the kinetics of the exit from a post-implantation epiblast-like pluripotency transcriptional state.

4.3.3 Mesodermal differentiation of Id1 overexpressing cells

4.3.3.1 Id1 overexpression cannot prevent EMT

Whilst BMP4 treatment can generate T-expressing cells in various differentiation systems (Johansson & Wiles 1995, Park *et al.* 2004), and favours mesendodermal gene expression in the SFEB culture experiments carried out by Zhang *et al.* (2010), it appears that in N2B27 adherent monolayer culture it results in a failure to downregulate Cdh1 and in the maintenance of Oct4-expressing colonies for at least 6 days of differentiation. Despite its delay on the downregulation of Cdh1 and Oct4, Id1 overexpression in N2B27 culture cannot fully replicate the effects of BMP4, and eventually results in an EMT step, the expression of Cdh2, and differentiation towards the mesoderm lineage.

Why this may be is unclear. Either other targets of BMP4 are critical for the maintenance of Cdh1 and Oct4 expression, or the downregulation of the *Id1* transgene during the course of the differentiation results in insufficient levels of Id1 protein to effect this block of EMT. It is interesting to note that the Id1 overexpression experiments in N2B27 carried out by Ying *et al.* (2003a) resulted in the generation of cells with two different morphologies. Constitutive Id1 overexpression resulted in the generation of the "big flat cells" which can be obtained after culture of wild-type cells in N2B27+BMP4 (Figure 4.1). Episomal transfection of supertransfectable ES cells (cells with a genetic modification which allows replication of circular plasmids) generated cells with a mesenchymal morphology after 6 days of differentiation (Ying *et al.* 2003a). Episomal overexpression is not stably maintained after differentiation (unpublished observations in the Lowell laboratory), suggesting that a difference in the maintenance of Id1 overexpression may have resulted in the generation of this mesenchymal cell type. This, coupled to the fact that large Cdh1-positive epithelial cells could be generated in the constitutive stable overexpression experiment performed by Ying *et al.* (2003a), suggests that an insufficient maintenance of Id1 expression may be the cause of the lack of long-term EMT inhibition in my own experiments.

If this were the case, the situation would be similar to that described by Kobayashi & Kageyama (2010) for the bHLH factor Hes1. Hes1 overexpression in N2B27 culture results in expression of *T*, but in order for cells to express markers of the mesodermal lineage (*Kdr*, *Nkx2.5*) Hes1 levels need to be reduced. If they are not, the cells acquire a morphology reminiscent of that seen after N2B27+BMP4 culture of wild-type ES cells (Kobayashi & Kageyama 2010, Zhou *et al.* 2013, Figure 4.17).

Notably, both Id1 and Hes1 act in other contexts as generic inhibitors of bHLH transcription factor activity, raising the possibility that their functions may converge upon the regulation of an as yet unidentified bHLH transcription factor in the post-implantation epiblast.

4.3.3.2 Id1 overexpression results in differentiation towards mesoderm

The Id1 overexpressing cells follow the normal sequence of marker expression seen during the specification of mesoderm *in vivo* (*T*, *Eomes*, *Gsc* at day 3; *Wnt3a*, *Snail*, *Pdgfra*, *Tcf15* at day 4, *Kdr* at day 5). A previous overexpression experiment carried out with Id2 overexpressing EpiSCs by Zhang *et al.* (2010) also showed that the differentiating Id2 overexpressing cells upregulated *T* and *Kdr* and had reduced upregulation of *Sox1* and *Sox2* compared to control EpiSCs, and that unlike BMP4 treated wild-type cells, they did not upregulate markers of endoderm and non-neural ectoderm upon differentiation.

In conclusion, Id1 overexpression in N2B27 favours the generation of mesodermal cells without forcing cells to differentiate into other non-neural lineages.

4.3.4 Single-cell transcriptional analysis of neural adherent monolayer differentiation

4.3.4.1 Id1 is not responsible for the heterogeneity in the capacity for neural differentiation of ES cells

The analysis of the levels of pluripotency and neural markers in single differentiating wild-type ES cells revealed that *Id1* was not present at detectable levels in any cell from a sample of approximately 100 cells at days 1 and 2 of neural differentiation. The cells differentiated asynchronously, with some activating a neural transcriptional programme as early as days 1 and 2 of differentiation, and others still expressing pluripotency markers after 4 days of differentiation. Whilst I hypothesised that this effect could have been due to the capability of Id1 of delaying the downregulation of Cdh1 and Oct4, as seen in the overexpression experiments, the fact that none of the tested cells expressed *Id1* at the timepoints in which these factors were lost suggests that Id1 is unlikely to play a major role in the differentiation of wild-type cells in

N2B27. This does not exclude the possibility that Id1 may be responsible for delaying differentiation in a rare subpopulation of cells, but we can conclude that Id1 is not likely to be a major factor in explaining the mechanisms behind heterogeneity in neural differentiation in monolayer cultures.

4.3.4.2 N2B27 differentiation follows an appropriate developmental pathway

The transcriptional analysis revealed that single differentiating ES cells exist in either of three stable states, a *Cdh1*- and *Oct4*-expressing ES cell, a *Cdh1*-, *Fgf5*- and *Oct4*-expressing post-implantation epiblast-like cell, or a *Cdh2*-, *Sox1*-, *Zeb1*- and *Zeb2*-expressing early neural cell. This in turn suggests that N2B27 neural adherent monolayer differentiation follows a similar pathway to what seen *in vivo* in embryonic development, and thus represents a valid model for studying early decisions in lineage segregation.

It is remarkable that differentiating cells succeed in stably adopting distinct coherent identities in culture, despite the lack of extrinsic segregation of positional clues. This may be due to the stability of alternate transcriptional networks, which would have evolved *in vivo* in response to signals from a three-dimensional environment patterned by signals from embryonic and extraembryonic tissues.

4.3.4.3 Zeb factors may be early neural inducers

In the single-cell analysis, I identified a number of cells at day 1 and 2 of differentiation which expressed *Zeb1* and *Zeb2* at levels similar to those seen on the third day of differentiation. The activation of these factors precedes the expression of *Sox1* and *Cdh2*, and the downregulation of *Cdh1* and *Oct4*, and could indicate that these pro-EMT factors play a crucial role in driving the loss of Cdh1 and generating a Cdh2-positive Sox1-positive neuroepithelium. The finding that Zeb2 is essential for

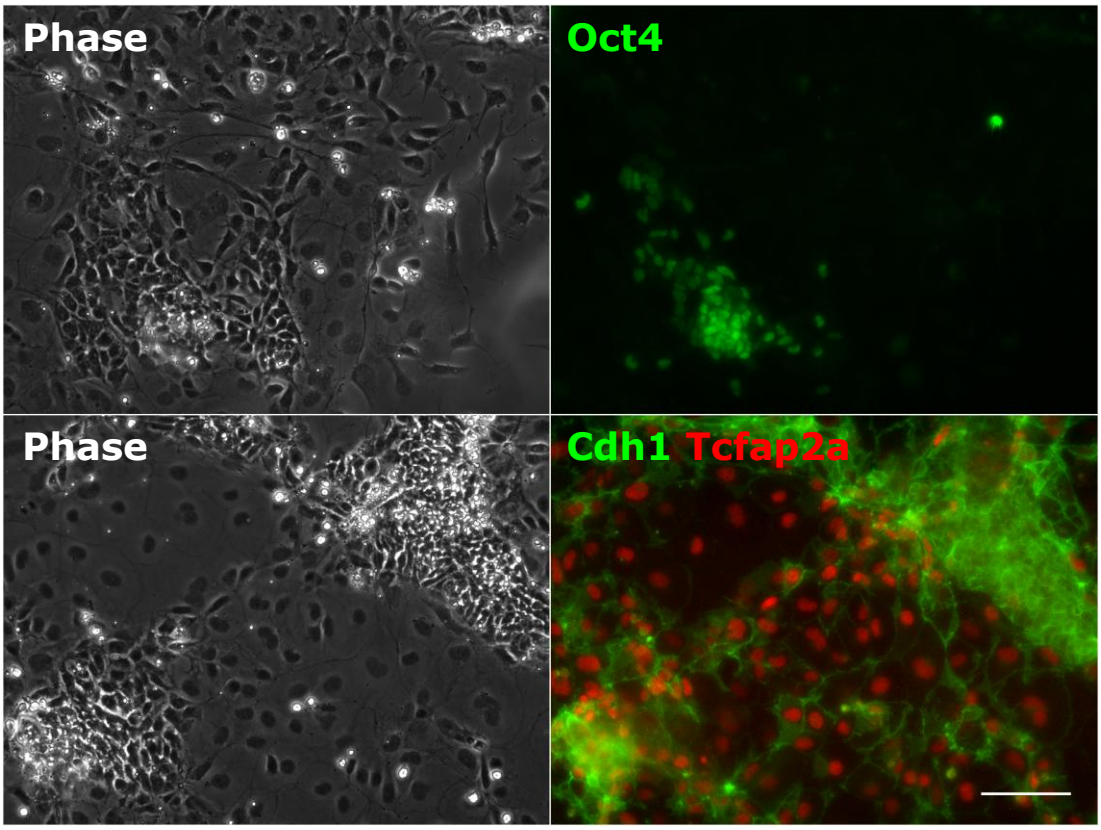
the neural differentiation of human ES cells provides support for this hypothesis (Chng *et al.* 2010).

4.3.4.4 Significance of the correlation between Cdh1 and Oct4 transcription

Throughout this chapter, I have presented data that show that Cdh1 and Oct4 are downregulated synchronously by differentiating ES cells. This was true irrespective of whether cells were differentiating to mesodermal or neural lineages, and of whether the analyses were performed at the population or at the single-cell level (Figures 4.7, 4.8, 4.9, 4.10, 4.13, 4.14, 4.15, 4.16). The two factors are co-expressed in pluripotent cells of the pre-implantation embryo (Fierro-González *et al.* 2013, Nichols *et al.* 1998, Palmieri *et al.* 1994), and they are downregulated synchronously upon the exit from pluripotency after implantation, both in the neural tube and in the posterior regions of the embryo (Malaguti *et al.* 2013, Perea-Gómez *et al.* 1999, Stemmler *et al.* 2005). Cdh1 expression will eventually be uncoupled from Oct4 expression in the surface ectoderm and in the endoderm, but it is intriguing that a transmembrane protein and a transcription factor would exhibit strong correlation in pluripotent and differentiating cells at the single-cell level. It is possible that Oct4 positively regulates the expression of *Cdh1* as well as its own, but in that case loss of *Oct4* mRNA should precede the loss of *Cdh1* mRNA, and this is not apparent. Other uncharacterised factors may therefore be regulating the expression of both genes.
The relationship between the two factors is not limited to a similar regulation of expression. Cdh1 can substitute for Oct4 in the generation of induced pluripotent stem cells (Redmer *et al.* 2011), and the two proteins have also been found to physically interact at the cell membrane (Faunes *et al.* 2013, Livigni *et al.* 2013).

The significance of the coupling of Cdh1 and Oct4 expression throughout development and of the crosstalk between these molecules is currently unclear, but these observations have paved a path for the future investigation of these intriguing phenomena.

FIGURE 4.1



N2B27+BMP4 Day 6

Figure 4.1 – Culture of ES cells in BMP4 generates different cell types

E14tg2a ES cells were cultured in N2B27 in the presence of 10ng/ml BMP4 for 6 days. The cells were then fixed, stained with antibodies against Oct4, Cdh1 or Tcfap2a and imaged on an Olympus IX51 microscope. Scale bar: 100µm.
Clusters of small cells express Oct4 whilst larger flattened cells express Tcfap2a. Both cell types express Cdh1.

FIGURE 4.2

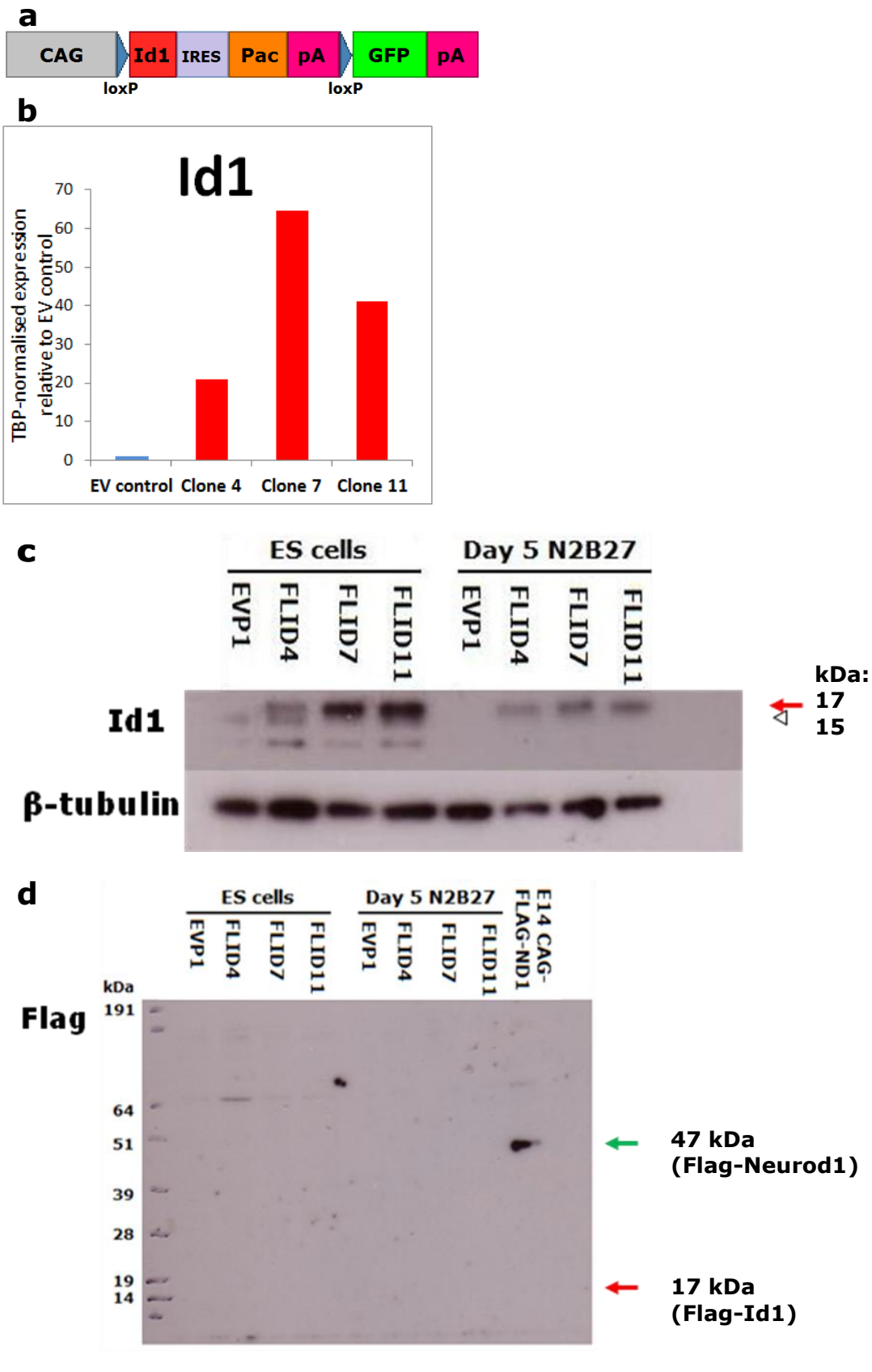

a
CAG
Id1
IRES
Pac
pA
GFP
pA
loxP
loxP
b
Id1
TBP-normalised expression relative to EV control
70
60
50
40
30
20
10
0
EV control
Clone 4
Clone 7
Clone 11
c
ES cells
Day 5 N2B27
EVP1
FLID4
FLID7
FLID11
EVP1
FLID4
FLID7
FLID11
kDa:
17
15
Id1
β-tubulin
d
ES cells
Day 5 N2B27
E14 CAG-FLAG-ND1
EVP1
FLID4
FLID7
FLID11
EVP1
FLID4
FLID7
FLID11
kDa
Flag
191
64
51
39
28
19
14
47 kDa
(Flag-Neurod1)
17 kDa
(Flag-Id1)

Figure 4.2 – Characterisation of *Id1* overexpressing ES cell lines

a. Plasmid transfected into E14tg2a ES cells to generate clonal cell lines overexpressing *Id1-IRES-Pac* under the control of the strong *CAG* promoter. The *Id1-IRES-Pac-pA* sequence is flanked by *loxP* sites, allowing excision of the cassette and expression of a downstream *GFP* transgene upon transfection with a plasmid encoding Cre recombinase. pA: polyadenylation signal.

b. *Id1* mRNA levels in three puromycin-resistant clones and in a control clone generated by transfecting E14tg2a ES cells with a *CAG-Pac-pA* (empty vector) plasmid. All three *Id1*-transfected clones overexpress *Id1*.

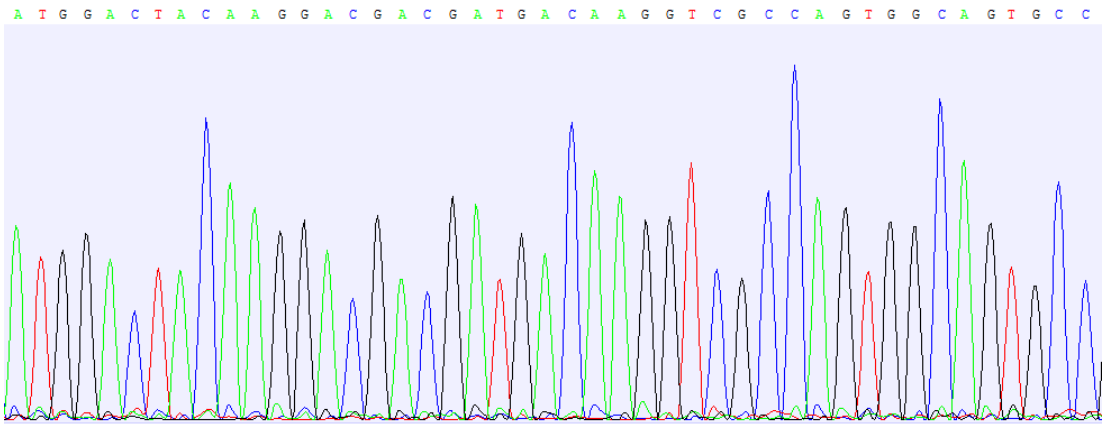
c. Western blots using anti-Id1 (SC-488) and anti-β-tubulin antibodies on cell lysates from the three *Id1* overexpressing (FLID4, FLID7, FLID11) and the control (EVP1) clone, cultured in LIF+FCS and in N2B27 for 5 days. Arrowhead: wild-type Id1 band, red arrow: Flag-tagged exogenous Id1 band. The approximate size of the bands is indicated. β-tubulin was used as a loading control.

d. Western blot using an anti-Flag antibody on the same cell lysates as in (c) and ES cells overexpressing Flag-Neurod1 as a positive control for Flag expression.

Red arrow: size of Flag-Id1 band, green arrow: Flag-Neurod1 band. The approximate size of the bands is indicated. No bands were detected in any of the Id1-overexpressing cells.

FIGURE 4.3

a



b

```
Plasmid      ATGGACTACAAGGACGACGATGACAAGGTCGCCAGTGGCAGTGCC
Id1 cDNA     ATG--------------------AAGGTCGCCAGTGGCAGTGCC
             ***                    **********************
```

c

```
Plasmid      MDYKDDDDKVASGSA
Id1 cDNA     M-------KVASGSA
             *       *******
```

Figure 4.3 – Sequence analysis of Id1 overexpression plasmid

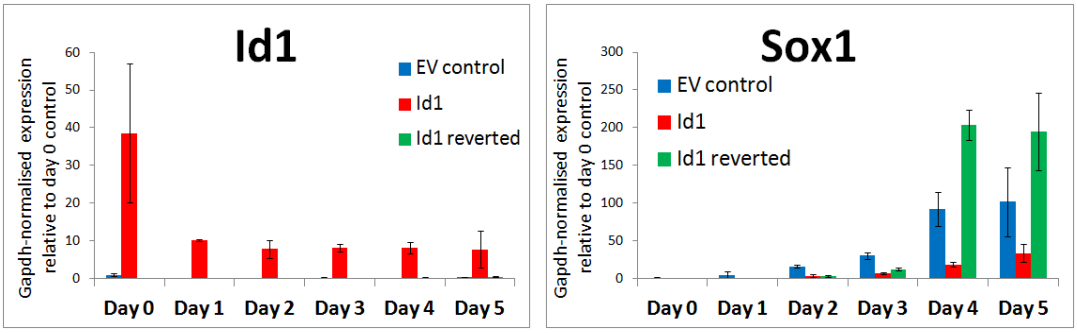
a. Chromatogram from a sequencing read of the start of the *Flag-Id1* open reading frame in the plasmid shown in Figure 4.2a.

b. Alignment of the sequencing read in (a) to the open reading frame of *Id1*.

c. Alignment of the translated sequencing read in (a) to the Id1 amino acid sequence (in one-letter code). The Flag epitope is boxed in yellow.

FIGURE 4.4

a


Id1
Gapdh-normalised expression relative to day 0 control
60
50
40
30
20
10
0
EV control
Id1
Id1 reverted
Day 0
Day 1
Day 2
Day 3
Day 4
Day 5
Sox1
Gapdh-normalised expression relative to day 0 control
300
250
200
150
100
50
0
EV control
Id1
Id1 reverted
Day 0
Day 1
Day 2
Day 3
Day 4
Day 5

b

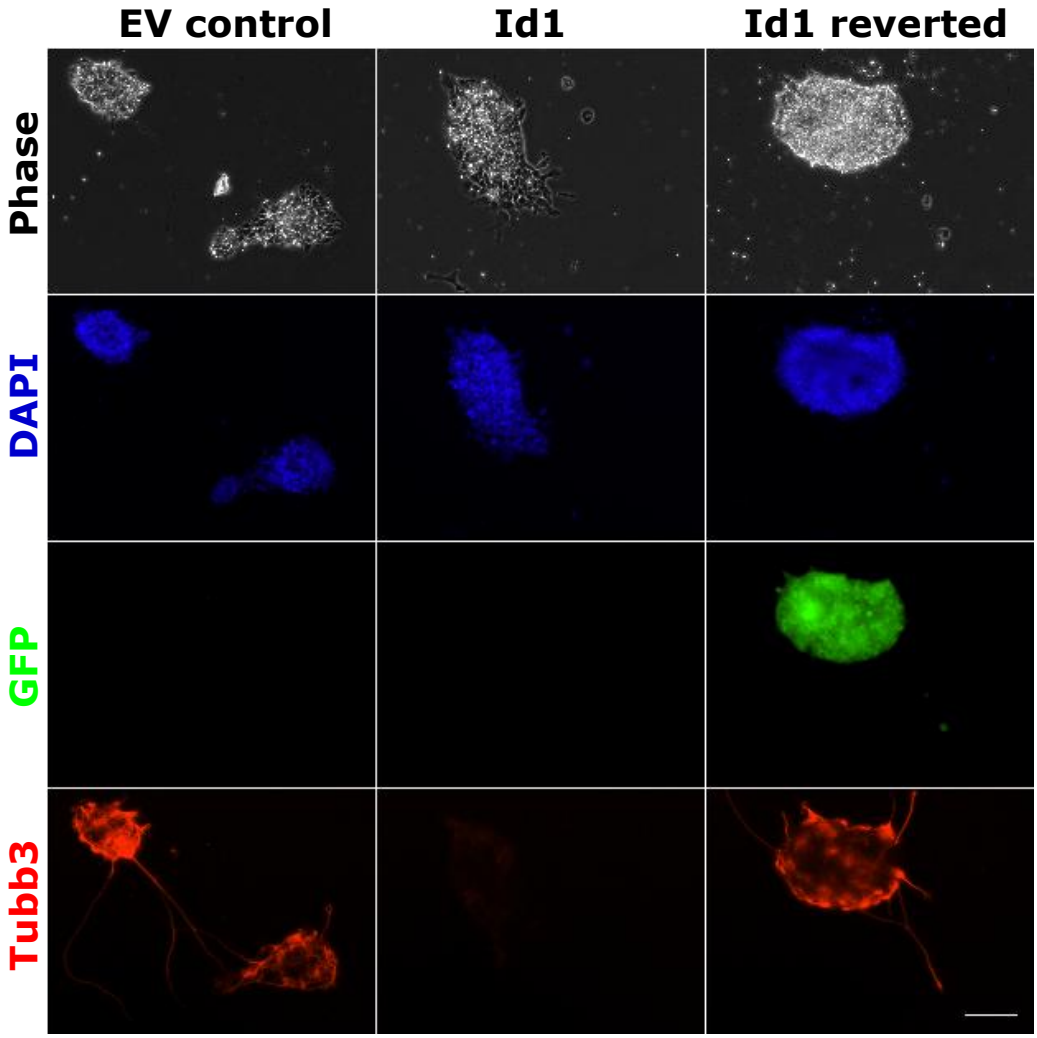

EV control
Id1
Id1 reverted
Phase
DAPI
GFP
Tubb3

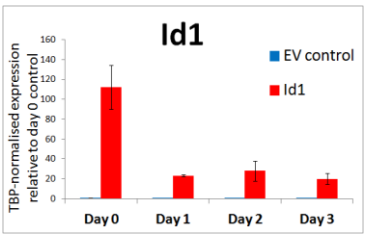
Figure 4.4 – Functional validation of *Id1* overexpressing cells

a. mRNA levels of *Id1* and the neural marker *Sox1* in a neural monolayer differentiation of an empty vector control (blue bars), *Id1* overexpressing cells (red bars) and a control rescued clonal cell line in which the *Id1-IRES-Pac* cassette was excised by Cre recombinase ("Id1 reverted", green bars).
Cells cultured in LIF+FCS (Day 0) were replated in N2B27 and RNA extracted at daily timepoints.
Error bars: standard deviation of two biological replicates.

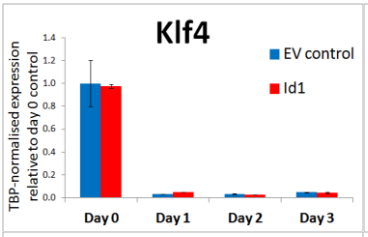
b. Empty vector control, *Id1* overexpressing and rescued cells ("Id1 reverted") after 5 days of culture in N2B27 medium. The cells were stained with an anti-Tubb3 antibody to detect neuronal differentiation, and nuclei were counterstained with DAPI. GFP expression indicates loss of the *Id1-IRES-Pac* cassette (see Fig.4.2a). The samples were imaged on an Olympus IX51 microscope. Scale bar: 100µm.

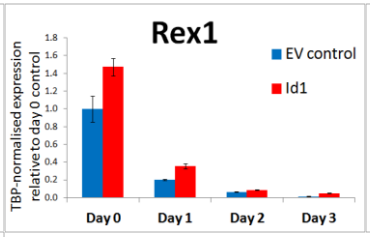
FIGURE 4.5

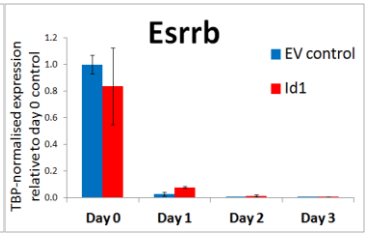
a – Id1



b - Naïve pluripotency







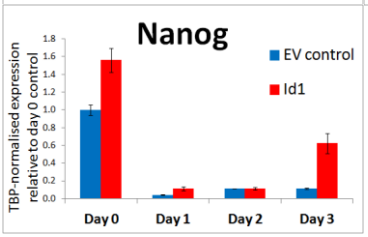
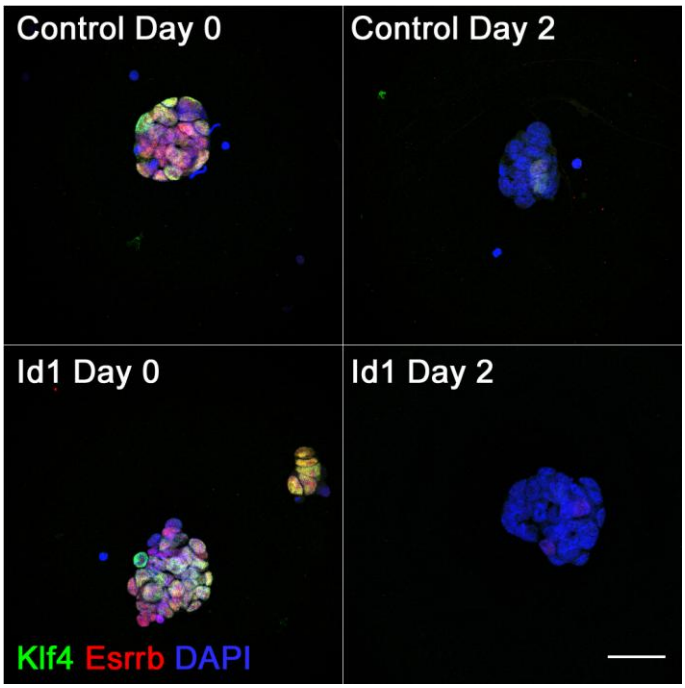


Figure 4.5 – Id1 does not delay the exit from naïve pluripotency

qRT-PCR on mRNA samples extracted from empty vector control (blue bars) and *Id1*-overexpressing (red bars) ES cells during the first three days of neural adherent monolayer differentiation.
ES cells cultured in LIF+FCS (Day 0) were replated in N2B27 and RNA extracted at daily timepoints.
Error bars: standard deviation of two biological replicates.

FIGURE 4.6

a


Control Day 0
Control Day 2
Id1 Day 0
Id1 Day 2
Klf4 Esrrb DAPI

b

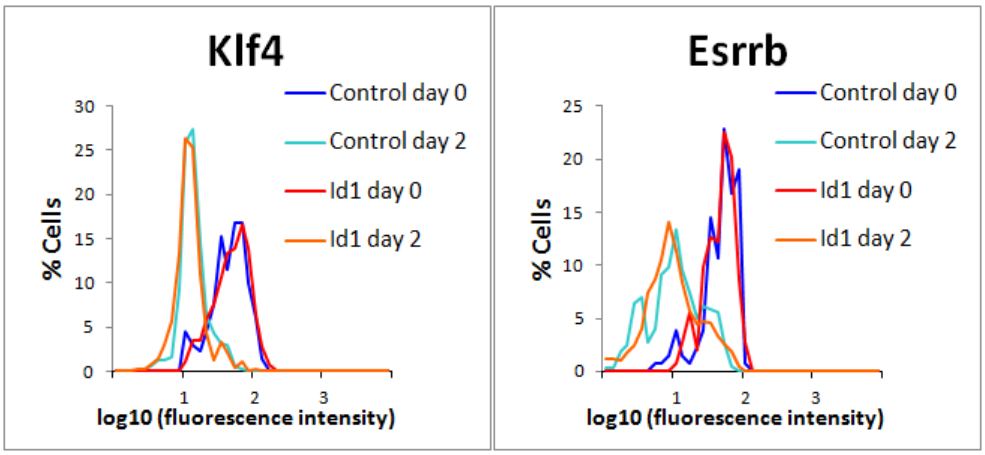

Klf4
Control day 0
Control day 2
Id1 day 0
Id1 day 2
% Cells
log10 (fluorescence intensity)
Esrrb
Control day 0
Control day 2
Id1 day 0
Id1 day 2
% Cells
log10 (fluorescence intensity)

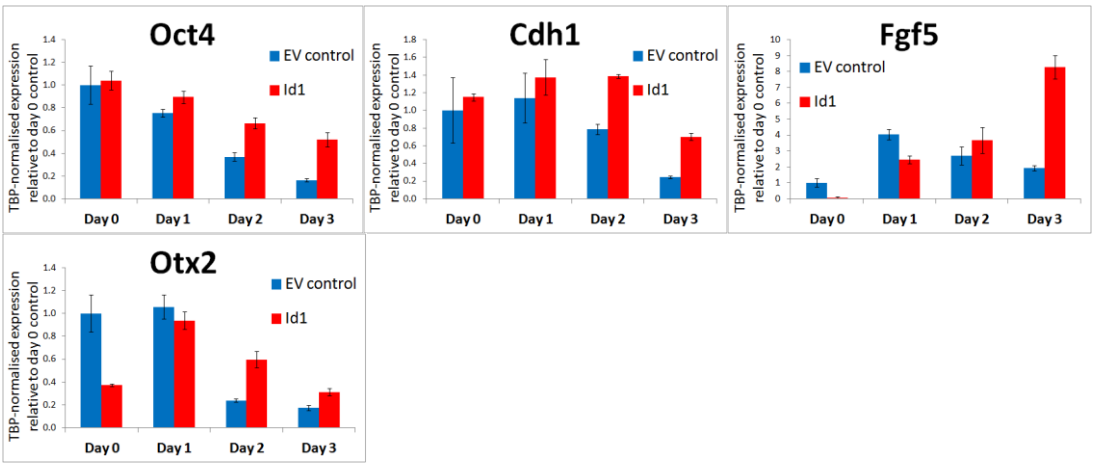
Figure 4.6 – Id1 does not prevent the loss of naïve pluripotency markers

a. Immunostaining for Klf4 and Esrrb in empty vector control and Id1-overexpressing ES cells cultured in N2B27+LIF+BMP4 (Day 0) and in N2B27 alone for 2 days (Day 2). Nuclei were counterstained with DAPI. Scale bar: 30μm. Images were acquired with a Leica TCS SPE confocal microscope.

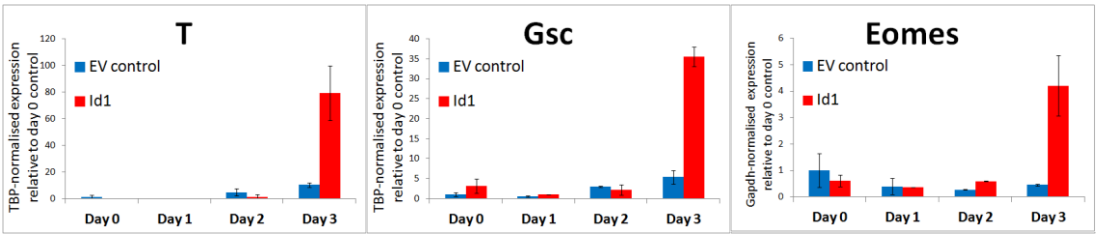
b. Immunostaining quantification of Klf4 and Esrrb nuclear signal in the samples shown in (a). A minimum of 130 nuclei were quantified for each condition.

