THE UNIVERSITY *of* EDINBURGH

SIGNALLING AND TRANSCRIPTIONAL REGULATION OF EARLY DEVELOPMENTAL LINEAGE DECISIONS

SOPHIE M. MORGANI

PhD
The University of Edinburgh
2014

To my family

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Josh Brickman for his guidance, support, patience and tutoring throughout my PhD in his lab, for playing such a significant role in my development as a scientist and challenging me to work to my full potential. Our scientific and non-scientific discussions have always renewed my enthusiasm during tough times. I would also like to thank all of my colleagues that have offered both technical and moral support throughout this time, both at the University of Edinburgh CRM and the Danstem Centre at the University of Copenhagen. Especially I would like to give a huge thank you to Dr. Maurice Canham for guiding me through the initial stages of this work, for his immense patience and dedication to helping and teaching me. Without this I would have been completely lost.

Thank you to our collaborators in the Lis lab at the University of Cornell and in the Ko lab at the NIA, without which this work would not be possible, particularly to Alexei Sharov for his incredible scientific insight. A big thank you to my viva examiners, Richard Meehan and Elizabeth Robertson for taking a lot of time to read my thesis and making my examination an extremely enjoyable one.

Additionally a big thank you to the flow cytometry and tissue culture technical staff, Simon Monard, Gelo de la Cruz and Jonathan Rans for making sure that everything always ran smoothly. Thank you to the transgenic facilities both at The University of Edinburgh and Copenhagen, Lynsey Robertson, Sally Inverarity, Javier Martin Gonzalez and Kasper Bonderup, for teaching me everything that I know about embryos and mice and for being incredibly supportive friends.

Most importantly, thank you to my close friends and family that have supported me throughout these years and been a constant motivation for me to reach the end of this period of my career and life.

This work was supported by the MRC and the Novo Nordisk Foundation.

DECLARATION

I declare that:

1. I composed this thesis.
2. The work presented in this thesis is of my own, unless otherwise stated.
3. This work has not been submitted for any other degree of professional qualification.

Sophie Morgani

ABSTRACT

Embryonic stem (ES) cells are cell lines isolated from the embryo at a time just prior to implantation into the uterus. In the right cocktail of medium and cytokines, these cell lines can be maintained indefinitely *in vitro* in a self-renewing state. Initially it was assumed that these cells represented a homogeneous population however, more recently it has been shown that there are a great number of genes that are expressed heterogeneously. ES cell cultures are therefore a mix of different subpopulations, some of which have distinct functional properties including a bias or 'lineage priming' towards a particular cell fate. These populations are also dynamic in nature, converting from one state to another with fairly rapid kinetics.

The main focus of this thesis was to gain a more in depth understanding of the mechanisms regulating heterogeneity and lineage priming in murine ES cells by asking which signalling pathways play a role in this phenomenon and how the switch between states is regulated at a transcriptional level. These questions were asked using an ES cell line containing a sensitive reporter for the endoderm marker *Hex*. This reporter, developed by a previous lab member, allowed the identification and separation of a population of ES cells primed towards a primitive endoderm fate.

Primarily, I assessed the effect of a defined culture system (2i) on the *Hex*-expressing population. This culture system contains inhibitors that block FGF signalling and the Wnt pathway component GSK3. Culturing ES cells in 2i has been suggested to generate a more homogeneous culture. Here, I have shown that culturing ES cells or pre-implantation embryos in 2i did not eliminate heterogeneity but maintained them in an early state prior to lineage segregation. When ES cells were cultured in standard serum-containing medium, *Hex* was expressed in a mutually exclusive manner with the embryonic marker NANOG, while in 2i a subpopulation of cells coexpressed both *Hex* and NANOG. This population was functionally primed towards extraembryonic endoderm and trophoblast. Furthermore, these ES cells could efficiently contribute to 2-cell embryos in chimaera assays. LIF signalling promoted this population through the JAK/STAT pathway.

I then asked how transcription was regulated during the switch between unprimed ES cells to those primed towards a primitive endoderm fate, as well as how regulation changes during further differentiation. To ask this, *Hex* positive (primed) and negative (unprimed) ES cell populations were sorted as well as a *Hex* positive differentiated sample. These samples were analysed by GRO-seq to determine the location, density and orientation of RNA-polymerase throughout the genome. Changes in gene expression between primed and unprimed states were regulated primarily through elongation whereas genes upregulated during differentiation were regulated at the point of *de novo* initiation.

LAY SUMMARY

Mouse embryonic development initiates from a single cell that ultimately divides and differentiates to form all of the cell types of the adult body. My PhD thesis is based on trying to understand how the cells of the early mouse embryo make decisions to adopt one cell fate over another. As the origin of the first asymmetries in mouse development remains unknown, and pre-implantation development consists of 3 well-characterised lineages, the early mouse embryo is a valuable model to untangle the mechanisms of cell specification. Using both pre-implantation mouse embryos and mouse embryonic stem cells as models, I ask how both external and internal factors can affect these decisions. I observed that, the conditions that embryonic stem cells are grown in greatly affects such fate choices. These findings have given us a greater insight into the signals that are important for normal mouse development.

ABBREVIATIONS

°C – degrees centigrade

µ - micro

2i – two inhibitor medium

$^{+}$ - positive

$^{-}$ - negative

A - ampere

ABI SOLiD – applied biosystems sequencing by oligonucleotide ligation and detection

ATP – adenosine triphosphate

AVE – anterior visceral endoderm

Bis-Tris - Bis(2-hydroxyethyl)amino-tris(hydroxymethyl)methane

bp – base pair

Br-UTP - Bromouridine-triphosphate

BSA – bovine serum albumin

C - carboxyl

Ca - calcium

cDNA – complementary DNA

ChIP-seq – chromatin immunoprecipitation sequencing

Ci – curie

Cl - chloride

cm - centimetre

CO_2 – carbon dioxide

Ct – cycle threshold

CTP – cytidine triphosphate

DAPI – 4',6-diamidino-2-phenylindole

DAVID – database for annotation, visualisation and integrated discovery

DEPC - Diethylpyrocarbonate

D-LIF – diffusible/soluble LIF

DMSO – dimethyl sulphoxide

DNA – deoxyribonucleic acid

dNTPs – deoxynucleotide triphosphates

DTT - Dithiothreitol

DVE – distal visceral endoderm

E – embryonic day

E14 – E14.Tg2A cells

EC cells – embryonic carcinoma

ECATS – ES cell-associated transcripts

ECL - enhanced chemilumescent

ECM – extracellular matrix

EDTA - Ethylenediaminetetraacetic acid

EMFIs – irradiated mouse embryonic fibroblasts

EPC – ectoplacental cone

ePCR – electronic PCR

EpiSCs – epiblast stem cells

ES cells – embryonic stem cells

ExE – extraembryonic ectoderm

F – Faradays

F1 – filial 1

FACS – fluorescence activated cell sorting

FCS – fetal calf serum

FDR – False discovery rate

g – gramme

GDP – guanosine di-phosphate

GEF – guanine nucleotide exchange factor

GFP – green fluorescent protein

GMEM – Glasgow modified Eagle's medium

GRO-seq – genome-wide run-on sequencing

GTP – guanosine tri-phosphate

H - hydrogen

H_2O – dihydrogen monoxide (water)

HRP – horseradish peroxidase

HSPs – heat shock proteins

HV – Hex-venus

HV^+ - Hex-venus high cells

HV^- - Hex-venus low cells

ICM – inner cell mass

IMS – industrial methylated spirit

IRES – internal ribosome entry site

JAKi – JAK inhibitor 1

K – potassium

kb - kilobases

KOSR – knockout serum replacement medium

KSOM – potassium simplex optimisation medium

L – litre

LIP – LacZ-IRES-Puromycin

LY – LY294002 PI3K inhibitor

m – milli

M – Molar

M2 – medium 2

MEF – mouse embryonic fibroblast

Mg - magnesium

min - minutes

M-LIF – matrix-bound LIF

mm - millimetre

mRNA – messenger RNA

N - amino

n - nano

N2B27 – NDiff 227 medium

Na - sodium

NIA – National Institute on Aging

NIH – National Insitutes of Health

NRO – nuclear run-on

OH - hydroxide

^{32}P – phosphorous 32

PAGE – polyacrylamide gel electrophoresis

PBS – phosphate buffered saline

PBST – PBS-Triton-X

PCR – polymerase chain reaction

PD – PD0325901 MEK inhibitor

PE – primitive endoderm

PFA – paraformaldehyde

PGC – primordial germ cell

pH – potential hydrogen

PIC – pre-initiation complex

PNK – polynucleotide kinase

PO_4 – phosphate

Pol II – RNA-polymerase II

PVP – polyvinylpyrrolidine

qRT-PCR – quantitative real time polymerase chain reaction

RA – RNA adaptor

rDNA – recombinant DNA

RNA – ribonucleic acid

RNA-seq – RNA sequencing

RP – RNA PCR primer

rpm – revolutions per minute

RPMI – Roswell Park Memorial Institute medium

RT – reverse transcription

SDS - Sodium dodecyl sulfate

SSPE - saline sodium phosphate EDTA

TAE - Tris-acetate-EDTA

TAP – tobacco acid phosphatase

TBE – Tris-borate-EDTA

TBST – Tris-buffered saline-Tween20

TCF – transgenic core facility

TE – Tris-EDTA

TET - Tris-EDTA Tween20

TGC – trophoblast giant cell

Tris - tris(hydroxymethyl)aminomethane

TS cells – trophoblast stem cells

U – units

UPL – universal probe library

UV - ultraviolet

V - volts

XEN cells – extraembryonic endoderm cells

X-gal - 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

ABBREVIATED GENES AND PATHWAYS

* multiple family members have been mentioned in this thesis

ActB – actin beta
Ada1 - Adenosine deaminase
Akt1/PKB – protein kinase B
Amot - angiomotin
aPKC – atypical protein kinase C
Axin – axis inhibition protein
APC – adenomatous polyposis coli gene
B-gal – beta-galactosidase
Bmp* – bone morphogenic protein
Calcr - Calcitonin receptor
Carm1 – coactivator-associated arginine methyltransferase 1
Cd55 - Complement decay-accelerating factor
Cdx2 – caudal-type homeobox 2
Cer1 – Cerberus
C-kit – Mast/stem cell growth factor receptor Kit
Cntf – ciliary neurotrophic factor
Col* – collagen
Ctf1 – cardiotrophin 1
Cubn - Cubilin
Cyp11a1 - Cholesterol side-chain cleavage enzyme, mitochondrial
Dab2 – disabled homolog 2
Dazl - Deleted in azoospermia-like
Dlk1 - Protein delta homolog 1
Dsc2 - Desmocollin-2
Dvl - dishevlled
eIF-4E – eukaryotic translation initiation factor 4E
Eomes – eomesodermin
Epn3 - Epsin-3
Eras – ES cell expressed Ras
Erk – extracellular signal-regulated kinase
Esrrb – estrogen-related receptor beta
Etv2 - ETS translocation variant 2
Fbxo15 - F-box only protein 15
Fgf* – fibroblast growth factor
Fgfr2 – fibroblast growth factor receptor 2

FoxA2 – forkhead box 2

Fzd – frizzled

Gapdh - Glyceraldehyde 3-phosphate dehydrogenase

Gata* – GATA-binding factor

Gbx2 – gastrulation brain homeobox 2

Gp130 – membrane glycoprotein 130

Grb2 – growth factor receptor-bound protein 2

Gsc – goosecoid

Gsk3 – glycogen synthase kinase

H2b - Histone H2B type

Hes1 – hairy and enhancer of split 1

Hex – haematopoietically-expressed homeobox

Hnf4α - Hepatocyte nuclear factor 4-alpha

Id* – inhibitor of differentiation protein

IL-6 – interleukin-6

Irs – insulin receptor substrate

Jak* – janus kinase

Klf* – Krueppel-like factor

Krt* - Keratin, type II cytoskeletal

Lama1/Lmn – laminin subunit alpha-1

Lats2 – large tumour suppressor 2

Lef – lymphoid-enhacing factor

Lefty – left-right determination factor

Lgl – lethal giant larvae

Lhx1 - LIM/homeobox protein Lhx1

Lif – leukaemia inhibitory factor

Lifr – leukaemia inhibitor factor receptor

Lin28 - Protein lin-28 homolog A

Lrp* – low-density lipoprotein receptor

Mapk – mitogen-activated protein kinase

Max – myc-associated factor X

Meg3 – maternally-expressed 3

Mek - Dual specificity mitogen-activated protein kinase kinase 1

Mirg – miRNA-containing gene

Mixl1 - Mix paired-like homeobox

Myc – myc proto-oncogene protein

Nf2 – neurofibromin-2

Ngn* - neurogenin

Nnat – Neuronatin

Notch3 - Neurogenic locus notch homolog protein 3

Nr0b1 - Nuclear receptor subfamily 0 group B member 1

Oct4 – octamer-binding protein 4

Osm – oncostatin-M

Par* – proteinase-activated receptor

Pdgfrα – platelet-derived growth factor receptor alpha

Pdk1 – 3-phosphoinositide-dependent protein kinase 1

Pecam-1 – platelet cell adhesion endothelial molecule 1

PI3K – phosphoinositide 3-kinase

PIP* – phosphatidylinositol phosphate

Piwil* – Piwi-like protein

Plau - Urokinase-type plasminogen activator

Prdm14 - PR domain zinc finger protein 14

Pygl - Glycogen phosphorylase, liver form

Ras – rat sarcoma viral oncogene homolog

Rex1/Zfp42 – reduced expression protein

Rian – RNA imprinted and accumulated in nucleus

Sall4 - Sal-like protein 4

Shp2 – tyrosine-protein phosphatase non-receptor type 11

Socs3 – suppressor of cytokine signalling 3

Sos – son of sevenless homolog 1

Sox – SRY (sex-determining region) box

Ssea-1 – stage-specific antigen 1

Stat – signal transducer and transcriptional activator

Stella/Dppa3 – developmental pluripotency-associated protein 3

Tbp - TATA-binding protein

Taz/WWTR1 – WWdomain-containing transcription regulator protein 1

Tbx3 – T-box transcription factor TBX3

Tcf – T-cell factor

Tcfap2a - Transcription factor AP-2-alpha

Tead4 – TEA domain family member 4

Tfcp2l1 – transcription factor CP2-like protein 1

Tle – transducin-like enhancer

Tuj1 – neuron-specific class III beta-tubulin

Utf1 - Undifferentiated embryonic cell transcription factor 1

Yap1 – yes-associated protein

Zscan4 – zinc finger and SCAN domain-containing protein 4

LIST OF FIGURES AND TABLES

Chapter 1: Introduction

Chapter 2: Materials and methods

Chapter 3: Results

2i cultured embryonic stem cells and embryos contain a totipotent cell population

Chapter 4: Results

LIF signalling promotes an extraembryonic cell population

CONTENTS

CHAPTER 1: INTRODUCTION

Chapter 1: Introduction

The work in this thesis is focused on understanding the primary lineage segregations that occur during mouse pre-implantation development and, in particular, how signalling pathways control and affect these processes. To model these early events I have used pre-implantation mouse embryos and embryonic stem (ES) cells, an *in vitro* culture system representative of the early mouse embryo (Section 1.3). In this introduction I will present an overview of early mouse development including the morphological changes, the function of different lineages and the molecular mechanisms of lineage specification (Sections 1.1 and 1.2). The defining properties of ES cells and the important molecular players in maintaining a self-renewing state are also discussed (Section 1.3). In this thesis, to study early fate choices, I have used a reporter construct for the endoderm marker *Hex*. In the final introduction section I will discuss this gene in more detail and the reporter construct that has been utilised (Section 1.4).

1.1 Introduction to murine embryonic development

1.1.1 The mouse as a developmental model

Mice are frequently used as a model in scientific research as their genome is highly homologous to that of the human and protein orthologues shared between these species are highly conserved at the amino acid level. However, as between all species, there is a higher level of divergence in genes that are under a strong selection pressure such as those involved in the immune response and reproduction (Emes et al., 2003). The small size of mice means that they are easy to handle, while still being large enough for dissection and manipulation of organs such as the uterus. Mice additionally have short gestation periods of 19-21 days, and give birth to an average of 10-12 pups per litter. Therefore large numbers of mice can be studied, across several generations, in a relatively short period of time. It is also possible to perform *in vitro* fertilisation of mouse oocytes to obtain large numbers of synchronously developing embryos.

There are multiple strains of mice that are commonly used in the laboratory. These can be categorised as either inbred or outbred. Inbred strains have been maintained for more than 20 generations by brother to sister matings and are consequently homozygous at most gene loci. Experimental phenotypes tend to be more reproducible in inbred mice due to their homogeneous genetic background. This is therefore advantageous when analysing subtle phenotypes that could otherwise be lost in the genetic noise of outbred strains. However, in relation to this, it is also possible that phenotypes observed in inbred mice could be a product of specific and rare mutations in these strains that are not translatable to other strains or species. In this thesis I have used C57BL/6 inbred mice.

The first mouse genome to be fully sequenced was that of the C57BL/6 (Gregory et al., 2002; Waterston et al., 2002). However, the Mouse Genomes Project (Wellcome Trust Sanger Institute) has now sequenced the genomes of 17 routinely used strains. Historically, C57BL/6 mice have proven difficult to genetically manipulate, therefore the more amenable 129 strain is habitually used to generate knockout mouse lines, reporter lines and to derive ES cells (Section 1.3). Outbred mouse strains are maintained using at least 25, and sometimes up to several hundred, different breeding pairs per generation to retain genetic variation. Outbred mice more accurately reflect the human population as similarly, the extent of heterozygosity and mutations within the population is unknown. The CD-1 outbred mouse strain has been used in this study. Evidence suggests that the timing and specific mechanisms of pre-implantation development vary between distinct strains (Kang et al., 2013; Krawchuk et al., 2013) and such diversity is likely to account for conflicting experimental results sometimes observed.

Mice are a particularly good model in which to study the earliest stages of development, prior to implantation, as embryos can be collected from the oviduct or uterus as early as embryonic day (E) 0.5 i.e. 0.5 days after fertilisation. At this point the embryo is just a single cell, and can be maintained *ex vivo* until E5.5, around the stage that implantation would occur. However, development occurs at a slightly slower rate *ex vivo* with E5.5 corresponding to E4.5 *in vivo* (Section 2.3.3). Live imaging of this time period has offered great insights into early mammalian development (Piliszek et al., 2011) (discussed in more detail in Section 1.2.3.3). It is also possible to return these embryos to the uterus of a host mother and for them to continue to develop normally.

1.1.2 Morphological changes during pre-implantation development

Figure 1 illustrates the early stages of mouse development. The mouse oocyte is formed from 2 consecutive rounds of meiotic divisions. During these divisions, a haploid egg and 2 ‘polar bodies’ are produced. Polar bodies are a by-product of meiosis that allow the egg to maintain a sufficient amount of cytoplasm, and therefore nutrients, while losing redundant sets of chromosomes by the extrusion of a small cell. The oocyte is surrounded by a membrane, the ‘zona pellucida’, comprised of glycoproteins important for binding the sperm during fertilisation. After fertilisation, the mouse oocyte completes meiosis II generating the second polar body (Fig. 1) that will degrade during further development. The zygote then begins to divide by mitotic cleavage divisions (Lehtonen, 1980). These are known as ‘reductive divisions’ as the size of the embryo remains fairly constant (approximately 80 μm) while the size of individual cells (also known as blastomeres) decreases (Lehtonen, 1980). It has been suggested that the orientation of these initial cleavage divisions relative to the position of the second polar body introduces a bias of cells towards particular lineages (Section 1.2.2). Zygotic transcription is initiated at low levels at the single cell stage (E0.5), but the main wave of zygotic genome activation occurs at the 2-cell stage (E1.5), concurrent with maternal mRNA degradation.

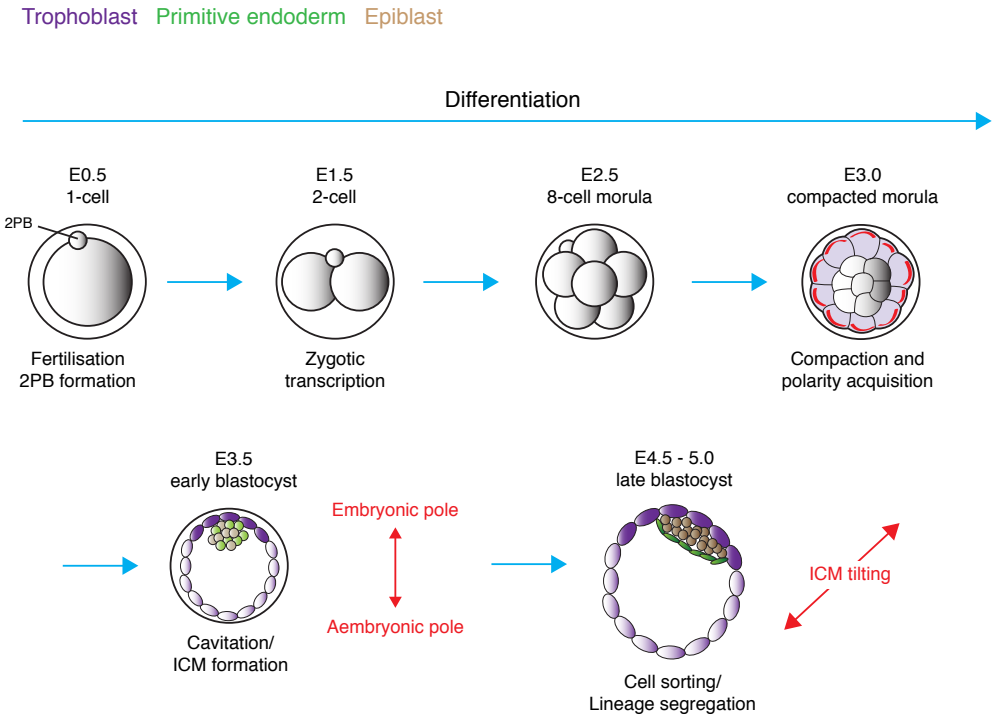


Figure 1. Schematic overview of mouse pre-implantation development. Illustration of mouse development between embryonic day (E) 0.5 and 5.0. The main events are listed below the stage at which they occur. After fertilisation, the oocyte completes meiosis II and extrudes the second polar body (2PB), a small cell containing the extra set of chromatids. The embryo then undergoes reductive divisions. At the 8-cell stage, the embryo undergoes compaction where tight adhesion junctions form between cells making it difficult to morphologically distinguish individual blastomeres. After compaction, blastomeres acquire apical polarity (illustrated by red apical regions) and subsequent asymmetric divisions generate outer polar and inner apolar cells. At the 16-cell stage (E3.0), outer and inner cells are already biased towards the extraembryonic and embryonic lineages respectively. Cavitation then occurs to form the early blastocyst with a layer of outer trophoblast cells (purple) and an ICM consisting of a heterogeneous mixture of epiblast (brown) and primitive endoderm (green) precursors. The asymmetric localisation of the ICM within the blastocyst generates the first axis of the embryo forming an embryonic and an aembryonic pole.The trophoblast cells directly overlying the epiblast form the polar trophectoderm (dark purple) and those at the opposite side, the mural trophectoderm (light purple). A day later these precursor populations have differentiated and segregated into distinct layers, the primitive endoderm now facing the blastocoel cavity. At the late blastocyst stage, the ICM tilts, the direction of the tilt being a predictor of the future anterior-posterior axis. Axis are shown by red arrows.

Between the 8-cell and 16-cell stage, the first morphological differences emerge between cells of the embryo. At the 8-cell stage 'compaction' occurs (Fig. 1) where adherens and tight junctions form between blastomeres resulting in an increase in cell-cell contact area. E-cadherin is one of the main mediators of this process (Riethmacher et al., 1995). Following compaction, blastomeres start to exhibit polarity, although the mechanisms that drive the acquisition of polarity are not understood. PAR3 and aPKC become localised to the apical membrane (Pauken and Capco, 2000; Plusa et al., 2005), and LGL and PAR1 to the basal membrane (Vinot et al., 2005) (Fig. 1). Subsequent division of these polarised cells generates the first intercellular heterogeneity in the embryo (other potentially earlier heterogeneities are discussion in Section 1.2.2). When cell division occurs perpendicular to the surface of the embryo, both daughter cells will inherit apical polarity proteins (conservative/symmetrical division) (Fig. 2A). However, if division occurs tangential to the surface they will be inherited by only one daughter cell (differentiative/asymmetrical division) (Fig. 2B).

The first round of asymmetric cell division precedes the earliest lineage segregation in the embryo, between the embryonic and extraembryonic lineages (Section 1.2.3.2). Following these divisions, at the 16-cell stage, the embryo is comprised of 2 populations, a polarized outer layer of cells and inner apolar cells. The inner cells will go on to contribute to the embryo proper, as well as to some extraembryonic tissues, while the outer cells form an epithelial monolayer called the trophoblast that will only generate extraembryonic tissues. A second round of asymmetric divisions also contributes more cells to the inside of the embryo.

Outer trophoblast cells express a Na^+/K^+ ATPase ion channel on the basal membrane that pumps Na^+ ions into the intercellular spaces (Watson and Kidder, 1988). This results in a trans-trophoblast ion gradient that causes diffusion of water in the same direction. The embryo consequently cavitates generating a central fluid-filled space known as the 'blastocoel' (Fig. 1). Cavitation initiates at E3.5, when the embryo is between 32 and 64 cells. The cavitated embryo is now referred to as the 'blastocyst'. The blastocyst maintains the outer layer of trophoblast cells while inside cells are now positioned in a cluster, the inner cell mass (ICM), at one extremity of the embryo. The asymmetrical localisation of the ICM marks the first axis in the embryo, the region containing the ICM signifying the embryonic pole and the opposite region the aembryonic pole. The trophoblast cells directly overlying the ICM are referred to as the polar trophectoderm, whereas those at the opposite side form the mural trophectoderm (Section 1.1.3). At this point, the ICM is a heterogeneous mix of precursors of the epiblast (cells that will generate the embryo proper) and the primitive endoderm (PE) (cells that will form the extraembryonic endoderm) (Chazaud et al., 2006).

Between E3.5 and E4.5, the cavity of the blastocyst expands. At E4.5, when the embryo amounts to more than 128 cells, the epiblast and PE precursors physically segregate into their respective lineages. The PE delaminates from the epiblast and forms an epithelial layer on the surface of the cavity. PE cells secrete matrix components that form a layer of extracellular matrix (ECM) between the PE and

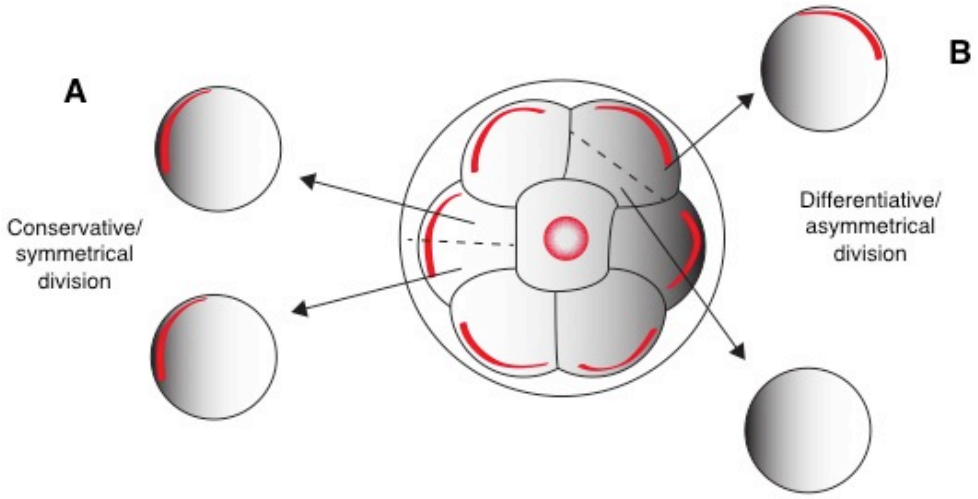


Figure 2. Schematic representation of symmetric and asymmetric divisions. At the 8-cell stage, just after compaction, the cells of the morula acquire apical polarity with proteins including PAR3 and aPKC localised to the apical surface (illustrated by red lines). During subsequent divisions, cells can divide perpendicular to the surface of the embryo **(A)**. Such divisions are symmetrical as the 2 resulting daughter cells both inherit apical polarity proteins. Alternatively a cell can divide transgential to the surface **(B)** in an asymmetric division where only the outer cell inherits polarity proteins. Dotted lines mark the plan of division.

the remaining inner cells that constitute the epiblast. Around this time, the blastocyst also hatches from the zona pellucida membrane so that it is able to implant into the uterus. The ICM tilts to one side (Fig. 1) and, following implantation, the polar trophectoderm expands in an asymmetrical manner causing the differentiated trophoblast to bend away from the proximal-distal axis of the embryo. The direction of the tilt corresponds to the orientation of the future anterior-posterior axis, although it does not predict which side will be anterior and which posterior (Gardner et al., 1992; Smith, 1980, 1985).

1.1.3 Morphological changes during early post-implantation development

At around E4.5, the cells of the mural trophectoderm halt proliferation and endoreduplicate to form mononuclear polyploid trophoblast giant cells (TGCs) (Dickson, 1963). These giant cells are highly migratory and invade into the uterus of the mother to facilitate implantation. If adverse environmental conditions, such as a lack of nutrients, are detected during this period, the embryo will arrest at the late blastocyst stage prior to implantation in a state known as 'diapause'. Diapause can also be artificially induced by removal of the ovaries, hence eliminating the source of hormones needed to maintain a receptive uterine environment.

If implantation occurs, cells of the polar trophectoderm rapidly proliferate into the blastocoel cavity to form the extraembryonic ectoderm (ExE), and externally to form the ectoplacental cone (EPC). This proliferation of polar trophectoderm, at approximately E5.5, combined with epiblast proliferation, causes elongation of the blastocyst into what is referred to as the 'egg cylinder' (Fig. 3). The EPC can also differentiate to form a secondary wave of TGCs (Ilgren, 1981a).

During this period, the PE continues to differentiate. After PE cells have sorted to the blastocoel cavity, they proliferate to line the surface of the epiblast and subsequently begin to grow along the inside of the trophoblast cells. Those that are in contact with the trophoblast generate the parietal endoderm (Ilgren, 1981a), while those remaining in contact with the epiblast form the visceral endoderm (Hogan and Tilly, 1981). A molecularly distinct region of the visceral endoderm, the anterior visceral endoderm (AVE), provides instructions to establish the future anterior-posterior axis of the embryo (Section 1.1.4) (Fig. 3).

At E5.5, a cavity is formed within the epiblast and ExE, referred to as the proamniotic cavity. The epiblast cells form an epithelial layer surrounding this cavity, with the basal sides attached to the visceral endoderm (Fig. 3). A day later at E6.5 the major morphological rearrangements begin in a process known as 'gastrulation', culminating in the formation of the 3 germ layers. Epiblast cells, at the proximal posterior of the epiblast, undergo an epithelial to mesenchymal transition whereby they migrate out of the epithelial layer and emerge between the epiblast and visceral endoderm (Fig. 3). This migratory region, called the 'primitive streak', extends distally as gastrulation proceeds. The emerging cells contribute to different lineages depending on their position within the primitive streak

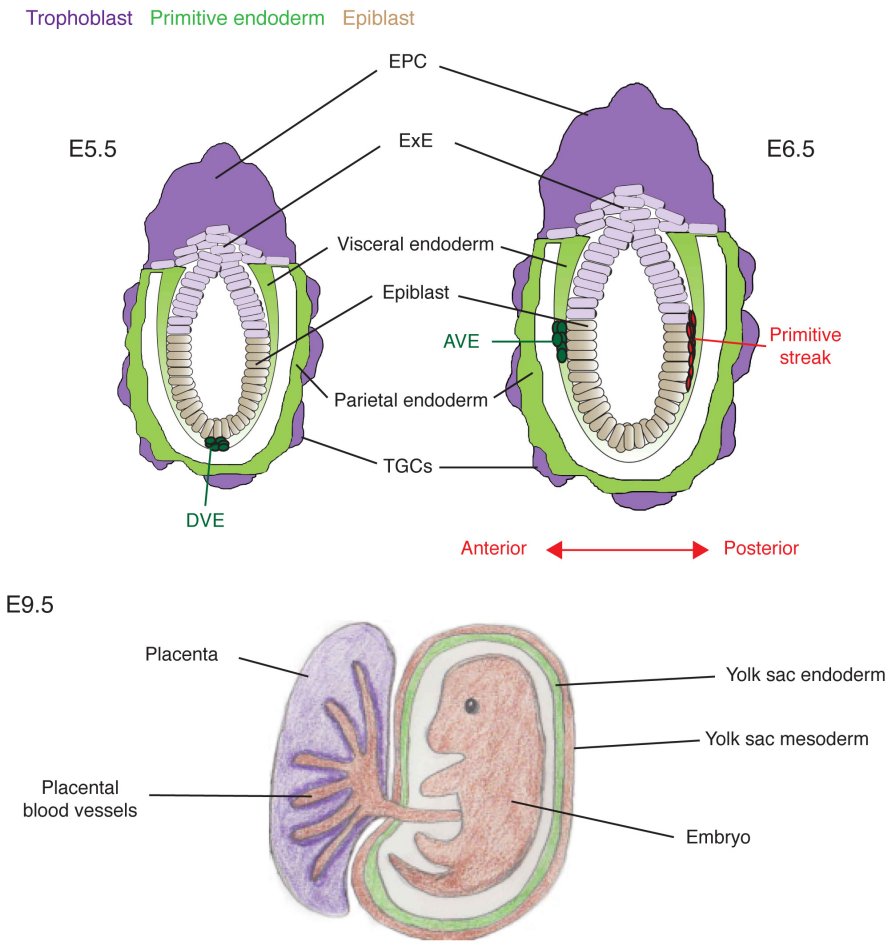


Figure 3. Schematic overview of mouse early post-implantation development. Illustration of mouse embryonic development at embryonic day (E) 5.5, 6.5 and 9.5. After implantation, the lineages of the early blastocyst (the trophoblast, epiblast and primitive endoderm) differentiate into more specified cell types. Derivatives of these lineages are shown in purple, brown and green respectively. The primitive endoderm forms the visceral and parietal endoderm. At E5.5 a proportion of cells at the distal tip of the embryo are molecularly distinct and are referred to as the distal visceral endoderm (DVE). The trophoblast differentiates into the extraembryonic ectoderm (ExE), the ectoplacental cone (EPC) and the trophoblast giant cells (TGCs). The epiblast is an epithelium surrounding a central cavity, attached at the basement membrane to the visceral endoderm. By E6.5 the DVE has migrated anteriorly to form the anterior visceral endoderm (AVE), a signalling centre in the embryo that maintains anterior identity. The position of the AVE marks the anterior-posterior axis of the embryo (marked by a red arrow). The proximal posterior of the epiblast begins to undergo morphological changes where cells move out of the epithelium in an epithelial-mesenchymal transition and emerge between the epiblast and visceral endoderm in a region called the primitive streak. These cells then differentiate into the 3 germ layers, the ectoderm, mesoderm and endoderm. By E9.5, more complex structures have developed. The visceral endoderm and parietal endoderm give rise to the yolk sac and the trophoblast forms the placenta. As well as generating the embryo proper, the epiblast also gives rise to some extraembryonic tissues including the yolk sac mesoderm and the placental blood vessels.

and the developmental stage of the embryo. The cells that first emerge from the posterior primitive streak form the extraembryonic mesoderm, while cells emerging later from the same position contribute to embryonic mesoderm. Cells emerging gradually anteriorly from the streak will form lateral mesoderm (mid streak), paraxial mesoderm (anterolateral streak), and notochord and definitive endoderm (anterior streak) respectively (Garbutt et al., 1987; Tam and Beddington, 1987). Epiblast cells that do not migrate through the streak generate the ectoderm. It was previously thought that the cells of the visceral endoderm overlying the epiblast were displaced as a sheet into the extraembryonic regions by the embryonic definitive endoderm as it emerges from the streak. However, it has now been shown that the definitive endoderm more subtly intercalates into the visceral endoderm layer and that some cells of the visceral endoderm contribute to the embryonic endoderm (Kwon et al., 2008).

1.1.4 Differentiation of early pre-implantation lineages

These morphological changes, in combination with the molecular changes discussed in section 1.2.3, result in the segregation of the trophoblast, PE and epiblast lineages, each giving rise to distinct cell types with specialised functions. A summary of the contribution of these 3 lineages to post-implantation embryos at E5.5, E6.5 and E9.5 is shown schematically in Figure 3. Additionally, a map of the derivatives of these 3 lineages throughout development is shown in Figure 4.

The first lineage of the embryo to segregate is the trophoblast (the outer cells of the embryo), at around E3.0. During the initial stages of development, the trophoblast is a simple single cell epithelium that mediates cavitation of the embryo. By the egg cylinder stage (E6.5), the trophoblast layer has differentiated into various extraembryonic tissues, including the ExE, EPC and post-mitotic TGCs (Fig. 3,4). The TGCs are the first terminally differentiated cell type of the mouse embryo and mediate implantation into the uterus by remodelling the ECM and promoting decidulisation (Bany and Cross, 2006). The EPC forms a secondary wave of TGCs and, in combination with the ExE, forms the placenta necessary for waste and nutrient exchange with the mother. The ExE, arising from the polar trophectoderm adjacent to the ICM, is where trophoblast stem cells are thought to reside (Ilgren, 1981b; Johnson and Rossant, 1981; Rossant and Tamura-Lis, 1981) (Section 1.3.8). These cells receive paracrine signals, of FGF4 and NODAL, from the epiblast (Arman et al., 1998; Chai et al., 1998; Guzman-Ayala et al., 2004; Ma et al., 2001; Xu et al., 1998) that maintain them in a highly proliferative state. The ExE therefore acts as a reservoir of relatively undifferentiated cells that can further differentiate.

After trophoblast-mediated cavitation of the embryo, the inner cells of the blastocyst segregate into the epiblast and PE. The PE is another extraembryonic lineage that gives rise to the visceral and parietal endoderm (Fig. 3,4). The visceral endoderm directly overlies the epiblast and acts as a signalling centre providing molecular cues critical for axis specification. At approximately E5.5, a region at the distal tip of the egg cylinder, the distal visceral endoderm (DVE), begins to express

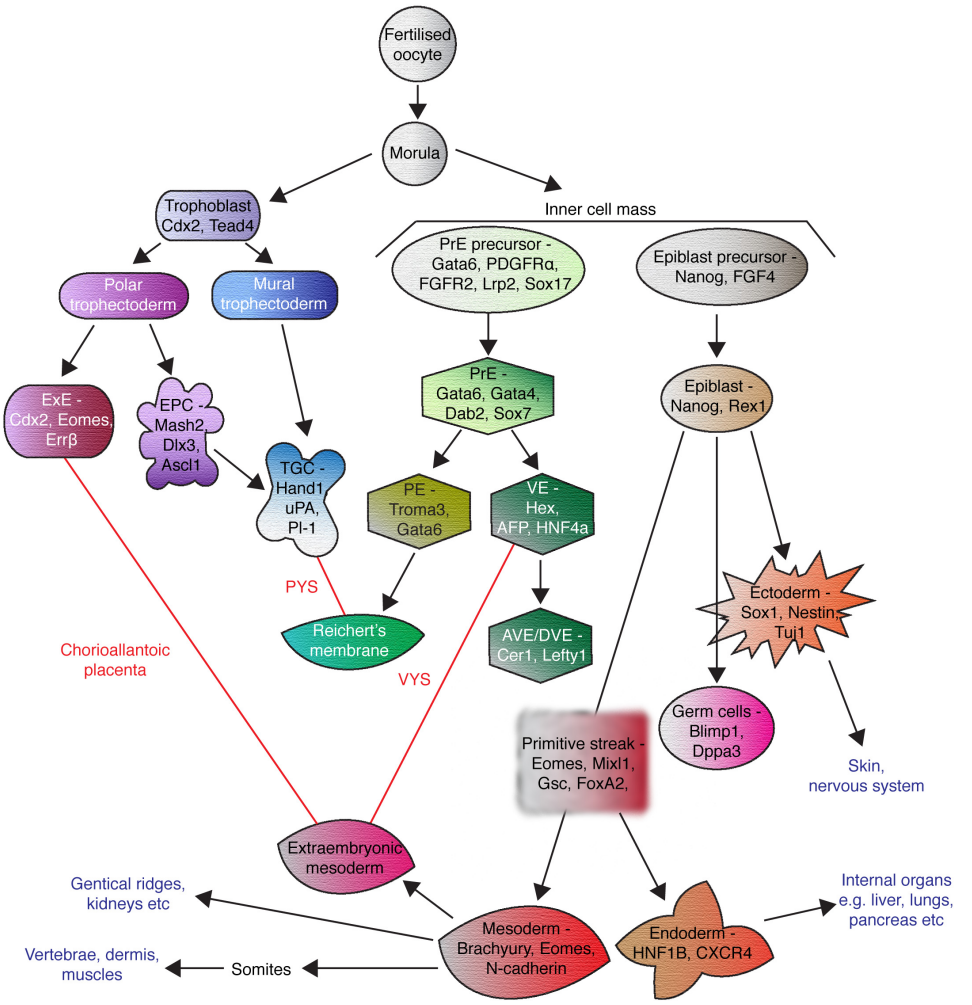


Figure 4. Lineage tree of mouse embryonic development. Schematic diagram showing how lineages progressively segregate during development to give rise to different specified tissues. Genetic markers are shown for some of the cell types. Arrows mark differentiation of one cell type into another. Some cell types combine to form complex embryonic structures (marked by red lines) such as the parietal yolk sac (PYS), visceral yolk sac (VYS) and chorioallantoic placenta. Upon fertilisation of the oocyte, the zygote undergoes cleavage divisions to form the morula. At the morula stage the first lineage segregation occurs where outside cells form the trophoblast and inside cells form the inner cell mass (ICM). The trophoblast is divided into polar trophectoderm, directly overlying the epiblast, and mural trophectoderm, at the opposite aembryonic region. The mural trophectoderm terminally differentiates into trophoblast giant cells (TGCs) while the polar trophectoderm segregates into the extraembryonic ectoderm (ExE) and the ectoplacental cone (EPC). The EPC also gives rise to TGCs. The ICM is made up of precursors of the epiblast and primitive endoderm (PrE) that differentiate into their respective lineages. The PrE segregates into the parietal (PE) and visceral endoderm (VE). The PE forms Reichert's membrane that, in conjunction with TGCs, generates the PYS. The VE forms the distal visceral endoderm (DVE) that migrates anteriorly to form the anterior visceral endoderm (AVE), a signalling centre involved in embryo axis formation. During gastrulation, the epiblast gives rise to the germ cells and ectoderm and, after epiblast cells migrate through the primitive streak, the mesoderm and endoderm are formed. Mesodermal cells also migrate into the extraembryonic regions and combine with the VE to form the VYS as well as with the ExE to form the chorioallantoic placenta. The three germ layers, the ectoderm, mesoderm and endoderm then differentiate further to form all cell types of the embryo (examples are shown in blue font). Adapted from Manipulating the Mouse Embryo, 2003.

endodermal markers including *Hex* and *Cer1* (Belo et al., 1997; Thomas and Beddington, 1996; Thomas et al., 1998). These cells subsequently migrate proximally to form the AVE, the first anterior-posterior asymmetry in the mouse embryo (Thomas et al., 1998). At later stages, the visceral endoderm contributes to the visceral yolk sac involved in nutrient and waste exchange (Gardner, 1997) (Fig. 3,4). A fraction of visceral endoderm cells also integrate into the embryonic definitive endoderm during gastrulation and subsequently contribute to the foregut (Kwon et al., 2008). The parietal endoderm is located more externally than the visceral endoderm layer and secretes many basement membrane factors forming the Reichert´s membrane that separates the yolk cavity from maternal tissues (Hogan et al., 1980; Semoff et al., 1982).

Although the trophoblast and PE generate the extraembryonic tissues, they are also both critical for development of the embryo proper, acting as sources of key paracrine factors. At approximately E5.5, BMP4 and BMP8b are expressed in the ExE and BMP2 in the visceral endoderm (Lawson et al., 1999; Ying et al., 2000; Ying and Zhao, 2001). These extraembryonic signals are critical for the production of primordial germ cells (PGCs), which arise as a small cluster of cells in the proximal epiblast at approximately E6.25. The PGCs later give rise to the sperm and oocytes. In BMP4 mutant embryos this process is reduced or completely abolished and cannot be rescued by BMP signalling from wild type embryonic cells (Hiiragi and Solter, 2005; Lawson et al., 1999; Lawson and Hage, 1994; Ying et al., 2000; Ying and Zhao, 2001).

Additionally, signals from the AVE and ExE combine to set out a precise molecular pattern in the epiblast that governs the formation of the anterior-posterior axis. *Nodal* is an important posterior specifier and in *Nodal* mutants the primitive streak cannot form (Conlon et al., 1994). The AVE maintains anterior identity in the adjacent epiblast by secreting antagonists of NODAL, such as CERBERUS and LEFTY (Conlon et al., 1994; Meno et al., 1999; Parfitt and Zernicka-Goetz, 2010; Perea-Gomez et al., 2002; Piotrowska-Nitsche et al., 2005; Piotrowska-Nitsche and Zernicka-Goetz, 2005), restricting its activity to the posterior side of the embryo. When AVE signalling is interrupted, NODAL induces ectopic primitive streaks in the anterior epiblast (Perea-Gomez et al., 2002). The ExE secretes proteases necessary to cleave pro-NODAL into an active form (Beck et al., 2002; Guzman-Ayala et al., 2004). These proteases diffuse into the proximal epiblast hence a domain of active NODAL is restricted to the posterior proximal region by a combination of AVE and ExE signalling.

The final lineage of the pre-implantation embryo, the epiblast, generates all tissues of the embryo proper. At E6.5 the embryo undergoes gastrulation (Section 1.1.3) leading to the formation of the 3 germ layers, endoderm, mesoderm and ectoderm, as well as the PGCs (Gardner, 2001; Hiiragi and Solter, 2005; Piotrowska et al., 2001). In addition to making the embryonic tissues, the epiblast also contributes to some extraembryonic tissues, including the chorion, allantois and mesodermal yolk sac (Fig. 3,4).

1.1.5 Cell lines derived from the mouse embryo

During early development, it is difficult to study distinct cell populations due to the limited amount of material available, particularly during pre-implantation development where there are just over 100 cells by the late blastocyst stage. Additionally, these populations are present only transiently before development proceeds. The derivation of self-renewing cell lines from the embryo, representative of each of the early lineages, has therefore contributed substantially to our understanding of early mouse development. Self-renewing cell lines have been derived that have gene expression profiles reminiscent of the 3 pre-implantation lineages; the epiblast, trophoblast and PE. Additionally these cell lines contribute to their respective lineages when reintroduced into embryos (discussed in more detail in Section 1.3.8).

The knowledge of the signalling pathways involved in the segregation and maintenance of these lineages *in vivo* (Section 1.2.3) was used to develop culture conditions that could sustain these cell types *in vitro*. ES cells, isolated from the mouse blastocyst, appear to represent the ICM, containing populations reminiscent of both epiblast and PE cell types (Section 1.3). However, stem cells have also been isolated representative of the later, post-implantation epiblast. These epiblast stem cells (EpiSCs) can be isolated both from blastocysts and post-implantation embryos (Najm et al., 2011; Tesar et al., 2007). Trophoblast stem (TS) cells can likewise be derived from both pre and post-implantation embryos and are thought to represent the proliferating ExE population (Tanaka et al., 1998). Cell lines can also be derived that represent the extraembryonic endoderm (XEN cells). XEN cells were initially thought to correspond to a PE-like state however there is evidence that they are more specified than originally thought and are comparable to the parietal endoderm rather than the PE that has the potential to make both visceral and parietal endoderm (Kunath et al., 2005). These cell lines are discussed in more detail in section 1.3.

1.2 Mechanisms of pre-implantation development

1.2.1 Regulative development

Although a lot is understood about early mouse development there are still unanswered questions, in particular, how asymmetries are generated within a group of seemingly identical starting cells. In many animals, including invertebrates and lower vertebrates, pattern is determined by the asymmetrical localisation of factors inside the mother's egg. During subsequent cell divisions, these factors are then differentially inherited by the daughter cells (Driever and Nusslein-Volhard, 1988a, b; Weeks and Melton, 1987). However, the origin of asymmetries and pattern formation in mammals is not known.

Mammalian development is extremely regulative as substantial perturbations can be made to the embryo without compromising the viability or further development. For example, a 2-cell embryo can be split into individual blastomeres that can each separately proceed in development and go on to form 2 individual mice (Tarkowski, 1959). Additionally, chimaeric mice can be generated by the aggregation of individual blastomeres from distinct mouse strains and even from different species (Rossant and Frels, 1980; Tarkowski, 1961). Such results suggest that either there is no early pre-defined pattern in the mouse embryo or, if a pattern does exist, that it can be regenerated upon disruption or is not necessary for development. In the following sections I will present what is currently known about the origin of heterogeneity in the mouse embryo as well as the molecular mechanisms that mediate lineage segregation.

1.2.2 Early lineage bias

As previously described (Sections 1.1.2 and 1.1.3), the morphological segregation of specific lineages is well characterized. However, there remains a significant debate as to whether lineage fate is already pre-determined at early cleavage stages. In many organisms, the patterning information needed for early development is contained within the oocyte (Section 1.2.1), hence zygotic transcription does not need to be initiated until late in development, in Drosophila not until after the 8th cell division. However, in mammals there is limited evidence to support the maternal inheritance model (Antczak and Van Blerkom, 1997). Furthermore, in mouse the zygotic genome is activated already at the 2-cell stage suggesting that, rather than a pre-existing pattern, *de novo* transcription is necessary for the initial stages of development. Additionally, mouse development is extremely regulative (Section 1.2.1) implying that there is not a strict pattern in place from the start of development. It is therefore generally accepted that asymmetries arise post-zygotically, although there is debate as to exactly when.

Three main categories of experimental techniques have been used to assess when the first lineage biases arise. The first class of experiments are disruptive i.e. those that involve significant physical manipulation of the embryo such as disaggregation of blastomeres. It is difficult to interpret the results of these experiments as they are carried out in the absence of a normal developmental context. The second class is invasive experiments, which involve minor manipulations of the embryo, for example labelling cells with dyes. Although these are more reflective of conventional development, there is still a possibility that such treatments can alter the experimental outcome. Finally there are non-invasive experiments that observe development, as it would normally occur, by tracking the fate of cells using genetically integrated markers such as fluorescent proteins.

In early experiments, single blastomeres of 2 to 8-cell embryos were labelled with horseradish peroxidase and their capacity to contribute to the later blastocyst analysed in chimaera assays. The majority of blastomeres contributed to both the ICM and trophoblast (Balakier and Pedersen, 1982), suggesting that they are not restricted to one particular lineage. Additionally, during early development, the expression of lineage-associated transcription factors is relatively homogeneous and distinct embryonic vs. extraembryonic transcriptional profiles have not been observed (Dietrich and Hiiragi, 2007; Guo et al., 2010).

However, at the 4-cell stage, variation has been observed in the level of histone modifications associated with transcriptional activation (H3R2me, H3R17me and H3R26me), as well as expression of the methyltransferase catalysing these modifications, CARM1 (Parfitt and Zernicka-Goetz, 2010). Cells overexpressing *Carm1* were biased in chimaera assays (Torres-Padilla et al., 2007). Cells with high *Carm1* contributed to the ICM and polar trophectoderm while those with low levels contributed to the mural trophectoderm (Torres-Padilla et al., 2007). Although, these experiments were carried out by artificially manipulating the expression levels of *Carm1* and hence may not represent endogenous mechanisms utilised to introduce lineage bias. In support of this, CARM1 null or kinase dead mutant mice die only at a late perinatal stage suggesting that this protein is not necessary for normal early development (Kim et al., 2010; Yadav et al., 2003). However, 2 independent, non-invasive cell tracking experiments also suggested that asymmetries arise at the 4-cell stage, although not all cells exhibited a bias at this point (Fujimori et al., 2003; Tabansky et al., 2013). Nevertheless, the stage at which blastomeres were labelled in these experiments was uncertain, as both studies used Cre-recombination labelling strategies and the time period at which recombination occurred was not clearly characterised.

As well as a lack of knowledge regarding the time point at which asymmetries arise, the mechanism by which they are initiated is also unclear. Such a mechanism would have to incorporate a level of flexibility to be able to explain the regulative nature of mouse development. One mechanism that has been proposed is the differential timing of cell divisions. The division of one blastomere before its partner could introduce temporal transcriptional differences between cells, generating lineage bias.

Early experiments suggested that the first blastomere of a 2-cell embryo to divide twice, contributed more descendants to the ICM than its neighbouring cell (Kelly et al., 1978). However, this was a disruptive experiment involving disaggregation and re-aggregation of embryos, hence does not mimic the normal developmental process. This division-dependent bias in cell fate was due to a difference in morphology between earlier and later dividing cells, the cells that divided first being flatter. Culturing blastomeres in medium with low calcium levels reversed cell flattening and eliminated any bias (Garbutt et al., 1987). The bias described in these experiments is relative rather than absolute, i.e. it depends on the properties of the neighbouring cells, and hence could reconcile the presence of an early bias with the capacity of single blastomeres to continue to develop upon disaggregation.

Another mechanism to introduce bias concerns the orientation of cleavage divisions. One model suggested that the embryonic-aembryonic axis is already determined in the oocyte and that the first cleavage division marks this axis (Gardner, 1997, 2001; Piotrowska et al., 2001). However, these experiments marked embryo orientation by labelling a position of the zona pellucida, though it is likely that the embryo is not stationary within this membrane. There has also been a more complex model proposed, suggesting that lineage bias arises according to the orientation of division of the 2 blastomeres at the 2-cell stage. For example, when the first dividing blastomere of a 2-cell embryo divided parallel to the second polar body, and the second dividing blastomere divided perpendicular, in the majority of cases the first blastomere contributed mostly to the embryonic and the second to the extraembryonic tissues (Piotrowska-Nitsche et al., 2005). On the other hand, when the first dividing blastomere divided perpendicular to the second polar body, and the second dividing blastomere divided parallel then, although one cell contributed mostly to embryonic tissues and the other to extraembryonic tissues, each cell had an equal chance of contributing to one or the other (Piotrowska-Nitsche et al., 2005). When chimaeras were generated by aggregating identically dividing cells from different embryos i.e. aggregation of all first cells to divide in a parallel manner or all second cells to divide in a parallel manner, each group demonstrated a difference in their potential to make viable chimaeras suggesting that, at this stage, cells may already have distinct functional properties (Piotrowska-Nitsche et al., 2005; Piotrowska-Nitsche and Zernicka-Goetz, 2005). However, this remains controversial (Alarcon and Marikawa, 2005) and it has now been shown that the second polar body moves towards the plane of the first division hence it cannot be used as a stable marker for embryo orientation (Hiiragi and Solter, 2005). Furthermore, these type of divisions occur in only 80% of embryos while the other 20%, where blastomeres both divide in the same orientation, have no bias towards either embryonic or extraembryonic lineages. This implies that, even if such bias is present, it is not necessary for development.

1.2.3 Signalling pathways and transcription factors in pre-implantation lineage segregation

1.2.3.1 Introduction to key lineage-associated transcription factors

One approach to unravelling the emergence of lineage bias in the mouse embryo is to understand the molecular components involved in early lineage segregation. The generation of knockout mouse models facilitated the characterisation of the functional role of genes during early development. During the first lineage segregation, *Cdx2* and *Tead4* are necessary to develop a functional trophoblast, whereas *Oct4* is thought to specify inner cells. The next lineage segregation between epiblast and PE is mediated by 2 transcription factors, *Nanog* promoting an epiblast fate and *Gata6* promoting a PE fate.

A prevailing model for these first 2 lineage segregations is a battle between lineage-specific transcription factors that mutually antagonise one another; in the first segregation *Cdx2* and *Oct4* and in the second segregation *Nanog* and *Gata6*. Consistent with this model, a knockdown of *Oct4* in ES cells results in trophoblast differentiation whereas a knockdown of *Nanog* leads to PE differentiation (Hough et al., 2006; Niwa et al., 2000). Lineage-specific transcription factors are initially coexpressed at low levels (Dietrich and Hiiragi, 2007), perhaps not sufficient for repression of one another. An unknown event prompts an increase in the expression of one gene, allowing it to reach a sufficient level to antagonise the other and subsequently activate downstream target genes that reinforce a specific developmental programme. The particular roles of these key transcription factors, and evidence for such a mechanism, will be discussed in the following sections.

1.2.3.2 Trophoblast and inner cell mass segregation

The first lineage segregation in development is that of the trophoblast from the ICM and is dependent upon the position of cells within the embryo. In the blastocyst, the transcription factor CDX2 is expressed specifically in outside cells whereas OCT4 is expressed in inner cells (Palmieri et al., 1994; Strumpf et al., 2005). This led to the hypothesis that these 2 factors mutually antagonise one another (Smith, 2005). CDX2 can bind to an upstream regulatory region of *Oct4* and inhibit its function as a transcriptional activator (Niwa et al., 2005), although there is no evidence that OCT4 directly blocks CDX2 function or expression. In ES cells, knockdown of *Oct4*, or overexpression of *Cdx2*, resulted in trophoblast differentiation (Hough et al., 2006; Niwa et al., 2000; Niwa et al., 2005) indicating that segregation of the ICM and trophoblast is connected to the relative levels of these genes. However, *in vivo*, OCT4 is coexpressed with CDX2 in outer cells for a considerable length of time after trophoblast specification (Dietrich and Hiiragi, 2007) arguing against a direct interaction of these factors.

A number of key transcription factors play a role in specifying and maintaining the trophoblast lineage, including *Cdx2*, *Eomes*, *Gata3*, *Id2* and *Klf5* (Frankenberg et al., 2011; Lin et al., 2010; Ralston et al., 2010; Strumpf et al., 2005; Tabansky et al., 2013). The expression of CDX2 is initiated soon after polarisation of the outer cells of the morula (Beck et al., 1995; Ralston and Rossant, 2008) (Section 1.1.2). At this time CDX2 is expressed ubiquitously, but at higher levels in outside cells, and is subsequently downregulated in inner cells of the early blastocyst (Dietrich and Hiiragi, 2007; Ralston and Rossant, 2008). This led to the hypothesis that the specification of the trophoblast is determined according to the relative position of cells within the embryo, the polarised outer cells forming extraembryonic trophoblast and the inner cells forming the ICM. However, cell location alone is not enough to bias cells towards a trophoblast fate. When the apical polarity markers, *Par3* and *aPKC*, were knocked down, outside cells contributed a higher number of progeny to the ICM than controls (Plusa et al., 2005). Hence, outer cell polarity seems to be the primary stimulus of the first lineage decision.

The function of *Cdx2* during trophoblast segregation was further explored using knockout mouse lines. *Cdx2* null embryos formed morphologically normal expanded blastocysts although epithelial integrity was compromised leading to defects in zona pellucida hatching and implantation (Blij et al., 2012; Strumpf et al., 2005; Wu et al., 2010). The outer cells of these embryos did not express trophoblast markers, such as EOMES, but instead ectopically expressed the ICM markers OCT4 and NANOG (Strumpf et al., 2005). Likewise, although *Cdx2* null cells contributed to the trophoblast at a normal frequency in chimaera assays, the mutant trophoblast cells expressed inner cell markers (Ralston and Rossant, 2008). This indicates that *Cdx2* is important for supressing an ICM fate but is not critical for the primary morphological segregation of inner and outer cells. Consistent with this idea, when the ICM marker *Oct4* was downregulated, ES cells could differentiate towards trophoblast even in the absence of *Cdx2* (Niwa et al., 2005)

When the transcription factor *Tead4* was knocked out, null embryos demonstrated a more severe phenotype than *Cdx2* mutant mice, suggesting that this factor is genetically upstream in trophoblast segregation. In *Tead4* null embryos, outer trophoblast cells did not form and embryos maintained a morula-like morphology that could not cavitate or express trophoblast transcription factors such as CDX2, GATA3 or EOMES (Ralston et al., 2010; Yagi et al., 2007). Chromatin immunoprecipitation sequencing (ChIP-seq) confirmed that TEAD4 directly binds to the promoters of multiple trophoblast genes in TS cells (Home et al., 2012) and overexpression of *Tead4* in ES cells triggered trophoblast differentiation (Nishioka et al., 2009). Additionally, when mRNA of *Tead4* and its 2 cofactors, *Yap1* and *Taz*, was injected into the inner cells of embryos, CDX2 expression was induced (Nishioka et al., 2009) indicating that these factors are sufficient to drive trophoblast gene expression.

TEAD4, YAP1 and TAZ are localised specifically in the nuclei of outer cells at the 16-cell stage. Studies to understand the mechanism of activation of *Tead4* and its cofactors revealed that the Hippo

signalling pathway is involved. The Hippo pathway regulates cell proliferation and growth and is activated at high densities by cell-cell contacts (Zhao et al., 2007). The Hippo pathway is active in inner cells of the embryo that possess a greater number of cell-cell contacts. Upon Hippo pathway activation, LATS2 phosphorylates YAP1 causing it to be sequestered in the cytoplasm by the scaffold protein 14-3-3 (Basu et al., 2003; Dong et al., 2007; Nishioka et al., 2009; Zhao et al., 2007). As YAP1 cannot translocate to the nucleus, it cannot interact with TEAD4 to activate target genes. *Amot* and *Nf2* were also identified as upstream components in this pathway, necessary for phosphorylation of YAP1 by LATS2 (Cockburn et al., 2013; Hirate et al., 2013).

In inside cells, Hippo pathway components, including AMOT, LATS2 and NF2, are localised at basolateral adherens junctions (Gladden et al., 2010). These components are in close association with one another hence the Hippo pathway is active. In outside polarised cells, actin fibres are present at the apical surface and AMOT is recruited away from the complex of pathway components at adherens junctions to bind to apical actin, inactivating Hippo signalling (Hirate et al., 2013). Hence, polarisation of outer cells plays a critical role in Hippo activation. When apical polarity was disrupted, YAP1 was excluded from the nucleus in outer cells (Hirate et al., 2013). Additionally, when cells of a 32-cell embryo were disaggregated, eliminating all cell-cell contacts, the former outer cells lost nuclear YAP1. However, it is still not understood what initiates cell polarisation.

FGF signalling also plays a role during trophoblast development. FGF4 is necessary for trophoblast maintenance *in vivo* (Takahashi et al., 2003) and is used in the derivation and maintenance of TS cells *in vitro* (Tanaka et al., 1998). Furthermore, it has been proposed that between the 8 and 16-cell stage, concomitant with the first asymmetrical cell division (Section 1.1.2, Fig. 2), the FGF receptor *Fgfr2* becomes more highly expressed in outer cells, perhaps establishing a bias towards the trophoblast lineage (Morris et al., 2013). However, maternal zygotic *Fgf4* mutant embryos developed a normal trophoblast layer (Kang et al., 2013) suggesting that, at least *Fgf4* is not needed for trophoblast specification.

1.2.3.3 Epiblast and primitive endoderm segregation

The next lineage segregation to occur is between the epiblast and PE and, in contrast to trophoblast and ICM segregation, is mediated independently of cell position within the ICM. *Nanog* and *Gata6* are the transcription factors most strongly implicated in the formation of epiblast and PE respectively. In the absence of *Gata6*, the PE layer does not form (Cai et al., 2008; Kang et al., 2013) and overexpression of *Gata6* in ES cells promotes endoderm differentiation (Takahashi and Yamanaka, 2006). In contrast, *Nanog* null embryos have no epiblast (Mitsui et al., 2003).