FIGURE 4.7

a – Pluripotent epiblast



b – Proximal posterior epiblast



c – Exit from epiblast

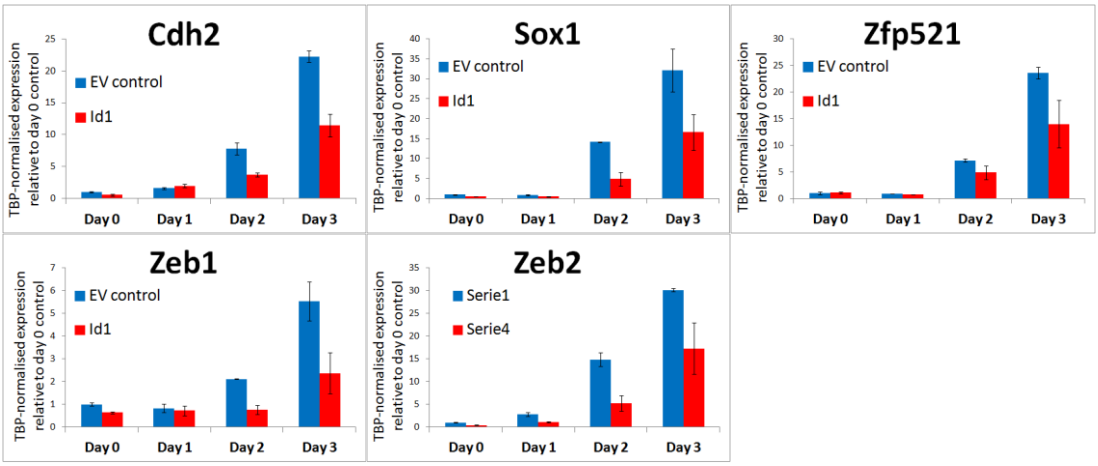


Figure 4.7 – Id1 delays the exit from a pluripotent epiblast-like state while promoting a proximal posterior identity

qRT-PCR on mRNA samples extracted from empty vector control (blue bars) and *Id1*-overexpressing (red bars) ES cells during the first three days of neural adherent monolayer differentiation (same RNA samples as in Figure 4.5).
ES cells cultured in LIF+FCS (Day 0) were replated in N2B27 and RNA extracted at daily timepoints.
Error bars: standard deviation of two biological replicates.

FIGURE 4.8

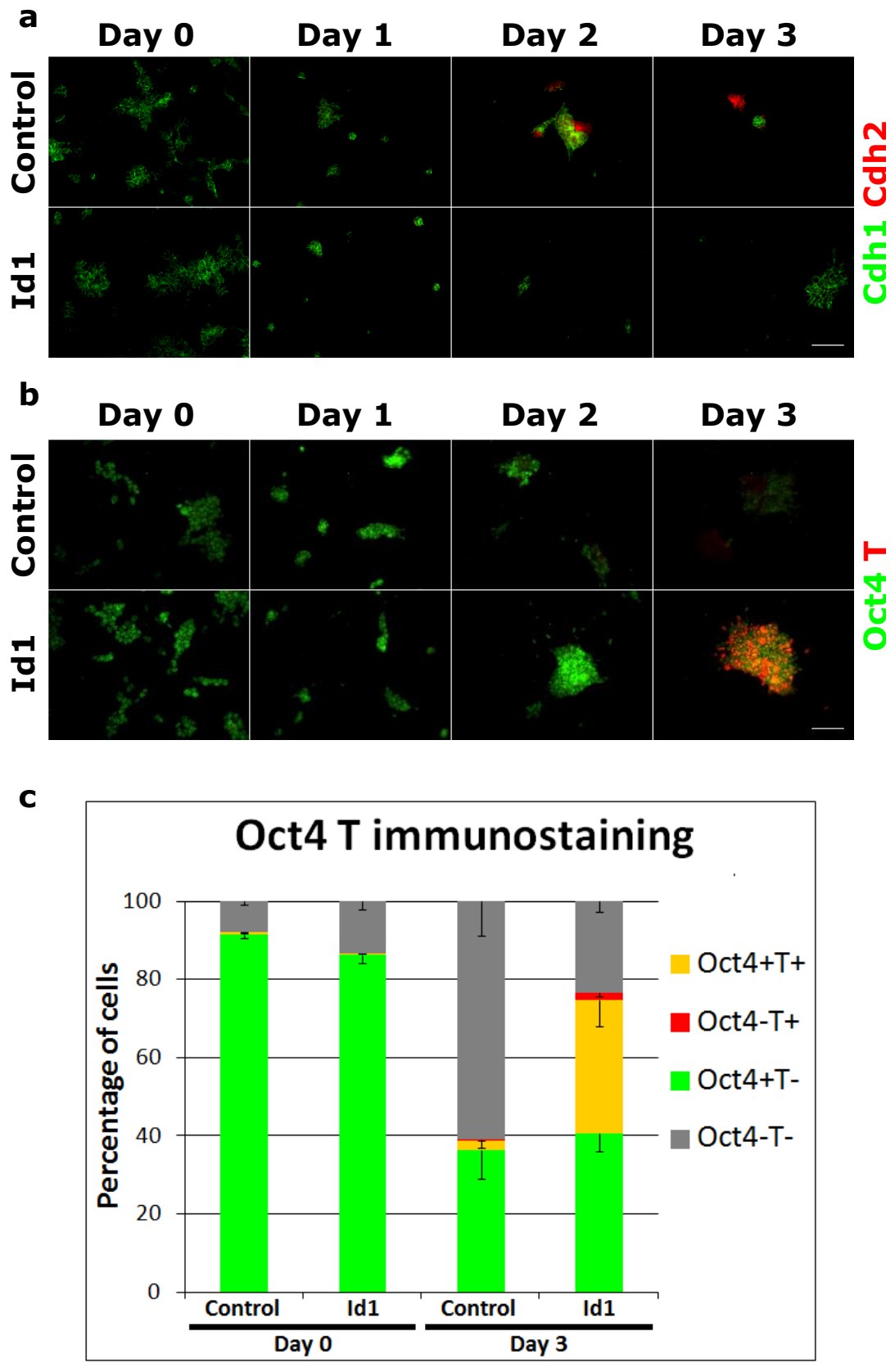

a
Day 0
Day 1
Day 2
Day 3
Control
Id1
Cdh1 Cdh2
b
Day 0
Day 1
Day 2
Day 3
Control
Id1
Oct4 T
c
Oct4 T immunostaining
Percentage of cells
100
80
60
40
20
0
Control
Id1
Control
Id1
Day 0
Day 3
Oct4+T+
Oct4-T+
Oct4+T-
Oct4-T-

Figure 4.8 – Id1 delays a switch from E- to N-cadherin and the loss of Oct4 whilst promoting upregulation of T

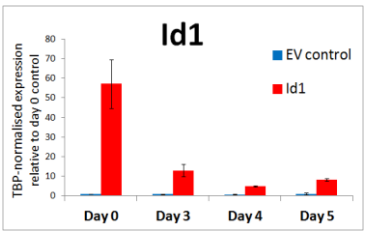
a. Immunostaining for Cdh1 and Cdh2 in control and Id1-overexpressing ES cells cultured in LIF+FCS (Day 0) and in N2B27 for 1, 2 and 3 days. The cells were imaged on an Olympus IX51 microscope. Scale bar: 100µm.

b. Immunostaining for Oct4 and T in control and Id1-overexpressing ES cells cultured in LIF+FCS (Day 0) and in N2B27 for 1, 2 and 3 days. The cells were imaged on an Olympus IX51 microscope. Scale bar: 100µm.

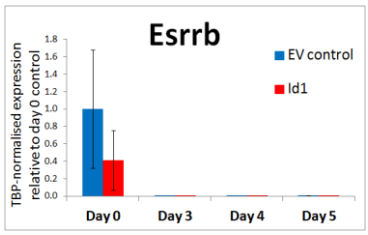
c. Immunostaining quantification of Oct4 and T nuclear signal in control and Id1-overexpressing ES cells cultured in LIF+FCS (Day 0) or in N2B27 for 3 days (Day 3). Five random fields were quantified for each of three biological replicates for each of the four conditions. A minimum of 3000 nuclei were quantified for each condition.
Error bars: standard error of the mean of 3 biological replicates.

FIGURE 4.9

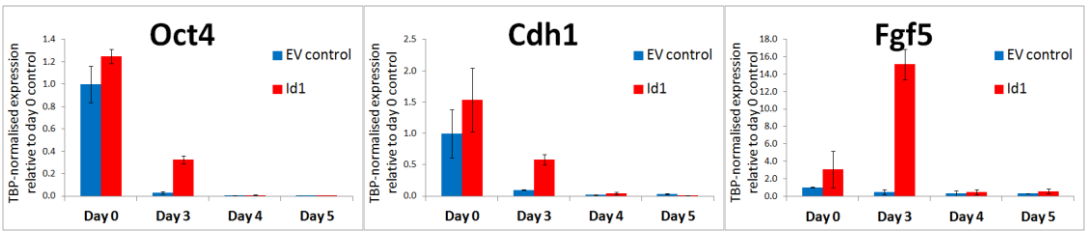
a – Id1



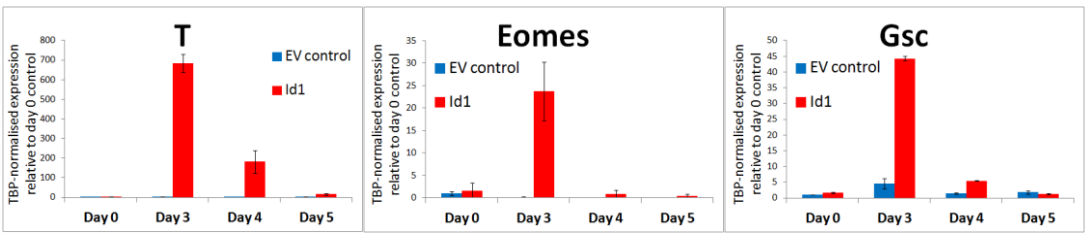
b – Naïve pluripotency



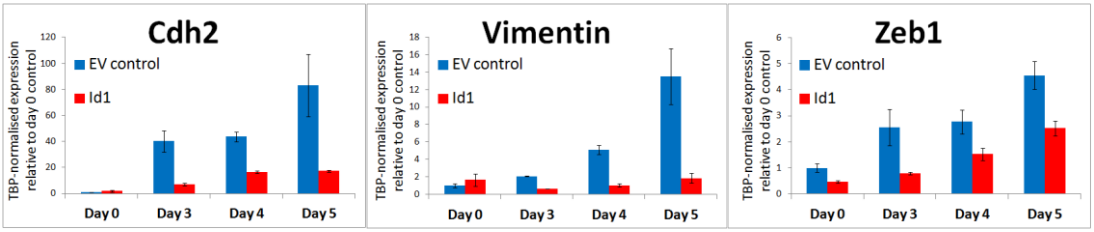
c – Pluripotent epiblast



d – Proximal posterior epiblast



e – Exit from epiblast (multilineage markers)



f – Exit from epiblast (neural ectoderm)

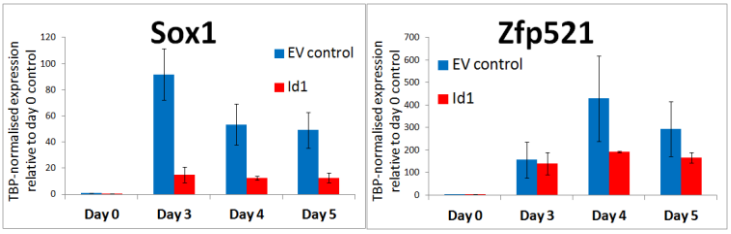
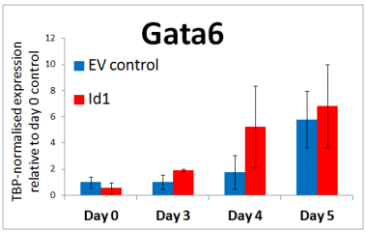


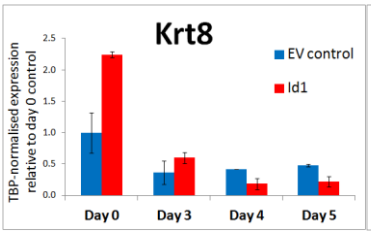
FIGURE CONTINUED ON THE NEXT PAGE

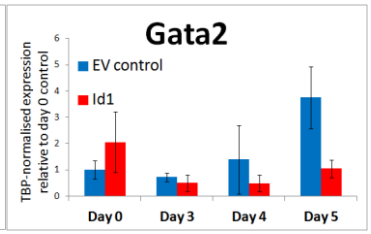
FIGURE CONTINUED FROM THE PREVIOUS PAGE

g – Exit from epiblast (endoderm)

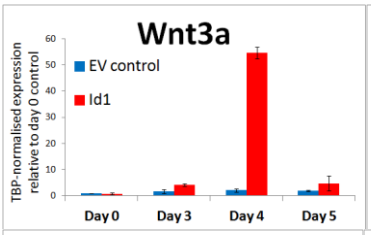


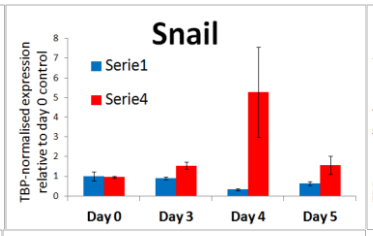
h – Exit from epiblast (surface ectoderm)

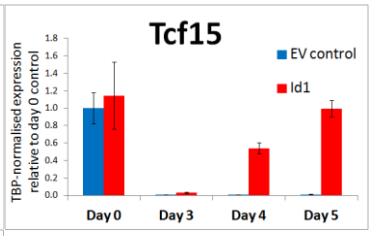


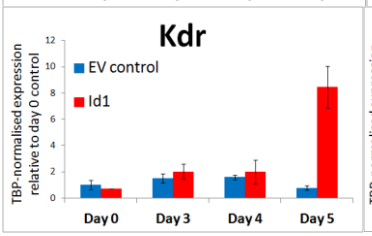


i – Exit from epiblast (mesoderm)









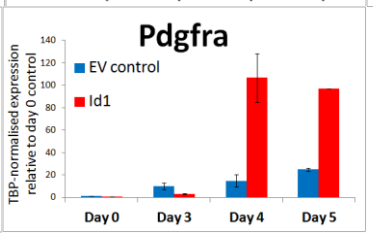


Figure 4.9 – Id1 ultimately favours mesodermal differentiation

qRT-PCR on empty vector control (blue bars) and *Id1*-overexpressing (red bars) ES cells at days 3, 4 and 5 of neural adherent monolayer differentiation.
ES cells cultured in LIF+FCS (Day 0) were replated in N2B27 and RNA extracted after 3, 4 and 5 days.
Error bars: standard deviation of two biological replicates.

FIGURE 4.10

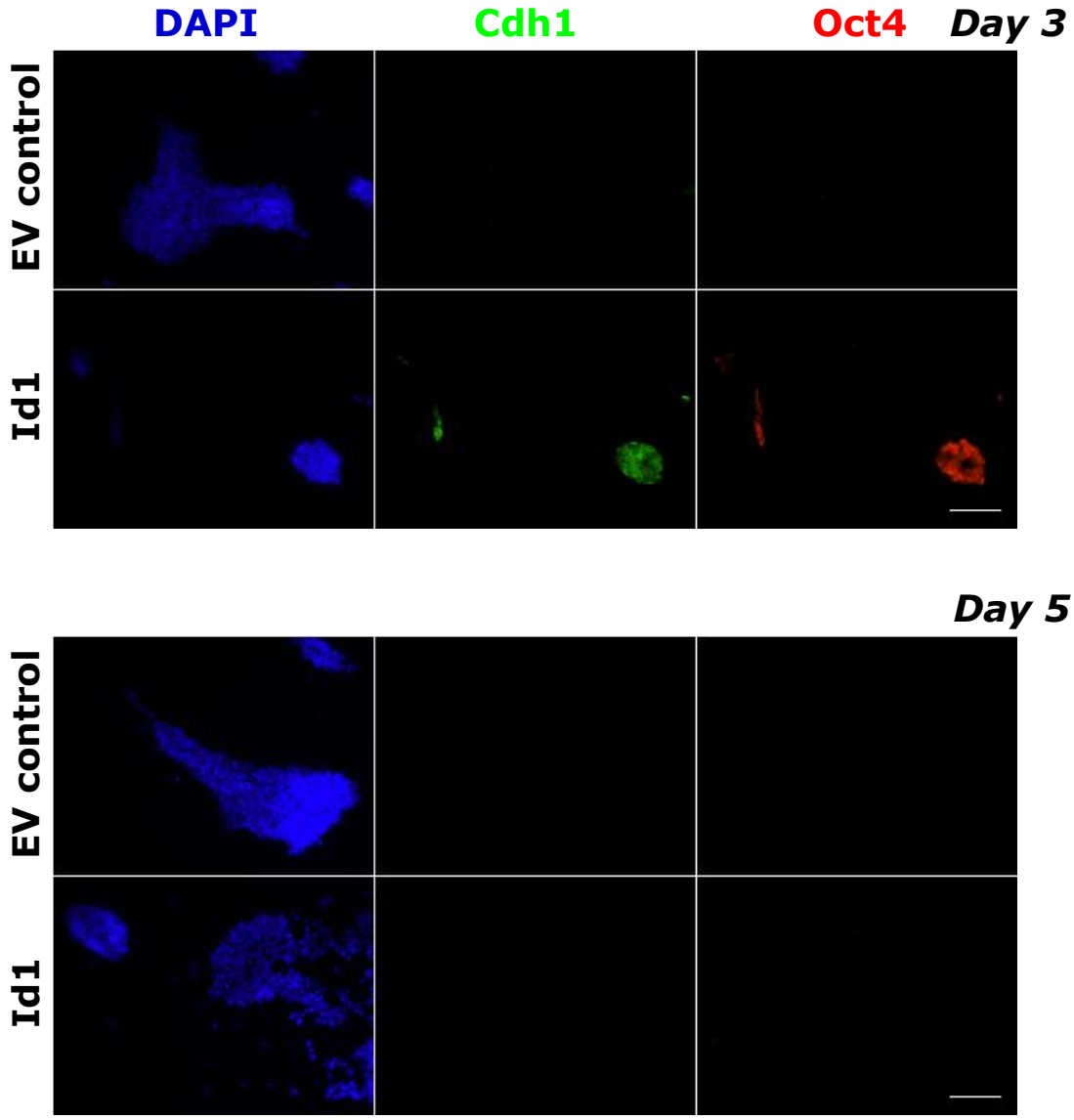


Figure 4.10 – Id1 delays but does not prevent the loss of Cdh1 and Oct4

Empty vector control and Id1 overexpressing ES cells differentiated for 3 and 5 days in N2B27 and immunostained for Cdh1 and Oct4. Nuclei were counterstained with DAPI. The cells were imaged on an Olympus IX51 microscope. Scale bars: 100µm.

FIGURE 4.11

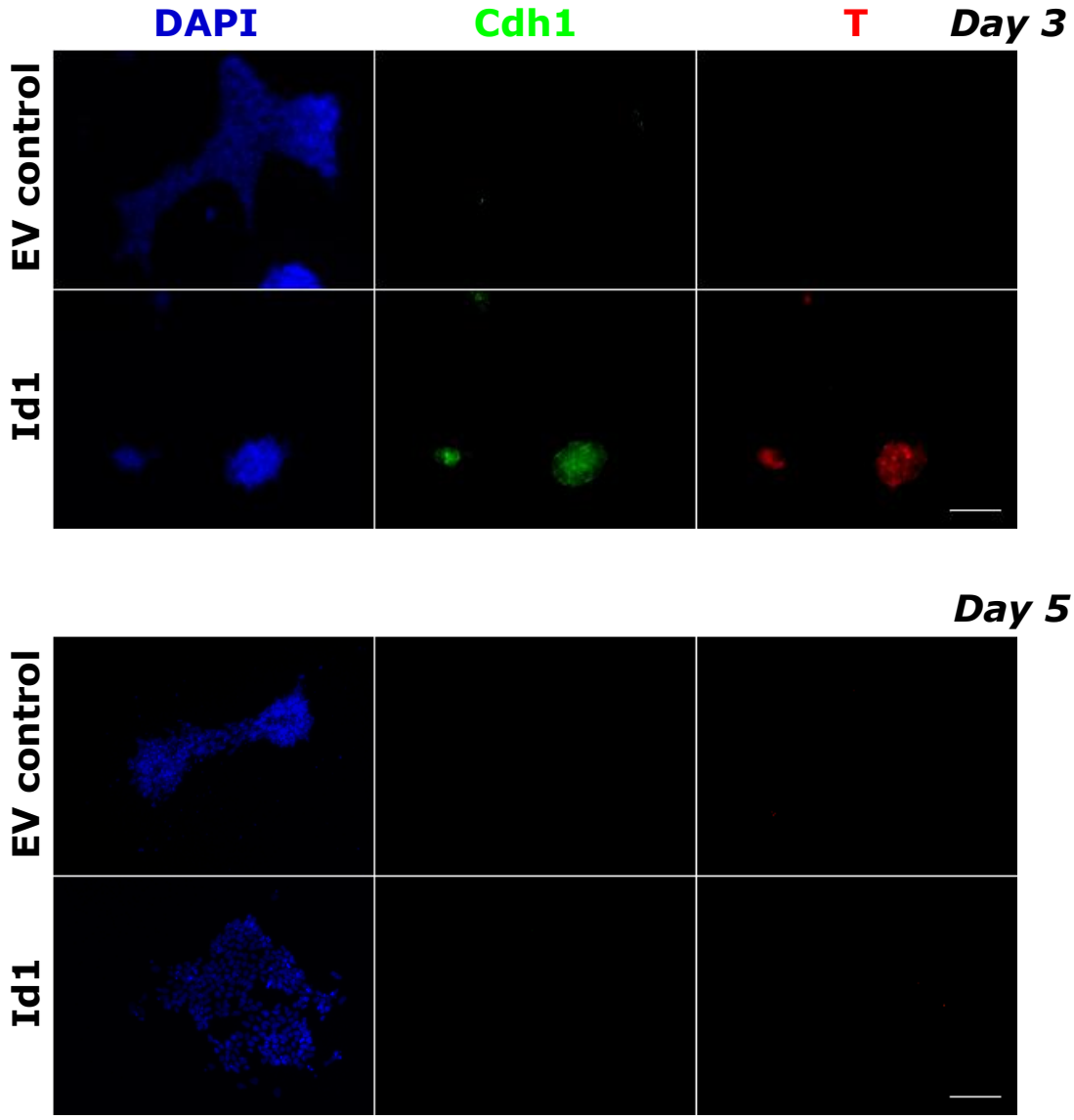


Figure 4.11 – Id1 promotes the generation of a Cdh1+T+ proximal posterior epiblast-like cell type but cannot maintain it

Empty vector control and Id1 overexpressing ES cells differentiated for 3 and 5 days in N2B27 and immunostained for Cdh1 and T. Nuclei were counterstained with DAPI. The cells were imaged on an Olympus IX51 microscope. Scale bars: 100μm.

FIGURE 4.12

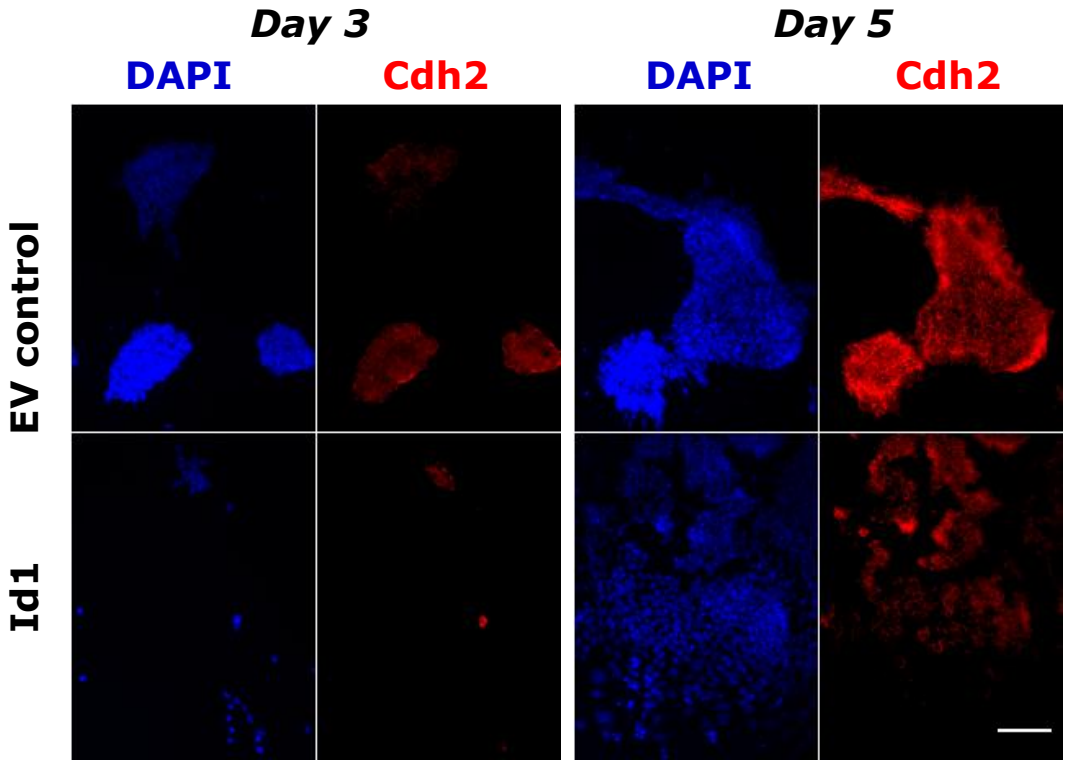
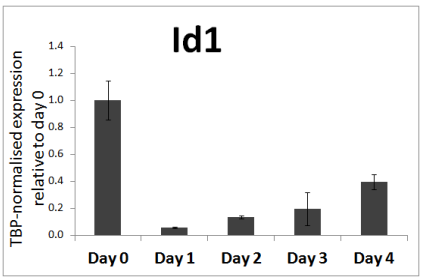


Figure 4.12 – Id1-overexpressing cells eventually acquire Cdh2 expression

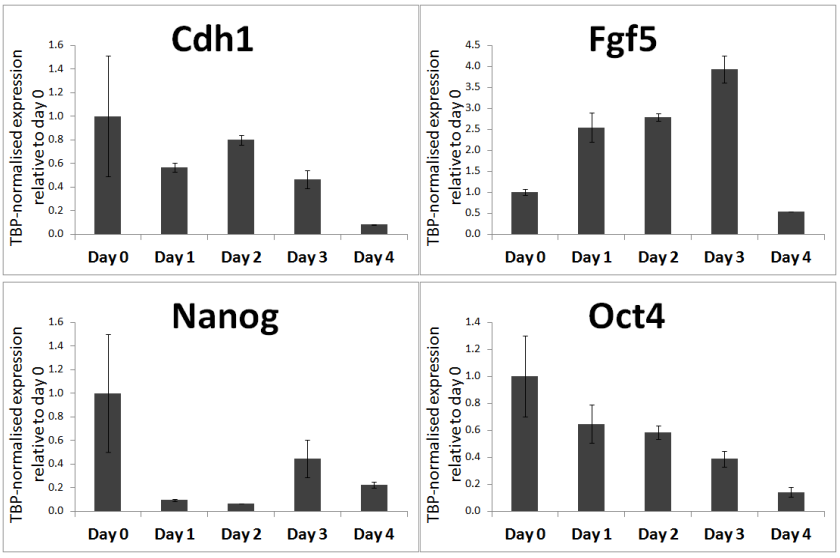
Empty vector control and Id1 overexpressing ES cells differentiated for 3 and 5 days in N2B27 and immunostained for Cdh2. Nuclei were counterstained with DAPI. The cells were imaged on an Olympus IX51 microscope. Scale bar: 100µm.

FIGURE 4.13

a – Id1



b – Pluripotent epiblast



c – Early neural and EMT

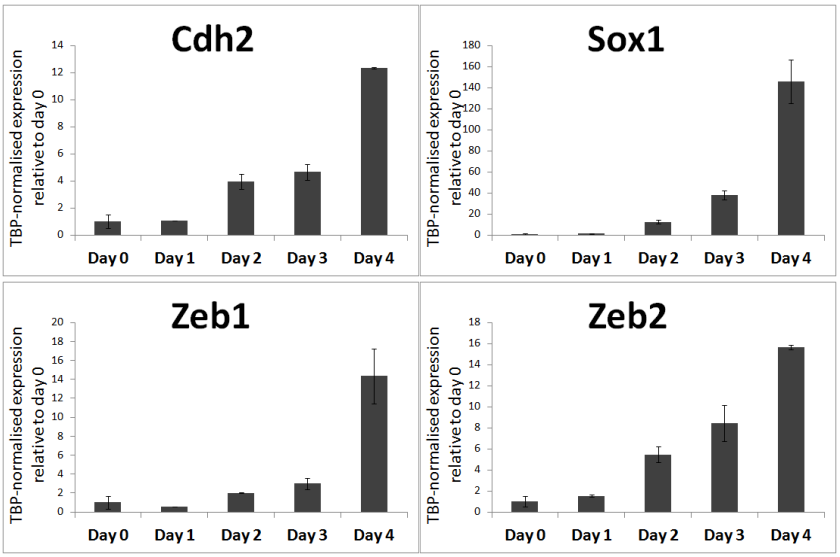
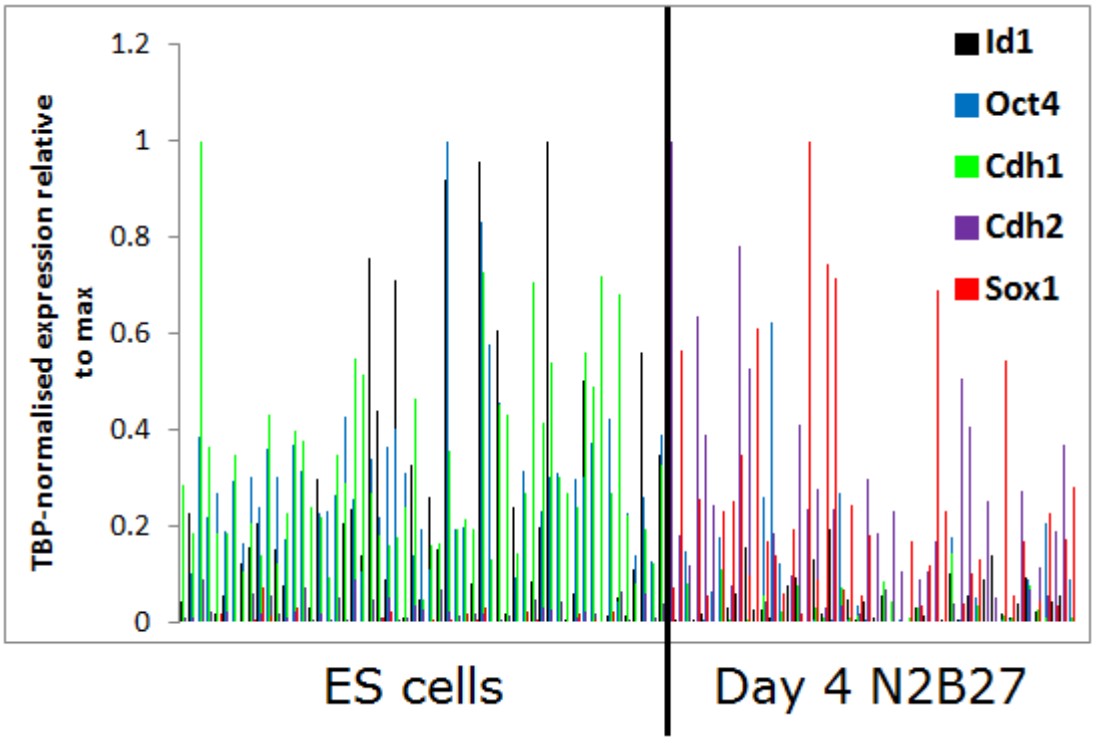


Figure 4.13 – Behaviour of markers selected for single-cell qRT-PCR on a whole population level in neural differentiation of ES cells

qRT-PCR on nine early neural and pluripotent epiblast markers during the first four days of neural adherent monolayer differentiation. E14tg2a ES cells cultured in LIF+FCS (Day 0) were replated in N2B27 and RNA extracted daily for 4 days.
Error bars: standard deviation of two biological replicates.

FIGURE 4.14

a



b

ρ	Oct4	Id1	Cdh1	Cdh2	Sox1
Oct4	1	0.32*	0.81*	-0.63*	-0.46*
Id1		1	0.39*	-0.18	-0.34*
Cdh1			1	-0.59*	-0.60*
Cdh2				1	0.50*
Sox1					1

*p-value < 0.05

Figure 4.14 – qRT-PCR on single ES and neural differentiated cells

a. mRNA levels for five genes in single E14tg2a ES cells cultured in LIF+FCS and in cells cultured for 4 days in N2B27. The levels of the normalising gene *Tbp* were measure in 90 cells for each timepoint. The cells with a *Tbp* value within a range of 0.5-1.5-fold the value of a 1:32 dilution of cDNA generated from 32 sorted cells were analysed for expression of the other factors. The final cell numbers for each timepoint were: ES cells: 57 cells, Day 4: 48 cells.

b. Spearman's rank correlation coefficients (ρ) for the five genes across all cells. A value of +1 would indicate a perfect positive correlation, a value of -1 a perfect negative correlation. Asymptotic p-values of less than 0.05 are plotted as asterisks.

FIGURE 4.15

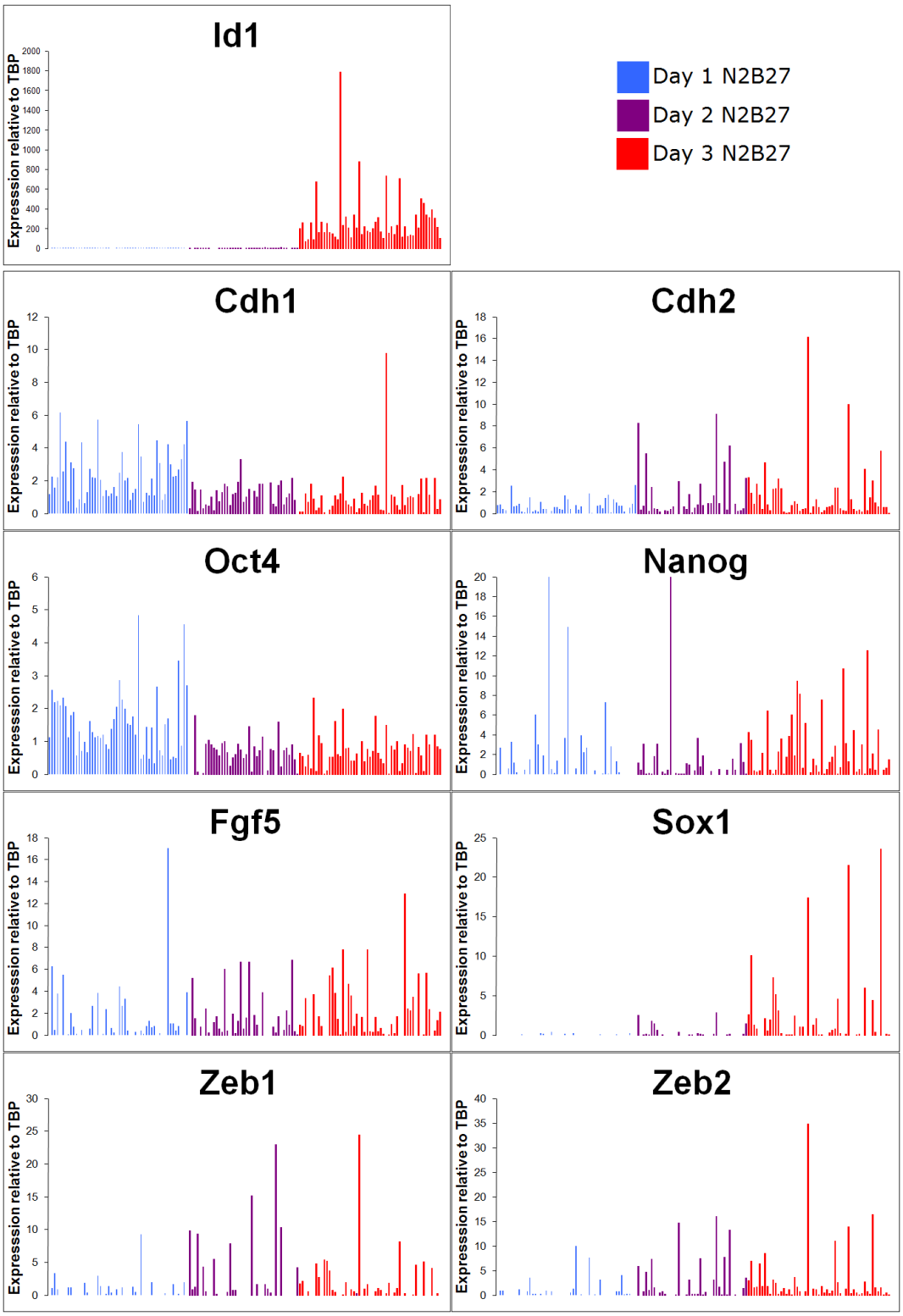

Id1
Cdh1
Cdh2
Oct4
Nanog
Fgf5
Sox1
Zeb1
Zeb2
Expresssion relative to TBP
Day 1 N2B27
Day 2 N2B27
Day 3 N2B27

Figure 4.15 – qRT-PCR on single cells during early neural differentiation

mRNA levels for nine genes in single E14tg2a ES cells cultured in N2B27 for 1, 2 or 3 days. The levels of the normalising gene *Tbp* were measured in 60 cells for each timepoint. The cells with a *Tbp* value within a range of 0.5-1.5-fold the value of a 1:32 dilution of cDNA generated from 32 sorted cells were analysed for expression of the other factors. The final cell numbers for each timepoint were: Day 1: 52, Day 2: 41, Day 3: 53.

FIGURE 4.16

Day 1	Id1	Cdh1	Oct4	Fgf5	Nanog	Cdh2	Sox1	Zeb1	Zeb2
Id1	1	0.44*	0.32*	0.05	0.20	0.13	0.11	0.04	0.23
Cdh1		1	0.42*	0.30*	0.04	0.27	0.01	-0.05	0.19
Oct4			1	0.15	0.10	0.08	0.05	0.03	-0.11
Fgf5				1	-0.21	0.20	0.05	-0.01	0.32*
Nanog					1	0.02	-0.06	-0.15	-0.24
Cdh2						1	0.04	0.08	0.12
Sox1							1	-0.18	0.19
Zeb1								1	0.29*
Zeb2									1

Day 2	Id1	Cdh1	Oct4	Fgf5	Nanog	Cdh2	Sox1	Zeb1	Zeb2
Id1	1	-0.18	-0.12	-0.19	0.21	0.56*	0.37*	0.36*	0.39*
Cdh1		1	0.47*	0.52*	0.19	-0.20	-0.50*	-0.46*	-0.44*
Oct4			1	0.53*	0.45*	-0.28	-0.26	-0.58*	-0.50*
Fgf5				1	0.36*	-0.41*	-0.34*	-0.49*	-0.51*
Nanog					1	-0.01	0.15	-0.26	-0.19
Cdh2						1	0.54*	0.47*	0.61*
Sox1							1	0.43*	0.66*
Zeb1								1	0.68*
Zeb2									1

Day 3	Id1	Cdh1	Oct4	Fgf5	Nanog	Cdh2	Sox1	Zeb1	Zeb2
Id1	1	0.03	-0.12	-0.39*	-0.06	0.38*	0.24	0.31*	0.32*
Cdh1		1	0.58*	0.21*	0.03	-0.46*	-0.65*	-0.72*	-0.42*
Oct4			1	0.60*	0.35*	-0.57*	-0.59*	-0.47*	-0.37*
Fgf5				1	0.47*	-0.47*	-0.36*	-0.32*	-0.40*
Nanog					1	-0.29*	-0.25	-0.14	-0.09
Cdh2						1	0.65*	0.61*	0.62*
Sox1							1	0.71*	0.52*
Zeb1								1	0.55*
Zeb2									1

*p-value < 0.05

Positive correlation between epiblast markers

Positive correlation between neural markers

Negative correlation between epiblast and neural markers

Figure 4.16 – Distinct pluripotent epiblast and neural gene expression signatures in differentiating ES cells

Spearman's rank correlation coefficients for the nine genes plotted in Fig.4.15 at day 1 (blue table), 2 (red table) and 3 (purple table) of neural differentiation.
A value of 1 would indicate a perfect positive correlation, a value of -1 a perfect negative correlation. Asymptotic p-values of less than 0.05 are plotted as asterisks.
Green boxes: statistically significant positive correlations between pluripotent epiblast markers, red boxes: statistically significant positive correlations between neural markers, blue boxes: statistically significant negative correlations between pluripotent epiblast and neural markers.

FIGURE 4.17

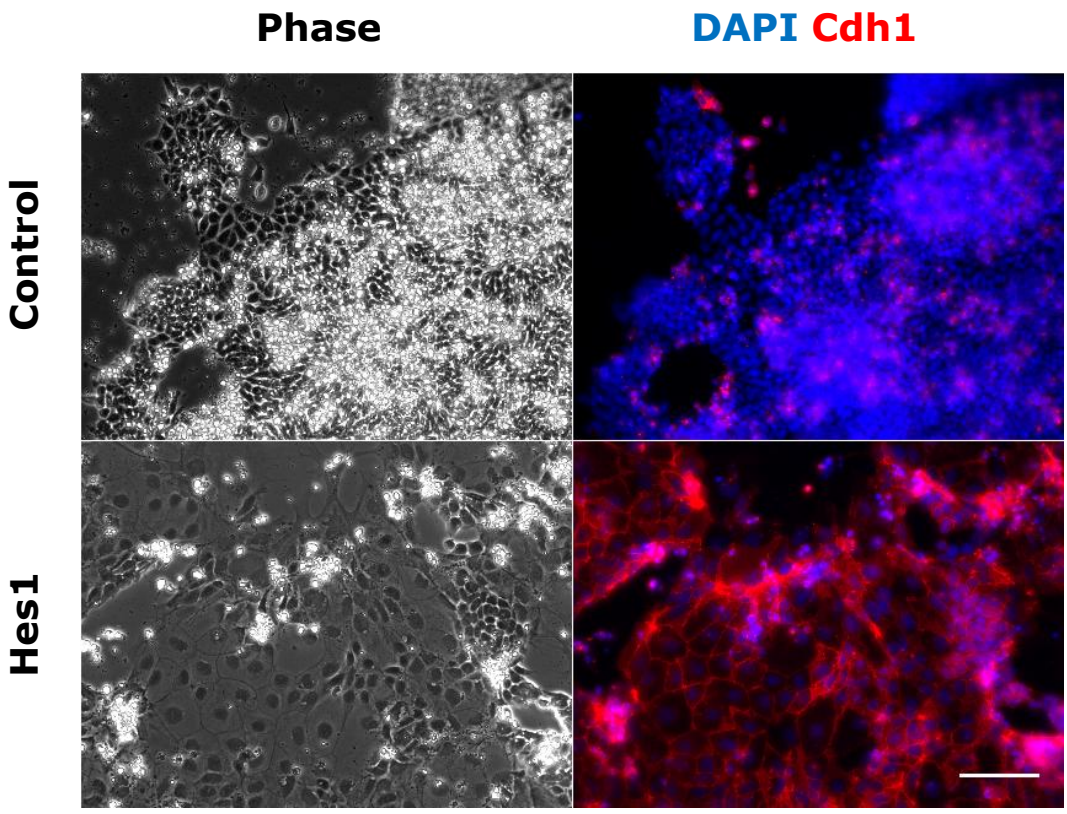


Figure 4.17 – Hes1 overexpression recapitulates the effects of BMP4 in N2B27 culture

TRE-Hes1 ES cells (cells with a doxycycline-inducible *Hes1* transgene, see Zhou *et al.* 2013) were cultured in N2B27 in the presence or absence of 1µg/ml doxycycline for 6 days. The cells were then fixed, stained with an antibody against Cdh1 and imaged on an Olympus IX51 microscope. Scale bar: 100µm.

CHAPTER 5

The expression pattern of Id1 in early post-implantation embryonic development

5.1 Introduction

The expression pattern of Id1 in the developing embryo has been addressed by means of *in situ* hybridisation by numerous research groups ever since the factor was first discovered (Duncan *et al.* 1992, Duncan *et al.* 1994, Evans & O'Brien 1993, Gray *et al.* 2004, Jen *et al.* 1996, Jen *et al.* 1997, Li *et al.* 2013, Wang *et al.* 1992).

Until recently, only two of these studies include images of *Id1* expression in early post-implantation embryos (Jen *et al.* 1997, Wang *et al.* 1992). *Id1* is expressed in the proximal regions of the embryo from pre-streak stages, in accordance with the expression of BMP4 in proximal extraembryonic tissues, and remains expressed throughout embryonic development.

In both of the aforementioned studies, the researchers made use of radioactively labelled hybridisation probes, resulting in low-definition images with unclear expression boundaries between embryonic and extraembryonic tissues and between different germ layers. It is for example hard to interpret whether *Id1* message is detected in the embryo proper or in visceral endoderm at pre- and peri-gastrulation stages.

A recent publication by Li *et al.* (2013) confirmed the proximal expression pattern of *Id1* between E6.5 and E7.5 and appears to indicate that *Id1* is expressed in the embryo proper, but did not fully clarify how *Id1* expression is distributed amongst the three germ layers.

In order to characterise the embryonic expression pattern of Id1 in more detail, I made use of Id1-Venus reporter ES cell lines, which express an Id1-Venus fusion

protein from one endogenous *Id1* locus (see sections 2.2.4.1.5-2.2.4.1.7 and 3.2.4 in previous chapters for more detail on these cells), to generate high-contribution chimaeric embryos. The analysis of Venus fluorescence in chimaeras dissected at different stages should allow the painting of a high-resolution picture of the localisation of Id1 protein in early post-implantation embryonic development, as well as allowing the study of the co-expression of Id1 with markers of different cell identities.

5.2 Results

5.2.1 Id1 is expressed in the proximal regions of E6.5 chimaeras

In order to verify the feasibility of the use of chimaeric embryos to study the expression pattern of Id1 in the early stages of post-implantation development, I set out to assess the degree of contribution of ES cells to chimaeras and to verify whether the pattern of Id1-Venus expression reproduces previously published *in situ* hybridisation data.

I made use of Id1-Venus reporter ES cells (IdV) stably expressing the red fluorescent protein mKate2, fused at its C-terminus to three SV40 nuclear localisation signals, under the control of the strong *CAG* promoter. mKate2-NLS serves as a constitutive lineage label to allow all ES-derived cells to be identified within chimaeric embryos. This clonal cell line was named IdVnK1. For further information on this cell line, please refer to section 2.2.4.1.7 of the Materials and Methods chapter.

ES cells were aggregated to morulae overnight and blastocysts were transferred into pseudopregnant females the next day. Embryos were dissected out of the impregnated uteri after 4 days, at an approximate equivalent of embryonic day 6.5 for a normal pregnancy. The embryos were fixed and stained with a polyclonal anti-GFP antibody capable of recognising the Venus epitope, and imaged by confocal

microscopy (Figure 5.1). The staging of the embryos was performed by assessing their morphological features, and comparing them to those described by Downs & Davies (1993), whose staging system will be used throughout this thesis.

The contribution of the ES cells to the conceptuses was very high for 13 out of 15 embryos, as judged by the expression of mKate2-NLS in the embryo proper.

In accordance with the *in situ* hybridisation data published by Jen *et al.* (1997) and Li *et al.* (2013), Id1-Venus expression is detected in the proximal regions of the embryos. This pattern is compatible with the published expression pattern for *Bmp4* in these early post-implantation stages. *Bmp4* is expressed in the regions of the extraembryonic ectoderm adjacent to the proximal epiblast at E6.5 and E7.5 (Coucouvanis & Martin 1999, Norris *et al.* 2002, Perea-Gómez *et al.* 1999, Perea-Gomez *et al.* 2001, Tremblay *et al.* 2001).

The smallest pre-streak stage embryos (Figure 5a,b) appear to harbour some Venus-expressing cells in more distal regions than the bigger early- and mid-streak stage embryos (Figure 5e-o).

Some of the bigger early- and mid-streak stage embryos (Figure 5g-k) appear to have a domain of high Id1-Venus expression in the proximal anterior epiblast, the epithelial cells in this region appearing to be brighter than their counterparts in the proximal posterior epiblast. This matches the *in situ* hybridisation data for *Id1* at E7.0 (Li *et al.* 2013).

5.2.2 Id1 is co-expressed with Nanog and T in the mid-streak stage epiblast

Having established that chimaeras generated from morula aggregation display high levels of ES cell contribution, I decided to investigate the relationship between Id1-Venus and other proteins in E6.5 embryos. The non-lineage labelled IdV ES cell line was used for morula aggregation, so that expression of other proteins could be

visualised with red and far red secondary antibodies. It should be kept in mind that while conclusions regarding co-expression of Id1-Venus with other factors may be drawn for Id1-Venus-expressing cells, there is no way of discerning whether Id1-Venus-negative cells within the chimaeras are not expressing Id1 or are not ES-cell derived, so no conclusion may be drawn regarding these cells.

A mid-streak stage chimaera was sectioned transversally. Seven adjacent 7µm-thick sections of the proximal regions of the epiblast were stained with antibodies against GFP, Nanog and T (Figure 5.2). Nanog is expressed in the posterior regions of E6.5 embryos (Morkel *et al.* 2003, Osorno *et al.* 2012); T is expressed in the primitive streak, both in the epiblast and in the ingressing mesoderm (Beddington *et al.* 1992, Rivea-Pérez & Magnuson 2005, Wilkinson *et al.* 1990, Willison 1990).

Id1-Venus is co-expressed with both Nanog and T, and is found in the anterior and posterior regions of the embryo (Figure 5.2a,b). In order to quantify the expression levels and patterns of the three factors in the anterior and posterior epiblast, I split each section into anterior and posterior regions of interest based on expression of Nanog. I took care to exclude visceral endoderm and other extraembryonic tissue from the analysis. The regions of interest for one of the seven sections are shown in Figure 5.2b.

I used Multicell3D (see Chapter 2 section 2.2.3.2.8) to quantify the fluorescence intensity signals for the three factors in each nucleus across all sections. I used the background fluorescence intensity signals detected in the nuclei of two 7µm-thick sections of extraembryonic ectoderm to establish a threshold to discriminate between cells expressing or not expressing the three factors of interest (Figure 5.3a). I generated dot plots to compare the level of each factor against the others in the whole of the proximal epiblast or in the anterior and posterior regions of interest I defined (Figure 5.3a). The absence of T-expressing cells within the anterior region of interest suggests a correct demarcation of anterior and posterior epiblast.

Many Id1-Venus-positive cells express Nanog, and there is no obvious negative relationship between the two factors at high levels of expression. Id1-Venus-positive cells do not appear to be enriched in, or depleted of, Nanog expression in comparison with the proximal epiblast as a whole: approximately 65% of Id1-Venus-positive cells are Nanog-positive, and approximately 62% of cells within the whole proximal epiblast are Nanog-positive (Figure 5.3b). The percentage of Nanog-positive Id1-Venus-positive cells increases to approximately 83% in the posterior epiblast, in keeping with the higher proportion of Nanog-expressing cells in this region as a whole (Figures 5.3c).

Similarly, Id1-Venus-positive cells can express high levels of T, although the majority of Id1-Venus-expressing cells is T-negative, as is expected by the broad distribution of Venus fluorescence throughout the embryo and the restriction of T to the primitive streak region (Figure 5.3d).

There appears to be a positive relationship between T and Nanog, as over 85% of T-expressing cells co-express Nanog (Figure 5.3e). This suggests that at this developmental stage T is primarily marking the region of the epiblast where the primitive streak is located, rather than mesodermal cells ingressing through the streak, since Nanog is downregulated as epiblast cells start to ingress (Morkel *et al.* 2003).

The pattern of Id1-Venus expression is similar in the whole, anterior and posterior regions of the proximal epiblast. Nanog and T, on the other hand, are enriched in the posterior epiblast, with complete lack of expression of T in the anterior epiblast (Figure 5.3f).

5.2.3 Id1 is co-expressed with Nanog and T in the early bud stage epiblast

I set out to investigate whether the co-expression of Id1-Venus and Nanog in the epiblast and the positive relationship between Nanog and T perdured later in development.

Chimaeras were generated by morula aggregation of IdV ES cells, and embryos were dissected after 5 days of blastocyst transfer to pseudopregnant females, at a stage approximately equivalent to wild-type E7.5 embryos.

I stained an early bud stage embryo with antibodies against GFP, Nanog and T and imaged the conceptus by confocal microscopy (Figure 5.4a). Id1-Venus is expressed proximally, but its domain of expression is expanded and includes the migrating mesoderm and endoderm. Nanog is still expressed at this stage and it is enriched, but not rigidly confined to, the posterior half of the embryo. T is expressed in the primitive streak and along the developing notochord.

The embryo was then sectioned in 7µm-thick sections, some of which were restained with antibodies against the same three markers (Figure 5.4b). Id1-Venus is expressed by the majority of cells in the proximal section shown in Figure 5.4b, both in the epiblast and in the migrating mesodermal and endodermal layers. Nanog is expressed in the posterior half of the epiblast, and is downregulated after migration through the primitive streak. It is interesting to observe that the downregulation of Nanog protein does not occur until after the migratory event, in contrast to *Nanog* mRNA, which is downregulated immediately prior to EMT (Morkel *et al.* 2003). T is expressed in the primitive streak region of the epiblast and by cells migrating through it. The expression of each of these markers does not appear to exclude the expression of the other two, so that Id1-Venus is co-expressed with Nanog in the posterior half of the epiblast, and both of these are co-expressed with T at the primitive streak.

5.2.4 Id1 is expressed proximally in headfold stage embryos

I proceeded with my investigation into the expression pattern of Id1 in later stages of embryonic development by making use of the lineage-labelled IdVnK1 ES cells. The analysis described in the previous section made use of IdV ES cells lacking the mKate2-NLS lineage label, in order to allow 3-colour marker analysis.
IdVnK1 ES cells were aggregated to morulae and chimaeras were dissected after 5 days of blastocyst transfer to pseudopregnant females, at a stage approximately equivalent to E7.5 embryos from a natural pregnancy.

I stained five embryos with an anti-GFP antibody and imaged them for expression of Id1-Venus and mKate2-NLS by confocal microscopy (Figure 5.5). The contribution of ES cells to all five chimaeras was very high, with the majority of embryonic cells expressing mKate2-NLS. Id1-Venus was expressed proximally in all five embryos, both in the anterior, lateral and posterior regions of the chimaeras. This expression pattern is once again in accordance to the expression domains of *Bmp2* and *Bmp4* in the extraembryonic ectoderm adjacent to the proximal regions of the embryo proper (Du *et al.* 2010, García-García & Anderson 2003, Madabhushi & Lacy 2011, Zakin & De Robertis 2004). Interestingly, in one of the embryos Id1-Venus appears to be expressed in a ring of cells surrounding the node (Figure 5.5c), an expression pattern reminiscent of that of *Nodal* (Nakaya *et al.* 2005).

5.2.5 Id1 is expressed in the foregut primordium, cardiac crescent, lateral mesoderm and allantois

The latest stage of embryonic development I analysed was the late headfold/cardiac crescent stage. Chimaeras were generated from IdV cells and dissected after 5.5 days of blastocyst transfer to pseudopregnant females, at a stage approximately equivalent to E7.5-E8.0 embryos from a natural pregnancy.

I stained seven embryos with an anti-GFP antibody and imaged them by confocal microscopy (Figure 5.7a). The expression pattern of Id1-Venus is well defined, and confined anteriorly to the cardiac crescent mesoderm and the underlying foregut primordium endoderm, and posteriorly to the lateral mesoderm and allantois. Id1-Venus expression can once again be detected at the node of some of the embryos (Figure 5.7a embryos i, iv, v, vii). It is impossible to draw a conclusion on whether this expression is transitory, and hence can only be observed for some embryos, or whether the chimaeras which do not exhibit Id1-Venus expression in the node are imaged from the wrong angle or have no ES cell contribution to this embryonic structure.

I then stained one embryo with antibodies against GFP and Nkx2.5, a transcription factor expressed in the cardiac crescent and foregut primordium (Lints *et al.* 1993) (Figure 5.7b). Id1-Venus and Nkx2.5 are co-expressed in the foregut primordium and in the cardiac crescent mesoderm. Nkx2.5 is expressed at similar levels in both regions, whilst Id1-Venus is expressed at a higher level in the foregut primordium.

To confirm this expression pattern with lineage-labelled cells, chimaeras generated from IdVnK1 ES cells were dissected at the same stage of development. I stained them with polyclonal antibodies against GFP and tagRFP, this second antibody being capable of recognising the mKate2 epitope. I imaged three embryos by confocal microscopy (Figure 5.7c,d). The contribution of the ES cells was very high, with the majority of cells in the embryos expressing mKate2-NLS. I confirmed the expression pattern seen for IdV chimaeras, Id1-Venus being expressed in the cardiac crescent, the foregut primordium, lateral mesoderm and, for two out of three embryos (Figure 5.7c embryos i, iii), the node. None of the embryos had an intact allantois so I was unable to assess Id1-Venus expression in this tissue.

5.3 Discussion

I made use of morula aggregation to generate chimaeras harbouring an Id1-Venus reporter allele in the ES-cell derived tissues. This allowed me to assess the expression of Id1 at high resolution at the single cell level within the embryo proper. The use of two different ES cell lines for the morula aggregation experiments offered the possibility to address different aspects of the expression of Id1 *in vivo*: the lineage-labelled line expressing mKate2-NLS from the ubiquitous *CAG* promoter was used to assess the status of Id1 expression both in Id1-Venus expressing and non-expressing labelled cells throughout the embryo; the non-lineage labelled line was used to assess whether Id1 was co-expressed with other transcription factors in the early post-implantation epiblast and across all germ layers.

5.3.1 Id1 expression in pre- and early-streak stage epiblast

I initially assessed the expression of Id1 in pre- and early-streak stage embryos by using the lineage-labelled Id1-Venus reporter ES cell line (IdVnK1). I observed that Id1-Venus is expressed in the proximal regions of the epiblast, both anteriorly and posteriorly (Figure 5.1). This expression pattern complements the data reported by Jen *et al.* (1997), who had observed *Id1* mRNA expression in the proximal regions of the embryo, without fully clarifying whether the expression was detected in the embryonic epiblast or limited to the visceral endoderm, and is in accordance with the recent *in situ* hybridisation data of Li *et al.* (2013). The expression pattern also fits with *in situ* hybridisation experiments analysing *Bmp4* expression, a key inducer of *Id1* transcription. *Bmp4* is expressed by extraembryonic tissue adjacent to the proximal regions of the epiblast at E5.5 and E6.5 (Coucouvanis & Martin 1999, Norris *et al.* 2002, Perea-Gomez *et al.* 2001, Polydorou & Georgiades 2013).

The embryos of the smallest size appeared to exhibit Id1-Venus expression not only in the proximal regions, but to a lower level, in some cells located more distally. It is an intriguing possibility that this may be due to the distance which can be covered by the morphogen BMP4 from its lieu of expression in the extraembryonic ectoderm. Alternatively, other molecules of the BMP family may be responsible for Id1 expression in these more distal cells: *Bmp2* is expressed throughout the visceral endoderm at E5.5 before localising to the proximal anterior epiblast at E6.5, *Bmp7* is expressed at the distal tip and in the anterior primitive streak at E6.5 (Arkell & Beddington 1997, Coucouvanis & Martin 1999, Madabhushi & Lacy 2011, Solloway & Robertson 1999). Whatever BMP factor or combination of factors may be responsible for Id1-Venus induction, it is reassuring that this pattern coincides with the localisation of phosphorylated Smad1 protein, which is observed throughout the embryo at the pre-streak stage, and is subsequently lost in the distal tip and detected only in the proximal regions of gastrulating embryos (Aoyama *et al.* 2012, Di-Gregorio *et al.* 2007).

5.3.2 Id1 expression in mid- and late-streak stage epiblast

I proceeded to analyse the expression pattern of Id1-Venus in mid- and late-streak stage embryonic cells making use of chimaeras derived from both the lineage-labelled reporter ES cells and from the non-lineage labelled reporter ES cells, co-staining the latter to detect expression of Nanog and T in a mid-streak stage chimaera (Figures 5.1, 5.2).

The proximal expression of Id1-Venus was conserved in mid- and late-streak stage chimaeric embryos, the HLH factor being expressed both in the epiblast and in the migrating mesoderm. This is again fitting with the observations of Jen *et al.* (1997), and with the reported expression of *Bmp4*. In some of the embryos, the expression of Id1-Venus appeared to be more intense in the anterior than in the posterior proximal

epiblast, coinciding with an area of the embryos that at the late-streak stage is characterised by expression of *Bmp2* and *Dlx5* and fated for surface ectoderm (Cajal *et al.* 2012, Madabhushi & Lacy 2011, Tam & Behringer 1997). It is an interesting possibility that BMP2 could be responsible for inducing Id1 expression in the anterior to complement the Id1-inducing action of BMP4 in the posterior.

Bmp7 is expressed around the node at this stage (Arkell & Beddington 1997, Solloway & Robertson 1999), but Id1-Venus does not appear to be expressed in this structure in any of the imaged embryos in Figure 5.1 (see below for discussion on this point). *Bmp5* displays a similar pattern of expression to *Bmp4* at these embryonic stages, and could also play a role in driving Id1 expression (Solloway & Robertson 1999).

Immunostaining of sections of a chimaera generated with non-lineage labelled Id1-Venus reporter ES cells with antibodies against GFP, Nanog and T revealed that Id1-Venus expression is localised equally in the anterior and posterior proximal epiblast and does not appear to be preferentially associated with expression of Nanog or T (Figures 5.2, 5.3). Nanog was expressed by the majority of Id1-Venus-expressing cells, suggesting that there is no negative relationship between the two factors in the gastrulating epiblast, and that the majority of Id1-Venus-expressing cells reside in the pluripotent epiblast at the mid-streak stage. T-expressing cells were mostly Nanog-positive, suggesting they represent cells of the epiblast surrounding the primitive streak as opposed to migrating mesoderm. Approximately 13% of Id1-Venus-expressing cells co-expressed T, compared with approximately 14% of all epiblast cells. This suggests once more that Id1-Venus expression is not preferentially associated with the expression of T, and suggests that Id1 expression is not restricted to the primitive streak but is rather widely expressed throughout the proximal epiblast.

5.3.3 Id1 expression in early bud stage epiblast

I stained an early bud stage chimaera generated from Id1-Venus non-lineage labelled ES cells with antibodies against GFP, Nanog and T, and imaged it both as an intact embryo and after sectioning (Figure 5.4). I showed that Id1-Venus was co-expressed in a proportion of Nanog- and T-expressing cells at this stage, and that it was no longer confined to the epiblast, but also expressed in the migrating mesoderm and endoderm, an observation that had not been clarified by the *in situ* hybridisation images of Jen *et al.* (1997) and Li *et al.* (2013). The expression of Id1-Venus declined from proximal to distal epiblast, confirming published observations. This stronger proximal expression is in accordance to the expression of *Bmp4* at this stage (Du *et al.* 2010, García-García & Anderson 2003, Madabhushi & Lacy 2011, Perea-Gómez *et al.* 1999, Tremblay *et al.* 2001). Id1-Venus expression in the migrating cells may reflect the phosphorylation of Smad1 in the proximal epiblast and visceral endoderm of early bud stage embryos (Di-Gregorio *et al.* 2007). It appears the expression of Id1-Venus, Nanog and T do not directly influence each other: Id1-Venus is expressed proximally and in mesendoderm, Nanog is expressed in the posterior half of the embryo, and T is expressed in the primitive streak and notochord. There is no apparent rise or decline in levels of any of the three factors when the others are or are not expressed.

It was interesting to notice that Nanog protein appears to persist shortly after cells migrate through the primitive streak (Figure 5.4b), suggestive of a dismantling of the pluripotency network of the cells after their delamination from the epiblast, as opposed to prior to migration through the primitive streak.

5.3.4 Id1 expression in early headfold stage chimaeras

I imaged chimaeras generated from lineage-labelled IdVnK1 ES cells at the early headfold stage (Figure 5.5). The expression of Id1-Venus, which at the early bud

stage had been detected in vast regions of migrating mesoderm and epiblast, appears to have been once again restricted to the proximal regions of the embryo, both in the anterior and in the posterior. At this stage, *Bmp4* is expressed in the extraembryonic regions adjacent to proximal posterior ectoderm, as well as in the surface ectoderm surrounding the neural folds (Perea-Gómez *et al.* 1999, Tremblay *et al.* 2001, Zakin & De Robertis 2004), so the pattern of expression of Id1-Venus fits the expression pattern of the morphogen.

5.3.5 Id1 expression in late headfold stage chimaeras

I imaged chimaeras generated from both lineage-labelled IdVnK1 ES cells and non-lineage labelled IdV ES cells at the late headfold stage (Figure 5.6). The expression of Id1-Venus was once again restricted to defined embryonic regions, the fluorescent reporter allele being expressed in cardiac crescent, lateral and allantoic mesoderm, and in the endoderm of the foregut primordium. The expression in the cardiac crescent and lateral mesoderm can also be observed in the *in situ* hybridisation experiment of Li *et al.* (2013). The expression pattern of Id1-Venus directly reflects that of *Bmp2*, *Bmp4*, *Bmp5*, *Bmp7* and *Bmpr1a* at this stage (Arkell & Beddington 1997, Danesh *et al.* 2009, Madabhushi & Lacy 2011, Mishina *et al.* 1995, Solloway & Robertson 1999, Zakin & De Robertis 2004). Whilst mice lacking the BMP2/4 receptor Bmpr1a exhibit premature neural differentiation and fail to express markers of the primitive streak (Mishina *et al.* 1995, Di-Gregorio *et al.* 2007), mice in which the receptor is inactivated in a mosaic pattern within the epiblast progress passed gastrulation. These embryos exhibit defects in the formation of the allantois and of the cardiac crescent (Miura *et al.* 2006), suggesting BMP signalling, potentially through the action of Id1, is required autonomously for the establishment of these structures.

Id1-Venus appears to be absent from the *Sox1*-expressing neural plate (Cajal *et al.* 2012, Perea-Gómez *et al.* 1999) at this stage, an observation which is interesting in the light of the anti-neural effect of Id1 in ES cell differentiation. *Noggin*, a BMP signalling inhibitor, is expressed in this tissue, and could thus explain the lack of Id1-Venus expression (Danesh *et al.* 2009, Mine *et al.* 2008).

The most intriguing of discoveries was the expression of Id1-Venus in a ring of cells surrounding the node in many of the headfold stage embryos, an unexpected observation that has not been described in the literature. A similar ring of expression was previously reported for *Nodal* (Conlon *et al.* 1994, Nakaya *et al.* 2005, Norris & Robertson 1999), so a straightforward hypothesis would be that of Nodal signalling driving expression of *Id1* in this region. This, however, is an unlikely possibility. I have presented data in Chapter 3 showing that Activin/Nodal signalling can repress the expression of Id1-Venus in pluripotent cultured cells. Mutual antagonism between Nodal and BMP signalling have also been reported at this stage of development (Mine *et al.* 2008, Pereira *et al.* 2012, Yang *et al.* 2010), as well as in ES cells and other organisms (Dale *et al.* 1992, Galvin *et al.* 2010, Veerkamp *et al.* 2013, Yamamoto *et al.* 2009). It is therefore unlikely that Nodal sustains the expression of Id1.

Bmp2 is expressed at the node at the early bud stage, but it disappears by the late bud stage (Madabhushi & Lacy 2011), and is therefore unlikely to be capable of maintaining Id1-Venus expression at the node until the late headfold stage.

Bmp7 is expressed in cardiac crescent and allantoic mesoderm, but also in the node (Arkell & Beddington 1997, Danesh *et al.* 2009). It appears therefore the most likely candidate for the promotion of Id1-Venus expression at the node. However, the BMP inhibitors *Chordin* and *Noggin* are also expressed at the node in late headfold stage embryos (Danesh *et al.* 2009, McMahon *et al.* 1998, Mine *et al.* 2008, Rhinn *et al.* 1998), where they play a crucial role in establishment of this structure: *Nodal* is not expressed at the node in *Chordin*$^{-/-}$ *Nodal*$^{+/-}$ embryos (Yang *et al.* 2010). Noggin inhibits BMP signalling by directly binding to BMP molecules, and has strong

affinity for BMP2 and BMP4. It also binds BMP7, but with lower affinity (Chang & Hemmati-Brivanlou 1999, Groppe *et al.* 2002, Zhu *et al.* 2006, Zimmerman *et al.* 1996). It is therefore possible that if the levels of BMP7 were high enough they could overcome the inhibition by Noggin and promote Id1 expression at the node. Should this be the case, it is unclear why Id1 expression at the node is not detectable in earlier embryos, since *Bmp7* is expressed at the node from no later than late streak stage (Arkell & Beddington 1997). It is possible that the larger domain of expression of *Nodal* at these earlier stages (Di-Gregorio *et al.* 1997, Du *et al.* 2010, Faust *et al.* 1995, Norris *et al.* 2002, Perea-Gomez *et al.* 2001) is epistatically suppressing the potential BMP7-driven induction of Id1.

It would appear that the expression of Id1-Venus at the node might be controlled by the integration of stimuli from a multitude of positive and negative regulators. Whether Id1 is playing an active role in the processes occurring around the node at this stage, from primitive streak elongation to left-right patterning, can only be the subject of speculation at this point. Its involvement in many lineage specification events and in many self-renewal versus differentiation decisions throughout development would make the HLH factor a fitting candidate for the crucial processes taking place at the distal tip of the embryo at this early developmental stage.