NANOG and GATA6 are initially coexpressed within individual cells of the ICM but expression is gradually restricted to an exclusive ‘salt and pepper’ pattern (Chazaud et al., 2006; Frankenberg et al.,

2011). By the early blastocyst stage (E3.5), the ICM is a heterogeneous mix of PE precursors expressing genes such as *Hex*, *Gata6* and *PDGFRα*, and epiblast precursors expressing *Nanog*. Live imaging of early blastocysts has shown that PE cells can downregulate endoderm markers (Plusa et al., 2008), although, it is not known how dynamic *Nanog* and *Gata6* expression is and whether these precursor populations are able to interconvert.

At around E4.5, PE precursors sort to the face of the blastocoel cavity by a combination of cell migration, apoptosis and downregulation of the PE transcriptional programme in inner cells (Plusa et al., 2008). The membrane proteins aPKC and DAB2 are necessary for sorting and in mutant embryos the precursor populations remained disorganised within the ICM (Guo et al., 2010; Saiz et al., 2013). During the sorting of PE from epiblast, *Pdgfrα* is necessary for survival of endodermal cells (Artus et al., 2013; Artus et al., 2010). *Gata6* is the earliest known PE marker, being expressed from the 4-cell stage (Canham et al., 2010; Plusa et al., 2008). Following *Gata6* expression, later endodermal markers such as *Pdgfrα*, *Sox17*, *Gata4* and *Sox7* are expressed with *Sox7* being expressed only after physical sorting of the PE from the epiblast (Artus et al., 2011; Frankenberg et al., 2011). After PE cell sorting to the blastocoel, LRP2, DAB2 and aPKC become apically localized and components of the basal lamina, such as COL4A1 and LAMA1, are secreted from the basement membrane of PE cells forming an ECM layer between the epiblast and PE (Gerbe et al., 2008; Saiz et al., 2013).

The mutually exclusive expression pattern of PE and epiblast markers at early stages is, at least in part, mediated by direct antagonism of *Gata6* by NANOG. NANOG can bind to the promoter of *Gata6* (Singh et al., 2007) and PE genes were upregulated when Nanog was knocked down in cells of the late morula (Frankenberg et al., 2011). However, in addition to cell autonomously blocking *Gata6* expression, NANOG also plays a contradictory role in PE formation, demonstrated by the lack of both epiblast and PE in *Nanog* null embryos (Kelly et al., 1978; Messerschmidt and Kemler, 2010; Mitsui et al., 2003). Although a mature PE, expressing late markers such as GATA4, was absent in *Nanog* null embryos, in keeping with the cell-autonomous role of NANOG, GATA6 was now expressed in all cells of the ICM (Frankenberg et al., 2011). At later stages the ICM underwent apoptosis. In tetraploid complementation assays, when a wild type ES cell-derived epiblast was present, *Nanog* null host embryos generated functional extraembryonic endoderm (Messerschmidt and Kemler, 2010). This revealed a cell non-autonomous role for NANOG in PE specification, likely through to the provision of a necessary paracrine factor.

A likely candidate paracrine and PE-promoting pathway is FGF signalling. The FGF pathway components, *Fgf4* (epiblast) and *Fgfr2* (PE) mark the earliest expression differences between epiblast and PE precursors (Frankenberg et al., 2011; Guo et al., 2010; Ohnishi et al., 2014). Additionally, *Nanog* null embryos have reduced *Fgf4* expression and the expression of late PE markers can be rescued upon addition of exogenous FGF (Frankenberg et al., 2011). Blocking this pathway with small molecule inhibitors (Nichols et al., 2009), or knocking out the FGF pathway adaptor protein

Grb2 (Chazaud et al., 2006), prevented the formation of PE and generated an ICM comprised solely of epiblast-like precursors expressing *Nanog*. In my thesis I have analysed these NANOG-expressing inner cells promoted by FGF inhibition and shown that they are not entirely epiblast in nature (Chapter 3). Conversely, culturing embryos in an excess of FGF generates an ICM comprised entirely of PE (Yamanaka et al., 2010). When exogenous FGF was added to *Fgf4* null embryos it was shown in one study that the ratio of PE to epiblast cells within the ICM was dose-dependent and increasing the FGF concentration progressively increased the percentage of $SOX17^{+}$ PE cells within the ICM (Krawchuk et al., 2013). However, a contrasting study suggested that, in this situation, the switch between a PE and epiblast fate is binary and exogenous FGF treatment results in either 100% NANOG-expressing or 100% GATA6-expressing ICMs (Kang et al., 2013). This suggested that local heterogeneities that cannot be recapitulated upon the uniform application of FGF may be necessary to generate the 'salt and pepper' distribution of epiblast and PE precursors during normal development. Discrepancies could be due to the different mouse strains used in these studies, inbred C57BL/6 vs. outbred CD1, as well as the different FGF family members used during embryo culture, FGF4 vs. FGF2. Such discrepancies and potential differences between mouse strains during early embryonic development should be taken into consideration in studies, especially those with a basis in drug testing.

When FGF mutants were cultured with exogenous FGF4, embryos within the same experimental group elicited distinct responses (Krawchuk et al., 2013). A bimodal distribution was observed indicating that some embryos responding to FGF signalling, while others were more refractory. This phenomenon is discussed further in my thesis, in Chapter 4, in relation to embryo culture experiments with the cytokine LIF that likewise resulted in variable phenotypic penetrance.

Although FGF is important for PE differentiation and NANOG seems to be important for FGF expression (Chazaud et al., 2006; Frankenberg et al., 2011; Nichols et al., 2009), the hierarchy of these components in lineage segregation is not fully understood. When wild type embryos are cultured in the presence of an FGF receptor inhibitor, GATA6 expression is lost and all cells of the ICM express NANOG. FGF is therefore critical for PE formation but it is not known whether it acts directly on the PE lineage by regulating *Gata6* expression or indirectly by blocking NANOG-mediated repression of *Gata6*. To distinguish between these possibilities, *Nanog* null embryos, normally expressing GATA6 in all ICM cells, were cultured in the presence of an FGF receptor inhibitor. If FGF acts directly on *Gata6* then, upon FGF inhibition, GATA6 expression would be lost. If FGF promotes a PE fate indirectly by blocking NANOG, then in *Nanog* null embryos GATA6 would continue to be expressed. When embryos were treated from the early until the late blastocyst stage with the FGF receptor inhibitor, GATA6 expression was maintained in approximately 70% of ICM cells. This suggests that, during this period, FGF predominantly promotes a PE fate through inhibition of NANOG. Although, FGF could act directly in some cells as 30% of cells did lose GATA6 expression. Additionally, in ES cells, activation of the FGF pathway resulted in the

repression of *Nanog* and subsequent PE differentiation (Hamazaki et al., 2006). An indirect role of FGF signalling in PE differentiation also fits with the observation that *Fgf4* maternal zygotic knockout embryos still initiate expression of GATA6 (Kang et al., 2013).

It has recently been shown that the ICM factor, *Oct4* also plays a role in PE formation and, similarly, it can regulate the expression of FGF4 (Ambrosetti et al., 1997; Curatola and Basilico, 1990; Nichols et al., 1998). *Oct4* is expressed at higher levels within the PE than the epiblast prior to PE differentiation (Palmieri et al., 1994) and, when overexpressed in ES cells, causes differentiation towards a PE fate (Niwa et al., 2000). In *Oct4* maternal zygotic null embryos PE gene expression was progressively lost generating an ICM composed of NANOG-expressing cells and some cells that expressed neither GATA6 nor NANOG (Frum et al., 2013). When exogenous FGF was added to *Oct4* null embryos, PE marker expression was not induced and this was not be rescued by wild type ICM cells, indicating that *Oct4* plays a cell-autonomous role in PE segregation. Additionally, when FGF signalling was blocked in *Oct4* null embryos, both NANOG and GATA6 were coexpressed within ICM cells. It has therefore been suggested that OCT4 regulates the antagonism of *Gata6* by NANOG. However, abolishing *Oct4* from an early stage also promotes trophoblast differentiation, hence conditional knockouts were also developed to analyse the function of *Oct4* specifically during the time period of segregation of the epiblast from the PE (Chia Le Bin et al., 2014). When Oct4 was knocked out just prior to epiblast and PE segregation, as in maternal zygotic null embryos, PE markers were decreased (Chia Le Bin et al., 2014). Although surprisingly, treatment with exogenous FGF was now able to rescue PE gene expression but not further differentiation. Additionally, in these embryos, PE differentiation could be rescued by the provision of wild type ES cells prior to the point, but not after, *Oct4* deletion (Chia Le Bin et al., 2014). This suggests that *Oct4* may play both cell autonomous and non-autonomous roles in PE specification, at different time points during development. Additionally, if embryos were treated with exogenous FGF after *Oct4* deletion but before injection of ES cells, PE formation was not rescued (Chia Le Bin et al., 2014) suggesting that *Oct4* could provide a PE-promoting paracrine factor other than FGF.

Although our knowledge of the mechanisms of epiblast and PE segregation is increasing, there are still critical questions that remain unanswered. Firstly, how are the initial heterogeneities in *Fgf4* and *Fgfr2* instigated? It has recently been suggested that outside cells internalised in the second wave of asymmetric cell divisions (Section 1.1.2, Fig. 2) express higher levels of the *Fgfr2* receptor than the cells internalised in the first wave of asymmetric divisions (Morris et al., 2013) linking the first ICM heterogeneities to cell position and polarisation at earlier stages. Initial differences in developmental expression profiles could alternatively arise from stochastic bursts of transcription (Elowitz et al., 2002) or differential exposure of cells to nutrients (Morgani and Brickman, 2013). Another unanswered question is, what regulates the initiation of PE gene expression, as *Gata6* is expressed in *Fgf4* null embryos. Additionally, what are the mechanisms by which cells move from coexpressing NANOG and GATA6 to mutually exclusive expression? The expression of *Fgf4* and *Fgfr2* is

inversely correlated, in epiblast and PE cells respectively, prior to the mutually exclusive expression pattern of NANOG and GATA6 (Frankenberg et al., 2011) and, as *Fgf4* expression is increasing until the blastocyst stage, it is possible that a certain threshold needs to be reached before the downstream pathways are activated.

1.2.4 Lineage commitment and plasticity

Although lineages become physically segregated at quite an early stage, this does not mean that they are irreversibly specified. The ICM is morphologically segregated from the trophoblast at the 16-cell stage, but when ICM cells from the early blastocyst were aggregated with morulae, 32% could still contribute to the trophoblast (Rossant and Lis, 1979). Also, when multiple ICMs were aggregated they generated blastocysts that could implant and form normal egg cylinders (Rossant and Lis, 1979), indicating that they maintain the capacity to generate functional extraembryonic tissues. From these experiments, it was observed that the ICM loses the capacity to make trophoblast when it expands to more than 16-19 cells (Rossant and Lis, 1979). Similar observations were made based on the *in vitro* differentiation of ICM cells into trophoblast (Handyside, 1978; Hogan and Tilly, 1978; Spindle, 1978). As well as ICM cells, the outer trophoblast cells retain functional plasticity after morphological segregation. A large proportion (86%) of outer cells isolated from late morulae contributed to both ICM and trophoblast lineages in morula aggregations and aggregated outer cells were able to generate complete blastocysts (Rossant and Vijh, 1980).

The epiblast and PE precursor populations, although expressing lineage-specific transcription factors (Section 1.2.3.3), also exhibit a degree of functional plasticity. It has recently been shown that PE precursors retain a higher level of plasticity than their epiblast-fated counterparts (Grabarek et al., 2012). Embryos containing a *Pdgfrα*-GFP endodermal reporter, in combination with a constitutive reporter, were disaggregated and sorted into GFP negative (epiblast precursors) and GFP positive (PE precursors) cells, and subsequently were re-aggregated with morulae. At the early blastocyst stage, a high proportion of both cell types (GFP negative and positive) contributed to multiple tissues. The PE contributed, almost equally efficiently, to trophoblast (30%), PE (44%) and epiblast (26%) while epiblast precursors were less flexible, contributing significantly to only epiblast (57%) and trophoblast (40%) but not PE (3%). However, unlike the trophoblast and ICM, once the PE is morphologically segregated from the epiblast, the 2 cell types lose their potential to make other lineages.

An approximate time has been assigned to the lineage-restriction of ICM cells according to their ability to switch their fate in response to manipulation of the PE-promoting FGF pathway. When embryos were cultured in FGF pathway inhibitors from the 8-cell stage until E4.5, all cells expressed NANOG and are thought to represent an epiblast fated ICM (Section 1.2.3.3), although the nature of these NANOG-expressing cells is explored in more detail Chapter 3 of my thesis. However, when embryos were cultured from the 8-cell stage until E3.75, and then returned to control medium, they

developed normally with both epiblast and PE layers (Yamanaka et al., 2010). Additionally, when FGF pathway inhibitor treatment was initiated at E3.75, it was sufficient to drive all cells to express NANOG (Yamanaka et al., 2010). This suggests that responsiveness of ICM cells to FGF-promoted lineage choice is lost between E3.75 and E4.5, coinciding with loss of plasticity of epiblast and PE cells in transplantation experiments, that occurs between E3.25 and E4.5 (Grabarek et al., 2012).

1.3 Embryonic stem cells

Cell lines have been established from the pre-implantation embryo that can act as an *in vitro* model of early development. Embryonic stem (ES) cells in particular have a large therapeutic potential due to their capacity to generate all cell types of the embryo proper. ES cells can be maintained *in vitro* using a number of different culture conditions including feeders, cytokines and small molecule inhibitors (Section 1.3.4). The primary derivation and culture of ES cells was accomplished using a layer of feeder cells. It was later discovered that these factors could be replaced by the cytokine leukaemia inhibitor factor (LIF) (Smith et al., 1988) and serum, and finally that serum could be replaced by the addition of BMP4 (Ying et al., 2003a). Under these conditions ES cells can be maintained *in vitro* indefinitely.

1.3.1 Properties of embryonic stem cells

ES cells are karyotypically normal cell lines derived from pre-implantation embryos between the morula and blastocyst stages (Balakier and Pedersen, 1982; Tesar, 2005). ES cells can self-renew, meaning that they can divide to generate at least one cell of equal functional potency.

ES cells are also defined as pluripotent, able to contribute to all 3 germ layers of the embryo as well as the germ cells, but not to extraembryonic tissues such as the trophoblast or PE. However, this definition is problematic as it can only be confirmed retrospectively, either *in vitro* using ES cell differentiation protocols or *in vivo* by generating chimaeric mice (Balakier and Pedersen, 1982; Fleming, 1987). Protocols have been developed that promote the differentiation of ES cells to mesoderm (e.g. cardiac differentiation), endoderm (e.g. hepatic differentiation) and ectoderm (e.g. neural differentiation) *in vitro*, although many of these are still inefficient and do not always generate functional cell types.

The gold standard assessment of pluripotency is to generate chimaeric mice by aggregating or injecting ES cells into morulae or blastocysts. It is possible to introduce a constitutive marker, such as β-galactosidase or a fluorescent protein, into ES cells in order to determine which tissues they have contributed to during development. Another *in vivo* method to assess ES cell pluripotency is to inject ES cells into a well vascularised environment, such as under the kidney capsule of mice, and observe whether they are able to form teratomas. Teratomas are tumours usually comprised of a disorganised mix of cell types from all germ layers; hence in some way mimic normal development. As well as differentiated tissues, these tumours often contain an undifferentiated, malignant cell type called embryonic carcinoma (EC) cells (Diwan and Stevens, 1976). EC cells are stem cells as they can self-renew and differentiate (Kleinsmith and Pierce, 1964) hence have been studied as a model of development in their own right (Bradley et al., 1984).

Finally, ES cells can contribute to the germline (Bradley et al., 1984), meaning that they can make PGCs and thus the oocytes and sperm. Therefore, when chimaeric mice are mated, a proportion of their offspring should originate from ES-cell generated gametes and therefore carry the genetic information from these injected cells. For example, ES cells derived from the 129 mouse strain express a gene producing an agouti pigment in the fur. When these cells are injected into host C57BL/6 mice, with dark brown/black fur, the F1 offspring from the chimaera should include 100% agouti mice.

1.3.2 Embryonic stem cell totipotency

Although ES cells are referred to as pluripotent, there is a small body of evidence suggesting that they may have a wider developmental potential than is commonly thought and be able to contribute to both embryonic and extraembryonic lineages, representing a totipotent cell type. Early chimaera experiments showed that, on occasion, ES cells contributed to visceral and parietal endoderm as well as trophectoderm (Beddington and Robertson, 1989; Canham et al., 2010; Lallemand and Brulet, 1990; Suemori et al., 1990). In more recent work it was shown that a sorted sub-population of ES cells could also contribute to the extraembryonic endoderm (Canham et al., 2010). However, the results of chimaera experiments seem to be dependent upon the experimental design. For example, a higher proportion of ES cells contribute to the trophoblast in blastocyst injections than in morula aggregations (Lallemand and Brulet, 1990). The number of ES cells introduced could also affect the outcome (Section 1.3.7). Additionally, the majority of chimaera experiments involve generating full-term chimaeras hence, the contribution of ES cells to extraembryonic tissues is often not analysed. In this thesis I have analysed the ability of ES cells to contribute to both embryonic and extraembryonic tissues (Chapter 3) as well as how culture conditions affect this capacity.

Intriguingly, ES cells express genes specific to earlier developmental stages than the blastocyst from which they are derived, potentially corresponding to populations that, like these early embryonic stages, have fewer lineage restrictions. *Zscan4*, a telomere maintenance gene specifically expressed at the 2-cell stage, is expressed in around 5% of ES cells (Falco et al., 2007). Although *Zscan4* is expressed in only a small population of cells at any given time, almost all cells expressed this gene at some point during prolonged passaging. When *Zscan4* was knocked down, cells slowed proliferation and underwent a crisis (Zalzman et al., 2010) suggesting that it is necessary for cells to express this early gene for long-term maintenance. A population was also identified, based on the expression of the murine endogenous retrovirus (*MuERV-L*) that expressed genes specific to the 2-cell stage (Macfarlan et al., 2012). This population contributed to embryonic and extraembryonic tissues in chimaera assays. However, these cells cannot be considered as totipotent as their functional capacity was not analysed at the single cell level. In my thesis I have analysed the expression of a number of

these 2-cell genes within subpopulations of ES cells, as well as assessing the competence of single cells to contribute to the embryo (Chapter 3).

1.3.3 Key transcription factors regulating embryonic stem cell identity

Since ES cells were initially derived, much progress has been made to identify genes involved in the maintenance of the self-renewing state. Novel factors, referred to as ECATs (ES cell-associated transcripts) were systematically identified according to their occurrence in a number of ES cell expression data sets (Mitsui et al., 2003). As expected, many of these are common to pre-implantation development such as *Oct4* and *Nanog* (Section 1.2.3.1) and a large number, including *Nanog*, *Oct4* and *Sox2*, are also expressed in PGCs (Pesce and Scholer, 2001; Surani et al., 2007; Western et al., 2005).

These pluripotency-associated factors are self-reinforcing, in that they regulate their own expression, as well as sharing many target genes, that often need more than one of these factors for activation (Chen et al., 2008; Kim et al., 2008; Nakatake et al., 2006; Rodda et al., 2005; Tomioka et al., 2002). OCT-SOX motifs in particular are commonly found in the promoters of many pluripotency-related genes. The overexpression of a core group of ES cell genes, *Oct4*, *Sox2*, *Klf4* and *c-Myc*, can induce the expression of the pluripotency network within somatic cells and functionally reprogramme them to an ES cell state (induced pluripotent stem cells, iPSCs) (Kaji et al., 2009; Takahashi and Yamanaka, 2006; Wernig et al., 2007).

A number of these transcription factors were identified as key players in ES cell self-renewal by the phenotype observed in knockout embryos. *Oct4* knockout embryos developed into normal blastocysts but at post-implantation stages, and during diapause, the ICM was lost and only trophoblast cells were recovered (Nichols et al., 1998). Similarly *Nanog* and *Sox2* null embryos lacked an epiblast and differentiated entirely to extraembryonic tissue (Avilion et al., 2003; Mitsui et al., 2003). Hence, the function of these factors appears to be to support an embryonic fate by blocking extraembryonic differentiation. Consistent with these results, in ES cell derivation conditions *Oct4* and *Sox2* mutant embryos generated only trophoblast cells (Li et al., 2007; Nichols et al., 1998) while *Nanog* null ES cells differentiated towards extraembryonic endoderm (Mitsui et al., 2003). Although it is not possible to derive ES cells from *Nanog* null embryos, it is possible to knockout *Nanog* in ES cell lines and for self-renewal to be maintained, albeit with higher levels of differentiation (Chambers et al., 2007). This suggests that *Nanog* may be critical for the specification, but not maintenance, of ES cells.

As well as *Oct4*, *Sox2* and *Nanog*, that regulate pluripotent gene expression, another module of genes has been identified in ES cells based around the transcription factor *Myc* and it's interaction partners (Hu et al., 2009). This module predominantly regulates genes involved with cell cycle, metabolism and cell death and its binding sites are mostly distinct from the binding sites of OCT4, SOX2 and

NANOG (Chen et al., 2008; Loh et al., 2006). When the MYC-interacting partner, *Max*, was knocked out in ES cells, pluripotency-associated genes were downregulated leading to differentiation and subsequent apoptosis (Hishida et al., 2011). However, this only occurred after 6 days and seemed to be as an effect of elevated FGF signalling, suggesting that these genes are not direct targets of the MYC module. However, the requirement for *Myc* can be bypassed under certain conditions including the overexpression of *Nanog* or when FGF signalling is blocked (Hishida et al., 2011).

The overexpression of *Nanog* in ES cells can maintain self-renewal in the absence of the exogenous cytokine LIF (Chambers et al., 2003; Mitsui et al., 2003). However, the overexpression of *Oct4* and *Sox2* results in either differentiation or cell death respectively (Mitsui et al., 2003). This suggests that so-called pluripotency factors may play distinct roles in ES cells. Concordantly, during development these factors have distinct expression patterns and are often associated with multiple lineages. *Sox2*, *Nanog* and *Oct4* are all expressed within the ICM, epiblast and PGCs. However, *Sox2* is also expressed in the ExE and neural stem cells (Avilion et al., 2003; Li et al., 1998; Zappone et al., 2000) and *Oct4* is expressed in the PE where it is transiently upregulated prior to further differentiation (Palmieri et al., 1994). Additionally, *Nanog, Oct4* and *Sox2* mutant embryos have defects in both embryonic and extraembryonic tissues (Frum et al., 2013; Kelly et al., 1978; Messerschmidt and Kemler, 2010) indicating that they have roles in other lineages and are likely to influence the ES cell state via distinct mechanisms.

The canonical view of the ES cell state is that pluripotency-associated genes maintain self-renewal by blocking differentiation. However, an alternative mechanism has been proposed whereby ES cell-associated transcription factors promote differentiation towards different lineages and a combination of these genes pulling in opposite directions counteracts differentiation towards any single lineage (Loh and Lim, 2011). For example, although, within a certain range, *Oct4* maintains self-renewal, if it is expressed above this threshold ES cells differentiate towards endoderm and mesoderm differentiation and if expressed below this threshold trophoblast differentiation occurs (Niwa et al., 2000). Accordingly, ECATs can be divided into categories including those that play a role in differentiation towards specific lineages (Thomson et al., 2011). For example, the gene *Brachyury* is categorised as an ECAT (Mitsui et al., 2003) even though it is strongly associated with the mesoderm lineage. This model attempts to reconcile pluripotency with the capacity to undergo multi-lineage differentiation by suggesting that the pluripotent state is an unstable one mediated by a precarious balance of the expression levels of differentiation promoting transcription factors. It is also supported by the finding that lineage-specific genes can replace a number of these canonical pluripotency markers during the reprogramming of somatic cells (Shu et al., 2013).

1.3.4 Key signalling pathways regulating embryonic stem cell identity

1.3.4.1 BMP signalling

The previous section described the individual transcription factors involved in maintaining a self-renewing ES cell state. Here, I summarise some of the key signalling pathways that are also involved. Various signalling pathways must either be activated or supressed by cytokines or small molecule inhibitors for long-term ES cell culture. Some of these pathways are summarised in Figure 5. ES cells can be maintained in serum and LIF, but if serum is removed, ES cells differentiate. The function of serum can be recapitulated by the addition of the cytokine BMP4 (Ying et al., 2003a). BMP4 binds to a heterotetrameric serine/threonine kinase receptor complex. The activated receptor phosphorylates SMAD1, which is then able to bind to a co-smad, SMAD4, necessary for translocation to the nucleus and activation of downstream target genes (Fig. 5A). In ES cells, BMP signalling promotes the expression of the inhibitor of differentiation (Id) family of genes. When overexpressed, these genes can maintain ES cells with LIF in the absence of BMP4, by specifically blocking neural differentiation (Ying et al., 2003a). However, if BMP4 is added to ES cells in the absence of LIF, cells differentiate into flattened epithelial cells (Ying et al., 2003a). These findings suggest that a single signalling pathway is not sufficient to maintain an ES cell state.

A number of signalling pathways that are active in ES cells are discussed in more detail below. However, some pathways are active but do not have a clear role in self-renewal. For example, although components of the Notch pathway are present in ES cells, this pathway does not play a role in self-renewal as activation or inhibition did not affect the ES cell phenotype (Lowell et al., 2006; Schroeder et al., 2003). However, activation of the Notch pathway during differentiation was observed to direct cells towards a neural fate (Lowell et al., 2006). Likewise, the Nodal pathway is active in mouse ES cells but upon inhibition, although proliferation was reduced, there was no effect on self-renewal (James et al., 2005; Ogawa et al., 2007).

1.3.4.2 LIF Signalling and the JAK/STAT pathway

When ES cells were first derived, it was necessary to culture them on a feeder layer of fibroblasts to prevent differentiation (Evans and Kaufman, 1981). Later, the critical factor provided by feeders was identified as the cytokine LIF, hence ES cells can be grown without feeders in the presence of LIF or other members of the IL-6 cytokine family such as CTF1, OSM, CNTF or IL-6 (Conover et al., 1993; Nichols et al., 1994; Pennica et al., 1995; Rose et al., 1994; Smith et al., 1988; Stewart et al., 1992; Williams et al., 1988). LIF binds to a heterodimeric receptor complex of GP130 and LIFR resulting in a conformational change that facilitates the binding of JAKs, and subsequent phosphorylation by the receptor. Activated JAKs reciprocally phosphorylate the receptor creating docking sites for proteins that contain Src homology 2 (SH2) domains, leading to the activation of a number of downstream

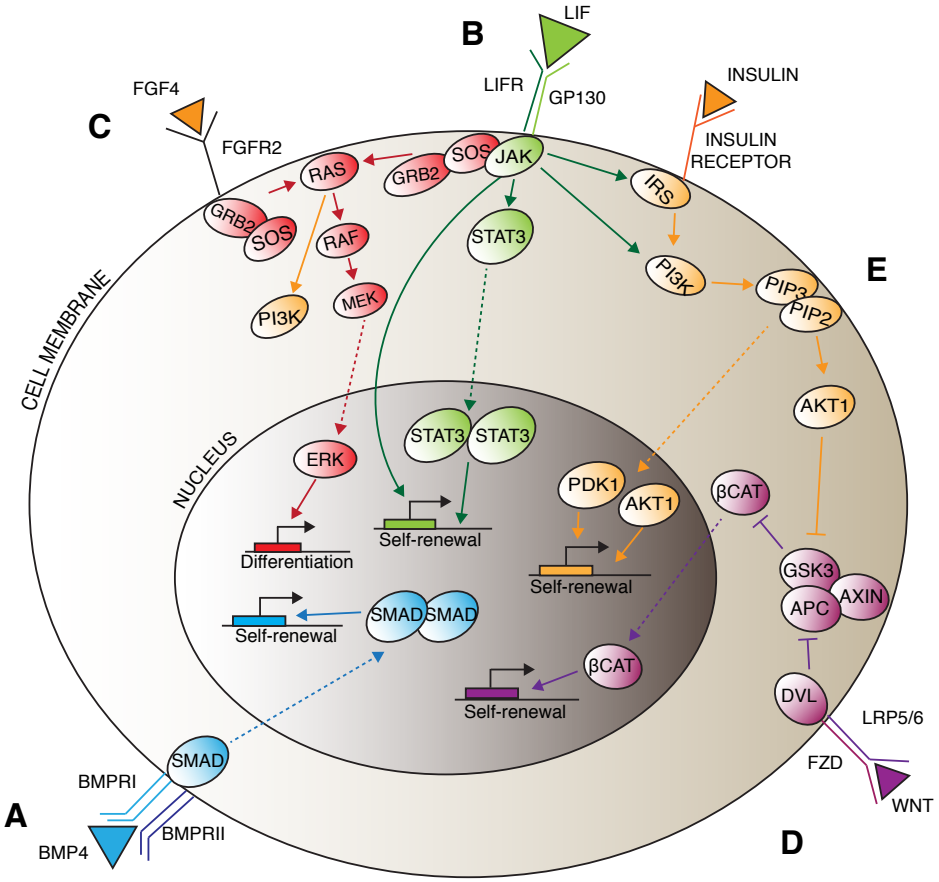


Figure 5. Overview of important signalling pathways in mouse ES cells. Dashed arrows indicate nuclear translocation whereas arrows indicate activation. **A.** BMP4 binds to a heterotetrameric receptor complex of 2 BMPRI and 2 BMPRII proteins. The type II receptor phosphorylates the type I receptor creating a docking site for Smads. Smads are then phosphorylated which facilitates dimerisation and subsequent translocation to the nucleus where they can activate target genes. **B.** LIF binds to a heterodimeric receptor complex of GP130 and LIFR. JAKs bound to the receptor subunits are then bought into close proximity where they can phosphorylate and activate one another. Activated JAKs then phosphorylate downstream STATs allowing them to dimerise and translocate to the nucleus to activate target genes. Downstream of LIF, JAK can also activate the PI3K pathway by binding to PI3K directly or by phosphorylating IRS, as well as binding to and directly effecting chromatin organisation. **C.** FGF4 binds to its receptor FGFR2 causing autophosphorylation that recruits the scaffold protein GRB2. GRB2 subsequently recruits the RAS-GEF, SOS so that it is in close proximity to its downstream component RAS. Sos activates RAS which then signals through the MAPK phosphorylation cascade activating RAF (MAP3K), which activates MEK (MAP2K) which activates ERK (MAPK). Activated ERK can then translocate to the nucleus to activate downstream target genes. RAS can also activate PI3K. **D.** Wnt binds to its heterodimeric receptor complex of FZD and LRP5/6. DVL, bound to the intracellular domain of the receptor complex, is then activated. DVL activation mediates the breakdown of a destruction complex consisting of GSK3, APC and AXIN. When WNT is not present, this complex causes the ubiquitination and subsequent degredation of β-catenin, hence in the presence of WNT, where the destruction complex is no longer functioning, β-CATENIN is stabilised and can translocate to the nucleus to activate target genes. **E.** PI3K is activated by various growth factor pathways including FGF, INSULIN and JAK/STAT signalling. PI3K phosphorylates phosphoinositides at the membrane to generate signalling molecules such as PIP2 and PIP3 that in turn activate PDK1 and AKT1 which can translocate to the nucleus and activate their target genes. AKT1 can also phosphorylate and inhibit GSK3.

pathways (Akira et al., 1994; Boeuf et al., 1997; Davis et al., 1993) (Fig. 5B). There are 3 well-characterised pathways downstream of LIF, the ERK/MAPK pathway, the PI3K pathway and the JAK/STAT pathway. SHP2 and GRB2, scaffold proteins in the MAPK pathway, are recruited to the GP130 receptor, which leads to subsequent activation of the downstream pathway components (Section 1.3.4.3). PI3K is also recruited to the activated receptor at the membrane through the SH2 domains in its p85 subunit (Section 1.3.4.4). Additionally JAK can bind to and activate PI3K directly as well as indirectly, by activating other pathways such as insulin signalling. Phosphorylated STAT proteins can also bind to the activated receptor and are phosphorylated by JAK. STAT proteins then dimerise via their SH2 domains and translocate to the nucleus to activate target genes (Fig. 5B).

The MAPK pathway promotes differentiation of ES cells while the PI3K and JAK/STAT pathways are involved in self-renewal, hence LIF activates pathways with contradictory functions. The outcome of LIF signalling seems to be the result of a balancing act between the relative activities of each pathway. For example, when the MAPK pathway was inhibited, ES cells could self-renew in the absence of LIF (Burdon et al., 1999; Ying et al., 2008) while inhibition of PI3K caused an increase in ERK signalling and consequent ES cell differentiation (Paling et al., 2004).

It has been proposed that the distinct branches of LIF signalling act upon different transcription factor networks. Immediate early genes in the LIF response in ES cells were analysed by LIF starvation for 21 hours before re-addition of the cytokine. The direct STAT3 target, *Socs3*, was upregulated within 0.5 hours of LIF addition (Niwa et al., 2009). *Klf4* follows this same rapid response, while *Nanog* and *Tbx3* were not upregulated at early stages. Inhibitor experiments revealed that LIF acts on *Nanog* and *Tbx3* through the PI3K pathway but on *Klf4* and *Socs3* through the JAK/STAT pathway (Niwa et al., 2009). Interestingly, ChIP-seq experiments revealed that STAT3 also binds to the promoters of differentiation-related genes as well as pluripotency genes (Kidder et al., 2008), in addition to the fact that the expression of LIF and its cognate receptor increase during ES cell differentiation (Aghajanova et al., 2006). In my thesis I have explored the relevance of the regulation of extraembryonic target genes by LIF (Chapter 4). Numerous LIF-responsive genes have been identified based on their ability to mediate LIF-independent ES cell self-renewal, including *Klf2*, *Klf4*, *Klf5*, *Tbx3*, *Tbx4, Myc* and *Nanog* (Cartwright et al., 2005; Chambers et al., 2003; Ema et al., 2008; Mitsui et al., 2003; Niwa et al., 2009; Toyooka et al., 2008). However, when these genes are deleted, ES cells can still be maintained in an undifferentiated state in the presence of LIF. This suggests that, although they may be sufficient, they are not necessary for self-renewal. One of the most interesting direct targets is *Tfcp2l1* as when it is knocked down in ES cells, self-renewal in the presence of LIF is compromised (Martello et al., 2013; Ye et al., 2013). Furthermore, it is a convergence point of numerous pathways involved in self-renewal and hence may allow an insight into how the input from many different signalling pathways may be integrated to promote a pluripotent state.

Although LIF supports ES cell self-renewal, this activity does not correlate clearly with its known activity during embryonic development. The most comprehensively studied role of LIF *in vivo* is in trophoblast maintenance and implantation of the embryo into the uterus (Poehlmann et al., 2005; Prakash et al., 2011; Takahashi et al., 2003) with very little being known about its pre-implantation role at the time that ES cells arise. This is further confounded by the fact that many pathway mutants, including *Lif*, *Lifr*, *Gp130* and *Stat3* null embryos all survive through these stages and exhibit later post-implantation defects (Do et al., 2013; Li et al., 1995; Takeda et al., 1997; Ware et al., 1995). However, it has now been shown that in *Lif* null embryos, IL-6 phosphorylates and activates STAT3, hence the lack of an obvious phenotype may be due to functional redundancy at these stages (Do et al., 2013). It was additionally shown that, although *Gp130* null embryos do not exhibit early defects during normal development, when diapause is induced the epiblast is not maintained (Nichols et al., 2001)

There are 2 distinct forms of LIF, a soluble form (D-LIF) and a matrix bound form (M-LIF) (Rathjen et al., 1990) expressed from alternative promoters. The M-LIF promoter is located in intron 1, 500bp downstream of the D-LIF promoter. Interestingly, M-LIF is not expressed in rat, pig, sheep or human and hence appear to be unique to mouse (Willson et al., 1992). When overexpressed from the morula stage, D-LIF gave rise to normal embryos while M-LIF gave rise to embryos that were blocked in development and could not differentiate towards mesoderm (Conquet et al., 1992). In egg cylinder stage embryos, the D form is most predominantly expressed in the trophoblast while the M form is expressed in the visceral endoderm (Robertson et al., 1993). Additionally, the matrix bound form is most highly expressed in ES cells, with the soluble form barely detectable (Rathjen et al., 1990). It has also recently been shown that self-renewal is dependent upon matrix remodelling by matrix metalloproteinases (MMPs) and that this involves the JAK/STAT pathway (Przybyla et al., 2013), as well as the fact that blocking the adhesion function of β-catenin leads to a decrease in LIFR, GP130 and STAT3 followed by a loss of self-renewal (del Valle et al., 2013). These observations strongly connect LIF and JAK/STAT signalling to the ECM and cell adhesion in addition to its role as a soluble factor.

The JAK/STAT pathway is necessary to promote self-renewal downstream of LIF. An inducible form of STAT3 can maintain ES cells in the absence of LIF (Matsuda et al., 1999) and conversely, ES cells expressing a dominant negative form of STAT3 cannot self-renew even in the presence of LIF (Niwa et al., 1998). Additionally, the immediate early transcriptional response occurring after 1 hour of LIF stimulation includes only 5 genes that are not dependent on STAT3 activation (Martello et al., 2013).

The activation of the JAK family of proteins by LIF also affects self-renewal independently of STAT3. In ES cells, a constitutively active form of JAK2 phosphorylated histones preventing the binding of proteins involved in heterochromatin formation at pluripotency-related genes, including *Sox2* and *Nanog*. Additionally, this JAK2 mutant could maintain self-renewal of ES cells in the

absence of growth factors (Griffiths et al., 2011). However, as mentioned above, the critical component downstream of LIF in this context is STAT3 and it is not clear whether JAK2 plays any role in self-renewal in wild type cells.

In the blastocyst, there is a complimentary expression of *Lif* in the trophoblast and the *Lifr* in the ICM implying a paracrine interaction between these 2 lineages (Nichols et al., 1996). STAT3 maternal null embryos make trophoblast successfully and form a blastocyst structure including ICM. However, just before implantation, the inner cells undergo apoptosis suggesting a role of JAK/STAT signalling in maintaining, although not establishing the epiblast and PE lineages (Do et al., 2013). Similarly, in porcine embryos JAK/STAT signalling is necessary to maintain ICM cells, but also for trophoblast function as blastocysts fail to hatch upon JAK/STAT inhibition (Rodriguez et al., 2012).

1.3.4.3 ERK/MAPK, Wnt signalling and 2i culture conditions

Two pathways important in the differentiation and self-renewal of ES cells are the ERK/MAPK and the Wnt pathway. LIF and FGF signalling can activate the ERK/MAPK pathway. In both cases, a complex of the adaptor proteins SHP2 and GRB2 bind to the activated receptor by their SH2 domains and recruit a guanine nucleotide exchange factor (GEF), SOS. The recruitment of SOS brings it into close proximity to its substrate, RAS, located at the membrane. RAS is activated by SOS-mediated exchange of guanosine di-phosphate (GDP) for guanosine tri-phosphate (GTP) and subsequently initiates a downstream phosphorylation cascade. RAS phosphorylates and activates the MAPKKK RAF, which in turn activates the MAPKK MEK and then the downstream MAPK ERK. Once phosphorylated, ERK can translocate to the nucleus and activate its target genes (Fig. 5C).

In the canonical Wnt pathway, the Wnt ligand binds to a heterodimeric receptor complex of FZD and LRP5/6. The conformational change of the receptor complex activates a protein, DVL, bound to the intracellular domain of FZD. DVL then inactivates a destruction complex, comprised of GSK3β, AXIN and APC, that targets β-catenin for proteosomal degradation. β-catenin consequently accumulates and translocates to the nucleus where it interacts with a DNA binding complex of TCF/LEF proteins to activate target genes (Fig. 5D).

As well as playing a role in lineage segregation of the epiblast and PE *in vivo* (Section 1.2.3.3), MAPK signalling promotes the differentiation of ES cells. When this pathway was blocked, either by chemical or genetic means, ES cells were defective in differentiation towards neural and mesodermal lineages (Kunath et al., 2007). Conversely, ES cells can be maintained without LIF when the FGF pathway is blocked. Serum-free ES cell culture conditions have been developed using a small molecule inhibitor (PD0325901) to inactivate the FGF pathway component, MEK, and a GSK3β inhibitor (CHIR99021)) (Ying et al., 2008). This inhibitor culture system is referred to as 2i (2 inhibitor).

In 2i, ES cells are shielded from FGF differentiation-promoting signals; by and large eliminating spontaneous differentiation and hence generating a more homogeneous ES cell culture (Antczak and Van Blerkom, 1997; Ying et al., 2008). Similarly, *Fgf4* null ES cells lose the expression of early differentiation marks such as DNMT3B (Findlay et al., 2013). In standard serum culture conditions, the pluripotency-associated gene NANOG is heterogeneously expressed but in 2i, the majority of ES cells express this marker (Wray et al., 2011). RNA-seq analysis of ES cells cultured in 2i revealed that markers of differentiated cell types, such as mesoderm and neurectoderm genes, were decreased compared to serum-cultured ES cells (Marks et al., 2012). However, both PE and germ cell markers were not decreased, and in some cases expression was enhanced, indicating that there was not a consistent effect on all lineages. The negligible effect of 2i on endodermal gene expression may, in part, be due to the significantly decreased expression of *c-Myc* in 2i, which is known to repress PE differentiation in standard culture conditions (Smith et al., 2010).

In 2i, in addition to an FGF pathway block, the Wnt pathway is activated by inhibition of GSK3β. The inhibition of GSKβ in these conditions is suggested to be necessary to maintain ES cells over many passages (Ying et al., 2008) as a means to overcome the metabolic stress caused by inhibition of a crucial metabolic pathway (FGF). However, activating Wnt signalling alone, in the absence of MEK inhibition, also promoted self-renewal of ES cells (Hao et al., 2006; Kirby et al., 2012; Sanchez-Ripoll et al., 2013; Sato et al., 2004) and it was later shown that GSK3β inhibition eliminates repressive TCF/LEF complexes on pluripotency-associated genes (Antczak and Van Blerkom, 1997; Martello et al., 2012). Nevertheless, abolishing WNT signalling in the embryo had no effect on pre-implantation development (Biechele et al., 2013) demonstrating that the same mechanisms are not necessary during *in vivo* development.

In 2i-cultured ES cells, the repressive chromatin mark, H3K27me3, is reduced on genes that are expressed at low levels (Marks et al., 2012), coincident with a global reduction in DNA methylation (Ficz et al., 2013; Leitch et al., 2013; Marks et al., 2012). This reduction does not result in an increase in expression of the demethylated genes hence the impact of this phenomenon is not clear. However, it has recently been shown that, in cells exhibiting global hypomethylation, Polycomb complexes can be titrated away from promoters to bind to intergenic regions (Reddington et al., 2013). The net effect of this would be the depression of extraembryonic promoters normally bound by Polycomb complexes (Boyer et al., 2006), allowing the upregulation of trophoblast and PE genes.

1.3.4.4 PI3K signalling

PI3K is activated by growth factor receptors including the LIF receptor and insulin receptor as well as by downstream components such as JAKs. Activated PI3K phosphorylates phospholipids in the plasma membrane to generate signalling factors such as $PI(3,4)P_2$ and $PI(3,4,5)P_3$. The

phosphorylated lipids can activate protein kinases including PDK1 and AKT that regulate the activity of downstream targets including GSK3β (Fig. 5E).

The PI3K pathway plays a role in self-renewal and cell cycle regulation in ES cells. ES cells with hyperactivated PI3K signalling exhibited increased viability and proliferation rates (Sun et al., 1999) whereas blocking PI3K resulted in a decrease in the expression of *CyclinD1* followed by an increase in the proportion of cells in the G1 phase of the cell cycle (Jirmanova et al., 2002). Additionally, ES cells expressing a constitutively active form of the downstream component *Akt1* could self-renew independent of LIF (Watanabe et al., 2006), while blocking this pathway induced ES cell differentiation (Paling et al., 2004). The role of PI3K in self-renewal may be a consequence of its cell cycle regulation. ES cells have a distinct cell cycle characterised by promiscuous expression of cyclins throughout all cell cycle phases as well as an unusually short G1 phase (Faast et al., 2004; Stead et al., 2002). These distinct cell cycle properties are rapidly lost upon differentiation hence modification of the cell cycle by manipulating the PI3K pathway may impact on self-renewal. However, PI3K signalling also regulates the expression of the pluripotency gene *Nanog*, although this seems to be via an indirect mechanism involving GSK3β. When the PI3K pathway was inhibited, *Nanog* expression also decreased. However, this was rescued upon inhibition of GSK3β (Watanabe et al., 2006) suggesting that phosphorylation of GSK3β by AKT1 may be an important mechanism of pluripotency regulation.

1.3.5 Heterogeneity of embryonic stem cells

As ES cells are isolated from a single embryonic location and are defined based on a common set of properties, it was initially assumed that they represented a homogeneous cell type. However, there is now a large amount of evidence showing that ES cells heterogeneously express many genes including *Nanog*, *Hex*, *Stella*, *Rex1* and *Zscan4* (Canham et al., 2010; Chambers et al., 2007; Falco et al., 2007; Hayashi et al., 2008; Kobayashi et al., 2009; Singh et al., 2007; Toyooka et al., 2008). ES cells are therefore a mixture of different subpopulations, perhaps recapitulating the heterogeneity in the early pre-implantation embryo from which they are derived. However, there are numerous genes that are expressed homogeneously in the blastocyst but become heterogeneous in ES cells (e.g. *Bmp4*) and vice versa (e.g. *C-myc*) (Tang et al., 2010), hence we must keep in mind that ES cells do not entirely recapitulate the embryonic environment. In the early blastocyst, lineage specific transcription factors such as *Nanog* and *Gata6* are expressed in a heterogeneous manner and are thought to represent precursor populations of the epiblast and PE (Section 1.2.3.3). *In vivo* this heterogeneity exists only for a limited time period before lineage commitment. However, in ES cell cultures, heterogeneity is preserved offering a model to study the earliest stages of lineage specification.

As in the early blastocyst, a number of ES cell subpopulations are biased in differentiation towards a particular lineage, referred to as ‘lineage-priming’. ES cells that express low levels of *Nanog*

coexpress pluripotency-associated factors and lineage-specific markers (MacArthur et al., 2012) and have an increased propensity to differentiate (Chambers et al., 2007; Singh et al., 2007). Lineage-priming also appears to be regulated by a Notch target, *Hes1*. *Hes1* low ES cells differentiate efficiently into neural cells while *Hes1* high cells differentiate efficiently into mesoderm (Kobayashi et al., 2009). A population of undifferentiated ES cells also express the endoderm gene *Hex* and are primed towards an extraembryonic endoderm fate (Canham et al., 2010) (Section 1.4.4).

It is becoming apparent that many different ES cell subpopulations exist and that these are not completely discrete, but show a degree of overlap i.e. *Nanog* and *Stella* expressing populations are both thought to represent epiblast-like precursors but *Nanog* expressing cells make up approximately 80% of ES cells, while *Stella* positive cells make up only 30% (Chambers et al., 2007; Hayashi et al., 2008; Singh et al., 2007). Additionally, quantitative immunostaining showed that 80% of cells express REX1, 75% KLF4, 50% NANOG and 30% TBX3. In only 4% of cells KLF4, NANOG and TBX3 were coexpressed (Niwa et al., 2009), even though all are pluripotency-associated genes. In this thesis I have analysed the expression patterns of a number of these genes in early blastocyst stage embryos also observing that many subpopulations are present (Chapter 4). These subpopulations may either represent a complex mix of functionally relevant cell types or else be a result of delayed gene expression between upstream and downstream targets within a single gene network.

ES cell populations are not stable and changes in gene expression profiles occur over relatively short periods of time. The expression of certain genes fluctuates over a regular time period, for example *Hes1* oscillations occur over 3-5 hours (Kobayashi et al., 2009). Although it is known that other subpopulations can interconvert, it is not known whether ES cells are continuously and regularly cycling from one state to another. *Nanog* high or low cells can switch to the opposing state within 24 hours (Kalmar et al., 2009), while *Stella* high and low expressing cells take between 5 and 7 days to regenerate mixed populations upon separation (Hayashi et al., 2008). As discussed in Section 1.2.3.3, it is not known whether the same changes can be observed in the early mouse embryo. Based on these observations, a model has been developed whereby ES cell cultures are a dynamic equilibrium of subpopulations, each primed towards specific fates in differentiation (Canham et al., 2010). If a cell is exposed to a differentiation-promoting signal while in a particular state (e.g. *Hex* positive), it is more likely to differentiate in that direction (e.g. towards endoderm).

Although heterogeneity exists, the origin is still unclear. Heterogeneity may arise through stochastic transcriptional bursts reinforced by negative and positive feedback mechanisms (Elowitz et al., 2002; Kaern et al., 2005; Raj et al., 2006). Additionally the random segregation of mitochondria at cell division introduces variability in the rate of transcription between individual cells (das Neves et al., 2010) that could result in different levels of gene expression. Alternatively, the fluctuations could be caused by a deterministic, intrinsic factor that oscillates over a regular time period, possibly implied by the regular fluctuations of *Hes1* (Kobayashi et al., 2009). A potential oscillatory regulator could be

the circadian rhythm, although canonical circadian genes have not been detected in ES cells (Yagita et al., 2010). Another regular oscillator is the cell cycle. The pluripotency-associated genes *Oct4*, *Nanog*, *Esrrb* and *Sox2* have all been implicated in cell cycle control or association with cell cycle regulators (Card et al., 2008; van der Laan et al., 2013; Yang et al., 2011; Zhang et al., 2009). Additionally, it has recently been shown in human ES cells that various signalling pathways become activated in distinct phases of the cell cycle and affect the capacity of cells within that phase to generate different lineages (Pauklin and Vallier, 2013; Singh et al., 2013). Nevertheless, as in the mouse embryo, decisive evidence as to the source of heterogeneity remains elusive.

1.3.6 Analysis of heterogeneity by single cell gene expression analysis

The observation that ES cells are not homogeneous highlights the importance for single cell research over population-based assays to truly understand the range of different cell types that are present within a culture. It is now possible to do high throughput single cell gene expression analysis in ES cells using systems such as the Biomark from Fluidigm, where 96 single cells can be analysed for the expression of 96 distinct genes at one time. However, normalisation of such data can be problematic. It has been observed that, consistently, 1/5 cells have significantly lower levels of gene expression than all other cells, indicating that there may be differences in cellular viability or cell cycle phase that should be taken into account (Tang et al., 2006). Additionally, due to the expense of equipment and protocols, there is currently only a fairly small set of single cell expression data in ES cells.

Single cells have been analysed from subpopulations sorted based on their level of *Nanog* or *Stella* expression. *Stella* high cells tended to express high levels of ES cell-associated genes such as *Pecam-1*, *Rex1*, *Nanog* and *Sox2* while *Stella* low cells expressed high levels of differentiation-associated genes including *Fgf5* and *Gbx2*. There was little variability in the expression of *Oct4* (Hayashi et al., 2008) and, in *Nanog* low cells *Oct4* was coexpressed with lineage-associated markers (MacArthur et al., 2012). Below a certain threshold of *Nanog* expression, the variation in the expression of other genes increased (Trott et al., 2012), likely corresponding to an increase in the promiscuous expression of lineage-specific markers.

However, in single cell analysis there are not always clear expression patterns within groups of associated genes as may be expected. There was no statistically significant correlation in the expression of 8 genes (*Sox2*, *Nanog*, *Oct4*, *Rex1*, *Gbx2*, *Stella*, *Pecam-1*, *Fgf5*) simultaneously analysed in ES cells, although a number of these are associated with the pluripotent state. However, in many cases, low levels of *Rex1* corresponded to high expression of *Fgf5* (Trott et al., 2012). Likewise, in other experiments single cell clustering did not generate any clearly defined populations (MacArthur et al., 2012) suggesting that heterogeneity could be stochastic rather than representing defined, physiologically relevant cell types.

1.3.7 Single cell *in vivo* functional potential

As ES cells are not homogeneous, it is essential to understand how gene expression differences within individual cells can affect their functional properties, for example their ability to contribute to the embryo in chimaeras. Experiments using genetically labelled ES cells found that in the majority of cases only 1 or 2 cells, from the 10-15 injected, contributed to the embryo (Wang and Jaenisch, 2004). When single ES cells were injected into blastocysts, they were seen to contribute to the embryo proper and trophoblast, and at low levels to the parietal endoderm (Beddington, 1983; Beddington and Robertson, 1989; Lallemand and Brulet, 1990).

However, a difficulty when studying the potential of single ES cells is to distinguish their intrinsic properties from the contribution of external factors. For example, the incorporation of 10 ES cells into an embryo could compromise its viability hence the number of contributing cells may be limited for this reason. In this situation, which particular cells contribute may be random rather than due to inherent differences in their functional capacities. In fact, the number of injected cells does influence the behaviour of both the ES cells and host embryo. When a single ES cell was introduced into a morula stage embryo, it underwent more cell divisions than when multiple cells were introduced (Saburi et al., 1997). In 24 hours a single cell divided to produce 2 cells, whereas when 2-5 cells were introduced they did not divide within the same time period. Therefore, the potential of an ES cell to contribute to the embryo is dependent on the relative properties of the surrounding ES cells. The question then arises as to whether it is better to study the potency of single cells by single cell injections or by the injection of multiple labelled cells. These experiments also suggest that cells introduced into the embryo communicate to one another. When there are multiple cells they may produce an optimum level of factors, such as LIF, that can be detected by other cells as sufficient to sustain an ES cell niche, or else a level of factors that is inhibitory to further growth. On the other hand when a single cell is present it could go through multiple divisions before this threshold is reached. Although proliferating more rapidly, when analysed at later developmental stages, single cell injections produced a lower percentage of chimaeras and degree of chimaerism (Saburi et al., 1997). Such discrepancies are difficult to explain hence I have studied in more detail the ability of ES cells to contribute to embryos when injected as single or multiple cells. This data can be found in Chapter 3 and in the appendix of this thesis.

The relationship of gene expression to the ability of single cells to contribute to the embryo has also been briefly analysed. ES cells were sorted based on their expression of the ES cell markers PECAM-1 and SSEA-1. Single PECAM-1 negative, SSEA-1 negative cells contributed to extraembryonic tissues and most likely corresponded to differentiated cell types (Furusawa et al., 2004). PECAM-1 positive cells, either SSEA-1 positive or negative, contributed to the epiblast, however SSEA-1 positive cells contributed more efficiently, even though both populations exhibited equivalent expression levels of the canonical ES cell markers *Oct4*, *Nanog* and *Rex1*.

1.3.8 Epiblast and extraembryonic stem cell lines

Much later than the isolation of ES cells came the isolation of distinct cell lines representative of other lineages of the pre-implantation and early post-implantation embryo. These cell lines include extraembryonic endoderm (XEN) cells, trophoblast stem (TS) cells and epiblast stem cells (EpiSCs). XEN and TS cells are of particular interest as, until recently, it was not possible to model the molecular events occurring in extraembryonic tissues *in vitro*.

EpiSCs can be derived either from post-implantation embryos between E5.5 and E7.5 or from pre-implantation blastocysts at E3.5 (Najm et al., 2011; Tesar et al., 2007). At E3.5 the epiblast is not yet formed hence cells from this stage must differentiate during the derivation process. To derive EpiSCs from post-implantation embryos, the epiblast was dissected from the extraembryonic trophoblast and endoderm and then grown on irradiated mouse embryonic fibroblast (EMFI) feeders in medium supplemented with ACTIVIN and FGF2, similar to the conditions used to maintain human ES cells. However, addition of ACTIVIN was not necessary and at high levels promoted differentiation (Tesar et al., 2007). EpiSCs could be derived from pre-implantation embryos by culture of the ICM in growth factor-free medium. In these conditions, some ICMs would generate EpiSC outgrowths and others ES cell outgrowths. This was dependent both on the culture conditions and the genetic background of the mice (Najm et al., 2011). After derivation, pre-implantation EpiSCs were expanded in medium containing FGF2, and grew as colonies that were bigger and flatter than ES cells. EpiSCs isolated by either method have comparable gene expression profiles and methylation states that correspond more closely to human than mouse ES cells. The expression of ICM markers including *Tbx3*, *Pecam-1*, *Gbx2* and germ cell markers are decreased, and epiblast markers *Eomes*, *Brachyury* and *FoxA2* are increased in EpiSCs relative to ES cells (Tesar et al., 2007). Although it was not possible to generate chimaeras with EpiSCs, these formed teratomas when injected under the kidney capsule of mice, hence are pluripotent (Tesar et al., 2007).

TS cells are self-renewing cell lines derived from the ExE and express markers common to early diploid trophoblast. Like EpiSCs, TS cells can be isolated from E3.5 blastocysts or E6.5 embryos (Tanaka et al., 1998). At E6.5 the extraembryonic region is isolated and disaggregated with trypsin before being plated on EMFIs or in MEF-conditioned medium. The conditions used for TS cell derivation are very similar to those used for EpiSCs, comprising of FGF4 and Heparin. In the place of MEF-conditioned medium, TS cells can be maintained in medium containing serum, FGF4 and TGF-ß (Erlebacher et al., 2004). Upon removal of any of these factors, the epithelial-like colonies begin to differentiate into flattened giant cells. TS cell derivation is not specific as colonies of cells with distinct morphologies also survive and proliferate hence TS cells must be selectively picked and expanded (Tanaka, 2006). The transcription factor *Cdx2*, important for trophoblast lineage segregation (Section 1.2.3.2), is also necessary for TS cell self-renewal (Niwa et al., 2005). However, cells can

still generate differentiated trophoblast in the absence of *Cdx2* highlighting key differences between the capacity of cells to give rise to a self-renewing trophoblast line and the capacity to differentiate into post-mitotic trophoblast.

When TS cells were introduced into embryos in chimaera assays, they contributed only to trophoblast-derived tissues (Tanaka et al., 1998). However, the degree of contribution was variable with only 25% of chimaeras showing contribution to all 3 early trophoblast derivatives, the ExE, the EPC and TGCs. A further 25% of chimaeras exhibited contribution to only 1 of these lineages, although the degree of contribution is not stated. The same percentage (25%) of ES cell generated chimaeras demonstrated contribution to TGCs, although generally at low levels and in combination with embryonic contribution (Beddington and Robertson, 1989).

XEN cells can be isolated in various ways; either by plating whole blastocysts or ICMs into TS cell conditions (see above) or else by plating E3.5 blastocysts into standard ES cell medium containing LIF (Kunath et al., 2005). Although these conditions also gave rise to TS cells, the majority of the resulting colonies were XEN cells. The cells can be grown on EMFIs or gelatin. If gelatin is removed XEN cells differentiate and halt proliferation (Kunath et al., 2005). XEN cell cultures are a mixture of different cell states with distinct morphologies, either rounded cells with large processes reminiscent of the DVE/AVE, or epithelial cells. These states are dynamic and single cells can convert between them without cell division (Kunath et al., 2005). XEN cells are mostly parietal endoderm in character, expressing high levels of ECM proteins associated with the Reichert's membrane as well as transcription factors such as *Gata4* and *Sox7*, but no expression of the visceral endoderm marker *AFP*. Additionally, in the majority of cases, XEN cells contributed only to parietal endoderm in chimaera assays, and in 1 case to the visceral endoderm (Kunath et al., 2005). However, the addition of laminins and BMP promotes a more epithelial morphology and upregulation of visceral endoderm markers (Paca et al., 2012).

In addition to XEN cells, other extraembryonic endoderm-like cell lines (visceral endoderm-like END2 and parietal endoderm-like PYS2 cells) have been derived, although from EC cell lines rather than directly from the embryo (Lehman et al., 1974; Mummery et al., 1991). All cell lines have distinct differences in gene expression although consistently no definitive endoderm markers are expressed (Brown et al., 2010). Only XEN cells have an intact canonical MAPK pathway, known to be critical in PE formation (Section 1.2.3.3) (Brown et al., 2010).

1.4 Hex

In this final introduction section I will introduce the endoderm marker *Hex*, used throughout this thesis to mark an ES cell population primed towards a PE fate. A fluorescent reporter construct for this gene, described in Section 1.4.4, has facilitated the study of early lineage-priming events in ES cells and pre-implantation mouse embryos.

1.4.1 Hex structure

Hex (Haematopoietically-expressed homeobox) is a divergent homeobox protein isolated from haematopoietic cells. The *Hex* gene contains 4 exons generating a protein of 30kDa with amino (N) and carboxyl (C) terminal domains as well as a homeodomain (Ghosh et al., 1999). The homeodomain is a mediator of DNA binding. However, HEX can repress transcriptional activity independent of the homeodomain (Bess et al., 2003; Guiral et al., 2001; Soufi et al., 2006; Swingler et al., 2004; Topcu et al., 1999). The N-terminal proline-rich domain allows oligomerisation of HEX proteins, as well as facilitating other protein interactions. The C-terminal acidic domain is involved in transcriptional activation (Kasamatsu et al., 2004).

1.4.2 Hex expression during mouse development

In situ hybridisation suggested that *Hex* is first expressed at E4.5 of mouse development, specifically in the PE (Thomas et al., 1998). However, more recent studies using more sensitive techniques, as well as the results presented in this thesis (Chapter 3), revealed that *Hex* is expressed already in the 2-cell embryo and at higher levels than *Oct4* (Guo et al., 2010).

Hex expression in the PE is rapidly restricted to the visceral endoderm at the distal tip of the embryo. These DVE cells will subsequently migrate anteriorly to form the AVE, a signalling centre important for generating the anterior-posterior mouse axis (Thomas et al., 1998) (Section 1.1.4).

At approximately E7.0, *Hex* is also expressed in the extraembryonic mesoderm within the nascent blood islands (Ghosh et al., 2000; Thomas et al., 1998). This expression is lost as the endothelial precursors differentiate. Later, *Hex* is also expressed in the embryonic mesoderm in the region that will form the heart and also in mesoderm fated to form the endothelial cells, particularly the cranial vasculature (Thomas et al., 1998).

Later in development, *Hex* is expressed in a wide variety of tissues, first in the anterior definitive endoderm, then in tissues such as liver, thyroid, pancreas and skin (Bort et al., 2004; Hunter et al., 2007; Keng et al., 2000; Martinez Barbera et al., 2000; Obinata et al., 2002; Puppin et al., 2004).

Hex expression is regulated by distinct enhancer regions, one 4.2kb upstream of *Hex* is necessary for expression within the liver, thyroid and endothelial cells while an intronic enhancer is necessary for correct expression within the AVE and anterior definitive endoderm (Rodriguez et al., 2001).

1.4.3 Hex function

HEX acts both as a transcriptional activator and repressor. It has been suggested that the repressor function is mediated by binding of HEX to target genes as an octamer, creating morphological distortions in DNA that prevent transcription (Williams et al., 2008). However, evidence for this was obtained *in vitro* hence it is unknown whether HEX acts in the same manner *in vivo*. As well as regulating transcription by DNA binding, HEX acts indirectly by binding to transcription factors, such as GATA2 and AP-1, and decreasing their affinity to bind and activate their downstream targets (Minami et al., 2004; Schaefer et al., 2001).

HEX can also act as a post-transcriptional modifier by inhibiting the transport of mRNAs from the nucleus to the cytoplasm, therefore preventing their translation. HEX binds to and disrupts the function of eIF-4E, a protein involved in mRNA transport, preventing the export of mRNAs important for growth and proliferation, including *CyclinD1* (Topisirovic et al., 2003).