FIGURE 5.1

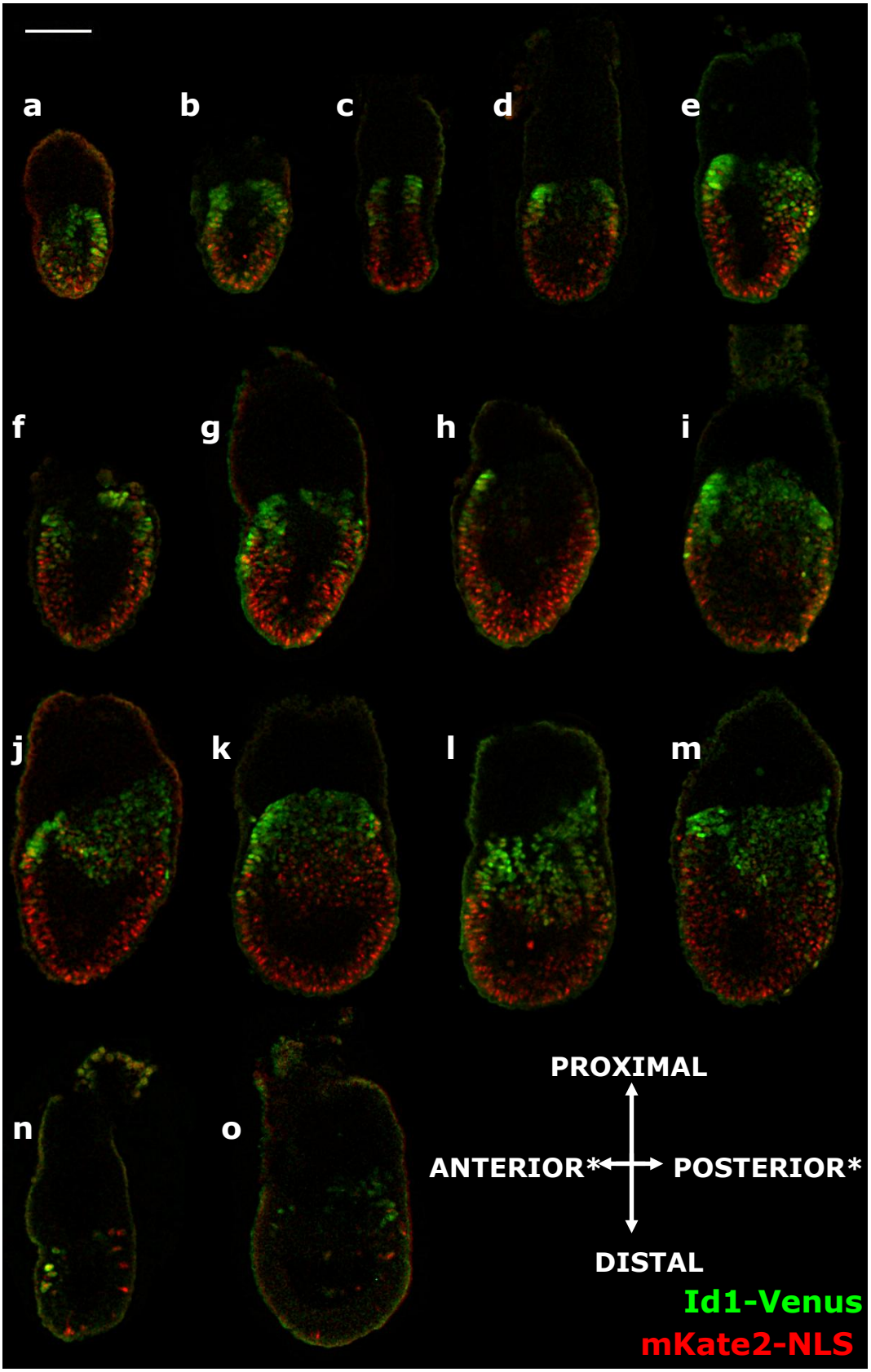

a
b
c
d
e
f
g
h
i
j
k
l
m
n
o
PROXIMAL
ANTERIOR*
POSTERIOR*
DISTAL
Id1-Venus
mKate2-NLS

Figure 5.1 – Id1-Venus expression at E6.5

Chimaeras generated from morula aggregation of IdV ES cells stably expressing a *CAG-mKate2-NLS-IRES-Pac* transgene (IdVnK1). The chimaeras were dissected 4 days after transferring the overnight cultured embryos to pseudopregnant females, at an approximate equivalent developmental stage to wild-type E6.5 embryos. All 15 analysed embryos are displayed in figure.

The embryos are ordered from smallest to biggest. The embryos on the last row are low contribution chimaeras. The proximal/distal and anterior/posterior embryonic axes are displayed. Embryos a, b, c, and d lacked an evident thickening of visceral endoderm on one side of the embryo, so the anterior/posterior axis may not be aligned from the left to the right of the image.

The images were acquired as a single focal plane on a Leica TCS SPE inverted confocal microscope. The fluorescence intensity was optimised singularly for each embryo, so the intensity of Id1-Venus and mKate2-NLS expression between embryos is not comparable. Scale bar: 100μm.

FIGURE 5.2

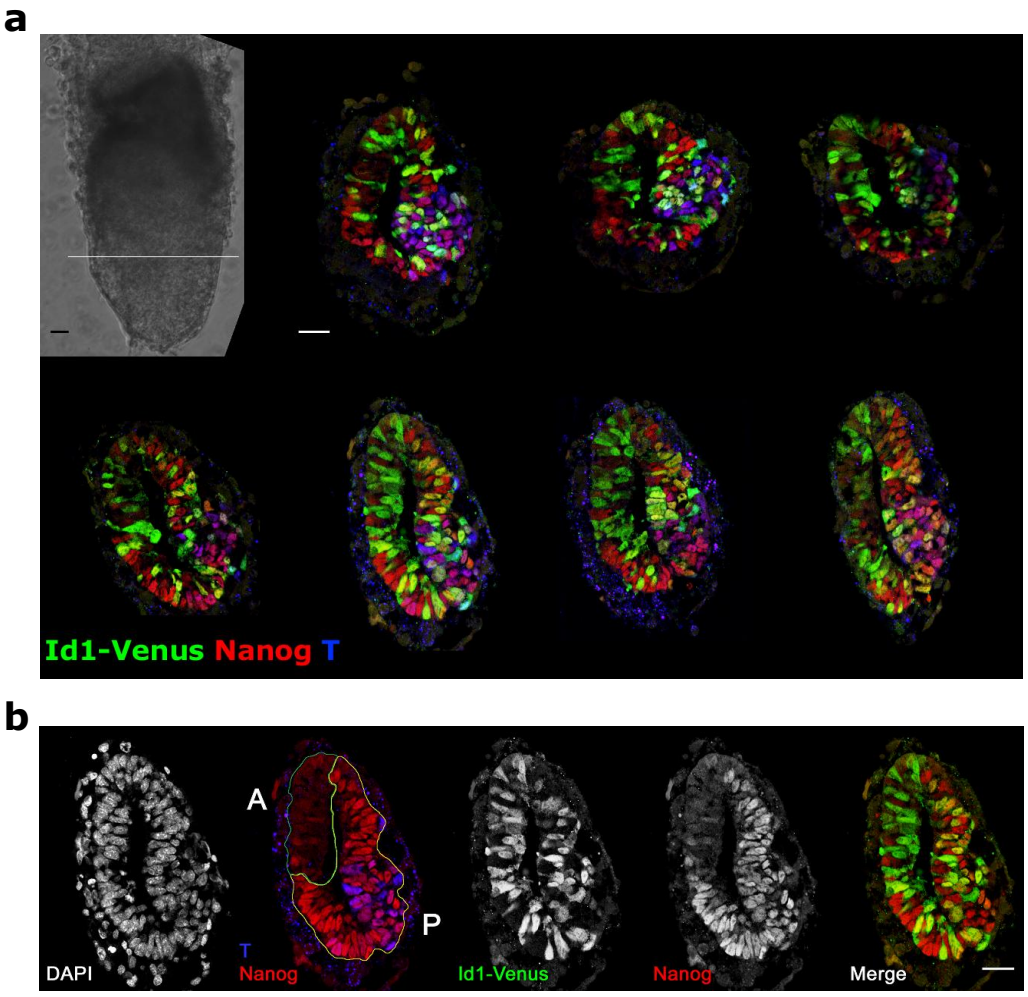


Figure 5.2 – Id1-Venus is coexpressed with Nanog and T in the proximal post-implantation epiblast

a. Mid-streak stage chimaera generated from morula aggregation of IdV ES cells. The embryo was sectioned transversally and stained with antibodies against GFP, Nanog and T. The nuclei were counterstained with DAPI. The sections are ordered from most proximal to most distal. The sections were imaged on a Leica TCS SPE inverted confocal microscope. Scale bar: 30µm.

b. Definition of anterior and posterior regions of interest for nuclear immunostaining quantification based on Nanog expression. Visceral endoderm was excluded from the regions of interest. A: anterior, P: posterior. Scale bar: 30µm.

FIGURE 5.3

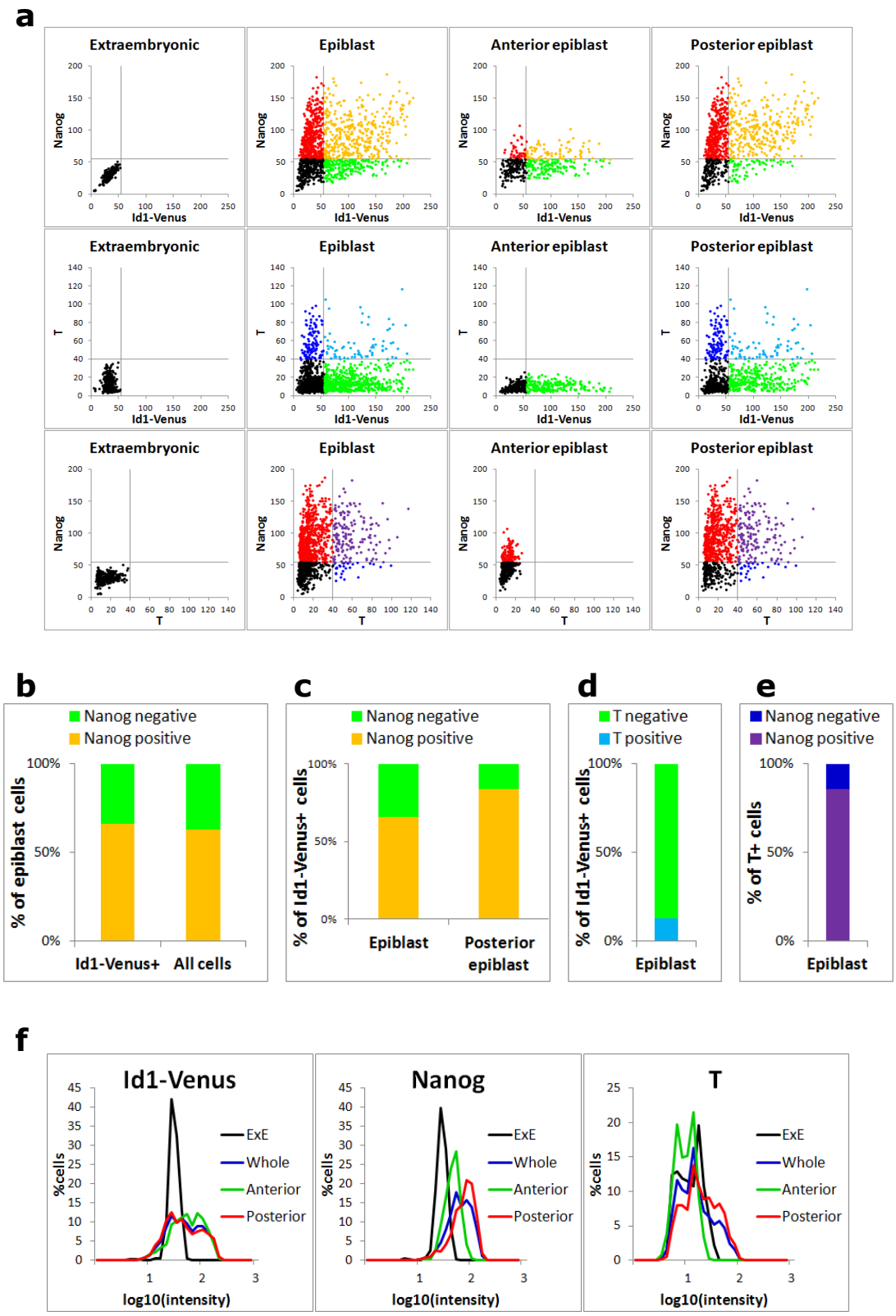

a
Extraembryonic
Epiblast
Anterior epiblast
Posterior epiblast
Nanog
T
Id1-Venus
b
Nanog negative
Nanog positive
% of epiblast cells
Id1-Venus+
All cells
c
% of Id1-Venus+ cells
Epiblast
Posterior epiblast
d
T negative
T positive
e
% of T+ cells
f
Id1-Venus
Nanog
T
%cells
log10(intensity)
ExE
Whole
Anterior
Posterior

Figure 5.3 – Immunostaining quantification of Id1-Venus, Nanog, T in E6.5 chimaera sections

a. Dot plots illustrating the levels of Id1-Venus, Nanog and T in the nuclei of single cells of whole, anterior and posterior regions of the proximal epiblast, in the sections shown in Fig.5.2. Stained sections of extraembryonic tissue were used to set expression gates. Number of nuclei scored: whole proximal epiblast: 1265, anterior epiblast: 395, posterior epiblast: 870, extraembryonic tissue: 363.

b. Quantification of Nanog positive cells amongst Id1-Venus positive cells and all cells in the proximal epiblast.

c. Quantification of Nanog positive cells amongst Id1-Venus positive cells in the whole and in the posterior proximal epiblast.

d. Quantification of T positive cells amongst Id1-Venus positive cells in the proximal epiblast.

e. Quantification of Nanog positive cells amongst T positive cells in the proximal epiblast.

f. Flow cytometry-like plots illustrating the intensity of Id1-Venus, Nanog and T in extraembryonic tissue (ExE), and in whole, anterior and posterior regions of the proximal epiblast.

FIGURE 5.4

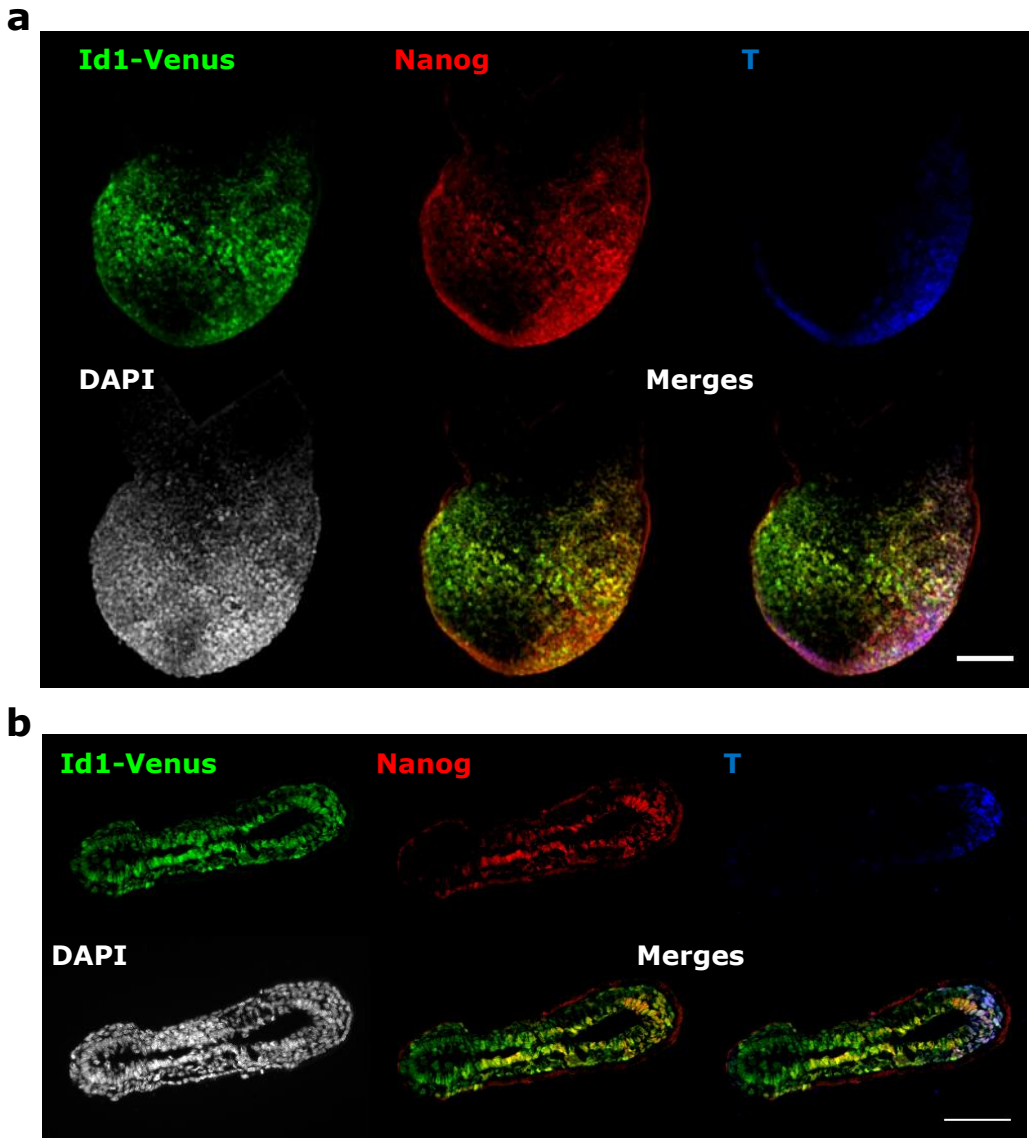

a
Id1-Venus
Nanog
T
DAPI
Merges
b
Id1-Venus
Nanog
T
DAPI
Merges

Figure 5.4 – Id1-Venus is coexpressed with Nanog and T in the early bud stage epiblast

a. Three-dimensional rendering of an early bud stage chimaera generated from morula aggregation of IdV ES cells, stained with anti-GFP, anti-Nanog and anti-T antibodies. The embryo was imaged on a Leica TCS SPE inverted confocal microscope. The anterior side of the embryo is on the left of the image. Scale bar: 100µm. The same Id1-Venus expression was observed in 3 IdV chimaeras and 6 IdVnK1 chimaeras.

b. Transverse section through the embryo in (a). The anterior side of the embryo in on the left of the image. The section was imaged on an Olympus IX51 microscope. Scale bar: 100µm.

FIGURE 5.5

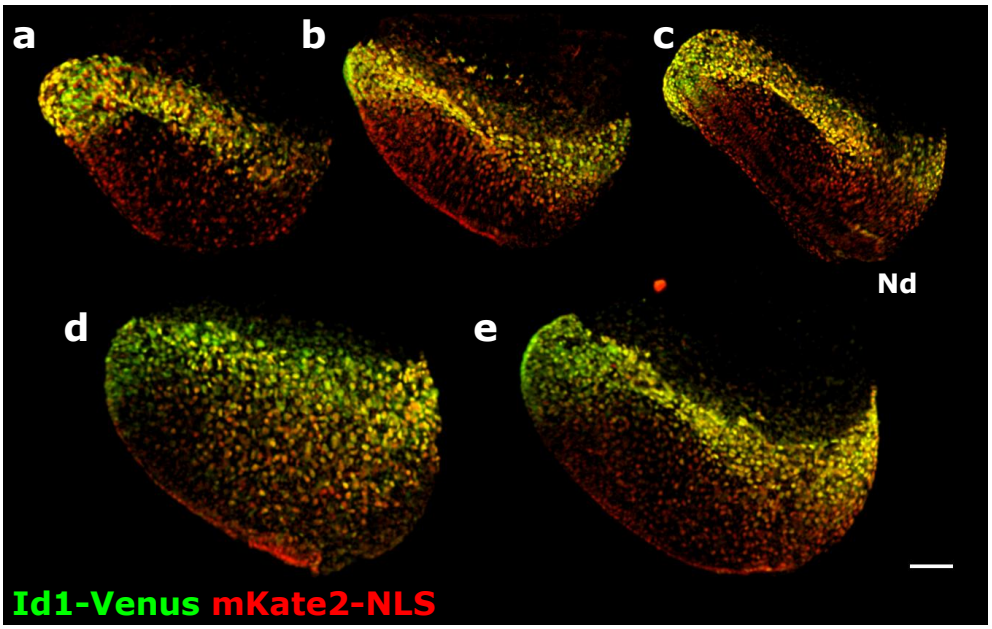


Figure 5.5 - Id1-Venus expression at the headfold stage

Three-dimensional rendering of chimaeras generated from morula aggregation of IdVnK1 ES cells, dissected 5 days after blastocyst transfer to pseudopregnant females, the approximate equivalent of wild-type E7.5 embryos, and stained with an anti-GFP antibody. All 5 analysed embryos are displayed in figure. The embryos are ordered from smallest to biggest and the anterior side is aligned on the left of the image. Embryo (c) is slightly tilted so that the node (Nd) is visible. The embryos were imaged on a Leica TCS SPE confocal microscope. Scale bar: 100µm.

FIGURE 5.6

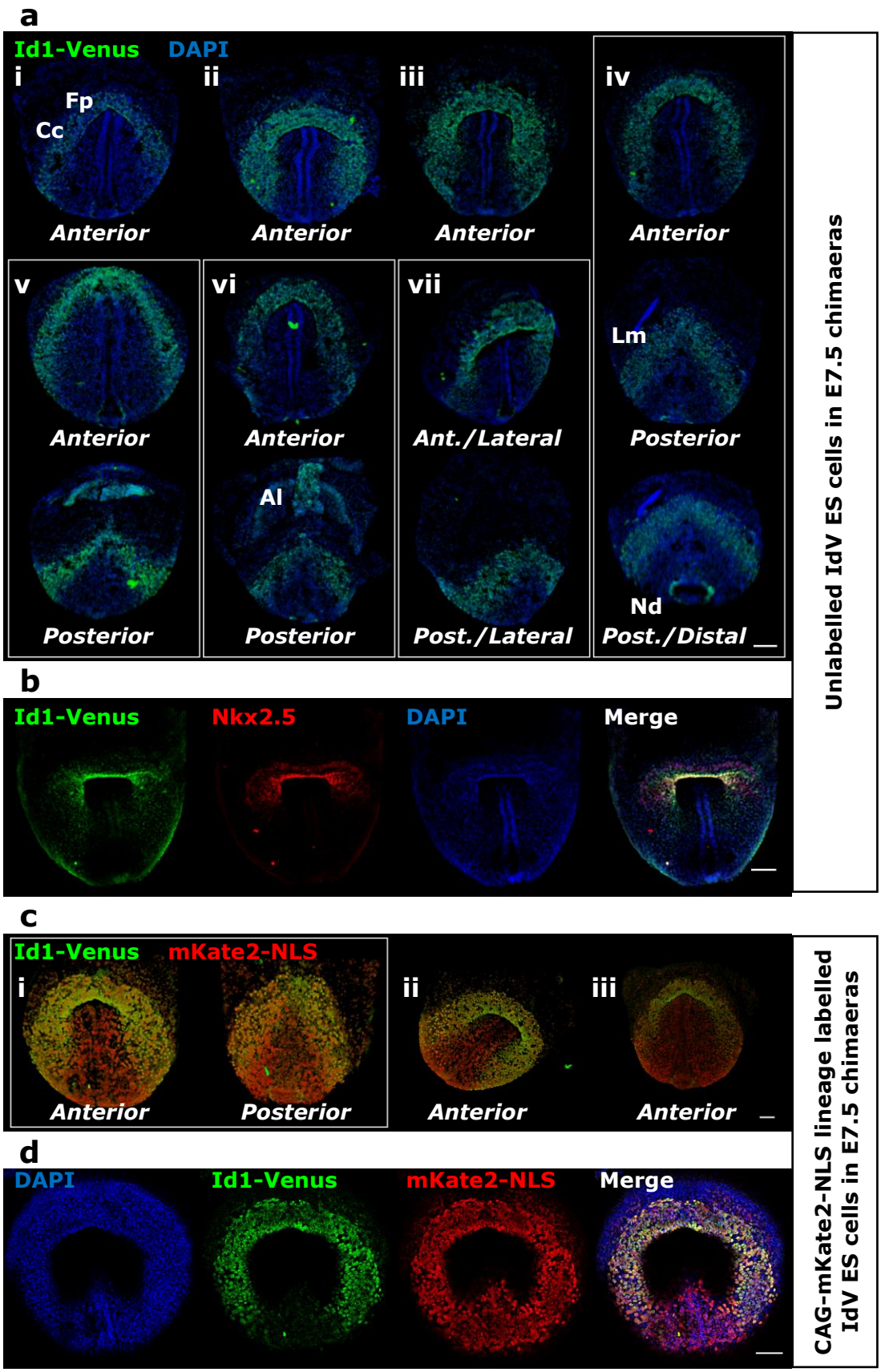

a
Id1-Venus DAPI
i
Fp
Cc
Anterior
ii
Anterior
iii
Anterior
iv
Anterior
Lm
Posterior
Nd
Post./Distal
v
Anterior
Posterior
vi
Anterior
Al
Posterior
vii
Ant./Lateral
Post./Lateral
b
Id1-Venus
Nkx2.5
DAPI
Merge
Unlabelled IdV ES cells in E7.5 chimaeras
c
Id1-Venus mKate2-NLS
i
Anterior
Posterior
ii
Anterior
iii
Anterior
d
DAPI
Id1-Venus
mKate2-NLS
Merge
CAG-mKate2-NLS lineage labelled IdV ES cells in E7.5 chimaeras

Figure 5.6 - Id1-Venus expression at the cardiac crescent stage

a. Three-dimensional rendering of chimaeras generated from morula aggregation of IdV ES cells, dissected at embryonic day 7.5 and stained with an anti-GFP antibody. Images within a white box are different views of the same embryo. The depicted side of the embryo is described under each image. Embryos are numbered in Roman numerals for ease of identification. Fg: foregut primordium, Cc: cardiac crescent, Lm: lateral mesoderm, Al: allantois, Nd: node. Scale bar: 100µm.

b. Three-dimensional rendering of the anterior view of a chimaera generated from morula aggregation of IdV ES cells, dissected at embryonic day 7.5 and stained with anti-GFP and anti-Nkx2.5 antibodies. Scale bar: 100µm.

c. Three-dimensional rendering of chimaeras generated from morula aggregation of IdVnK1 ES cells, dissected at embryonic day 7.5 and stained with anti-GFP and anti-tagRFP antibodies. The images within the white box are anterior and posterior views of the same embryo, the other two embryos are viewed from the anterior. Embryos are numbered in Roman numerals for ease of identification. Scale bar: 100µm.

d. Single focal plane of embryo “i” in (c). Scale bar: 100µm.

All embryos were imaged on a Leica TCS SPE inverted confocal microscope. All 11 embryos analysed at this stage are displayed in figure.

CHAPTER 6

General discussion

6.1 The regulation of Id1 expression

6.1.1 Id1 expression is regulated by a multitude of inputs

The data presented in this thesis indicate that the regulation of Id1 expression is a point of convergence of many signalling pathways involved in the regulation of pluripotency and differentiation (Figure 6.1).

Three major signalling pathways appear to converge on the expression of this factor: BMP, Activin/Nodal/TGFβ and mitogen-activated protein kinase.

6.1.1.1 Bone morphogenic protein

BMP is a positive regulator of *Id1* expression. Various BMP family members have been reported to induce *Id1* transcription in numerous cell types from different species (Galvin *et al.* 2010, Hollnagel *et al.* 1999, Kang *et al.* 2003, Katagiri *et al.* 1994, Katagiri *et al.* 2002, Korchynskyi & ten Dijke 2002, López-Rovira *et al.* 2002, Ogata *et al.* 1993, Ying *et al.* 2003a). My experiments confirmed that the stimulation of Id1 expression by BMP occurs in mouse ES cells cultured in LIF+FCS and 2i medium, as well as in epiblast stem cells. Furthermore, the expression pattern of Id1 *in vivo* was shown to recapitulate that of BMP signals present in the developing embryo.

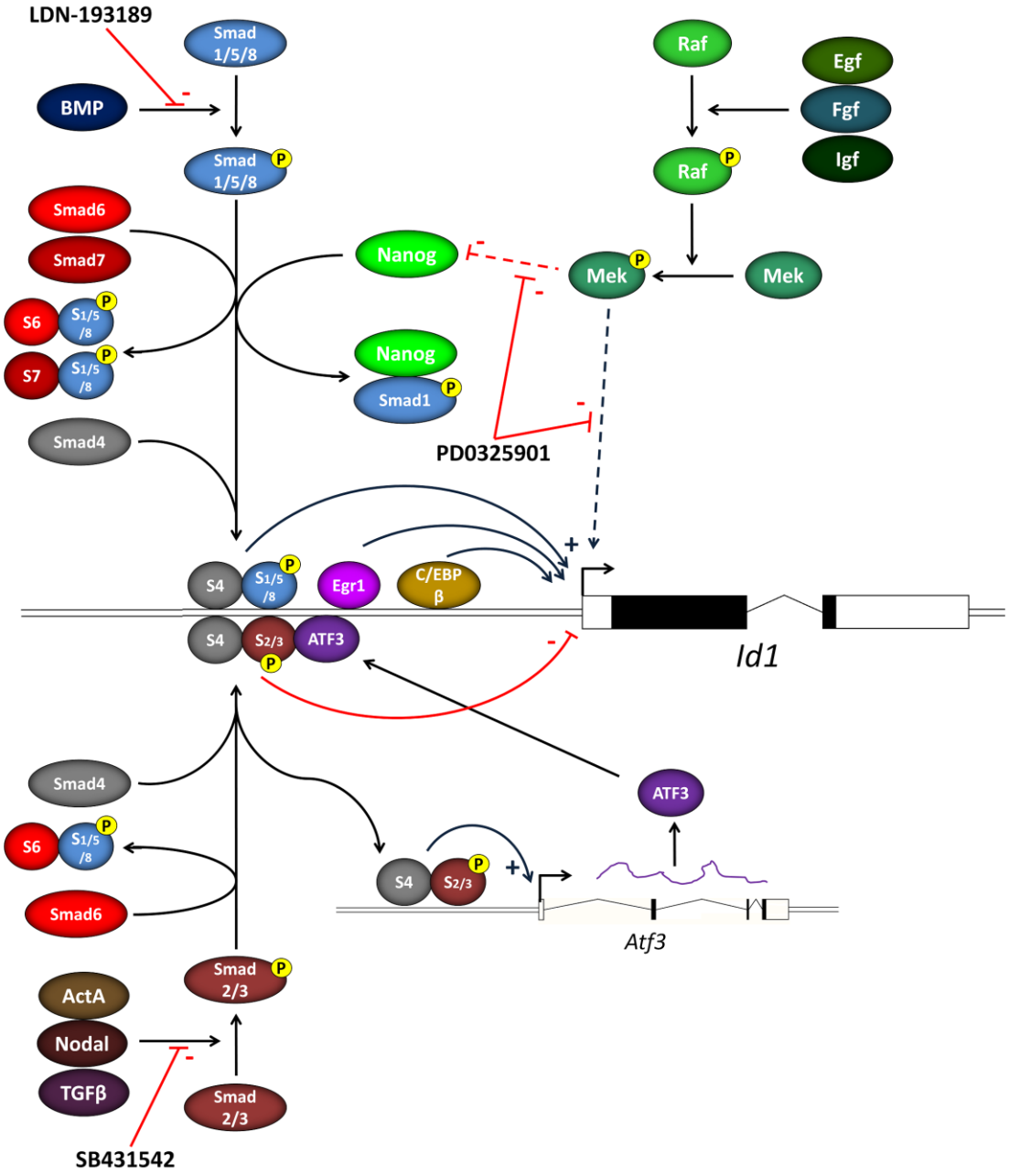

LDN-193189
Smad 1/5/8
BMP
Smad 1/5/8
Smad6
Smad7
S6
S1/5/8
S7
S1/5/8
Smad4
Nanog
Nanog
Smad1
PD0325901
Raf
Egf
Fgf
Igf
Raf
Mek
Mek
S4
S1/5/8
Egr1
C/EBP β
S4
S2/3
ATF3
Id1
Smad4
S6
S1/5/8
Smad6
ActA
Nodal
TGFβ
Smad 2/3
Smad 2/3
SB431542
ATF3
S4
S2/3
Atf3

Figure 6.1 – Regulation of *Id1* expression

Diagram illustrating the molecules involved in the regulation of *Id1* expression. BMP signalling induces *Id1* via phosphorylated Smad1/5/8-Smad4 complexes. Activin/Nodal/TGFβ signalling represses *Id1* via phosphorylated Smad2/3. The inhibition of *Id1* may involve DNA association in complex with Smad4, and in human cell lines it is achieved through the formation of a complex with the transcriptional repressor ATF3, a direct target of the Smad2/3-Smad4 complex. Smad6 can inhibit both BMP and Activin/Nodal/TGFβ signalling through sequestration of Smad1/5/8 and Smad2/3, Smad7 can inhibit BMP signalling through the same mechanism. Induction of the mitogen-activated protein kinase pathway results in Mek phosphorylation, which leads to the repression of Nanog. Nanog can interact with Smad1 and prevent its activation of *Id1* transcription. It is unclear whether Nanog can bind inactive, active or both forms of Smad1. Mek activity may also induce *Id1* transcription through alternative mechanisms. A complex involving Egr1 is responsible for *Id1* induction in response to serum. C/EBPβ can induce *Id1* expression in pro-B-cells. The depicted binding of factors to the promoter of Id1 is not representative of the actual position of their binding sites. The processes inhibited by the small molecule inhibitors LDN-193189, PD0325901 and SB431542 are also illustrated in figure. Abbreviations: ActA: Activin A, S: Smad.

6.1.1.2 Activin/Nodal/TGFβ

Activin/Nodal/TGFβ signalling results in repression of *Id1* transcription in mouse and human cells (Galvin *et al.* 2010, Kang *et al.* 2003, Lee *et al.* 2011b, Veerkamp *et al.* 2013). In human cells lines this repression is achieved through the interaction of SMAD3 with the SMAD2/3 target ATF3 (Kang *et al.* 2003). TGFβ has been reported to induce *ID1* transcription in human breast cancer cells through the action of SMAD3/4 and the CBP/p300 co-activator, suggesting that under certain conditions SMAD3 may change its interaction partner from ATF3 to CBP/p300 and activate *ID1* expression (Padua *et al.* 2008, Stankic *et al.* 2013). It is unclear whether this mechanism is conserved in healthy cells or is an artefact of cancer cell lines. In my experiments, the small molecule SB431542, an inhibitor of the Activin/Nodal/TGFβ receptors Acvr1b, Acvr1c and Tgfbr1, resulted in upregulation of *Id1* expression in pluripotent cultures, suggesting Activin/Nodal/TGFβ exerts a repressive role on *Id1* transcription in these cultures.

Interestingly, whilst SB431542 treatment of mouse ES cells results in a reduction of Smad2 phosphorylation (Galvin *et al.* 2010), it also leads to a relocalisation of Smad2 protein on the *Id1* locus, and stronger binding in the region of the promoter containing the Smad-responsive elements as well as in the *Id1* gene body (Lee *et al.* 2011b). This suggests that Smad2 may not be involved in the repression of *Id1* transcription in mouse ES cells, and that it may rather act as an activator; that Smad2 interaction with other regions of the locus may be responsible for its inhibitory activity; or that SB431542 treatment of ES cells may result in the downregulation of other factors required for the Smad2-mediated repression of *Id1* transcription. *Atf3* expression is not affected by SB431542 treatment in the microarray data of Lee *et al.* (2011b), but ATF3 was not identified as an interactor of SMAD2 in the study by Kang *et al.* (2003), suggesting Smad2 repression of *Id1* may occur through interaction with alternative factors.

SB431542 treatment of mouse ES cells results in an increase in Smad1/5 phosphorylation, that can be rescued by supplementation with the BMP signalling inhibitor LDN-193189 (Galvin *et al.* 2010), suggesting the SB431542 induction of *Id1* may also occur through the stimulation of BMP signalling, potentially by relieving the inhibitory cross-talk of Activin/Nodal/TGFβ signals on the BMP pathway. Whilst simultaneous addition of LDN-193189 and SB431542 to ES cells suppresses Smad1/5 phosphorylation (Galvin *et al.* 2010), my flow cytometry experiments have shown that this combination of molecules does not suppress the expression of Id1 in LIF+FCS culture (Figure 3.8c,d), implying that the inhibitory activity of SB431542 on the Activin/Nodal TGFβ pathway plays an important role in the induction of *Id1* expression.

6.1.1.3 Mitogen activated protein kinase

Inhibition of the mitogen-activated protein kinase transducer Mek with the small molecule PD0325901 results in the inhibition of *Id1* expression. This may be an indirect effect resulting from the upregulation of Nanog in cells treated with PD0325901, as Nanog has been characterised as a negative regulator of *Id1* transcription. Nanog can in fact bind and inhibit the transcriptional activity of Smad1 (Suzuki *et al.* 2006). It is however possible that Mek may induce *Id1* expression through alternative mechanisms. Although PD0352901 has been reported to decrease the phosphorylation of Smad1/5 and Smad2 (Galvin *et al.* 2010), its effects appears to be independent of BMP and Activin/Nodal/TGFβ signalling, as BMP stimulation promotes Id1 expression in the presence of the inhibitor, and both LDN-193189 and Activin A promote further downregulation of Id1 in PD0325901-treated cells.

The use of *Nanog*-null ES cells would allow us to address the mechanism of action of PD0325901 in the downregulation of *Id1* expression: should PD0325901 fail to inhibit *Id1* expression in these cells, we would be able to conclude that Nanog is the key mediator of the negative role of Mek inhibition on *Id1* expression.

Supplementation of ES cells with Fgf2 did not result in the upregulation of Id1. This is puzzling in the light of the strong effect of Mek inhibition on *Id1* expression. Possible explanations of this phenomenon are that Fgf signals may already be present at high levels in ES cell cultures, that Mek inhibition of Nanog expression cannot be increased by further stimulation of Mek phosphorylation, or that other mitogens are required for Mek induction in these cultures.

6.1.1.4 A direct role for Nanog in Id1 repression

The negative input of Nanog on *Id1* expression may not be limited to its interaction with Smad1. Festuccia *et al.* (2012) performed a microarray on *Nanog*-null ES cells following tamoxifen-induced nuclear translocation of a transgenic Nanog-ERT2 fusion protein. Analysis of this dataset reveals that *Id1* is the transcript displaying the strongest level of downregulation following the nuclear translocation of Nanog. Smad1 is phosphorylated in the cytoplasm before it enters the nucleus and drives transcription of *Id1*. If Nanog inhibited Id1 expression only through its association with Smad1, it should result in Id1 inhibition when forced to localise to the cytoplasm. The observation that it promotes a rapid and strong downregulation of *Id1* transcription upon nuclear localisation suggests it can inhibit Id1 expression in further ways, potentially through the direct binding and repression of the *Id1* locus. Analysis of the ChIP-Seq data of Chen *et al.* (2008), Marson *et al.* (2008) and Whyte *et al.* (2013) reveals the potential presence of Nanog binding sites at the *Id1* locus (Figure 6.2), confirming that Nanog repression of Id1 expression may occur through both direct and indirect mechanisms.

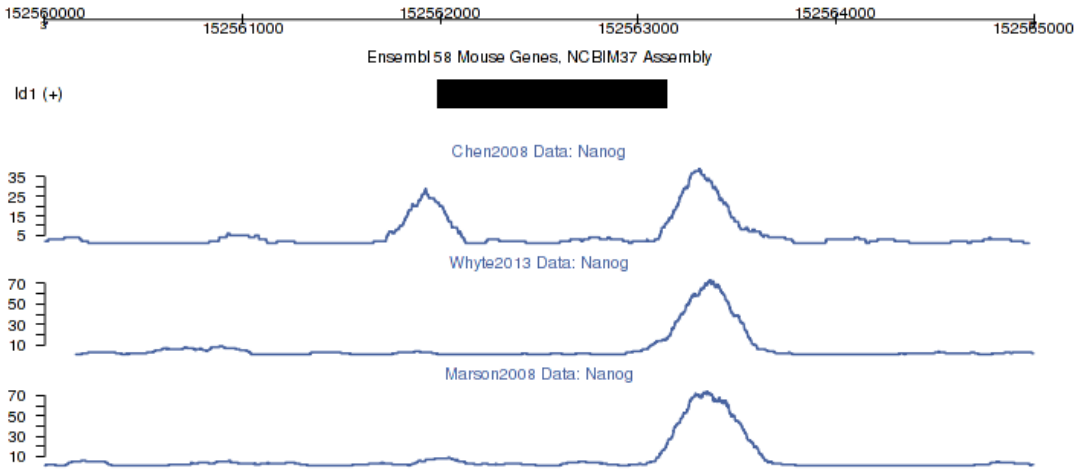


Figure 6.2 – ChIP-Seq identifies Nanog binding peaks in the *Id1* locus

Analysis of the Nanog ChIP-Seq data of Chen *et al.* (2008), Marson *et al.* (2008) and Whyte *et al.* (2013) with the GeneProf online resource (Halbritter *et al.* 2012) reveals peaks of Nanog binding upstream and downstream of the *Id1* exonic DNA (black box).

6.1.2 Id1 is expressed heterogeneously in pluripotent cells

The integration of multiple positive and negative signals resulted in a heterogeneous expression pattern for Id1 in LIF+FCS and epiblast stem cell cultures. In EpiSCs, the heterogeneity in Id1 expression became even more evident upon stimulation with BMP4, with only a proportion of cells capable of responding to the signal by upregulating Id1. The resolution of this bimodal distribution through the inhibition of Activin/Nodal signalling suggests that agonists of this pathway can inhibit the capability of pluripotent cells to respond to BMP stimulation, and suggests that expression of Id1 in pluripotent cultures reflects the balance of active Smad1/5/8 and Smad2/3 in individual cells. This raises an important question. How can cells

cultured as an adherent monolayer respond differently to uniform signals provided in the culture medium?

I believe the answer to lie in the cross-repressive action of Activin/Nodal/TGFβ and BMP, and in the heterogeneity in expression of other factors independent of these uniform signals. Nanog is expressed heterogeneously in LIF+FCS culture, and this heterogeneity is intrinsic to the culture system. Purification of cells expressing high levels of Nanog will result in the restoration of the original distribution of Nanog expression after approximately one week of culture (Chambers *et al.* 2007). Nanog can repress its own expression (Fidalgo *et al.* 2012, Navarro *et al.* 2012), implying that cells with high levels of Nanog protein will downregulate *Nanog* transcription and become cells with low levels of Nanog protein over time. The levels of *Nanog* transcript and protein in ES cells are not regulated by Nanog alone, which implies that the oscillatory mechanism described above is over-simplified, but it is a useful model nonetheless, as it allows us to understand how cells may establish heterogeneity in expression in a uniform culture system.

Nanog is a negative regulator of Smad1 activity, it may act as a direct repressor of *Id1* transcription (Figure 6.2), and I have shown that ES cells are rarely observed in a status of high Nanog and high Id1 expression. In human ES cells, the transcription of *NANOG* is induced by SMAD2/3, and NANOG can bind and modulate the activity of SMAD2/3 (Vallier *et al.* 2009, Xu *et al.* 2008). BMP-induced and Activin/Nodal/TGFβ-induced Smad's have been reported to inhibit the activity of each other (Yamamoto *et al.* 2009), and Smad2 and Smad3 can also antagonise one another (Míguez *et al.* 2013). This results in the establishment of a complex Smad1/5/8-Smad2/3-Nanog network in which every factor has the potential to influence the expression or the activity of every other factor, and where one of the components (Nanog) is heterogeneously expressed. It is therefore straightforward to hypothesise that the intracellular balance of the repressive activities of Smad2/3 and Nanog and the inductive action of Smad1/5/8 on *Id1* expression may vary between

different cells due to variable levels of Nanog and to the consequent variation in levels and activity phosphorylated Smad's.

In light of these considerations, it is encouraging to observe that in 2i medium Nanog is expressed homogeneously at a high level, and Id1 is not expressed. BMP and SB431542 can induce expression of Id1, but this inductive potential is repressed by Activin A and LDN-193189 respectively, implying that the balance of positive over negative regulation of Id1 is skewed in favour of the negative regulators.

Future studies on the regulation of Id1 expression in pluripotent cultures would gain considerable insight with the use of *Nanog*-null ES cells. These cells would allow us to test whether Id1 is expressed in a more homogeneous pattern in LIF+FCS and epiblast stem cell culture, and whether Id1 expression is de-repressed in 2i culture. They could be used to address whether simultaneous stimulation of BMP and inhibition of Activin/Nodal can result in a homogeneous state of high Id1 expression, as opposed to the increased Id1 expression and unaffected heterogeneity observed in wild-type cells. It would also be interesting to test whether full repression of Id1 can be achieved in the absence of Nanog, or whether this factor is essential for a full transcriptional shutdown of the *Id1* locus.

6.1.3 The regulation of Id1 expression in vivo

The expression of Id1 throughout development mirrors that of BMP signals. In small pre-streak stage embryos, I detected Id1 expression both proximally and in some cells in distal regions, probably as a result of the short distance between the distal portions of the embryos and the BMP source in extraembryonic ectoderm (Coucouvanis & Martin 1999). Id1 later remained expressed proximally and in migrating mesoderm, and at late headfold stage was found in the cardiac crescent, lateral mesoderm and allantois, close to sources of BMP (Perea-Gómez *et al.* 1999,

Tremblay *et al.* 2001, Zakin & De Robertis 2004). A few interesting points were raised by the expression pattern I characterised.

6.1.3.1 The role of Nodal in Id1 expression in vivo

Id1 was found to be expressed by cells in proximal and distal regions in small early post-implantation embryos (Figure 5.1a,b). At this stage, Nodal is expressed throughout the egg cylinder (Varlet *et al.* 1997), and it is surprising that it does not suppress the expression of Id1. It therefore likely that BMP stimulation of epiblast cells has a dominant effect over Nodal activity at this stage of development.

The embryo later expands, distancing the source of BMP from the distal tip. The DVE and the AVE then secrete BMP inhibitors, suppressing Id1 expression in the overlying epiblast. Nodal localises to the primitive streak and the node secretes BMP inhibitors (Arnold & Robertson 1999, Varlet *et al.* 1997). Whilst Id1 expression is still observed in the proximal posterior epiblast, suggesting BMP signalling is still dominant over repression of Id1 by Nodal, the highest levels of Id1 can be detected in the proximal anterior regions (Figure 5.1d-m), fated for surface ectoderm specification (Cajal *et al.* 2012, Madabhushi & Lacy 2011, Tam & Behringer 1997). This observation is confirmed by the *in situ* hybridisation data of Li *et al.* (2013), and suggests that Nodal acts as a repressor of *Id1* expression *in vivo*.

6.1.3.2 Id1 expression at the node in headfold stage embryos

I identified a previously uncharacterised expression pattern for Id1 in a ring of cells surrounding the node (Figures 5.5, 5.6). This expression was unexpected due to the expression of BMP inhibitors in this region and the expression of Nodal in a similar ring of cells (Conlon *et al.* 1994, Danesh *et al.* 2009, McMahon *et al.* 1998, Mine *et al.* 2008, Nakaya *et al.* 2005, Norris & Robertson 1999, Rhinn *et al.* 1998). BMP7 is

expressed at the node at this stage and may be responsible for this induction (Arkell & Beddington 1997, Danesh *et al.* 2009). It would be interesting to verify whether BMP and Nodal signalling are responsible for the establishment of this particular expression pattern. The dissection of embryos at E6.5/E7.0 followed by their culture in medium supplemented with agonists and antagonists of Nodal and BMP should allow to gain valuable insights into the roles of these signalling pathways in the establishment of this expression pattern.

6.1.3.3 The relationship between Nanog and Id1 in vivo

Nanog negatively regulates Id1 expression in pluripotent cultures, and the two factors are not co-expressed at high levels in ES cells. This negative relationship does not appear to be conserved in the post-implantation pluripotent epiblast. Id1 and Nanog are co-expressed in the proximal posterior epiblast, and the levels of Id1 do not appear to be affected by the expression of Nanog (Figures 5.2, 5.3, 5.4). It is unclear why this should be the case. One possibility is that BMP signalling induces the preferential phosphorylation of Smad5 and Smad8 over Smad1 in the proximal posterior epiblast, and the binding of Nanog to Smad1 is thus unable to exert is negative influence on *Id1* expression. If Nanog is able to directly bind to the *Id1* locus, as proposed in section 6.1.1.4, it is possible that the Nanog binding sites are inaccessible due to their occupancy by other factors or to epigenetic changes in the locus. A third possibility is that the levels of Nanog in these cells are low compared to the levels in ES cells, and are unable to overcome BMP stimulation of *Id1* transcription. The last possibility could be addressed by microdissection of proximal posterior epiblast followed by qRT-PCR comparing this sample to ES cells. The use of EpiSC lines to positively and negatively alter the levels of Nanog may provide some insight into whether Nanog is involved in the regulation of Id1 in the *in vitro* equivalent of post-implantation epiblast and into its role in relation to BMP signalling.

6.1.3.4 The expression of Id1 in pre-implantation embryos

Phosphorylated Smad1 and Smad2 have been detected in the blastocyst. In particular, phosphorylated Smad1 was found in a few cells, whereas phosphorylated Smad2 was detected in most if not all cells (James *et al.* 2005). The single-cell qRT-PCR experiments of Tang *et al.* (2010) suggest that *Id1* is expressed heterogeneously in the inner cell mass of E3.5 and E4.5 blastocysts, and this raises the question of whether Activin/Nodal/TGFβ and BMP are regulating the expression of the factor in the pre-implantation embryo. We are currently attempting to generate a mouse line from the Id1-Venus reporter ES cells, which will allow us to perform a detailed analysis of the expression pattern of Id1 in pre-implantation development and to measure the levels of Id1-Venus and phosphorylated Smad's in single cells of pre-implantation embryos. The culture of morula and blastocyst stage embryos in the presence of agonists and antagonists of BMP and Activin/Nodal/TGFβ will also aid in the understanding of the roles of these pathways in the regulation of *Id1* expression in pre-implantation development. *In situ* hybridisation experiments could also be used for the embryo culture experiments, but would not be amenable for the simultaneous measurement of the levels of Smad phosphorylation in Id1-expressing cells.

6.2 The roles of Id1 in lineage specification

6.2.1 Possible identity of the factors inhibited by Id1 in differentiation

In Chapter 4 I presented data indicating that Id1 delays the exit from a state of post-implantation epiblast pluripotency, and that it promotes the regionalisation of cells in culture to the equivalent of proximal posterior epiblast. Id1 is a dominant negative inhibitor of DNA binding, and its function is therefore dependent on the inductive or repressive transcriptional outputs of the factors it is inhibiting via direct binding or

sequestration of E proteins. What factor(s) is Id1 inhibiting during the differentiation process?

6.2.1.1 Neurod1, Tcf15, Twist1

A yeast two-hybrid screen was performed in our laboratory to identify the interacting partners of Id1 and E12/E47 in ES cells cultured in LIF+FCS. Id1 was found to interact with E12 and E47 only, whereas these proteins were found to interact with Id1, Id3 and the pro-differentiation bHLH factors Neurod1, Tcf15 and Twist1 (Davies *et al.* 2013). All three of these bHLH proteins have the capability to induce differentiation to a variable extent when episomally overexpressed in ES cell culture: Tcf15 drives the downregulation of *Nanog* and *Oct4* and is negatively associated with pluripotency markers (Davies *et al.* 2013), Twist1 drives an EMT event and the differentiation of ES cells into Vimentin-expressing mesenchyme when expressed in a fusion with E47 (but not as a monomer) (Malaguti *et al.* 2013), and Neurod1 can induce the formation of Tubb3-positive neurones (unpublished observations in the Lowell laboratory), which is unsurprising considering its overexpression is used to generate neurones from mesodermal cells (Pang *et al.* 2011). It is however unlikely that the inhibition of any of these factors is responsible for the delay in the exit from primed pluripotency and in the induction of proximal posterior epiblast gene expression in Id1 overexpressing cells. *Twist1* is not expressed in post-implantation embryos until the early bud stage, when its transcription is activated in mesodermal cells; knockdown of *Twist1* expression in N2B27 culture does not hinder neural specification (unpublished observations by Dr. Paul Nistor in our laboratory). *Neurod1* is expressed at low levels in a subset of ES cells characterised by low expression of *Id* genes, but its expression is not associated with lower levels of pluripotency factors nor with higher levels of neural markers (data not shown). This, coupled to the fact that Neurod1 is a terminal differentiation factor activated late in the cascade of pro-neural genes (Ma *et al.* 1999), suggests it is also unlikely to be

involved in the exit from pluripotency and the specification of neural ectoderm. *Tcf15* is downregulated by day 1 of N2B27 differentiation in both control and Id1 overexpressing cells, suggesting it cannot be responsible for the loss of Cdh1 and Oct4 that occurs between day 2 and day 4 of differentiation in control cells (Malaguti *et al.* 2013).

The relationship between Id1 and Tcf15 appears to be complicated by the interaction of both of these factors with Nanog. Tcf15 can repress *Nanog* expression in ES cell culture, *Id1* expression is inhibited by Nanog, and Id1 can inhibit the activity of Tcf15, resulting in the generation of a feedback loop between these factors that has the potential to regulate the pluripotent status of ES cells (Figure 6.3). It was interesting to note that in the sorted Id1-Venus Nanog-tagRFP reporter ES cells, *Tcf15* expression was enriched in the Nanog-tagRFP-low expressing subpopulations (both Id1-Venus-high and –low), and that the expression of its target *Otx2* (Davies *et al.* 2013), as well as that of other epiblast markers, was enriched in the Id1-Venus-low Nanog-tagRFP-low subpopulation (Figure 3.18), suggesting Id1 may be acting in a subset of Nanog-low ES cells to preserve a state of pre-implantation pluripotency.

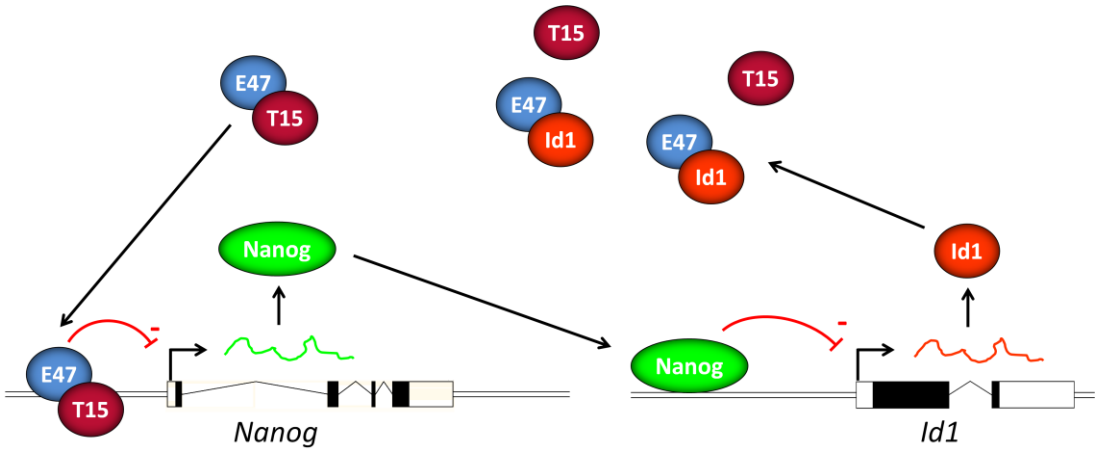


Figure 6.3 – Cross-regulation between Id1, Tcf15 and Nanog

Tcf15/E47 dimers can repress the expression of *Nanog*. Nanog can repress the expression of *Id1*. Id1 can inhibit the transcriptional activity of Tcf15 by sequestering E47. The direct binding of Tcf15/E47 at the *Nanog* promoter and of Nanog at the *Id1* promoter has not been demonstrated and is illustrated here in speculative terms only. T15: Tcf15.

6.2.1.2 E proteins

The factor(s) inhibited by Id1 to delay the loss of pluripotency markers and to promote a proximal posterior identity must therefore be other interactors of Id1. This leaves two possibilities amongst the known interactors of Id1 expressed in differentiating ES cells: E proteins and Ets family members.

E47 can directly bind to E-boxes in the *Cdh1* promoter and repress the transcription of the gene (Pérez-Moreno *et al.* 2001). E proteins can homodimerise, and thus require no other bHLH factor for their DNA binding activity (Murre *et al.* 1989a). It is therefore possible that the E47-mediated downregulation of *Cdh1* expression is carried out by E47 homodimers. As *Id1* is sharply downregulated at the onset of differentiation (Figure 1.2), E47 would no longer be prevented from heterodimerising by Id1 and could induce the downregulation of Cdh1. This could in turn drive the exit from pluripotency and the acquisition of a "default" neural state

(Figure 6.5a,b). The overexpression of Id1 would therefore account for the delayed exit from pluripotency, but it is unclear how it would cause a shift in the regional identity of the differentiating cells. It is possible that an E47-interacting bHLH factor not expressed in ES cells (or not detected by the yeast two-hybrid screen) is upregulated at the onset of differentiation and is capable of inhibiting proximal posterior epiblast specification, or it is possible that the extended period of pluripotent gene expression would result in a transcriptional equilibrium being established, with a number of cells activating the expression of markers of proximal posterior epiblast, as seen in epiblast stem cells, where T expression is observed (Tesar *et al.* 2007).

E47 regulates the downregulation of *Cdh1* and promotes epithelial-to-mesenchymal transitions in a number of cell types (Lee *et al.* 2011a, Slattery *et al.* 2006, Slattery & McMorrow 2008), and Id1 has been shown to inhibit this activity in cultured cells (Cubillo *et al.* 2013, Kondo *et al.* 2004). It is therefore possible that Id1 may act as an anti-EMT factor during various stages of embryonic development. For example, preliminary observations suggest Id1 may be involved in the regulation of the formation of the notochord by cells delaminating from the hindgut (Figure 6.4) (Inman & Downs 2006).

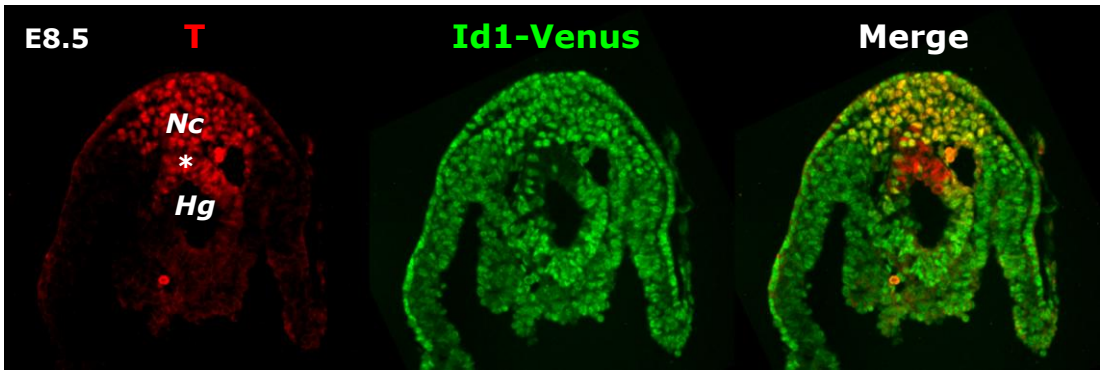


Figure 6.4 – Id1-Venus is not expressed by cells undergoing EMT in the hindgut

Transverse section through the caudal region of an E8.5 chimaera generated by morula aggregation of IdV ES cells. Id1-Venus expression is not detected in T-expressing cells that are delaminating (*) from the hindgut (Hg) to form the developing notochord (Nc).

6.2.1.3 Ets protein family

Id1 has also been shown to interact with the Ets family members Elk1 and Ets2 and to repress their transcriptional activity (Ohtani *et al.* 2001, Yates *et al.* 1999). The Ets protein family comprises many proteins involved in various biological processes, which include the transcriptional activation of Zeb factors. Ets1 can induce the transcription of *Zeb1* and *Zeb2*, and this activity can be inhibited by Id2 (Shirakihara *et al.* 2007). Zeb (zinc-finger E-box binding homeobox) factors can bind to the E-boxes present in the *Cdh1* promoter and repress the transcription of the gene (Comjin *et al.* 2001, Shirakihara *et al.* 2007). In my single-cell qRT-PCR analysis of ES cells undergoing neural differentiation, I observed that the upregulation of *Zeb1* and *Zeb2* in a subset of cells at day 1 of differentiation preceded the downregulation of *Cdh1* and the upregulation of *Sox1* (Figure 4.15), and that expression of the two *Zeb* genes displayed a weak but statistically significant correlation at this timepoint (Figure 4.16), suggesting that a common mechanism for the induction of the transcription of

both *Zeb* genes may be in place. *Ets1* and other Ets family members, such as the Fgf target *Etv4*, are expressed during neural adherent monolayer in N2B27 medium (Aiba *et al.* 2006, Aiba *et al.* 2009, data not shown). It is therefore possible that upon loss of Id1 expression in wild-type cells one or more Ets family members may induce the expression of Zeb1 and Zeb2, which could in turn downregulate *Cdh1* expression and induce neural differentiation. Zeb2 is required for the neural specification of human ES cells, consistent with a possible role of the transcription factor in this process (Chng *et al.* 2010). In this model, the overexpression of Id1 would inhibit the transcriptional activity of target Ets proteins and thus prevent the upregulation of Zeb factors. The level of *Zeb* transcripts in Id1 overexpressing cells is lower than in control cells throughout differentiation, in accordance with this hypothesis (Figures 4.7, 4.9). The repression of *Zeb1* and *Zeb2* transcription would in turn result in the maintenance of Cdh1 expression and pluripotency, and the inhibition of a neural specification transcriptional programme (Figure 6.5c,d).

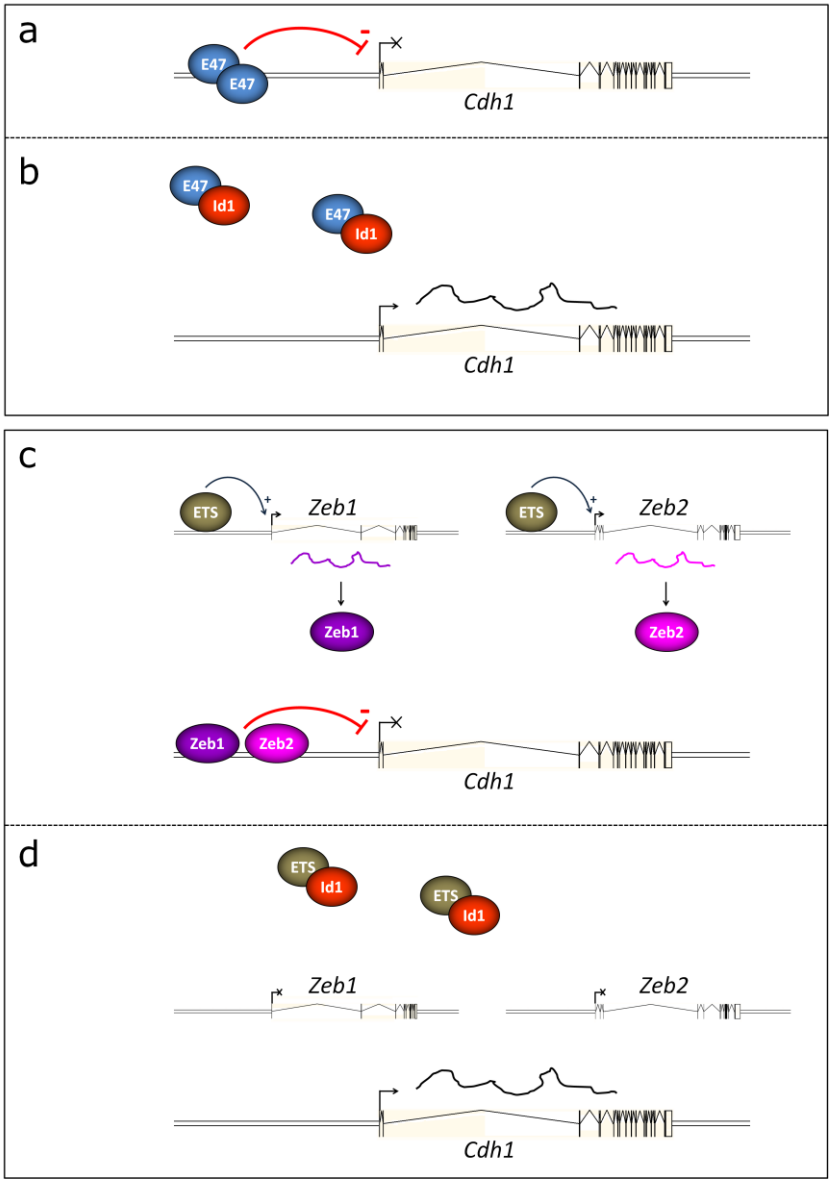


Figure 6.5 – Alternative models for the maintenance of Cdh1 expression by Id1

Id1 delays the downregulation of Cdh1 in ES cell differentiation. (a,b) It may achieve this by repressing the transcriptional downregulation of *Cdh1* by E47 homodimers (a: wild-type cell, b: Id1 overexpressing cell). (c,d) Inhibition of an Ets protein may prevent expression of Zeb1 and Zeb2, which can repress *Cdh1* expression (c: wild-type cell, d: Id1 overexpressing cell).

6.2.2 Id1 and the specification of surface ectoderm

The differentiation of ES cells in N2B27 in the presence of BMP leads to the formation of two types of Cdh1-expressing cells: pluripotent colonies expressing Oct4 and T, and differentiated "big flat cells" expressing the surface ectoderm and trophectoderm marker Tcfap2a (Figure 4.1, Malaguti *et al.* 2013). Ying *et al.* (2003a) demonstrated that Id1 overexpression in N2B27 culture can also generate these big flat cells. Furthermore, I observed high levels of Id1 expression in the proximal anterior regions of gastrulation stage embryos, where cells destined for surface ectoderm specification reside (Cajal *et al.* 2012, Madabhushi & Lacy 2011, Tam & Behringer 1997), an observation confirmed by the *in situ* hybridisation data of Li *et al.* (2013). This would suggest an ability for Id1 to induce surface ectoderm differentiation, possibly through the inhibition of an EMT event. I did not observe differentiation of Id1 overexpressing cells into surface ectoderm nor other epithelial cell types in my experiments. As suggested in paragraph 4.3.3.1, it is possible that different levels of Id1 expression influence the direction of differentiation, with high levels driving surface ectoderm formation, medium levels driving mesodermal specification and low levels allowing neural differentiation. The generation of novel cell lines that consent the expression of high levels of Id1 during differentiation, through the use of alternative promoters or multiple transgenes, may help to test the veracity of this hypothesis.

6.3 Concluding remarks

Id1 is a transcriptional modulator of tremendous potential. Its activity as a dominant negative regulator of gene expression implies that it has the capability to affect all biological processes involving the activity of its heterodimerisation partners, and of the molecules regulated by them.

CHAPTER 6 General discussion

This work has shown that Id1 is expressed in the post-implantation embryo at the appropriate locations in time and space for it to influence early lineage specification events, and that it is capable of affecting these developmental decisions in an *in vitro* model of embryonic development: embryonic stem cell differentiation. Furthermore, a novel expression pattern for Id1 has been observed at the node of headfold stage embryos, raising the question of whether it may play a role in the regulation of this signalling centre.

The regulation of Id1 expression by extracellular signalling pathways has also been characterised extensively in various types of pluripotent cultures, providing a clear picture of the effect of these pathways on Id1 expression in isolation and with respect to one another, thus increasing our understanding of the mechanisms that may regulate the expression of this molecule in the development of the mouse.

Many avenues remain open for exploration. The significance of the negative relationship between Nanog and Id1 is still unclear. The roles of Id1 in promoting pluripotency and regionalising cells in culture have not been uncoupled, as the factors inhibited by Id1 in differentiation have not been identified. The expression of Id1 at the node requires further characterisation, as it is currently unclear what signals drive the establishment of this expression pattern and what the significance of Id1 expression in this structure is. It is to be hoped that future research will allow us to answer these questions and thereby improve our understanding of the biology of this small but significant helix-loop-helix factor.

REFERENCES

Afrakhte M., Morén A., Jossan S., Itoh S, Sampath K., Westermark B., Heldin C.H., Heldin N.E., ten Dijke P. (1998). Induction of inhibitory Smad6 and Smad7 mRNA by TGF-β family members. *Biochemical and Biophysical Research Communications* 249:505-511.

Aiba K., Nedorezov T., Piao Y., Nishiyama A., Matoba R., Sharova L.V., Sharov A.A., Yamanaka S., Niwa H., Ko M.S.H. (2009). Defining developmental potency and cell lineage trajectories by expression profiling of differentiating mouse embryonic stem cells. *DNA Research* 16:73-80.

Aiba K., Sharov A.A., Carter M.G., Foroni C., Vescovi A.L., Ko M.S.H. (2006). Defining a developmental path to neural fate by global gene expression profiling of mouse embryonic stem cells and adult neural stem progenitor cells. *Stem Cells* 24:889-895.

Aksoy I., Jauch R., Chen J., Dyla M., Divakar U., Bogu G.K., Teo R., Leng Ng C.K., Herath W., Lili S., Hutchins A.P., Robson P., Kolatkar P.R., Stanton L.W. (2013). Oct4 switches partnering from Sox2 to Sox17 to reinterpret the enhancer code and specify endoderm. *The EMBO Journal* 32:938-953.

Allendorph G.P., Vale W.W., Choe S. (2006). Structure of the ternary signalling complex of a TGF-β superfamily member. *Proceedings of the National Academy of Sciences of the United States of America* 103:7643-7648.

Anand G., Yin X., Shahidi A.K., Grove L., Prochownik E.V. (1997). Novel regulation of the helix-loop-helix protein Id1 by S5a, a subunit of the 26 S proteasome. *The Journal of Biological Chemistry* 272:19140-19151.

Anokye-Danso F., Trivedi C.M., Juhr D., Gupta M., Cui Z., Tian Y., Zhang Y., Yang W., Gruber P.J., Epstein J.A., Morrisey E.E. (2011). Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. *Cell Stem Cell* 8:376-388.

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

Aoyama M., Sun-Wada G.H., Yamamoto A., Yamamoto M., Hamada H., Wada Y. (2012). Spatial restriction of bone morphogenetic protein signaling in mouse gastrula through the mVam2-dependent endocytic pathway. *Developmental Cell* 22:1163-1175.

Arkell R., Beddington R.S.P. (1997). BMP-7 influences pattern and growth of the developing hindbrain of mouse embryos. *Development* 124:1-12.

Arnold S.J., Hofmann U.K., Bikoff E.K., Robertson E.J. (2008). Pivotal roles for eomesodermin during axis formation, the epithelium-to-mesenchyme transition and endoderm specification in the mouse. *Development* 135:501-511.

Arnold S.J., Stappert J., Bauer A., Kispert A., Hermann B.G. Kemler R. (2000). *Brachyury* is a target gene of the Wnt/β -catenin signalling pathway. *Mechanisms of Development* 91:249-258.

Atchley W.R., Fitch W.M. (1997). A natural classification of the basic helix-loop-helix class of transcription factors. *Proceedings of the National Academy of Sciences of the United States of America* 94:5172-5176.

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

Bai G., Sheng N., Xie Z., Bian W., Yokota Y., Benezra R., Kageyama R., Guillemot F., Jing N. (2007). Id sustains *Hes1* expression to inhibit precocious neurogenesis by releasing negative autoregulation of *Hes1*. *Developmental Cell* 13:283-297.

Barndt R.J., Zhuang Y. (1999). Controlling lymphopoiesis with a combinatorial E-protein code. *Cold Spring Harbor Symposia on Quantitative Biology* 64:45-50.

Barone M.V., Pepperkok R., Peverali F.A., Philipson L. (1994). Id proteins control growth induction in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* 91:4985-4988.

Bayart E., Cohen-Haguenauer O. (2013). Technological overview of iPS induction from human adult somatic cells. *Current Gene Therapy* 13:73-92.

Beddington R.S.P., Rashbass P., Wilson V. (1992). *Brachyury* – a gene affecting mouse gastrulation and early organogenesis. *Development Supplement* 157-165.