HEX interacts with numerous signalling pathways. In Xenopus, HEX anteriorises embryos by acting synergistically with the canonical Wnt pathway through inhibition of the Wnt antagonist TLE4 (Zamparini et al., 2006). HEX also interacts with the TLE/GROUCHO family of co-repressors in haematopoietic cells (Swingler et al., 2004). Additionally HEX inhibits *Nodal* target genes both in ES cells and Xenopus (Zamparini et al., 2006).

Hex null embryos exhibit an anterior truncation, both in mouse and Xenopus models, as well as liver and thyroid dysplasia. HEX also plays a role in haematopoiesis and vascular development (Brickman et al., 2000; Martinez Barbera et al., 2000).

1.4.4 Hex reporter construct and embryonic stem cell heterogeneity

A sensitive *Hex* reporter was previously generated in our lab (Canham et al., 2010) and is used throughout this thesis (Fig. 6A). The construct contains *Hex* cDNA followed by 10 copies of a 9-nucleotide internal ribosome entry site motif (Gtx-IRES) (Chappell et al., 2000, 2004). This motif facilitates the translational amplification of a downstream yellow fluorescent protein, Venus. The Gtx-IRES motif includes a region of nucleotides homologous to the 18S ribosomal subunit, hence recruits ribosomes more efficiently for translation. Therefore, low levels of *Hex* transcription can be detected due to amplification of the fluorescent reporter.

The *Hex*-venus (HV) construct was inserted into the first exon of the *Hex* locus in ES cells, by homologous recombination, and these cell lines used to report on endogenous *Hex* expression in both ES cells and early differentiation. These ES cells were also used to generate HV reporter mice (Dr. M. Canham) by blastocyst injection.

Using this reporter, *Hex* expression in ES cells was shown to be heterogeneous, but within the $OCT4^{+}$ self-renewing ES cell population (Canham et al., 2010) (Fig. 6B). This contradictory coexpression of pluripotency-associated genes with a lineage-specific marker was unexpected and hence this population was analysed in more detail. *Hex* high and low sorted ES cells could interconvert to regenerate mixed cell populations. Additionally, FGF, a known PE-promoting factor (Section 1.2.3.3), was found to increase *Hex* expression and upon exit from the ES cell state towards differentiated endoderm, the level of *Hex* expression increased further.

Global gene expression analysis of *Hex* high and low ES cells revealed that *Hex* high cells slightly upregulated many PE genes. Additionally, these cells contributed to both to epiblast and extraembryonic endoderm when reintroduced into embryos. Conversely, *Hex* low cells expressed higher levels of epiblast markers, and contributed only to the epiblast in chimaera assays (Canham et al., 2010). This suggests that *Hex* high ES cells represent an early PE precursor that is not fully differentiated and retains the capacity to convert to an epiblast-primed (*Hex* low) state.

In this thesis, I follow up on these observations. A large part of this work involves fluorescence activated cell sorting (FACS) of HV low (HV^{-}) and HV high (HV^{+}) ES cell populations (routinely selected based on the lower and upper 25% of HV expression respectively). Figure 6C contains an example of this kind of analysis, and the gating methodology used. As I was interested in analysing early lineage-priming events occurring in ES cells, unless otherwise stated, HV^{-} and HV^{+} cells were sorted from a population of cells also expressing the ES cell marker PECAM-1 or SSEA-1. DAPI (4',6-diamidino-2-phenylindole) positive (non-viable cells) and PECAM-1/SSEA-1 negative (differentiating cells) were excluded during the sorting process (Fig. 6C). Purity checks were also carried out after sorts to confirm accurate separation of these populations (Fig. 6D).

In this thesis, I have used a combination of the HV reporter, different culture conditions and cell sorting to show that ES cells cultured in 2i contain a population of cells that can contribute to embryonic, extraembryonic endoderm and trophoblast lineages *in vivo*, that this cell population is promoted by LIF and finally that the switch between HV^{-} and HV^{+} ES cell states is transcriptionally regulated primarily at the point of elongation.

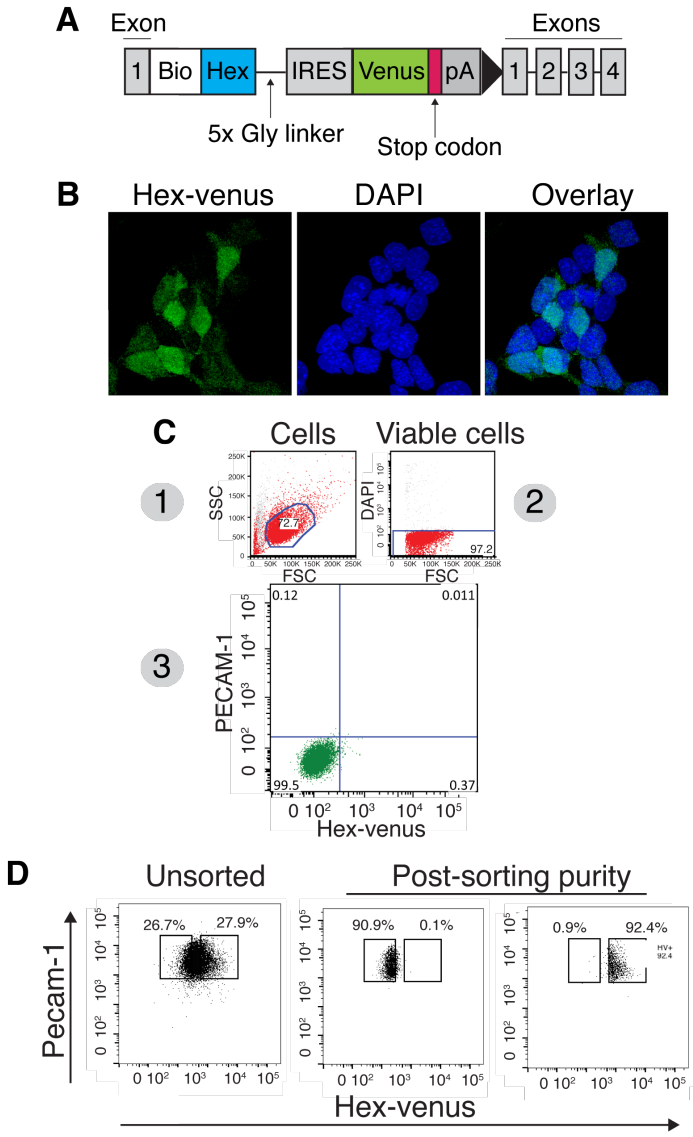


Figure 6. Hex-venus (HV) ES cell reporter cell line. A. Schematic diagram of the HV reporter construct inside the endogenous Hex locus. The construct consists of Biotin-tagged Hex cDNA follwed by an internal ribosome entry site (IRES) that causes the translational amplification of the downstream Venus fluorescent protein. The construct is inserted into the Hex locus, disrupting exon 1. As the construct contains Hex cDNA, endogenous expression levels are maintained. **B.** Confocal optical section of undifferentiated HV ES cells. **C.** Standard gating scheme used when sorting HV ES cells by FACS. Cells were separated from debris by gating on size based on forward (FSC) and side (SSC) scatter (1).DAPI negative cells were then gated to exclude any non-viable, $DAPI^+$ cells (2). Unstained control E14 cells (wild type parental cell line without any fluorescent reporter) were used to set HV negative, PECAM-1 negative gates (3). **D.** Routinely, the lower and upper 25% of HV expressing, PECAM-1 positive cells were sorted to define HV low (HV^-) and HV high (HV^+) populations respectively. PECAM-1 is a marker of undifferentiated ES cells hence allows the exclusion of differentiated cells. Purity checks were carried out after sorting to confirm the correct separation of these populations.

CHAPTER 2: MATERIALS AND METHODS

Chapter 2: Materials and methods

2.1 Cell culture

2.1.1 Cell lines

The following mouse ES cell lines were used in this study: E14Tg2a (E14), HV, HV LacZ-IRES-puromycin (LIP), HV H2B-Tomato and *Stat3*$^{-/-}$ ES cells (kindly supplied by Dr. Jennifer Nichols). All of the above ES cell lines are from a 129 background. In addition, the Transgenic Core Facility (TCF) at the University of Copenhagen (Appendix) derived F1 (C57BL/6, 129) hybrid lines in various culture conditions.

For derivation in 2i, E2.5 embryos were flushed from oviducts in M2 medium and cultured in KSOM (Millipore) for approximately 8 hours until the evening where they were transferred to KSOM + 2i (1 μM PD0325901, Stemgent, 3 μM CHIR99021, Biovision). Embryos were cultured overnight until E3.5 blastocysts then transferred to individual wells of gelantinised 96-well plates containing 2i/LIF ES cell culture medium (2i inhibitors in N2B27 medium). Embryos were cultured for 5-8 days at 37°C, 5% CO_2, 90% humidity, until a cluster of cells generated from the ICM emerged. Colonies were picked and dissociated into smaller clusters in Accutase for 3-5 minutes at 37°C before being plated into fresh medium for expansion. The same protocol was used to derive ES cells in serum and KOSR conditions but with pre-culture of embryos in KSOM alone and then transfer to 96-well plates containing either serum/LIF or KOSR/LIF.

HV and HV LIP cells were generated by Dr. M. Canham (Brickman lab) (Canham et al., 2010). HV H2B-Tomato cells were generated by the introduction of an H2B-Tomato fusion protein (a kind gift from Heiko Lickert) into the previously described HV cell line (see 2.1.9).

2.1.2 Standard embryonic stem cell culture conditions

ES cell lines were cultured in tissue culture flasks or plates (IWAKI, Corning) coated with 0.1% gelatin (Sigma) for at least 10 min. ES cells were cultured in Glasgow modified Eagle's medium (GMEM, Sigma) containing, non-essential amino-acids (Gibco), glutamine and sodium pyruvate (Gibco), 0.1 mM β-mercaptoethanol (Sigma) and 10% Foetal Calf Serum (FCS) with 1000 U/mL LIF. Each time a new batch of medium was made, an overnight test for bacterial contamination was carried out adding 5 mL of medium to 5 mL tryptose phosphate broth solution (Sigma). The next day the tester was checked for cloudy bacterial growth. Cells were cultured at 37°C, 5% CO_2 and 90% humidity (BBD 6220 incubators, Thermo Scientific).

ES cells were passaged every 2 days upon reaching approximately 80% confluency. Cells were passaged with either 0.1% trypsin (Life Technologies) when cultured in serum conditions, or with Accutase (Millipore) when cultured in serum-free conditions. Cells were washed in PBS (Sigma) then either trypsin or Accutase added and incubated at 37°C for 3 min. Cells were then collected and the reaction neutralized either by serum-containing medium (in the case of trypsin) or PBS (in the case of Accutase). Cells were pelleted at 1300 rpm for 3 min and 1/5 of cells re-plated into fresh medium.

ES cells were frozen for long-term storage in the GMEM medium described above containing 10% DMSO (Sigma). Cells were frozen overnight in cryovials (Thermo Scientific) at -80°C to allow gradual freezing to prevent damaging crystal formation with the cells. Frozen vials of cells were then transferred to liquid nitrogen storage (CBS Isothermal Carousel V3000EH-AB/C). Cells were thawed rapidly by placing cryovials in a 37°C incubator and then immediately pipetting the cells into 18 mL of pre-warmed medium to remove residual DMSO. Cells were centrifuged and re-suspended in fresh medium for plating.

2.1.3 2i culture of embryonic stem cells

ES cells were cultured in defined serum-free medium in the presence of 2 inhibitors (2i) according to the protocol developed in the Smith lab (Ying et al., 2008). 2i medium is comprised of N2B27 basal medium (Stem Cell Sciences), 1 μM PD0325901 (MEK inhibitor, Stemgent) and 3 μM CHIRON99021 (GSK3 inhibitor, BioVision). LIF can also be added to 2i medium at a standard concentration of 1000 U/m. Medium was replaced daily and cells were cultured in these conditions for at least 5 days before beginning experiments. *Stat3*$^{-/-}$ ES cells were routinely cultured in 2i/LIF to prevent differentiation.

2.1.4 Small molecule inhibitors and cytokines

The details of small molecule inhibitors and cytokines used in this thesis can be found in Table 1. All inhibitors were diluted, stored and used according to manufacturers instructions.

2.1.5 Fluorescence-activated cell sorting (FACS) and flow cytometry

ES cells were collected as in 2.1.2 and washed in PBS. ES cells were then re-suspended in FACS buffer (PBS with 10% FCS) with the appropriate dilution of primary antibody (Table 2) and incubated on ice for 10 min. Cells were then washed 3 times with FACS buffer and re-suspended in FACS buffer containing the appropriate dilution of secondary antibody (1:1000 for Alexa fluors, Life Technologies) and incubated on ice for a further 10 min. ES cells were washed 3 times in FACS buffer and re-suspended in FACS buffer containing 100 ng/mL DAPI (Life Technologies) to exclude dead cells. Flow cytometry analysis was carried out on a FACS Calibur or Fortessa (BD Bioscience).

Cell sorting was completed on a FACS Aria Cell Sorter II SORP (BD Bioscience). Unstained E14 ES cells were used as a non-fluorescent control cell line in all sorting experiments. Cells were collected in the same medium as they would be plated in, or else in the same medium that they were cultured in when being sorted for RNA extraction. When sorting single cells, these were sorted directly into gelatinised 96-well plates containing the appropriated medium or else into lysis buffer for single cell qRT-PCR (Section 2.2.9). Flow cytometry data was analysed using FlowJo software (Tree Star).

Inhibitor	In text	Target	Working Concentration	Company	Catalogue Number
CHIR99021	CHIR	GSK3	3 µM	Biovision	1677-5
InSolution JAK inhibitor I	JAKi	JAK1-3	100 nM - 5 µM	Millipore	420097
LY294002	LY	PI3K	5 µM	Cell Signalling	9901
PD0325901	PD	MEK	1 µM	Stemgent	04-0006

Table 1. Details of small molecule inhibitors used in this study. Table shows the details of all small molecule inhibitors as well as the name that is used to refer to these in the text of the results section.

Antibody	Dilution	Use	Company	Catalogue Number
Brachyury	1:200	Immunostaining	R&D	AF2085
Cdx2	1:200	Immunostaining	Biogenex	MU392A-UC
Cytokeratin7	1:100	Immunostaining	Santa Cruz	sc70936
Phospho-ERK	1:1000	Western blot	Cell Signaling	4370
Total-ERK	1:1000	Western blot	Cell Signaling	9102
Gata4	1:200	Immunostaining	Santa cruz	sc1237
Gata6	1:100	Immunostaining	R&D	AF1700
Gata6	1:200	Immunostaining	Cell Signaling	5851
GFP Alexa488 conjugated	1:200	Immunostaining	Life Technologies	A21311
Ki67	1:200	Immunostaining		
Ki67-APC conjugated	1:50	Immunostaining	eBioscience	51-5698
Klf4	1:200	Immunostaining	R&D	AF3158
LIFR	1:200	Flow cytometry	R&D	MAB5990
Nanog	1:200	Immunostaining	eBioscience	14-5761
PDGFRα APC conjugated	1:200	Flow cytometry	eBioscience	12-1401
Pecam-1 APC conjugated	1:500	Flow Cytometry	BD Bioscience	551262
Phospho-Stat3	1:200	Immunostaining	Cell Signaling	9145
Tubulin	1:5000	Western blot	Abcam	ab6160
Tuj1	1:1000	Immunostaining	Covance	mms-435p

Table 2. Details of antibodies used in this study. Table showing antibodies used, for which application and at which dilution.

2.1.6 LIF withdrawal differentiation

LIF withdrawal differentiation was carried out as previously described (Nishikawa et al., 1998). 3 x 10^3 ES cells were plated into gelatin-coated 6-well plates in serum-containing medium without LIF. After 5 days, the majority of cells had differentiated. Fresh medium was added daily.

2.1.7 Neural differentiation

Monolayer neural differentiation was carried out as previously described (Ying et al., 2003b). 10^4 ES cells were plated into a gelatin-coated 6-well plate in N2B27 medium. Medium was replaced as needed due to cell death. Cells were cultured for 7-9 days in these conditions.

2.1.8 Trophoblast differentiation

10^4 ES cells were plated onto gelatin-coated 6-well plates in TS cell conditions. Medium comprised of 70% MEF-conditioned medium (R&D) and 30% TS cell medium (RPMI - Gibco, glutamine and sodium pyruvate, 0.1 mM β-mercaptoethanol, 20% FCS). FGF4 (Peprotech) was added to medium at 25 ng/mL along with 1 μg/mL Heparin sulphate (Sigma). Cells were cultured for 7 days before analysis of differentiation levels by immunostaining.

2.1.9 Electroporation for stable integration of histone H2B-Tomato constitutive fluorescent reporter

The H2B-Tomato plasmid (Lickert lab) was digested overnight (as described in 2.2.2). 2-10 x 10^7 cells were trypsinised and collected then washed twice in PBS. Cells were counted with a haemocytometer and 10^7 cells were re-suspended in 800 μL of PBS. 800 μL cells were added to 2 electroporation cuvettes (BioRad). As a control, no DNA was added to one of the cuvettes while 100 μg of linearized DNA was added to the other. Electroporation was carried a Gene Pulser (BioRad) at 800 V voltage, 3 μF capacitance, 0.1 second time constant. Cells were transferred to pre-warmed medium and plated to 5 x 10^5 cells in 10 cm gelatinized dishes. Plates were incubated overnight and, the following day, 2 μg/mL Puromycin (Life Technologies) was added to the media. ES cells were expanded for approximately 2 weeks until colonies were large enough to pick. Cells were mechanically picked and dissociated in trypsin, which was then inactivated by transferring the cells to gelatinised 96-well plates containing pre-warmed medium. Cells were then further expanded and stocks frozen.

2.1.10 Immunostaining

ES cells were cultured on gelatin-coated plastic for standard fluorescence imaging or on gelatinised glass coverslips (VWR) for confocal imaging. ES cells were washed twice with PBS and fixed with 4% paraformaldehyde (PFA, Sigma) for 15 min at room temperature. Cells were washed a further 2 times and permeabilised for 10 min with PBS containing 0.1% Triton-X (Sigma) (PBST). Staining using the phospho-STAT3 (pSTAT3) antibody required permeabilisation in methanol (Sigma) at -20°C for 15 min. ES cells were blocked for 15 min in PBST containing 1% BSA (Life Technologies) and 3% serum (Sigma) of the host for which the secondary antibody is derived. The primary antibody was diluted to the appropriate concentration (Table 2) in PBST containing 1% BSA and incubated overnight at 4°C. The following day, cells were washed 3 times for 15 min with PBST before adding the Alexa fluor secondary antibody diluted in PBST with 1% BSA to a concentration of 1:800 and incubated at room temperature for 2 hours. All directly conjugated antibodies (such as Alexa488 conjugated anti-GFP, used to detect HV expression) were added at the same time as the secondary antibodies. ES cells were then washed a further 3 times for 15 min with 100 ng/mL DAPI being added to the final wash as a nuclear stain. For standard microscopy, on tissue culture plastic, a Nikon Eclipse Ti microscope was used. Glass coverslips were mounted onto glass slides using Vectashield mounting medium (Vector labs) and sealed with nail varnish. Slides were then imaged by confocal microscopy using a Leica TCS SP8.

2.2 Molecular biology

2.2.1 Preparation of genomic DNA

Genomic DNA for genotyping was extracted from mouse tails, which were digested overnight in lysis buffer (100 mM Tris, pH 8.5, 5 mM EDTA, 200 mM NaCl, 2% SDS) with 100 µg/µL proteinase K (Roche). Tails were centrifuged at full speed and the supernatant taken off and used to extract DNA with the DNeasy Blood & Tissue Kit (Qiagen) as per manufacturers instructions.

2.2.2 DNA restriction digest

100 µg plasmid DNA was digested overnight at 37°C in a volume of 450 µL with 45 µL NEB reaction buffer (NEB) and 5 µL of the appropriate linearising enzyme (NEB). DNA was then ethanol precipitated to remove any contaminating enzyme. 45 µL of 3 M sodium acetate and 1 mL of 100% ethanol (Sigma) was added before incubating on dry ice for 30 min. The sample was then centrifuged at maximum speed for 15 min and the liquid gently aspirated. 1 mL of 70% ethanol was added and the sample was again centrifuged at maximum speed for 5 min. The ethanol was aspirated and the sample allowed to air dry. Finally the sample was re-suspended in 50 µL pre-warmed nuclease-free H_2O (Ambion).

2.2.3 Agarose gel electrophoresis of DNA

1 or 2% agarose gels were made by diluting agarose powder in TAE buffer (50x buffer – 2 M Tris pH 8.0, 50 mM EDTA, 57.1 mL/L glacial acetic acid, all reagents from Sigma) and 1:10,000 dilution of SYBR Safe DNA gel stain (Life Technologies). DNA was mixed with 1x loading buffer (Fermentas) and loaded onto the gel. 1 kb plus DNA ladder (Life Technologies) was used to determine the size of DNA fragments. The gel electrophoresis was run at 100 V in TAE buffer for approximately 20 min until the loading dye reached the end of the gel. DNA was visualized with UV light using a Gel Doc XR+ (BioRad).

2.2.4 RNA extraction

RNA was extracted using the RNeasy Mini Kit (Qiagen) with on column DNA digestion using an RNase-Free DNase Set (Qiagen) as per manufacturers instruction so that RNA preparations were not contaminated with genomic DNA.

2.2.5 cDNA synthesis

cDNA was synthesized using SuperscriptIII reverse transcriptase (Life Technologies) according to manufacturers guidelines. 1 µg RNA was incubated at 65°C for 5 min with 1 µL of 1 µg random primers (Promega) and 1 µL of 10 mM dNTPs (Roche) in a volume of 14 µL. Samples were then incubated on ice for 1 min before adding 4 µL First strand buffer (Life Technologies), 1 µL of 0.1 M DTT (Life Technologies), 1 µL 40,000 U/mL RNaseOUT (NEB) and 1 µL of SuperscriptIII enzyme (Life Technologies). Samples were then incubated for 1 hour at 50°C then inactivated at 70°C for 15 min. 1 µL of 5000 U/mL RNase H (NEB) was added at 37°C for 20 min to remove the template RNA and then samples were diluted 1:100 with nuclease-free H_2O.

2.2.6 Nucleic acid quantification

Nucleic acid concentrations were determined using a Nanodrop spectrophotometer (Fisher Scientific). A 260/280 ratio above 1.8 for DNA and 1.5 for RNA indicated that the samples were not contaminated with organic compounds from the extraction process.

2.2.7 Polymerase chain reaction (PCR)

PCR for genotyping was carried out using the primers below. 15 µL of genomic DNA, isolated from mouse tails, was mixed with 5 µL 10x Coral Load PCR Buffer (Qiagen), 3 µL 5x Solution Q (Qiagen), 1 µL Taq Polymerase (Qiagen), 2.5 µL of each primer from a 10 µM stock (forward and reverse) and 3 µL nuclease-free H_2O. The PCR conditions were as follows:

1. 98°C – 5 min
2. 98°C – 30 sec
3. 59°C – 30 sec
4. 72°C – 30 sec
5. Go to step 2 30 times
6. 72°C – 10 min

2.2.8 Quantitative real time polymerase chain reaction (qRT-PCR)

qRT-PCR was carried out using either the SybrGreen (Life Technologies) or Universal Probe Library (UPL) system (Roche). Primers were designed using the UPL assay design centre (Roche). Primers were designed to span introns so that any remaining genomic DNA would not be amplified. Probes for each primer set were selected from the UPL library (Roche), which covers the whole of the mouse transcriptome. The level of probe fluorescence corresponds to the level of product generated in the PCR reaction. For the UPL system, reactions were carried out in 384-well plates (Roche) with 0.99 µL

of each 10 µM primer stock (forward and reverse), 0.1 µL probe, 5 µL LightCycler480 Probes Master Mix (Roche), 3 µL diluted cDNA. For SybrGreen, reactions were carried out with 5 µL LightCycler480 Master Mix (Life Technologies), 1 µL of each primer and 3 µL diluted cDNA. All cDNA samples were analysed in triplicates so that outlying values could be removed. Standard curves for each gene assay were generated by making serial dilutions of plasmids containing the gene of interest. Dilutions of 10^8, 10^6, 10^4 and 10^2 plasmid copies were used. This allowed an estimation of the copy number of each product within the samples. qRT-PCR reactions were carried out in a LightCycler480 (Roche). The samples were normalized for the amount of cDNA in each sample by calculating the values of each gene relative to the housekeeping gene TBP. Details of the primers used in this study can be found in Table 3.

2.2.9 Single cell qRT-PCR using the Biomark system

Single cells were sorted into 96-well PCR plates containing 5 µL CellsDirect reaction mix, 0.2 µL SuperScriptIII/Platinum Taq mix (CellsDirect One-Step qRT-PCR kit, Life Technologies), 2.8 µL DNA suspension buffer (TEKnova), and 1 µL 500 nM primer mix containing a mix of 48 DELTAgene Assays (Fluidigm) (Table 4). Controls of 100 and 1,000 cells were included. RT reaction conditions were 50°C, 15 min; 95°C, 2 min; and 22× (95°C, 15 s; 60°C, 4 min). An exonuclease step was performed to remove unincorporated primers at 37°C, 30 min and 80°C, 15 min. Amplification products were then diluted 5-fold in TE buffer (10mM Tris pH 8.0, 1mM EDTA). Amplified cDNA was mixed with SsoFast EvaGreen SuperMix with Low ROX (Bio-Rad). The same DELTAgene assays were used in qRT-PCR. Samples and assays were mixed with appropriate loading reagents and loaded onto a 96.96 gene expression Dynamic Array (Fluidigm). Samples were loaded in technical replicates. Arrays were read using a BioMark HD genetic analysis system (Fluidigm). Downstream analysis was completed in Microsoft Excel. Cells that expressed no or low levels of ACTB and GAPDH housekeeping genes, or with a Ct over 30, were excluded from further analysis. DELTAgene assays were custom designed by Fluidigm to cross introns and avoid amplifying genomic DNA. Assays showing poor melting curves were also excluded from analysis. Data were analysed without normalization or also normalized to the median expression of all genes across the array. No significant difference was observed in the data generated from either analysis method.

Gene	Forward Primer	Reverse Primer	Probe
Brachyury	ACTGGTCTAGCCTCGGAGTG	TTGCTCACAGACCAGAGACTG	27
Cdx2	CACCATCAGGAGGAAAAGTGA	CTGCGGTTCTGAAACCAAAT	34
Cytokeratin7	ACCCTCAACAACAAATTCGCGTCC	TGCTCTTGGCTGACTTCTGTTCCT	
Dab2	CAGTACCAGTTCCCACTCCAG	CTGTGCATTTACACCGACAAC	
FoxA2	GAGCAGCAACATCACCACAG	CGTAGGCCTTGAGGTCCAT	
Gata4	TTCGCTGTTTCTCCCTCAAG	CAATGTTAACGGGTTGTGGA	60
Gata6	GGTCTCTACAGCAAGATGAATGG	TGGCACAGGACAGTCCAAG	40
Goosecoid	GGAGACGAAGTACCCAGACG	CGGCGGTTCTTAAACCAG	32
Hex	CTACACGCACGCCCTACTC	CAGAGGTCGCTGGAGGAA	50
HNF4α	CTTGGCTCTGCGGTTCTG	CCGAGGGACGATGTAGTCAT	
Klf4	CGGGAAGGGAGAAGACACT	GAGTTCCTCACGCCAACG	62
Myc	CCTAGTGCTGCATGAGGAGA	TCTTCCTCATCTTCTTGCTCTTC	77
Nanog	GGACAGGTTTCAGAAGCAGAA	GGTTTTGAAACCAGGTCTTAACC	75
Nestin	CTGCAGGCCACTGAAAAGT	TTCCAGGATCTGAGCGATCT	2
Neurogenin1	TTTGGTCATGTTTCCACTTCC	CGGAGAAGGCAAGGTGTC	
Neurogenin2	ACATCTGGAGCCGCGTAG	CCCAGCAGCATCAGTACCTC	69
Oct4	GTTGGAGAAGGTGGAACCAA	CTCCTTCTGCAGGGCTTTC	95
Rex1	GCGGTGTGTACTGTGGTGTC	CCCTCAGCTTCTTCTTGCAC	67
Serpine2	TTGGGTCAAAAATGAGACCAG	CCTTGAAATACACTGCATTAACGA	64
Sox2	GGCAGAGAAGAGAGTGTTTGC	TCTTCTTTCTCCCAGCCCTA	34
TBP	GGGGAGCTGTGATGTGAAGT	CCAGGAAATAATTCTGGCTCA	97
Tead4	TCAGCCAGGCATATCACCAAGACA	TCAGCCAGGCATATCACCAAGACA	25

Table 3. List of primers used in standard qRT-PCR in this study. Forward and reverse sequences are listed for all primers used in this study. Where the UPL system was used, a probe number is also stated. Those without probe numbers were used with the SYBR green system.

Gene	Forward Primer	Reverse Primer
ACTB	CCCTAAGGCCAACCGTGAAA	CAGCCTGGATGGCTACGTAC
Dlk1	ATGGATTCTGCGAGGCTGAC	GGGGCAGTTACACACTTGTCA
Eomes	GCGGCAAAGCGGACAATAAC	ATCCAGTGGGAGCCAGTGTTA
GAPDH	AGACGGCCGCATCTTCTT	TTCACACCGACCTTCACCAT
Gata3	CCTACCGGGTTCGGATGTAA	CCGCAGTTCACACACTCC
Klf4	CAGGCTGTGGCAAAACCTATAC	CGTCCCAGTCACAGTGGTAA
Nanog	TCTGGGAACGCCTCATCAA	GAGGCAGGTCTTCAGAGGAA
Oct4	TCCCTACAGCAGATCACTCAC	CGCCGGTTACAGAACCATAC
Serpine2	CTCTGCCTCTGAGTCCATCA	CGATCAGATTTGGGGAAAGCA
Sox2	TGAAGGAGCACCCGGATTATA	CGGGAAGCGTGTACTTATCC
Stella	TGAAGAGGACGCTTTGGATGA	CCGGGGTTTAGGGTTAGCTT
Tcfap2a	CCAATGAGCAAGTGGCAAGAAA	GAGCCAGCAGGTCAGTGAA

Table 4. List of primers used in single-cell qRT-PCR in this study. The list shows the forward and reverse sequences of primers with defined melt-curves. The data for these genes was used for further analysis. Where melt-curves were not defined i.e. showed multiple peaks, data was not used for further analysis. All primers were designed using the Fluidigm DELTAgene assay design service.

2.2.10 Western blotting

ES cells were grown to approximately 80% confluency in a 6-well plate. Cells were washed quickly with ice cold PBS and lysed directly with 200 µL Laemmli buffer (BioRad). The sample was scraped using a pipette tip and collected into an eppendorf tube. Samples were kept on ice as much as possible. The samples were then sonicated using a Branson Digital Sonifier, following by boiling for 5 min. Samples were centrifuged at full speed for 5 min and a small amount of bromophenol blue (Sigma) added in order to visualise samples on the gel. The supernatants were loaded onto precast Novex 4-12% Bis-Tris acrylamide gels (Life Technologies) alongside 15 µL of Kaleidoscope protein standard marker (BioRad) to show the size of proteins bands present. Gel electrophoresis was carried out at 200 V for 45 min. 500 µL of NuPage antioxidant (Life Technologies) was added to the gel tank during electrophoresis to ensure that sensitive amino acids did not become oxidized. The separated proteins were transferred to Hybond ECL nitrocellulose membranes (Amersham) by wet transfer at 360 mA for 1 hour at 4°C. The blot was washed for 10 min in TBST (25 mM Tris, pH 7.4, 3 mM KCl, 140 mM NaCl, 0.1% Tween-20 (Sigma)) then blocked for 1 hour in TBST with 5% milk powder (Marvel). The primary antibody was diluted in TBST with 5% milk and incubated overnight at 4°C on a rolling platform. The membrane was then washed 3 times for 15 min in TBST. The secondary HRP conjugated antibody (Cell Signaling) was diluted 1:2000 in TBST with 5% milk overnight at 4°C on a rolling platform. The membrane was washed 3 times for 15 min in TBST and then developed using an ECL Western Blot Substrate (Pierce) and visualized by exposure to X-ray film (Kodak). The membrane can be stripped with 100 mM β-mercaptoethanol, 2% SDS and 62.5 mM Tris, pH 6.7 at 80°C for 30 min and then re-probed with another antibody to detect a housekeeping protein such as total ERK for the loading control.

2.3 Mouse experiments

2.3.1 Maintenance of mouse lines

Mouse lines were maintained under 12 hour light/dark cycles in the designated facilities at the University of Edinburgh and the University of Copenhagen. All animal work was carried in accordance with UK and European legislation and in particular according to the regulations described in the Animals (Scientific Procedures) Act of 1986 (UK). All work was authorized by and carried out under Project License 60/3715 issued by the UK Home Office.

2.3.2 Genotyping of Hex-Venus (HV) mouse line

Two HV mouse lines were generated from distinct ES cell clones, HVJu5.1 and HVJu9.3 by a previous PhD student in the lab, Dr. M. Canham. These lines were both genotyped following the procedures outlined in 2.2.1 and 2.2.7 for genomic DNA extraction and PCR using the forward primer 5'-CGGAGGCGAATCTGAAGCCAGC-3' in combination with the reverse primer 5'-GCATACAGCGGGACTCCCACG-3'. DNA fragments were separated on a 2% agarose gel (see 2.2.3). The wild type band was 220 bp and the transgenic band 300 bp.

2.3.3 Culture of pre-implantation embryos (E0.5-E4.5)

Embryos were flushed from either the oviducts (E0.5-E2.5) or the uterus (E3.5-E4.5) of pregnant female mice in PB1 solution (137 mM NaCl, 2.7 mM KCl, 0.8 mM $CaCl_2$, 1.5 mM KH_2PO_4, 0.5 mM $MgCl_2$ $6H_2O$, 8 mM Na_2HPO_4, 0.3 mM Na-pyruvate, 5.5 mM Glucose, 0.2 mM Penicillin, 0.07 mM Streptomycin, 0.3% BSA, all reagents from Sigma). Embryos were washed in PB1 and then transferred to KSOM medium (Millipore) where they can be cultured at 37°C, 5% CO_2 and 90% humidity up until E5.5.

It was observed that embryos cultured *ex vivo* showed delayed development. Instead of the PE segregating from the epiblast at E4.5, as occurs *in vivo*, segregation occurred at E5.5. Throughout this thesis in *ex vivo* experiments, embryos cultured until E5.5 will be referred to as E4.5, the *in vivo* stage that they correlate to.

2.3.4 Dissection of post-implantation embryos (E6.5 and E9.5)

E6.5 and E9.5 embryos were dissected from the uterus of pregnant female mice in M2 medium (Sigma) using a Leica M165C dissection microscope. The deciduae of E6.5 embryos were opened using forceps to separate the 2 halves and the embryos gently extracted. All extraembryonic

membranes were kept attached to the embryo in order to assess ES cell contribution to these tissues. The deciduae of E9.5 embryos were peeled away using forceps. Embryos were kept attached to the placenta for analysis.

2.3.5 Whole mount immunostaining of pre-implantation embryos (E0.5-E4.5)

The zona pellucida was removed from embryos by incubation in pre-warmed acid tyrodes solution (Sigma). Embryos were then fixed in 4% PFA for 15 min at room temperature then washed in PBS with 3 mg/mL of polyvinylpyrrolidine (Sigma, PBS/PVP). Embryos were permeabilised in PBS/PVP with 0.25% Triton-X for 30 min and then blocked for 15 min in PBS/PVP containing 2% donkey serum (Sigma), 0.1% BSA and 0.01% Tween-20. The primary antibody was then diluted in blocking solution and embryos were incubated overnight at 4°C (see Table 2 for antibody details). The following day embryos were washed 3 times for 15 min in blocking solution before being incubated for 2 hours at room temperature in the secondary antibody diluted in blocking solution. Alexa fluor secondary antibodies were used at a dilution of 1:500. Embryos were washed 3 times for 15 min in blocking solution with the last wash containing 100 ng/mL DAPI. Embryos were then mounted on glass slides using Vectashield mounting medium or imaged in an Attofluor chamber (Life Technologies) on a 25 mm glass coverslip.

2.3.6 Whole mount immunostaining of gastrulation stage embryos (E6.5)

Embryos were dissected and fixed in 4% PFA for 30 min at room temperature then washed in PBS/PVP. Embryos were permeabilised in PBST with 3, 5 min washes followed by one, 20 min wash. They were then blocked in PBST, rocking at 4˚C for 2 days. The primary antibody was then added, diluted in the blocking solution (PBST, 3% serum of secondary antibody host), for 2 days, rocking at 4˚C. Embryos were then washed 6 times for 5 min in PBST before adding the secondary antibody, diluted in blocking solution, for 5 hours, rocking at room temperature. Embryos were then washed 6 times for 5 min in PBST, with the last 3 washes containing DAPI, before being imaged by confocal microscopy using a Leica TCS SP8.

2.3.7 X-gal staining of E6.5 embryos

Embryos were washed with PBS and fix with X-gal fix (0.2% gluteraldehyde, 0.1 M phosphate buffer, pH 7.3, 2 mM $MgCl_2$, 5 mM EGTA, 0.1 M disodium hydrogen orthophosphate, 0.1 M sodium dihydrogen orthophosphate) at 4°C for 20 min. Embryos were washed 3 times for 15 min in X-gal wash (0.1 M phosphate buffer, pH 7.3, 2 mM MgCl2, 0.1% sodium deoxycholate, 0.02% Igepal (Sigma)) and stained overnight in the dark, at room temperature in X-gal staining solution (X-gal wash containing, 5 mM potassium ferricyanide (Sigma), 5 mM potassium ferrocyanide (Sigma), 15

μM NaCl, 1 mg/mL X-gal (Promega)). Embryos were washed 3 times for 15 min in X-gal wash and fixed in 4% PFA for 15 min.

2.3.8 Cryosectioning of embryos

E6.5 X-gal stained embryos were fixed in 4% PFA overnight at 4°C. They were washed in PBS and transferred to 15% sucrose solution in PBS for 2 hours at 4°C. Embryos were then incubated in 15% sucrose and 7% gelatin in PBS at 37°C until the embryos sink. They were then transferred to an aluminium mould in the gelatin, sucrose solution to be orientated and allowed to set on ice. The mould was then snap frozen in liquid nitrogen and stored at -80°C until sectioning where it was allowed to equilibrate at -24°C.

2.3.9 Wax sectioning of embryos

E9.5 embryos were washed in PBS and fixed overnight at 4°C in 4% PFA. Embryos were then transferred to 70% IMS for 1 hour at 4°C before being taken through a series of 80% IMS, 90% IMS, 95% IMS then 100% ethanol for 30 min at 4°C in each solution. Embryos were then incubated in Xylene (Sigma) until embryos were cleared as judged by microscope observation. The embryos were then washed 3 times for 1 hour in paraffin wax before being put into the final orientation where the wax was allowed to set. Wax blocks were trimmed and mounted on plastic cassettes. 7 μM sections were cut and placed onto glass slides. Wax was removed from sections by a dilution series from Xylene (10 min), 100% ethanol, 95% IMS, 90% IMS, 70% IMS, 50% IMS, 30% IMS and water for 5 min each. DAPI was added to slides before they were mounted with Vectashield.

2.3.10 Chimaera generation

Chimaeras were generated by morula aggregation with clusters of approximately 8 ES cells. After aggregation, morulae were cultured overnight with ES cells before being transferred to pseudopregnant female mice. Only successfully compacted morulae and blastocysts were transferred. Alternatively single cells were injected into morula or 2-cell embryos. For single cell injections, embryos were first decompacted for 20 min at room temperature in PB1 medium without calcium and magnesium and transferred the same day. In multiple cell morula injections (Appendix), 5 ES cells were introduced. Aggregations and injections were performed by Lynsey Robertson, Sally Inverarity (TCF, CRM, University of Edinburgh), Javier Martin Gonzalez and Kasper Bonderup (TCF, Danstem Centre, University of Copenhagen).

2.4 RNA-seq

RNA was prepared from ES cell samples sorted by FACS as in 2.2.4. ES cells were previously cultured in various culture conditions prior to sorting, including serum/LIF, 2i and 2i/LIF. ES cells were sorted into tubes containing a small amount of the medium in which they were cultured. Samples were sequenced by the Aboobaker lab at the University of Nottingham. There, approximately 10 µg of total RNA underwent 2 rounds of mRNA enrichment with Dynalbeads Oligo$(dT)_{25}$ (Life technologies). SOLiD whole transcriptome libraries were made according to the SOLiD Total RNA-seq protocol with the exception of RNA fragmentation (Life technologies). RNA was fragmented by chemical hydrolysis; heating to 95°C, 10 min in 1x RNAse III buffer (Life technologies) and snap cooled on ice. ATP (0.83 mM, Roche) and 10 U of T4 polynucleotide kinase (PNK) (NEB) were added and incubated at 37°C, 30 min. RNA was purified using Purelink RNA Micro Kit (Life technologies). Equimolar pools of RNA-seq libraries were made following qPCR quantification using Kapa Library Quantification kit (Kapa Biosystems). ePCR and templated bead enrichment was carried out with SOLiD EZ bead system according to manufacturer's guidelines. Enriched beads were sequenced on an ABi SOLiD 4 analyser according to the manufacturer's instructions to generate 50 bp reads in colour space.

RNA-seq reads were mapped to the mouse RefSeq genes (NCBI) using the Bowtie program (http://bowtie-bio.sourceforge.net). A maximum of 2 mismatches were allowed. Reads that matched to the negative strand or to >10 locations were not used in the analysis. If multiple RefSeq sequences corresponded to the same gene (Entrez), we selected the sequence with the maximum number of aligned RNA-seq reads. The total number of RNA-seq reads (HV^+ plus HV^-) was used for normalization. Statistical analysis was done using NIA Array Analysis (PMID: 15734774). The difference was considered statistically significant based on FDR <0.05, and change > 1.5 fold to account for multiple comparison tests.

2.5 GRO-seq analysis

2.5.1 GRO-seq sample collection and nuclei isolation

HV ES cells were sorted by FACS into HV low (lowest 25% of HV expression) and HV high (highest 25% of HV expression) cells. Cells both positive and negative for the expression of the ES cell marker SSEA-1 were collected in order to analyse differentiated and undifferentiated samples. SSEA-1 negative cells were collected from samples differentiated by LIF withdrawal, as described in Section 2.1.6. After FACS to isolate cell nuclei, cells were re-suspended in 5 mL LYSO-A buffer (35 mM Tris-Cl, pH 7.5, 150 mM sucrose, 80 mM KCl, 5 mM $MgCl_2$, 0.5 mM DTT, all reagents from Sigma) and centrifuged for 5 min at 1000 g. They were then re-suspended in 5 mL LYSO-A buffer containing 0.05% Tween-20. Cells were incubated on ice for 1-5 min, with constant checking to ensure that 90% of cells are permeabilised using a haemocytometer and tryphan blue (Life Technologies). Cells were washed in 5 mL LYSO-A buffer and transferred to siliconized tubes (Eppendorf) in 1 mL LYSO-A. Cells were centrifuged at 900 g for 5 min and re-suspended in 500 μL F80 buffer (50 mM Tris, pH8.3, 40% glycerol, 80 mM KCl, 5 mM $MgCl_2$, 0.1 mM EDTA, 0.5 mM DTT). Cells were counted and then re-suspended at 5 x 10^6 cells/100 μL before being snap-frozen in liquid nitrogen in aliquots of 100 μL. All steps were carried out at 4°C and using wide-bore pipette tips to minimize the loss of sample. Samples were then sent to out collaborator in the Lis lab, Leighton Core.

2.5.2 GRO-seq library preparation (Lis Lab)

Libraries were prepared as in (Core et al., 2008), with the modifications stated below.

NRO reaction:
Nuclei (100 μL) were mixed with an equal volume of reaction buffer (10 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$, 1 mM DTT, 300 mM KCL, 1% sarkosyl, 500 μM each of ATP, GTP, and Br-UTP, 2 μM CTP and 0.33 μM α-^{32}P-CTP (3000Ci/mM), all reagents from Sigma, 20 units of SUPERaseIn (Life Technologies). The reaction was allowed to proceed for 5 min at 30°C. The reaction was stopped by the addition of 2 mL (10x volume) of Trizol (Life Technologies). The phases were separated by addition of 400 μL of chloroform as per manufacturers instruction. An additional acid-phenol and then chloroform extraction were carried out, followed by precipitation with 2.5 volumes of ethanol. The pellet was washed in 75% ethanol before re-suspending in 20 μL of DEPC-treated water. Base hydrolysis was performed on ice by addition of 5 μL 1 M NaOH (Sigma) and incubated on ice for 30 min. The reaction was neutralized by addition of 25 μL 1 M Tris-HCl (pH 6.8). The reaction was then run through a p-30 RNAse-free spin column (BioRad), according to the manufacturer's instructions. The column flow through was brought to 100 μL with DEPC water and EDTA was added to a final concentration of 1 mM.

Bead pre-wash**:**

All buffers used in bead enrichment steps were kept on ice and were supplemented with 4 U/mL of SUPERaseIN. Anti-deoxyBrU beads (Santa Cruz) were first washed three times with a pre-wash buffer: 0.25x SSPE, 500 mM NaCl, 1 mM EDTA, 0.05% Tween-20 (all reagents from Sigma) for 5 min; washed twice in binding buffer: 0.25x SSPE, 37.5 mM NaCl, 1 mM EDTA, 0.05% Tween-20 for 5 min; blocked in bead blocking buffer: 0.25x SSPE, 1 mM EDTA, 0.05% Tween-20, 0.1% PVP, and 1 mg/mL ultrapure BSA (Ambion) for 1 hour; followed by one wash in binding buffer for 5 min. The ratio of beads to volume did not exceed 1:8 for any wash or blocking step. The beads were re-suspended in a 25% slurry (original concentration).

Bead enrichment:

NRO RNA was heat denatured at 70°C for 3 min and placed on ice for 2 min. 350 µL of binding buffer and 50 µL of bead slurry were added to the RNA, and the samples were incubated for 30 min on a rotating stand (8 rpm). The beads were washed once in binding buffer; once in low salt wash buffer: 0.2x SSPE, 1 mM EDTA, 0.05% Tween-20; once in high salt wash buffer: 0.25% SSPE, 137.5 mM NaCl, 1 mM EDTA, 0.05% Tween-20; and twice in TET: 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, .05% Tween-20. The NRO RNA was eluted three times (2x 125 µL, 1x 250 µL) with elution buffer: 20 mM DTT, 150 mM NaCl, 5 mM Tris-HCl (pH 7.5), 1 mM EDTA, and 0.1% SDS. The NRO RNA was then isolated by a standard extraction-precipitation method: One acid-phenol extraction, one chloroform extraction, addition of NaCl to 300 mM and 1 µL of glycoblue (Ambion) to the aqueous phase, precipitation with 2.5 volumes of cold ethanol, and a wash of the resulting pellet with 75% ethanol. The pellet was re-suspended in DEPC water at volumes appropriate for the subsequent step.

End Repair:

Pelleted RNA from the first bead binding was re-suspended in 20 µL, and heated to 70^{o}C for 5 min, followed by incubation on ice for 2 min. 3 µL tobacco acid pyrophosphatase (TAP) buffer, 4.5 µL water, 1 µL SUPERaseIn, and 1.5 µL TAP (Epicentre) were then added and the reaction incubated at 37^{o}C for 1.5 hours. 1 µL 300 mM $MgCl_2$ and 1 µL T4 PNK were added to the reaction for an additional 30 min. for phosphorylating the 5'-ends, 20 µL T4 PNK buffer, 2 µL 100 mM ATP, 145 µL water, 1 µL SUPERaseIn, and an additional 2 µL of PNK were added for 30 min at 37^{o}C. The reaction was then stopped by addition of 20 mM EDTA followed by acid phenol extraction and precipitation.

Adapter ligations:

For adapter ligations the RNA was re-suspended in 8.5 µL, and incubated with 2.5 µL of either the 5'- or 3'- adapter oligo (RA5 or RA3, see below), 1 µL SUPERase In, 2 µL RNA ligase-1 buffer, 5 µL 50% PEG 8000, and 1.5 µL of T4 RNA ligase-1 (NEB). The reactions were incubated on the lab

bench for 4 hours. After both the first and second adapter ligations the RNAs were enriched over anti-deoxy-BrU beads as described above.

Reverse transcription and amplification and PAGE purification of NRO-RNA libraries:
The RNAs were reverse transcribed in 25 µL reactions, with 1 µL 25 µM RPI-# (see below), and 1 µL SIII reverse transcriptase (Life Technologies), at 44°C for 15 min, followed by 52°C for 45 min. The RNAs were degraded by addition of RNAse cocktail (Ambion), and RNAse H (Ambion), and amplified 15 cycles, with Phusion high fidelity DNA polymerase (Finnzymes) using the RNA PCR Primer, RP1 (see below). The NRO-cDNA libraries were then run on a non-denaturing 1x TBE, 8% acrylamide gel, and amplicons with inserts greater than 20 nucleotides were excised from the gel and eluted by incubating in TE containing 300 mM NaCl, overnight while rotating. The library was then extracted, precipitated, and submitted for sequencing.

2.5.3 Data acquisition and mapping to the mouse genome

GRO-seq libraries were sequenced on the Illumina HiSeq2000, using standard protocol at the Cornell bioresources center (http//www.BRC.cornell.edu). Bowtie (Langmead et al., 2009) was used to map 30mers with up to two mismatches to the mm9 genome. Reads were also mapped to a representative of repetitive genes transcribed specifically by Pol I (rRNA gene; GenBank accession #: BK000964.1). The rDNA included the extragenic spacers. Biological replicates had a Pearson product-moment correlation coefficient (r value) of 0.98 or higher. To examine the efficiency of the FACS sorting procedure reads were also mapped to the endogenous *Hex* cDNA and the exogenous Venus cDNA and results were as follows, $HV^{+}SSEA\text{-}1^{+}$: 14 reads, $HV^{+}SSEA\text{-}1^{+}$: 97 reads, $HV^{+}SSEA\text{-}1^{-}$: 60 reads.

2.5.4 Identification and clustering of up and downregulated genes

Up and downregulated genes were identified using the edgeR statistical package (Robinson et al., 2010) with the FDR cutoff set to 0.01. Signal in the gene body regions (+500 to gene ends) from replicate experiments were used as input and were normalized by the number of mapped uniquely mapped reads. The number of uniquely mapped reads were as follows, $HV^{+}SSEA\text{-}1^{+}$: 1.07×10^{8} reads, $HV^{+}SSEA\text{-}1^{+}$: 1.11×10^{8} reads, $HV^{+}SSEA\text{-}1^{-}$: 2.04×10^{7} reads. To cluster the regulated genes, separate matrices were constructed for up and down regulated genes that composed of the correlation coefficients for each gene across the time points. 1 - r^2 values for each gene were then used to generate groups used agglomerative with hierarchical clustering. Clustering was performed using r statistical package clustering tools, as described in (Danko and Pertsov, 2009), with method set to 'Ward'.

2.5.5 Oligos

RNA 5’ Adapter (RA5), part # 15013205
GUUCAGAGUUCUACAGUCCGACGAUC

RNA 3’ Adapter (RA3), part # 15013207
P-TGGAATTCTCGGGTGCCAAGG-inv-dT

RNA PCR Primer (RP1), part # 15005505
AATGATACGGCGACCACCGAGATCTACACGTTCAGAGTTCTACAGTCCGA

RNA PCR Primer, Index 1 (RPI1)
CAAGCAGAAGACGGCATACGAG|NNNNNNNN|GTGACTGGAGTTCCTTGGCACCCGAGAA
TTCCA

(#’ represents one of the 48 Illumina 6-base barcodes in the middle of the oligo, shown above as “NNNNNN”),

CHAPTER 3:

2i CULTURED EMBRYONIC STEM CELLS AND EMBRYOS CONTAIN A TOTIPOTENT CELL POPULATION

Chapter 3:

2i cultured embryonic stem cells and embryos contain a totipotent cell population

The aim of my PhD was to gain a more comprehensive understanding of the regulation of heterogeneity in ES cells and pre-implantation mouse embryos. To this aim, I have explored the role of different signalling pathways in controlling the balance and functional potential of discrete ES cell subpopulations and on early embryonic development.

ES cells can be maintained in the presence of LIF and either BMP4 or serum (Ying et al., 2003a) (Section 1.3.4). Under these conditions, numerous genes are expressed in a heterogeneous manner (Canham et al., 2010; Chambers et al., 2007; Hayashi et al., 2008; Kobayashi et al., 2009; Singh et al., 2007; Toyooka et al., 2008) (Section 1.3.5), suggesting that ES cell cultures may harbour cells with distinct functional potentials. ES cells can also be cultured in minimal serum-free medium with an inhibitor of the FGF pathway component, MEK (mitogen-activated protein or extracellular signal-regulated kinase) and an inhibitor of GSK3 (glycogen synthase kinase 3) (2i culture conditions) (Ying et al., 2008) (Section 1.3.4.3). FGF is known to promote differentiation of ES cells as well as being involved in lineage segregation *in vivo* (Sections 1.3.4.3 and 1.2.3.3), hence inhibition of FGF signalling in 2i shields ES cells from differentiation-inducing signals and is thought to generate a homogeneous early epiblast-like state that can contribute to embryonic but not extraembryonic development (Nichols et al., 2009; Wray et al., 2010; Ying et al., 2008).

I have previously described a sensitive reporter for the endoderm marker *Hex* utilizing a reiterated IRES element to translationally amplify the expression of the fluorescent protein Venus, encoded downstream of *Hex* in the endogenous locus (Section 1.4.4) (Canham et al., 2010). Here, I have utilized ES cells containing this reporter, and a transgenic reporter mouse line derived from them, to explore the nature of 2i-cultured ES cells. I show that embryos and ES cells cultured in 2i are heterogeneous and contain a fraction of cells coexpressing markers of both embryonic and extraembryonic lineages. This population demonstrated an enhanced capacity to generate extraembryonic cell types, including trophoblast, *in vitro*, and single cells from this fraction were totipotent when assessed by morula aggregation *in vivo*.

3.1 2i culture of pre-implantation embryos prevents lineage segregation

In 2i culture conditions, ES cells express both endodermal and pluripotency markers (Marks et al., 2012). However, it was previously reported that the culture of pre-implantation embryos in 2i, from the morula until the late blastocyst stage, prevented the formation of the PE as evidenced by a lack of GATA4 expression. Resulting embryos showed a 100% NANOG positive ICM (Nichols et al., 2009). Here, I used transgenic mice, containing the same sensitive endodermal *Hex* reporter as used in ES cells (Canham et al., 2010) (Fig. 7A), to determine whether endoderm gene expression was abolished when embryos were cultured in 2i.

I first characterised the early expression of the *Hex* transgene. Although *in situ* hybridisation on early embryos previously suggested that *Hex* was not expressed until E4.5 (Thomas et al., 1998), more sensitive single cell transcriptional analysis revealed that it was expressed from the 2-cell stage (Guo et al., 2010). Embryos were flushed from the oviduct at E2.5, and cultured for 3 days (the equivalent of E4.5 *in vivo*) in KSOM medium. At pre-implantation stages, it was possible to clearly distinguish transgenic HV embryos from wild type littermates by fluorescence microscopy (Fig. 7B). HV was expressed from E2.5 in a heterogeneous manner, present in approximately 50% of blastomeres (Fig. 7C). At the early blastocyst stage (E3.5) HV was expressed heterogeneously within the ICM and also in the trophoblast while at the late blastocyst stage (E4.5) HV expression became restricted to the PE where it was coexpressed with the later endodermal gene, GATA4 (Fig. 7C,D).

When embryos were cultured in 2i, although HV expression was still observed, the PE did not segregate and *Hex* continued to be expressed heterogeneously within the ICM, reminiscent of the early blastocyst (Fig. 7E). However, as previously reported (Nichols et al., 2009), these *Hex*-expressing cells never progressed to the point where they expressed the later endoderm marker GATA4 (Fig. 7D). Thus, although 2i culture did not block PE gene expression, it did prevent PE segregation and further differentiation. Taken together these observations suggest that 2i-cultured embryos are blocked at a stage of development prior to lineage segregation.

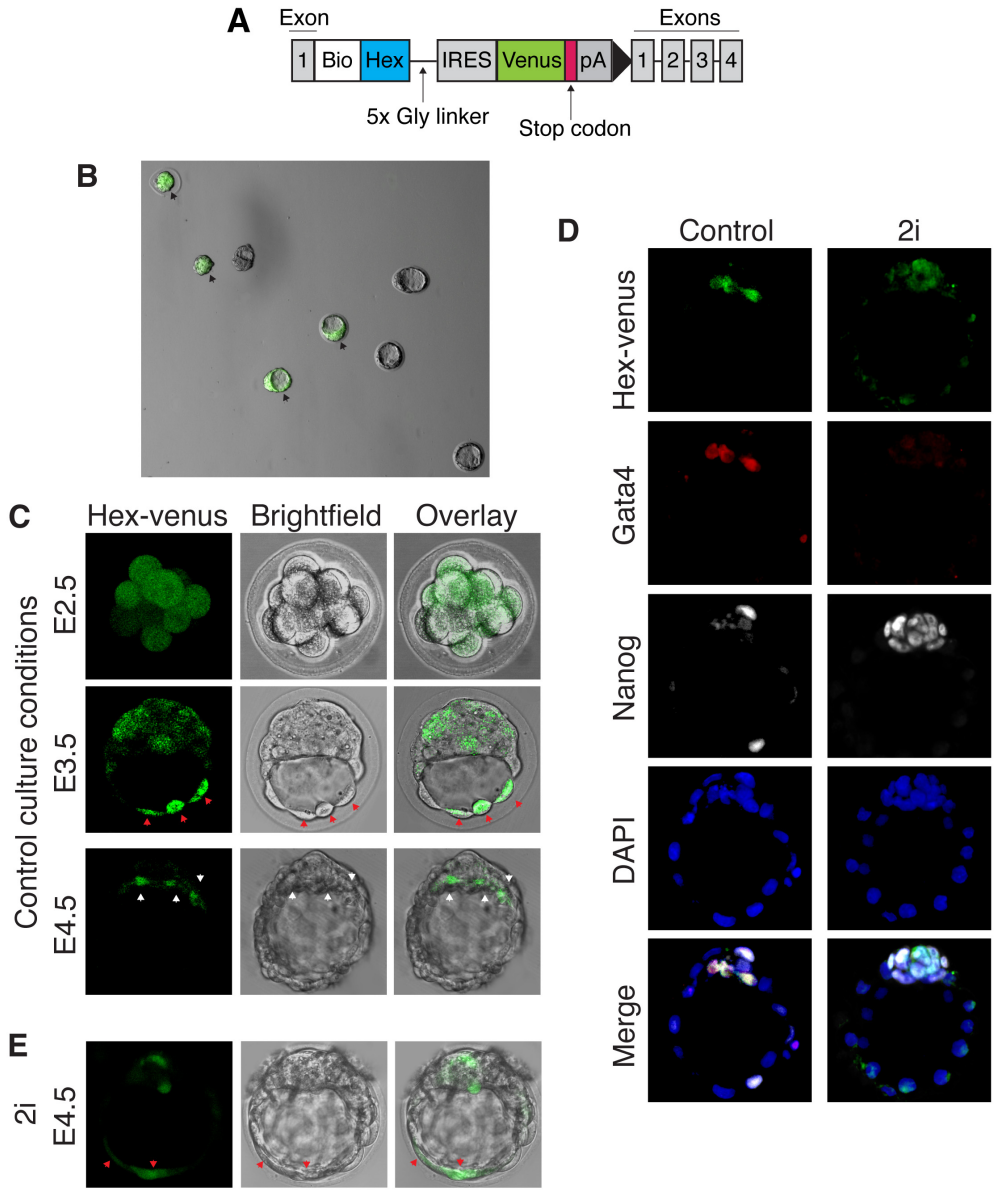


Figure 7. Culture of transgenic Hex-venus (HV) embryos in 2i. A. Schematic diagram of the HV reporter construct inside the endogenous Hex locus. The construct consists of Biotin-tagged Hex cDNA follwed by an internal ribosome entry site (IRES) that facilitates the translational amplification of the downstream Venus fluorescent protein. The construct is inserted into the Hex locus, disrupting exon 1. As the construct contains Hex cDNA, endogenous expression levels are maintained. **B.** Fluorescence microscopy of wild type and transgenic HV (black arrowheads) pre-implantation embryos. **C.** Endogenous expression pattern of HV throughout pre-implantation development. Red arrowheads mark HV expression in the trophoblast while white arrowheads mark HV expression in the primitive endoderm. Embryos were imaged live up until E3.5 due to loss of the weak fluorescence upon fixation. **D.** E4.5 control-cultured or 2i-cultured HV embryos stained for the late primitive endoderm marker GATA4 and the embryonic marker NANOG. **E.** E4.5 HV transgenic embryo following culture in 2i from E2.5 until E4.5. Blastocyst stage images represent confocal optical sections through the ICM while morula images are extended images of the entire embryo.

3.2 Heterogeneity of embryonic stem cells in 2i culture conditions

I then asked whether ES cells cultured in 2i similarly continued to express HV. HV ES cells were cultured in serum/LIF, 2i or 2i/LIF and analysed by flow cytometry. Consistent with our observations in blastocysts, I noted that the endoderm gene *Hex* was still expressed in all conditions (Fig. 8A-C). In ES cells cultured in 2i, HV expression was more homogeneous than in ES cells cultured under standard serum/LIF conditions (Fig. 8A-C). Notably, the addition of LIF to 2i promoted an increase in HV expression (Fig. 8A-C, Chapter 4). ES cells were monitored over a period of time to determine how rapidly 2i reduced the distribution of HV expression. After 3 days, *Hex* expression became more homogeneous in 2i than in serum/LIF (Fig. 9A). However, the difference was most apparent after 3 passages (Fig. 9A). When single HV^- or HV^+ cells were sorted by FACS, re-plated and expanded, in 2i they retained their ability to interconvert and regenerate mixed populations (Fig. 9B). However, in all culture conditions, HV^- cells had a tendency to regenerate mixed populations with a higher proportion of the seeding cell type and vice versa for HV^+ cells.

In standard serum/LIF conditions, HV and NANOG were expressed in a mutually exclusive manner (Canham et al., 2010) (Fig. 10A), reminiscent of the 'salt and pepper' expression pattern in the early blastocyst (Chazaud et al., 2006) (Section 1.2.3.3). However, in 2i and 2i/LIF conditions, as NANOG was expressed in the majority of cells, HV and NANOG were now predominantly coexpressed (Fig. 10A,B). This supports the observation that embryos cultured in 2i retained the expression of both HV and NANOG (Fig. 7D,E). Based on these stainings, the HV reporter appeared to mark different ES cell populations in 2i compared to in serum culture. To determine the molecular signature of these cells types, HV ES cells were cultured in serum/LIF or 2i/LIF for 3 passages (a time period ascertained to affect the expression pattern of *Hex*, Fig. 9A)and sorted into HV^- and HV^+ populations (represented as the lower and upper 25% of *Hex* expression) and next generation sequencing of RNA (RNA-seq) was carried out using ABi SOLiD technology obtaining 4-9 x 10^7 raw reads (Gene Expression Omnibus accession number GSE45182). Analysis was carried out using the National Institute on Aging (NIA) online array analysis tool in collaboration with Alexei Sharov and Minoru Ko (NIA, Baltimore, USA). Although, by flow cytometry, *Hex* expression was more homogeneous in 2i/LIF-cultured cells than serum/LIF-cultured cells (Fig. 8A-C), HV^- and HV^+ populations from 2i/LIF had more distinct gene expression profiles than cells cultured in serum (Fig. 11A). Although the endoderm marker *Gata6* was expressed at lower levels in 2i, it was expressed more highly in the HV^+ than the HV^- population (Fig. 11B). Moreover, the PE marker *Dab2* and the epiblast marker *Rex1* showed more profound differences in the HV^- and HV^+ populations in 2i/LIF than in serum/LIF, when analysed by qRT-PCR (Fig. 11C).

Specific gene classes were then analysed in more detail in the RNA-seq dataset. ES cell-associated genes did not show significant expression differences between HV^- and HV^+ populations, although a significant number were downregulated in 2i/LIF compared to serum/LIF, including *Myc*, *Utf1*, *Lin28*,

Id1 and *Id2* (Fig. 12A). I observed that more than 20 imprinted genes were enriched in the 2i/LIF HV^+ compared to the 2i/LIF HV^- population; for example the *Dlk3-Dio3* cluster (Fig. 12B) associated with efficient reprogramming (Liu et al., 2010). I also observed increased levels of trophoblast gene expression in the 2i/LIF HV^+ population, including markers specifically expressed in TS cells (Rugg-Gunn et al., 2012) (Fig. 12C). In addition, genes that are regulated by endogenous retroviral elements and enriched in an ES cell population comparable to the 2-cell stage embryo (Macfarlan et al., 2012) were upregulated in the 2i/LIF HV^+ population (Fig. 12D).

In summary, ES cells cultured in 2i are heterogeneous and maintain the expression of *Hex* although now coexpressed with NANOG. This coexpressing population is molecularly distinct from cells that mutually exclusively express HV and NANOG in serum-containing medium.

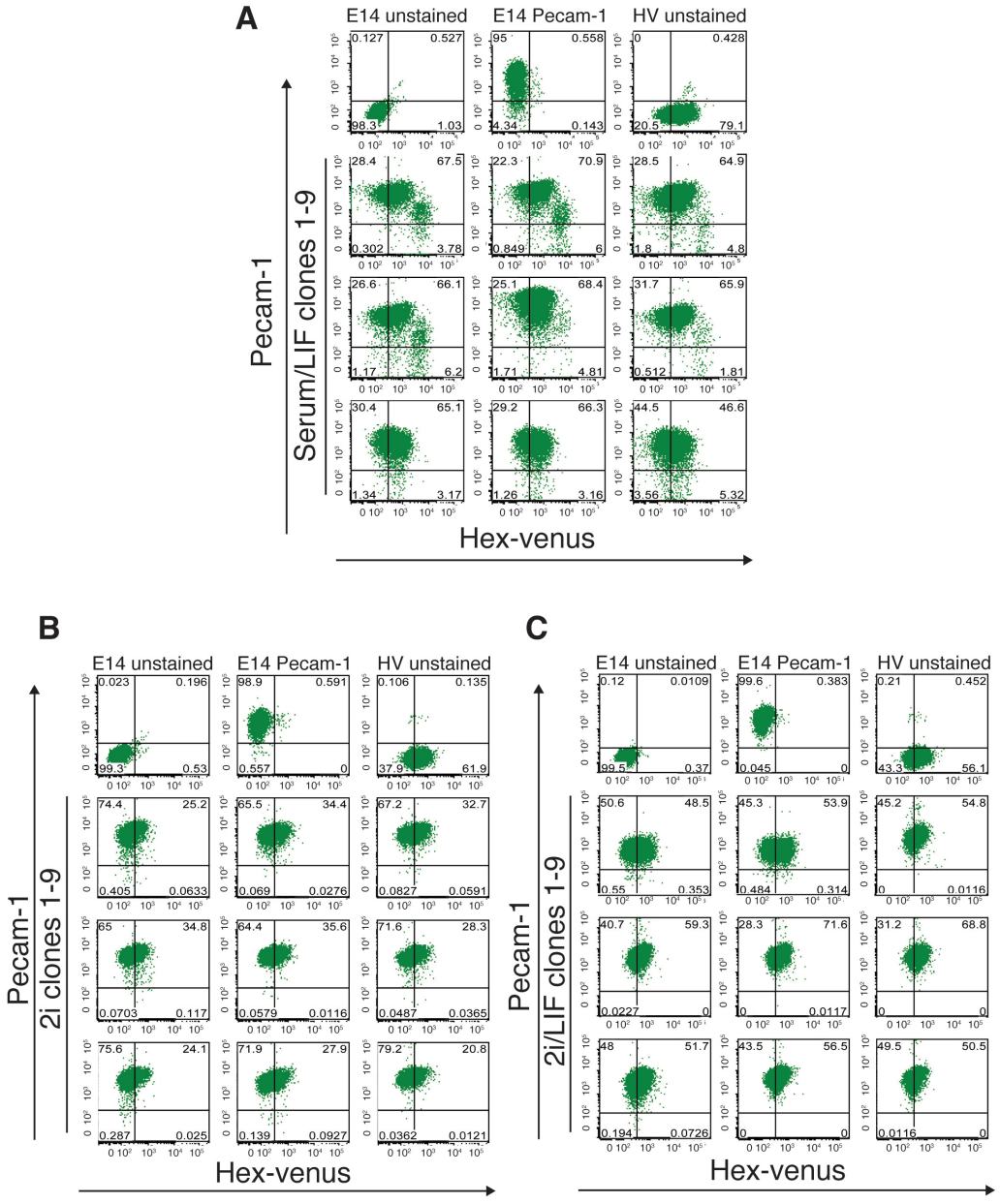


Figure 8. Hex-venus (HV) expression in ES cells. Flow cytometry analysis showing examples of unsorted clones expanded from single cells cultured in various conditions, in **A.** serum/LIF, **B.** 2i, **C.** 2i/LIF. The number in the corners of each plot represents the percentage of cells within that particular gate. Gates were set based on unstained E14 control cells (see Fig. 6C). PECAM-1 marks undifferentiated ES cells.

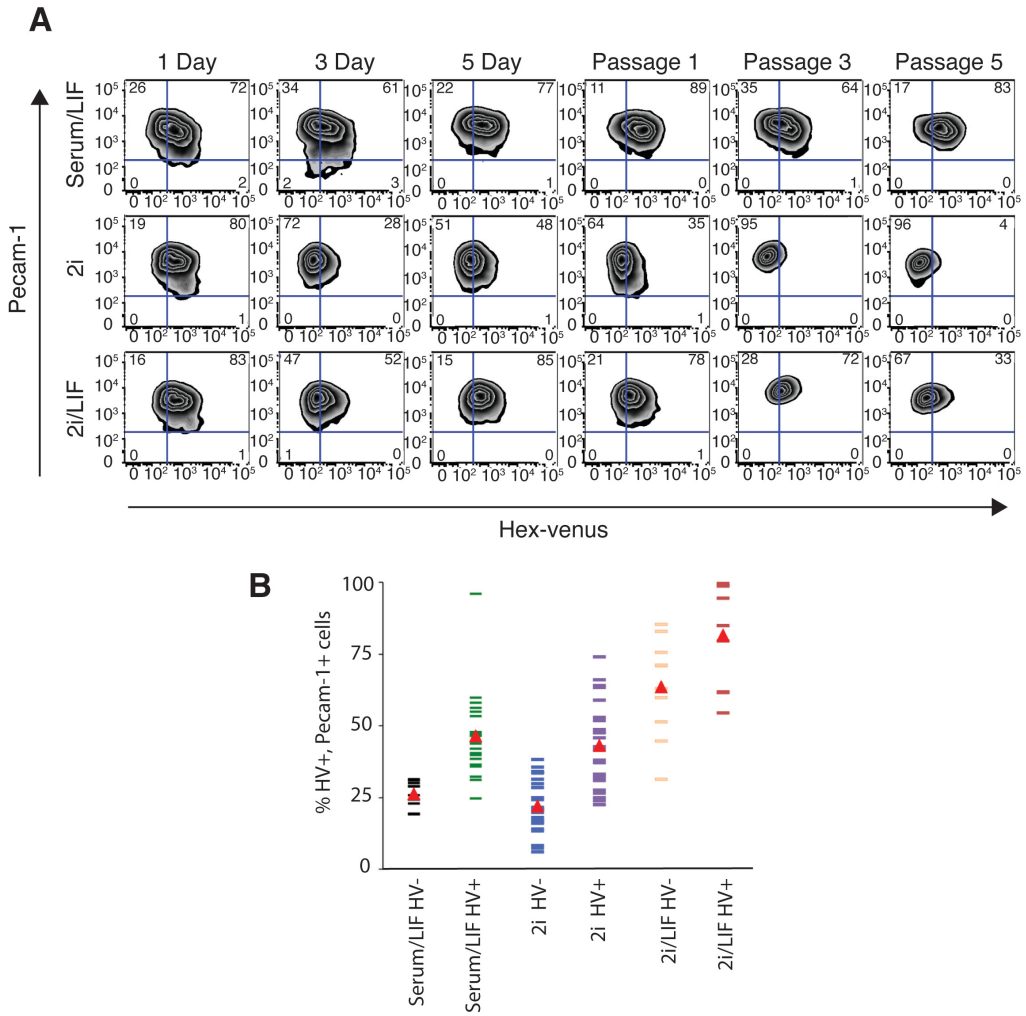


Figure 9. Dynamics of Hex-venus (HV) expression in 2i medium. A. Flow cytometry data showing HV ES cells during several passages, either cultured continuously in serum/LIF or else switched at day 0 to 2i or 2i/LIF medium. The number in the corners of each plot represent the percentage of cells within that particular gate. Gates were set according to unstained E14 control cells (see Fig. 6C). **B.** Summary of flow cytometry of single sorted HV^+ or HV^- cells from different culture conditions. Cells were sorted and plated back into the original condition to expand and were analysed 7-10 days later. Each line represents an individual clone. Red triangles indicate mean. PECAM-1 marks undifferentiated ES cells.

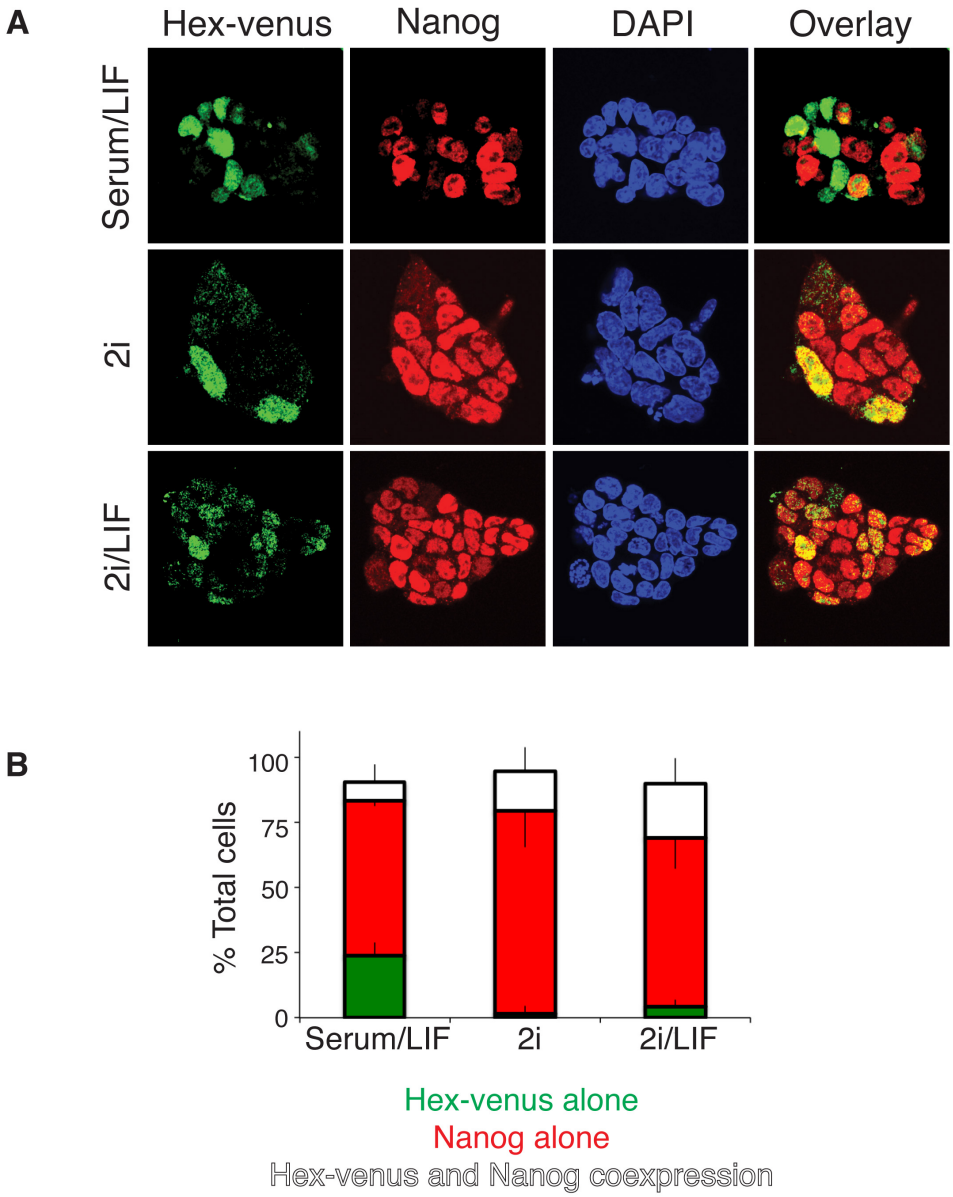


Figure 10. NANOG and Hex-venus (HV) expression in 2i medium. A. Confocal optical sections through undifferentiated HV ES cell colonies cultured in either serum/LIF, 2i or 2i/LIF immunostained for NANOG. An anti-GFP antibody was utilised to detect HV expression. **B.** Quantification of NANOG and HV coexpression in each culture condition. Error bars indicate standard deviation of the mean of 5 ES cell colonies, where the whole z-stack was analysed.

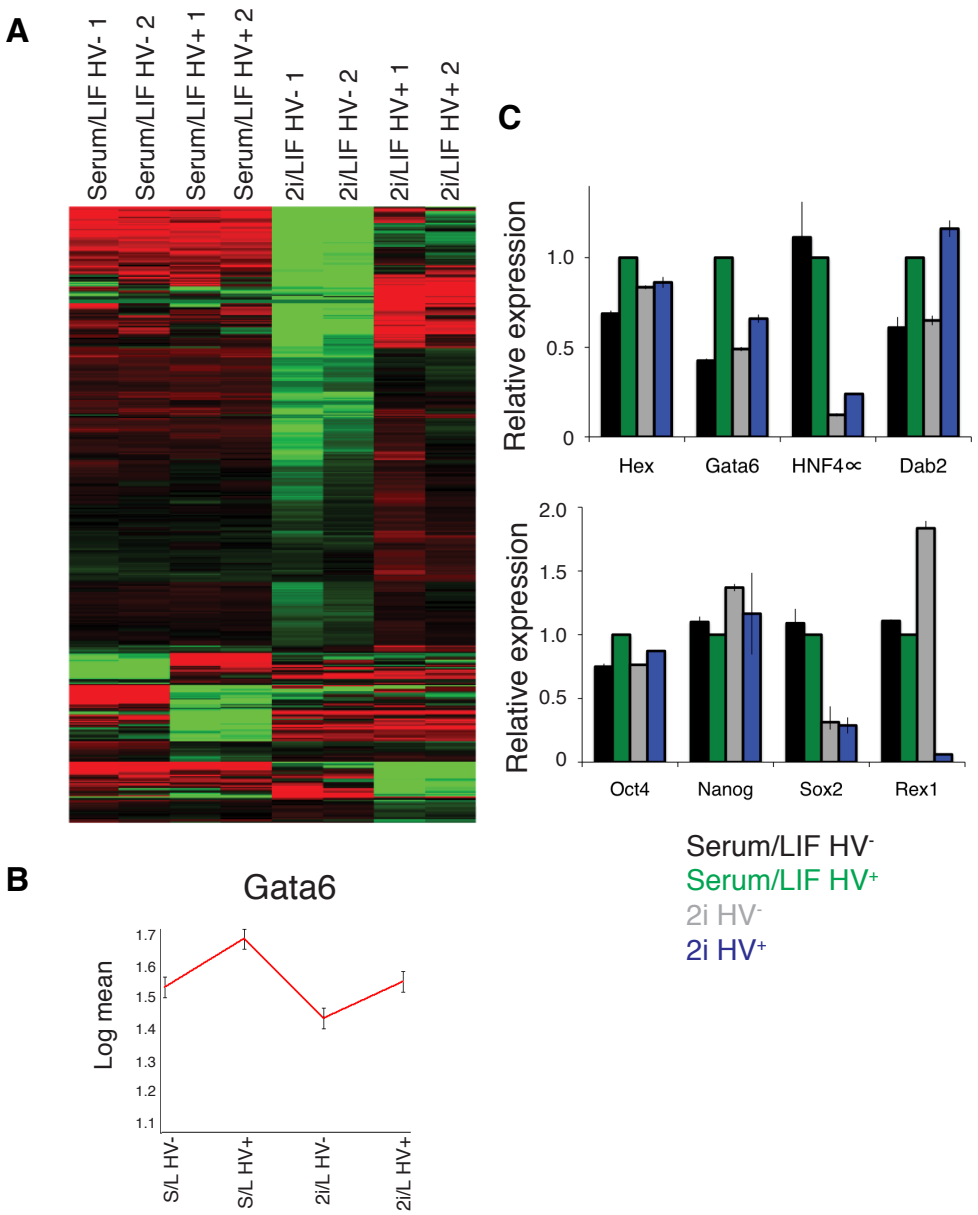


Figure 11. RNA-seq of Hex-venus (HV) ES cells cultured in 2i/LIF compared to serum/LIF medium. A. Heat map of sorted HV$^-$ and HV$^+$ populations from serum/LIF and 2i/LIF culture, based on gene expression data from RNA-seq. Heat map shows differentially expressed genes identified by pair-wise comparison of all sorted fractions. Data was normalised by subtracting the average log expression from all samples. Genes are hierarchically clustered by average Euclidean distance. 2 biological replicates are shown per sample. Red = upregulated genes, Green = downregulated genes. **B.** RNA-seq expression profile of the critical endodermal gene Gata6 in each sample, generated using the NIA array analysis tool. **C.** qRT-PCR of endodermal and ES cell-associated genes in sorted HV$^-$ and HV$^+$ populations from cells cultured in either serum/LIF or 2i. Expression levels are normalised to the housekeeping gene TBP and shown relative to HV$^+$, serum/LIF. Error bars indicate standard deviation of the mean from 3 biological repeats.

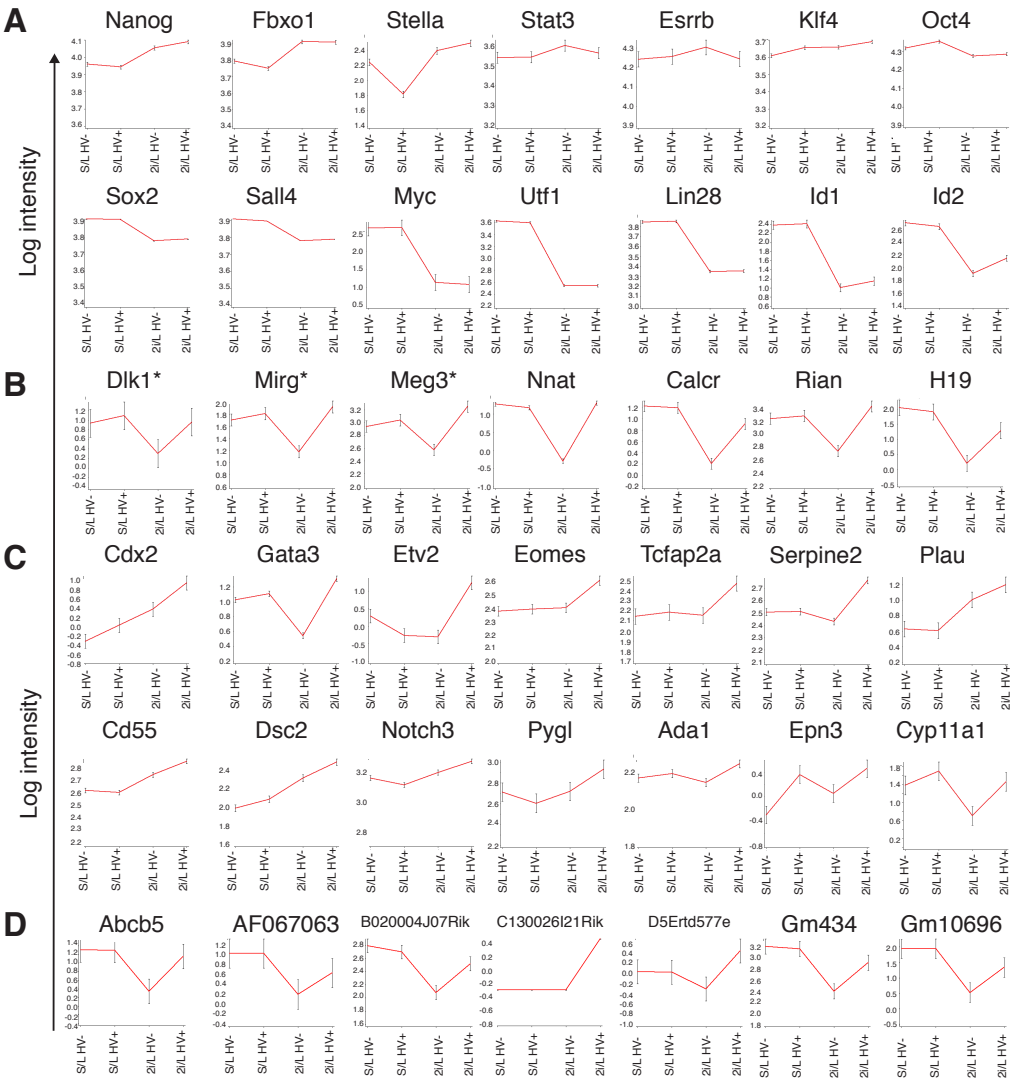


Figure 12. Common RNA-seq signatures in specific gene classes. A-D. The behaviour of the individual genes belonging to the classes described in the text is shown. Plots compare mean log intensity values for individual genes among four sorted populations; serum/LIF (S/L) HV$^-$, serum/LIF HV$^+$, 2i/LIF (2i/L) HV$^-$, 2i/LIF HV$^+$. Error bars represent standard deviation of the mean of 2 biological replicates. Classes include **A.** markers of undifferentiated ES cells, **B.** imprinted genes (* = member of Dlk1-Dio3 cluster), **C.** trophoblast markers, **D.** 2-cell stage embryo genes under the control of retroviral promoter elements.

3.3 *In vitro* assessment of lineage-priming in 2i-cultured embryonic stem cells

As there were clear gene expression differences between the HV^- and HV^+ populations present in 2i, I wanted to determine whether there were also functional differences. It was previously observed that, when cultured in serum/LIF, HV^+ ES cells exhibited a bias towards endoderm and HV^- ES cells towards epiblast, particularly in the context of cell competition assays such as blastocyst injections (Canham et al., 2010). To ask about the extent to which the phenomenon represented lineage-priming and if it also existed in 2i, I developed assays to detect bias in ES cell differentiation.

HV ES cells were cultured in serum/LIF or 2i for at least 3 passages then sorted by FACS into HV^- and HV^+ populations. Sorted ES cells were immediately plated into differentiation-promoting conditions, either LIF withdrawal (where endoderm and mesoderm lineages are predominantly generated), neural differentiation or trophoblast differentiation. ES cells were differentiated for 5 days by LIF withdrawal and the extent of endoderm differentiation was quantified by counting the number and size of GATA6 positive endodermal clusters (Fig. 13A-D), qRT-PCR (Fig. 15A), and flow cytometry for the endodermal cell surface marker PDGFRα (Fig. 15B). HV^+ cells generated more, and larger, endodermal clusters than HV^- cells (Fig. 13A-D). Although the ratio of endoderm generated was the same between HV^- and HV^+ cells from either serum/LIF or 2i, the overall levels of differentiation were much higher when cells were previously cultured in 2i (Fig. 13C). This is in part due to the increased rate of differentiation of ES cells cultured in 2i compared to serum/LIF. The ES cell marker PECAM-1 was downregulated at a more rapid rate, upon LIF withdrawal, from ES cells cultured in 2i than those cultured in serum/LIF (Fig. 14). However, irrespective of the rate of differentiation, the expression of endoderm genes, including markers such as *Gata6*, *Gata4*, *Hex, Dab2* and PDGFRα, was enhanced in differentiated cells derived from the HV^+ compared to the HV^- population (Fig. 15A,B).

The opposite outcome was observed in neural differentiation. After 9 days of differentiation in N2B27 medium, the HV^- population showed an increased capacity to generate extended $TUJ1^+$ neurons and upregulate neural markers, most notably *Ngn1*, compared to the HV^+ population (Fig. 16A-C). Although the ratio of endoderm generated by the HV^- and HV^+ populations remained constant between cells cultured in serum/LIF and those cultured in 2i (Fig. 13C), this was not the case in neural differentiation. HV^- cells previously cultured in 2i generated more neurons than those previously cultured in serum (Fig. 16B). However HV^+ cells cultured under either condition could not efficiently generate neurons (Fig. 16B).

Finally, as higher levels of trophoblast markers were expressed in HV^+ than HV^- 2i-cultured ES cells (Fig. 12C), I asked whether these cells were capable of generating trophoblast *in vitro*. As there is no defined trophoblast differentiation protocol for murine ES cells, I started by sorting HV^- and HV^+ cells cultured in serum/LIF, 2i or 2i/LIF and plating them into TS cell culture conditions (FGF4, Heparin

sulphate, and MEF-conditioned medium). After 5 days, colonies emerged from HV^+ cells that had been cultured in 2i with a flattened TS cell-like morphology. These subsequently differentiated into post-mitotic TGCs or migratory mesenchymal cells upon passaging. I therefore decided to analyse the trophoblast potential of each population prior to passaging. HV^- and HV^+ populations from serum/LIF, 2i or 2i/LIF were plated into TS cell conditions for 7 days and then fixed and immunostained for CDX2, GATA6 and BRACHYURY (Fig. 17A,B). As CDX2 is expressed in the trophoblast, mesoderm and endoderm, BRACHYURY in the mesoderm and GATA6 in the endoderm, expression of CDX2 alone is likely to represent a trophoblast cell type. I observed no expression of BRACHYURY in these cultures and, although GATA6 was expressed, it was never found in the same cells as CDX2 (Fig. 17A,B), suggesting that $CDX2^+$ cells most likely corresponded to trophoblast. Quantification of the immunostaining demonstrated that few $CDX2^+$ cells were generated from cells previously cultured in serum/LIF. However, in 2i or 2i/LIF there were a high number of $CDX2^+$ cells generated from the HV^+ population (Fig. 17C). Additionally, upon LIF withdrawal, only the HV^+ population from 2i showed an upregulation of trophoblast markers (Fig. 17D). In summary, the HV^+ population of cells, present in both serum and 2i cultures are lineage-primed towards an endodermal fate. However, the HV^+ population in 2i also has the capacity to generate trophoblast-like cells. Repeated attempts to derive TS cells under these conditions resulted in the expansion of ES cell-like colonies expressing OCT4, that dominated the culture upon prolonged passaging. However, as discussed in Section 1.3.8, the capacity of cells to generate a self-renewing TS cell line is distinct from the ability to generate differentiated trophoblast cells.

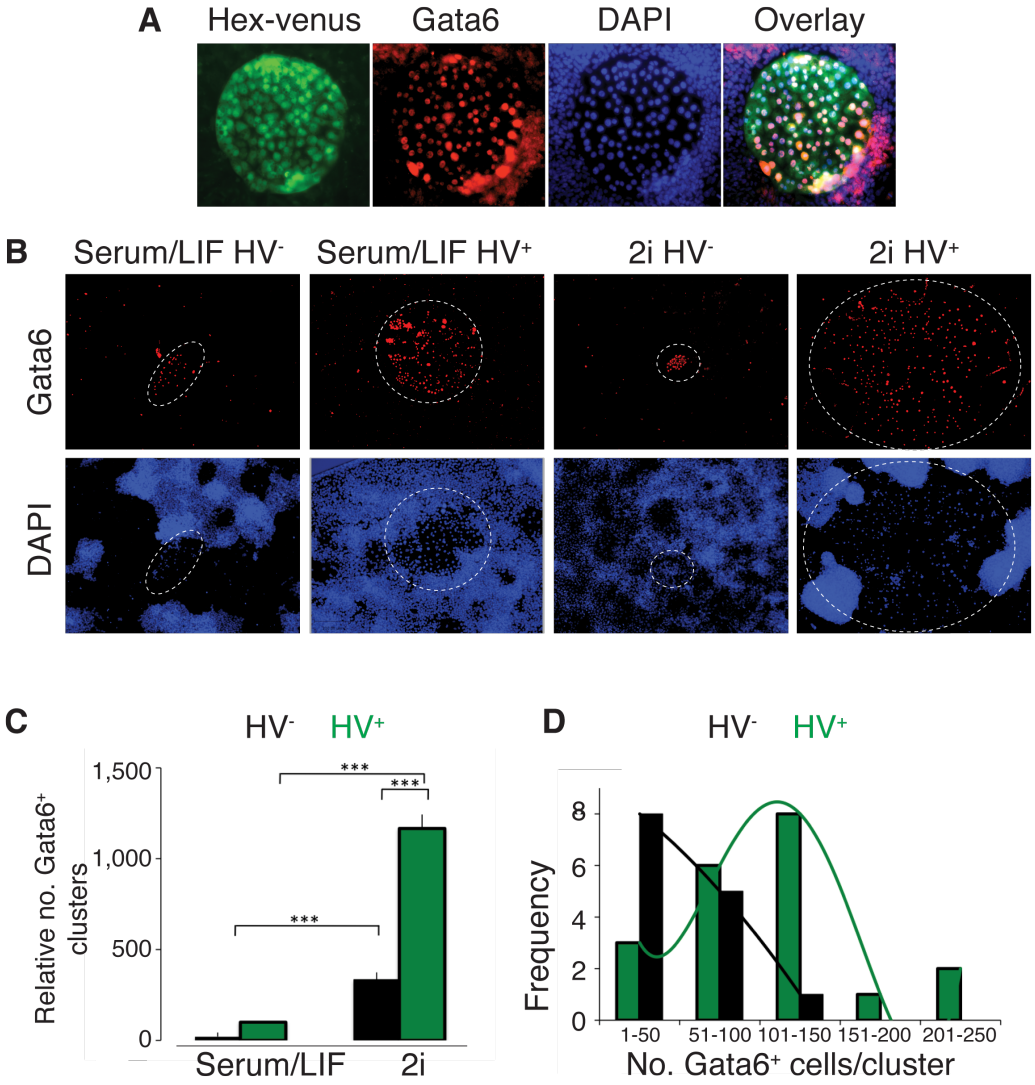


Figure 13. Lineage-priming of Hex-venus (HV) ES cells in LIF withdrawal differentiation. A. An example of a differentiated endodermal colony expressing both HV and GATA6. Colonies tended to exhibit a defined, round morphology making them easily quanitifiable. **B.** Sorted HV ES cells (sorted based on upper (HV^+) and lower (HV^-) 25% of Hex expression) which have been differentiated by LIF withdrawal and immunostained for the endoderm marker GATA6. Dotted white circles highlight endodermal colonies of cells. **C.** Quantification of endodermal colonies (shown in B) in each condition. Values shown relative to HV^- serum/LIF. Error bars indicate standard deviation of the mean of 3 biological replicates (*** = $p<0.001$). **D.** Quantification of the number of $GATA6^+$ cells per endodermal cluster of ES cells differentiated from serum/LIF.

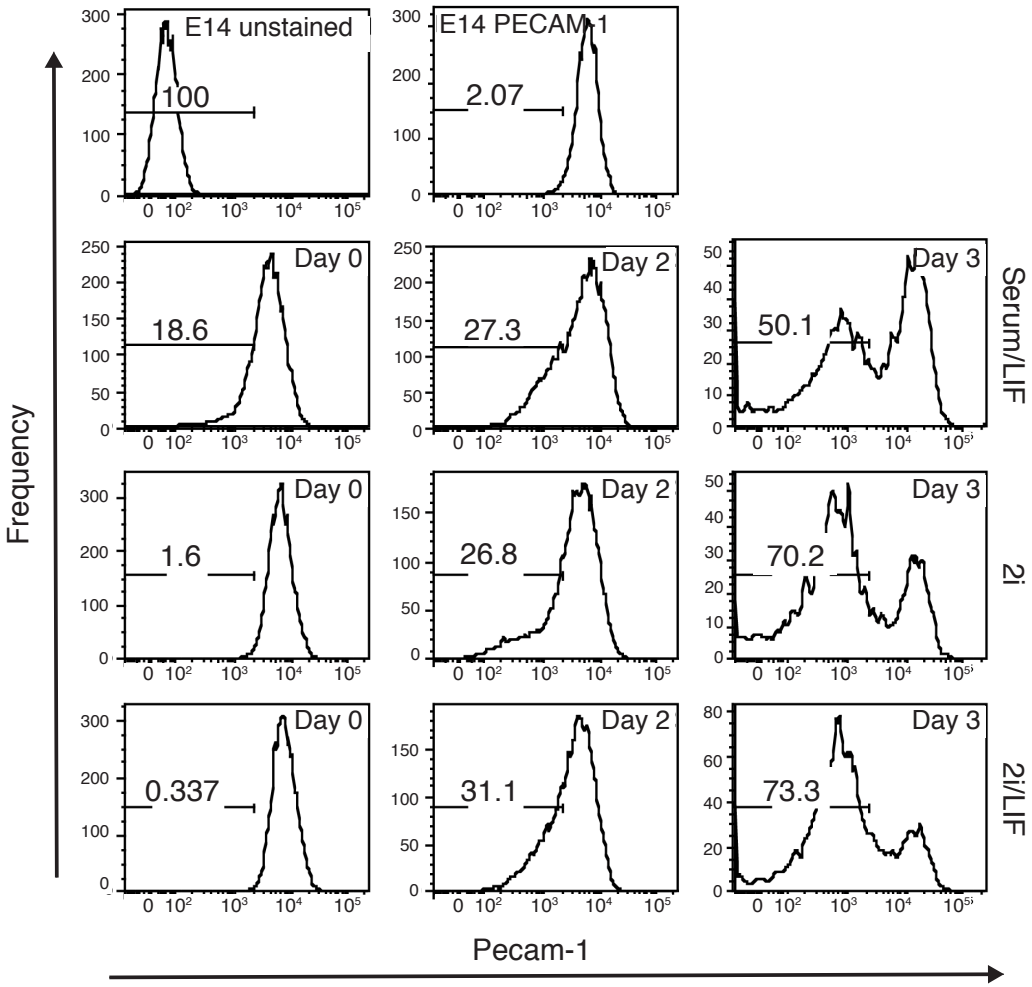


Figure 14. ES cells previously cultured in 2i rapidly downregulate PECAM-1 upon differentiation. Flow cytometry showing the rate of downregulation of the ES cell marker PECAM-1 in differentiation-promoting LIF withdrawal conditions. ES cells had been cultured in either serum and LIF, 2i or 2i/LIF prior to differentiation. Gates were set using undifferentiated E14 ES cells stained for PECAM-1. Numbers shown above the gate indicate the percentage of PECAM-1 negative cells.

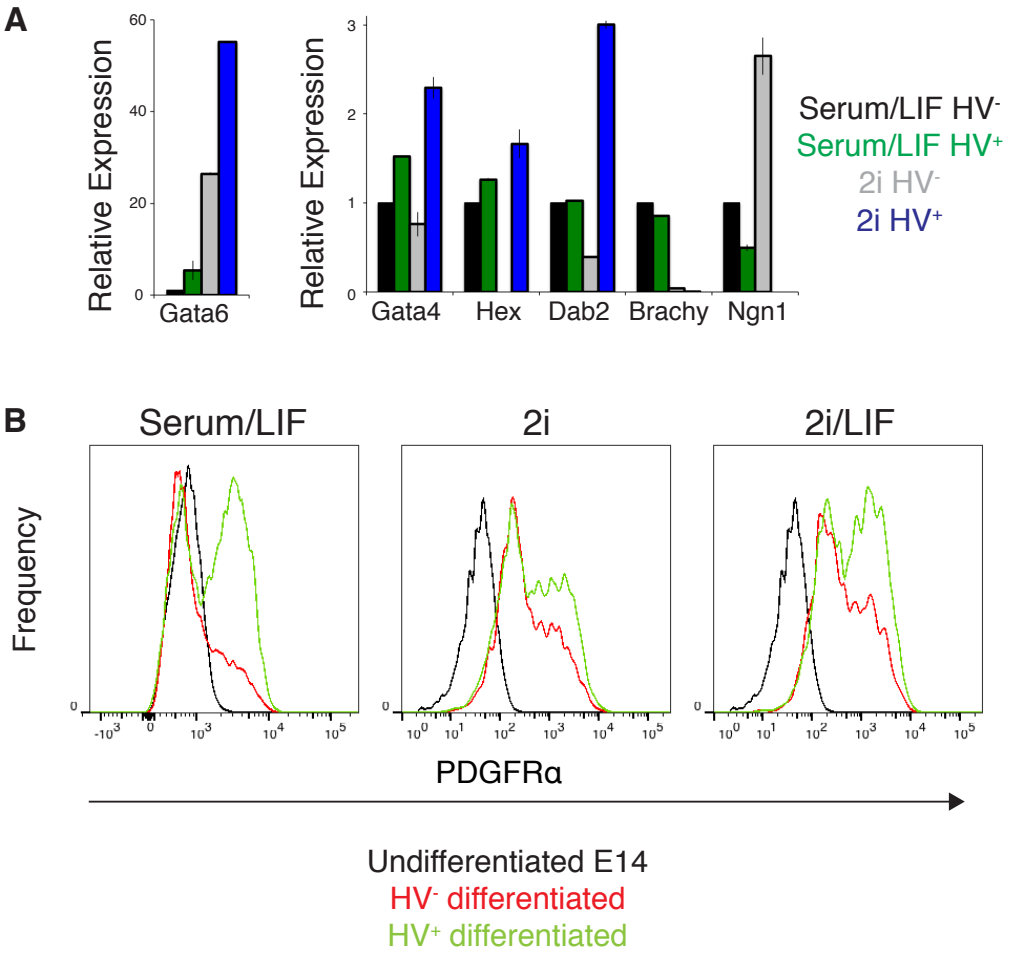


Figure 15. Quantification of lineage-priming of Hex-venus (HV) ES cells upon differentiation by LIF withdrawal. A. qRT-PCR of sorted ES cells differentiated by LIF withdrawal. ES cells were previously cultured in serum/LIF or 2i and sorted into HV$^-$ and HV$^+$ populations by FACS. Gene expression is normalised to the expression of the housekeeping gene TBP. Values are shown relative to serum/LIF HV$^-$. Error bars indicate standard deviation of the mean. **B.** Flow cytometry analysis of HV$^+$ and HV$^-$ ES cells, cultured in serum/LIF, 2i or 2i/LIF and subsequently differentiated by LIF withdrawal, for the endoderm cell surface marker PDGFRα.

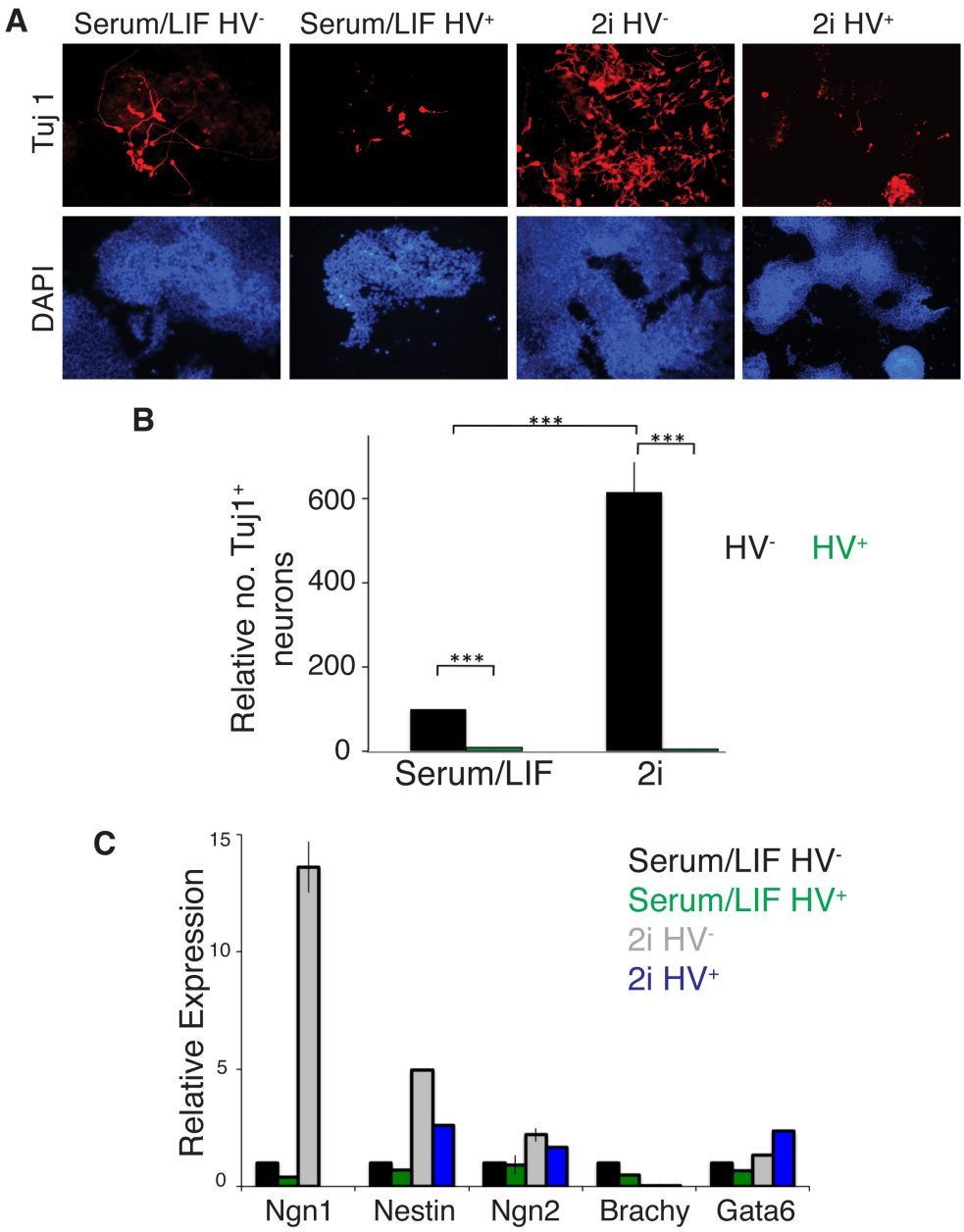


Figure 16. Quantification of lineage-priming of Hex-venus (HV) ES cells in neural differentiation. A. HV- and HV+ ES cells (sorted based on upper and lower 25% of Hex expression) were differentiated in N2B27 medium for 9 days and immunostained for the neural marker TUJ1. Prior to differentiation, ES cells were cultured in serum/LIF or 2i. **B.** Immunostaining quantification of elongated TUJ1 positive neurons in each condition. Values shown relative to serum/LIF HV+. Error bars indicate standard deviation of the mean of 3 biological replicates (*** = p<0.001). **C.** qRT-PCR of differentiated samples. Gene expression is normalised to the expression of the housekeeping gene TBP. Error bars indicate standard deviation of the mean. Values are shown relative to serum/LIF HV-.

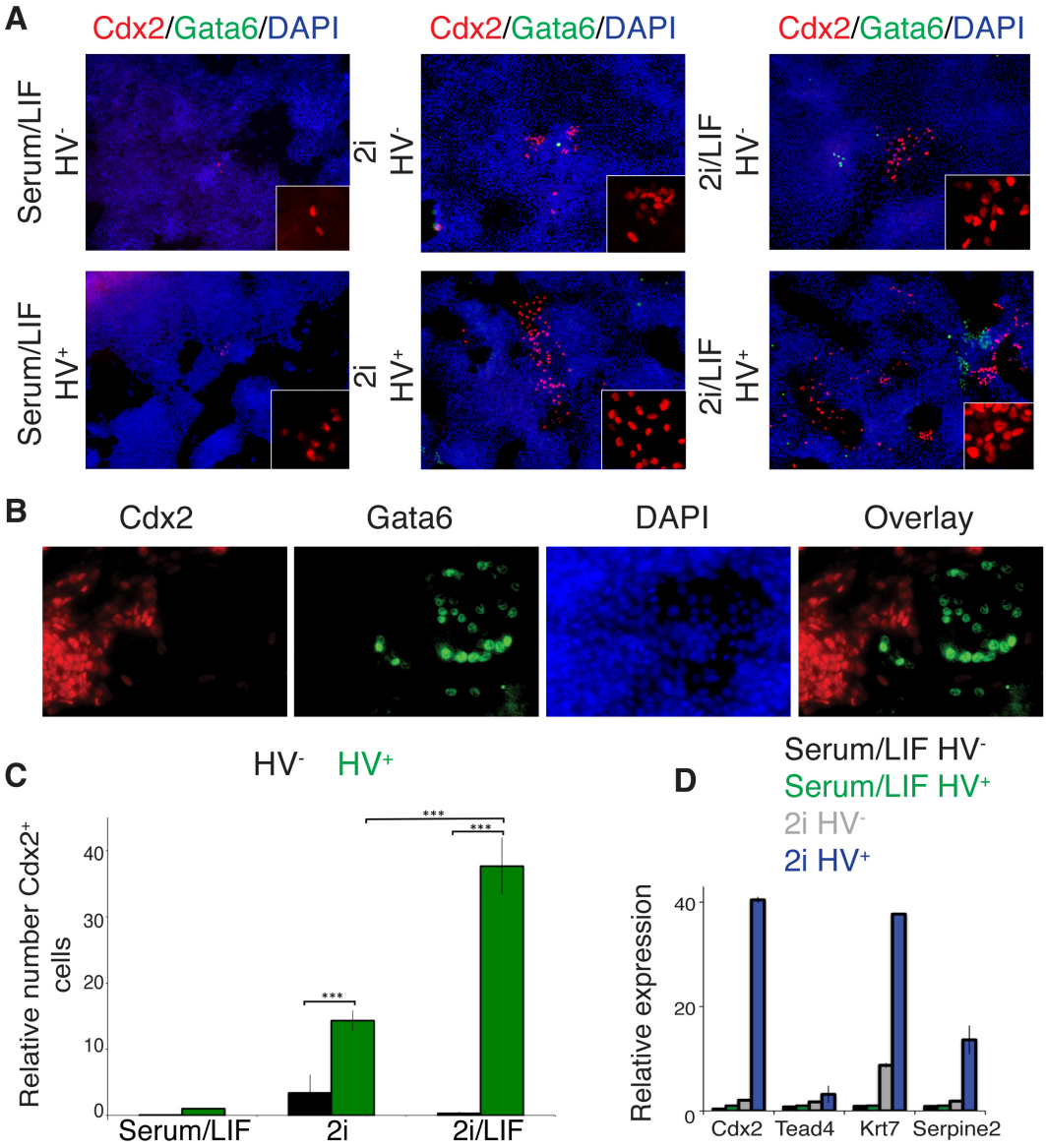


Figure 17. Quantification of lineage-priming of Hex-venus (HV) ES cells in trophoblast differentiation. **A.** Immunostaining of ES cells differentiated towards trophoblast. HV ES cells were cultured in serum/LIF, 2i or 2i/LIF before being sorted by FACS. The lower (HV$^-$) and upper (HV$^+$) 10% of Hex-expressing ES cells were sorted. After sorting cells were cultured in trophoblast stem cell conditions for 7 days before being fixed and immunostained for CDX2, GATA6 and BRACHYURY. No BRACHYURY protein was present. Insets show higher magnification images of CDX2 positive areas. **B.** High magnification of differentiated cells immunostained for GATA6 and CDX2. **C.** Quantification of the number of CDX2 positive cells generated after differentiation in trophoblast stem cell conditions. Values shown relative to serum/LIF HV$^+$. Error bars indicate standard deviation of the mean of 3 biological replicates (*** = $p<0.05$). **D.** qRT-PCR for trophoblast genes on sorted HV$^-$ and HV$^+$ cells, previously cultured in serum/LIF or 2i, after differentiation by LIF withdrawal for 4 days. Data is normalised to the housekeeping gene TBP and shown relative to serum/LIF HV$^-$. Error bars indicate standard deviation of the mean.

3.4 *In vivo* assessment of lineage-priming in 2i-cultured embryonic stem cells

It has previously been shown in chimaera assays that HV^+ ES cells cultured in serum/LIF contribute less efficiently to the epiblast than HV^- ES cells, but have the added capacity to contribute the extraembryonic endoderm (both visceral and parietal) (Canham et al., 2010). In the previous section (Section 3.3) it was shown that HV^- and HV^+ ES cells cultured in 2i also exhibit a lineage bias *in vitro*. Here I ask whether these subpopulations found in 2i show functional differences *in vivo*.

HV ES cells were cultured in 2i and then sorted by FACS into HV^- and HV^+ populations. Sorted populations were introduced into morula stage embryos that were allowed to develop to various stages to assess chimaerism. The injected cells constitutively expressed either LacZ or an H2B-Tomato fusion to identify the progeny of injected cells. When HV^+ cells were introduced into morulae, and embryos cultured for 3 days *ex vivo* until the late blastocyst stage, descendants of HV^+ ES cells could be found in the trophoblast and cells facing the blastocoel cavity (most likely PE) (Fig. 18A). Cells in the trophoblast either expressed CDX2 and NANOG or NANOG alone (Fig. 18A). However, the expression of NANOG in the trophoblast was also observed in wild type embryos, both alone and coexpressed with CDX2 (Fig. 18B). Injected embryos were also transferred to recipient mice and allowed to develop until later stages. At E6.5 HV^- cells contributed efficiently to the epiblast while HV^+ cells contributed efficiently to the epiblast, but also gave rise to progeny in the extraembryonic endoderm and trophoblast (Fig. 19A-D). In 2i/LIF, a proportion of the HV^- cells also contributed to extraembryonic lineages, although less efficiently than HV^+ cells (Fig. 19C,D), likely to be due to the fact that LIF increases the overall expression level of *Hex*, and therefore the endodermal character of the culture as a whole (Fig. 8B,C). At E9.5, when extraembryonic tissues were more differentiated and the placenta was fully formed, I observed that HV^+ H2B-Tomato cells had colonised tissues of the embryo, placenta and yolk sac (Fig. 20A,B).

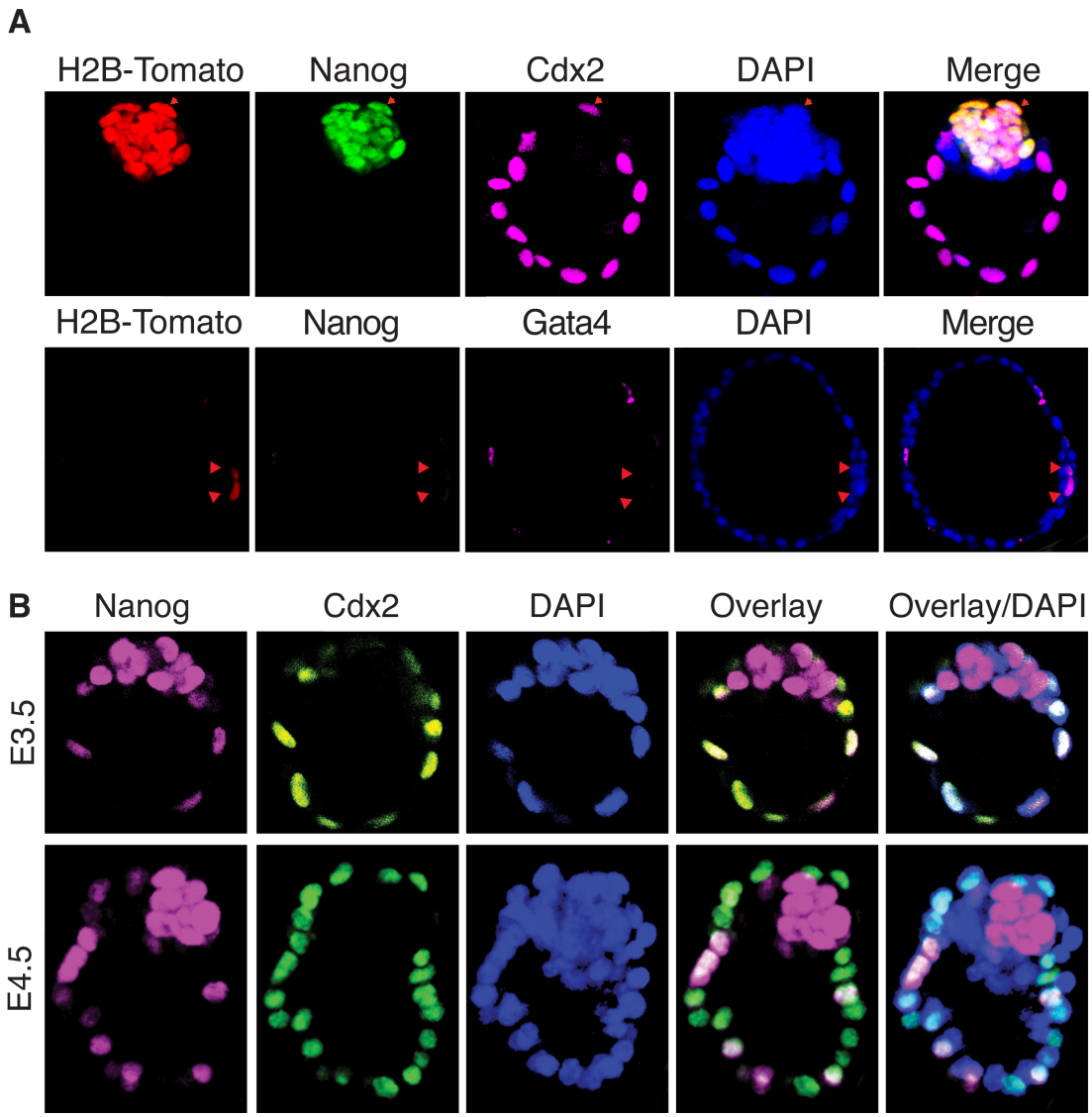


Figure 18. Contribution of Hex-venus-expressing (HV⁺) ES cells to blastocysts. A. Sorted HV⁺ ES cells expressing a constitutive H2B-Tomato fusion protein, cultured in 2i/LIF, were aggregated with wild type morulae. Embryos were cultured until E4.5 and immunostained. Red arrowheads indicate H2B-Tomato cells within the trophoblast. **B.** Immunostaining of E3.5 and E4.5 wild type blastocysts. Embryos were imaged by confocal microscopy and images represent optical sections through the ICM.

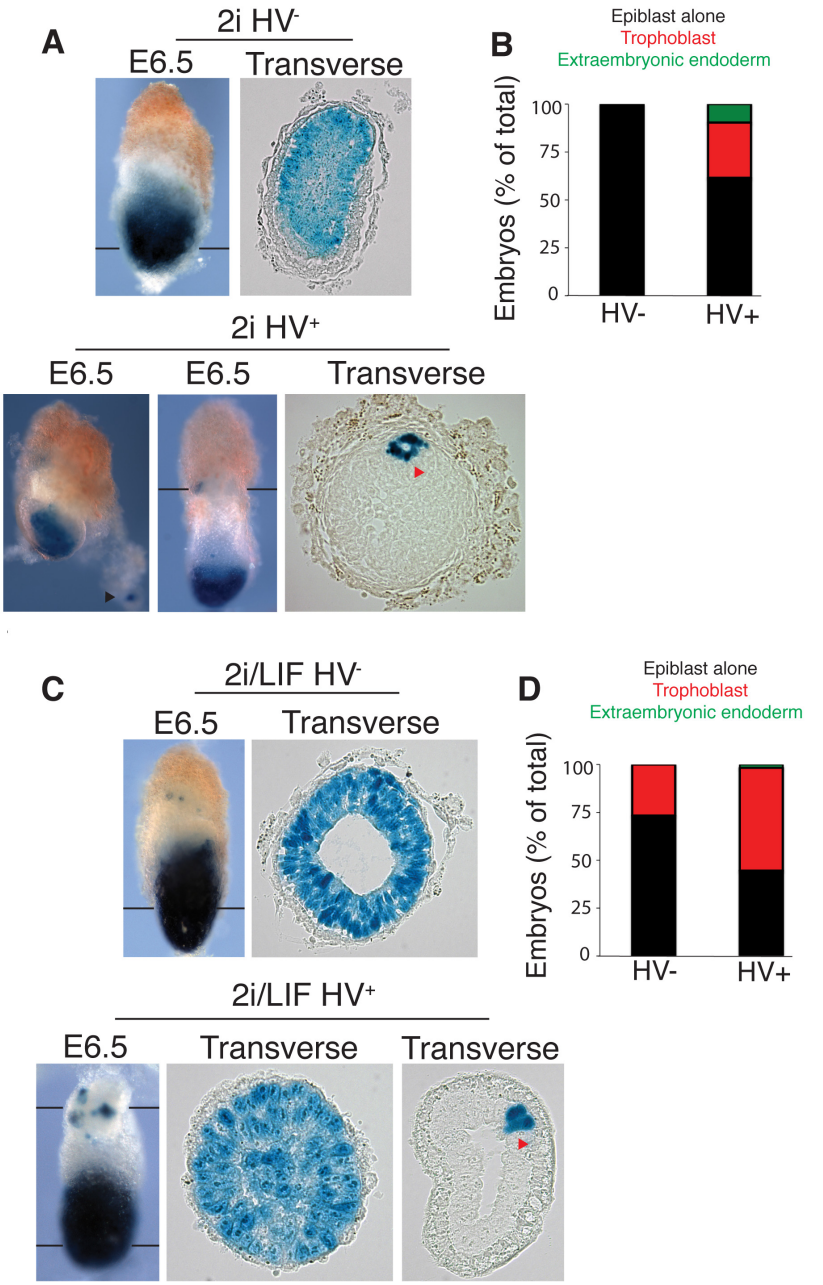


Figure 19. Contribution of sorted Hex-venus low (HV$^-$) and high (HV$^+$) expressing ES cells to gastrulation stage embryos. A. and C. Chimaeric embryos generted by morula aggregations with HV ES cells expressing a constitutive LacZ lineage tracer. Prior to aggregations ES cells were cultured in 2i or 2i and LIF and the top ($^+$) and bottom ($^-$) 25% of Hex-expressing ES cells were sorted by FACS. Embryos were collected and X-gal stained at E6.5. Black lines indicate the plane of the adjacent cryosection shown. Black arrowheads mark parietal endoderm contribution. Red arrowheads mark trophoblast. **B. and D.** Quantification of ES cell contribution to embryos based on location. 2i HV$^-$ n = 18, 2i HV$^+$ n = 26, 2i/LIF HV$^-$ n = 60, 2i/LIF HV$^+$ n = 55.

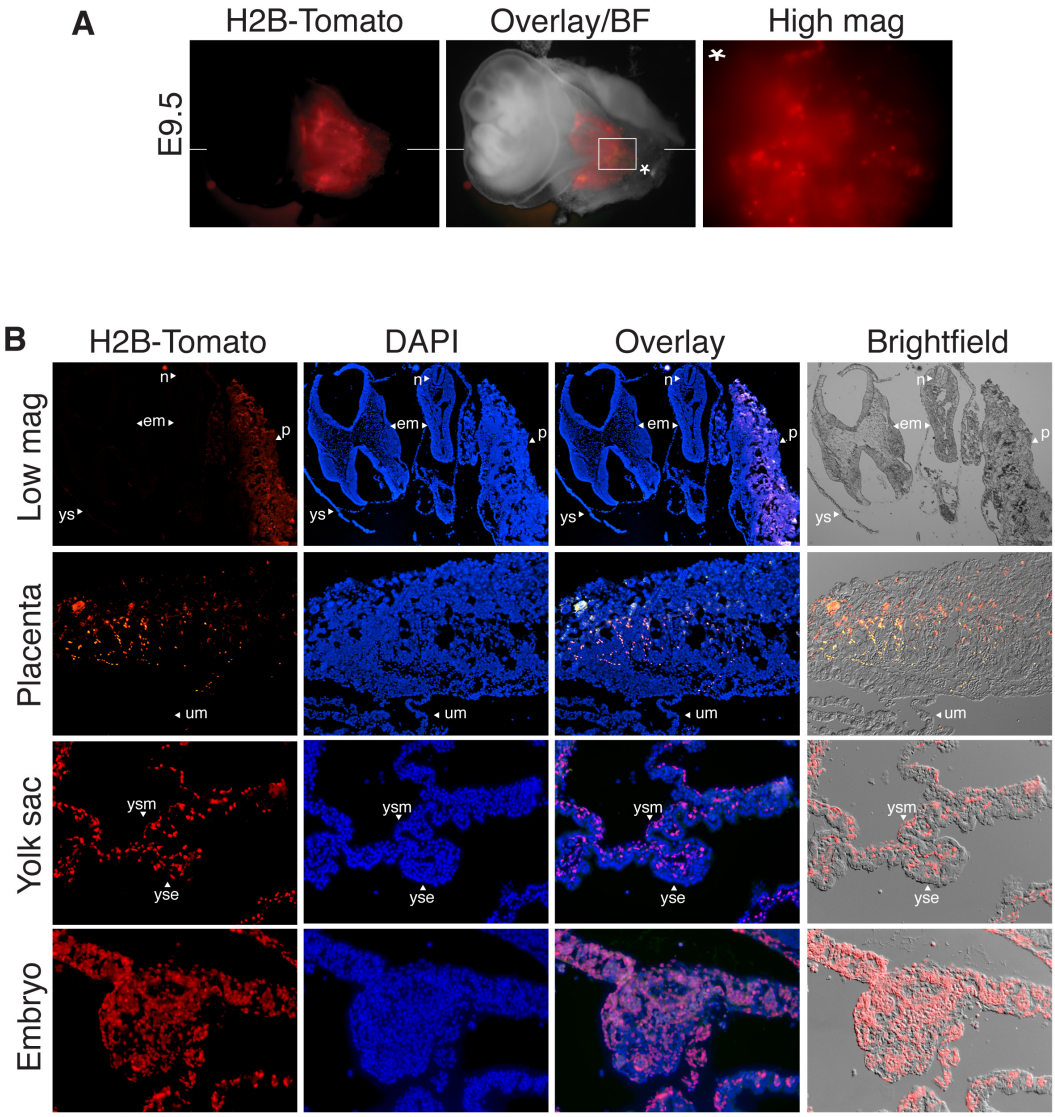


Figure 20. Contribution of Hex-venus expressing (HV$^+$) ES cells to E9.5 embryos. A. Chimaeras were generated from sorted HV$^+$ ES cells cultured in 2i/LIF and aggregated with morula stage wild type embryos. Embryos were transferred to recipient mothers and dissected at E9.5. Image shows whole mount embryo inside yolk sac and attached to the placenta. Image marked with asterisk shows high magnification of H2B-Tomato positive ES cells within the yolk sac. **B.** Wax sections of an E9.5 chimaeric embryo showing HV H2B-Tomato cells in the placenta, yolk sac and embryo. Uppermost panels show a low magnification section of the whole embryo within yolk sac attached to the placenta. ys = yolk sac, yse = yolk sac endoderm, ysm = yolk sac mesoderm, em = embryo, n = notochord, p = placenta, um = umbilical cord.