Beddington R.S.P., Robertson E.J. (1999). Axis development and early asymmetry in mammals. *Cell* 96:195-209.

Bedford L., Walker R., Kondo T., van Crüchten I., King E.R., Sablitzky F. (2005). Id4 is required for the correct timing of neural differentiation. *Developmental Biology* 280:386-395.

Bedzhov I., Alotaibi H., Basilicata M.F., Ahlborn K., Liszewska E., Brabletz T., Stemmler M.P. (2013). Adhesion, but not a specific cadherin code, is indispensable for ES cell and induced pluripotency. *Stem Cell Research* 11:1250-1263.

Bedhzov I., Liszewska E., Kanzler B., Stemmler M.P. (2012). Igf1r signalling is indispensable for preimplantation development and is activated via a novel function of E-cadherin. *PLoS Genetics* 8:e1002609.

Belo J.A., Bachiller D., Agius E., Kemp C., Borges A.C., Marques S., Piccolo S., De Robertis E.M. (2000). *Cerberus-like* is a secreted BMP and nodal antagonist not essential for mouse development. *Genesis* 26:265-270.

Ben-Haim N., Lu C., Guzman-Ayala M., Pescatore L., Mesnard D., Bischofberger M., Naef F., Robertson E.J., Constam D.B. (2006). The Nodal precursor acting via Activin receptors induces mesoderm by maintaining a source of its convertases and BMP4. *Developmental Cell* 11:313-323.

Benezra R., Davis R.L., Lockshon D., Turner D.L., Weintraub H. (1990). The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. *Cell* 61:49-59.

Beppu H., Kawabata M., Hamamoto T., Chytil A., Minowa O., Noda T., Miyazono K. (2000). BMP type II receptor is required for gastrulation and early development of mouse embryos. *Developmental Biology* 221:249-258.

Bertrand N., Castro D.S., Guillemot F. (2002). Proneural genes and the specification of neural cell types. *Nature Reviews Neuroscience* 3:517-530.

Blackwood E.M., Eisenman R.N. (1991). Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. Sc*ience* 251:1211-1217.

Bottenstein J.E., Sato G.H. (1979). Growth of a rat neuroblastoma cell line in serum-free supplemented medium. *Proceedings of the National Academy of Sciences of the United States of America* 76:514-517.

Bounpheng M.A., Dimas J.J., Dodds S.G., Christy B.A. (1999). Degradation of Id proteins by the ubiquitin-proteasome pathway. *The FASEB Journal* 13:2257-2264.

Bourillot P.Y., Aksoy I., Schreiber V., Wianny F., Schulz H., Hummel O., Hubner N., Savatier P. (2009). Novel STAT3 target genes exert distinct roles in the inhibition of mesoderm and endoderm differentiation in cooperation with Nanog. *Stem Cells* 27:1760-1771.

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

Brennan J., Lu C.C., Norris D.P., Rodriguez T.A., Beddington R.S.P., Robertson E.J. (2001). *Nodal* signalling in the epiblast patterns the early mouse embryo. *Nature* 411:965-969.

Brewer G.J., Torricelli J.R., Evege E.K., Price P.J. (1989). Optimal survival of hippocampal neurons in B27-supplemented Neurobasal™, a new serum-free medium combination. *Journal of Neuroscience* 35:567-576.

Brinster R.L. (1974). The effect of cells transferred into the mouse blastocyst on subsequent development. *The Journal of Experimental Medicine* 140:1049-1056.

Brodin G., Åhgren A., ten Dijke P., Heldin C.H., Heuchel R. (2000). Efficient TGF-β induction of the Smad7 gene required cooperation between AP-1, Sp1, and Smad proteins on the mouse Smad7 promoter. *The Journal of Biological Chemistry* 275:29023-29030.

Brons I.G.M., Smithers L.E., Trotter M.W.B., Rugg-Gunn P., Sun B., Chuva de Sousa Lopes S.M., Howlett S.K., Clarkson A., Ahrlund-Richter L., Pedersen R.A., Vallier L. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* 448:191-195.

Burdon T., Chambers I., Stracey C., Niwa H., Smith A. (1999). Signaling mechanisms regulating self-renewal and differentiation of pluripotent embryonic stem cells. *Cells Tissues Organs* 165:131-143.

Burdon T., Stracey C., Chambers I., Nichols J., Smith A. (1999). Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. *Developmental Biology* 210:30-43.

Cajal M., Lawson K.A., Hill B., Moreau A., Rao J., Ross A., Collignon J., Camus A. (2012). Clonal and molecular analysis of the prospective anterior neural boundary in the mouse embryo. *Development* 139:423-436.

Cambray N., Wilson V. (2002). Axial progenitors with extensive potency are localised to the mouse chordoneural hinge. *Development* 129:4855-4866.

Camus A., Perea-Gomez A., Moreau A., Collignon J. (2006). Absence of Nodal signaling promotes precocious neural differentiation in the mouse embryo. *Developmental Biology* 295:743-755.

Candia A.F., Watabe T., Hawley S.H.B., Onichtchouk D., Zhang Y., Derynck R., Niehrs C., Cho K.W.Y. (1997). Cellular interpretation of multiple TGF-β signals: intracellular antagonism between activin/BVg1 and BMP-2/4 signaling mediated by Smads. *Development* 124:4467-4480.

Canham M.A., Sharov A.A., Ko M.S.H., Brickman J.M. (2010). Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcript. *PLoS Biology* 8:e100379.

Carver E.A., Jiang R., Lan Y., Oram K.F., Gridley T. (2001). The mouse Snail gene encodes a key regulator of the epithelial-mesenchymal transition. *Molecular and Cellular Biology* 21:8184-8188.

Casey E.S., O'Reilly M.A.J., Conlon F.L., Smith J.C. (1998). The T-box transcription factor Brachyury regulates expression of *eFGF* through binding to a non-palindromic response element. *Development* 125:3887-3894.

Cattaneo E., Cerbai E., Garagna S. (2010). Italy's stem cell challenge gaining momentum. *Nature* 463:729.

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

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

Chambers S.M., Fasano C.A., Papapetrou E.P., Tomishima M., Sadelain M., Studer L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nature Biotechnology* 27:275-280.

Chambers S.M., Qi Y., Mica Y., Lee G., Zhang X.J., Niu L., Bilsland J., Cao L., Stevens E., Whiting P., Shi S.H., Studer L. (2012). Combined small-molecule inhibition accelerates developmental timing and converts human pluripotent stem cells into nociceptors. *Nature Biotechnology* 30:715-720.

Chan H.Y., Sivakamasundari V., Xing X., Kraus P., Yap S.P., Ng P., Lim S.L., Lufkin T. (2011). Comparison of IRES and F2A-based locus-specific multicistronic expression in stable mouse lines. *PLoS ONE* 6:e28885.

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

Chang C., Hemmati-Brivanlou A. (1999). Xenopus GDF6, a new antagonist of noggin and a partner of BMPs. *Development* 126:3347-3357.

Chappell S.A., Edelman G.M., Mauro V.P. (2000). A 9-nt segment of a cellular mRNA can function as an internal ribosome entry site (IRES) and when present in linked multiple copies greatly enhances IRES activity. *Proceedings of the National Academy of Sciences of the United States of America* 97:1536-1541.

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

Chen C., Shen M.M. (2004). Two modes by which Lefty proteins inhibit Nodal signaling. *Current Bioogy* 14:618-624.

Chen C.W., Liu C.S., Chiu I.M., Shen S.C., Pan H.C., Lee K.H., Lin S.Z., Su H.L. (2010). The signals of FGFs on the neurogenesis of embryonic stem cells. *Journal of Biomedical Science* 17:33.

Chen F., Weinberg R.A. (1995). Biochemical evidence for the autophosphorylation and transphosphorylation of transforming growth factor β receptor kinases. *Proceedings of the National Academy of Sciences of the United States of America* 92:1565-1569.

Chen H., Thiagalingam A., Chopra H., Borges M.W., Feder J.N., Nelkin B.D., Baylin S., Ball D.W. (1997a). Conservation of the *Drosophila* lateral inhibition pathway in human lung cancer: a hairy-related protein (HES-1) directly represses achaete-scute homolog-1 expression. *Proceedings of the National Academy of Sciences of the United States of America* 5355-5360.

Chen J., Liu J., Yang J., Chen Y., Chen J., Ni S., Song H., Zeng L., Ding K., Pei D. (2011). BMPs functionally replace Klf4 and support efficient reprogramming of mouse fibroblasts by Oct4 alone. *Cell Research* 21:205-212.

Chen X., Xu H., Yuan P., Fang F., Huss M., Vega V.B., Wong E., Orlov Y.L., Zhang W., Jiang J., Loh Y.H., Yeo H.C., Yeo Z.X., Narang V., Govindarajan K.R., Leong B., Shahab A., Ruan Y., Bourque G., Sung W.K., Clarke N.D., Wei C.L., Ng H.H. (2008). Integration of external signalling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133:1106-1117.

Chen Y., Bhushan A., Vale W. (1997b). Smad8 mediates the signalling of the receptor serine kinase. *Proceedings of the National Academy of Sciences of the United States of America* 64:12938-12943.

Cheng S.K., Olale F., Bennett J.T., Brivanlou A.H., Schier A.F. (2003). EGF-CFC proteins are essential coreceptors for the TGF-β signals Vg1 and GDF1. *Genes & Development* 17:31-36.

Chew J.L., Loh Y.H., Zhang W., Chen X., Tam W.L., Yeap L.S., Li P., Ang Y.S., Lim B., Robson P., Ng H.H. (2005). Reciprocal transcriptional regulation of *Pou5f1* and *Sox2* via the Oct4/Sox2 complex in embryonic stem cells. *Molecular and Cellular Biology* 25:6031.

Chng Z., Teo A., Pedersen R.A., Vallier L. (2010). SIP1 mediates cell-fate decisions between neuroectoderm and mesendoderm in human pluripotent stem cells. *Cell Stem Cell* 6:59-70.

Christy B.A., Sanders L.K., Lau L.F., Copeland N.G., Jenkins N.A., Nathans D. (1991). An Id-related helix-loop-helix protein encoded by a growth factor-inducible gene. *Proceedings of the National Academy of Sciences of the United States of America* 88:1815-1819.

Ciruna B.G., Rossant J. (1999). Expression of the T-box gene *Eomesodermin* during mouse development. *Mechansisms of Development* 81:199-203.

Ciruna B.G., Rossant J. (2001). FGF signaling regulates mesoderm cell fate specification and morphogenetic movement at the primitive streak. *Developmental Cell* 1:37-49.

Clevers H., Nusse R. (2012). Wnt/β-catenin signaling and disease. *Cell* 149:1192-1205.

Colvin J.S., Bohne B.A., Harding G.W., McEwen D.G., Ornitz D.M. (1996). Skeletal overgrowth and deafness in mice lacking fibroblast growth factor receptor 3. *Nature Genetics* 12:390-397.

Comjin J., Berx G., Vermassen P., Vershueren K., van Grunsven L., Bruyneel E., Mareel M., Hylebroeck D., van Roy F. (2001). The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. *Molecular Cell* 7:1267-1278.

Conlon F.L., Lyons K.M., Takaesu N., Barth K.S., Kispert A., Herrmann B., Robertson E.J. (1994). A primary requirement for *nodal* in the formation and maintenance of the primitive streak in the mouse. *Development* 120:1919-1928.

Copp A.J. (1979). Interaction between inner cell mass and trophectoderm of the mouse blastocyst. II. The fate of the polar trophectoderm. *Journal of Embryology and Experimental Morphology* 51:109-120.

Coraux C., Hilmi C., Rouleau M., Spadafora A., Hinnrasky J., Ortonne J.P., Dani C., Aberdam D. (2003). Reconstituted skin from murine embryonic stem cells. *Current Biology* 13:849-853.

Costa Y., Ding J., Theunissen T.W., Faiola F., Hore T.A., Shliaha P.V., Fidalgo M., Saunders A., Lawrence M., Dietmann S., Das S., Levasseur D.N., Li Z., Xu M., Reik W., Silva J.C.R., Wang J. (2013). Nanog-dependent function of Tet1 and Tet2 in establishment of pluripotency. *Nature* 495:370-374.

Coucouvanis E., Martin G.R. (1995). Signals for death and survival: a two-step mechanism for cavitation in the vertebrate embryo. *Cell* 83:279-287.

Cross J.C., Werb Z., Fisher S.J. (1994). Implantation and the placenta: key pieces of the development puzzle. *Science* 266:1508-1518.

Crossley P.H., Martin G.R. (1995). The mouse *Fgf8* gene encodes a family of polypeptides and is expressed in regions that direct outgrowth and patterning in the developing embryo. *Development* 121:439-451.

Cubillo E., Diaz-Lopez A., Cuevas E.P., Moreno-Bueno G., Peinado H., Montes A., Santos V., Portillo F., Cano A. (2013). E47 and Id1 interplay in epithelial-mesenchymal transition. *PLoS ONE* 8:e59948.

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

Dai J.P., Lu J.Y., Zhang Y. (2010). Jmjd3 activates *Mash1* gene in RA-induced neuronal differentiation of P19 cells. *Journal of Cellular Biochemistry* 110:1457-1463.

Dale L., Howes G., Price B.M., Smith J.C. (1992). Bone morphogenetic protein 4: a ventralizing factor in early Xenopus development. *Development* 115:573-585.

Dalerba P., Kalisky T., Sahoo D., Rajendran P.S., Rothenberg M.E., Levrat A.A., Sim S., Okamoto J., Johnston D.M., Qian D., Zabala M., Bueno J., Neff N.F., Wang J., Shelton A.A., Visser B., Hisamori S., Shimono Y., van de Wetering M., Clevers H., Clarke M.F., Quake S.R. (2011). Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nature Biotechnology* 29:1120-1127.

Danesh S.M., Villasenor A., Chong D., Soukup C., Cleaver O. (2009). BMP and BMP receptor expression during murine organogenesis. *Gene Expression Patterns* 9:255-265.

Davies O.R., Lin C.Y., Radzisheuskaya A., Zhou X., Taube J., Blin G., Waterhouse A., Smith A.J.H., Lowell S. (2013). Tcf15 primes pluripotent cells for differentiation. *Cell Reports* 3:472-484.

Davis S., Aldrich T.H., Stahl N., Pan L.,Taga T., Kishimoto T., Ip N.Y., Yancopoulos G.D. (1993). LIFR beta and gp130 as heterodimerizing signal transducers of the tripartite CNTF receptor. *Science* 260:1805-1808.

Dawson S.R., Turner D.L., Weintraub H., Parkhurst S.M. (1995). Specificity for the Hairy/Enhancer of split basic helix-loop-helix (bHLH) proteins maps outside the bHLH domain and suggests two separable modes of transcriptional repression. *Molecular and Cellular Biology* 15:6923-6931.

De Robertis E.M., Kuroda H. (2004). Dorsal-ventral patterning and neural induction in *Xenopus* embryos. *Annual Reviews of Cell and Developmental Biology* 20:285-308.

de Vries W.N., Evsikov A.V., Haac B.E., Fancher K.S., Holbrook A.E., Kemler R., Solter S., Knowles B.B. (2004). Maternal β-catenin and E-cadherin in mouse development. *Development* 131:4435-4445.

Deed R.W., Armitage S., Norton J.D. (1996). Nuclear localization and regulation of Id protein through an E protein-mediated chaperone mechanism. *The Journal of Biological Chemistry* 271:23603-23606.

Deed R.W., Hara E., Atherton G.T., Peters G., Norton J.D. (1997). Regulation of Id3 cell cycle function by Cdk-2-dependent phosphorylation. *Molecular and Cellular Biology* 17:6815-6821.

del Valle I., Rudloff S., Carles A., Li Y., Liszewska E., Vogt R., Kemler R. (2013). E-cadherin is required for the proper activation of the Lifr/Gp130 signaling pathway in mouse embryonic stem cells. *Development* 140:1684-1692.

Di-Gregorio A., Sancho M., Stuckey D.W., Crompton L.A., Godwin J., Mishina Y., Rodriguez T.A. (2007). BMP signalling inhibits premature neural differentiation in the mouse embryo. *Development* 134:3359-3369.

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

Doetschman T.C., Eistetter H., Katz M., Schmidt W., Kemler R. (1985). The *in vitro* development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *Journal of Embryology and Experimental Morphology* 87:27-45.

Donnelly M.L.L., Luke G., Mehrotra A., Li X., Hughes L.E., Gani D., Ryan M.D. (2011). Analysis of the aphthovirus 2a/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. *Journal of General Virology* 82:1013-1025.

Downs K.M., Davies T. (1993). Staging of gastrulating mouse embryos by morphological landmarks in the dissecting microscope. *Development* 118:1255-1266.

Du J., Takeuchi H., Leonhard-Melief C., Shroyer K.R., Dlugosz M., Haltiwanger R.S., Holdener B.C. (2010). O-fucosylation of thrombospondin type 1 repeats restricts epithelial to mesenchymal transition (EMT) and maintains epiblast pluripotency during mouse gastrulation. *Developmental Biology* 346:25-38.

Duncan M., DiCicco-Bloom M., Xiang X., Benezra R., Chada K. (1992). The gene for the helix-loop-helix protein, Id, is specifically expressed in neural precursors. *Developmental Biology* 154:l-10.

Duncan M.K., Shimamura T., Chada K. (1994). Expression of the helix-loop-helix protein, Id, during branching morphogenesis in the kidney. *Kidney International* 46:324-332.

Dvorak P., Dvorakova D., Koskova S., Vodinska M., Najvirtova M., Krekac D., Hampl A. (2005). Expression and potential role of fibroblast growth factor 2 and its receptors in human embryonic stem cells. *Stem Cells* 23:1200-1211.

Ellenberger T., Fass D., Arnaud M., Harrison S.C. (1994). Crystal structure of transcription factor E47: E-box recognition by a basic region helix-loop-helix dimer. *Genes & Development* 8:970-980.

Ellmeier W., Weith A. (1995). Expression of the helix-loop-helix gene Id3 during murine embryonic development. *Developmental Dynamics* 203:163-173.

Ema M., Mori D., Niwa H., Hasegawa Y., Yamanaka Y., Hitoshi S., Mimura J., Kawabe Y., Hosoya T., Morita M., Shimosato D., Uchida K., Suzuki N., Yanagisawa J., Sogawa K., Rossant J., Yamamoto M., Takahashi S., Fujii-Kuriyama Y. (2008). *Krüppel-like factor 5* is essential for blastocyst development and the normal self-renewal of mouse ESCs. *Cell Stem Cell* 3:555-567.

Ephrussi A., Church G.M., Tonegawa S., Gilbert W. (1985). B lineage-specific interactions of an immunoglobulin enhancer with cellular factors in vivo. *Science* 227:134-140.

Esteban M.A., Wang T., Qin B., Yang J., Qin D., Cai J., Li W., Weng Z., Chen J., Ni S., Chen K., Li Y., Liu X., Xu J., Zhang S., Li F., He W., Labuda K., Song Y., Peterbauer A., Wolbank S., Redl H., Zhong M., Cai D., Zeng L., Pei D. (2009). Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. *Cell Stem Cell* 6:71-79.

Evans A.L., Faial T., Gilchrist M.J., Down T., Vallier L., Pedersen R.A., Wardle F.C., Smith J.C. (2012). Genomic targets of Brachyury (T) in differentiating mouse embryonic stem cells. *PLoS ONE* 7:e33346.

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

Evans S.M., O'Brien T.X. (1993). Expression of the helix-loop-helix factor Id during mouse embryonic development. *Developmental Biology* 159:485-499.

Faunes F., Hayward P., Muñoz Descalzo S., Chatterjee S.S., Balayo T., Trott J., Christoforou A., Ferrer-Vaquer A., Hadjantonakis A.K., Dasgupta R., Martinez Arias A. (2013). A membrane-associated β-catenin/Oct4 complex correlates with ground-state pluripotency in mouse embryonic stem cells. *Development* 140:1171-1183.

Faust C., Schumacher A., Holdener B., Magnuson T. (1995). The eed mutation disrupts anterior mesoderm production in mice. *Development* 121:273-285.

Fei T., Zhu S., Xia K., Zhang J., Li Z., Han J.D.J., Chen Y.G. (2010). Smad2 mediates Activin-Nodal signaling in mesendoderm differentiation of mouse embryonic stem cells. *Cell Research* 20:1306-1318.

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

Feng B., Jiang J., Kraus P., Ng J.H., Heng J.C.D., Chan Y.S., Yaw L.P., Zhang W., Loh Y.H., Han J., Vega V.B., Cacheux-Rataboul V., Lim B., Lufkin T., Ng H.H. (2009). Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nature Cell Biology* 11:197-203.

Feng X.H., Derynck R. (1997). A kinase subdomain of transforming growth factor-β (TGF-β) type I receptor determines the TGF-β intracellular signaling specificity. *The EMBO Journal* 16:3912-3923.

Ferré-D'Amaré A.R., Prendergast G.C., Ziff E.B., Burley S.K. (1993). Recognition by Max of its cognate DNA through a dimeric b/HLH/Z domain. *Nature* 363:38-45.

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

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

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

Finley M.F.A., Devata S., Huettner J.E. (1999). BMP-4 inhibits neural differentiation of embryonic stem cells. *Journal of Neurobiology* 40:271-287.

Fraidenraich D., Lang R., Basilico C. (1998). Distinct regulatory elements govern *Fgf4* gene expression in the mouse blastocyst, myotomes and developing limb. *Developmental Biology* 204:197-209.

Fraidenraich D., Stillwell E., Romero E., Wilkes D., Manova K., Basson C.T., Benezra R. (2004). Rescue of cardiac defects in Id knockout embryos by injection of embryonic stem cells. *Science* 306:247-252.

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

Frum T., Halbisen M.A., Wang C., Amiri H., Robson P., Ralston A. (2013). Oct4 cell-autonomously promotes primitive endoderm development in the mouse blastocyst. *Developmental Cell* 25:610-622.

Gagliardi A., Mullin N.P., Tan Z.Y., Colby D., Kousa A.I., Halbritter F., Weiss J.T., Felker A., Bezstarosti K., Favaro R., Demmers J., Nicolis S.K., Tomlinson S.R., Poot R.A., Chambers I. (2013). A direct physical interaction between Nanog and Sox2 regulates embryonic stem cell self-renewal. *The EMBO Journal* 32:2231-2247.

Gafni O., Weinberger L., Mansour A.A., Manor Y.S., Chomsky E., Ben-Yosef D., Kalma Y., Viukov S., Maza I., Zviran A., Rais Y., Shipony Z., Mukamel Z., Krupalnik V., Zerbib M., Geula S., Caspi I., Schneir D., Shwartz T., Gilad S., Amann-Zalcenstein D., Benjamin S., Amit I., Tanay A., Massarwa R., Novershtem N., Hanna J.H. (2013). Derivation of novel human ground state naive pluripotent stem cells. *Nature* 504:282-286.

Galvin K.E., Travis E.D., Yee D., Magnuson T., Vivian J.L. (2010). Nodal signaling regulates the bone morphogenic protein pluripotency pathway in mouse embryonic stem cells. *The Journal of Biological Chemistry* 285:19747-19756.

García-García M.J., Anderson K.V. (2003) Essential role of glycosaminoglycans in Fgf signalling during mouse gastrulation. *Cell* 114:727-737.

Gardner R.L. (2000). Flow of cells from polar to mural trophectoderm is polarized in the mouse blastocyst. *Human Reproduction* 15:694-701.

Gardner R.L., Davies T.J. (2002). Trophectoderm growth and bilateral symmetry of the blastocyst in the mouse. *Human Reproduction* 17:1839-1845.

Gardner R.L., Meredith M.R., Altman D.G. (1992). Is the anterior-posterior axis of the fetus specified before implantation in the mouse? *Journal of Experimental Zoology* 264:437-443.

Gearing D.P., Comeau M.R., Friend D.J., Gimpel S.D.,Thut C.J., McGourty J., Brasher K.K., King J.A., Gillis S., Mosley B., Ziegler S.F., Cosman D. (1992). The IL-6 signal transducer, gp130: an oncostatin M receptor and affinity converter for the LIF receptor. *Science* 255:1434-1437.

Giagtzoglou N., Alifragis P., Koumbanakis K.A.,, Delidakis C. (2003). Two modes of recruitment of E(spl) repressors onto target genes. *Development* 130:259-270.

Gontan C., Achame E.M., Demmers J., Barakat T.S., Rentmeester E., van IJcken W., Grootegoed J.A., Gribnau J. (2012). RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. *Nature* 458:386-390.

Gouon-Evans V., Boussemart L., Gadue P., Nierhoff D., Koehler C.I., Kubo A., Shafritz D.A., Keller G. (2006). BMP-4 is required for hepatic specification of mouse embryonic stem cell-derived definitive endoderm. *Nature Biotechnology* 24:1402-1411.

Gray P.A., Fu H., Luo P., Zhao Q., Yu J., Ferrari A., Toyoaki T., Yuk D., Tsung E.F., Xai Z., Alberta J.A., Cheng L., Liu Y., Stenman J., Valerius M.T., Billings N., Kim H.A., Greenberg M.E., McMahon A.P., Rowitch D.H., Stiles C.D., Ma Q. (2004). Mouse brain organization revealed through direct genome-scale TF expression analysis. *Science* 306: 2255-2257.

Greber B., Lehrach H., Adjave J. (2007). Fibroblast growth factor 2 modulates - transforming growth factor β signalling in mouse embryonic fibroblasts and human ESCs (hESCs) to support hESC self-renewal. *Stem Cells* 25:455-464.

Greber B., Wu G., Bernemann C., Joo J.Y., Han D.W., Ko K., Tapia N., Sabour D., Sterneckert J., Tesar P., Schöler H. (2010). Conserved and divergent roles of FGF signaling in mouse epiblast stem cells and human embryonic stem cells. *Cell Stem Cell* 6:215-226.

Greenwald J. Groppe J., Gray P. Wiater E. Kwiatkowski W., Vale W., Choe S. (2003). The BMP7/ActRII extracellular domain complex provides new insights into the cooperative nature of receptor assembly. *Molecular Cell* 11:605-617.

Griffith D.L., Keck P.C., Sampath T.K., Rueger D.C., Carlson W.D. (1996). Three-dimensional structure of recombinant human osteogenic protein 1: Structural paradigm for the transforming growth factor β superfamily. *Proceedings of the National Academy of Sciences of the United States of America* 93:878-883.

Groppe J., Greenwald J., Wiater E., Rodriguez-Leon J., Economides A.N., Kwiatkowski W., Affolter M., Vale W.W., Izpisua Belmonte J.C., Choe S. (2002). Structural basis of BMP signalling inhibition by the cystine knot protein Noggin. *Nature* 420:636-642.

Grunz H., Tacke L. (1989). Neural differentiation of *Xenopus laevis* ectoderm takes place after disaggregation and delayed reaggregation without inducer. *Cell Differentiation and Development* 28:211-218.

Guallar D., Pérez-Palacios R., Climent M., Martínez-Abadía I., Larraga A., Fernández-Juan M., Vallejo C., Muniesa P., Schoorlemmer J. (2012). Expression of endogenous retroviruses is negatively regulated by the pluripotency marker *Rex1*/*Zfp42*. *Nucleic Acids Research* 40:8993-9007.

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

Halbritter F., Vaidya H., Tomlinson S.R. (2012). GeneProf: analysis of high-throughput sequencing experiments. *Nature Methods* 9:7-8.

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

Hamazaki T., Kehoe S.M., Nakano T., Terada N. (2006). The Grb2/Mek pathway represses Nanog in murine embryonic stem cells. *Molecular and Cellular Biology* 26:7539-7549.

Hamilton W.B., Kaji K., Kunath T. (2013). ERK2 suppresses self-renewal capacity of embryonic stem cells, but is not required for multi-lineage commitment. *PLoS ONE* 8:e60907.

Han D.W., Tapia N., Joo J.Y., Greber B., Araúzo-Bravo M.J., Bernemann C., Ko K., Wu G., Stehling M., Do J.T., Schöler H.R. (2010). Epiblast stem cell subpopulations represent mouse embryos of distinct pregastrulation stages. *Cell* 143:617-627.

Hanna J., Cheng A.W., Saha K., Kim J., Lengner C.J., Soldner F., Cassady J.P., Muffat J., Carey B.W., Jaenisch R. (2010). Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proceedings of the National Academy of Sciences of the United States of America* 107:9222-9227.

Hansson M., Olesen D.R., Peterslund J.M.L., Engberg N., Kahn M., Winzi M., Klein T., Maddox-Hyttel P., Serup P. (2009). A late requirement for Wnt and FGF signaling during activin-induced formation of foregut endoderm from mouse embryonic stem cells. *Developmental Biology* 330:286-304.

Hao J., Li T.G., Qi X., Zhao D.F., Zhao G.Q. (2006). WNT/β-catenin pathway up-regulates *Stat3* and converges on LIF to prevent differentiation of mouse embryonic stem cells. *Developmental Biology* 290:81-91.

Hassani S.N., Totonchi M., Farrokhi A., Taei A., Larijani M.R., Gourabi H., Baharvand H. (2012). Simultaneous suppression of TGF-β and ERK signaling contributes to the highly efficient and reproducible generation of mouse embryonic stem cells from previously considered refractory and non-permissive strains. *Stem Cell Reviews and Reports* 8:472-481.

Hassani S.N., Totonchi M., Sharifi-Zarchi A., Mollamohammadi S., Pakzad M., Moradi S., Samadian A., Masoudi N., Mirshahvaladi S., Farrokhi A., Greber B., Araúzo-Bravo M.J., Sabour D., Sadeghi M., Salekdeh G.H., Gourabi H., Schöler H.R., Baharvand H. (2013). Inhibition of TGFβ signaling promotes ground state pluripotency. *Stem Cell Reviews and Reports* [Epub ahead of print] DOI 10.1007/s12015-013-9473-0.

Hara E., Uzman J.A., Dimri G.P., Nehlin J.O., Testori A., Campisi J. (1996). The helix-loop-helix protein Id-1 and a retinoblastoma protein binding mutant of SV40 T antigen synergize to reactivate DNA synthesis in senescent human fibroblasts. *Developmental Genetics* 18:161-172.

Hata A., Lagna G., Massagué J., Hemmati-Brivanlou A. (1997). Smad6 inhibits BMP/Smad1 signaling by specifically competing with the Smad4 tumour suppressor. *Genes & Development* 12:186-197.

Hawkins K., Mohamet L., Ritson S., Merry C.L.R., Ward C.M. (2012). E-cadherin and, in its absence, N-cadherin promotes Nanog expression in mouse embryonic stem cells via STAT3 phosphorylation. *Stem Cells* 30:1842-1851.

Hemmati-Brivanlou A., Melton D. (1997). Vertebrate embryonic cells will become nerve cells unless told otherwise. *Cell* 88:13-17.

Heng J.C.D., Feng B., Han J., Jiang J., Kraus P., Ng J.H., Orlov Y.L., Huss M., Yang L., Lufkin T., Lim B., Ng H.H. (2009). The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. *Cell Stem Cell* 6:167-174.

Hernandez M.C., Andres-Barquin P.J., Israel M. (1996). Molecular cloning of the cDNA encoding a helix-loop-helix protein, ID1B: tissue specific expression of ID1A and ID1B genes. *Biochimica et Biophysica Acta* 1308:28-30.

Herrera B., Inman G.J. (2009). A rapid and sensitive biomassa for the simultaneous measurement of multiple bone morphogenetic proteins. Identification and quantification of BMP4, BMP6 and BMP9 in bovine and human serum. *BMC Cell Biology* 10:20.

Hirano M., Hashimoto S., Yonemura S., Sabe H., Aizawa S. (2008). EPB41L5 functions to post-transcriptionally regulate cadherin and integrain during epithelial-mesenchymal transition. *The Journal of Cell Biology* 182:1217-1230.

Hollnagel A., Oehlmann V., Heymer J., Rüther U., Nordheim A. (1999). *Id* genes are direct targets of bone morphogenetic protein induction in embryonic stem cells. *The Journal of Biological Chemistry* 274:19838-19845.

Hoffmeyer K., Raggioli A., Rudloff S., Anton R., Hierholzer A., Del Valle I., Hein K., Vogt R., Kemler R. (2012). Wnt/β-catenin signaling regulates telomerase in stem cells and cancer cells. *Science* 336:1549-1554.

Hooker C.W., Hurlin P.J. (2006). Of Myc and Mnt. *Journal of Cell Science* 119:208-216.

Hooper M., Hardy K., Handyside A., Hunter S., Monk M. (1987). HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature* 326:292-295.

Hosler B.A., LaRosa G.J., Grippo J.F., Gudas L.J. (1989). Expression of *REX-1*, a gene containing zinc finger motifs, is rapidly reduced by retinoic acid in F9 teratocarcinoma cells. *Molecular and Cellular Biology* 9:5623-5629.

Hou P., Li Y., Zhang X., Liu C., Guan J., Li H., Zhao T., Ye J., Yang W., Liu K., Ge J., Xu J., Zhang Q., Zhao Y., Deng H. (2013). Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science* 651-654.

Hua X., Miller Z.A., Benchabane H., Wrana J.L., Lodish H.F. (2000). Synergism between transcription factors TFE3 and Smad3 in Transforming Growth Factor-β-induced transcription of the *Smad7* gene. *The Journal of Biological Chemistry* 275:33205-33208.

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

Huang Z.J., Edery I., Rosbash M. (1993). PAS is a dimerization domain common to *Drosophila* Period and several transcription factors. *Nature* 364:259-262.

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

Hutchins A.P., Choo S.H., Mistri T.K., Rahmani M., Woon C.T., Ng C.K.L., Jauch R., Robson P. (2013). Co-motif discovery identifies an Esrrb-Sox2-DNA ternary complex as a mediator of transcriptional differences between mouse embryonic and epiblast stem cells. *Stem Cells* 31:269-281.

Iavarone A., Garg P., Lasorella A., Hsu J., Israel M.A. (1994). The helix-loop-helix protein Id-2 enhances cell proliferation and binds to the retinoblastoma protein. *Genes & Development* 8:1270-1284.

Ichida J.K., Blanchard J., Lam K., Son E.Y., Chung J.E., Egli D., Loh K.M., Carter A.C., Di Giorgio F.P., Koszka K., Huangfu D, Akutsu H., Liu D.R., Rubin L.L., Eggan K. (2009). A small-molecule inhibitor of Tgf-β signaling replaces *Sox2* in reprogramming by inducing *Nanog*. *Cell Stem Cell* 5:491-503.

Illmensee K., Mintz B. (1976). Totipotency and normal differentiation of single teratocarcinoma cells cloned by injection into blastocysts. *Proceedings of the National Academy of Sciences of the United States of America* 73:549-553.

Imamura T., Takase M., Nishihara A., Oeda E., Hanai J., Kawabata M., Miyazono K. (1997). Smad6 inhibits signalling by the TGF-β superfamily. *Nature* 389:622-626.

Imayoshi I., Isomura A., Harima Y., Kawaguchi K., Kori H., Miyachi H., Fujiwara T., Ishidate F., Kageyama R. (2013). Oscillatory control of factors determining multipotency and fate in mouse neural progenitors. *Science* 342:1203-1208.

Inman K.E., Downs K.M. (2006). Localization of Brachyury (T) in embryonic and extraembryonic tissues during mouse gastrulation. *Gene Expression Patterns* 6:783-793.

Inman G.J., Nicolás F.J., Callahan J.F., Harling J.F., Gaster L.M., Reith A.D., Laping N.J., Hill C.S. (2002). SB-431542 is a potent and specific inhibitor of Transforming Growth Factor-β superfamily Type I Activin Receptor-Like Kinase (ALK) receptors ALK4, ALK5, and ALK7. *Molecular Pharmacology* 62:65-74.

Ivanova N., Dobrin R., Lu R., Kotenko I., Levorse J., DeCoste C., Schafer X., Lun Y., Lemischka I.R. (2006). Dissecting self-renewal in stem cells with RNA interference. *Nature* **442**:533-538.

Jacob F. (1978). The Leeuwenhoek Lecture, 1977: Mouse Teratocarcinoma and Mouse Embryo. *Proceedings of the Royal Society of London B* 201:249-270.

James D., Levine A.J., Besser D., Hemmati-Brivanlou A. (2005). TGFβ/activin/nodal signalling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development* 132:1273-1282.

Jedlicka P., Gutierrez-Hartmann A. (2008). Ets transcription factors in intestinal morphogenesis, homeostasis and disease. *Histology and Histopathology* 23:1417-1424.

Jen Y., Manova K., Benezra R. (1996). Expression patterns of *Id1*, *Id2*, and *Id3* are highly related but distinct from that of *Id4* during mouse embryogenesis. *Developmental Dynamics* 207:235-252.

Jen Y., Manova K., Benezra R. (1997). Each member of the Id gene family exhibits a unique expression pattern in mouse gastrulation and neurogenesis. *Developmental Dynamics* 208:92-106.

Jiang J., Chan Y.S., Loh Y.H., Cai J., Tong G.Q., Lim C.A., Robson P., Zhong S., Ng H.H. (2008). A core Klf circuitry regulates self-renewal of embryonic stem cells. *Nature Cell Biology* 10:353-360.

Jhas S., Ciura S., Belanger-Jasmin S., Dong Z., Llamosas E., Theriault F.M., Joachim K., Tang Y., Liu L., Liu J., Stifani S. (2006). Hes6 inhibits astrocyte differentiation and promotes neurogenesis through different mechanisms. *The Journal of Neuroscience* 26:11061-11071.

Jögi A., Persson P., Grynfeld A., Påhlman S., Axelson H. (2002). Modulation of basic helix-loop-helix transcription complex formation by Id proteins during neuronal differentiation. *The Journal of Biological Chemistry* 277:9118-9126.

Johansson B.M., Wiles M.V. (1995). Evidence for involvement of Activin A and Bone Morphogenetic Protein 4 in mammalian mesoderm development. *Molecular and Cellular Biology* 15:141-151.

Jones S. (2004). An overview of the basic helix-loop-helix proteins. *Genome Biology* 5:226.

Kageyama R., Ohtsuka T., Kobayashi T. (2007). The Hes gene family: repressors and oscillators that orchestrate embryogenesis. *Development* 134:1243-1251.

Kageyama R., Ohtsuka T., Kobayashi T. (2008). Roles of *Hes* genes in neural development. *Development, Growth & Differentiation* 50:S97-S103.

Kahan B.W., Ephrussi B. (1970). Developmental potentialities of clonal in vitro cultures of mouse testicular teratoma. *Journal of the National Cancer Institute* 44:1015-1036.

Kamiya D., Banno S., Sasai N., Ohgushi M., Inomata H., Watanabe K., Kawada M., Yakura R., Kiyonari H., Nakao K., Jakt L.M., Nishikawa S.I., Sasai Y. (2011). Intrinsic transition of embryonic stem-cell differentiation into neural progenitors. *Nature Neurosience* 470:503-509.

Kan N.G., Stemmler M.P., Junghans D., Kanzler B., de Vries W.N., Dominis M., Kemler R. (2007). Gene replacement reveals a specific role for E-cadherin in the formation of a functional trophectoderm. *Development* 164:31-41.

Kang Y., Chen C.R., Massagué J. (2003). A self-enabling TGFβ response coupled to stress signalling: Smad engages stress response factor ATF3 for *Id1* repression in epithelial cells. *Molecular Cell* 11:915-926.

Karwacki-Neisius V., Göke J., Osorno R., Halbritter F., Ng J.H., Weiβe A.Y., Wong F.C.K., Gagliardi A., Mullin N.P., Festuccia N., Colby D. Tomlinson S.R., Ng H.H., Chambers I. (2013) Reduced Oct4 expression directs a robust pluripotent state with distinct signaling activity and increased enhancer occupancy by Oct4 and Nanog. *Cell Stem Cell* 12:531-545.

Katagiri T., Imada M., Yanai T., Suda T., Takahashi N., Kamijo R. (2002). Identification of a BMP-responsive element in *Id1*, the gene for inhibition of myogenesis. *Genes to Cells* 7:949-960.

Katagiri T., Yamaguchi A., Komaki M., Abe E., Takahashi N., Ikeda T., Rosen V., Wozney J.M., Fujisawa-Sehara A., Suda T. (1994). Bone morphogenetic protein-2 converts the differentiation pathway of C2C12 myoblasts into the osteoblast lineage. *The Journal of Cell Biology* 127:1755-1766.

Kaufman M.H., Evans M.J., Robertson E.J., Bradley A. (1984). Influence of injected pluripotential (EK) cells on haploid and diploid parthenogenetic development. *Journal of Embryology and Experimental Morphology* 80:75-86.

Kawasaki H., Mizuseki K., Nishikawa S., Kaneko S., Kuwana Y., Nakanishi S., Nishikawa S.I., Sasai Y. (2000). Induction of midbrain dopaminergic neurons from ES cells by stromal cell-derived inducing activity. *Neuron* 28:31-40.

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

Keller G., Kennedy M., Papayannopoulou T., Wiles M.V. (1993). Haematopoietic commitment during embryonic stem cell differentiation in culture. *Molecular and Cellular Biology* 13:473-486.

Kewley R.J., Whitelaw M.L., Chapman-Smith A. (2004). The mammalian basic helix-loop-helix/PAS family of transcriptional regulators. *The International Journal of Biochemistry and Cell Biology* 36:189-204.

Kielman M.F., Rindapää M., Gaspar C., van Poppel N., Breukel C., van Leeuwen S., Taketo M.M., Roberts S., Smits R., Fodde R. (2002). *Apc* modulates embryonic stem-cell differentiation by controlling the dosage of β-catenin signaling. *Nature Genetics* 32:594-605.

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

Kim K., Doi A., Wen B., Ng K., Zhao R., Cahan P., Kim J., Aryee M.J., Ji H., Ehrlich L.I.R., Yabuuchi A., Takeuchi A., Cunniff K.C., Hongguang H., Mckinney-Freeman S., Naverias O., Yoon T.J., Irizarry R.A. Jung N., Seita J., Hanna J., Murakami P., Jaenisch R., Weissleder R., Orkin S.H., Weissman I.L., Feinberg A.P., Daley G.Q. (2010). Epigenetic memory in induced pluripotent stem cells. *Nature* 467:285-290.

Kim W., Kim M., Jho E. (2013). Wnt/β-catenin signalling: from plasma membrane to nucleus. *Biochemical Journal* 450:9-21.

Kimura-Yoshida C., Nakano H., Okamura D., Nakao K., Yonemura S., Belo J.A., Aizawa S., Matsui Y., Matsuo I. (2005). Canonical Wnt signaling and its antagonist regulate anterior-posterior axis polarization by guiding cell migration in mouse visceral endoderm. *Developmental Cell* 9:639-650.

Kinder S.J., Tsang T.E., Quinlan G.A., Hadjantonakis A.K., Nagy A., Tam P.P.L. (1999). The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo. *Development* 126:4691-4701.

Kirby D.R.S., Potts D.M., Wilson I.B. (1967). On the orientation of the implanting blastocyst. *Journal of Embryology and Experimental Morphology* 17:527-532.

Kirsch T., Sebald W., Dreyer M.K. (2000). Crystal structure of the BMP2-BRIA ectodomain complex. *Nature Structural Biology* 7:492-496.

Kleinsmith L.J., Pierce G.B. (1964). Multipotentiality of single embryonal carcinoma cells. *Cancer Research* 24:1544-1551.

Kobayashi T., Kageyama R. (2010). Hes1 regulates embryonic stem cell differentiation by suppressing Notch signaling. *Genes to Cells* 15:689-698.

Kobayashi T., Mizuno H., Imayoshi I., Furusawa C., Shirahige K., Kageyama R. (2009). The cyclic gene *Hes1* contributes to diverse differentiation responses of embryonic stem cells. *Genes & Development* 23:1870-1875.

Kondo M., Cubillo E., Tobiume K., Shirakihara T., Fukuda N., Suzuki H., Shimizu K., Takehara K., Cano A., Saitoh M., Miyazono K. (2004). A role for Id in the regulation of TGF-β-induced epithelial-mesenchymal transdifferentiation. *Cell Death and Differentiation* 11:1092-1101.

Korchynskyi O., ten Dijke P. (2002). Identification and functional characterization of distinct critically important bone morphogenetic protein specific response elements in the Id1 promoter. *The Journal of Biological Chemistry* 277:4883-4891.

Kretzchmar M., Liu F., Hata A., Doody J., Massagué J. (1997). The TGF-β family mediator Smad1 is phosphorylated directly and activated functionally by the BMP receptor kinase. *Genes & Development* 11:984-995.

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

Kunath T., Strumpf D., Rossant J. (2004). Early trophoblast determination and stem cell maintenance in the mouse - a review. *Placenta* 25 Supplement A:S32-S38.

Kuo C.T., Veselits M.L., Leiden J.M. (1997). LKLF: a transcriptional regulator of single-positive T cell quiescence and survival. *Science* 277:1986-1990.

Kwon G.S., Viotti M., Hadjantonakis A.K. (2008). The endoderm of the mouse embryo arises by dynamic widespread intercalation of embryonic and extraembryonic lineages. *Developmental Cell* 15:509-520.

Lagna G., Hata A., Hemmati-Brivanlou A., Massagué J. (1996). Partnership between DPC4 and SMAD proteins in TGF-β signalling pathways. *Nature* 383:832-836.

Langlands K., Yin X., Anand G., Prochownik E.V. (1997). Differential interactions of Id proteins with basic-helix-loop-helix transcription factors. *The Journal of Biological Chemistry* 272:19785-19793.

Larue L., Antos C., Butz S., Huber O., Delmas V., Dominis M., Kemler R. (1996). A role for cadherins in tissue formation. *Development* 122:3185-3194.

Larue L., Ohsugi M., Hirchenhain J., Kemler R. (1994). E-cadherin null mutant embryos fail to form a trophectoderm epithelium. *Proceedings of the National Academy of Sciences of the United States of America* 91:8263-8267.

Lawson K.A. (1999). Fate mapping the mouse embryo. *International Journal of Developmental Biology* 43:773-775.

Lawson K.A., Dunn N.R., Roelen B.A.J., Zeinstra L.M., Davis A.M., Wright C.V.E., Korving J.P.W.F.M., Hogan B.L.M. (1999). *Bmp4* is required for the generation of primordial germ cells in the mouse embryo. *Genes & Development* 13:424-436.

Ledent V., Paquet O., Vervoort M. (2002). Phylogenetic analysis of the human basic-helix-loop-helix proteins. *Genome Biology* 3:research0030.1-0030.18.

Lee J.E., Pintar J., Efstratiadis A. (1990). Pattern of the insulin-like growth factor II gene expression during mouse embryogenesis. *Development* 110:151-159.

Lee K.A., Gjorevski N., Boghaert E., Radisky D.C., Nelson C.M. (2011a). Snail1, Snail2 and E47 promote mammary epithelial branching morphogenesis. *The EMBO Journal* 30:2662-2674.

Lee K.L., Lim S.K., Orlov Y.L., Yit L.Y., Yang H., Ang L.T., Poellinger L., Lim B. (2011b). Graded Nodal/Activin signaling titrates conversion of quantitative phospho-Smad2 levels into qualitative embryonic stem cell fate decisions. *PLoS Genetics* 7:e1002130.

Lewis T.C., Prywes R. (2013). Serum regulation of Id1 expression by a BMP pathway and BMP responsive element. *Biochimica et Biophysica Acta* 1829:1147-1159.

Li L., Liu C., Biechele S., Zhu Q., Song L., Lanner F., Jing N., Rossant J. (2013). Location of transient ectodermal progenitor potential in mouse development. *Development* 140:4533-4543.

Li P., Tong C., Mehrian-Shai R., Jia L., Wu N., Yan Y., Maxson R.E., Schulze E.N., Song H., Hsieh C.L., Pera M.F., Ying Q.L. (2008). Germline competent embryonic stem cells derived from rat blastocysts. *Cell* 135:1299-1310.

Li V.S.W., Ng S.S., Boersema P.J., Low T.Y., Karthaus W.R., Gerlach J.P., Mohammed S., Heck A.J.R., Maurice M.M., Mahmoudi T., Clevers H. (2012). Wnt signaling through inhibition of β-catenin degradation in an intact Axin1 complex. *Cell* 149:1245-1256.

Li W., Nishimura R., Kashishian A., Batzer A.G., Kim W.J.H., Cooper J.A., Schlessinger J. (1994). A new function for a phosphotyrosine phosphatase: linking GRB2-Sos to a receptor tyrosine kinase. *Molecular and Cellular Biology* 14:509-517.

Li Y., McClintick J., Zhong L., Edenberg H.J., Yoder M.C., Chan R.J. (2005). Murine embryonic stem cell differentiation is promoted by SOCS-3 and inhibited by the zinc finger transcription factor Klf4. *Blood* 105:635-637.

Li Y., Zhang Q., Yin X., Yang W., Du Y., Hou P., Ge J., Liu C., Zhang W., Zhang X., Wu Y., Li H., Liu K., Wu C., Song Z., Zhao Y., Shi Y., Deng H. (2011). Generation of iPSCs from mouse fibroblasts with a single gene, *Oct4*, and small molecules. *Cell Research* 21:196-204.

Liang J., Wan M., Zhang Y., Gu P., Xin H., Jung S.Y., Qin J., Wong J., Cooney A.J., Liu D., Songyang Z. (2008). Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. *Nature Cell Biology* 10:731-739.

Lin G., Adiga U., Olson K., Guzowski J.F., Barnes C.A., Roysam B. (2003). A hybrid 3D watershed algorithm incorporating gradient cues and object models for automatic segmentation of nuclei in confocal image stacks. *Cytometry Part A* 56:23-36.

Lingbeck J.M., Trausch-Azar J.S., Ciechanover A., Schwartz A.L. (2005). E12 and E47 modulate cellular localization and proteasome-mediated degradation of MyoD and Id1. *Oncogene* 24:6376-6384.

Lints T.J., Parsons L.M., Hartley L., Lyons I., Harvey R.P. (1993). Nkx2.5: a novel homeobox gene expressed in early heart progenitor cells and their myogenic descendants. *Development* 119:419-431.

Liu P., Wakamiya M., Shea M.J., Albrecht U., Behringer R.R., Bradley A. (1999). Requirement for *Wnt3* in vertebrate axis formation. *Nature Genetics* 22:361-365.

Livigni A., Peradziryi H., Sharov A.A., Chia G., Hammachi F., Portero Migueles R., Sukparangsi W., Pernagallo S., Bradley M., Nichols J., Ko M.S.H., Brickman J.M. (2013). A conserved Oct4/POUV-dependent network links adhesion and migration to progenitor maintenance. *Current Biology* 23:2233-2244.

Lo R.S., Chen Y.G., Shi Y., Pavletich N.P., Massagué J. (1998). The L3 loop: a structural motif determining specific interactions between SMAD proteins and TGF-β receptors. *The EMBO Journal* 17:996-1005.

Loh Y.H., Wu Q., Chew J.L., Vega V.B., Zhang W., Chen X., Bourque G., George J., Leong B., Liu J., Wong K.Y., Sung K.W., Lee C.W.H., Zhao X.D., Chiu K.P., Lipovich L., Kuznetsov V.A., Robson P., Stanton L.W., Wei C.L., Ruan Y., Lim B., Ng H.H. (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nature Genetics* 38:431-440.

Longo A., Guanga G.P., Rose R.B. (2008). Crystal structure of E47-NeuroD1/Beta2 bHLH domain-DNA complex: heterodimer selectivity and DNA recognition. *Biochemistry* 47:218-229.

Lopes F.L., Desmarais J.A., Murphy B.D. (2004). Embryonic diapause and its regulation. *Reproduction* 128:669-678.

López-Rovira T., Chalaux E., Massagué K, Rosa J.L., Ventura F. (2002). Direct binding of Smad1 and Smad5 to two distinct motifs mediates bone morphogenetic protein-specific transcriptional activation of *Id1* gene. *The Journal of Biological Chemistry* 277:3176-3185.

Lovell-Badge R.H., Bygrave A.E., Bradley A., Robertson E., Evans M.J., Cheah K.S.E. (1985). Transformation of embryonic stem cells with the human type-II collagen gene and its expression in chimeric mice. *Cold Spring Harbor Symposia on Quantitative Biology* 50:707-711.

Loveys D.A., Streiff M.B., Kato G.J. (1996). E2A basic-helix-loop-helix transcription factors are negatively regulated by serum growth factor and by the Id3 protein. *Nucleic Acids Research* 24:2813-2820.

Lowell S., Benchoua A., Heavey B., Smith A.G. (2006). Notch promotes neural lineage entry by pluripotent embryonic stem cells. *PLoS Biology* 4:e121.

Luo J., Sladek R., Bader J.A., Matthyssen A., Rossant J., Giguère V. (1997). Placental abnormalities in mouse embryos lacking the orphan nuclear receptor ERR-β. *Nature* 388:778-782.

Lyden D., Young A.Z., Zagzag D., Yan W., Gerald W., O'Reilly R., Bader B.L., Hynes R.O., Zhuang Y., Manova K., Benezra R.(1999). Id1 and Id3 are required for neurogenesis, angiogenesis and vascularisation of tumour xenografts. *Nature* 401:670-677.

Ma Q., Fode C., Guillemot F., Anderson D.J. (1999). NEUROGENIN1 and NEUROGENIN2 control two distinct waves of neurogenesis in developing dorsal root ganglia. *Genes & Development* 13:1717-1728.

Ma Y.G., Rosfjord E., Huebert C., Wilder P., Tiesman J., Kelly D., Rizzino A. (1992). Transcriptional regulation of the murine k-FGF gene in embryonic cell lines. *Developmental Biology* 154:45-54.

Madabhushi M., Lacy E. (2011) Anterior visceral endoderm directs ventral morphogenesis and placement of head and heart via BMP2 expression. *Developmental Cell* 21:907-919.

Maherali N., Hochedlinger K. (2009). Tgfβ signal inhibition cooperates in the induction of iPSCs and replaces Sox2 and cMyc. *Current Biology* 19:1718-1723.

Mahmood R., Bresnik J., Hornbruch A., Mahony C., Morton N., Colquhoun K., Martin P., Lumsden A., Dickson C., Mason I. (1995). A role for FGF-8 in the initiation and maintenance of vertebrate limb bud outgrowth. *Current Biology* 5:797-806.

Makita J., Kurooka H., Mori K., Akagi Y., Yokota Y. (2006). Identification of the nuclear export signal in the helix-loop-helix inhibitor Id1. *FEBS Letters* 580:1812-1816.

Malaguti M., Nistor P.A., Blin G., Pegg A., Zhou X., Lowell S. (2013). Bone morphogenic protein signalling suppressed differentiation of pluripotent cells by maintaining expression of E-cadherin. *eLife* 2:e01197.

Marchand M., Schroeder I.S., Markossian S., Skoudy A., Nègre D., Cosset F.L., Real P., Kaiser C., Wobus A.M., Savatier P. (2009). Mouse ES cells over-expressing the transcription factor NeuroD1 show increased differentiation towards endocrine lineages and insulin-expressing cells. *The International Journal of Developmental Biology* 53:569-578.

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

Marson A., Levine S.S., Cole M.F., Frampton G.M., Brambrink T., Johnstone S., Guenther M.G., Johnston W.K., Wernig M., Newman J., Calabrese J.M., Dennis L.M., Volkert T.L., Gupta S., Love J., Hannett N., Sharp P.A., Bartel D.P., Jaenisch R., Young R.A. (2008). Connecting microRNA gnes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* 134:521-533.

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

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

Martin G.R., Evans M.J. (1974). The morphology and growth of a pluripotent teratocarcinoma cell line and its derivatives in tissue culture. *Cell* 2:163-172.

Martin G.R., Evans M.J. (1975). Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies *in vitro*. *Proceedings of the National Academy of Sciences of the United States of America* 72:1441-1445.

Maruyama M., Ichisaka T., Nakagawa M., Yamanaka S. (2005). Differential roles for Sox15 and Sox2 in transcriptional control in mouse embryonic stem cells. *The Journal of Biological Chemistry* 280:24371-24379.

Massari M.E., Murre C. (2000). Helix-loop-helix proteins: reglators of transcription in eucaryotic organisms. *Molecular and Cellular Biology* 20:429-440.

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

Mattaj I.W., Englmeier L. (1998). Nucleocytoplasmic transport: the soluble phase. *Annual Review of Biochemistry* 67:265-306.

McMahon J.A., Takada S., Zimmerman L.B., Fan C.M., Harland R.M., McMahon A.P. (1998). Noggin-mediated antagonism of BMP signaling is required for growth and patterning of the neural tube and somite. *Genes & Development* 12:1438-1452.

Meek S., Wei J., Sutherland L., Nilges B., Buehr M., Tomlinson S.R., Thomson A.J., Burdon T. (2013). Tuning of β-catenin activity is required to stabilize self-renewal of rat embryonic stem cells. *Stem Cells* 31:2104-2115.

Mesnard D., Guzman-Ayala M., Constam D.B. (2006). Nodal specifies embryonic visceral endoderm and sustains pluripotent cells in the epiblast before overt axial patterning. *Development* 133:2497-2505.

Míguez D.G., Gil-Guiñón E., Pons S., Martí E. (2013). Smad2 and Smad3 cooperate and antagonize simultaneously in vertebrate neurogenesis. *Journal of Cell Science* 126:5335-5343.

Mintz B., Illmensee K. (1975). Normal genetically mosaic mice produced from malignant teratocarcinoma cells. *Proceedings of the National Academy of Sciences of the United States of America* 72:3585-3589.

Mine N., Anderson R.M., Klingensmith J. (2008). BMP antagonism is required in both the node and lateral plate mesoderm for mammalian left-right axis establishment. *Development* 135:2425-2434.

Mishina Y., Suzuki A., Ueno N., Behringer R.R. (1995). *Bmpr* encoder a type I bone morphogenetic protein receptor that is essential for gastrulation during mouse embryogenesis. *Genes & Development* 9:3027-3037.

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

Mitchell P.J., Timmons P.M., Hébert J.M., Rigby P.W.J., Tijan R. (1991). Transcription factor AP-2 is expressed in neural crest cell lineages during mouse embryogenesis. *Genes & Development* 5:105-119.

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

Miura S., Davis S., Klingensmith J., Mishina Y. (2006). BMP signaling in the epiblast is required for proper recruitment of the prospective paraxial mesoderm and development of the somites. *Development* 133:3767-3775.

Miyoshi N., Ishii H., Nagano H., Haraguchi N., Dewi D.L., Kano Y., Nishikawa S., Tanemura M., Mimori K., Tanaka F., Saito T., Nishimura J., Takemasa I., Mizushima T., Ikeda M., Yamamoto H., Sekimoto M., Doki Y., Mori M. (2011). Reprogramming of mouse and human cells to pluripotency using mature microRNAs. *Cell Stem Cell* 8:633-638.

Mori S., Nishikawa S.I., Yokota Y. (2000). Lactation defect in mice lacking the helix-loop-helix inhibitor *Id2*. *The EMBO Journal* 19:5772-5781.

Morkel M., Huelsken J., Wakamiya M., Ding J., van de Wetering M., Clevers H., Taketo M.M., Behringer R.R., Shen M.M., Birchmeier W. (2003). β-Catenin regulates Cripto- and Wnt3-dependent gene expression programs in mouse axis and mesoderm formation. *Development* 130:6283-6294.

Mullin N.P., Yates A., Rowe A.J., Nijmeijer B., Colby D., Barlow P.N., Walkinshaw M.D., Chambers I. (2008). The pluripotency rheostat Nanog functions as a dimer. *Biochemical Journal* 411:227-231.

Muñoz Descalzo S., Rué P., Faunes F., Hayward P., Jakt L.M., Balayo T., Garcia-Ojalvo J., Martinez Arias A. (2013). A competitive protein interaction network buffers Oct4-mediated differentiation to promote pluripotency in embryonic stem cells. *Molecular Systems Biology* 9:694.

Muñoz Descalzo S., Rué P., Garcia-Ojalvo J., Martinez Arias A. (2012). Correlations between the levels of Oct4 and Nanog as a signature for naïve pluripotency in mouse embryonic stem cells. *Stem Cells* 30:2683-2691.

Muñoz-Sanjuán I., Brivanlou A.H. (2002). Neural induction, the default model and embryonic stem cells. *Nature Reviews Neuroscience* 3:271-280.

Murre C., McCaw P.S., Baltimore D. (1989a). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, *daughterless*, *MyoD*, and *myc* proteins. *Cell* 56:777-783.

Murre C., McCaw P.S., Vaessin H., Caudy M., Jan L.Y., Jan Y.N., Cabrera C.V., Buskin J.N., Hauschka S.D., Lassar A.B., Weintraub H., Baltimore D. (1989b) Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell* 58:537-544.

NagaiT., Ibata K., Park E.S., Kubota M., Mikoshiba K., Miyawaki A. (2002). A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nature Biotechnology* 20:87-90.

Nagaso H., Suzuki A., Tada M., Ueno N. (1999). Dual specificità of activin type II receptor ActRIIb in dorso-ventral patterning during zebrafish embryogenesis. *Development, Growth & Differentiation* 41:119-133.

Nagata Y., Shoji W., Obinata M., Todokoro K. (1995). Phosphorylation of helix-loop-helix proteins Id1, Id2 and Id3. *Biochemical and Biophysical Research Communications* 207:916-926.

Najm F.J., Chenoweth J.G., Anderson P.D., Nadeau J.H., Redline R.W., McKay R.D.G., Tesar P.J. (2011). Isolation of epiblast stem cells from preimplantation mouse embryos. *Cell Stem Cell* 8:318-325.

Nakamura H., Cook R.N., Justice M.J. (2013). Mouse *Tenm4* is required for mesoderm induction. *BMC Developmental Biology* 13:9.

Nakano T., Kodama H., Honjo T. (1994). Generation of lymphohematopoietic cells from embryonic stem cells in culture. *Science* 265:1098-1101.

Nakao A., Imamura T., Souchelnytskyi S., Kawabata M., Ishisaki A., Oeda E., Tamaki K., Hanai J., Heldin C.H., Miyazono K., ten Dijke P. (1997). TGF-β receptor-mediated signalling through Smad2, Smad4 and Smad4. *The EMBO Journal* 16:5353-5362.

Nakatake Y., Fukui N., Iwamatsu Y., Masui S., Takahashi K., Yagi R., Yagi K., Miyazaki J., Matoba R., Ko M.S.H., Niwa H. (2006). Klf4 cooperates with Oct3/4 and Sox2 to activate the *Lefty1* core promoter in embryonic stem cells. *Molecular and Cellular Biology* 26:7772-7782.

Nakaya M., Biris K., Tsukiyama T., Jaime S., Rawls J.A., Yamaguchi T.P. (2005). Wnt3a links left-right determination with segmentation and anteroposterior axis elongation. *Development* 132:5425-5436.

Nam H.S., Benezra R. (2009). High levels of Id1 expression define B1 type adult neural stem cells. *Cell Stem Cell* 5:515-526.

Navarro P., Festuccia N., Colby D., Gagliardi A., Mullin N.P., Zhang W., Karwacki-Neisius V., Osorno R., Kelly D., Robertson M., Chambers I. (2012). OCT4/SOX2-independent *Nanog* autorepression modulates heterogeneous *Nanog* gene expression in mouse ES cells. *The EMBO Journal* 31:4547-4562.

Navarro P., Oldfield A., Legoupi J., Festuccia N., Dubois A., Attia M., Schoorlemmer J., Rougeulle C., Chambers I., Avner P. (2010). Molecular coupling of *Tsix* regulation and pluripotency. *Nature* 468:457-460.

Nemir M., Croquelois A., Pedrazzini T., Radtke F. (2006). Induction of cardiogenesis in embryonic stem cells via downregulation of Notch1 signaling. *Circulation Research* 98:1471-1478.

Nichols J., Chambers I., Taga T., Smith A. (2001). Physiological rationale for responsiveness of mouse embryonic stem cells to gp130 cytokines. *Development* 128:2333-2339.

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

Nichols J., Smith A. (2012). Pluripotency in the embryo and in culture. *Cold Spring Harbor Perspectives in Biology* 4:a008128.

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

Nishikawa S.I., Nishikawa S., Hirashima M., Matsuyoshi N., Kodama H. (1998). Progressive lineage analysis by cell sorting and culture identifies $FLK1^{+}VE\text{-}cadherin^{+}$ cells at a diverging point of endothelial and hemopoietic lineages. *Development* 125:1747-1757.

Nishimoto M., Fukushima A., Okuda A., Muramatsu M. (1999). The gene for the embryonic stem cell coactivator UTF1 carries a regulatory element which selectively interacts with a complex composed of Oct-3/4 and Sox2. *Molecular and Cellular Biology* 19:5453-5465.

Niswander L., Martin G.R. (1992). *Fgf-4* expression during gastrulation, myogenesis, limb and tooth development. *Development* 114:755-768.

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

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

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

Niwa H., Toyooka Y., Shimosato D., Strumpf D., Takahashi K., Yagi R., Rossant J. (2005). Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* 123:917-929.

Niwa H., Yamamura K., Miyazaki J. (1991). Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* 108:193-200.

Norris D.P., Brennan J., Bikoff E.K., Robertson E.J. (2002). The Foxh1-dependent autoregulatory enhancer controls the level of Nodal signals in the mouse embryo. *Development* 129:3455-3468.

Norris D.P., Robertson E.J. (1999). Asymmetric and node-specific *nodal* expression patterns are controlled by two distinct *cis*-acting regulatory elements. *Genes & Development* 13:1575-1588.

Norton J.D. (2000). ID helix-loop-helix proteins in cell growth, differentiation and tumorigenesis. *Journal of Cell Science* 113:3897-3905.

Nowotschin S., Hadjantonakis A.K. (2010). Cellular dynamics in the early mouse embryo: from axis formation to gastrulation. *Current Opinion in Genetics & Development* 20:420-427.

Ogata T., Wozney J.M., Benezra R., Noda M. (1993). Bone morphogenetic protein 2 transiently enhances expression of a gene, Id (inhibitor of differentiation), encoding a helix-loop-helix molecule in osteoblast-like cells. *Proceedings of the National Academy of Sciences of the United States of America* 90:9219-9222.

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

Ogawa K., Saito A., Matsui H., Suzuki H., Ohtsuka S., Shimosato D., Morishita Y., Watabe T., Niwa H., Miyazono K. (2007). Activin-Nodal signaling is involved in propagation of mouse embryonic stem cells. *Journal of Cell Science* 120:55-65.

Ohi Y., Qin H., Hong C., Blouin L., Polo J.M., Guo T., Qi Z., Downey S.L., Manos P.D., Rossi D.J., Yu J., Hebrok M., Hochedlinger K., Costello J.F., Song J.S.,

Ramalho-Santos M. (2011). Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nature Cell Biology* 13:541-549.

Ohtani N., Zebedee Z., Huot T.J.G., Stinson J.A., Sugimoto M., Ohashi Y., Sharrocks A.D., Peters G., Hara E. (2001). Opposing effects of Ets and Id proteins on $p16^{INK4a}$ expression during cellular senescence. *Nature* 409:1067-1070.

Ohtsuka S., Nishikawa-Torikai S., Niwa H. (2012). E-cadherin promotes incorporation of mouse epiblast stem cells into normal development. *PLoS ONE* 7:e45220.

Ohtsuka T., Ishibashi M., Gradwohl G., Nakanishi S., Guillemot F., Kageyama R. (1999). *Hes1* and *Hes5* as Notch effectors in mammalian neuronal differentiation. *The EMBO Journal* 18:2196-2207.

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

Orsulic S., Huber O., Aberle H., Arnold S., Kemler R. (1999). E-cadherin binding prevents β-catenin nuclear localization and β-catenin/LEF-1-mediated transactivation. *Journal of Cell Science* 112:1237-1245.

Osorno R., Chambers I. (2011). Transcription factor heterogeneity and epiblast pluripotency. *Philosophical Transactions of the Royal Society B* 366:2230-2237.

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

Palmenberg A.C., Parks G.D., Hall D.J., Ingraham R.H., Seng T.W., Pallai P.V. (1992). Proteolytic processing of the cardioviral P2 region: primary 2A/2B cleavage in clone-derived precursors. *Virology* 190:754-762.

Palmieri S.L., Peter W., Hess H., Schöler H. (1994). Oct-4 transcription factor is differentially expressed in the mouse embryo during establishment of the first two extraembryonic cell lineages involved in implantation. *Developmental Biology* 166:259-267.

Pan L., Sato S., Frederick J.P., Sun X.H., Zhuang Y. (1999). Impaired immune responses and B-cell proliferation in mice lacking the *Id3* gene. *Molecular and Cellular Biology* 19:5969-5980.

Pang Z.P., Yang N., Vierbuchen T., Ostermeier A., Fuentes D.R., Yang T.Q., Citri A., Sebastiano V., Marro S., Südhof T.C., Wernig M. (2011). Induction of human neuronal cells by defined transcription factors. *Nature* 476:220-223.

Papaioannu V.E., Gardner R.L., McBurney M.W., Babinet C., Evans M.J. (1978). Participation of cultured teratocarcinoma cells in mouse embryogenesis. *Journal of Embryology and Experimental Morphology* 44:93-104.

Papaioannu V.E., McBurney M.W., Gardner R.L., Evans M.J. (1975). Fate of teratocarcinoma cells injected into early mouse embryos. *Nature* 258:70-73.

Parisi S., Passaro F., Aloia L., Manabe I., Nagai R., Pastore L., Russo T. (2008). Klf5 is involved in self-renewal of mouse embryonic stem cells. *Journal of Cell Science* 121:2629-2634.

Park C., Afrikanova I., Chung Y.S., Zhang W.J., Arentson E., Fong G., Rosendahl A., Choi K. (2004). A hierarchical order of factors in the generation of FLK1- and SCL-expressing hematopoietic and endothelial progenitors from embryonic stem cells. *Development* 131:2749-2762.

Padua D., Zhang X.H.F., Wang Q., Nadal C., Gerald W.L., Gomis R.R., Massagué J. (2008). TGFβ primes breast tumors for lung metastasis seeding through angiopoietin-like 4. *Cell* 133:66-77.

Park C., Afrikanova I., Chung Y.S., Zhang W.J., Arentson E., Fong G., Rosendahl A., Choi K. (2004). A hierarchical order of factors in the generation of FLK1- and SCL-expressing hematopoietic and endothelial progenitors from embryonic stem cells. *Development* 131:2749-2762.

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

Pera E.M., Ikeda A., Eivers E., De Robertis E.M. (2003). Integration of IGF, FGF and anti-BMP signals via Smad1 phosphorylation in neural induction. *Genes & Development* 17:3023-3028.

Pera E.M., Wessely O., Li S.Y., De Robertis E.M. (2001). Neural and head induction by insulin-like growth factor signals. *Developmental Cell* 1:655-665.