3.5 Single HV$^+$ cells in 2i coexpress embryonic and extraembryonic markers

It was observed that the HV$^+$ population of ES cells in 2i expressed extraembryonic (*Hex*) and embryonic (*Nanog*) markers and could contribute to all lineages *in vivo* (Sections 3.2-3.4). However, to determine whether HV$^+$ cells were individually totipotent or represented a mixed population of cells with different functional potentials, I analysed the expression of a number of genes within single cells. Single cell gene expression analysis was carried out using the Biomark HD platform (Fluidigm) on 96.96 microfluidics chips (Fig. 21). HV ES cells were cultured in serum/LIF or 2i then single HV$^-$ or HV$^+$ cells (from the lower and upper 10% of HV expression) were sorted by FACS directly into lysis buffer. Target-specific amplification of cDNA was carried out before the qRT-PCR reactions. Only genes with primers that had clearly defined melt curves were used for further analysis (Fig. 22). Genes known to be heterogeneously expressed in ES cells, such as *Nanog* and *Stella* (Chambers et al., 2003; Hayashi et al., 2008; Mitsui et al., 2003), varied in expression levels between single cells, while genes thought to be more homogeneous, including *Oct4, Sox2* and housekeeping genes such as *Actin*, showed little variation (Fig. 21). In the HV$^+$ population in serum/LIF, 10% of cells showed no *Nanog* expression while maintaining high levels of other pluripotency markers such as *Oct4* and *Sox2*. However, these cells also showed lower levels of housekeeping genes hence were excluded from further analysis (data not shown). In all populations, extraembryonic genes *Gata3*, *Serpine2* and *Tcfap2a* were expressed at low levels and in a heterogeneous manner (Fig. 21). However, the HV$^+$ population in 2i showed an upregulation of these genes, and at the same time maintained high expression of embryonic genes (Fig. 21). Interestingly, the expression of *Stella* became more heterogeneous in 2i HV$^-$ cells than in ES cells in serum. Additionally, in serum, the ES cell-associated gene *Klf4* was expressed more homogeneously and at higher levels in the HV$^+$ than the HV$^-$ population (Fig. 21).

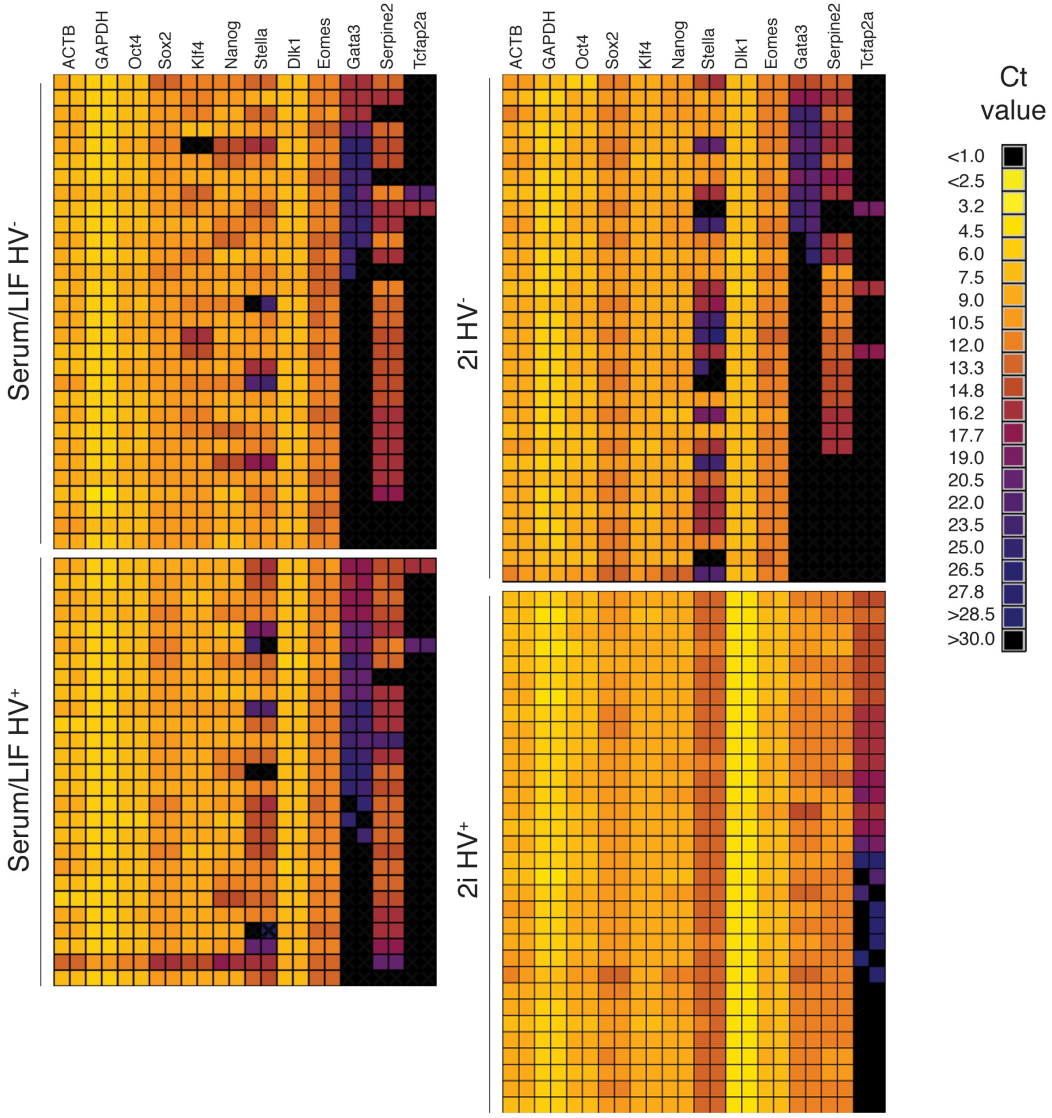


Figure 21. Single cell qRT-PCR of Hex-venus (HV) sorted populations. Heat maps of single cell qRT-PCR performed on 96.96 microfluidics chips using the Biomark HD system (Fluidigm). HV ES cells were cultured in serum/LIF or 2i before being sorted by FACS into HV^- and HV^+ populations. ES cells were sorted directly into 96 well plates containing lysis buffer. cDNA was generated and went through a round of target-specific amplification before the qRT-PCR reaction. ES cells containing low levels of housekeeping genes were discarded from the analysis. Only genes with primers giving a clearly defined melt-curve were analysed. Ct value key shown, low Ct value corresponds to high expression level. In this case, no samples showed a Ct value <1.0.

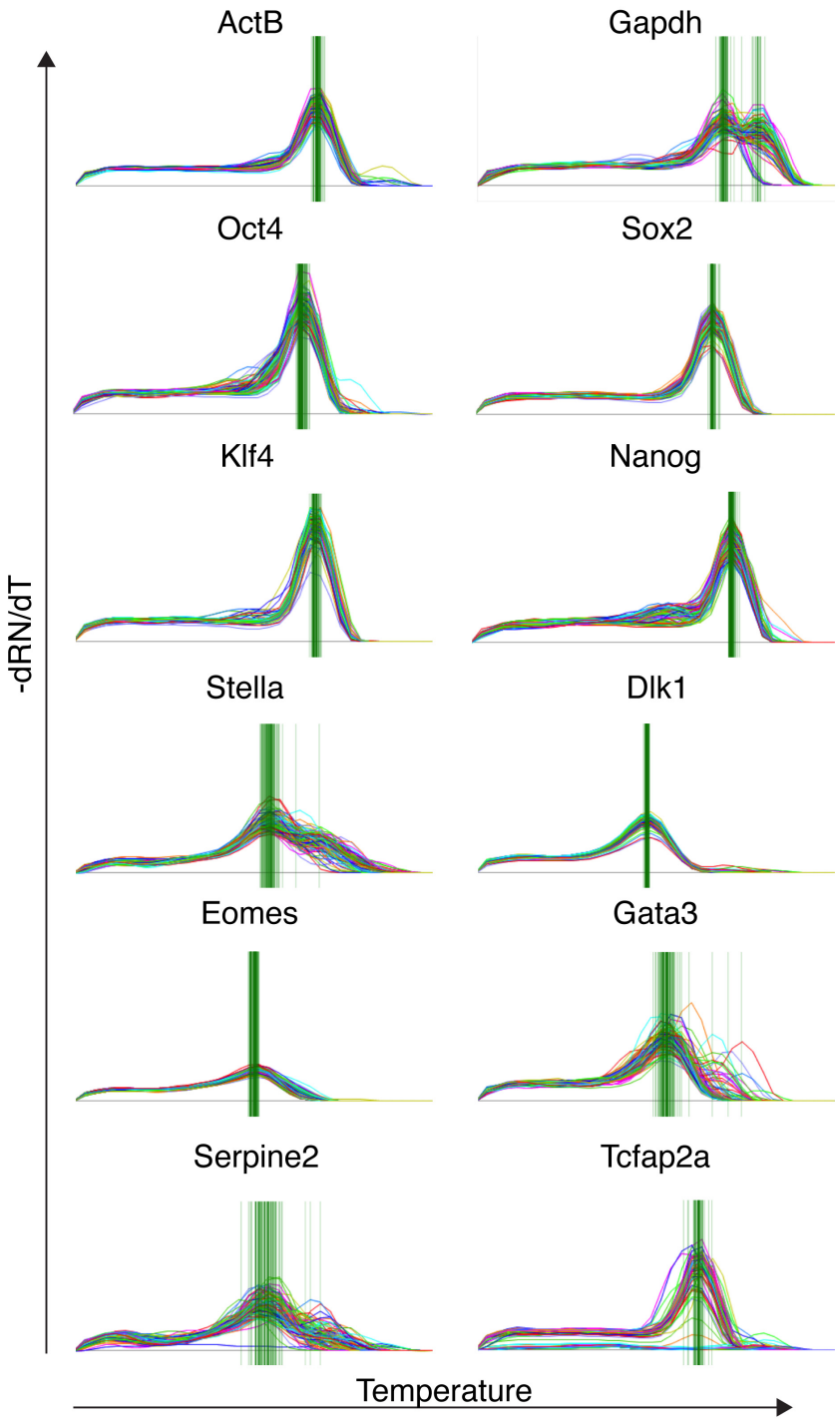


Figure 22. Single cell qRT-PCR primer melt curve controls. Primer melt curves from DELTAGene assays used in single cell qRT-PCR on the Biomark HD system (Fluidigm). Each line represent a single cell. Green vertical lines mark the peak of each single cell melt curve.

3.6 Single HV$^+$ cells cultured in 2i/LIF can contribute both to embryonic and extraembryonic tissues

As it was observed that single ES cells cultured in 2i coexpressed embryonic and extraembryonic markers, I wanted to ask whether a single cell could also contribute to both embryonic and extraembryonic tissues. Single HV$^-$ or HV$^+$ ES cells, previously cultured in serum/LIF, 2i or 2i/LIF, were sorted directly into individual wells of a 96-well plate in differentiation-promoting conditions in the absence of LIF. Clonally differentiated cells were immunostained for the lineage markers GATA6, CDX2 and BRACHYURY to analyse multi-lineage differentiation. Although these genes are expressed in various cell types, coexpression of these markers within the same cell is likely to represent a mesendoderm cell type, while the presence of all three in separate cells would suggest the presence of PE, trophoblast, and epiblast/mesoderm respectively. Cells that had previously been cultured in serum/LIF formed differentiated colonies that either exclusively expressed one or two markers or coexpressed multiple markers within the same cell (Fig. 23A,B). Conversely, cells cultured in 2i or 2i/LIF showed a decrease in single-marker expression and coexpression in favour of mutually exclusive expression of all three genes within the same colony (Fig. 23A,B). Thus, individual 2i-cultured and, in particular, HV$^+$ cells have the capacity to generate all three lineages: epiblast, trophoblast, and PE. Additionally, cells cultured in 2i before differentiation showed an increase in the expression of CDX2 alone, indicative of trophoblast differentiation, compared to cells previously cultured in serum. However, due to the overlap of expression of these markers in certain cell types it is difficult to assign totipotency based solely on marker expression in the absence of an embryonic context, hence this experiment gives only a crude estimate of ES cell potency.

To investigate the differentiation potential of single ES cells in more detail, single sorted HV$^+$ ES cells cultured in 2i/LIF were injected into wild type morula stage embryos. ES cells constitutively expressed an H2B-Tomato lineage tracer. Embryos were dissected at E6.5 showing that 56% (13/23) of single cells contributed both to embryonic and extraembryonic tissues, hence may be considered to be totipotent (Fig. 24A). In one embryo, ES cell progeny were observed only in extraembryonic tissues. Whole-mount immunostaining showed that progeny of injected ES cells (marked by H2B-Tomato) found in the extraembryonic regions, expressed GATA6 (endoderm) or KRT7 (trophoblast) (Fig. 24B). Interestingly, there was large variation in the ability of ES cells to contribute to the embryo. Some cells were able to colonise the whole epiblast while others colonised only a specific region (Fig. 24A).

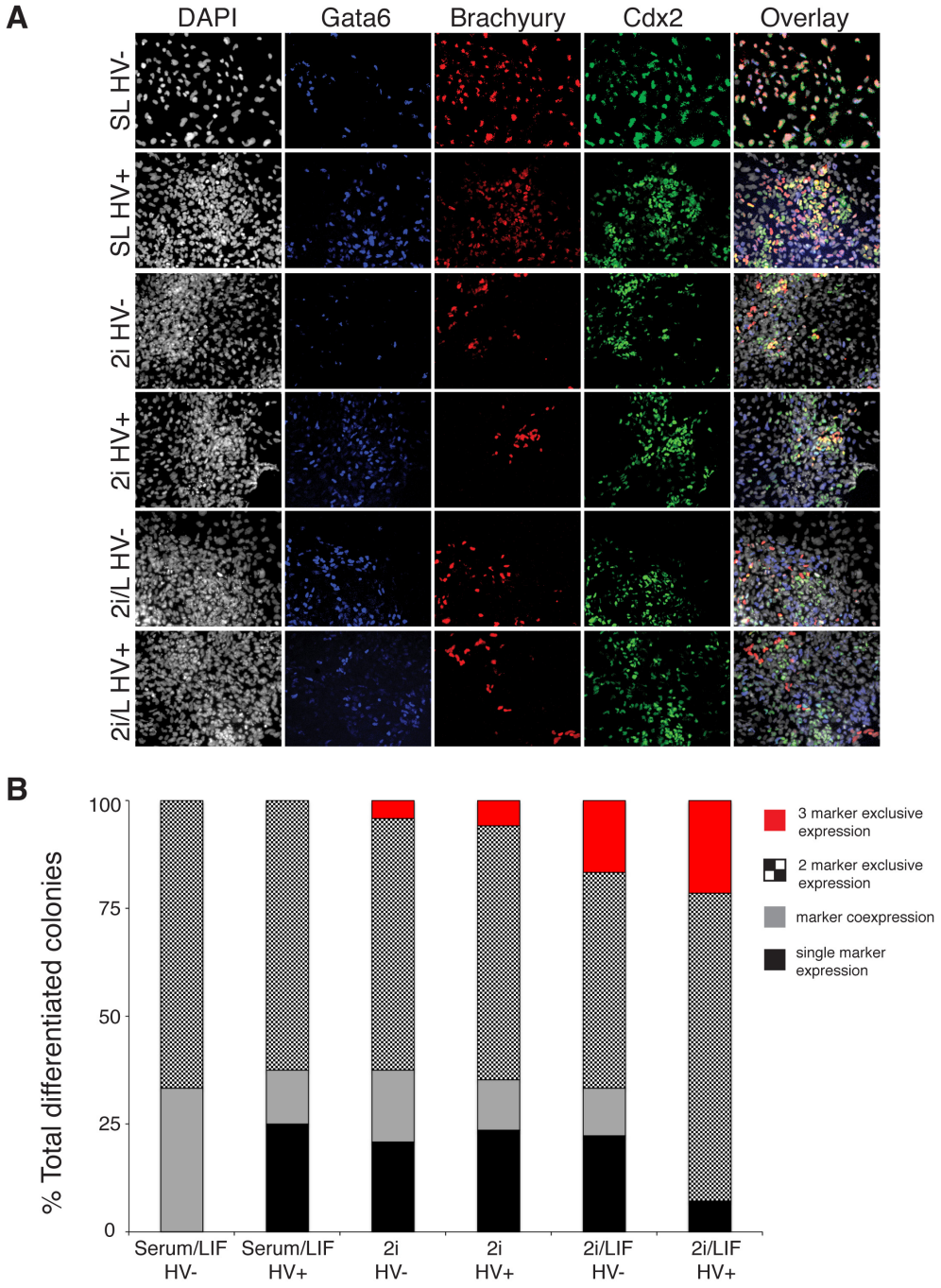


Figure 23. Clonal differentiation of single cells in LIF withdawal conditions. A. Immunostaining of ES cells cultured in serum/LIF (SL), 2i or 2i/LIF (2iL) that were sorted into HV$^-$ or HV$^+$ populations and single cells plated directly by FACS into 96-well plates in LIF withdrawal conditions. Cells were fixed and analysed after 7 days of LIF withdrawal. **B.** Quantification of the expression of different lineage markers within single differentiated colonies. Red bars represent a colony where a single cell has given rise to cells mutually exclusively expressing GATA6, CDX2 and BRACHYURY.

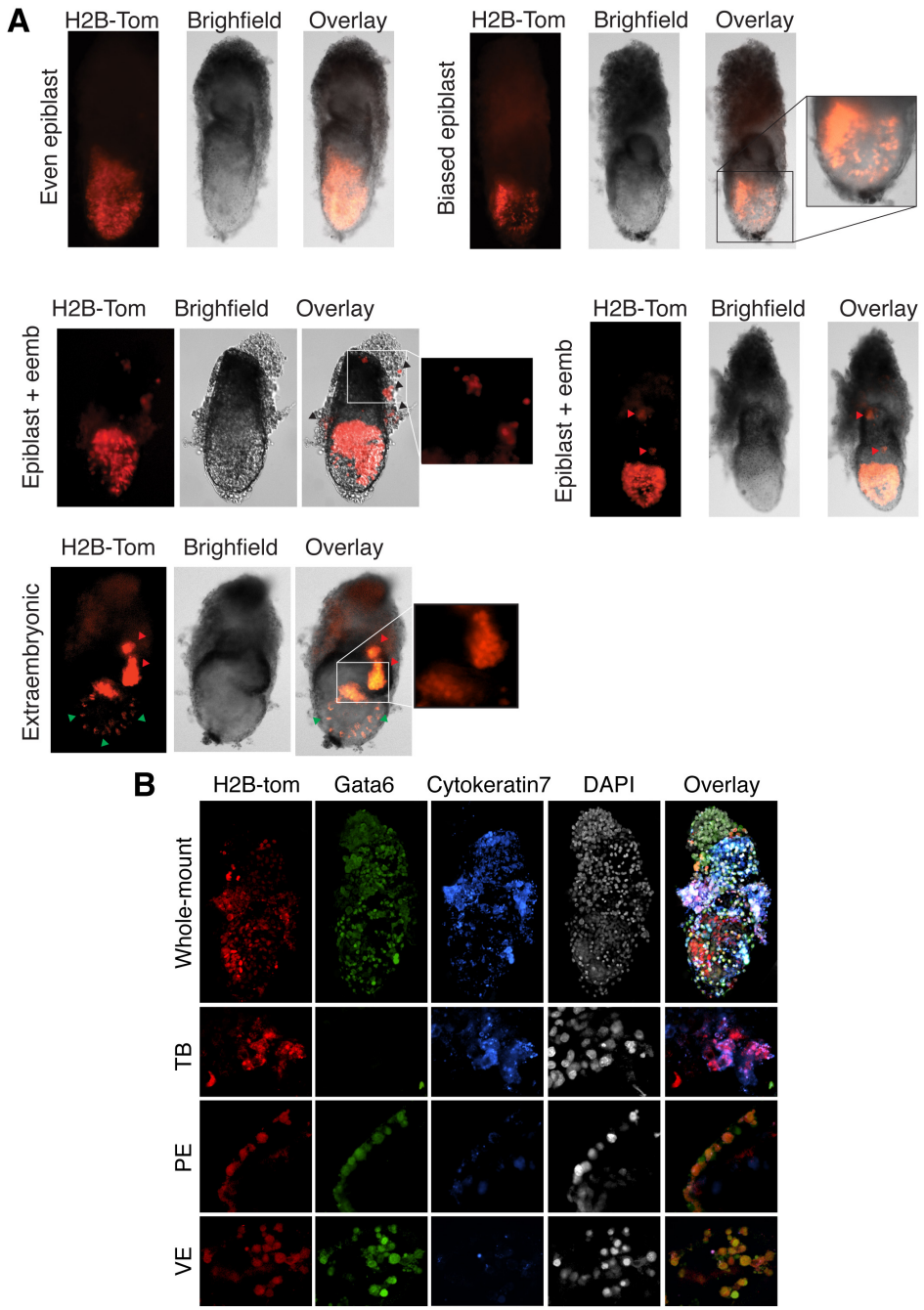


Figure 24. Contribution of single Hex-venus expressing (HV⁺) ES cells to gastrulation stage embryos. A. HV ES cells containing a constitutive H2B-Tomato lineage marker were cultured in 2i/LIF and sorted into HV⁻ and HV⁺ before single cells were injected into wild type morulae and dissected at E6.5. Example images are shown of chimaeras generated from HV⁺ cells. Red arrowheads indicate cells in trophoblast, green arrowheards indicate cells in visceral endoderm. Black arrowheads indicate cells in parietal endoderm. **B.** Whole-mount immunostaining of an E6.5 chimaera showing H2B-Tomato cells in the endoderm, trophoblast and epiblast.

3.7 Discussion

In this chapter I have shown that pre-implantation embryos cultured in 2i appear to be arrested at a stage of development prior to segregation of the epiblast and PE, where markers of both of these lineages continue to be expressed within the ICM. Likewise, ES cells cultured in 2i are heterogeneous and contain a population of cells that coexpress both extraembryonic and embryonic markers. When reintroduced into embryos, this population is able to contribute to embryonic lineages as well as the extraembryonic endoderm and trophoblast. A model of these findings is presented in Figure 25.

Culture in 2i is purported to maintain a "naïve" pluripotent state by shielding ES cells from differentiation-promoting signals (Nichols et al., 2009; Wray et al., 2011), but despite more homogeneous NANOG expression, there is little evidence that cells in 2i represent a single cell type. As well as heterogeneous expression of *Hex* in 2i-cultured ES cells, the PGC and ES cell marker *Stella* was also highly heterogeneous in 2i (Fig. 21).

Pre-implantation embryos cultured in 2i retained the early heterogeneous expression of *Hex*, and ES cells in 2i contained a population capable of generating PE and trophoblast. This suggests that 2i blocks the commitment of totipotent cells to embryonic or extraembryonic lineages. Cells and embryos maintain the coexpression of epiblast and extraembryonic markers that occurs in early pre-implantation development and has been observed for the embryonic markers NANOG and OCT4 and the extraembryonic marker CDX2 (Dietrich and Hiiragi, 2007). The 2i MEK inhibitor PD0325901 blocks FGF signalling, which is important for the segregation of epiblast and PE (Chazaud et al., 2006; Hamazaki et al., 2006; Nichols et al., 2009; Yamanaka et al., 2010). It is also important for the support and expansion of trophoblast (Quinn et al., 2006; Tanaka et al., 1998); thus, blocking this signal could prevent embryonic-extraembryonic lineage segregation. I found that the low-level expression of extraembryonic markers correlated with totipotent function in ES cells, reminiscent of the greater lineage flexibility observed in PE-primed blastomeres compared to their epiblast-primed counterparts (Grabarek et al., 2012). Similarly, the expression of the early PE marker PDGFR-α was maintained in embryos cultured in 2i (Yamanaka et al., 2010). This continued expression was originally attributed to the persistence of the stable GFP reporter. However, it has since been shown that, while inhibitor treated embryos maintained *Pdgfr-α* expression (Kang et al., 2013), FGF null embryos recapitulated the early wild type expression of *Pdgfr-α* but did not maintain expression at later stages (Kang et al., 2013). This suggests that the inhibitors mediate a differential effect upon lineage segregation than in FGF pathway knockout mice.

Although it has been shown previously that ES cells can contribute to extraembryonic lineages (Beddington and Robertson, 1989; Canham et al., 2010; Lallemand and Brulet, 1990; Macfarlan et al., 2012; Suemori et al., 1990), these events occurred at a low frequency. Thus, totipotent cells could exist in most ES cell cultures, but because they are relatively rare, they are frequently not

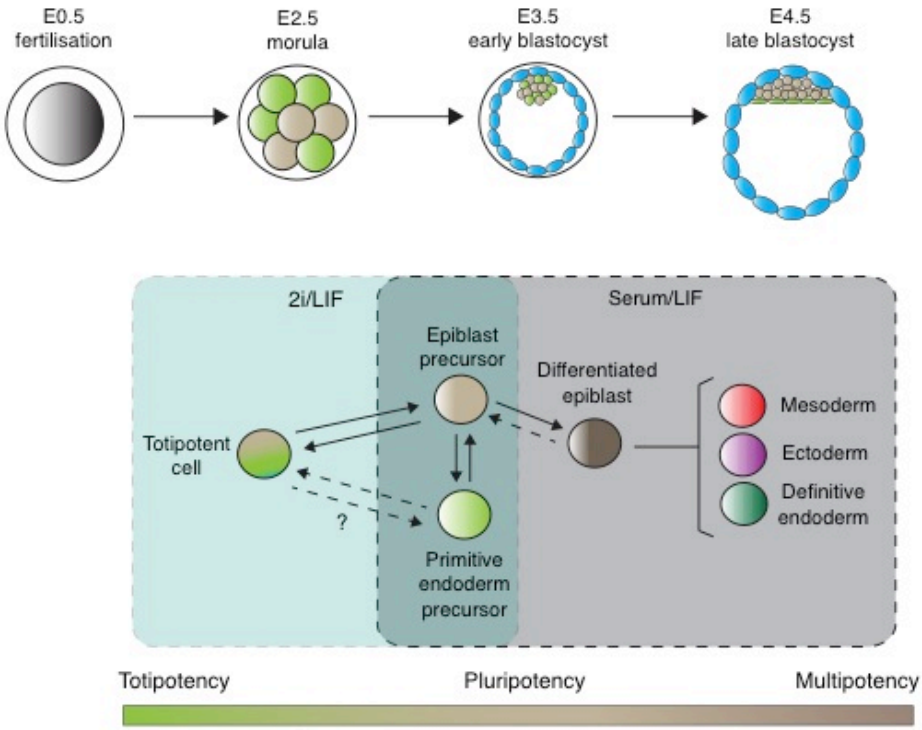


Figure 25. Model summarising the key findings of Chapter 3. The model relates serum/LIF and 2i ESC cultures to embryonic development. 2i/LIF cultures contain totipotent single cells that are comparable to early pre-implantation stages when embryonic cells are still totipotent. As FGF is needed for later lineage segregation, inhibiting FGF signaling in 2i seems to block embryos and ESCs at a point prior to this commitment event. Serum/LIF ESC cultures are reminiscent of later embryonic stages when cells are pluripotent and contain more restricted cell types. Although partially differentiated mesoderm and ectoderm appear to be lost in 2i/LIF, endoderm is maintained. Based on our data, totipotent ESCs are able to dynamically interconvert to a pluripotent epiblast precursor state and may also exist in equilibrium with a similar primitive endoderm precursor, although we have not been able to distinguish this population in this study.

acknowledged in literature. Although relatively rare in standard culture conditions, our data suggests that some individual 2i-cultured ES cells are totipotent. Remarkably single cells demonstrated the capacity to generate the majority of the epiblast. While the extent of extraembryonic contribution that was observed was never as great as that which was observed in the epiblast, the level of contribution to trophoblast was similar to that observed in chimaeras generated from TS cells (Quinn et al., 2006) and the number of chimaeras with visceral endoderm contribution, comparable to that observed in chimaeras generated from XEN cells (Kunath et al., 2005). The low-level of contribution to extraembryonic regions compared to embryonic lineages could be a general experimental issue with reintroducing extraembryonically primed cells into the embryo. One possibility is that the extraembryonic environment rapidly promotes differentiation so that proliferation and cell division are slowed, whereas cells in the epiblast can continue to divide and colonise a greater area.

It has been noted that some cells within standard ES cell cultures express embryonic two-cell-stage transcripts and can contribute to both embryonic and extraembryonic lineages (Macfarlan et al., 2012). These cells lack expression of epiblast markers such as OCT4, NANOG, and SOX2, although these genes are expressed from early in mammalian development and throughout the period that mammalian embryos are considered totipotent. Rossant et al. demonstrated that, up until the early blastocyst stage, embryonic cells posses a similar level of flexibility to 2i-cultured HV^{+} ES cells. ICM cells contribute to the trophoblast in morula aggregations and isolated early ICMs implant into the uterus (Rossant and Lis, 1979). Additionally, the majority of outer trophoblast progenitor cells can contribute to the ICM and epiblast in blastocyst injections (Rossant and Vijh, 1980). I also observed that, at this early stage of blastocyst formation, a subset of trophoblast cells expressed high levels of NANOG in conjunction with CDX2 (Fig. 18B). Similarly, NANOG, OCT4, and CDX2 have been observed to be coexpressed throughout early pre-implantation development in both ICM and trophoblast cells (Dietrich and Hiiragi, 2008). Thus, at these stages, the ICM and trophoblast appear to retain plasticity, which is lost as embryos progress to the late blastocyst (Handyside, 1978; Hogan and Tilly, 1978; Rossant and Lis, 1979; Spindle, 1978), where certain ICM cells only exhibit pluripotency (Gardner and Rossant, 1979). Our findings suggest that 2i/LIF promotes the expansion of a totipotent population of cells, reminiscent of these early developmental stages *in vivo*.

CHAPTER 4:

LIF SIGNALLING PROMOTES AN EXTRAEMBRYONIC CELL POPULATION

Chapter 4:

LIF signalling promotes an extraembryonic cell population

Although LIF is routinely used to maintain ES cells in the absence of feeders, the physiological rationale for this is not fully understood (Section 1.3.4.2). Interestingly the addition of LIF to 2i medium increased the expression of the endoderm marker *Hex* (Section 3.2). This observation seems counter-intuitive and at odds with the idea of LIF maintaining ES cells in an undifferentiated state. However, it fits more appropriately with the *in vivo* role of LIF in supporting extraembryonic tissues involved in implantation (Stewart et al., 1992). Here, I probe in more detail the role of LIF in ES cells and pre-implantation mouse development.

4.1 LIF promotes a HV$^+$ extraembryonic-primed embryonic stem cell population

Increasing the level of LIF in serum and 2i cultures increased the expression of *Hex* in a dose-dependent manner. However, it had no effect on the level of *Nanog* expression (Fig. 26A). At 1000U, the dose routinely used to culture ES cells, the effect of LIF saturated and the HV population could not be further increased (Fig. 26A). This effect was not an immediate early response, observed by monitoring the induction of HV upon the addition of LIF to ES cells previously cultured in 2i alone (Fig. 26B). After 1 day of culture in LIF, cells already showed an increase in the proportion of cells expressing HV, although the cultures did not reach the same point as those continuously cultured in 2i/LIF until day 4 (Fig. 26B).

To confirm that LIF was acting more generally on an extraembryonic population rather than specifically on the HV reporter, I first analysed the expression of trophoblast and endoderm genes by qRT-PCR. I observed that ES cells cultured in 2i/LIF expressed higher levels of genes including *Gata6*, *Gata4*, *Dab2*, *Cdx2* and *Tead4* than ES cells cultured in 2i alone (Fig. 26C). I additionally analysed the expression of genes of the endoderm, mesoderm and neurectoderm lineages, as well as ES cell-associated genes, in ES cells where the LIF target *Stat3* was knocked out compared to *Stat3* wild type ES cells. *Stat3* null ES cells showed a downregulation of endoderm markers compared to wild type cells, while the expression of markers of other lineages or ES cell-associated genes was not compromised (Fig. 26D).

I then looked at how LIF globally affects the molecular signature of ES cells cultured in 2i, by carrying out RNA-seq analysis of unsorted ES cells cultured either in 2i alone or 2i with LIF for at least 3 passages. RNA-seq was carried out as previously described in Chapter 3. Here, LIF also promoted the upregulation of endoderm markers and downregulation of the majority of mesoderm and neurectoderm markers. ES cell markers were not affected other than the direct STAT3 target *Klf4* (Fig. 27A). Additionally, when the list of genes that were most significantly upregulated upon the

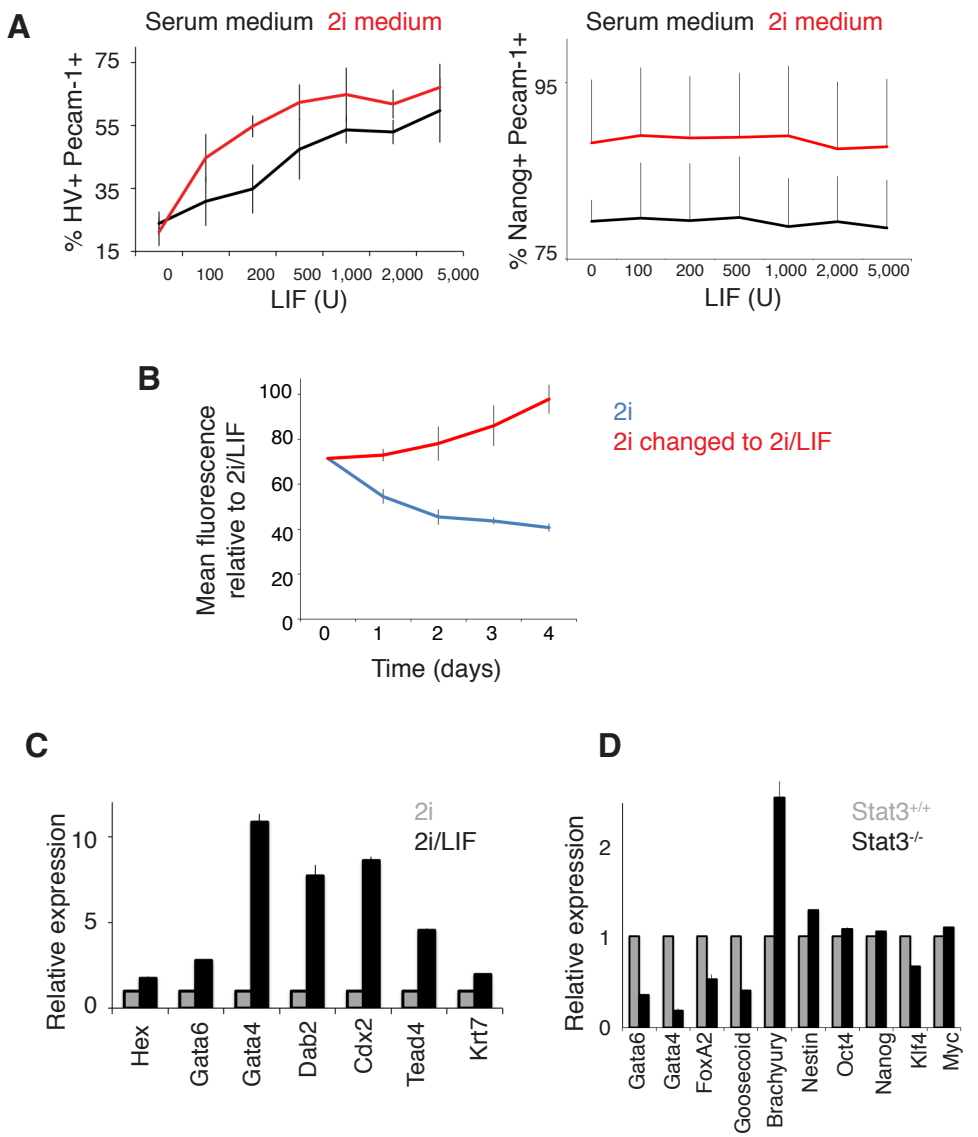


Figure 26. The effect of LIF on endoderm gene expression in ES cells. A. Summary of flow cytometry data showing the effect of increasing LIF doses on the proportion of Hex-venus positive (HV^+), PECAM-1 positive ES cells (left panel) or Nanog-GFP positive, PECAM-1 positive ES cells (right panel). Error bars indicate standard deviation of the mean of 3 biological replicates. in either serum or 2i medium. **B.** HV ES cells were continuously cultured in 2i or 2i/LIF or else cultured in 2i and then changed to 2i/LIF at Day 0. The graphs shows flow cytometry data of the mean HV fluorescence (shown relative to the mean fluorescence of HV ES cells cultured in 2i/LIF). Error bars indicate standard deviation of the mean of 3 biological replicates. **C.** qRT-PCR comparing gene expression levels in ES cells cultured in either 2i or 2i/LIF. Error bars indicate standard deviation of the mean. All samples are normalised to the level of the housekeeping gene TBP. Values are shown relative to 2i. **D.** qRT-PCR comparing gene expression levels in wild type ES cells or $Stat3^{-/-}$ ES cells derived at the same time. Error bars indicate standard deviation of the mean. All samples are normalised to the level of the housekeeping gene TBP. Values are shown relative to $Stat3^{+/+}$

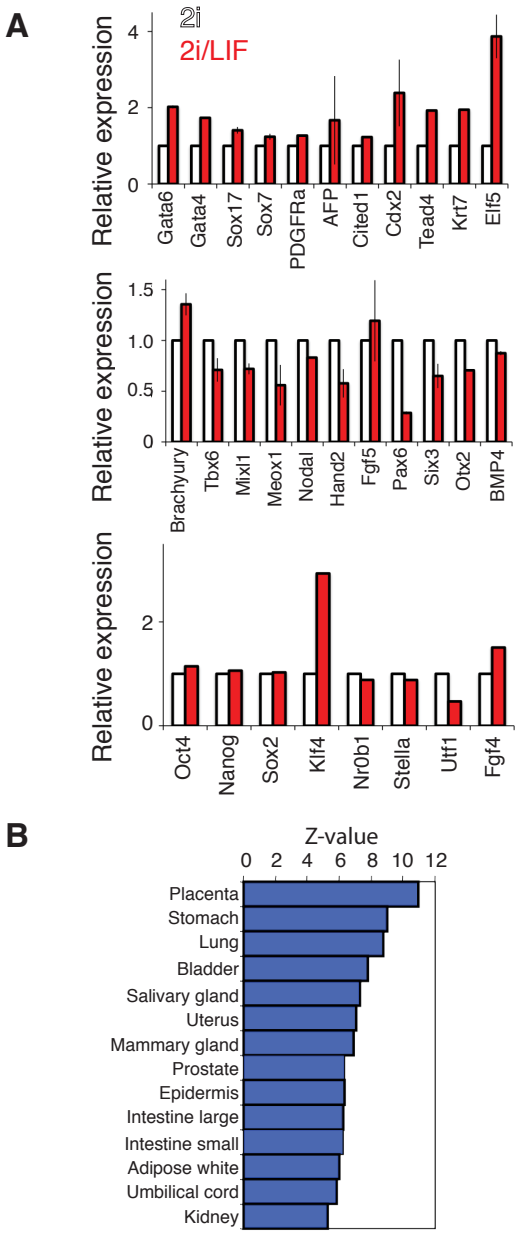


Figure 27. RNA-seq data to determine the global effect of LIF on ES cell cultures. A. RNA-seq data from unsorted ES cells cultured either in 2i or 2i/LIF. Genes are separated into classes of either extraembryonic (upper), mesoderm/neurectoderm (middle) or ES cell (lower) markers. Error bars indicate standard deviation. **B.** Statistical significance of correlation of gene expression differences between ES cells cultured in the presence and absence of LIF and tissue-specific gene expression in the GNF Ver.3 database. Correlation estimated from data on 5085 genes, which represent the overlap between 10,000 most informative genes in our RNA-seq data and 10,000 most informative genes in the GNF database. Organs/tissues sorted by decreasing correlation. Results are shown as a z-value i.e. a score of the standard deviation away from the results observed in 2i alone.

addition of LIF to 2i, were compared to the GNF tissue specific expression database these correlated most strongly with placental tissue, derived from the extraembryonic trophoblast lineage (Fig. 27B).

It has been previously shown that members of the IL-6 cytokine family other than LIF can maintain ES cells in an undifferentiated state (Nichols et al., 1994). I cultured HV ES cells in increasing doses of IL-6 to determine whether other IL-6 family members mimicked the effect of LIF to promote the endoderm-primed ES cell population. IL-6 did promote HV expression although, unlike LIF, this effect did not become saturated at higher doses (Fig. 28).

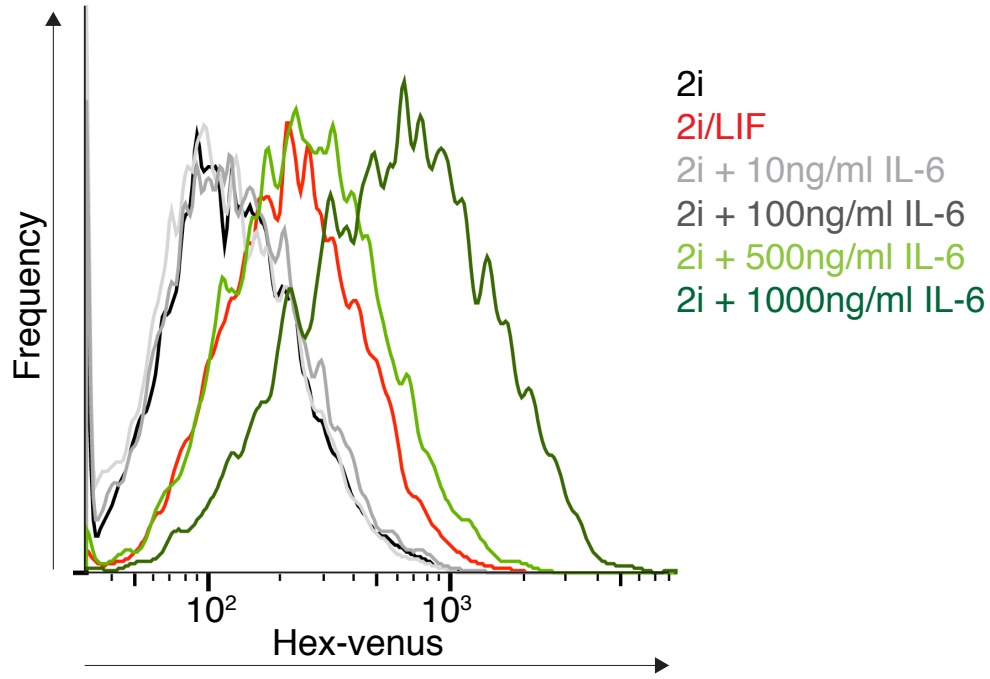


Figure 28. IL-6 effect on Hex-venus (HV) expression in ES cells. A. Histogram showing flow cytometry analysis of HV PECAM-1 positive ES cells cultured for 3 days in increasing concentrations of the cytokine IL-6.

4.2 LIF promotes a HV$^+$ embryonic stem cell state through the JAK/STAT pathway

As the binding of LIF to its receptor can activate multiple downstream pathways, namely the MAPK, PI3K and JAK/STAT pathways (Section 1.3.4.2), I wanted to distinguish which pathway or combination of pathways, was responsible for the increase in the HV$^+$ extraembryonic-primed population. HV ES cells were cultured with LIF in the presence of inhibitors of these pathways; PD (MEK inhibitor, MAPK pathway), LY (PI3K inhibitor) and JAKi (JAK1 inhibitor, JAK/STAT pathway). In serum culture conditions, LY and JAKi both promoted differentiation, as observed by a decrease in PECAM-1 expression, while PD increased PECAM-1 expression (Fig. 29). This data was therefore difficult to interpret as I could not separate changes in HV expression directly mediated by the inhibitors, from those occurring as a result of differentiation. I therefore assessed the effect of these inhibitors on HV expression in ES cells cultured in 2i. HV expression was still promoted upon the addition of LIF to 2i suggesting that the MAPK pathway did not mediate this effect. Only inhibition of the JAK/STAT pathway prevented the LIF-induced increase in HV expression (Fig. 29). Although JAK can also activate the PI3K and MAPK pathways (Fig. 5), as inhibiting these 2 pathways independently did not block the induction of HV expression by LIF, it suggests that LIF mediates this effect through the JAK/STAT pathway.

4.3 LIF increases proliferation of HV$^+$ embryonic stem cells

As LIF was not having an immediate early effect on HV expression (Fig. 26B), I asked how it promoted an increase in the proportion of HV$^+$ ES cells in 2i. One possibility was that it could bypass the blockade to FGF signalling, known to be important for PE specification (Frankenberg et al., 2011; Nichols et al., 2009; Yamanaka et al., 2010) (Section 1.2.3.3). I tested this hypothesis by carrying out Western blots for phospho-ERK on unsorted cells in serum/LIF, 2i and 2i/LIF. As expected, there was a strong phospho-ERK signal in serum/LIF, which was absent in 2i, (due to the presence of the MEK inhibitor) (Fig. 30A). However, LIF did not bypass the blockade to FGF signalling, as in 2i/LIF phospho-ERK was still absent (Fig. 31A).

Another possibility was that LIF promoted the expansion of an endodermal population by increasing proliferation of HV$^+$ cells. HV ES cells cultured in serum/LIF, 2i or 2i/LIF were immunostained for the proliferation marker Ki67 and the number of Ki67$^+$ (proliferating) cells that were either HV$^+$ or HV$^-$ was manually quantified. Upon the addition of LIF there was an approximately 50% increase in the number of non-proliferating HV$^-$ cells (from 4 to 11%) alongside a 50% decrease in the number of non-proliferating HV$^+$ cells (from 11 to 4%) (Fig. 30B,C). Over a prolonged period of time, this small % change would manifest as a gradual increase in the HV$^+$ population, as was observed (Fig. 26B).

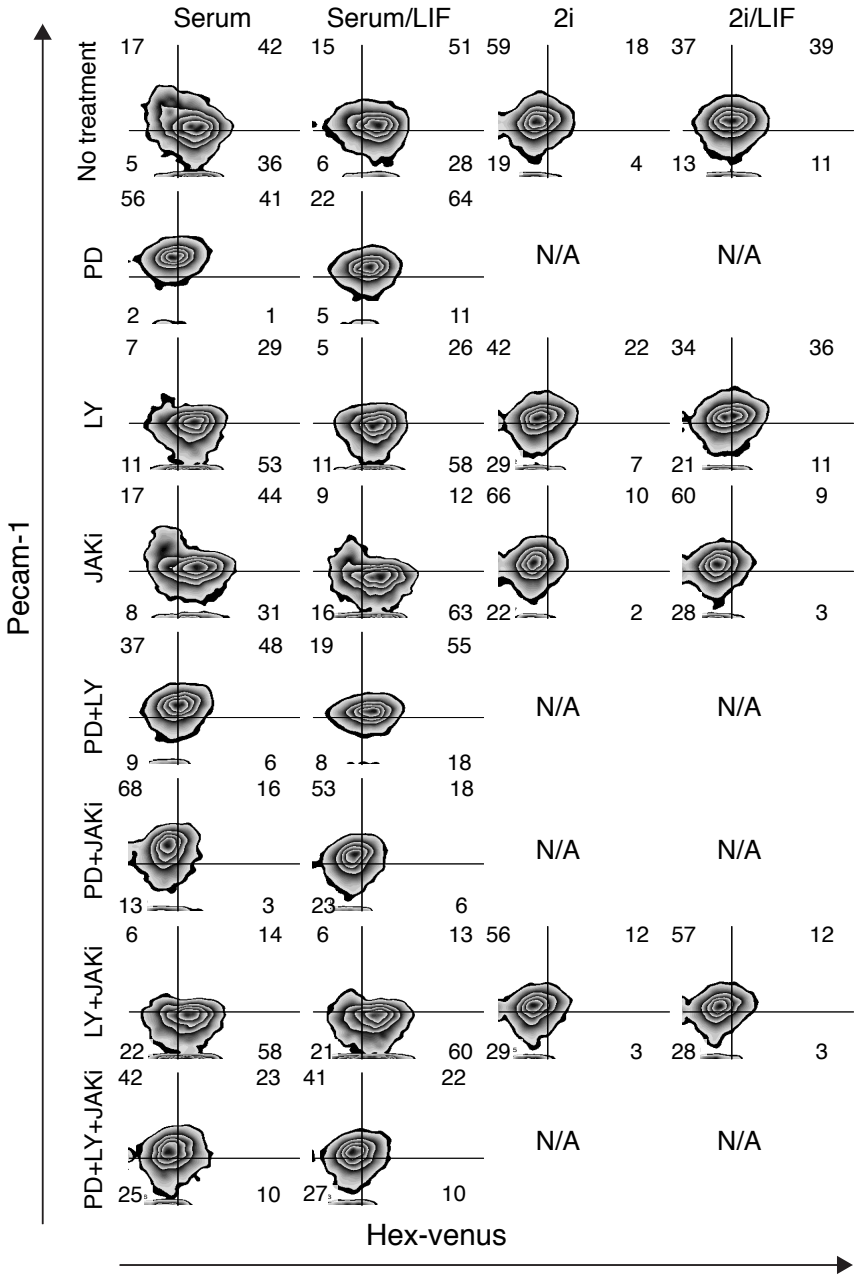


Figure 29. LIF promotes Hex-venus (HV) expression through the JAK/STAT pathway. HV ES cells were cultured in serum, serum/LIF, 2i or 2i/LIF for 3 days in the presence of inhibitors of the MAPK (PD), PI3K (LY) or JAK/STAT (JAKi) pathway, and then analysed by flow cytometry. The numbers in the corners of plots indicate the percentage of cells within that particular gate. Some combinations are not applicable (N/A) as 2i already contains the PD inhibitor. Gates were set based on unstained E14 control cells as in Fig. 6C. PECAM-1 marks undifferentiated ES cells.

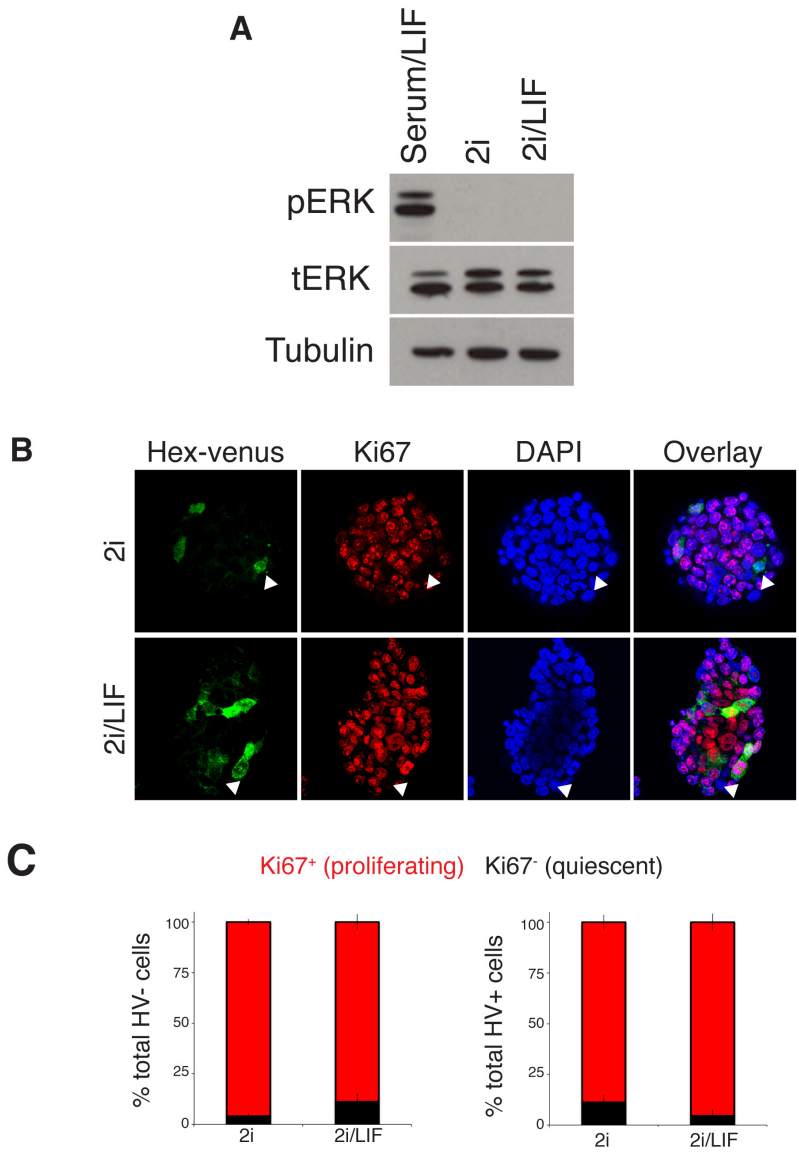


Figure 30. LIF promotes proliferation of Hex-venus-expressing (HV$^+$) ES cells. A. Western blot for phospho ERK of ES cells cultured in serum/LIF, 2i or 2i/LIF. Loading controls of both total ERK and tubulin are shown. **B.** HV ES cells cultured in serum/LIF, 2i or 2i/LIF and immunostained for the proliferation marker Ki67. Images are confocal optical sections through ES cell colonies. White arrows indicate Ki67 negative quiescent cells. **C.** Quantification of Ki67 localisation in HV$^+$ and HV$^-$ cells in each culture condition, from immunostaining. Error bars indicate standard deviation of the mean of 5 ES cell colonies per condition. Quantification was carried out on z-stacks of whole colonies.

I then asked whether LIF differentially affected the HV^- and HV^+ populations due to a difference in their ability to detect LIF. Using flow cytometry, I analysed the expression of the LIF receptor on HV^- and HV^+ cells by flow cytometry. When the lower and upper 10% of *Hex*-expressing cells was selected, HV^+ cells expressed a 20% higher level of the LIF receptor than HV^- cells (Fig. 31A,B).

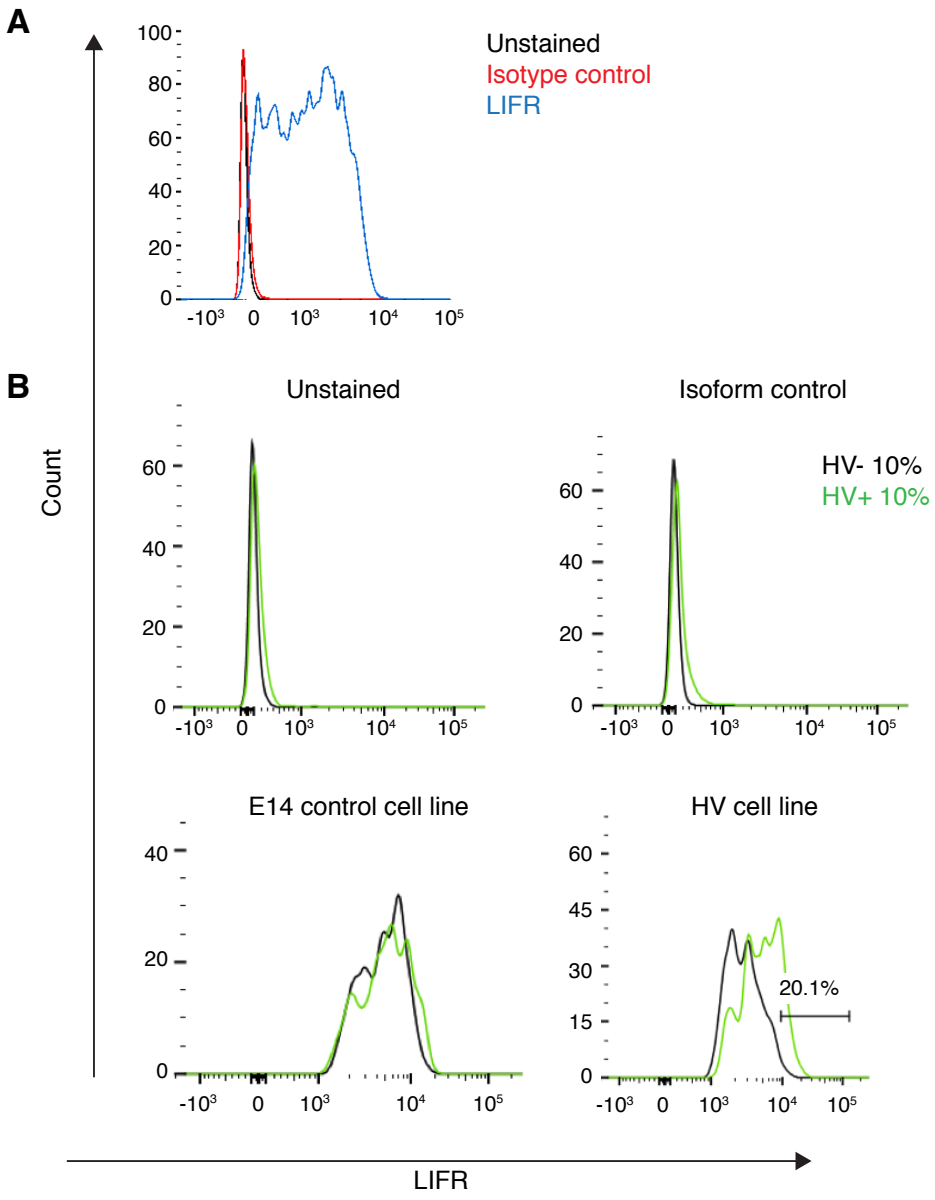


Figure 31. LIF receptor (LIFR) expression within Hex-venus (HV) ES cells. A. Histograms showing flow cytometry analysis of ES cells antibody stained for the LIFR. Only cells expressing the marker of undifferentiated ES cells, SSEA-1, were analysed for the expression of the LIFR. Unstained and isotype antibody controls were also analysed to determine specificity of the signal. **B.** HV cells were used for the analysis so the top and bottom 10% of HV expression was selected by gating and analysed for expression of the LIF receptor. The top and bottom 10% of HV expression was also selected in control unstained samples, isotype controls and also the E14 cell line without the HV reporter, where any signal would correspond to autofluorescence.

4.4 Expression of LIF targets *in vitro*

As LIF elicited different effects, and the LIFR was expressed at different levels, on HV^+ and HV^- ES cells I asked whether the LIF pathway was differentially activated in these populations. To do this, I carried out immunostaining of downstream targets of the LIF pathway, KLF4 and phospho-STAT3 (pSTAT3), in HV ES cells and analysed their expression by confocal microscopy. As previously shown, the epiblast marker, NANOG and the PE marker, HV are expressed in a mutually exclusive manner in ES cells cultured in serum (Fig. 32A). However, KLF4 was expressed in $NANOG^+$ and HV^+ cells to a similar extent (Fig. 32A). This seems to suggest that KLF4 is not strictly associated with either the embryonic or extraembryonic-primed ES cell populations. The correct antibody combination could not be used to stain for both HV and pSTAT3 at the same time but the expression of pSTAT3 was investigated in relation to NANOG and KLF4. Although pSTAT3 showed a strong correlation with KLF4, the 2 markers did not completely overlap. This contrasted with NANOG and pSTAT3 that were predominantly expressed in a mutually exclusive manner (Fig. 32B), suggesting that pSTAT3 correlates more strongly with an extraembryonic rather than an embryonic state.

Upon treatment for 24 hours with the JAK/STAT inhibitor, JAKi, a loss of pSTAT3 was observed as well as a decrease in the expression of both KLF4 and NANOG (Fig. 33A,B), despite the fact that pSTAT3 is not expressed in $NANOG^+$ cells. This suggests that this effect could be mediated indirectly through the effect of JAKi on other pathways, such as the PI3K or MAPK pathway (Fig. 5). As discussed earlier (Sections 1.3.4.2 and 1.3.4.4) the PI3K pathway can indirectly regulate the expression of *Nanog*. The expression of *Nanog* is also directly regulated by the LIF target KLF4 (Zhang et al., 2010), hence a loss of KLF4 expression through JAK/STAT inhibition could consequently lead to a downregulation of NANOG. However, the expression of OCT4 and HV was unchanged during this time period (Fig. 33A-C). This again suggests that the effect of LIF signalling on the HV population is by an indirect mechanism, such as proliferation (Section 4.3) rather than direct transcriptional regulation.

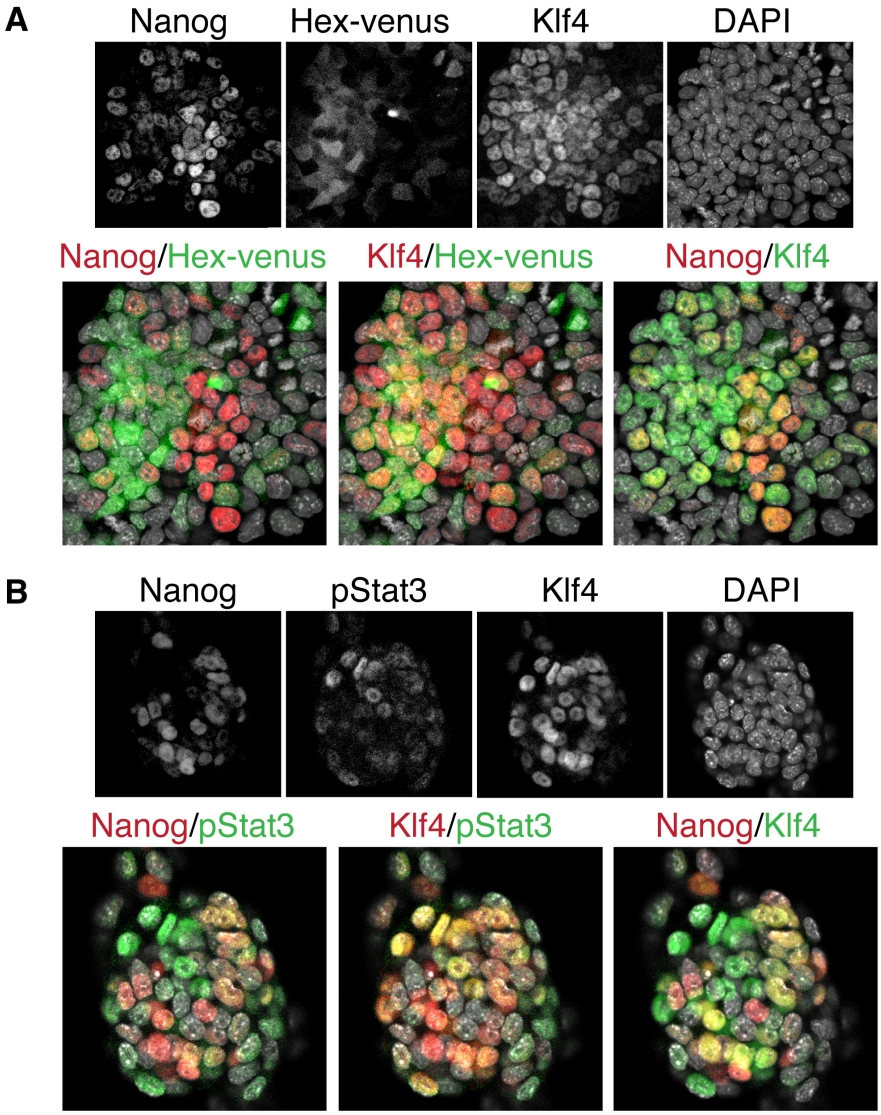


Figure 32. In vitro expression of targets downstream of LIF signalling. Confocal optical sections of undifferentiated ES cell colonies immunostained for **A.** NANOG, Hex-venus (HV) and KLF4 or **B.** NANOG, KLF4and phospho-STAT3. Overlay images are shown for each marker combination to clearly demonstrate coloclalisation of genes.