Pera M.F., Andrade J., Houssami S., Reubinoff B., Trounson A., Stanley E.G., Ward-van Oostwaard D., Mummery C. (2004). Regulation of human embryonic stem cell differentiation by BMP-2 and its antagonist noggin. *Journal of Cell Science* 117:1269-1280.

Perea-Gomez A., Lawson K.A., Rhinn M., Zakin L., Brûlet P., Mazan S., Ang S. (2001). Otx2 is required for visceral endoderm movement and for the restriction of posterior signals in the epiblast of the mouse embryo. *Development* 128:753-765.

Perea-Gómez A., Shawlot W., Sasaki H., Behringer R.R., Ang S. (1999). HNF3β and Lim1 interact in the visceral endoderm to regulate primitive streak formation and anterior-posterior polarity in the mouse embryo. *Development* 126:4499-4511.

Perea-Gomez A., Vella F.D.J., Shawlot W., Oulad-Abdelghani M., Chazaud C., Meno C., Pfister V., Chen L., Robertson E., Hamada H., Behringer R.R., Ang S.L. (2002). Nodal antagonists in the anterior visceral endoderm prevent the formation of multiple primitive streaks. *Developmental Cell* 3:745-756.

Pereira P.N.G., Dobreva M.P., Maas E., Cornelis F.M., Moya I.M., Umans L., Verfaillie C.M., Camus A., Chuva de Sousa Lopes S.M., Huylebroeck D., Zwijsen A. (2012). Antagonism of Nodal signaling by BMP/Smad5 prevents ectopic primitive streak formation in the mouse amnion. *Development* 139:3343-3354.

Pérez-Moreno M.A., Locascio A., Rodrigo I., Dhondt G., Portillo F., Nieto M.A., Cano A. (2001). A new role for E12/E47 in the repression of *E-cadherin* expression and epithelial-mesenchymal transitions. *The Journal of Biological Chemistry* 276:27424-27431.

Perk J., Gil-Bazo I., Chin Y., de Candia P., Chen J.J.S., Zhao Y., Chao S., Cheong W., Ke Y., Al-Ahmadie H., Gerald W.L., Brogi E., Benezra R. (2006). Reassessment of Id1 protein expression in human mammary, prostate, and bladder cancers using a monospecific rabbit monoclonal anti-Id1 antibody. *Cancer Research* 66:10870-10877.

Pesce S., Benezra R. (1993). The loop region of the helix-loop-helix protein Id1 is critical for its dominant negative activity. *Molecular and Cellular Biology* 13:7874-7880.

Pfendler K.C., Catuar C.S., Meneses J.J., Pedersen R.A. (2005). Overexpression of *Nodal* promotes differentiation of embryonic stem cells into mesoderm and endoderm at the expense of neuroectoderm formation. *Stem Cells and Development* 14:162-172.

Piek E., Afrakhte M., Sampath K., Van Zoelen E.J., Heldin C.H., ten Dijke P. (1999). Functional antagonism between Activin and Osteogenic Protein-1 in human embryonal carcinoma cells. *Journal of Cellular Physiology* 180:141-149.

Pollard S.M., Benchoua A., Lowell S. (2006). Neural stem cells, neurons and glia. *Methods in Enzymology* 418:151-169.

Polydorou C., Georgiades P. (2013). Ets2-dependent trophoblast signalling is required for gastrulation progression after primitive streak initiation. *Nature Communications* 4:1658.

Prendergast G.C., Lawe D., Ziff E.B. (1991). Association of Myn, the murine homolog of Max, with c-Myc stimulates methylation-sensitive DNA binding and Ras cotransformation. *Cell* 65:395-407.

Radice G.L., Rayburn H., Matsunami H., Knudsen K.A., Takeichi M., Hynes R.O. (1997). Developmental defects in mouse embryos lacking N-cadherin. *Developmental Biology* 181:64-78.

Radzisheuskaya A., Chia G.L.B., dos Santos R.L., Theunissen T., Castro L.F.C., Nichols J., Silva J.C.R. (2013). A defined Oct4 level governs cell state transitions of pluripotency entry and differentiation into all embryonic lineages. *Nature Cell Biology* 15:579-590.

Ramkumar N., Anderson K.V. (2011). SnapShot: Mouse primitive streak. *Cell* 146:488.

Rappolee D.A., Basilico C., Patel Y., Werb Z. (1994). Expression and function of FGF-4 in peri-implantation development in mouse embryos. *Development* 120:2259-2269.

Redmer T., Diecke S., Grigoryan T., Quiroga-Negreira A, Birchmeier W., Besser D. (2011). E-cadherin is crucial for embryonic stem cell pluripotency and can replace OCT4 during somatic cell reprogramming. *EMBO Reports* 12:720-726.

Reissman E., Jörnvall H., Blokzijl A., Andersson O., Chang C., Minchiotti G. Persico M.G., Ibáñez C.F., Brivanlou A.H. (2001). The orphan receptor ALK7 and the Activin receptor ALK4 mediate signaling by Nodal proteins during vertebrate development. *Genes & Development* 15:2010-2022.

Reubinoff B.E., Pera M.F., Fong C.Y., Trounson A., Bongso A. (2000). Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. *Nature Biotechnology* 18:399-404.

Rhinn M., Dierich A., Shawlot W., Behringer R.R., Le Meur M., Ang S.L. (1998). Sequential roles for Otx2 in visceral endoderm and neuroectoderm for foebrain and midbrain induction and specification. *Development* 125:845-856.

Ribeiro J.D., Morey L., Mas A., Gutierrez A., Luis N.M., Mjetta S., Richly H., Benitah S.A., Keyes W.M., Di Croce L. (2013). ZRF1 controls oncogene-induced senescence through the *INK4-ARF* locus. *Oncogene* 32:2161-2168.

Riechmann V., van Crüchten I., Sablitzky (1994). The expression pattern of *Id4*, a novel dominant negative helix-loop-helix protein, is distinct from *Id1*, *Id2* and *Id3*. *Nucleic Acids Research* 22:749-755.

Rivea-Pérez J.A., Magnuson T. (2005). Primitive streak formation in mice is preceded by localized activation of *Brachyury* and *Wnt3*. *Developmental Biology* 288:363-371.

Robb L., Tam P.P.L. (2004). Gastrula organiser and embryonic patterning in the mouse. *Seminars in Cell & Developmental Biology* 15:543-554.

Roberts E.C., Deed R.W., Inoue T., Norton J.D., Sharrocks A.D. (2001). Id helix-loop-helix proteins antagonize Pax transcription factor activity by inhibiting DNA binding. *Molecular and Cellular Biology* 21:524-533.

Robertson E., Bradley A., Kuehn N., Evans M. (1986). Germline transmission of genes introduced into cultured pluripotential cells by retroviral vector. *Nature* 323:445-448.

Rodda D.J., Chew J.L., Lim L.H., Loh Y.H., Wang B., Ng H.H., Robson P. (2005). Transcriptional regulation of *Nanog* by OCT4 and SOX2. *The Journal of Biological Chemistry* 280:24731-24737.

Rodriguez T.A., Srinivas S., Clements M.P., Smith J.C., Beddington R.S.P. (2005). Induction and migration of the anterior visceral endoderm is regulated by the extra-embryonic ectoderm. *Development* 132:2513-2520.

Rogers M.B., Hosler B.A., Gudas L.J. (1991). Specific expression of a retinoic acid-regulated, zinc-finger gene, Rex-1, in preimplantation embryos, trophoblast and spermatocytes. *Development* 113:815-824.

Romero-Lanman E.E., Pavlovic S., Amlani B., Chin Y., Benezra R. (2012). Id1 maintains embryonic stem cell self-renewal by up-regulation of Nanog and repression of Brachyury expression. *Stem Cells and Development* 21:384-393.

Rosenthal M.D., Wishnow R.M., Sato G.H. (1970). In vitro growth and differentiation of clonal populations of multipotential mouse cells derived from a transplantable testicular teratocarcinoma. *Journal of the National Cancer Institute* 44:1015-1036.

Rosner M.H., Vigano M.A., Ozato K., Timmons P.M, Poirier F., Rigby P.W.J., Staudt L.M. (1990). A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. *Nature* 345:686-692.

Ross S.E., Greenberg M.E., Stiles C.D. (2000). Basic helix-loop-helix factors in cortical development. *Neuron* 39:13-25.

Rossant J. (2004). Lineage development and polar asymmetries in the peri-implantation mouse blastocyst. *Seminars in Cell & Developmental Biology* 15:573-581.

Rossant J., Tam P.P.L. (2004). Emerging asymmetry and embryonic patterning in early mouse development. *Developmental Cell* 7:155-164.

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

Russ A.P., Wattler S., Colledge W.H., Aparicio S.A.J.R., Carlton M.B.L., Pearce J.J., Barton S.C., Surani M.A., Ryan K., Nehls M.C., Wilson V., Evans M.J. (2000). *Eomesodermin* is required for mouse trophoblast development and mesoderm formation. *Nature* 404:95-99.

Ruzinova M.B., Benezra R. (2003). Id proteins in development, cell cycle and cancer. *TRENDS in Cell Biology* 13:410-418.

Ryan M.D., King A.M.Q., Thomas G.P. (1991). Cleavage of foot-and-mouth disease virus polyprotein is mediated by residues located within a 19 amino acid sequence. *Journal of General Virology* 72:2727-2732.

Saisanit S., Sun X.H. (1995). A novel enhancer, the pro-B enhancer, regulates Id1 gene expression in progenitor B cells. *Molecular and Cellular Biology* 15:1513-1521.

Saisanit S., Sun X.H. (1997). Regulation of the pro-B-cell-specific enhancer of the *Id1* gene involves the C/EBP family of proteins. *Molecular and Cellular Biology* 17:844-850.

Sasai Y., Kageyama R., Tagawa Y., Shigemoto R., Nakanishi S. (1992). Two mammalian helix-loop-helix factors structurally related to *Drosophila hairy* and *Enhancer of split*. *Genes & Development* 6:2620-2634.

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

Schindelin J., Arganda-Carreras I., Frise E., Kaynig V., Longair M., Pietzsch T., Preibisch S., Rueden C., Saalfeld S., Schmid B., Tinevez J.Y., White D.J., Hartenstein V., Eliceiri K., Tomancak P., Cardona A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods* 9:676-682.

Schmid B., Schindelin J., Cardona A., Longair M., Heisenberg M. (2010). A high-level 3D visualization API for Java and ImageJ. *BMC Bioinformatics* 11:274.

Schroeder T., Meier-Stiegen F., Schwanbeck R., Eilken H., Nishikawa S., Häsler R., Schreiber S., Bornkamm G.W., Nishikawa S.I., Just U. (2006). Activated Notch1 alters differentiation of embryonic stem cells into mesodermal cell lineages at multiple stages of development. *Mechanisms of Development* 123:570-579.

Scotland K.B., Chen S., Sylvester R., Gudas L.J. (2009). Analysis of Rex1 (Zfp42) function in embryonic stem cell differentiation. *Developmental Dynamics* 238:1863-1877.

Segre J.A., Bauer C., Fuchs E. (1999). Klf4 is a transcription factor required for establishing the barrier function of the skin. *Nature Genetics* 22-356-360.

Shaner N.C., Steinbach P.A., Tsien R.Y. (2005). A guide to choosing fluorescent proteins. *Nature Methods* 2:905-909.

Shirakihara T., Saitoh M., Miyazono K. (2007). Differential regulation of epithelial and mesenchymal markers by δEF1 proteins in epithelial-mesenchymal transition induced by TGF-β. *Molecular Biology of the Cell* 18:3533-3544.

Shoji W. Inoue T., Yamamoto T., Obinata M. (1995). MIDA1, a protein associated with Id, regulates cell growth. *The Journal of Biological Chemistry* 270:24818-24825.

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

Silva S.S., Rowntree R.K., Mekhoubad S., Lee J.T. (2008). X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America* 105:4820-4825.

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

Singla D.K., Schneider D.J., LeWinter M.M., Sobel B.E. (2006). wnt3a but not wnt11 supports self-renewal of embryonic stem cells. *Biochemical and Biophysical Research Communications* 345:789-795.

Skinner M.K., Rawls A., Wilson-Rawls J., Roalson E.H. (2010). Basic helix-loop-helix transcription factor gene family phylogenetics and nomenclature. *Differentiation* 80:1-8.

Slattery C., McMorrow T. (2008). E2A proteins: regulators of cell phenotype in normal physiology and disease. *The International Journal of Biochemistry & Cell Biology* 40:1431-1436.

Slattery C., McMorrow T., Ryan M.P. (2006). Overexpression of E2A proteins induces epithelial-mesenchymal transition in human renal proximal tubular epithelial cells suggesting a potential role in renal fibrosis. *FEBS Letters* 580:4021-4030.

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

Smith A.G., Hooper M.L. (1987). Buffalo rat liver cells produce a diffusible activity which inhibits the differentiation of murine embryonal carcinoma and embryonic stem cells. *Developmental Biology* **121**:1-9.

Smith J.L., Gesteland K.M., Schoenwolf G.C. (1994). Prospective fate map of the mouse primitive streak at 7.5 days of gestation. *Developmental Dynamics* 201:279-289.

Smith L.J. (1980). Embryonic axis orientation in the mouse and its correlation with blastocyst relationships to the uterus. Part 1. Relationships between 82 hours and 4¼ days. *Journal of Embryology and Experimental Morphology* 55:257-277.

Smith L.J. (1985). Embryonic axis orientation in the mouse and its correlation with blastocyst relationships to the uterus. II. Relationships between 4¼ to 9½ days. *Journal of Embryology and Experimental Morphology* 89:15-35.

Soares M.L., Torres-Padilla M.E., Zernicka-Goetz M. (2008). Bone morphogenetic protein 4 signaling regulates development of the anterior visceral endoderm in the mouse embryo. *Development, Growth & Differentiation* 50:615-621.

Solloway M.J., Robertson E.J. (1999). Early embryonic lethality in Bmp5;Bmp7 double mutant mice suggests functional redundancy within the 60A subgroup. *Development* 125:1753-1765.

Soncin F., Mohamet L., Eckardt D., Ritson S., Eastham A.M., Bobola N., Russell A., Davies S., Kemler R., Merry C.L.R., Ward C.M. (2009). Abrogation of E-cadherin-mediated cell-cell contact in mouse embryonic stem cells results in reversible LIF-independent self-renewal. *Stem Cells* 27:2069-2080.

Soncin F., Mohamet L., Ritson S., Hawkins K., Bobola N., Zeef L., Merry C.L.R., Ward C.M. (2011). E-cadherin acts as a regulator of transcripts associated with a wide range of cellular processes in mouse embryonic stem cells. *PLoS ONE* 6:e21463.

Stahl N., Boulton T.G., Farruggella T., Ip N.Y., Davis S., Witthuhn B.A., Quelle F.W., Silvenhoinen O., Barbieri G., Pellegrini S., Ihle J.N., Yancopoulos G.D. (1994). Association and activation of Jak-Tyk kinases by CNTF-LIF-OSM-IL-6 beta receptor components. *Science* 263:92-95.

Stahl N., Farrugella T.J., Boulton T.G., Zhong Z., Darnell J.E., Yancopoulos G.D. (1995). Choice of STATs and other substrates specified by modular tyrosine-based motifs in cytokine receptors. *Science* 267:1349-1353.

Stankic M., Pavlovic S., Chin Y., Brogi E., Padua D., Norton L., Massagué J., Benezra R. (2013). TGF-β-Id1 signaling opposes Twist1 and promotes metastatic colonization via a mesenchymal-to-epithelial transition. *Cell Reports* 5:1228-1242.

Stavridis M.P., Lunn J.S., Collins B.J., Storey K.G. (2007). A discrete period of FGF-induced Erk1/2 signalling is required for vertebrate neural specification. *Development* 134:2889-2894.

Stemmler M.P. (2008). Cadherins in development and cancer. *Molecular BioSystems* 4:835-850.

Stemmler M.P., Hecht A., Kemler R. (2005). E-cadherin intron 2 contains cis-regulatory elements essential for gene expression. *Development* 132:965-976.

Stephenson R.O., Yamanaka Y., Rossant J. (2010). Disorganized epithelial polarity and excess trophectoderm cell fate in preimplantation embryos lacking E-cadherin. *Development* 137:3383-3391.

Stern C. (1936). Somatic crossing over and segregation in Drosophila melanogaster. *Genetics* 21:625-730.

Stern C.D. (2006). Neural induction: 10 years on since the ‘default model’. *Current Opinion in Cell Biology* 18:692-697.

Sterneckert J., Stehling M., Bernemann C., Araúzo-Bravo M.J., Greber B., Gentile L., Ortmeier C., Sinn M., Wu G., Ruau D., Zenke M., Brintrup R., Klein D.C., Ko K., Schöler H.R. (2010). Neural induction intermediates exhibit distinct roles of Fgf signalling. *Stem Cells* 28:1772-1781.

Stevens J.D., Roalson E.H., Skinner M.K. (2008). Phylogenetic and expression analysis of the basic helix-loop-helix transcription factor gene family: genomic approach to cellular differentiation. *Differentiation* 76:1006-1022.

Stevens L.C., Little C.C. (1954). Spontaneous testicular teratomas in an inbred strain of mice. *Proceedings of the National Academy of Sciences of the United States of America* 40:1080-1087.

Stewart C.L., Kaspar P., Brunet L.J., Bhatt H., Gadi I., Köntgen F., Abbondanzo S.J. (1992). Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. *Nature* 359:76-79.

Streit A., Berliner A.J., Papanayotou C., Sirufnik A., Stern C.D. (2000). Initiation of neural induction by FGF signalling before gastrulation. *Nature* 406:74-78.

Streit A., Lee K.J., Woo I., Roberts C., Jessell T.M., Stern C.D. (1998). Chordin regulates primitive streak development and the stability of induced neural cells, but is not sufficient for neural induction in the chick embryo. *Development* 125:507-519.

Sudheer S., Bhushan R., Fauler B., Lehrach H., Adjaye J. (2012). FGF inhibition directs BMP4-mediated differentiation of human embryonic stem cells to syncytiotrophoblast. *Stem Cells and Development* 21:2987-3000.

Sun H., Ghaffari S., Taneja R. (2007). bHLH-Orange transcription factors in development and cancer. *Translational Oncogenomics* 2:107-120.

Sun X., Meyers E.N. Lewandoski M., Martin G.R. (1999). Targeted disruption of *Fgf8* causes failure of cell migration in the gastrulationg mouse embryo. *Genes & Development* 13:1834-1846.

Sun X.H., Copeland N.G., Jenkins N.A., Baltimore D. (1991). Id proteins Id1 and Id2 selectively inhibit DNA binding by one class of helix-loop-helix proteins. *Molecular and Cellular Biology* 11:5603-5611.

Suzuki A., Chang C., Yingling J.M., Wang X.F., Hemmati-Brivanlou A. (1997). Smad5 induces ventral fates in *Xenopus* embryo. *Developmental Biology* 184:402-405.

Suzuki A. Raya Á., Kawakami Y., Morita M., Matsui T., Nakashima K., Gage F.H., Rodríguez-Esteban C., Izpisúa Belmonte J.C. (2006). Nanog binds to Smad1 and blocks bone morphogenetic protein-induced differentiation of embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America* 103:10294-10299.

Tada M., Smith J.C. (2000). *Xwnt11* is a target of *Xenopus* Brachyury: regulation of gastrulation movements via Dishevelled, but not through the canonical Wnt pathway. *Development* 127:2227-2238.

Takahashi K., Tanabe K., Ohnuki M., Narita M., Ichisaka T., Tomoda K., Yamanaka S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131:861-872.

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

Takahashi-Tezuka M., Yoshida Y., Fukada T., Ohtani T., Yamanaka Y., Nishida K., Nakajima K., Hibi M., Hirano T. (1998). Gab1 acts as an adapter molecule linking the cytokine receptor gp130 to ERK mitogen-activated protein kinase. *Molecular and Cellular Biology* 18:4109-4117.

Tam P.P.L., Behringer R.R. (1997). Mouse gastrulation: the formation of a mammalian body plan. *Mechanisms of Development* 68:3-25.

Tam P.P.L., Loebel D.A.F. (2007). Gene function in mouse embryogenesis: get set for gastrulation. *Nature Reviews Genetics* 8:368-381.

Takase M., Imamura T., Sampath T.K., Takeda K., Ichijo H., Miyazono K., Kawabata M. (1998). Induction of Smad6 mRNA by Bone Morphogenetic Proteins. *Biochemical and Biophysical Research Communiations* 244:26-29.

Tang F., Barbacioru C., Bao S., Lee C., Nordman E., Wang X., Lao K., Surani M.A. (2010). Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis. *Cell Stem Cell* 6:468-478.

Tarkowski A.K. (1959). Experiments on the development of isolated blastomeres of mouse eggs. *Nature* 184:1286-1287.

Tarkowski A.K., Suvińska A., Czolowska R., Ożdżeński W. (2010). Individual blastomeres of 16- and 32-cell mouse embryos are able to develop into foetuses and mice. *Developmental Biology* 348:190-198.

Tarkowski A.K., Wróblewska J. (1967). Development of blastomeres of mouse eggs isolated at the 4- and 8-cell stage. *Journal of Embryology and Experimental Morphology* 18:155-180.

Telford N.A., Hogan A., Franz C.R., Schultz G.A. (1990). Expression of genes for insulin and insulin-like growth factors and receptors in early postimplantation mouse embryos and embryonal carcinoma cells. *Molecular Reproduction and Development* 27:81-92.

Tesar P.J., Chenoweth J.G., Brook F.A., Davies T.J., Evans E.P., Mack D.L., Gardner R.L., McKay R.D.G. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 448:196-199.

ten Berge D., Kurek D., Blauwkamp T., Koole W., Maas A., Eroglu E., Siu R.K., Nusse R. (2011). Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. *Nature Cell Biology* 13:1070-1075.

Thisse B., Thisse C. (2005). Functions and regulations of fibroblast growth factor signaling during embryonic development. *Developmental Biology* 287:390-402.

Thomas K.R., Capecchi M.R. (1986). Introduction of homologous DNA sequences into mammalian cells induces mutations in the cognate gene. *Nature* 324:34-38.

Thomson J.A., Itskovitz-Eldor J., Shapiro S.S., Waknitz M.A., Swiergiel J.J., Marshall V.S., Jones J.M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science* 282:1145-1147.

Tomioka M., Nishimoto M., Miyagi S., Katayanagi T., Fukui N., Niwa H., Muramatsu M., Okuda A. (2002). Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex. *Nucleic Acids Residues* **30**:3202-3213.

Tournay O., Benezra R. (1996). Transcription of the dominant-negative helix-loop-helix protein Id1 is regulated by a protein complex containing the immediate-early response gene *Egr*-1. *Molecular and Cellular Biology* 16:2418-2430.

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

Trausch-Azar J.S., Lingbeck J., Ciechanover A., Schwartz A.L. (2004). Ubiquitin-proteasome-mediated degradation of Id1 is mediated by MyoD. *The Journal of Biological Chemistry* 31:32614-32619.

Tremblay K.D., Dunn N.R., Robertson E.J. (2001). Mouse embryos lacking Smad1 signals display defects in extra-embryonic tissues and germ cell formation. *Development* 128:3609-3621.

Trott J., Martinez Arias A. (2013). Single cell lineage analysis of mouse embryonic stem cells at the exit from pluripotency. *Biology Open* 2:1049-1056.

Tzouanacou E., Wegener A., Wymeersch F.J., Wilson V., Nicolas J.F. (2009). Redefining the progression of lineage segregations during mammalian embryogenesis by clonal analysis. *Developmental Cell* 17:365-376.

Vallier L., Mendjan S., Brown S., Chng Z., Teo A., Smithers L.E., Trotter M.W.B., Cho C.H.H., Martinez A., Rugg-Gunn P., Brons G., Pedersen R.A. (2009). Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. *Development* 136:1339-1349.

van den Berg D.L.C., Snoek T., Mullin N.P., Yates A., Bezstarosti K., Demmers J., Chambers I., Poot R.A. (2010). An Oct4-centered protein interaction network in embryonic stem cells. *Cell Stem Cell* 6:369-381.

van den Berg D.L.C., Zhang W., Yates A., Engelen E., Takacs K., Bezstarosti K., Demmers J., Chambers I., Poot R.A. (2008). Estrogen-related receptor beta interacts with Oct4 to regulate *Nanog* gene expression. *Molecular and Cellular Biology* 28:5986-5995.

Varlet I., Collignon J., Robertson E.J. (1997). *nodal* expression in the primitive endoderm is required for specifying the anterior axis during mouse gastrulation. *Development* 124:1033-1044.

Villares R., Cabrera C.V. (1987). The *achaete-scute* gene complex of D. melanogaster: conserved domains in a subset of genes required for neurogenesis and their homology to *myc*. *Cell* 50:415-424.

Villegas S.N., Canham M., Brickman J.M. (2010). FGF signalling as a mediator of lineage transitions – Evidence from embryonic stem cell differentiation. *Journal of Cellular Biochemistry* 110:10-20.

Wadman M. (2013). High Court ensures continued US funding of human embryonic-stem-cell research. *Nature* doi:10.1038/nature.2013.12171.

Waldrip W.R., Bikoff E.K., Hoodless P.A., Wrana J.L., Robertson E.J. (1998). Smad2 signaling in extraembryonic tissues determines anterior-posterior polarity of the early mouse embryo. *Cell* 92:797-808.

Wang J., Levasseur D.N., Orkin S.H. (2008). Requirement of Nanog dimerization for stem cell self-renewal and pluripotency. *Proceedings of the National Academy of Sciences of the United States of America* 105:6326-6331.

Wang K., Rao S., Chu J., Shen X., Levasseur D.N., Theunissen T.W., Orkin S.H. (2006). A protein interaction network for pluripotency of embryonic stem cells. *Nature* 444:364-368.

Wang Y., Benezra R., Sassoon D.A. (1992). Id expression during mouse development: a role in morphogenesis. *Developmental Dynamics* 194:222-230.

Watanabe K., Kamiya D., Nishiyama A., Katayama T., Nozaki S., Kawasaki H., Watanabe Y., Mizuseki J., Sasai Y. (2005). Directed differentiation of telencephalic precursors from embryonic stem cells. *Nature Neuroscience* 8:288-296.

Weber D., Kotzsch A., Nickel J., Harth S., Seher A., Mueller U., Sebald W., Mueller T.D. (2007). A silent H-bond can be mutationally activated for high-affinity interaction of BMP-2 and activin type IIB receptor. *BMC Structural Biology* 7:6.

Wen Z., Zhong Z., Darnett J.E. (1995). Maximal activation of transcription by Stat1 and Stat3 requires both tyrosine and serine phosphorylation. *Cell* 80:241-250.

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

Whyte W.A., Orlando D.A., Hnisz D., Abraham B.J., Lin C.Y., Kagey M.H., Rahl P.B., Lee T.I., Young R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153:307-319.

Wibley J., Deed R., Jasiok M., Douglas K., Norton J. (1996). A homology model of the Id-3 helix-loophelix domain as a basis for structure-function predictions. *Biochimica et Biophysica Acta* 1294:138-146.

Wilder P.J., Kelly D., Brigman K., Peterson C.L., Nowling T., Gao Q.S., McComb R.D., Capecchi M.R., Rizzino A. (1997). Inactivation of the FGF-4 gene in embryonic stem cells alters the growth and/or the survival of their early differentiated progeny. *Developmental Biology* 192:614-629.

Wilkinson D.G., Bhatt S., Hermann B.G. (1990). Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature* 343:657-659.

Williams R.L., Hilton D.J., Pease S., Willson T.A., Stewart C.L., Gearing D.P., Wagner E.F., Metcalf D., Nicola N.A., Gough N.M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* **336**:684-688.

Willis S.A., Zimmerman C.M., Li L., Mathews L.S. (1996). Formation and activation by phosphorylation of Activin receptor complexes. *Molecular Endocrinology* 10:367-379.

Willison K. (1990) The mouse Brachyury gene and mesoderm formation. *Trends in Genetics* 6:104-106.

Wilson P.A., Hemmati-Brivanlou A. (1995). Induction of epidermis and inhibition of neural fate by Bmp-4. *Nature* 376:331-333.

Wilson V., Beddington R. (1997). Expression of T protein in the primitive streak is necessary and sufficient for posterior mesoderm movement and somite differentiation. *Developmental Biology* 192:45-58.

Wilson V., Beddington R.S.P. (1996). Cell fate and morphogenetic movement in the late mouse primitive streak. *Mechanisms of Development* 55:79-89.

Wilson V., Manson L., Skarnes W.C., Beddington R.S.P. (1995). The *T* gene is necessary for normal mesodermal morphogenetic movements during gastrulation. *Development* 121:877-886.

Winnier G., Blessing M., Labosky P.A., Hogan B.L.M. (1995). Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. *Genes & Development* 9:2105-2116.

Wrana J.L., Attisano L., Wieser R., Ventura F., Massagué J. (1994). Mechanism of activation of the TGF-β receptor. *Nature* 370:341-347.

Wray J., Kalkan T., Smith A.G. (2009). The ground state of pluripotency. *Biochemical Society Transactions* 38:1027-1032.

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

Wu D., Pan W. (2009). GSK3: a multifaceted kinase in Wnt signaling. *Trends in Biochemical Sciences* 35:161-168.

Wu R.Y., Zhang Y., Feng X.H., Derynck R. (1997). Heteromeric and homomeric interactions correlate with signaling activity and functional cooperativity of Smad3 and Smad4/DPC4. *Molecular and Cellular Biology* 17:2521.

Wu W., Glinka A., Delius H., Niehrs C. (2000). Mutual antagonism between *dickkopf1* and *dickkopf2* regulates Wnt/β-catenin signalling. *Current Biology* 10:1611-1614.

Xi Q., Wang Z., Zaromytidou A.I., Zhang X.H.F., Chow-Tsang L.F., Liu J.X., Kim H., Barlas A., Manova-Todorova K., Kaartinen V., Studer L., Mark W., Patel D.J., Massagué J. (2011). A poised chromatin platform for TGF-β access to master regulators. *Cell* 147:1511-1524.

Xu C., Rosler E., Jiang J., Lebkowski J.S., Gold J.D., O'Sullivan C., Delavan-Boorsma K., Mok M., Bronstein A., Carpenter M.K. (2005). Basic fibroblast growth factor supports undifferentiated human embryonic stem cell growth without conditioned medium. *Stem Cells* 23:315-323.

Xu R.H., Chen X., Li D.S., Li R., Addicks G.C., Glennon C., Zwaka T.P., Thomson J.A. (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nature Biotechnology* 20:1261-1264.

Xu R.H., Sampsell-Barron T.L., Gu F., Root S., Peck R.M., Pan G., Yu J., Antosiewicz-Bourget J., Tian S., Stewart R., Thomson J.A. (2008). *NANOG* is a direct target of TGFβ/Activin-mediated SMAD signalling in human ESCs. *Cell Stem Cell* 3:196-206.

Yamaguchi T.P., Harpal K., Henkemeyer M., Rossant J. (1994). *fgfr-1* is required for embryonic growth and mesodermal patterning during mouse gastrulation. *Genes & Development* 8:3032-3044.

Yamaguchi T.P., Takada S., Yoshikawa Y., Wu N., McMahon A.P. (1999). *T* (Brachyury) is a direct target of Wnt3a during paraxial mesoderm specification. *Genes & Development* 13:3185-3190.

Yamamoto M., Beppu H., Takaoka K., Meno C., Li E., Miyazono K., Hamada H. (2009). Antagonism between Smad1 and Smad2 signaling determines the site of distal visceral endoderm formation in the mouse embryo. *The Journal of Cell Biology* 184:323-334.

Yamamoto M., Saijoh Y., Perea-Gomez A., Shawlot W., Behringer R.R., Ang S.L., Hamada H., Meno C. (2004). Nodal antagonists regulate formation of the anteroposterior axis of the mouse embryo. *Nature* 428:387-392.

Yan W., Young A.Z., Soares V.C., Kelley R., Benezra R., Zhuang Y. (1997). High incidence of T-cell tumors in E2A-null mice and E2A/Id1 double-knockout mice. *Molecular and Cellular Biology* 17:7317-7327.

Yang Y.P., Anderson R.M., Klingensmith J. (2010). BMP antagonism protects Nodal signalling in the gastrula to promote the tissue interactions underlying mammalian forebrain and craniofacial patterning. *Human Molecular Genetics* 19:3030-3042.

Yi F., Pereira L., Hoffman J.A., Shy B.R., Yuen C.M., Liu D.R., Merrill B.J. (2011). Opposing effects of Tcf3 and Tcf1 control Wnt stimulation of embryonic stem cell self-renewal. *Nature Cell Biology* 13:762-770.

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

Ying Q.L., Stavridis M., Griffiths D., Li M., Smith A. (2003b). Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nature Biotechnology* 21:183-186.

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

Yokota Y., Mansouri A., Mori S., Sugawara S., Adachi S., Nishikawa S.I., Gruss P. (1999). Development of peripheral lymphoid organs and natural killer cells depends on the helix-loop-helix inhibitor Id2. *Nature* 397:702-706.

Yu P., Pan G., Yu J., Thomson J.A. (2011). FGF2 sustains *NANOG* and switches the outcome of BMP4-induced human embryonic stem cell differentiation. *Cell Stem Cell* 8:326-334.

Yu P.B., Deng D.Y., Lai C.S., Hong C.C., Cuny G.D., Bouxsein M.L., Hong D.W., McManus P.M., Katagiri T., Sachidanandan C., Kamiya N., Fukuda T., Mishina Y., Peterson R.T., Bloch K.D. (2008). BMP type I receptor inhibition reduces heterotopic ossification. *Nature Medicine* 14:1363-1369.

Yuan H., Corbi N., Basilico C., Daley L. (1995). Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Genes & Development* 9:2635-2645.

Yun K., Mantani A., Garel S., Rubenstein J., Israel M.A. (2004). *Id4* regulates neural progenitor proliferation and differentiation in vivo. *Development* 131:5441-5448.

Zakin L., De Robertis E.M. (2004). Inactivation of mouse Twisted gastrulation reveals its role in promoting Bmp4 activity during forebrain development. *Development* 131:413-424.

Zhang K., Dai X., Wallingford M.C., Mager J. (2013). Depletion of Suds3 reveals an essential role in early lineage specification. *Developmental Biology* 373:359-372.

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

Zhang P., Andrianakos R., Yang Y., Liu C., Lu W. (2010). Kruppel-like Factor 4 (Klf4) prevents embryonic stem cell differentiation by regulating Nanog gene expression. *The Journal of Biological Chemistry* 285:9180-9189.

Zhang X., Zhang J., Wang T., Esteban M.A., Pei D. (2008). *Esrrb* activates *Oct4* transcription and sustains self-renewal and pluripotency in embryonic stem cells. *The Journal of Biological Chemistry* 283:35825-35833.

Zhou X., Smith A.J.H., Waterhouse A., Blin G., Malaguti M., Lin C.Y., Osorno R., Chambers I., Lowell S. (2013). Hes1 desynchronyzes differentiation of pluripotent cells by modulating STAT3 activity. *Stem Cells* doi:10.1002/stem.1426.

Zhu W., Kim J., Cheng C., Rawlins B.A., Boachie-Adjei O., Crystal R.G., Hidaka C. (2006). Noggin regulation of bone morphogenetic protein (BMP) 2/7 heterodimer activity in vitro. *Bone* 39:61-71.

Zimmerman L.B., De Jesús-Escobar J.M., Harland R.M. (1996). The Spemann Organizer signal noggin binds and inactivates bone morphogenetic protein 4. *Cell* 86:599-606.