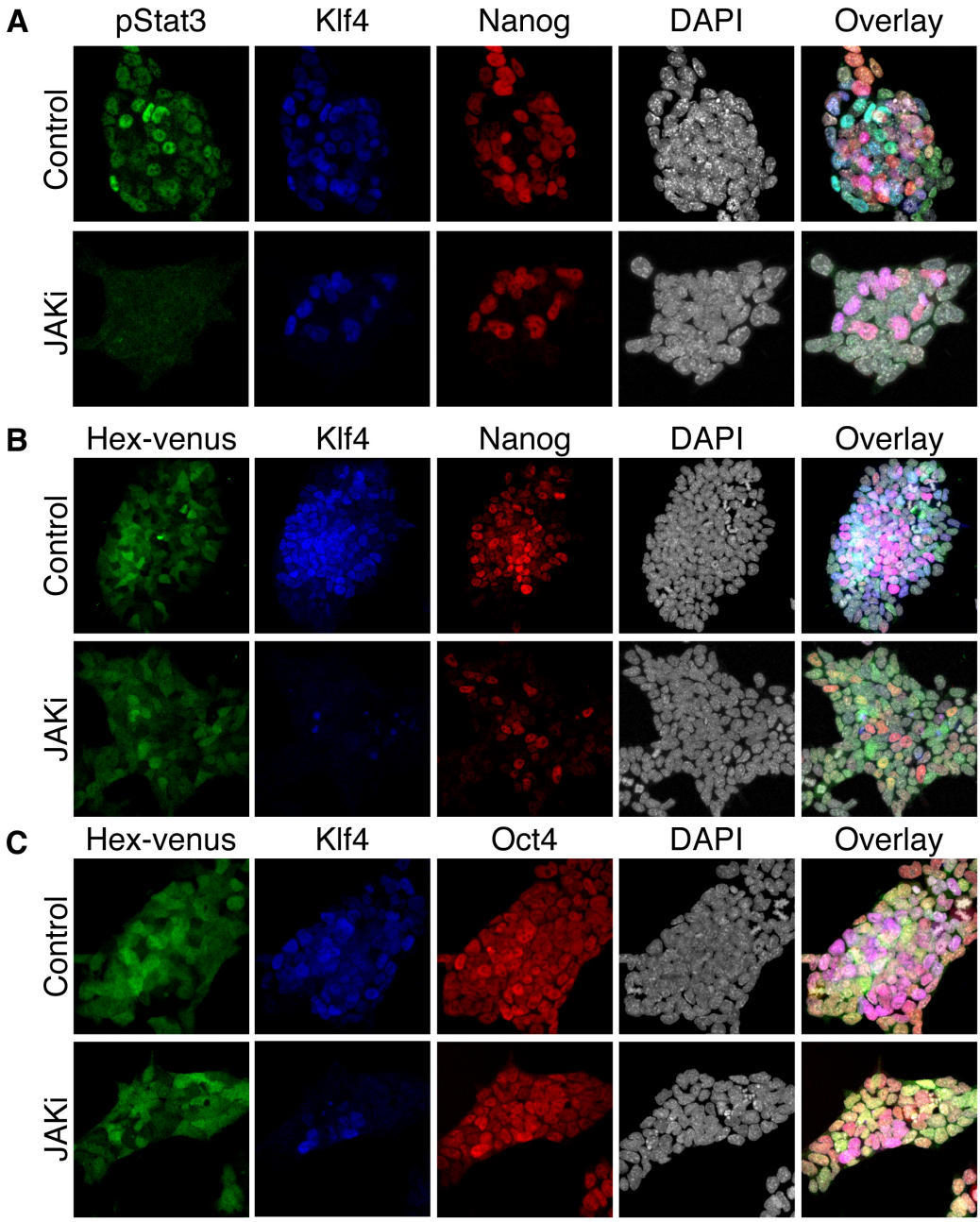


Figure 33. Blocking JAK/STAT signalling causes a rapid decrease in phospho-STAT3, KLF4 and NANOG expression. Confocal optical sections of ES cell colonies in control conditions or after culture in the JAK/STAT inhibitor, JAKi for 24 hours. Cells were immunostained for **A.** pSTAT3, KLF4, NANOG, **B.** Hex-venus (HV), KLF4, NANOG or **B.** HV, KLF4, OCT4.

4.5 Expression of LIF targets *in vivo*

As the role of LIF *in vivo* during pre-implantation development is not fully understood, I sought to ask where and when this pathway is active in the early embryo. To do this, I examined the localisation of its known downstream targets, KLF4 and pSTAT3 in relation to the extraembryonic marker GATA6 and epiblast marker NANOG. Embryos at different stages of pre-implantation development were immunostained and quantified using the open access software, CellProfiler (developed at MIT and the Broad Institute, USA, www.cellprofiler.org). Immunostaining showed that in C57BL/6 mice, up until the compacted morula stage, the expression of KLF4 and pSTAT3 corresponded more strongly to the extraembryonic marker GATA6 than the embryonic marker NANOG (Fig. 34-36). This was due to the fact that NANOG expression was initiated at the morula stage in only a subset of cells. Outbred CD1 mice were also used to confirm these expression patterns. In CD1 mice, at this stage NANOG showed homogeneous expression throughout the morula hence all markers were equally correlated. At later stages C57BL/6 and CD1 mice showed identical expression patterns hence data was pooled. Throughout the blastocyst stages, pSTAT3 was correlated with GATA6 and was specifically expressed in extraembryonic lineages by the late blastocyst stage (Fig. 35,36). The expression pattern of KLF4 was more complex. In the early blastocyst KLF4 was expressed heterogeneously and its expression overlapped with both $GATA6^{+}$ extraembryonic precursors as well as $NANOG^{+}$ embryonic precursors (Fig. 34,36). At later stages KLF4 was expressed highly in the embryonic epiblast, but also expressed at lower levels in the PE (Fig. 34,36). Analysis of these trends by CellProfiler was validated initially by quantifying the coexpression of the markers, NANOG, GATA6 and CDX2 that have known expression patterns (Fig. 37,38). The expected trend was observed where, by the late blastocyst stage, when the trophoblast, PE and epiblast lineages are segregated, $CDX2^{+}$ trophoblast cells showed only low expression of GATA6 and NANOG (Fig. 37,38). The quantification of the coexpression of KLF4 and pSTAT3 with these markers also revealed the patterns expected from direct observation of the immunostaining. As development progressed, more coherent correlations emerged (Fig. 39-43). At E3.5 a large number of $NANOG^{+}$ cells did not express pSTAT3 (Fig. 42). Additionally many $NANOG^{-}$ cells showed low levels of KLF4 expression (Fig. 42). These patterns were further reinforced by E4.5 where NANOG and GATA6 were expressed mutually exclusively in distinct populations and KLF4 and NANOG showed a strong correlation. A proportion of cells expressing low levels of KLF4 were $NANOG^{-}$ and $GATA6^{+}$ hence seemed to correspond to the PE (Fig. 43). At the late blastocyst stage pSTAT3 was expressed exclusively in extraembryonic tissues, the PE and trophoblast (Fig. 35,36). Although KLF4 and pSTAT3 are both regulated by LIF, they were not strongly correlated at the late pre-implantation stages (Fig. 36). Additionally, although LIF signalling through pSTAT3 is correlated with self-renewal (Matsuda et al., 1999; Niwa et al., 1998) (Section 1.3.4.1), a correlation with epiblast identity was not readily apparent.

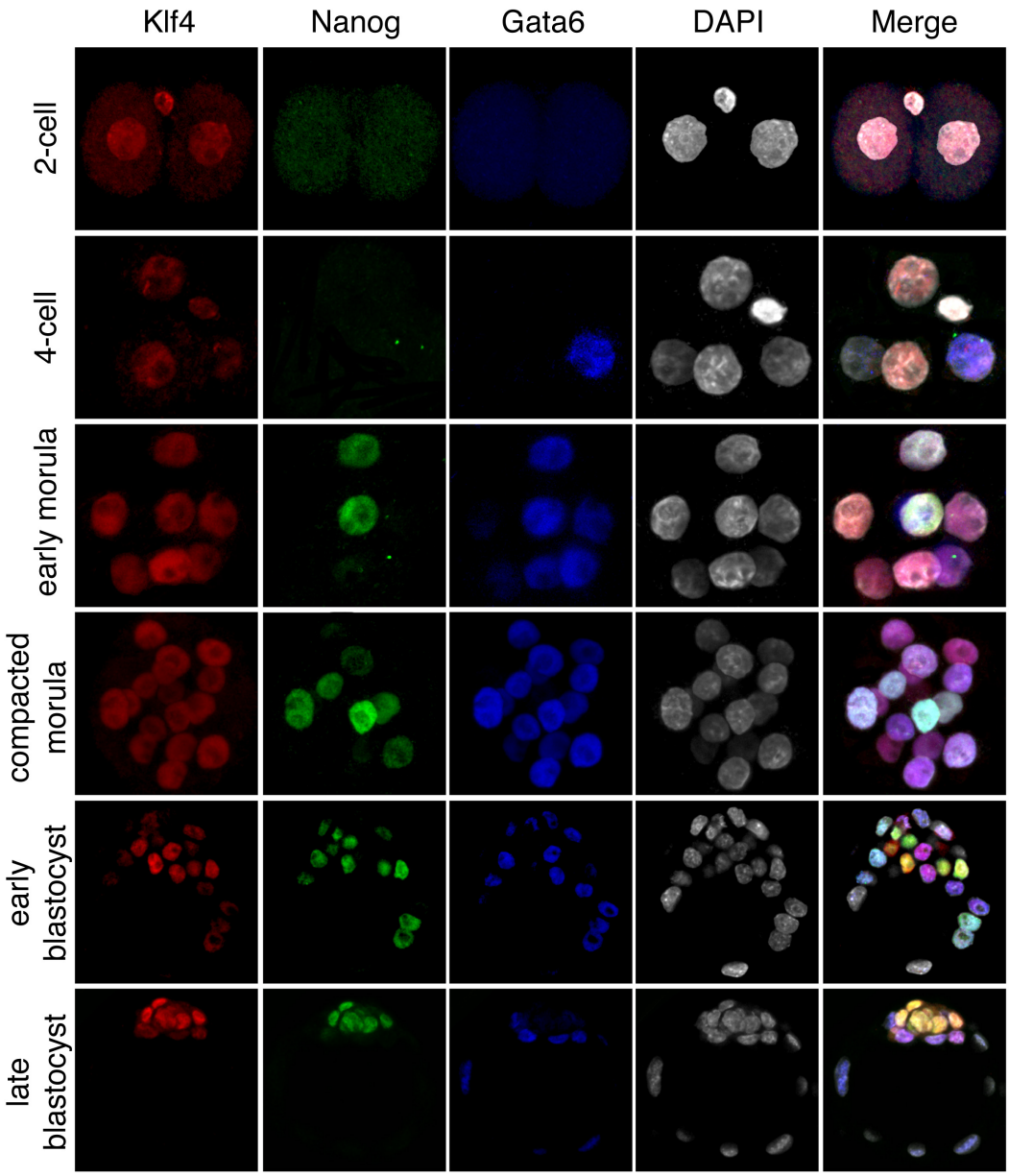


Figure 34. Pre-implantation expression pattern of NANOG, GATA6 and KLF4. Immunostaining of pre-implantation embryos at different stages of development. Images from blastocyst stage embryos represent confocal optical sections through the ICM while images of earlier stages are extended images showing the entire embryo.

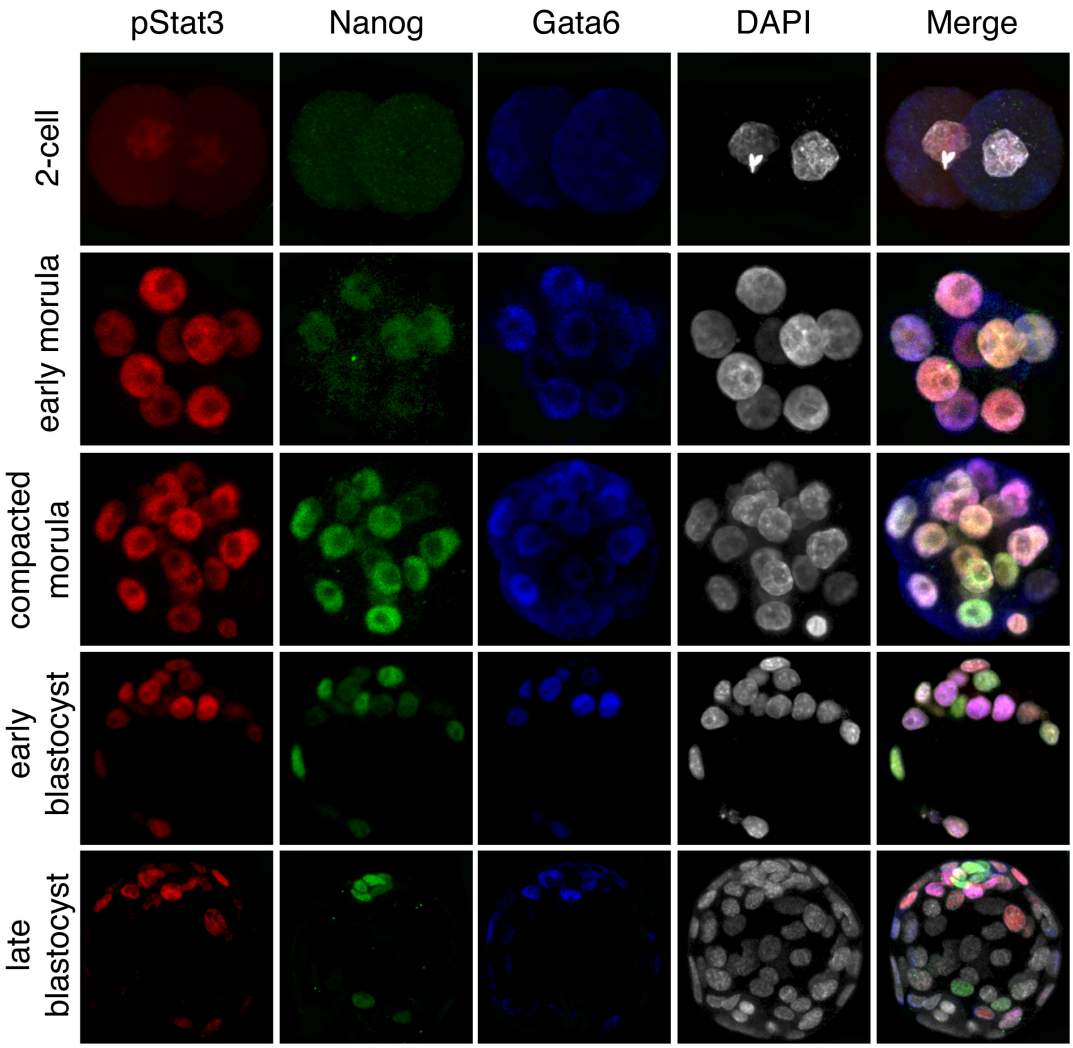


Figure 35. Pre-implantation expression pattern of NANOG, GATA6 and phospho-STAT3. Immunostaining of pre-implantation embryos at different stages of development. Images from blastocyst stage embryos represent confocal optical sections through the ICM while images of earlier stages are extended images showing the entire embryo.

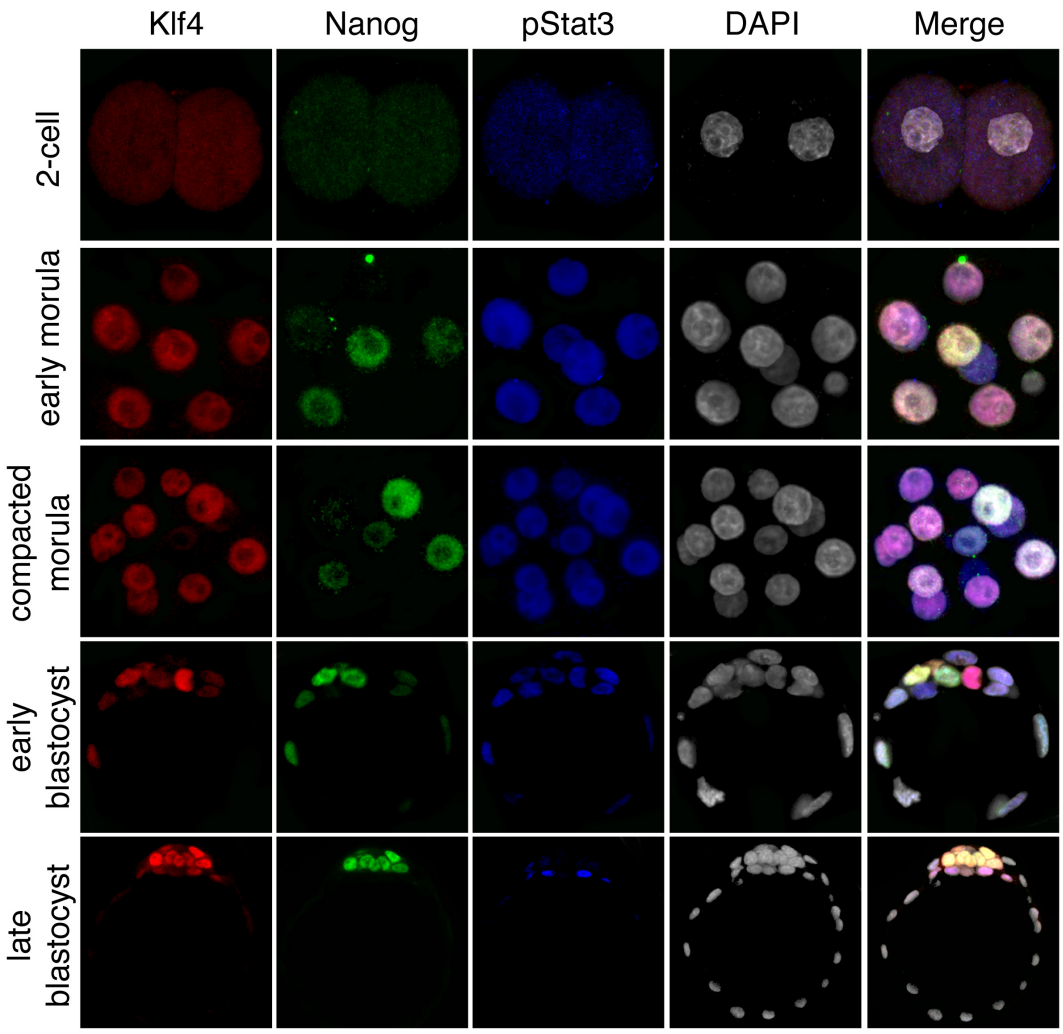


Figure 36. Pre-implantation expression pattern of NANOG, KLF4 and phospho-STAT3.
Immunostaining of pre-implantation embryos at different stages of development. Images from blastocyst stage embryos represent confocal optical sections through the ICM while images of earlier stages are extended images showing the entire embryo.

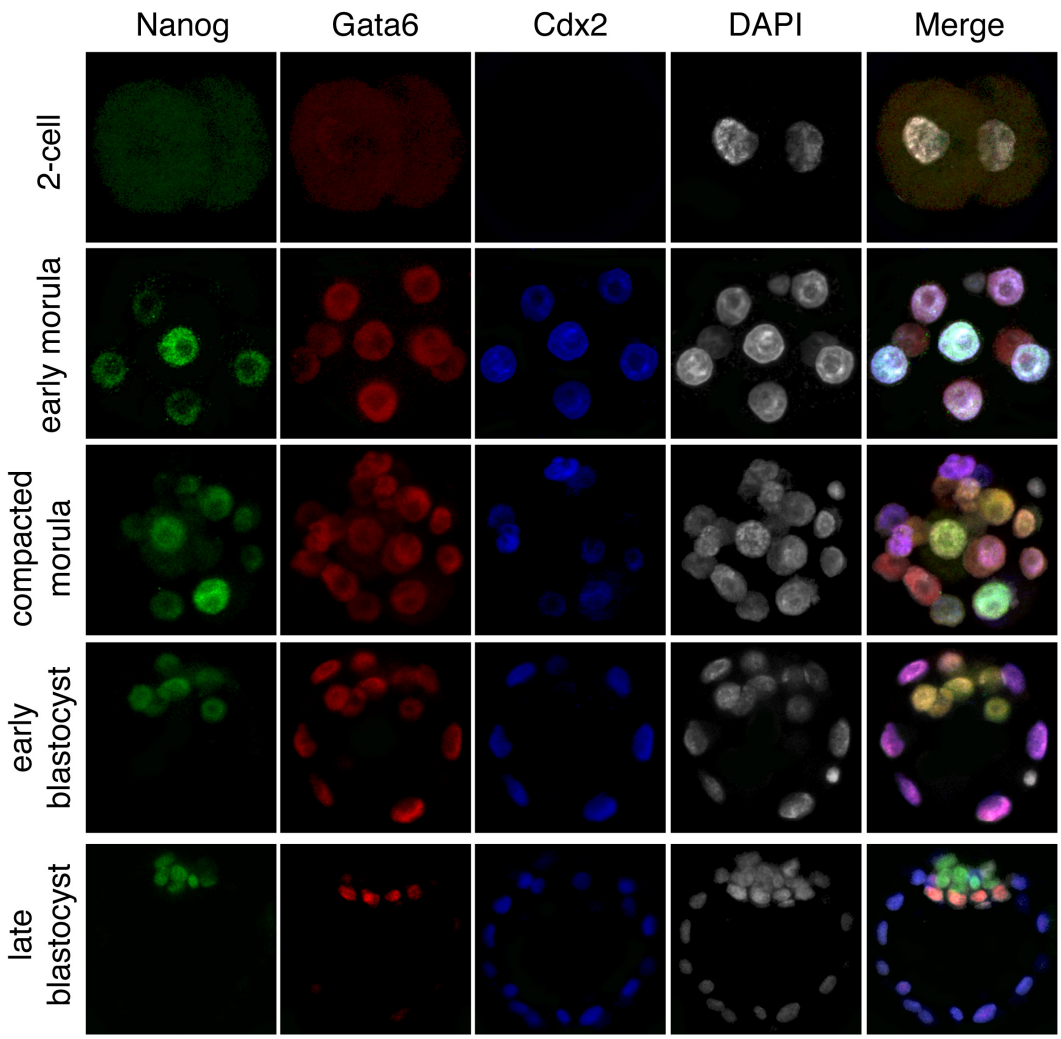


Figure 37. Pre-implantation expression pattern of NANOG, GATA6 and CDX2. Immunostaining of pre-implantation embryos at different stages of development. Images from blastocyst stage embryos represent confocal optical sections through the ICM while images of earlier stages are extended images showing the entire embryo.

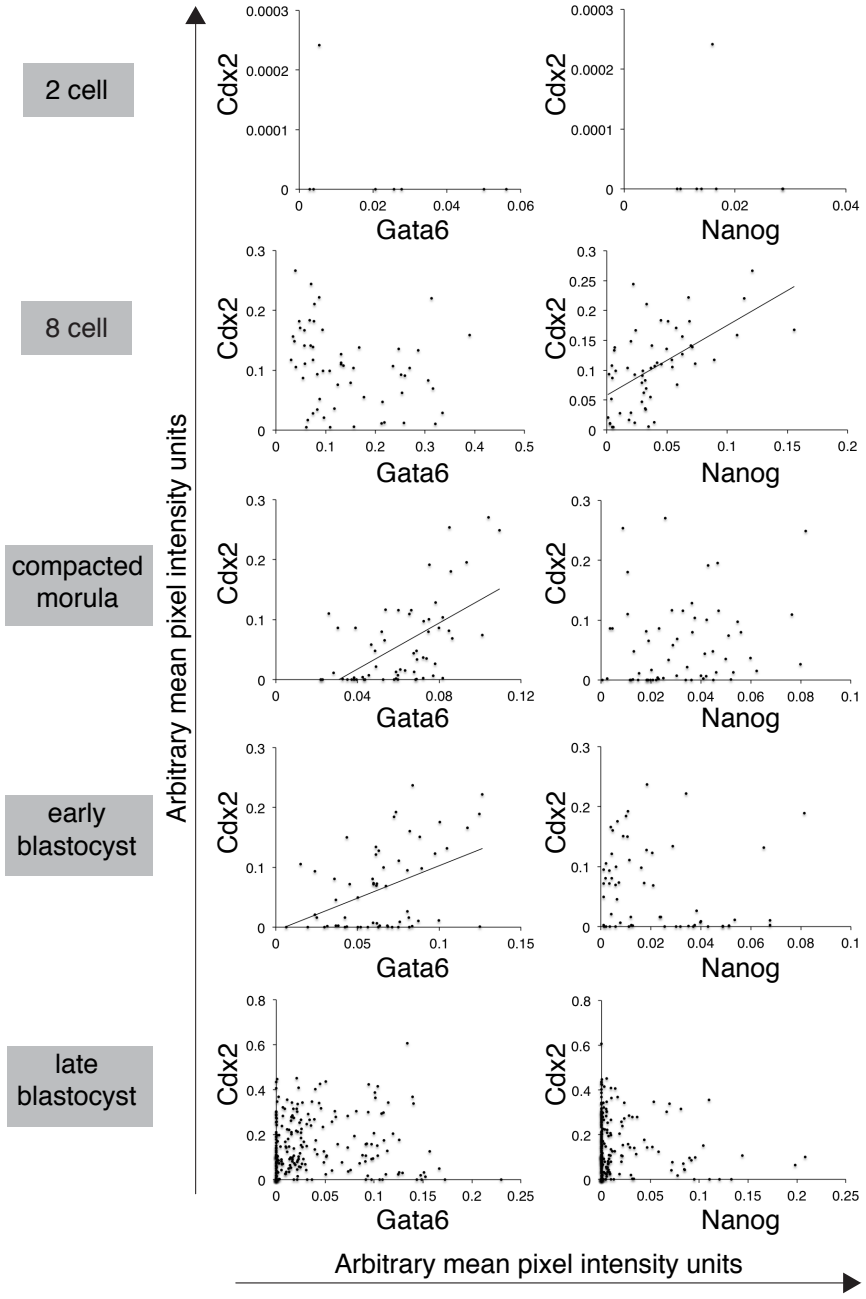


Figure 38. Quantification of colocalisation of NANOG, GATA6 and CDX2 during pre-implantation development. Open source image analysis software, CellProfiler was used to quantify immunostaining in individual cells. Cell nuclei were identified by manual selection and the mean pixel intensity measured in arbitrary units of intensity. For cleavage-stage embryos (stages prior to E3.5) every cell was selected for analysis. For blastocyst stage embryos, 2 optical sections through the ICM were selected and each nuclei in those planes was analysed. Each point on the graph represents the intensity of the noted markers within a single nucleus. Lines of best fit are shown where a trend was evident.

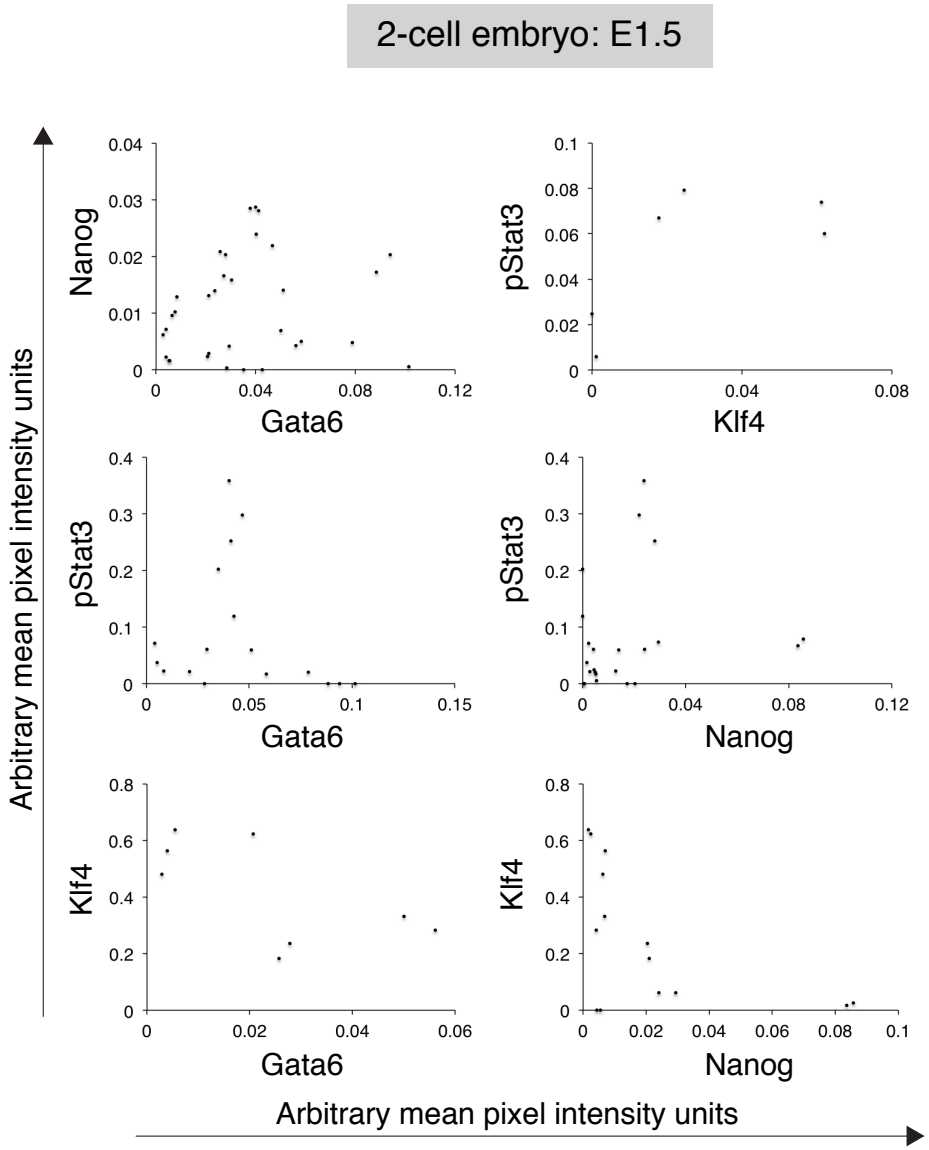


Figure 39. Quantification of colocalisation of various genes at the 2-cell embryo stage. Open source image analysis software, CellProfiler was used to quantify immunostaining in individual cells. Cell nuclei were identified by manual selection and the mean pixel intensity measured in arbitrary units of intensity. Each point on the graph represents the intensity of the noted markers within a single nucleus. Lines of best fit are shown where a trend was evident.

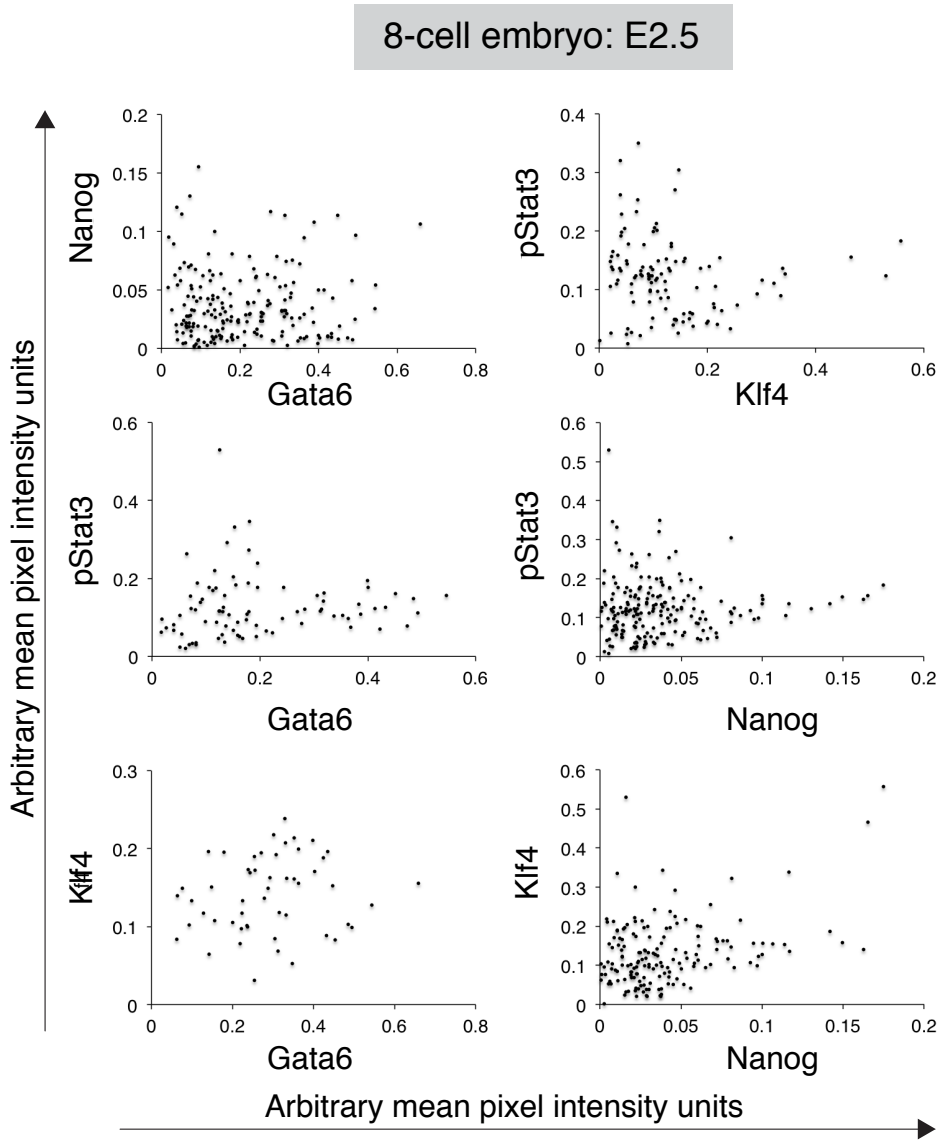


Figure 40. Quantification of colocalisation of various genes at the 8-cell embryo stage. Open source image analysis software, CellProfiler was used to quantify immunostaining in individual cells. Cell nuclei were identified by manual selection and the mean pixel intensity measured in arbitrary units of intensity. Each point on the graph represents the intensity of the noted markers within a single nucleus. Lines of best fit are shown where a trend was evident.

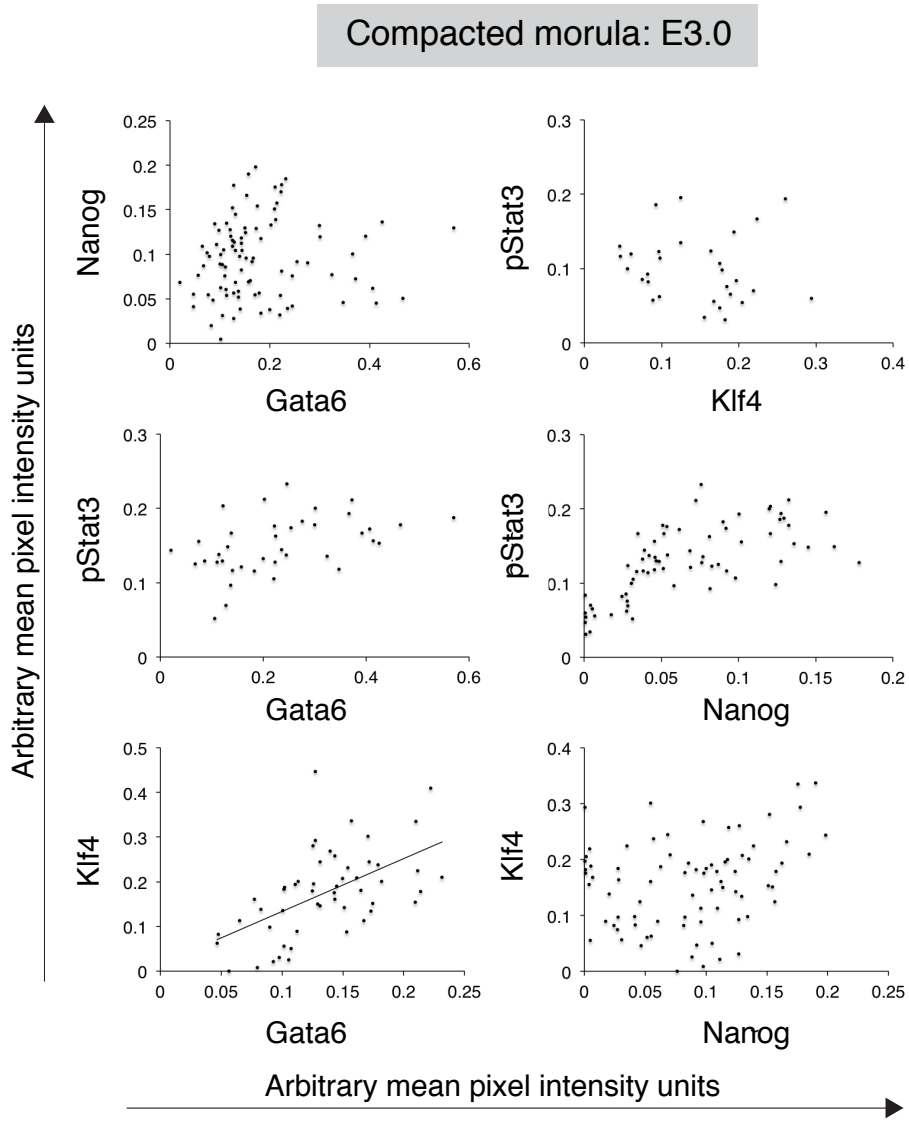


Figure 41. Quantification of colocalisation of various genes at the compacted morula stage. Open source image analysis software, CellProfiler was used to quantify immunostaining in individual cells. Cell nuclei were identified by manual selection and the mean pixel intensity measured in arbitrary units of intensity. Each point on the graph represents the intensity of the noted markers within a single nucleus. Lines of best fit are shown where a trend was evident.

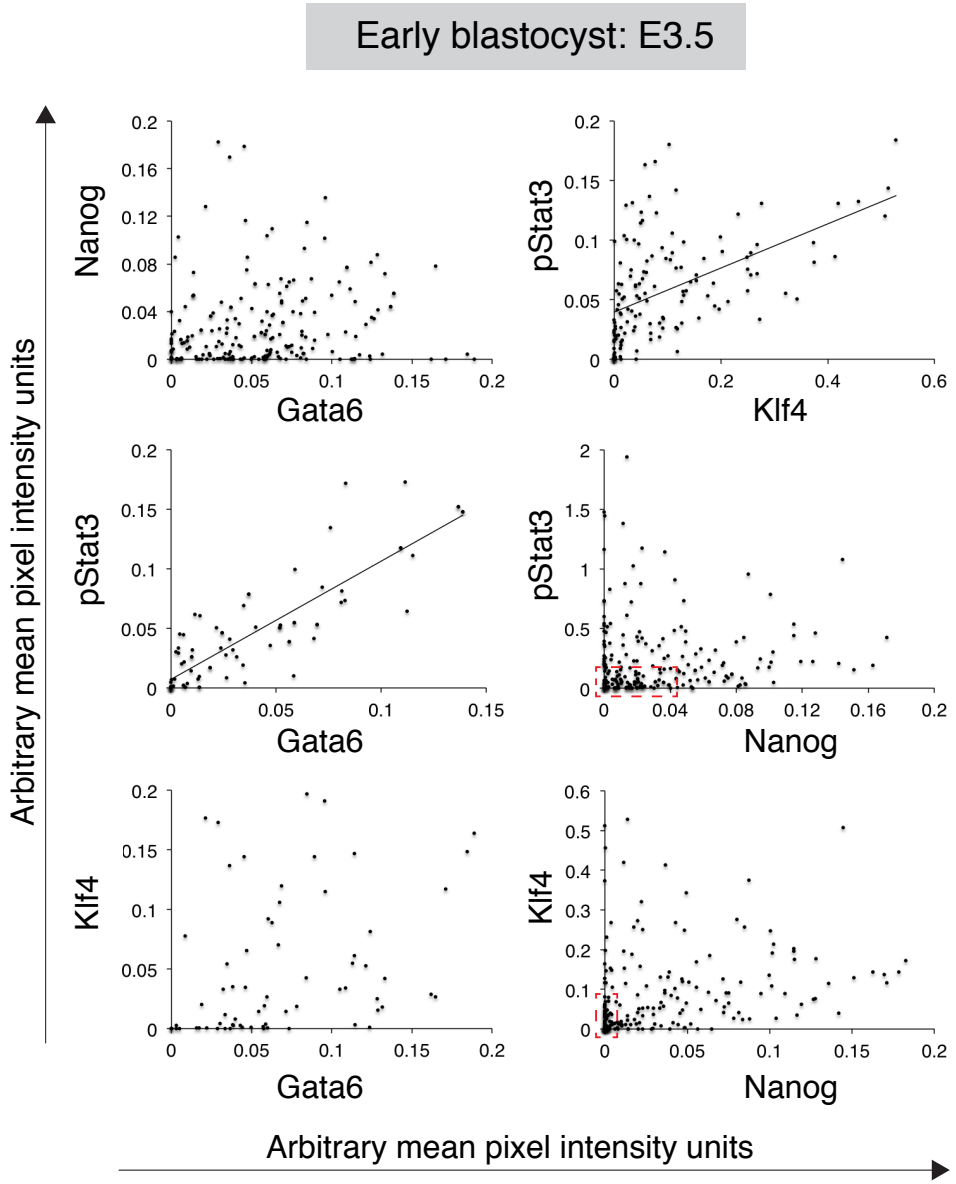


Figure 42. Quantification of colocalisation of various genes at the early blastocyst stage. Open source image analysis software, CellProfiler was used to quantify immunostaining in individual cells. Cell nuclei were identified by manual selection and the mean pixel intensity measured in arbitrary units of intensity. 2 optical sections through the ICM were selected and each nuclei in those planes was analysed. Each point on the graph represents the intensity of the noted markers within a single nucleus. Lines of best fit are shown where a trend was evident. Emerging patterns mentioned in the text are highlighted by dashed red boxes.

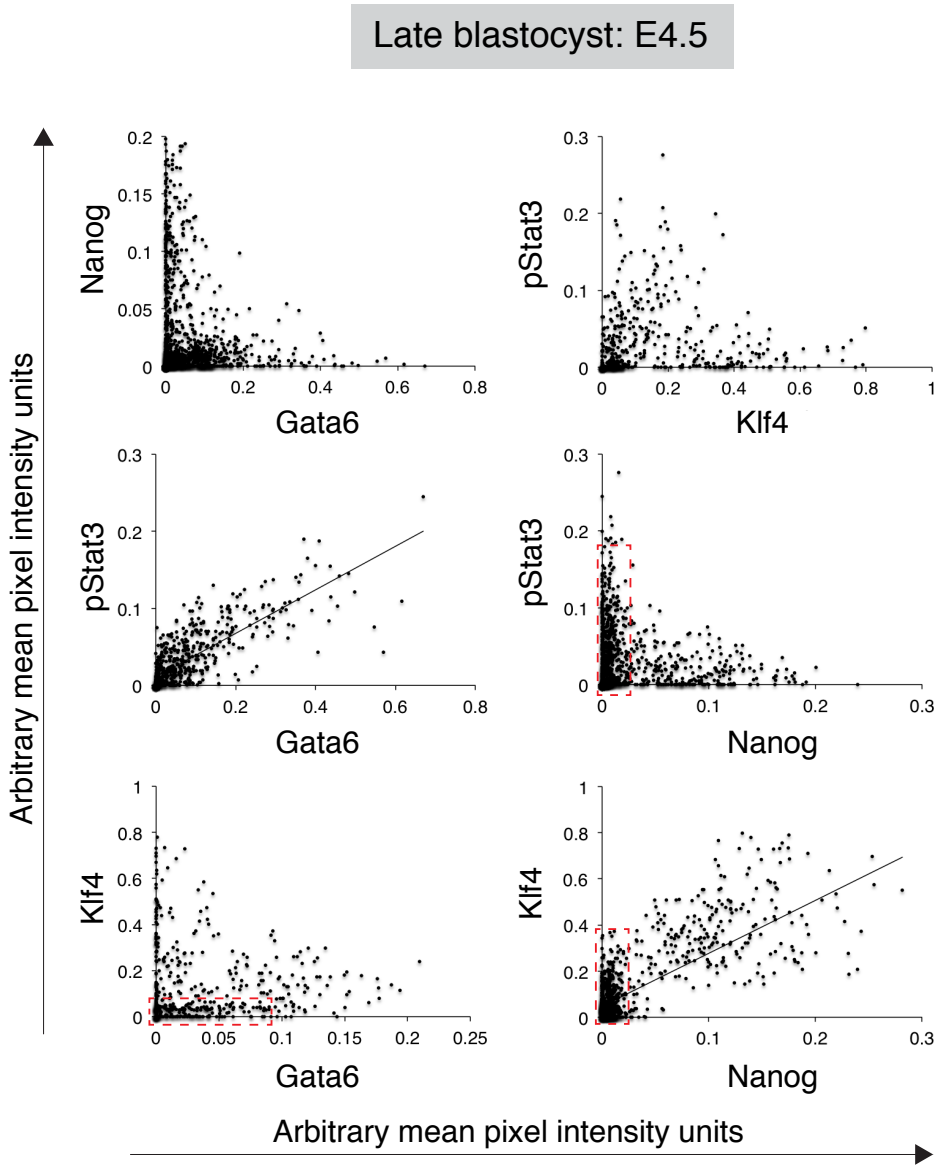


Figure 43. Quantification of colocalisation of various genes at the late blastocyst stage. Open source image analysis software, CellProfiler was used to quantify immunostaining in individual cells. Cell nuclei were identified by manual selection and the mean pixel intensity measured in arbitrary units of intensity. 2 optical sections through the ICM were selected and each nuclei in those planes was analysed.Each point on the graph represents the intensity of the noted markers within a single nucleus. Lines of best fit are shown where a trend was evident. Emerging patterns mentioned in the text are highlighted by dashed red boxes.

4.6 The role of LIF in embryonic stem cell differentiation

As LIF promoted an endoderm-primed population of ES cells, I asked whether it also affected their differentiation capacity. Unsorted ES cells were cultured in either 2i or 2i/LIF and subsequently differentiated by LIF withdrawal. ES cells pre-cultured in LIF were more competent to differentiate towards endoderm (Fig. 44A,B).

I then asked, whether blocking the JAK/STAT pathway downstream of LIF, which mediates the increase in the HV population (Section 4.2), could also affect ES cell differentiation. ES cells were differentiated for 5 days by LIF withdrawal either with or without the JAK/STAT inhibitor, JAKi. Cells were then fixed and immunostained for the lineage markers GATA6, CDX2 and BRACHYURY. While JAKi treatment is compatible with standard ES cell culture, the inclusion of JAKi during differentiation resulted in extensive cell death. Cells differentiated in the presence of JAKi showed a marked decrease in the fraction of surviving $GATA6^{+}$ cells (Fig. 45). However, this could be compensated for by increasing the plating density at the start of the experiment (Fig. 45).

Additionally unsorted cells cultured in 2i alone or 2i/LIF, expressing a constitutive β-gal marker, were introduced into embryos by morula aggregation. At E6.5, contribution was observed to the epiblast and in some cases to the extraembryonic tissues. Although the number of embryos where ES cells contributed to the extraembryonic region was the same after culture in 2i or 2i/LIF, the extent of contribution to the extraembryonic region was higher from cells that had been cultured in the presence of LIF (Fig. 46A-C).

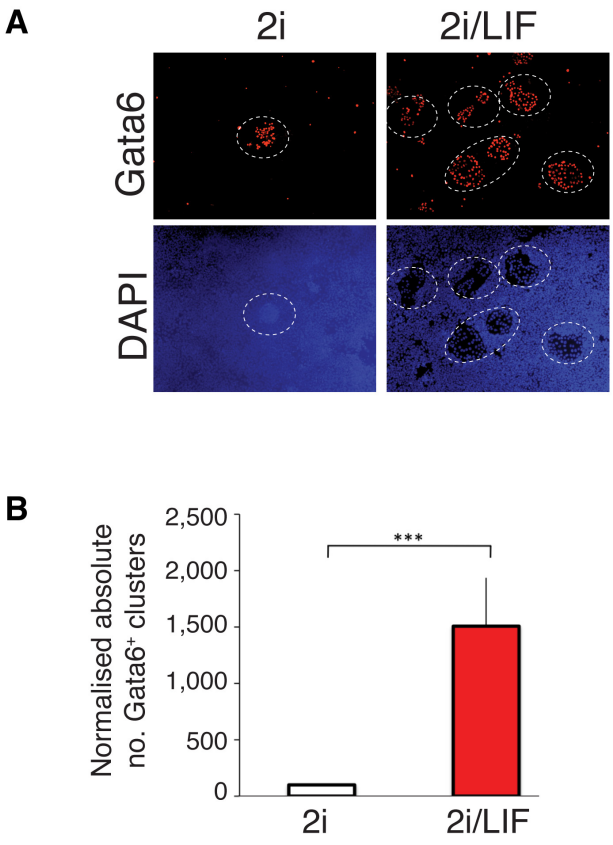


Figure 44. Culturing ES cells in LIF before differentiation increases the extent of endoderm differentiation. A. ES cells cultured in either 2i or 2i/LIF medium were differentiated for 5 days by LIF withdrawal and immunostained for the endoderm marker GATA6. **B.** Quantification of GATA6 positive colonies of endodermal cells observed by immunostaining, normalised to the 2i condition. Error bars indicate standard deviation of the mean of 3 biological replicates (*** = p<0.001).

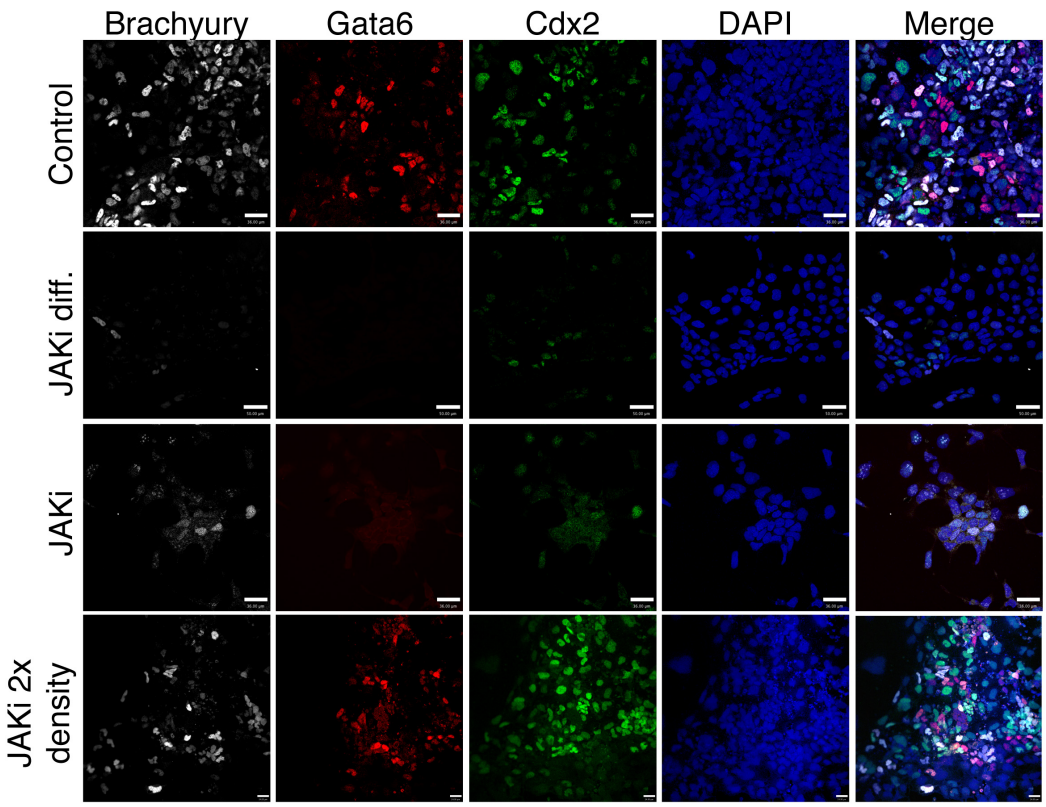


Figure 45. Inhibition of the JAK/STAT pathway causes cell death during differentiation. ES cells were differentiated by LIF withdrawal in various conditions, either control medium, medium where JAKi was added only during differentiation (JAKi diff), medium where JAKi was added for 1 day prior as well as throughout differentiation (JAKi) or where ES cells were plated at 2x the standard density before differentiation. Cells were then stained for different lineage markers to assess differentiation.

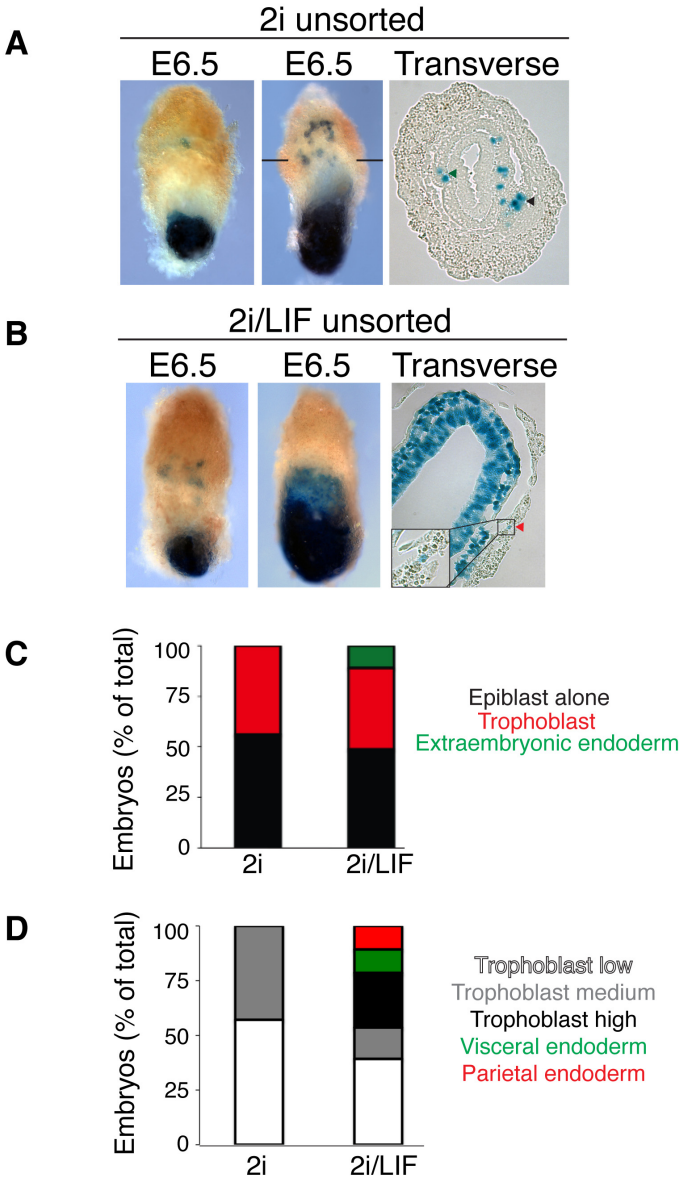


Figure 46. ES cells cultured in 2i/LIF contribute to extraembryonic regions to a greater extent to than when cultured in 2i alone. A. and B Chimaeric embryos were generated by morula aggregations with Hex-venus ES cells expressing a constitutive LacZ lineage tracer. Prior to aggregations, ES cells were cultured in 2i or 2i/LIF. Embryos were collected and X-gal stained at E6.5. Black lines indicate the plane of section in the embryo. Black arrowheads mark parietal endoderm contribution. Green arrowheads mark visceral endoderm. Red arrowheads mark trophoblast. **C.** Quantification of ES cell contribution to embryos based on location. 2i n = 18, 2i/LIF n = 43. **D.** Quantification of the level of contribution of ES cells to the extraembryonic regions of E6.5 embryos. Low = 1-5 cells, medium = 5-10 cells, high = more than 10 cells.

4.7 LIF and JAKi culture of pre-implantation embryos

Finally I wanted to ask whether LIF also promoted an endodermal population during pre-implantation development. Previous reports have suggested that LIF is involved in maintaining the ICM-derived lineages, particularly during diapause (Do et al., 2013; Nichols et al., 2001). However, in depth analysis of lineage specification has not been carried out.

Embryos were cultured from different stages (E0.5, 1.5, 2.5 and 3.5) up until the late blastocyst stage in increasing concentrations of LIF and immunostained for markers of the 3 pre-implantation lineages, NANOG (epiblast), GATA6 (PE) and CDX2 (trophoblast), to determine whether there was an effect on lineage segregation. As previously discussed (Section 2.3.3), embryos cultured *ex vivo* develop more slowly than those *in vivo*, hence although the late blastocyst stage is labelled as its *in vivo* equivalent E4.5, in these experiments embryos were actually cultured until E5.5. Culturing embryos in the presence of LIF from E0.5 until the late blastocyst stage increased the proportion of embryos that had no inner cells, whereas culturing embryos from E1.5 and later promoted an increase in the proportion of $GATA6^+$ ICM cells (Fig. 47). This phenomenon was observed in C57BL/6 and CD1 embryos and hence is not due to genetic background. Data obtained from both mouse strains was therefore combined. No effect on the trophoblast lineage was observed. The most substantial increase in PE was observed when embryos were cultured with LIF from E2.5-4.5 (Fig. 47) hence this time period was analysed in more detail.

Embryos were cultured in 1000U, 2000U or 5000U (1x, 2x and 5x respectively of the standard concentration used to maintain ES cells) of LIF. At all concentrations, LIF increased the proportion of embryos that showed elevated levels of PE (Fig. 48A). *In vivo* experiments are traditionally scored according to the penetrance of the phenotype i.e. the number of abnormal embryos in each condition, as well as by expressivity i.e. the extent of the phenotype in affected embryos. In embryos cultured from E2.5-E4.5, I observed that the phenotypic penetrance of LIF was approximately 40%. Embryos were scored as phenotypically abnormal if the proportion of $GATA6^+$ cells within the ICM was more than one standard deviation outside of that observed in control embryos. Scoring of phenotypic penetrance is shown in Figure 48B. If the proportion of PE cells was higher than in controls, these were categorised as 'GATA6 high'. If the proportion of epiblast cells was higher than in control these were classified as 'NANOG high'. Embryos were also observed that contained only epiblast or only PE cells or no inner cells at all. These were categorised as 'NANOG only' or 'GATA6 only' or 'no inner cells'. Scoring revealed an increase in the number of embryos that showed high levels of PE, from 15% in control conditions to an average of 39% at different doses of LIF (more than 2.5-fold) (Fig. 48B). Although increasing the dose of LIF did not affect the penetrance, it did increase the expressivity as a higher proportion of embryos were observed that showed only GATA6 expression within the ICM (Fig. 48B). At higher concentrations there were also cells that expressed neither

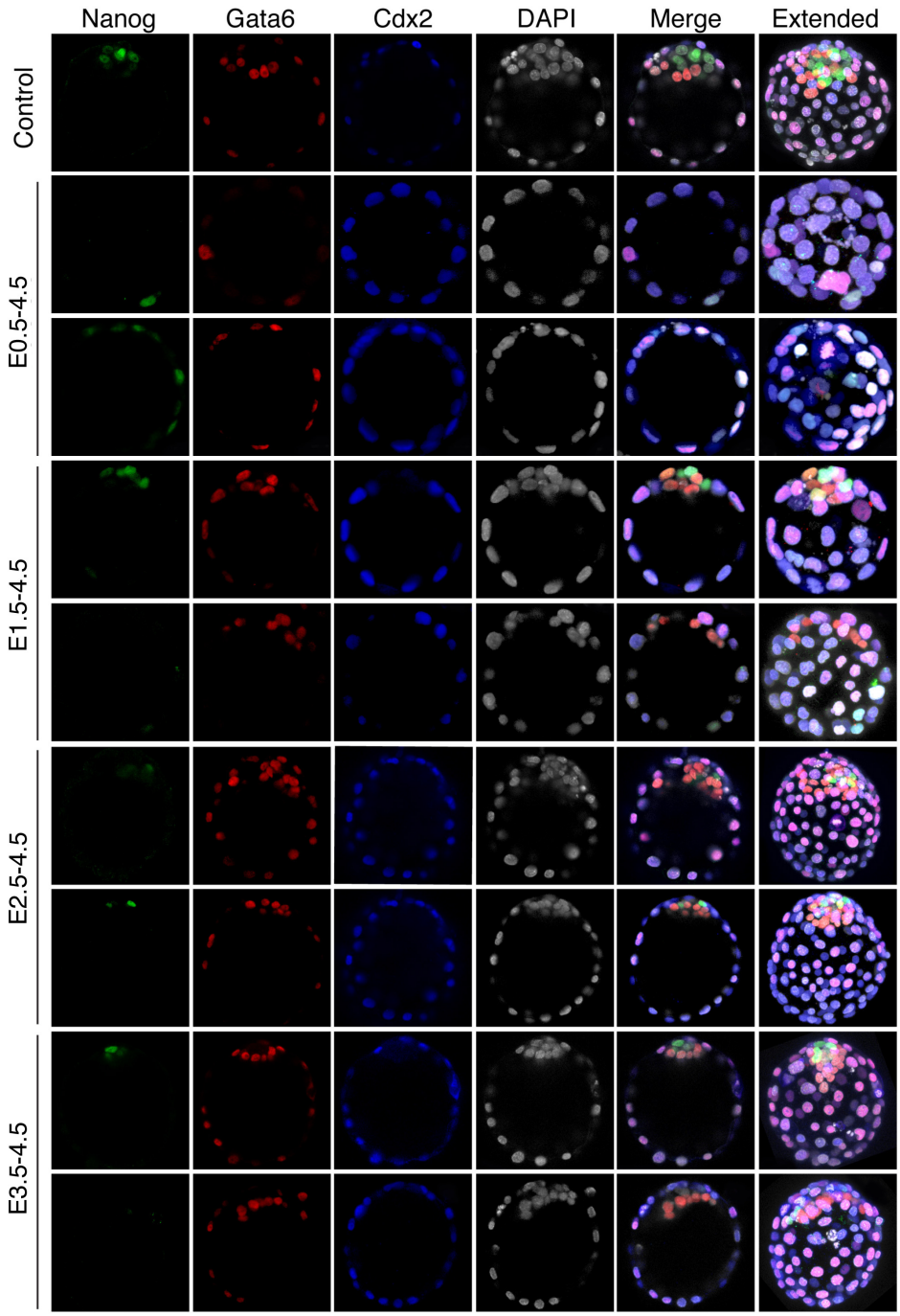


Figure 47. Culturing embryos for different time periods in the presence of LIF affects lineage segregation. Pre-implantation embryos were flushed from oviducts or the uterus of wild type C57BL/6 and CD1 mice and cultured until the late blastocyst stage in KSOM (control) or KSOM + LIF. Embryos were immunostained for the markers of the 3 pre-implantation lineages, NANOG (epiblast), GATA6 (PE) and CDX2 (trophoblast). Images represent confocal optical sections through the ICM, 2 examples are shown per condition.

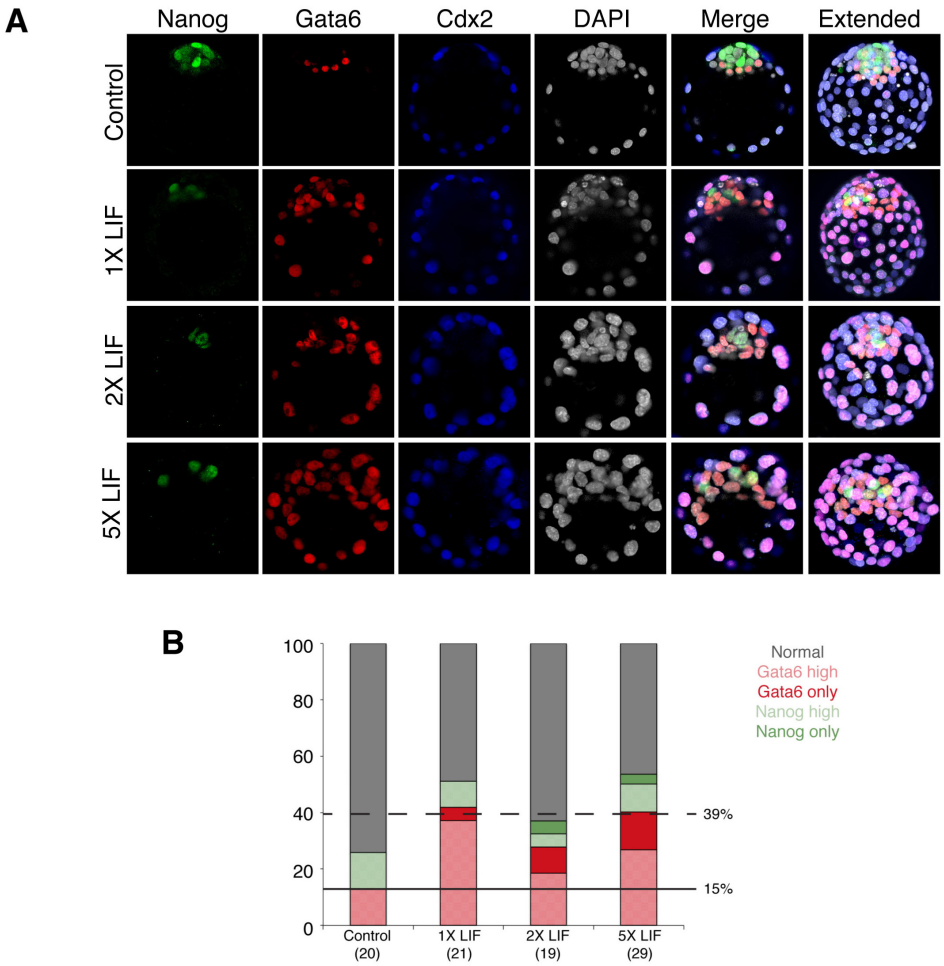


Figure 48. Culturing embryos from E2.5-4.5 in different doses of LIF increases the proportion of embryos that show elevated levels of primitive endoderm (PE). Embryos were cultured from E2.5-E4.5 in KSOM or KSOM with 1000U (1X), 2000U (2X) or 5000U (5X) of LIF and immunostained for the 3 lineage markers NANOG (epiblast), GATA6 (PE) and CDX2 (trophoblast). **A.** Images show confocal optical sections through the ICM as well as an extended image through the whole embryo. **B.** Graph showing categorisation of immunostained embryos. The ICM of embryos was analysed based on the proportion of GATA6 and NANOG-expressing cells and categorised as above. Normal embryos are those with the same proportions of PE and epiblast cells as the average proportions +/- the standard deviation in control embryos. GATA6 high or NANOG high categories correspond to embryos that fell outside of the average control proportions +/- the standard deviation, due to having more GATA6 or NANOG positive cells. The number in brackets under each bar indicates the number of embryos analysed. The black line indicates the proportion of GATA6 high embryos in control conditions. The dashed line indicates the average GATA6 high combined with GATA6 alone embryos across all LIF treatments.

GATA6 nor NANOG. However, as approximately 60% of embryos were considered 'normal', when the total number of cells was scored across all embryos, no significant experimental difference was observed compared to the controls.

As LIF appeared to promote PE through the JAK/STAT pathway (Section 4.2), I asked whether inhibition of JAK/STAT signalling would impact on blastocyst lineage segregation. Using doses defined in ES cells, I cultured embryos from E2.5 until the late blastocyst in the presence of JAKi. Surprisingly, this treatment also caused an increase in the levels of PE (15% in controls vs. 57% of embryos in treated embryos) (Fig. 49A-C). Multiple morphological phenotypes were observed; either the PE and epiblast failed to segregate or the epiblast was lost entirely leaving a PE encapsulated cavity (Fig. 49A). This variety of phenotypes could be due to the fact that JAKi may block both the JAK/STAT and the PI3K pathway (discussed in Section 4.4). Additionally ICM cells were observed that expressed neither NANOG, GATA6 nor CDX2. When total cell numbers were quantified, JAKi treatment resulted in a slight decrease in the number of epiblast cells as well as an increase in the number of PE cells (Fig. 49B). As with increasing LIF doses, increasing the concentration of the inhibitor increased the expressivity, but also the penetrance of the phenotype.

4.8 LIF culture increases the expression of extracellular matrix components in embryonic stem cells and pre-implantation embryos

As I observed that culture of both ES cells and pre-implantation embryos in the presence of LIF promoted a population of extraembryonic endoderm-primed cells, I wanted to investigate the mechanism by which LIF mediated its effect. To do this, I analysed the RNA-seq data previously described (Section 4.1) that compared ES cells cultured in 2i alone to those cultured in 2i/LIF. I carried out functional clustering analysis, using the Database for Annotation, Visualisation and Integrated Discovery (DAVID, NIH) online bioinformatics tool, on genes that were expressed at levels more than 1.5-fold higher in the presence of LIF (FDR 0.05). Functional clustering is similar to gene ontology analysis but groups redundant annotations together to simplify the results. The top 10 results are shown in Table 5. Many of these genes are involved with ECM and cell adhesion, including numerous collagen and laminin genes. Additionally, there was an enrichment of components of the focal adhesion pathway, which can affect proliferation rate by regulating the expression of *CyclinD1*, a gene that was upregulated in response to LIF. I therefore wanted to ask whether culture in the presence of LIF also affected the expression of ECM genes *in vivo*.

I cultured embryos from E2.5 until the late blastocyst stage in the presence of LIF and immunostained them using an antibody against all forms of LAMININ, expressed both in the trophoblast and at the PE-epiblast junction. I observed that the expression of LAMININ was increased upon culture in the presence of LIF (Fig. 50) suggesting that the effect of LIF may be mediated through the ECM.

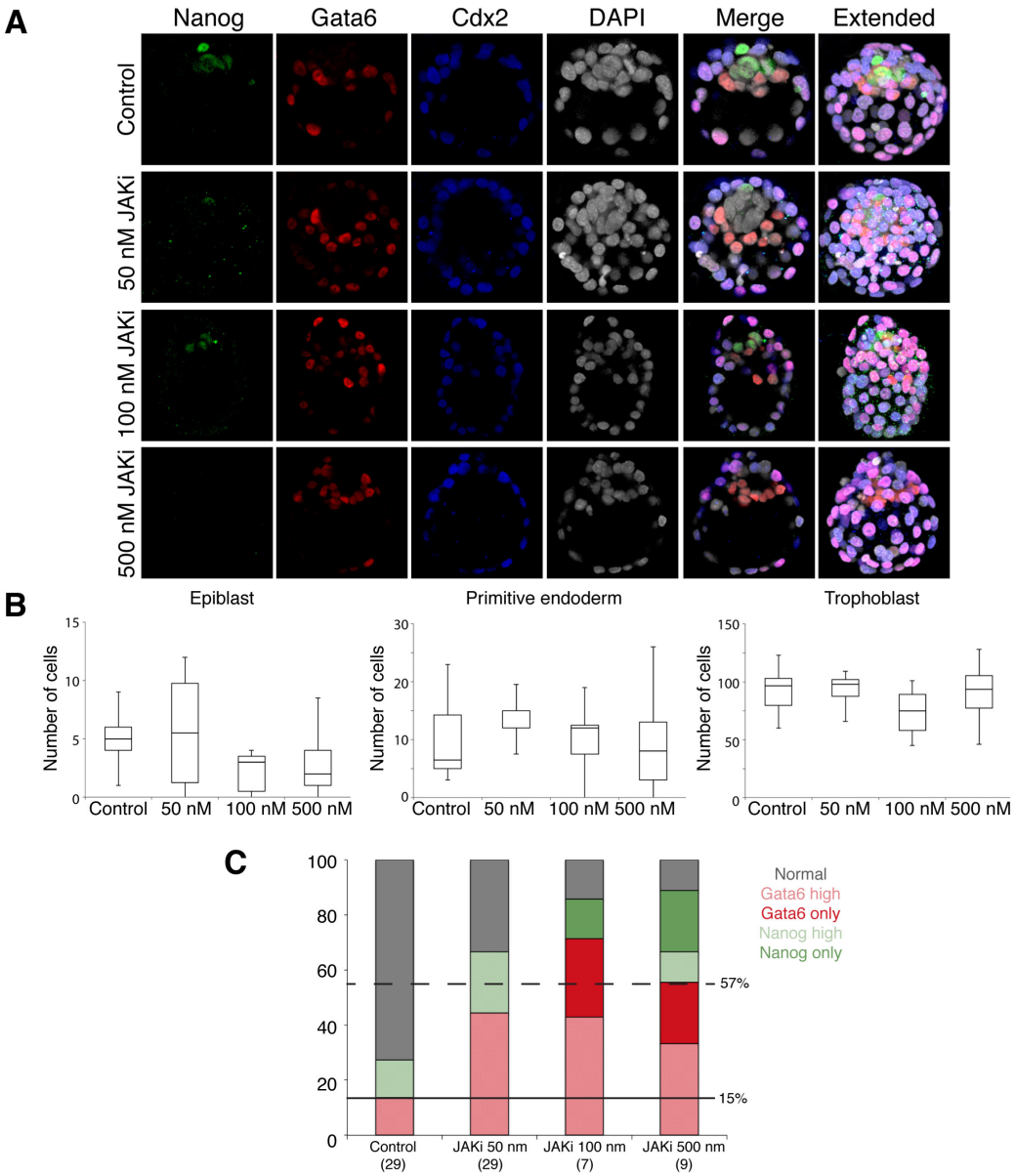


Figure 49. Culturing embryos from E2.5-4.5 in different doses of JAKi increases the proportion of embryos that show elevated levels of primitive endoderm (PE). Embryos were cultured from E2.5-4.5 in KSOM OR KSOM + JAKi and immunostained for the 3 lineage markers NANOG (epiblast), GATA6 (PE) and CDX2 (trophoblast). **A.** Images show confocal optical sections through the ICM as well as an extended image through the whole embryo. **B.** Box plots showing the absolute number of cells in each lineage after treatment with different doses of JAKi. **C.** Graph showing categorisation of immunostained embryos. The ICM of embryos was analysed based on the proportion of GATA6 and NANOG-expressing cells and categorised as above. Normal embryos are those that showed the same proportions of PE and epiblast cells as the average proportions +/- the standard deviation in control embryos. GATA6 high or NANOG high categories correspond to embryos which fell outside of the average control proportions +/- the standard deviation due to having more GATA6 or NANOG positive cells. The number in brackets under each bar indicates the number of embryos analysed. Black lines indicate the proportion of GATA6 high embryos in control conditions. Dashed lines indicate the average GATA6 high, combined with GATA6 alone, embryos across all JAKi treatment doses.

ANNOTATION	ENRICHMENT
Secreted factor	5.7
Epithelial morphogenesis	3.9
Cell adhesion	3.9
ECM	3.7
Lysosome	3.0
Response to external stimulus	2.9
Vesicle	2.8
Focal adhesion	2.8
Placenta development	2.8
Peptidase inhibitor	2.8

Table 5. Functional clustering annotation of genes expressed more highly in ES cells cultured in 2i/LIF than in 2i alone. Genes were identified based on a more than one-fold increase in expression when ES cells were cultured in 2i/LIF compared to 2i alone (FDR 0.05). Functional annotation clustering, using DAVID online bioinformatic analysis resource (NIH), was carried out. The top 10 results are shown for each cluster along with the enrichment score, indicating the enrichment of each functional annotation group over what is observed in the genome as a whole.

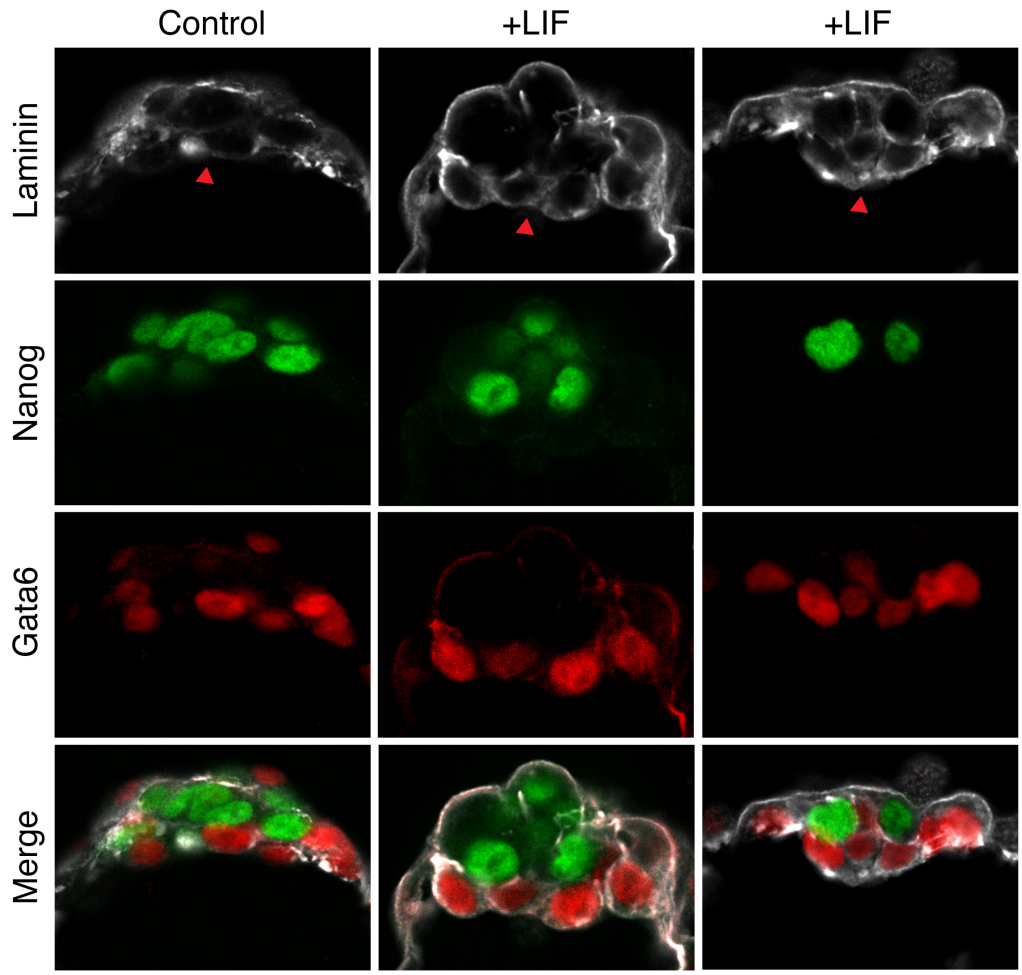


Figure 50. LIF culture increases the expression of LAMININ in pre-implantation embryos. Embryos were cultured from E2.5 until the late blastocyst in KSOM (control) or KSOM+LIF. Embryos were fixed and immunostained for LAMININ, NANOG and GATA6. Images show confocal optical sections through the ICM. Arrowheads mark LAMININ surrounding the primitive endoderm at the blastocoel cavity.

4.9 Discussion

In this chapter we have shown that LIF supports an extraembryonic-primed HV^+ population of ES cells that was previously characterised in Chapter 3 of this thesis. This finding has been added to the model I have developed for ES cell culture (Fig. 51).

Both serum-containing and 2i ES cell culture conditions, are routinely supplemented with LIF (Nichols et al., 2009; Ying et al., 2008) and, while LIF is known to upregulate pluripotency markers (Hall et al., 2009), its role *in vivo* is to support extraembryonic development (Stewart et al., 1992; Takahashi et al., 2003). In embryoid body differentiation, LIF selectively blocks primitive ectoderm differentiation while permitting PE differentiation (Shen and Leder, 1992). Additionally, the downstream effector of LIF, STAT3, binds directly to extraembryonic gene promoters such as *Gata6*, *Gata3*, and *Eomes* (Kidder et al., 2008). Thus, the role of LIF in supporting ES cell culture may be mediated via a totipotent, extraembryonically primed cell type capable of effectively expanding and dynamically generating the heterogeneous distribution of cells normally observed in ES cell culture. This is supported by the observation that a common pathway, the JAK/STAT pathway, promotes the expansion of the HV^+ population and also mediates self-renewal.

Additionally, the culture of ES cells and embryos in the presence of LIF caused an increase in the expression of ECM components and an enrichment of focal adhesion pathway components that can affect proliferation rate by regulating *CyclinD1* expression. It is therefore possible that LIF mediates the expansion of an extraembryonic population through the ECM. LIF is expressed in 2 forms, a soluble form and a matrix-bound form (Rathjen et al., 1990; Smith et al., 1992). The matrix-bound form is predominantly expressed in ES cells and prevents cells from differentiating when overexpressed during early development (Conquet et al., 1992). Thus, as well as increasing the expression of matrix components, LIF may also play a direct role in matrix organisation during development.

It was also observed that downstream targets of LIF were not restricted to the embryonic cells of the pre-implantation embryo, with pSTAT3 in particular showing a strong correlation to the PE lineage. However, KLF4 expression was not as strongly correlated with the expression of pSTAT3 as expected. This is likely to be because KLF4 is a more downstream component of the JAK/STAT pathway and is known, in other cell types, to be regulated by other pathways (Ghaleb et al., 2008).

The analysis of early KLF4 expression also suggested that the pre-implantation embryo is perhaps more complex than originally thought. It is commonly stated that the early ICM is a mix of PE and epiblast precursors (Section 1.2.3.3). However, KLF4 was not expressed in all cells but also showed a degree of overlap with each of these populations. This indicates that the early ICM is comprised of more than just these 2 discrete cell populations.

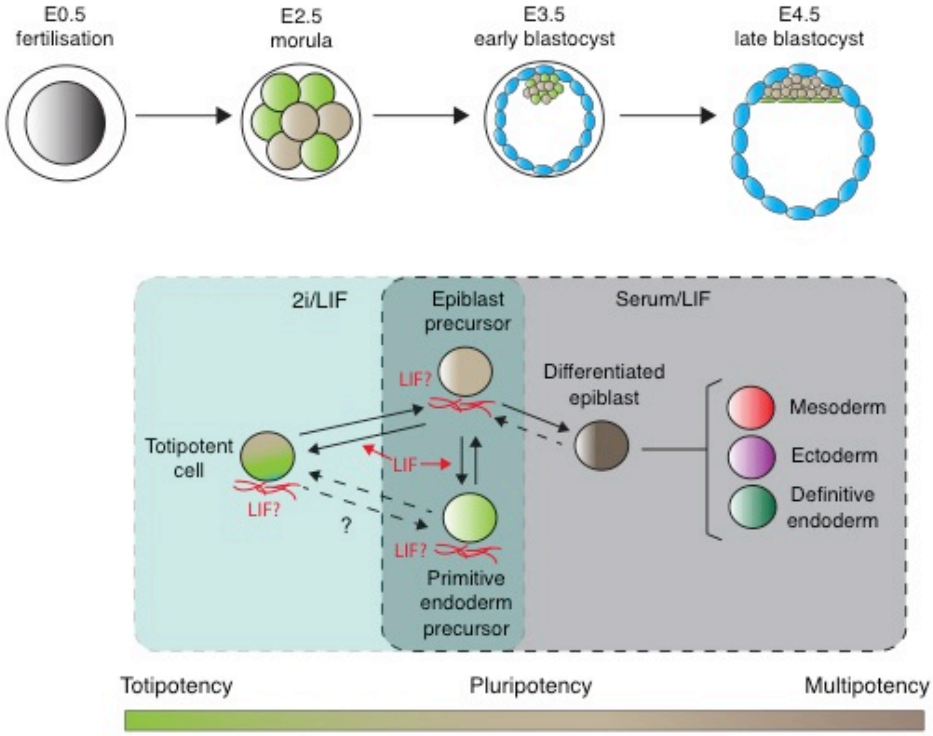


Figure 51. Model summarising the key findings of Chapters 3 and 4. The model relates serum/LIF and 2i ESC cultures to embryonic development. 2i/LIF cultures contain totipotent single cells that are comparable to early pre-implantation stages when embryonic cells are still totipotent. As FGF is needed for later lineage segregation, inhibiting FGF signaling in 2i seems to block embryos and ESCs at a point prior to this commitment event. Serum/LIF ESC cultures are reminiscent of later embryonic stages when cells are pluripotent and contain more restricted cell types. Although partially differentiated mesoderm and ectoderm appear to be lost in 2i/LIF, endoderm is maintained. Based on our data, totipotent ESCs are able to dynamically interconvert to a pluripotent epiblast precursor state and may also exist in equilibrium with a similar primitive endoderm precursor, although we have not been able to distinguish this population in this study. LIF supports the expansion of the totipotent ESC state and primitive endoderm precursors in serum, perhaps by modulating the expression of extracellular matrix factors important for self-renewal and proliferation. The addition of LIF both increases the expression of ECM components such as LAMININ but may also be incorporated into the matrix itself in its matrix-bound form.

CHAPTER 5:

MECHANISMS OF TRANSCRIPTIONAL REGULATION IN EMBRYONIC STEM CELL PRIMING AND DIFFERENTIATION

Chapter 5:

Mechanisms of transcriptional regulation in ES cell priming and differentiation

ES cells are cell lines derived from the ICM of the pre-implantation embryo. Although they are isolated from a specific embryonic stage and region, they are a mixed cell population with subgroups defined based on the heterogeneous expression of particular markers (Canham et al., 2010; Chambers et al., 2007; Hayashi et al., 2008; Singh et al., 2007; Toyooka et al., 2008). These subpopulations are not discrete units but represent unstable states that interconvert over relatively short periods of time. There is also evidence that they signify physiologically relevant cell types primed towards particular paths of differentiation. This so-called 'lineage-priming' has been demonstrated both at the gene expression and functional level (Canham et al., 2010; Chambers et al., 2007; Kobayashi et al., 2009). However, the mechanisms that control such rapid and reversible transcriptional changes, and how these mechanisms may vary as cells move through differentiation towards their fated cell type, have not been analysed.

Transcription can be regulated both at *de novo* initiation, i.e. recruitment and assembly of the pre-initiation complex (PIC) and transcription of 20-50 nucleotides, and also elongation, progression of transcription through the gene body. A fairly common mechanism of initiation regulation, observed at approximately 30% of human promoters, is 'polymerase pausing' where RNA-polymerase II (Pol II) becomes stalled after clearing the promoter region (Bentley, 1995; Core et al., 2008; Rougvie and Lis, 1988). Currently, various functions of polymerase pausing have been proposed including a physical block of nucleosome binding to maintain an open promoter structure (Gilchrist et al., 2010; Gilchrist et al., 2008), an insulator to prevent promiscuous gene activation from inappropriate enhancer elements (Chopra et al., 2009) or a checkpoint (Adelman and Lis, 2012). Although, the role of pausing in transcriptional regulation is not fully understood, in Drosophila this mark is associated with genes that are only transiently repressed, to be turned back on later in differentiation (Saunders et al., 2013). Hence, pausing may poise genes for activation by completing the PIC binding step prior to the point at which activation of the gene is necessary.

Here, I utilised the same sensitive HV reporter system (Canham et al., 2010) as used in previous chapters to isolate a sub-population of ES cells primed towards extraembryonic endoderm as well as cells that represent fully differentiated endoderm. In collaboration with Leighton Core and John Lis (University of Cornell), we carried out genome-wide run-on sequencing (GRO-seq) to determine the location, orientation and density of transcriptionally engaged Pol II in these sorted populations (see Methods) to track changes in transcriptional regulation during lineage-priming and early

differentiation. We showed that the priming event in ES cells is regulated at the point of elongation whereas further differentiation is regulated at *de novo* initiation.

5.1 GRO-seq analysis and clustering of samples

HV ES cells were sorted into HV^- and HV^+, $SSEA\text{-}1^+$ populations i.e. unprimed and primed ES cells (based on the lower and upper 25% of HV expression). SSEA-1 is a marker of undifferentiated ES cells. Additionally, HV ES cells were differentiated by removing LIF from the culture medium for 5 days and the $HV^+SSEA\text{-}1^-$ population was also sorted, representing differentiated endoderm, either the extraembryonic PE or embryonic mesendoderm/primitive streak (Fig. 52A,B). Throughout this chapter, these populations will be referred to as HV^-S^+ (ES cell state), HV^+S^+ (primed PE state) and HV^+S^- (differentiated endoderm state). After sorting, cells were immediately processed to isolate nuclei and cryopreserved for further analysis in collaboration with the Lis lab at Cornell University. Our collaborator, Leighton Core, performed GRO-seq on all of these samples. In this process nuclear run-on assays allow the extension of RNA from already transcriptionally active polymerases, but do not allow novel binding or initiation of inactive Pol II. The RNA generated by the nuclear run-on assays can then be sequenced to determine the location, orientation and density of Pol II across the whole genome (Section 2.5). The sequencing of these samples returned between 2×10^7 and 1×10^8 raw reads depending on the sample. In the paper from the Lis lab originally describing this technique, 2.5×10^7 raw reads were generated (Core et al., 2008), hence a reasonable depth of sequencing was obtained.

A schematic diagram of the bioinformatics analysis carried out can be found in Figure 53. The expression level of genes was first determined according to the density of Pol II within the gene body. To enable analysis of transcriptional states, genes were then classified as significantly up or downregulated between the undifferentiated unprimed (HV^-S^+) and primed (HV^+S^+) populations and differentiated endoderm (HV^+S^-) using the edge R statistical package (FDR 0.01) (Robinson et al., 2010) (see Methods). This generated a list of 2861 upregulated genes and 1363 downregulated genes. Genes within these categories were then clustered according to their level of expression within each population, based on Ward clustering methodology (Ward Jr., 1963) to ask whether genes that showed similar transcriptional changes were regulated by similar mechanisms. This resulted in 5 clusters of upregulated and 5 clusters of downregulated genes (Fig. 54A,B, Supplementary Tables 1,2).

When ES cells are differentiated in the absence of LIF they give rise to both primitive streak-like mesendoderm and extraembryonic primitive endoderm HV^+ populations (Fig. 52A). I asked whether genes associated with these lineage choices showed different expression patterns during ES cell priming and differentiation by assessing their membership in specific clusters. I carried out functional clustering analysis of 9 of these gene groups, using the online DAVID bioinformatics resource (NIH) to determine whether these groups could be separated. Functional clustering is similar to gene

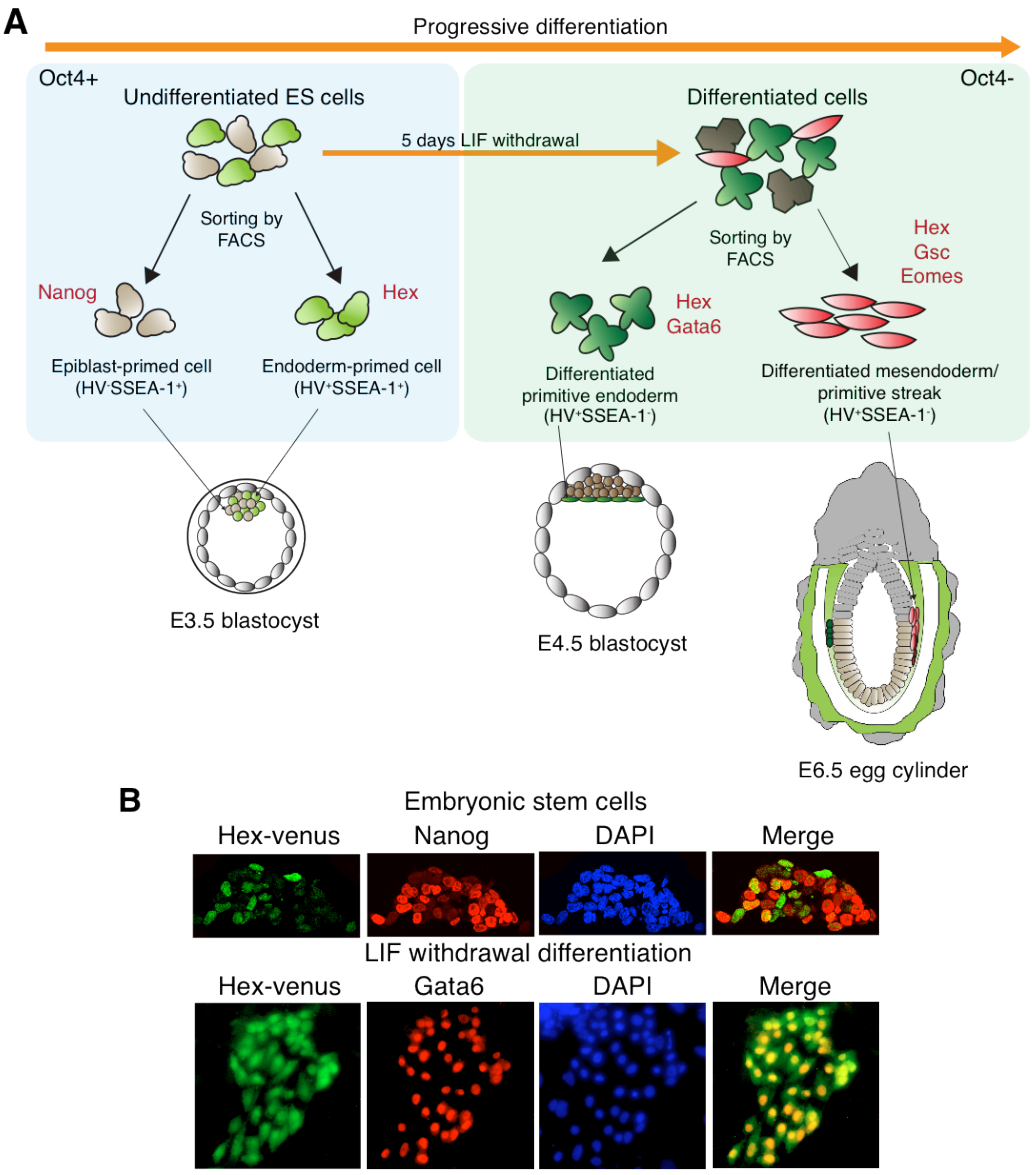


Figure 52. Experimental scheme. A. ES cells containing a Hex-venus (HV) reporter were sorted into HV low (HV^-) and HV (HV^+) fractions by FACS. ES cells were also stained for the marker of undifferentiated ES cells, SSEA-1, to exclude any cells that had undergone spontaneous differentiation. HV^- cells have been shown to represent an epiblast-like state while HV^+ ES cells are primed towards a primitive endoderm fate, similar to precursor populations present in the early mouse blastocyst. ES cells were then differentiated by LIF withdrawal for 5 days, downregulating pluripotency-associated genes such as Oct4 and SSEA-1. $HV^+SSEA\text{-}1^-$ differentiated cells were sorted. These may represent either the differentiated primitive endoderm of the late blastocyst or a mesendoderm/primitive streak-like population from the egg cylinder stage. **B.** Immunostaining showing the 2 ES cell populations, the epiblast-like state marked by Nanog and the primitive endoderm primed state marked by Hex. The lower image shows ES cells after differentiation where Hex-expressing cells also coexpress the endoderm marker GATA6. Although Gata6 transcripts are observed in ES cells, GATA6 protein is not present.

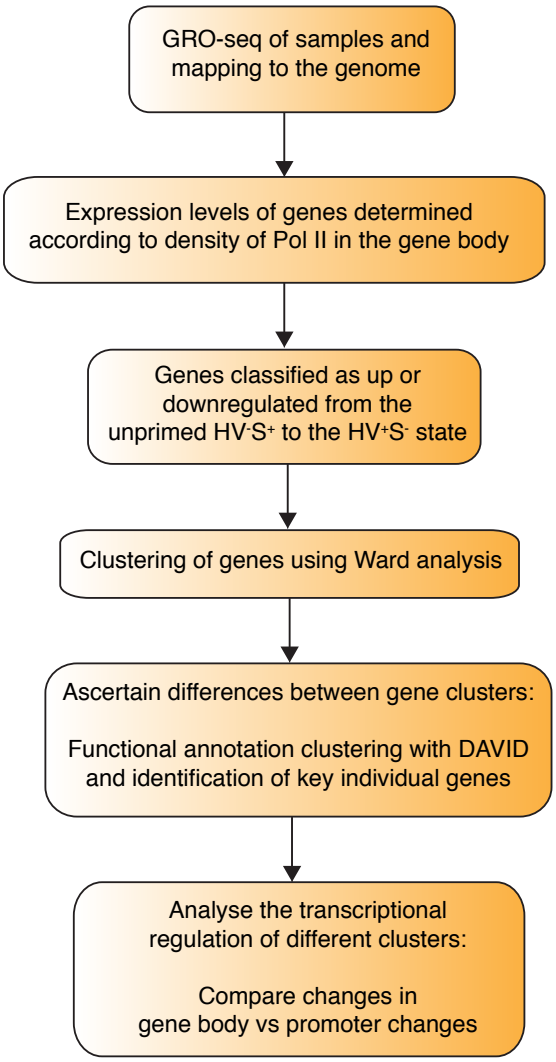


Figure 53. Scheme of bioinformatics analysis of GRO-seq data. GRO-seq data was mapped to the genome and the expression level of each gene determined according to the density of Pol II within the gene body. Genes were then classified as up or downregulated when comparing $HV^{-}S^{+}$ unprimed ES cells to $HV^{+}S^{-}$ differentiated endoderm. Up and downregulated genes were then clustered based on Ward clustering analysis (Section 2.5.4), generating 5 clusters of upregulated and 5 clusters of downregulated genes. The association of gene clusters with distinct developmental functions was then explored by carrying out functional annotation clustering using the online DAVID bioinformatics tool (NIH) as well as individually searching for key developmental regulators within the clusters. Finally, the transcriptional regulation of genes within each cluster was analysing by plotting the changes in Pol II density at the promoter regions and within the gene body.

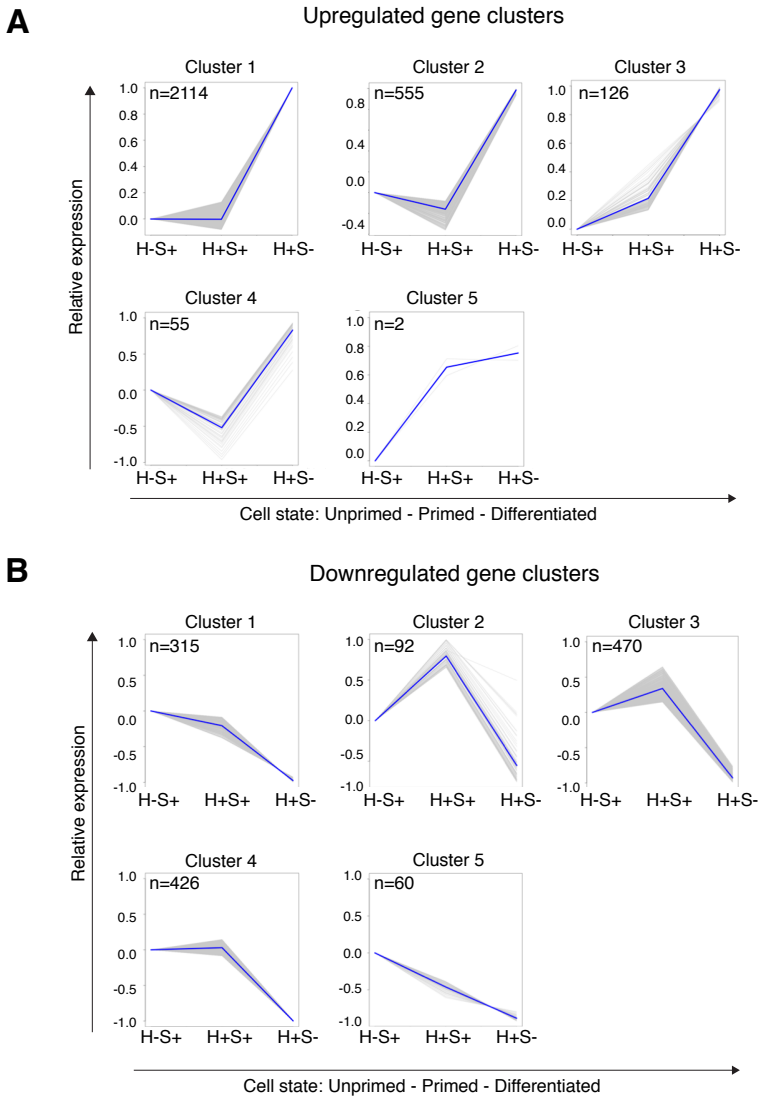


Figure 54. Expression-based gene clustering. Genes that were significantly upregulated or downregulated from the unprimed ES cell population ($HV^{-}S^{+}$) to the differentiated endoderm population ($HV^{+}S^{-}$) were clustered by Ward clustering methodology. Graphs depict the expression of genes in each cell state, unprimed ES cells ($H^{-}S^{+}$), endoderm primed ES cells ($H^{+}S^{+}$) and differentiated endoderm ($H^{+}S^{-}$) within the defined clusters. Grey lines show the expression of individual genes. The blue line indicates the average expression of all genes within the cluster. The number of genes in each cluster is shown in the top left corner of each graph. Clusters are shown for genes that are **A.** upregulated and **B.** downregulated from the unprimed ES cell to differentiated endoderm states.

ontology analysis however redundant annotations are represented together to simplify the results. As upregulated Cluster 5 contained only 2 genes (*Lamin A* and *Hsp70.3*), it was not possible to carry out further analysis on this cluster. The top 5 results for each cluster are shown in Table 6. Upregulated genes in Cluster 1 showed stronger enrichment scores than any other cluster (Table 6). This indicates that these genes represent the most functionally distinct group. Genes within this cluster are strongly associated with adhesion and the ECM, including 17 collagen genes. ECM components are highly expressed in the PE and are also important for the migration of mesendodermal and primitive streak cells in the embryo (Cheng et al., 2013; Hogan et al., 1980; Hogan et al., 1982; Smith and Strickland, 1981) and, therefore it is not surprising that these genes are upregulated during differentiation. Conversely, genes involved in gamete formation are downregulated upon differentiation. This is consistent with the fact that germ cell markers are expressed in ES cells but not in differentiated endodermal cell types (Chu et al., 2011; Zwaka and Thomson, 2005).

I then carried out further analysis by looking at individual genes that were represented within these clusters. Genes upregulated in Cluster 1 were expressed at similar levels between the $HV^{-}S^{+}$ and $HV^{+}S^{+}$ populations but were increased during differentiation, indicating *de novo* transcription (Fig. 54A). This cluster was enriched for mesendodermal/primitive streak genes including *Eomes*, *Brachyury*, *Gsc*, *Mixl1*, *Lhx1*, *Fgf8* and *Nodal*. Conversely, upregulated Cluster 3, comprised of genes that were priming in ES cells (i.e. significantly higher in primed $HV^{+}S^{+}$ than unprimed $HV^{-}S^{+}$ ES cells). These included PE markers such as *Gata6*, *Lrp2*, *Cubn*, *Col4a1*, *Col4a2* and *LmnA*. This suggested that groups of genes, clustered based on their changing expression during differentiation, corresponded to genes associated with distinct cell fates.

All significantly downregulated clusters included pluripotency-associated factors such as *Sox2*, *Esrrb*, *Rex1*, *Pecam-1*, *Fbxo15*, *Tfcp2l1*, *Nr0b1* and *Eras.* As less is known about the expression of these factors *in vivo*, it was difficult to associate individual clusters with distinct developmental roles. However, as identified by functional clustering (Table 6), downregulated Clusters 3 and 4 included specific germ cell markers such as *Prdm14*, *Dazl*, *Piwil2* and *c-Kit*.

Upregulated clusters

CLUSTER 1	
Extracellular matrix	13.7
Adhesion	11.6
Secreted factor	11.0
Tissue morphogenesis	10.0
Vasculature development	8.1
CLUSTER 2	
Vasculature development	1.8
Catalytic activity	1.5
Ion binding	1.4
Oxidation-reduction	1.4
Glycosylation	1.3
CLUSTER 3	
Membrane	2
Regulation of developmental process	1.9
Cardiomyopathy	1.8
Cell death	1.8
Extracellular matrix	1.7
CLUSTER 4	
Embryonic morphogenesis	1.1
Kinase activity	0.8
Cell cycle	0.8
Apoptosis	0.6
Vesicle	0.6

Downregulated clusters

CLUSTER 1	
Ion channel	3.5
Ribosome	2.0
Nucleotide binding	1.9
Ion transport	1.8
Aging	1.7
CLUSTER 2	
Migration	2.0
Tube development	1.6
Adhesion	1.5
Extracellular matrix	1.4
Glycosylation	1.3
CLUSTER 3	
Gamete development	3.1
Metabolic process	2.2
Amino acid metabolism	1.9
Transcription factor activity	1.9
Cytoskeleton	1.9
CLUSTER 4	
Ion binding	3.4
Glycosylation	2.3
Transmembrane protein	1.9
Gamete development	1.6
Translational regulation	1.5
CLUSTER 5	
Lysosome	1.1
Ion binding	1.0
Ion homeostasis	0.9
Apoptosis	0.8
Oxidation-reduction	0.8

Table 6. Functional clustering annotation of up and downregulated gene clusters. Up and downregulated genes were clustered based on the Ward clustering system (Section 2.5.4). 5 clusters were identified for upregulated genes and 5 for downregulated genes. Functional annotation clustering, using DAVID online bioinformatic analysis resource (NIH), was carried out on all clusters apart from cluster 5 of upregulated genes, due to the fact that it contains only 1 gene. The top 5 results are shown for each cluster along with the enrichment score, indicating the enrichment in each functional annotation group over what is observed in the genome as a whole.

5.2 Changes in transcriptional regulation from unprimed to primed embryonic stem cell states

As the expression of genes within clusters changed in the same way during development, we asked whether their transcription was also regulated by common mechanisms. To do this we divided the differentiation process into 2 steps, initial priming (HV^-S^+ vs. HV^+S^+) followed by differentiation (HV^-S^+ and HV^+S^+ vs. HV^+S^-). The regulation of genes was examined by analysing changes in Pol II density both in the promoter region of genes and in the gene body. A 'pausing index' (PI) was also calculated based on the ratio of these 2 values.

In the upregulated clusters, 2 modes of gene expression changes were observed during priming. The expression of genes in Cluster 1 did not change between HV^-S^+ and HV^+S^+ populations (Fig. 54A) therefore no change was observed in the Pol II density at the promoter region or within the gene body. Clusters 3 and 4 both showed a greater change in Pol II density within the gene body than at the promoter resulting in a change in the PI (Fig. 55A,B). Genes within Cluster 2 exhibited changes in Pol II density both within their gene body and promoter regions, although the change within the gene body was more appreciable (Fig. 55A,B). Therefore, genes that were decreased in expression during priming showed an increased PI while those that increased in expression showed a decreased PI.

A similar pattern was observed in the clusters of genes downregulated during priming, where changes within the gene body were larger than those at the promoter. Clusters 2-4, that show an increase in transcription during priming (very slight in Cluster 4) but a decrease in further differentiation, showed an increase in the amount of Pol II in the gene body but not the promoter (Fig. 55C). Additionally Cluster 5 genes, that decrease in expression both in priming and differentiation, show a decrease in the density of Pol II within the gene body during priming (Fig. 55C). Again there was little change in either the gene body or promoter region of Cluster 1.

Taken together, these observations show that genes changing in expression during ES cell priming are predominantly regulated at the point of transcriptional elongation into the gene body rather than initiation.

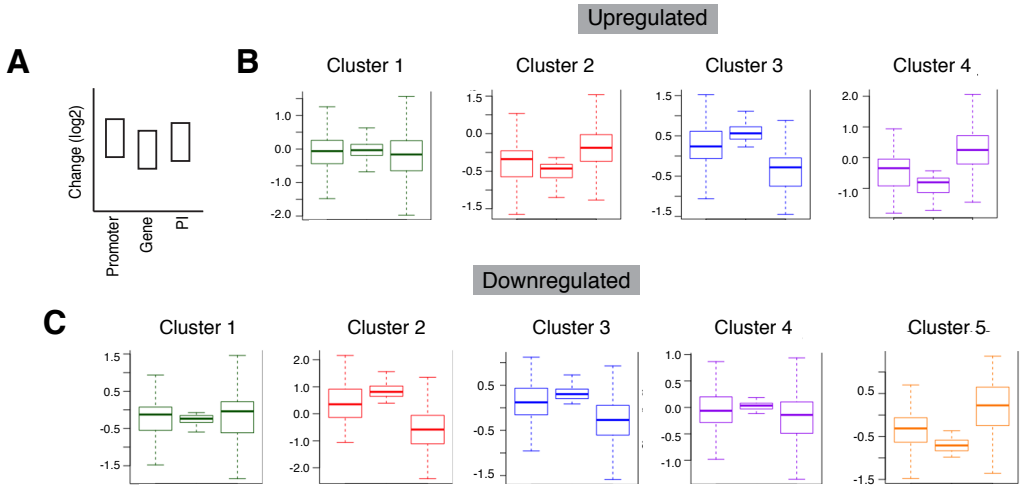


Figure 55. Transcription is regulated mainly by elongation during ES cell priming. Box plots showing the change in engaged RNA-polymerase II (Pol II) occupancy at the gene and promoter regions of specific gene clusters. Genes were clustered based on their pattern of expression between the 3 populations, unprimed ES cells (HV$^-$SSEA-1$^+$), primed ES cells (HV$^+$SSEA-1$^+$) and early differentiated endoderm (HV$^+$SSEA-1$^-$). Analysis of expression level was based on the total amount of Pol II within the gene body. A scheme of the box plot arrangement is shown in **A.** Changes between the unprimed and primed ES cell states are shown. Data is separated into upregulated (**B.**) and downregulated (**C.**) genes.

5.3 Changes in transcriptional regulation from embryonic stem cells to differentiated endoderm

During differentiation towards endoderm, in most clusters, genes that were upregulated showed a larger change in Pol II density at the promoter than in the gene body. This suggests that these genes were predominantly regulated at the initiation step (Fig. 56A,B). Additional effects were also observed on elongation as genes upregulated in Cluster 1 had an equally increased Pol II density at the promoter and within the gene body.

Genes that were downregulated during differentiation showed a significant change in Pol II density within the gene body but almost no change at the promoter. This leads to an increase in the pausing index of genes that are being turned off (Fig. 56A,C) and suggests that lineage specification involves the progressive reduction in transcriptional elongation of ES cell genes prior to the complete shut down of transcription.

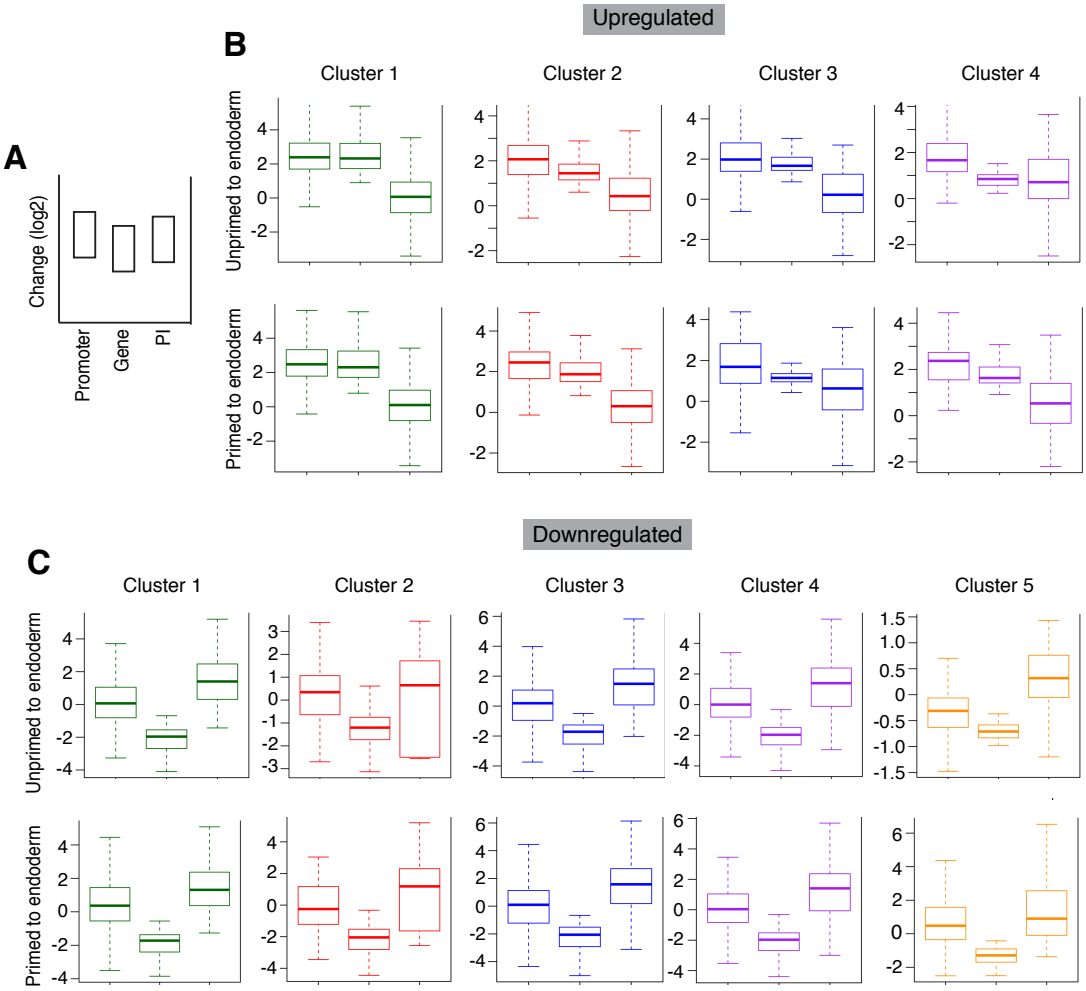


Figure 56. Transcription is regulated mainly by initiation during early differentiation. Box plots showing the change in engaged RNA-polymerase (Pol II) occupancy at the gene and promoter regions of specific gene clusters. Genes were clustered based on their pattern of expression between the 3 populations, unprimed ES cells ($HV^{-}SSEA\text{-}1^{+}$), primed ES cells ($HV^{+}SSEA\text{-}1^{+}$) and early differentiated endoderm ($HV^{+}SSEA\text{-}1^{-}$). Analysis of expression level was based on the total amount of Pol II within the gene body. A scheme of the box plot arrangement is shown in **A.** Changes between the unprimed and endoderm states and the primed and endoderm states are shown. Data is separated into upregulated (**B**) and downregulated (**C**) genes.

5.4 Discussion

In this chapter we showed that clusters of genes exhibiting particular expression patterns during ES cell priming and differentiation were associated with distinct cell fates. Genes that were not associated with priming, i.e. had a similar level of expression between primed and unprimed ES cells, were associated with a differentiated cell type reminiscent of the primitive streak or mesendoderm, while genes involved in priming were those associated with extraembryonic PE. This could be explained by the fact that ES cell cultures represent the ICM of the blastocyst at a point when it is a heterogeneous mixture of epiblast and PE precursors (Section 1.2.3.3). Hence, lineage priming in ES cells is likely to be associated with these precursors rather than priming towards a differentiated cell type that does not arise until later in development.

While I observed transcriptional regulation at the level of initiation during differentiation, the major transcriptional response during the priming stages appeared to be at the level of elongation. Changes in transcriptional elongation were observed in genes both increasing and decreasing in expression during priming, and for genes that were repressed in early differentiation.

It has been hypothesised that polymerase pausing facilitates rapid gene activation by pre-recruiting Pol II to important loci. Hence, when expression needs to be turned on, the polymerase needs only to progress into the gene body rather than being recruited *de novo*. In support of this, in Drosophila, Pol II pausing is present at the promoter regions of many heat shock proteins (HSPs) (Li et al., 1996; O'Brien and Lis, 1991; Rasmussen and Lis, 1993) needed for swiftly relaying information of environmental stress to the cell. Additionally, HSPs feed back onto Pol II pausing (Sawarkar et al., 2012; Teves and Henikoff, 2011). However, it has also been shown that Pol II pausing is important in a developmental context. In Drosophila, Pol II is enriched at the promoter-proximal region of genes that are only transiently repressed and will be needed again later in development (Saunders et al., 2013).

Despite much evidence to suggest that pausing may play a role in critical regulatory and developmental processes, it was also observed in ES cells that Pol II pausing was associated with housekeeping genes, involved in catabolic processes, translation and cell cycle regulation and that unpaused genes were associated with developmental processes and signalling (Min et al., 2011). Consistent with this, ES cells cultured in 2i, had increased Pol II pausing at Myc target genes including many cell cycle regulators (Marks et al., 2012). However, there is also a general increase in pausing in genes that are upregulated in 2i-cultured compared to serum-cultured ES cells (Marks et al., 2012). Thus, although there are now many datasets analysing pausing in different contexts, the meaning of this phenomenon, and whether it is a conserved mechanism of transcriptional regulation in different organisms, is still not understood. However, my observations that elongation is inhibited as a

first step in turning off gene expression suggests that apparent pausing in some contexts may represent an intermediate, but reversible state in the repression of the ES cell programme during differentiation.

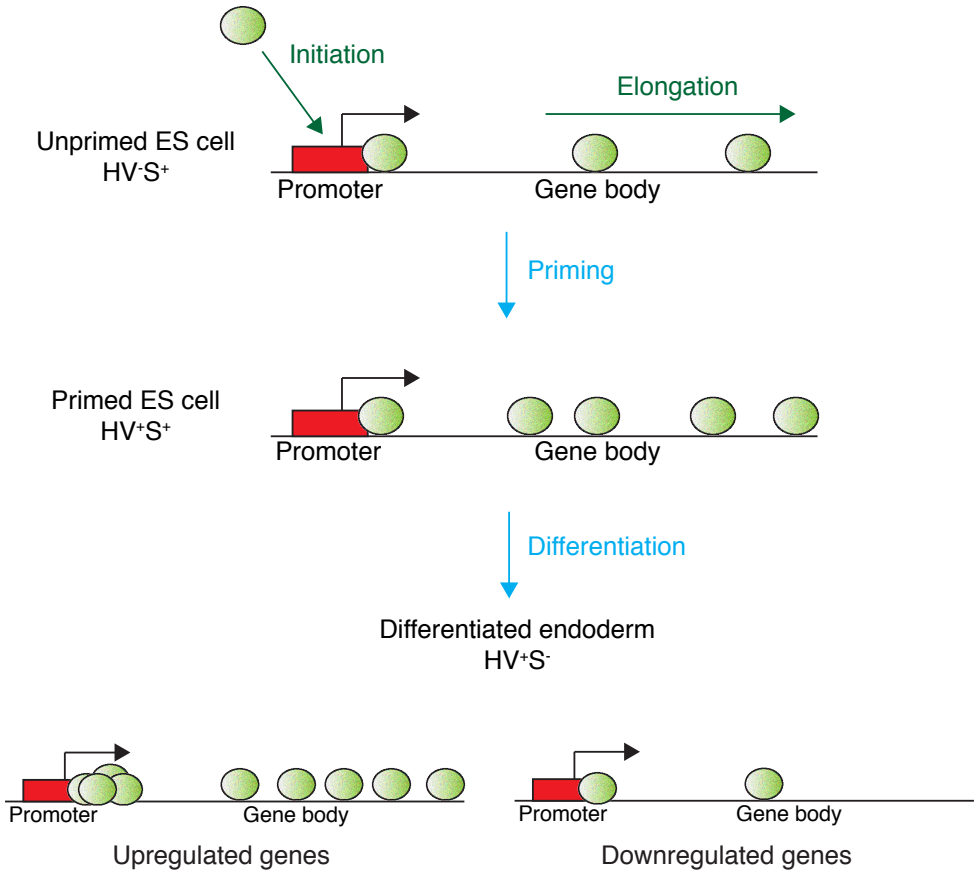


Figure 57. Schematic representation of transcriptional regulation during ES cell priming and endoderm differentiation. To activate gene expression, Pol II (green circle) is recruited to the promoter region (red box) of a gene and subsequently initiates transcription. After transcribing 20-50 nucleotides, Pol II may become paused at a promoter-proximal position, or continue transcription of the gene body in a process known as elongation. During the priming event in ES cells, a change in the density of Pol II was observed within the gene body but not within the promoter indicating that transcriptional changes are regulated at the point of elongation. Coversely, during differentiation from an ES cell state to an endoderm state, although changes were observed in the density of polymerase within the gene body, in upregulated genes, the change of Pol II density at the promoter-proximal region was more significant. Hence, these upregulated genes are regulated at the point of initiation during differentiation. Downregulated genes conversely had no change in Pol II density at the promoter, but a significant decrease in Pol II density within the gene body. Hence these genes are also regulated at the point of elongation.

CHAPTER 6: FINAL DISCUSSION

Chapter 6: Final discussion

The results in this thesis have shown that the nature of ES cells, including their gene expression profile and functional potential, can be significantly altered by the culture conditions in which they are maintained. By analysing the nature of ES cells in these conditions in further detail, we may determine if there is an *in vitro* state that is truly representative of the *in vivo* pre-implantation embryo. Such information is also useful for more technical purposes such as selecting which condition can give rise to the highest contribution chimaeras for mouse line generation.

Growing ES cells in defined serum-free medium, with inhibition of FGF signalling and GSK3 (2i), yielded populations that coexpressed both embryonic and extraembryonic markers and contributed to embryonic and extraembryonic lineages *in vivo* (Chapter 3). As FGF signalling seems to be important in the specification or maintenance of all 3 early lineages of the pre-implantation embryo, this increased plasticity of 2i-cultured ES cells may be due to the fact that, without FGF lineage segregation is not possible and hence cells remain in an early uncommitted, perhaps totipotent, state. A model of these findings is presented in Figure 58.

The notion of totipotent ES cells, able to contribute both to embryonic and extraembryonic lineages, has been examined briefly in the past, but these experimental findings have been largely ignored as such cells seemed to be rare. Furthermore, as discussed in Sections 1.3.2 and 1.3.7, the techniques that we use to assess ES cell potency, appear to bias the outcome. For example, ES cells can contribute more efficiently to embryos in morula aggregations than in blastocyst injections (Lallemand and Brulet, 1990). This is also an issue in adult stem cells where potency is often assessed based on injury models that are likely to employ self-renewal mechanisms distinct from those occurring during development or tissue homeostasis. For example, in the gut, the generally accepted model was that intestinal stem cells divided asymmetrically to generate one daughter stem cell and one differentiated cell type (Morrison and Kimble, 2006; Quyn et al., 2010). However, when the system was modelled in a standard tissue homeostasis context, it was observed that stem cells could divide both asymmetrically and symmetrically, to produce 2 stem cells or 2 differentiated cells, and that these divisions occurred at equal frequency (de Navascues et al., 2012; Klein and Simons, 2011). Consequently, our definition of cell potency and understanding of the mechanisms of self-renewal is limited by our experimental design.

Therefore, as well as empirically assessing cell potency, we should try to understand the phenotype of totipotent cells. It has been shown that chromatin state corresponds with the functional potential of cells. In particular, a reduction in DNA methylation is important for attaining a totipotent/pluripotent state. Cell types that are able to self-renew and show a wide differentiation capacity exhibit DNA hypomethylation, including ES cells, cancer cells and PGCs (Gama-Sosa et al., 1983; Grabole et al.,

2013). The PGC-associated gene, *Prdm14*, maintains a pluripotent state by inhibiting FGF signalling and repressing DNA methylation (Grabole et al., 2013). Additionally, in 2i-cultured ES cells, the repressive chromatin mark H3K27me3 is reduced (Marks et al., 2012), coincident with a global reduction in DNA methylation (Ficz et al., 2013; Leitch et al., 2013; Marks et al., 2012). It has recently been shown that DNA methylation regulates the binding of Polycomb complexes to DNA. Genomic regions with high levels of DNA methylation show low levels of histone methylation marks (Reddington et al., 2013). In ES cells, Polycomb complexes catalyse the addition of the repressive H3K27me3 histone modification to the promoters of developmentally important genes (Boyer et al., 2006). In Polycomb mutant cells, the expression of these genes is elevated, hence it is thought that Polycomb complexes maintain a pluripotent state by supressing differentiation-associated gene expression (Boyer et al., 2006). In globally hypomethylated cells, histone methylation is diluted at the promoters of genes as Polycomb complexes can now also bind to, and methylate, intragenic regions (Reddington et al., 2013). As in the promoter region, intragenic H3K27me3 also represses gene expression (Reddington et al., 2013). Therefore, global hypomethylation is associated with a reduction in the expression of differentiation-associated genes and therefore could be an indicator of an elevated functional potency. Consistently, a hypomethylated state is present during the early stages of embryonic development before lineage restrictions occur (Smith et al., 2012).

Another characteristic that may define a totipotent cell is an elevated rate of proliferation. It is possible that the difference between a pluripotent and totipotent cell, as assessed by chimaera contribution, could be attributed to how many cell divisions it can undergo before differentiating. A 'totipotent' cell could be a single cell that, for example, has shorter cell divisions hence could produce a greater number of progeny before responding to differentiation-inducing signals. The progeny of this cell would therefore colonise the embryo to a greater extent than a cell with slower divisions. Additionally, it is known that the rate of cell cycle divisions decreases after the blastocyst stage when cells begin to become more specified and lose their functional potential (Molls et al., 1983). When ES cells were cultured in 2i, they generated chimaeras more efficiently in single cell injections (Appendix). This was concurrent with an increased rate of proliferation of these cells when reintroduced into the embryo.

A totipotent cell could also be more resistant to differentiation-promoting signals. As with an increased proliferation rate, cells that can self-renew for a longer period of time before differentiating could also be able to efficiently colonise the embryo. The pre-culture of ES cells in 2i prior to injection into the embryo could result in them being more refractory to differentiation-promoting FGF signalling than ES cells cultured in standard serum conditions. An increase in the resistance to differentiation may also be due to the ability of a cell to make its own niche. In Chapter 4 of this thesis, I have shown that LIF is able to promote a totipotent cell type. At the same time, LIF promotes an increase in ECM components such as laminin and collagen that could act as a niche to maintain ES cells in an undifferentiated state.

Ultimately, the idea of a ‘totipotent’ cell remains a confusing one and it has been suggested that the functional capacity of a single cell to generate all cell types of the body does not necessarily define totipotency (Condic, 2014). Instead, it has been proposed that a cell can only be described as totipotent if it can be proven to form an entire, structurally defined, organism in a developmental context (Condic, 2014). However, regardless of the definition of totipotency, we must strive to understand the molecular basis of cell potency as well as the functional outcome.

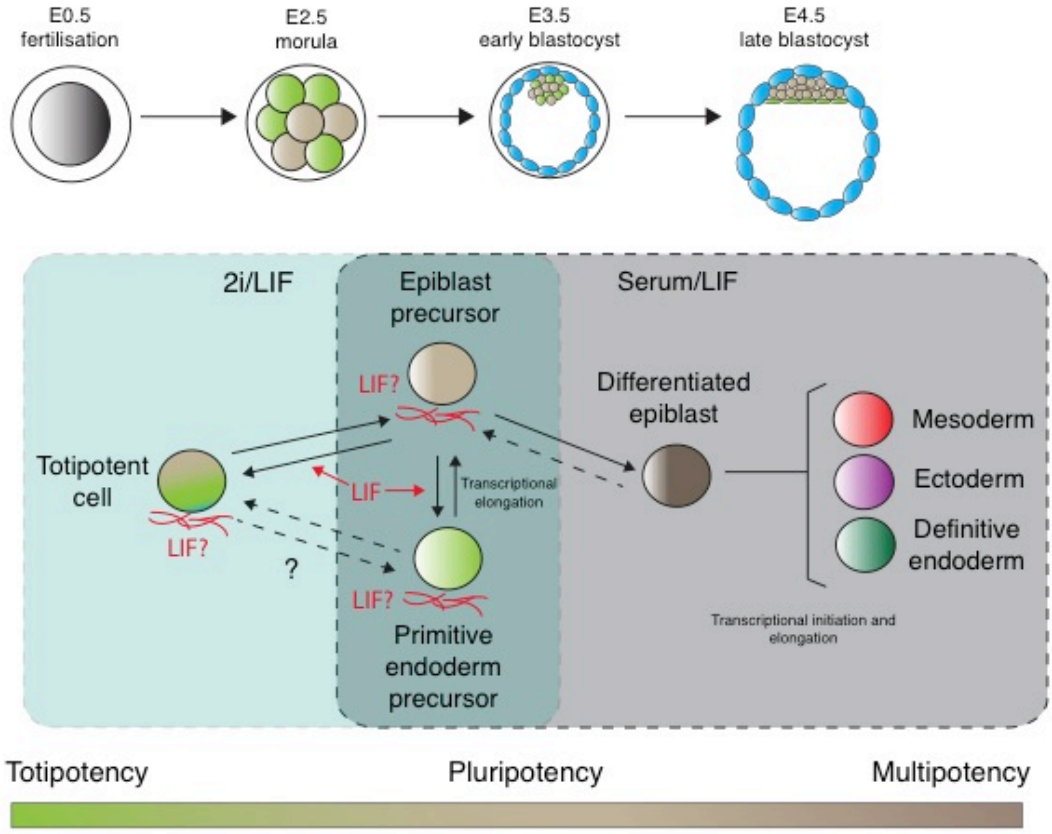


Figure 58. Model summarising the key findings of Chapters 3 and 4. The model relates serum/LIF and 2i ESC cultures to embryonic development. 2i/LIF cultures contain totipotent single cells that are comparable to early pre-implantation stages when embryonic cells are still totipotent. As FGF is needed for later lineage segregation, inhibiting FGF signaling in 2i seems to block embryos and ESCs at a point prior to this commitment event. Serum/LIF ESC cultures are reminiscent of later embryonic stages when cells are pluripotent and contain more restricted cell types. Although partially differentiated mesoderm and ectoderm appear to be lost in 2i/LIF, endoderm is maintained. Based on our data, totipotent ESCs are able to dynamically interconvert to a pluripotent epiblast precursor state and may also exist in equilibrium with a similar primitive endoderm precursor, although we have not been able to distinguish this population in this study. In serum/LIF cultures, ES cell conversion from a HV^{-} to a HV^{+} state is transcriptionally regulated at the point of elongation. During further differentiation of these cells, transcriptional changes are regulated both at the point of elongation and initiation. LIF supports the expansion of the totipotent ESC state and primitive endoderm precursors in serum, perhaps by modulating the expression of extracellular matrix factors important for self-renewal and proliferation. The addition of LIF both increases the expression of ECM components such as LAMININ but may also be incorporated into the matrix itself in its matrix-bound form.

CHAPTER 7: APPENDIX

Generating monoclonal 100% embryonic stem cell-derived mice

Chapter 7: Appendix

Generating monoclonal 100% embryonic stem cell-derived mice

Embryonic stem (ES) cells are cell lines derived from the pre-implantation mammalian embryo (Evans and Kaufman, 1981). After prolonged periods of time in culture, they retain the capacity to participate in embryonic development when introduced into morulae or blastocysts in chimaera assays (Evans and Kaufman, 1981; Fleming, 1987). Initially ES cells were grown on feeders, but now are routinely cultured on gelatin in medium containing serum and the cytokine LIF. While these conditions permit self-renewal, it is unknown how the many and variable components of serum affect the nature of ES cells. During ES cell derivation in serum, PGC-associated genes are turned on although this does not occur during epiblast formation *in vivo* (Chu et al., 2011; Tang et al., 2010). Hence it is uncertain if ES cells in these conditions are still representative of the embryonic stage from which they were derived.

More recently, a novel serum-free culture system has been developed that facilitates ES cell self-renewal under defined conditions. These conditions, referred to as '2i', consist of a MEK inhibitor to block FGF-mediated differentiation signals, along with a GSK3 inhibitor to stimulate the expression of pluripotency factors through the Wnt pathway (Antczak and Van Blerkom, 1997; Ying et al., 2008). ES cells cultured in 2i are more homogeneous morphologically than those in serum and transcriptional analysis has shown that they express lower levels of differentiation-associated factors (Antczak and Van Blerkom, 1997; Dietrich and Hiiragi, 2007). Additionally, they do not transit through a germ cell state during derivation.

Although much *in vitro* characterisation has been carried out on 2i-cultured ES cells, how derivation and culture in 2i affects contribution to the embryo relative to ES cells cultured in standard serum conditions is not entirely understood. Such experiments may allude to whether these ES cells still correlate to the embryo itself. In Chapter 3 I showed that a population of cells in 2i coexpress both embryonic and extraembryonic markers and can contribute both to embryonic and extraembryonic tissues in chimaeras assays. It has therefore been suggested that they represent an early stage of development prior to lineage commitments.

Here, we show that ES cells cultured in 2i can contribute to a stage of development earlier than previously thought to be possible. Single 2i-cultured ES cells introduced into 2-cell embryos, efficiently generated 100% ES cell-derived mice. This was also observed for ES cells cultured in knockout serum replacement (KOSR) medium, but not those cultured in serum suggesting that serum culture negatively impacts the functional potential of ES cells. However, derivation of ES cells in serum did not irreversibly affect their functional capacity. ES cells, derived and cultured in serum for a prolonged period of time, contributed to 2-cell embryos after being switched to 2i of KOSR culture

for 3 passages. Single ES cells from serum cultures were also not able to contribute to morulae. However, in multiple cell injections, they contributed with equal efficiency to morulae as 2i cultured cells. This suggests that 2i culture increases the clonogenicity of ES cells *in vivo*. This work was carried out in collaboration with Dr. Javier Martin Gonzalez and Kasper Bonderup from the University of Copenhagen TCF.

7.1 Embryonic stem cells can contribute to 2-cell embryos

In previous chapters it was observed that HV embryos and ES cells cultured in 2i seemed to be maintained in a state reminiscent of an early stage of development prior to lineage commitment (Section 3.1). I therefore tested the ability of ES cells cultured in 2i to contribute to earlier embryonic stages in chimaera assays. Typically chimaeras are generated by introducing ES cells into either morula or blastocyst stage embryos, but there is no reason that they could not be generated using earlier stage embryos. Here, single unsorted H2B-Tomato ES cells, cultured in serum/LIF or 2i/LIF, were injected into 2-cell stage embryos (E1.5) (Fig. 59A). Embryos were then cultured *ex vivo* until the late blastocyst stage (E4.5) and the ES cell contribution assessed.

ES cells cultured in 2i/LIF generated a higher percentage of chimaeras compared to ES cells cultured in serum/LIF (Fig. 59B 50% vs. 18% and 58% vs. 38% in two separate experiments). However, no chimaeras generated from serum ES cells developed until term. At this early stage there was also a difference in the level of contribution of ES cells to embryos that was mediated by a more rapid expansion of 2i/LIF cultured ES cells (Fig. 59C).

As ES cells are being introduced into embryos at a much earlier stage than in standard chimaera assays, it may be expected that they would now be able to contribute more efficiently to both embryonic and extraembryonic tissues, perhaps due to the increased time period that they spend in the embryo or due to the embryonic environment that they are in. However, although ES cells could be found both in embryonic and extraembryonic tissues, in the majority of cases they only contributed to the epiblast lineage, even after culture in 2i/LIF (Fig. 59D,E). In a large percentage of chimaeras generated from injecting a single ES cell into 2-cell embryos (78% from ES cells cultured in 2i/LIF and 50% from ES cells cultured in serum/LIF), all NANOG positive cells of the ICM were generated from the ES cells introduced and not from the host embryo (Fig. 59E).

Chimaeras were also generated using a distinct hybrid F1 ES cell line and transferred to recipient females to determine whether they were viable. Mice developed to term with the majority of those that showed ES cell contribution, being 100% ES cell-derived (Fig. 59F).

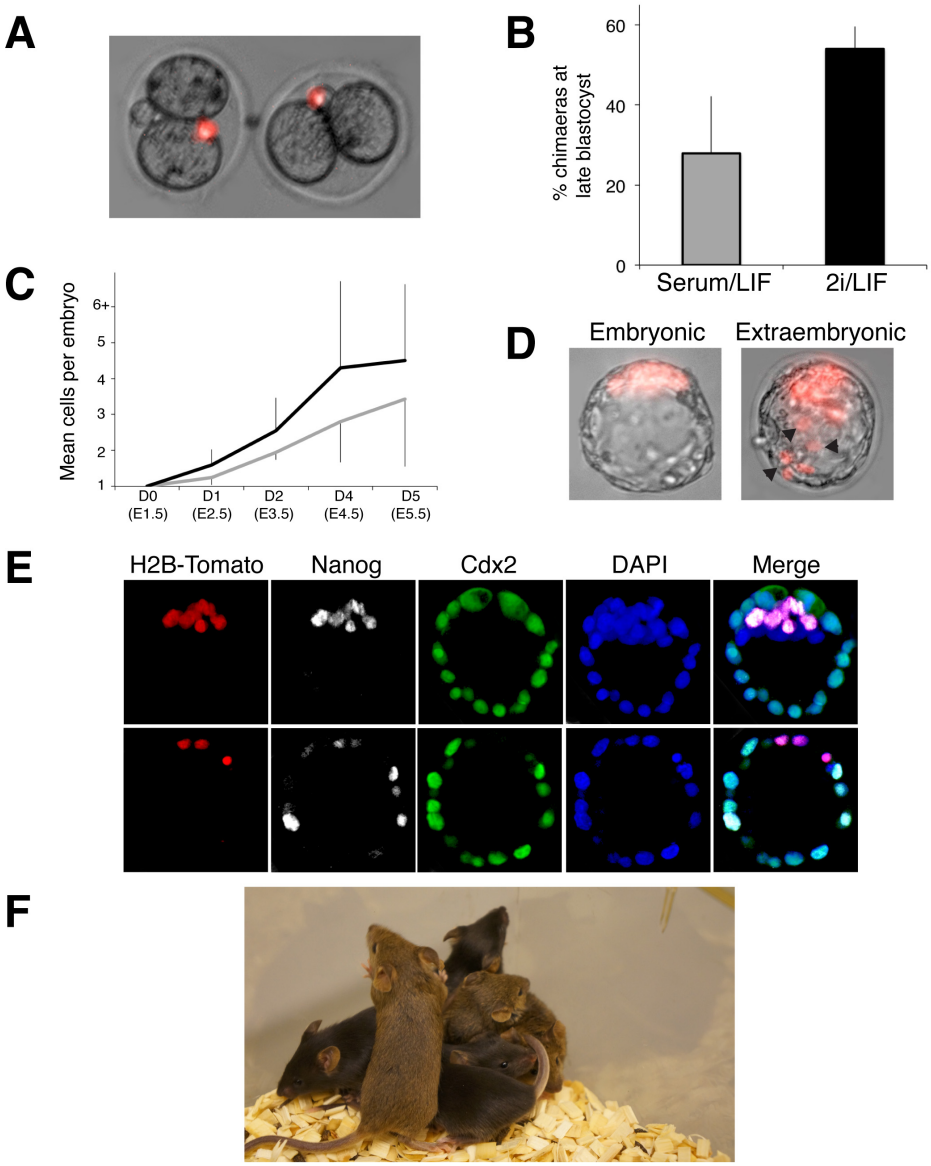


Figure 59. ES cells injected into 2-cell embryos can generate chimaeras. Single ES cells, previously cultured in serum/LIF or 2i/LIF were injected into 2-cell embryos and cultured ex vivo in KSOM for 3 days until the late blastocyst. ES cells contained a constitutive H2B-Tomato marker. **A.** Representative 2-cell stage embryos immediately after injection with a single ES cell. **B.** Quantification of the percentage of chimaeric embryos generated (serum/LIF n = 65, 2i/LIF n = 41). Error bars indicate standard deviation of the mean of 2 biological replicates. **C.** Graph showing the average number of ES cell-derived cells per embryo at each day of culture. At later stages, it was not always possible to distinguish the number of individual cells by standard microscopy hence at these time points embryos were scored in a category of 6 or more cells (6+). **D.** Representative images of chimaeric embryos showing ES cell contribution to the embryonic epiblast alone or both the embryonic and extraembryonic lineages. Arrowheads indicate ES cells contributing to the extraembryonic regions. **E.** Confocal images showing Z-planes of immunostained late blastocysts resulting from embryos injected with 2i-cultured ES cells, marked by H2B-Tomato at the 2 cell stage. **F.** Image of 100% ES-derived agouti mice generated from the injection of a single 2i-cultured cell into a 2-cell embryo, along wild type C57BL/6 littermates.

7.2 Effect of embryonic stem cell derivation protocol on embryo contribution

As the conditions that ES cells were cultured in affected their ability to contribute to embryos in chimaera assays, we also wanted to ask whether the derivation method had an impact. F1 hybrid ES cell lines were derived either in serum/LIF, 2i/LIF or KOSR/LIF culture conditions and cultured either in the same medium as they were derived or switched to the opposing culture condition (either serum/LIF or 2i/LIF). Single ES cells from each condition were injected into 2-cell stage embryos and also into morulae as a control. Embryos were transferred to recipients and allowed to develop until term, where the percentage chimaerism was assessed by coat colour.

ES cells that were derived in serum/LIF and cultured in serum/LIF did not contribute to embryos when single cells were introduced into either morulae or 2-cell stage embryos (Fig. 60A,B). After at least 12 passages in serum, ES cells derived in serum were transferred to 2i for 3 passages. These ES cells were now able to generate chimaeras from single cell injections (Fig. 60A-D). 2i derived ES cells were also able to generate chimaeras when cultured in 2i (Fig. 60A,B). Using 2 different ES cell lines derived in 2i, cells underwent a crisis upon transfer to serum-containing medium, either by an immediate change or else through gradual adaptation. These cells could no longer contribute to embryos as single cells. This data suggests that it is primarily the culture medium and not the derivation conditions that affect the functional potential of ES cells to contribute to embryos.

7.3 2i culture increases the clonogenicity of embryonic stem cells in embryos

As serum cultured ES cells were not able to contribute even to morula stage embryos when injected as single cells, we also injected multiple cells either into morulae or blastocysts as a control experiment. We observed that ES cells from all conditions then generated chimaeras with high efficiency (Fig. 61A-C). Hence, serum and 2i cultured ES cells do not differ in their ability to contribute to chimaeras in standard injection procedures (i.e. multiple cells into morulae) but only when single cells are injected. One explanation could be that a higher percentage of ES cells are competent to contribute to chimaeras when cultured in 2i compared to those cultured in serum. It has previously been shown that, in serum ES cell cultures, approximately 2/10 ES cells have the capacity to contribute to the embryo (Wang and Jaenisch, 2004). Hence, if this was the only influencing factor, our sampling number (n=11 in 2 cell injections and n=13 in morula injections) should have generated some chimaeras. A more likely possibility is that ES cells cultured in serum culture show a reduced clonogenicity compared to 2i-cultured cells when injected into embryos.

The number of chimaeras generated in each experiment is shown in Table 7.

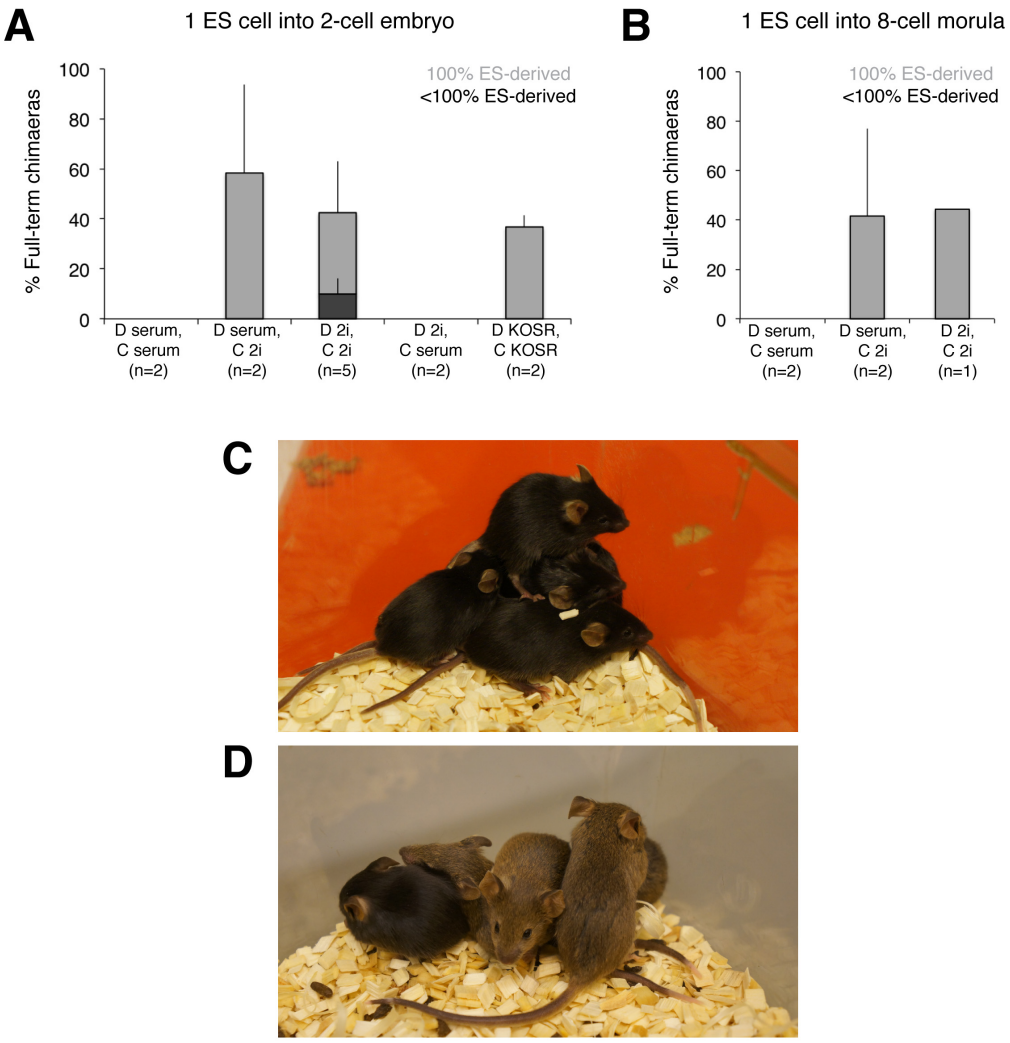


Figure 60. Culture in serum-containing medium inhibits the capactiy of single ES cells to contribute to the embryo. ES cells, were derived in and cultured under various conditions; derived (D) in serum and cultured (C) in serum, D in serum and C in 2i, D in 2i and C in 2i, D in 2i and C in serum and D and C in knockout serum replacement medium (KOSR). **A.** Single ES cells from each of these conditions were injected into 2-cell embryos and their contribution to full-term chimaeras assessed according to coat colour. **B.** Single ES cells from each of these conditions were injected into 8-cell morulae and their contribution to full-term chimaeras assessed according to coat colour. The majority of chimaeras were 100% ES-derived (light grey) although some had less than 100% ES contribution (dark grey). n values refer to the number of independently derived cell lines that were tested for each condition. Error bars indicate standard deviation of the mean between biological replicates of different cell lines. **C.** Representative image of full-term mice generated from the injection of single cell ES cells, D and C in serum, into 2-cell embryos. Only wild type C57BL/6 mice were born. **D.** Representative image of full-term mice generated from single cell injections into 2-cell embryos of ES cells D in serum and C in 2i, generating 100% ES-derived agouti mice, shown alongside wild type C57BL/6 littermate.

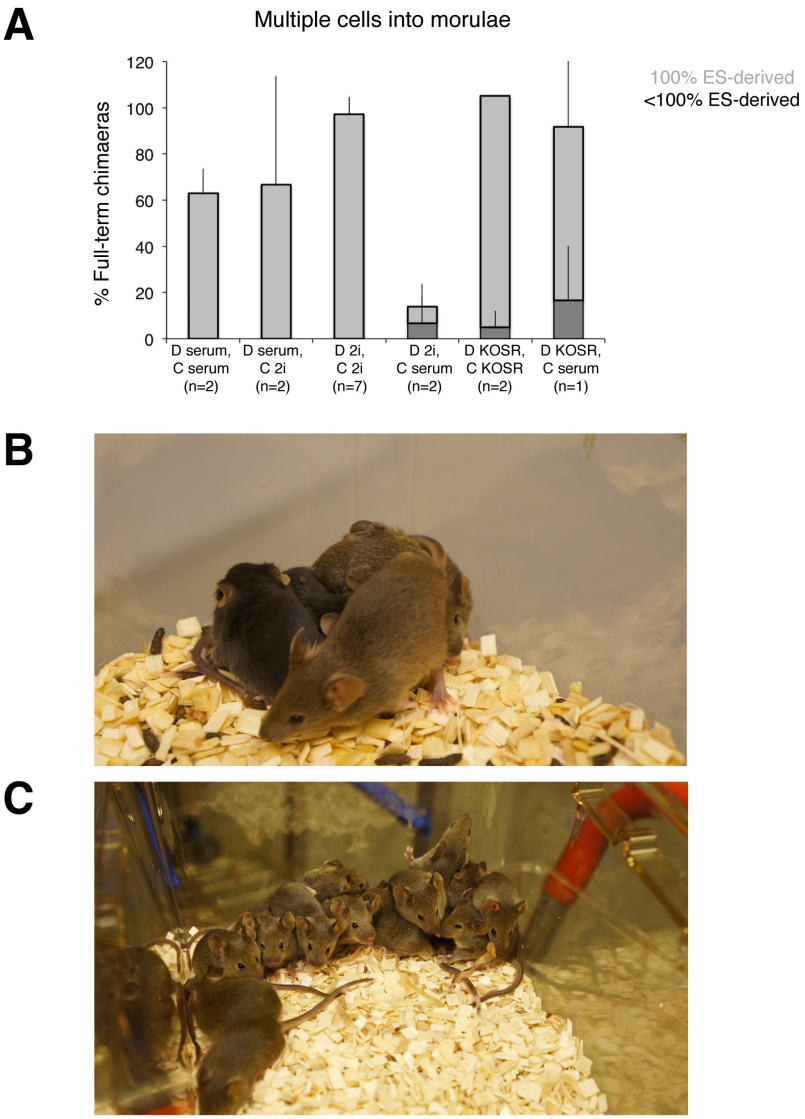


Figure 61. Culture in serum-containing medium reduces the clonogenicity of ES cells in embryo injections. ES cells, were derived in and cultured under various conditions; derived (D) in serum and cultured (C) in serum, D in serum and C in 2i, D in 2i and C in 2i, D in 2i and C in serum and D and C in knockout serum replacement medium (KOSR). **A.** Multiple (5) ES cells from each of these conditions were injected into morulae and their contribution to full-term chimaeras assessed according to coat colour.The majority of chimaeras were 100% ES-derived (light grey) although some had less than 100% ES contribution (dark grey). n values refer to the number of independently derived cell lines that were tested for each condition. Error bars indicate standard deviation of the mean between biological replicates of different cell lines. **B.** Representative image of a litter of mice generated from injection of multiple cells D and C in serum into morulae. **C.** Representative image of a litter of mice generated from injection of multiple cells derived and cultured in 2i into morulae.

1 ES cell into 2-cell embryo

CONDITION	LINE	CHIMAERAS
D serum, C serum	3.4	0/6
	3.2	0/5
D serum, C 2i	3.4	5/6
	3.2	2/6
D 2i, C 2i	2.5	2/18
	2.2	6/11
	2.6	3/18 (1 <100%)
	2.2H	7/14 (2 <100%)
	2.2V	2/4
D 2i, C serum	2.5	0/6
	2.5	0/5
D KOSR, C KOSR	5.3	2/5
	5.45	1/3

1 ES cell into 8-cell morula

CONDITION	LINE	CHIMAERAS
D serum, C serum	3.4	0/11
	3.2	0/2
D serum, C 2i	3.4	6/9
	3.2	1/6
D 2i, C 2i	2.2H	4/9

5-10 ES cells into 8-cell morula

CONDITION	LINE	CHIMAERAS
D serum, C serum	3.4	4/7
	3.2	5/7
D serum, C 2i	3.4	4/4
	3.2	2/6
D 2i, C 2i	2.2	4/4
	2.2H	35/35
	2.2V	4/5
	2.2A1 5	1/1
	2.2A1 11	6/6
	2.2A1 28	6/6
D 2i, C serum	2.5	0/1
	2.2	1/7
D KOSR, C KOSR	5.3	14/14
	5.45	3/3
D KOSR, C serum	5.3	4/6
	5.45	5/6 (5<100%)

Table 7. Tables showing the number of chimaeras generated under each condition. ES cells were derived (D) and cultured (C) in either serum, 2i of knockout serum replacement (KOSR) medium. The number of chimeras generated from each condition is shown out of the total number of pups born. Multiple independently derived cell lines were used e.g. 2.2. Where extensions are added to the end of these cell lines e.g. 2.2H, these represent subclones of the initial cell line.

7.4 Discussion

In this chapter it has been shown that ES cells can generate chimaeras by contributing to embryos at earlier stages than originally believed. In chimaera assays, ES cells are routinely injected into morulae or blastocysts. There is also limited data to show that chimaeras can be generated when ES cells are injected into 4 cell embryos (Huang et al., 2008; Poueymirou et al., 2007). Additionally, previous studies involved injecting multiple cells into embryos hence this is the first time it has been shown that a single cell injected at this early stage of development can generate a 100% ES cell-derived mouse. Mice were scored as 100% ES-derived based on coat colour chimaerism however, we have also derived ES cell lines containing a constitutive Venus marker so that the extent of contribution to internal organs can also be assessed.

As ES cells are known to be heterogeneous (Section 1.3.5), the efficient derivation of mice from a single ES cell could facilitate the study of the distinct functional properties of individual cells, even those that occur at a low frequency. It has recently been shown that even colonies of ES cells derived at clonal density cannot be truly defined as clonal, as it is possible for cells to migrate between colonies. Hence, the most conservative assessment of single cell function is by injection of a single cell rather than multiple ES cells that have been 'clonally' re-derived.

Currently we are continuing research in this area to answer some remaining questions. Single ES cells cultured in serum could not contribute to 2-cell embryos. However, as single ES cells from this condition also could not contribute to the morula, we could not distinguish between the possibilities that serum-cultured cells are not able to contribute to earlier stages due to the fact that they represent a later developmental stage than ES cells cultured in 2i, or as a result of their decreased clonogenicity. To distinguish between these possibilities, we are presently carrying out injections of multiple serum-cultured ES cells into 2-cell embryos to see if this can rescue an effect mediated by reduced clonogenicity. Preliminary results show that the injection of 5 cells cultured in serum into 2-cell embryos cannot rescue this phenotype (D serum, C serum: 0/3 chimaeras, D serum, C 2i: 6/7 100% ES cell-derived mice).

The decreased clonogenicity of ES cells cultured in serum could be due to a number of reasons. For example, ES cells may need neighbouring cells to create a survival niche, perhaps involving secreted factors. As 2i-cultured ES cells showed an increased rate of proliferation compared to serum-cultured cells after introduction to 2-cell embryos, it is possible that they generate neighbouring cells and hence a survival niche more rapidly. To support this, ES cells cultured in the presence of the 2i inhibitors as well as an FGF receptor inhibitor (3i) exhibit increased clonogenicity compared to serum-cultured cells in vitro (Ying et al., 2008).

Additionally, as KOSR-cultured cells showed a similar level of clonogenicity to ES cells cultured in 2i, we are now trying to determine if they share molecular properties that are different between serum and 2i-cultured cells, and therefore whether these differential properties are a result of the lack of serum rather than the presence of the 2 inhibitors. For example, the pluripotency-associated gene *Nanog* is expressed heterogeneously in serum-cultured cells but more homogeneous in 2i-cultured cells. Do KOSR-cultured cells therefore also have homogeneous *Nanog* expression? Experiments are continuing in this area to try to understand the molecular basis of the functional differences that we observe.

BIBLIOGRAPHY

Bibliography

Adelman, K., and Lis, J.T. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat Rev Genet *13*, 720-731.

Aghajanova, L., Skottman, H., Stromberg, A.M., Inzunza, J., Lahesmaa, R., and Hovatta, O. (2006). Expression of leukemia inhibitory factor and its receptors is increased during differentiation of human embryonic stem cells. Fertility and sterility *86*, 1193-1209.

Akira, S., Nishio, Y., Inoue, M., Wang, X.J., Wei, S., Matsusaka, T., Yoshida, K., Sudo, T., Naruto, M., and Kishimoto, T. (1994). Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related transcription factor involved in the gp130-mediated signaling pathway. Cell *77*, 63-71.

Alarcon, V.B., and Marikawa, Y. (2005). Unbiased contribution of the first two blastomeres to mouse blastocyst development. Molecular reproduction and development *72*, 354-361.

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

Antczak, M., and Van Blerkom, J. (1997). Oocyte influences on early development: the regulatory proteins leptin and STAT3 are polarized in mouse and human oocytes and differentially distributed within the cells of the preimplantation stage embryo. Molecular human reproduction *3*, 1067-1086.

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

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

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

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

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

Balakier, H., and Pedersen, R.A. (1982). Allocation of cells to inner cell mass and trophectoderm lineages in preimplantation mouse embryos. Developmental biology *90*, 352-362.

Bany, B.M., and Cross, J.C. (2006). Post-implantation mouse conceptuses produce paracrine signals that regulate the uterine endometrium undergoing decidualization. Developmental biology *294*, 445-456.

Basu, S., Totty, N.F., Irwin, M.S., Sudol, M., and Downward, J. (2003). Akt phosphorylates the Yes-associated protein, YAP, to induce interaction with 14-3-3 and attenuation of p73-mediated apoptosis. Molecular cell *11*, 11-23.

Beck, F., Erler, T., Russell, A., and James, R. (1995). Expression of Cdx-2 in the mouse embryo and placenta: possible role in patterning of the extra-embryonic membranes. Developmental dynamics : an official publication of the American Association of Anatomists *204*, 219-227.

Beck, S., Le Good, J.A., Guzman, M., Ben Haim, N., Roy, K., Beermann, F., and Constam, D.B. (2002). Extraembryonic proteases regulate Nodal signalling during gastrulation. Nature cell biology *4*, 981-985.

Beddington, R.S. (1983). Histogenetic and neoplastic potential of different regions of the mouse embryonic egg cylinder. Journal of embryology and experimental morphology *75*, 189-204.

Beddington, R.S., and Robertson, E.J. (1989). An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. Development *105*, 733-737.

Belo, J.A., Bouwmeester, T., Leyns, L., Kertesz, N., Gallo, M., Follettie, M., and De Robertis, E.M. (1997). Cerberus-like is a secreted factor with neutralizing activity expressed in the anterior primitive endoderm of the mouse gastrula. Mechanisms of development *68*, 45-57.

Bentley, D.L. (1995). Regulation of transcriptional elongation by RNA polymerase II. Current opinion in genetics & development *5*, 210-216.

Bess, K.L., Swingler, T.E., Rivett, A.J., Gaston, K., and Jayaraman, P.S. (2003). The transcriptional repressor protein PRH interacts with the proteasome. The Biochemical journal *374*, 667-675.
Biechele, S., Cockburn, K., Lanner, F., Cox, B.J., and Rossant, J. (2013). Porcn-dependent Wnt signaling is not required prior to mouse gastrulation. Development *140*, 2961-2971.
Blij, S., Frum, T., Akyol, A., Fearon, E., and Ralston, A. (2012). Maternal Cdx2 is dispensable for mouse development. Development *139*, 3969-3972.
Boeuf, H., Hauss, C., Graeve, F.D., Baran, N., and Kedinger, C. (1997). Leukemia inhibitory factor-dependent transcriptional activation in embryonic stem cells. J Cell Biol *138*, 1207-1217.
Bort, R., Martinez-Barbera, J.P., Beddington, R.S., and Zaret, K.S. (2004). Hex homeobox gene-dependent tissue positioning is required for organogenesis of the ventral pancreas. Development *131*, 797-806.
Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K.*, et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature *441*, 349-353.
Bradley, A., Evans, M., Kaufman, M.H., and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature *309*, 255-256.
Brickman, J.M., Jones, C.M., Clements, M., Smith, J.C., and Beddington, R.S. (2000). Hex is a transcriptional repressor that contributes to anterior identity and suppresses Spemann organiser function. Development *127*, 2303-2315.
Brown, K., Legros, S., Artus, J., Doss, M.X., Khanin, R., Hadjantonakis, A.K., and Foley, A. (2010). A comparative analysis of extra-embryonic endoderm cell lines. PloS one *5*, e12016.
Burdon, T., Stracey, C., Chambers, I., Nichols, J., and Smith, A. (1999). Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. Developmental biology *210*, 30-43.
Cai, K.Q., Capo-Chichi, C.D., Rula, M.E., Yang, D.H., and Xu, X.X. (2008). Dynamic GATA6 expression in primitive endoderm formation and maturation in early mouse embryogenesis. Developmental dynamics : an official publication of the American Association of Anatomists *237*, 2820-2829.
Canham, M.A., Sharov, A.A., Ko, M.S., and Brickman, J.M. (2010). Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcript. PLoS biology *8*, e1000379.
Card, D.A., Hebbar, P.B., Li, L., Trotter, K.W., Komatsu, Y., Mishina, Y., and Archer, T.K. (2008). Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells. Molecular and cellular biology *28*, 6426-6438.
Cartwright, P., McLean, C., Sheppard, A., Rivett, D., Jones, K., and Dalton, S. (2005). LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. Development *132*, 885-896.
Chai, N., Patel, Y., Jacobson, K., McMahon, J., McMahon, A., and Rappolee, D.A. (1998). FGF is an essential regulator of the fifth cell division in preimplantation mouse embryos. Developmental biology *198*, 105-115.
Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell *113*, 643-655.
Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. Nature *450*, 1230-1234.
Chappell, S.A., Edelman, G.M., and 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.
Chappell, S.A., Edelman, G.M., and Mauro, V.P. (2004). Biochemical and functional analysis of a 9-nt RNA sequence that affects translation efficiency in eukaryotic cells. Proceedings of the National Academy of Sciences of the United States of America *101*, 9590-9594.
Chazaud, C., Yamanaka, Y., Pawson, T., and Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Developmental cell *10*, 615-624.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J.*, et al.* (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell *133*, 1106-1117.
Cheng, P., Andersen, P., Hassel, D., Kaynak, B.L., Limphong, P., Juergensen, L., Kwon, C., and Srivastava, D. (2013). Fibronectin mediates mesendodermal cell fate decisions. Development *140*, 2587-2596.
Chia Le Bin, G., Munoz-Descalzo, S., Kurowski, A., Leitch, H., Lou, X., Mansfield, W., Etienne-Dumeau, C., Grabole, N., Mulas, C., Niwa, H.*, et al.* (2014). Oct4 is required for lineage priming in the developing inner cell mass of the mouse blastocyst. Development.
Chopra, V.S., Cande, J., Hong, J.W., and Levine, M. (2009). Stalled Hox promoters as chromosomal boundaries. Genes & development *23*, 1505-1509.
Chu, L.F., Surani, M.A., Jaenisch, R., and Zwaka, T.P. (2011). Blimp1 expression predicts embryonic stem cell development in vitro. Current biology : CB *21*, 1759-1765.
Cockburn, K., Biechele, S., Garner, J., and Rossant, J. (2013). The Hippo pathway member Nf2 is required for inner cell mass specification. Current biology : CB *23*, 1195-1201.
Condic, M.L. (2014). Totipotency: What It Is and What It Is Not. Stem cells and development.
Conlon, F.L., Lyons, K.M., Takaesu, N., Barth, K.S., Kispert, A., Herrmann, B., and 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.
Conover, J.C., Ip, N.Y., Poueymirou, W.T., Bates, B., Goldfarb, M.P., DeChiara, T.M., and Yancopoulos, G.D. (1993). Ciliary neurotrophic factor maintains the pluripotentiality of embryonic stem cells. Development *119*, 559-565.
Conquet, F., Peyrieras, N., Tiret, L., and Brulet, P. (1992). Inhibited gastrulation in mouse embryos overexpressing the leukemia inhibitory factor. Proceedings of the National Academy of Sciences of the United States of America *89*, 8195-8199.
Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science *322*, 1845-1848.
Curatola, A.M., and Basilico, C. (1990). Expression of the K-fgf proto-oncogene is controlled by 3' regulatory elements which are specific for embryonal carcinoma cells. Molecular and cellular biology *10*, 2475-2484.
Danko, C.G., and Pertsov, A.M. (2009). Identification of gene co-regulatory modules and associated cis-elements involved in degenerative heart disease. BMC medical genomics *2*, 31.
das Neves, R.P., Jones, N.S., Andreu, L., Gupta, R., Enver, T., and Iborra, F.J. (2010). Connecting variability in global transcription rate to mitochondrial variability. PLoS biology *8*, e1000560.
Davis, S., Aldrich, T.H., Stahl, N., Pan, L., Taga, T., Kishimoto, T., Ip, N.Y., and Yancopoulos, G.D. (1993). LIFR beta and gp130 as heterodimerizing signal transducers of the tripartite CNTF receptor. Science *260*, 1805-1808.
de Navascues, J., Perdigoto, C.N., Bian, Y., Schneider, M.H., Bardin, A.J., Martinez-Arias, A., and Simons, B.D. (2012). Drosophila midgut homeostasis involves neutral competition between symmetrically dividing intestinal stem cells. The EMBO journal *31*, 2473-2485.
del Valle, I., Rudloff, S., Carles, A., Li, Y., Liszewska, E., Vogt, R., and 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.
Dickson, A.D. (1963). Trophoblastic Giant Cell Transformation of Mouse Blastocyts. Journal of reproduction and fertility *6*, 465-466.
Dietrich, J.E., and Hiiragi, T. (2007). Stochastic patterning in the mouse pre-implantation embryo. Development *134*, 4219-4231.
Dietrich, J.E., and Hiiragi, T. (2008). Stochastic processes during mouse blastocyst patterning. Cells, tissues, organs *188*, 46-51.
Diwan, S.B., and Stevens, L.C. (1976). Development of teratomas from the ectoderm of mouse egg cylinders. Journal of the National Cancer Institute *57*, 937-942.
Do, D.V., Ueda, J., Messerschmidt, D.M., Lorthongpanich, C., Zhou, Y., Feng, B., Guo, G., Lin, P.J., Hossain, M.Z., Zhang, W.*, et al.* (2013). A genetic and developmental pathway from STAT3 to the OCT4-NANOG circuit is essential for maintenance of ICM lineages in vivo. Genes & development *27*, 1378-1390.
Dong, J., Feldmann, G., Huang, J., Wu, S., Zhang, N., Comerford, S.A., Gayyed, M.F., Anders, R.A., Maitra, A., and Pan, D. (2007). Elucidation of a universal size-control mechanism in Drosophila and mammals. Cell *130*, 1120-1133.

Driever, W., and Nusslein-Volhard, C. (1988a). The bicoid protein determines position in the Drosophila embryo in a concentration-dependent manner. Cell *54*, 95-104.
Driever, W., and Nusslein-Volhard, C. (1988b). A gradient of bicoid protein in Drosophila embryos. Cell *54*, 83-93.
Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002). Stochastic gene expression in a single cell. Science *297*, 1183-1186.
Ema, M., Mori, D., Niwa, H., Hasegawa, Y., Yamanaka, Y., Hitoshi, S., Mimura, J., Kawabe, Y., Hosoya, T., Morita, M.*, et al.* (2008). Kruppel-like factor 5 is essential for blastocyst development and the normal self-renewal of mouse ESCs. Cell stem cell *3*, 555-567.
Emes, R.D., Goodstadt, L., Winter, E.E., and Ponting, C.P. (2003). Comparison of the genomes of human and mouse lays the foundation of genome zoology. Hum Mol Genet *12*, 701-709.
Erlebacher, A., Price, K.A., and Glimcher, L.H. (2004). Maintenance of mouse trophoblast stem cell proliferation by TGF-beta/activin. Developmental biology *275*, 158-169.
Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154-156.
Faast, R., White, J., Cartwright, P., Crocker, L., Sarcevic, B., and Dalton, S. (2004). Cdk6-cyclin D3 activity in murine ES cells is resistant to inhibition by p16(INK4a). Oncogene *23*, 491-502.
Falco, G., Lee, S.L., Stanghellini, I., Bassey, U.C., Hamatani, T., and Ko, M.S. (2007). Zscan4: a novel gene expressed exclusively in late 2-cell embryos and embryonic stem cells. Developmental biology *307*, 539-550.
Ficz, G., Hore, T.A., Santos, F., Lee, H.J., Dean, W., Arand, J., Krueger, F., Oxley, D., Paul, Y.L., Walter, J.*, et al.* (2013). FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. Cell Stem Cell *13*, 351-359.
Findlay, G.M., Smith, M.J., Lanner, F., Hsiung, M.S., Gish, G.D., Petsalaki, E., Cockburn, K., Kaneko, T., Huang, H., Bagshaw, R.D.*, et al.* (2013). Interaction domains of Sos1/Grb2 are finely tuned for cooperative control of embryonic stem cell fate. Cell *152*, 1008-1020.
Fleming, T.P. (1987). A quantitative analysis of cell allocation to trophectoderm and inner cell mass in the mouse blastocyst. Developmental biology *119*, 520-531.
Frankenberg, S., Gerbe, F., Bessonnard, S., Belville, C., Pouchin, P., Bardot, O., and Chazaud, C. (2011). Primitive endoderm differentiates via a three-step mechanism involving Nanog and RTK signaling. Developmental cell *21*, 1005-1013.
Frum, T., Halbisen, M.A., Wang, C., Amiri, H., Robson, P., and Ralston, A. (2013). Oct4 cell-autonomously promotes primitive endoderm development in the mouse blastocyst. Developmental cell *25*, 610-622.
Fujimori, T., Kurotaki, Y., Miyazaki, J., and Nabeshima, Y. (2003). Analysis of cell lineage in two- and four-cell mouse embryos. Development *130*, 5113-5122.
Furusawa, T., Ohkoshi, K., Honda, C., Takahashi, S., and Tokunaga, T. (2004). Embryonic stem cells expressing both platelet endothelial cell adhesion molecule-1 and stage-specific embryonic antigen-1 differentiate predominantly into epiblast cells in a chimeric embryo. Biology of reproduction *70*, 1452-1457.
Gama-Sosa, M.A., Slagel, V.A., Trewyn, R.W., Oxenhandler, R., Kuo, K.C., Gehrke, C.W., and Ehrlich, M. (1983). The 5-methylcytosine content of DNA from human tumors. Nucleic acids research *11*, 6883-6894.
Garbutt, C.L., Johnson, M.H., and George, M.A. (1987). When and how does cell division order influence cell allocation to the inner cell mass of the mouse blastocyst? Development *100*, 325-332.
Gardner, R.L. (1997). The early blastocyst is bilaterally symmetrical and its axis of symmetry is aligned with the animal-vegetal axis of the zygote in the mouse. Development *124*, 289-301.
Gardner, R.L. (2001). Specification of embryonic axes begins before cleavage in normal mouse development. Development *128*, 839-847.
Gardner, R.L., Meredith, M.R., and Altman, D.G. (1992). Is the anterior-posterior axis of the fetus specified before implantation in the mouse? The Journal of experimental zoology *264*, 437-443.
Gardner, R.L., and Rossant, J. (1979). Investigation of the fate of 4-5 day post-coitum mouse inner cell mass cells by blastocyst injection. Journal of embryology and experimental morphology *52*, 141-152.
Gerbe, F., Cox, B., Rossant, J., and Chazaud, C. (2008). Dynamic expression of Lrp2 pathway members reveals progressive epithelial differentiation of primitive endoderm in mouse blastocyst. Developmental biology *313*, 594-602.

Ghaleb, A.M., Aggarwal, G., Bialkowska, A.B., Nandan, M.O., and Yang, V.W. (2008). Notch inhibits expression of the Kruppel-like factor 4 tumor suppressor in the intestinal epithelium. Molecular cancer research : MCR *6*, 1920-1927.
Ghosh, B., Ganea, G.R., Denson, L.A., Iannucci, R., Jacobs, H.C., and Bogue, C.W. (2000). Immunocytochemical characterization of murine Hex, a homeobox-containing protein. Pediatric research *48*, 634-638.
Ghosh, B., Jacobs, H.C., Wiedemann, L.M., Brown, A., Bedford, F.K., Nimmakayalu, M.A., Ward, D.C., and Bogue, C.W. (1999). Genomic structure, cDNA mapping, and chromosomal localization of the mouse homeobox gene, Hex. Mammalian genome : official journal of the International Mammalian Genome Society *10*, 1023-1025.
Gilchrist, D.A., Dos Santos, G., Fargo, D.C., Xie, B., Gao, Y., Li, L., and Adelman, K. (2010). Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. Cell *143*, 540-551.
Gilchrist, D.A., Nechaev, S., Lee, C., Ghosh, S.K., Collins, J.B., Li, L., Gilmour, D.S., and Adelman, K. (2008). NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. Genes & development *22*, 1921-1933.
Gladden, A.B., Hebert, A.M., Schneeberger, E.E., and McClatchey, A.I. (2010). The NF2 tumor suppressor, Merlin, regulates epidermal development through the establishment of a junctional polarity complex. Developmental cell *19*, 727-739.
Grabarek, J.B., Zyzynska, K., Saiz, N., Piliszek, A., Frankenberg, S., Nichols, J., Hadjantonakis, A.K., and Plusa, B. (2012). Differential plasticity of epiblast and primitive endoderm precursors within the ICM of the early mouse embryo. Development *139*, 129-139.
Grabole, N., Tischler, J., Hackett, J.A., Kim, S., Tang, F., Leitch, H.G., Magnusdottir, E., and Surani, M.A. (2013). Prdm14 promotes germline fate and naive pluripotency by repressing FGF signalling and DNA methylation. EMBO reports *14*, 629-637.
Gregory, S.G., Sekhon, M., Schein, J., Zhao, S., Osoegawa, K., Scott, C.E., Evans, R.S., Burridge, P.W., Cox, T.V., Fox, C.A.*, et al.* (2002). A physical map of the mouse genome. Nature *418*, 743-750.
Griffiths, D.S., Li, J., Dawson, M.A., Trotter, M.W., Cheng, Y.H., Smith, A.M., Mansfield, W., Liu, P., Kouzarides, T., Nichols, J.*, et al.* (2011). LIF-independent JAK signalling to chromatin in embryonic stem cells uncovered from an adult stem cell disease. Nat Cell Biol *13*, 13-21.
Guiral, M., Bess, K., Goodwin, G., and Jayaraman, P.S. (2001). PRH represses transcription in hematopoietic cells by at least two independent mechanisms. The Journal of biological chemistry *276*, 2961-2970.
Guo, G., Huss, M., Tong, G.Q., Wang, C., Li Sun, L., Clarke, N.D., and Robson, P. (2010). Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. Developmental cell *18*, 675-685.
Guzman-Ayala, M., Ben-Haim, N., Beck, S., and Constam, D.B. (2004). Nodal protein processing and fibroblast growth factor 4 synergize to maintain a trophoblast stem cell microenvironment. Proceedings of the National Academy of Sciences of the United States of America *101*, 15656-15660.
Hall, J., Guo, G., Wray, J., Eyres, I., Nichols, J., Grotewold, L., Morfopoulou, S., Humphreys, P., Mansfield, W., Walker, R.*, et al.* (2009). Oct4 and LIF/Stat3 additively induce Kruppel factors to sustain embryonic stem cell self-renewal. Cell stem cell *5*, 597-609.
Hamazaki, T., Kehoe, S.M., Nakano, T., and Terada, N. (2006). The Grb2/Mek pathway represses Nanog in murine embryonic stem cells. Molecular and cellular biology *26*, 7539-7549.
Handyside, A.H. (1978). Time of commitment of inside cells isolated from preimplantation mouse embryos. Journal of embryology and experimental morphology *45*, 37-53.
Hao, J., Li, T.G., Qi, X., Zhao, D.F., and Zhao, G.Q. (2006). WNT/beta-catenin pathway up-regulates Stat3 and converges on LIF to prevent differentiation of mouse embryonic stem cells. Developmental biology *290*, 81-91.
Hayashi, K., Lopes, S.M., Tang, F., and Surani, M.A. (2008). Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. Cell stem cell *3*, 391-401.
Hiiragi, T., and Solter, D. (2005). Mechanism of first cleavage specification in the mouse egg: is our body plan set at day 0? Cell Cycle *4*, 661-664.
Hirate, Y., Hirahara, S., Inoue, K., Suzuki, A., Alarcon, V.B., Akimoto, K., Hirai, T., Hara, T., Adachi, M., Chida, K.*, et al.* (2013). Polarity-dependent distribution of angiomotin localizes Hippo signaling in preimplantation embryos. Current biology : CB *23*, 1181-1194.

Hishida, T., Nozaki, Y., Nakachi, Y., Mizuno, Y., Okazaki, Y., Ema, M., Takahashi, S., Nishimoto, M., and Okuda, A. (2011). Indefinite self-renewal of ESCs through Myc/Max transcriptional complex-independent mechanisms. Cell stem cell *9*, 37-49.
Hogan, B., and Tilly, R. (1978). In vitro development of inner cell masses isolated immunosurgically from mouse blastocysts. I. Inner cell masses from 3.5-day p.c. blastocysts incubated for 24 h before immunosurgery. Journal of embryology and experimental morphology *45*, 93-105.
Hogan, B.L., Cooper, A.R., and Kurkinen, M. (1980). Incorporation into Reichert's membrane of laminin-like extracellular proteins synthesized by parietal endoderm cells of the mouse embryo. Developmental biology *80*, 289-300.
Hogan, B.L., Taylor, A., and Cooper, A.R. (1982). Murine parietal endoderm cells synthesise heparan sulphate and 170K and 145K sulphated glycoproteins as components of Reichert's membrane. Developmental biology *90*, 210-214.
Hogan, B.L., and Tilly, R. (1981). Cell interactions and endoderm differentiation in cultured mouse embryos. Journal of embryology and experimental morphology *62*, 379-394.
Home, P., Saha, B., Ray, S., Dutta, D., Gunewardena, S., Yoo, B., Pal, A., Vivian, J.L., Larson, M., Petroff, M.*, et al.* (2012). Altered subcellular localization of transcription factor TEAD4 regulates first mammalian cell lineage commitment. Proceedings of the National Academy of Sciences of the United States of America *109*, 7362-7367.
Hough, S.R., Clements, I., Welch, P.J., and Wiederholt, K.A. (2006). Differentiation of mouse embryonic stem cells after RNA interference-mediated silencing of OCT4 and Nanog. Stem Cells *24*, 1467-1475.
Hu, G., Kim, J., Xu, Q., Leng, Y., Orkin, S.H., and Elledge, S.J. (2009). A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. Genes & development *23*, 837-848.
Huang, J., Deng, K., Wu, H., Liu, Z., Chen, Z., Cao, S., Zhou, L., Ye, X., Keefe, D.L., and Liu, L. (2008). Efficient production of mice from embryonic stem cells injected into four- or eight-cell embryos by piezo micromanipulation. Stem Cells *26*, 1883-1890.
Hunter, M.P., Wilson, C.M., Jiang, X., Cong, R., Vasavada, H., Kaestner, K.H., and Bogue, C.W. (2007). The homeobox gene Hhex is essential for proper hepatoblast differentiation and bile duct morphogenesis. Developmental biology *308*, 355-367.
Ilgren, E.B. (1981a). The initiation and control of trophoblastic growth in the mouse: binucleation and polyploidy. Placenta *2*, 317-332.
Ilgren, E.B. (1981b). On the control of the trophoblastic giant-cell transformation in the mouse: homotypic cellular interactions and polyploidy. Journal of embryology and experimental morphology *62*, 183-202.
James, D., Levine, A.J., Besser, D., and Hemmati-Brivanlou, A. (2005). TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. Development *132*, 1273-1282.
Jirmanova, L., Afanassieff, M., Gobert-Gosse, S., Markossian, S., and Savatier, P. (2002). Differential contributions of ERK and PI3-kinase to the regulation of cyclin D1 expression and to the control of the G1/S transition in mouse embryonic stem cells. Oncogene *21*, 5515-5528.
Johnson, M.H., and Rossant, J. (1981). Molecular studies on cells of the trophectodermal lineage of the postimplantation mouse embryo. Journal of embryology and experimental morphology *61*, 103-116.
Kaern, M., Elston, T.C., Blake, W.J., and Collins, J.J. (2005). Stochasticity in gene expression: from theories to phenotypes. Nature reviews Genetics *6*, 451-464.
Kaji, K., Norrby, K., Paca, A., Mileikovsky, M., Mohseni, P., and Woltjen, K. (2009). Virus-free induction of pluripotency and subsequent excision of reprogramming factors. Nature *458*, 771-775.
Kalmar, T., Lim, C., Hayward, P., Munoz-Descalzo, S., Nichols, J., Garcia-Ojalvo, J., and Martinez Arias, A. (2009). Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells. PLoS biology *7*, e1000149.
Kang, M., Piliszek, A., Artus, J., and Hadjantonakis, A.K. (2013). FGF4 is required for lineage restriction and salt-and-pepper distribution of primitive endoderm factors but not their initial expression in the mouse. Development *140*, 267-279.
Kasamatsu, S., Sato, A., Yamamoto, T., Keng, V.W., Yoshida, H., Yamazaki, Y., Shimoda, M., Miyazaki, J., and Noguchi, T. (2004). Identification of the transactivating region of the homeodomain protein, hex. Journal of biochemistry *135*, 217-223.
Kelly, S.J., Mulnard, J.G., and Graham, C.F. (1978). Cell division and cell allocation in early mouse development. Journal of embryology and experimental morphology *48*, 37-51.

Keng, V.W., Yagi, H., Ikawa, M., Nagano, T., Myint, Z., Yamada, K., Tanaka, T., Sato, A., Muramatsu, I., Okabe, M.*, et al.* (2000). Homeobox gene Hex is essential for onset of mouse embryonic liver development and differentiation of the monocyte lineage. Biochemical and biophysical research communications *276*, 1155-1161.
Kidder, B.L., Yang, J., and Palmer, S. (2008). Stat3 and c-Myc genome-wide promoter occupancy in embryonic stem cells. PloS one *3*, e3932.
Kim, D., Lee, J., Cheng, D., Li, J., Carter, C., Richie, E., and Bedford, M.T. (2010). Enzymatic activity is required for the in vivo functions of CARM1. The Journal of biological chemistry *285*, 1147-1152.
Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S.H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. Cell *132*, 1049-1061.
Kirby, L.A., Schott, J.T., Noble, B.L., Mendez, D.C., Caseley, P.S., Peterson, S.C., Routledge, T.J., and Patel, N.V. (2012). Glycogen synthase kinase 3 (GSK3) inhibitor, SB-216763, promotes pluripotency in mouse embryonic stem cells. PloS one *7*, e39329.
Klein, A.M., and Simons, B.D. (2011). Universal patterns of stem cell fate in cycling adult tissues. Development *138*, 3103-3111.
Kleinsmith, L.J., and Pierce, G.B., Jr. (1964). Multipotentiality of Single Embryonal Carcinoma Cells. Cancer research *24*, 1544-1551.
Kobayashi, T., Mizuno, H., Imayoshi, I., Furusawa, C., Shirahige, K., and Kageyama, R. (2009). The cyclic gene Hes1 contributes to diverse differentiation responses of embryonic stem cells. Genes & development *23*, 1870-1875.
Krawchuk, D., Honma-Yamanaka, N., Anani, S., and Yamanaka, Y. (2013). FGF4 is a limiting factor controlling the proportions of primitive endoderm and epiblast in the ICM of the mouse blastocyst. Developmental biology *384*, 65-71.
Kunath, T., Arnaud, D., Uy, G.D., Okamoto, I., Chureau, C., Yamanaka, Y., Heard, E., Gardner, R.L., Avner, P., and Rossant, J. (2005). Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. Development *132*, 1649-1661.
Kunath, T., Saba-El-Leil, M.K., Almousailleakh, M., Wray, J., Meloche, S., and 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.
Kwon, G.S., Viotti, M., and 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.
Lallemand, Y., and Brulet, P. (1990). An in situ assessment of the routes and extents of colonisation of the mouse embryo by embryonic stem cells and their descendants. Development *110*, 1241-1248.
Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology *10*, R25.
Lawson, K.A., Dunn, N.R., Roelen, B.A., Zeinstra, L.M., Davis, A.M., Wright, C.V., Korving, J.P., and Hogan, B.L. (1999). Bmp4 is required for the generation of primordial germ cells in the mouse embryo. Genes & development *13*, 424-436.
Lawson, K.A., and Hage, W.J. (1994). Clonal analysis of the origin of primordial germ cells in the mouse. Ciba Foundation symposium *182*, 68-84; discussion 84-91.
Lehman, J.M., Speers, W.C., Swartzendruber, D.E., and Pierce, G.B. (1974). Neoplastic differentiation: characteristics of cell lines derived from a murine teratocarcinoma. Journal of cellular physiology *84*, 13-27.
Lehtonen, E. (1980). Changes in cell dimensions and intercellular contacts during cleavage-stage cell cycles in mouse embryonic cells. Journal of embryology and experimental morphology *58*, 231-249.
Leitch, H.G., McEwen, K.R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W., Nashun, B., Knezovich, J.G., Smith, A.*, et al.* (2013). Naive pluripotency is associated with global DNA hypomethylation. Nat Struct Mol Biol *20*, 311-316.
Li, B., Weber, J.A., Chen, Y., Greenleaf, A.L., and Gilmour, D.S. (1996). Analyses of promoter-proximal pausing by RNA polymerase II on the hsp70 heat shock gene promoter in a Drosophila nuclear extract. Molecular and cellular biology *16*, 5433-5443.
Li, J., Pan, G., Cui, K., Liu, Y., Xu, S., and Pei, D. (2007). A dominant-negative form of mouse SOX2 induces trophectoderm differentiation and progressive polyploidy in mouse embryonic stem cells. The Journal of biological chemistry *282*, 19481-19492.
Li, M., Pevny, L., Lovell-Badge, R., and Smith, A. (1998). Generation of purified neural precursors from embryonic stem cells by lineage selection. Current biology : CB *8*, 971-974.

Li, M., Sendtner, M., and Smith, A. (1995). Essential function of LIF receptor in motor neurons. Nature *378*, 724-727.
Lin, S.C., Wani, M.A., Whitsett, J.A., and Wells, J.M. (2010). Klf5 regulates lineage formation in the pre-implantation mouse embryo. Development *137*, 3953-3963.
Liu, L., Luo, G.Z., Yang, W., Zhao, X., Zheng, Q., Lv, Z., Li, W., Wu, H.J., Wang, L., Wang, X.J*., et al.* (2010). Activation of the imprinted Dlk1-Dio3 region correlates with pluripotency levels of mouse stem cells. The Journal of biological chemistry *285*, 19483-19490.
Loh, K.M., and Lim, B. (2011). A precarious balance: pluripotency factors as lineage specifiers. Cell stem cell *8*, 363-369.
Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J*., et al.* (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nature genetics *38*, 431-440.
Lowell, S., Benchoua, A., Heavey, B., and Smith, A.G. (2006). Notch promotes neural lineage entry by pluripotent embryonic stem cells. PLoS biology *4*, e121.
Ma, G.T., Soloveva, V., Tzeng, S.J., Lowe, L.A., Pfendler, K.C., Iannaccone, P.M., Kuehn, M.R., and Linzer, D.I. (2001). Nodal regulates trophoblast differentiation and placental development. Developmental biology *236*, 124-135.
MacArthur, B.D., Sevilla, A., Lenz, M., Muller, F.J., Schuldt, B.M., Schuppert, A.A., Ridden, S.J., Stumpf, P.S., Fidalgo, M., Ma'ayan, A*., et al.* (2012). Nanog-dependent feedback loops regulate murine embryonic stem cell heterogeneity. Nature cell biology *14*, 1139-1147.
Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O., Trono, D., and Pfaff, S.L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature *487*, 57-63.
Marks, H., Kalkan, T., Menafra, R., Denissov, S., Jones, K., Hofemeister, H., Nichols, J., Kranz, A., Stewart, A.F., Smith, A*., et al.* (2012). The transcriptional and epigenomic foundations of ground state pluripotency. Cell *149*, 590-604.
Martello, G., Bertone, P., and Smith, A. (2013). Identification of the missing pluripotency mediator downstream of leukaemia inhibitory factor. The EMBO journal *32*, 2561-2574.
Martello, G., Sugimoto, T., Diamanti, E., Joshi, A., Hannah, R., Ohtsuka, S., Gottgens, B., Niwa, H., and Smith, A. (2012). Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. Cell Stem Cell *11*, 491-504.
Martinez Barbera, J.P., Clements, M., Thomas, P., Rodriguez, T., Meloy, D., Kioussis, D., and Beddington, R.S. (2000). The homeobox gene Hex is required in definitive endodermal tissues for normal forebrain, liver and thyroid formation. Development *127*, 2433-2445.
Matsuda, T., Nakamura, T., Nakao, K., Arai, T., Katsuki, M., Heike, T., and Yokota, T. (1999). STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. The EMBO journal *18*, 4261-4269.
Meno, C., Gritsman, K., Ohishi, S., Ohfuji, Y., Heckscher, E., Mochida, K., Shimono, A., Kondoh, H., Talbot, W.S., Robertson, E.J*., et al.* (1999). Mouse Lefty2 and zebrafish antivin are feedback inhibitors of nodal signaling during vertebrate gastrulation. Molecular cell *4*, 287-298.
Messerschmidt, D.M., and Kemler, R. (2010). Nanog is required for primitive endoderm formation through a non-cell autonomous mechanism. Developmental biology *344*, 129-137.
Min, I.M., Waterfall, J.J., Core, L.J., Munroe, R.J., Schimenti, J., and Lis, J.T. (2011). Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. Genes & development *25*, 742-754.
Minami, T., Murakami, T., Horiuchi, K., Miura, M., Noguchi, T., Miyazaki, J., Hamakubo, T., Aird, W.C., and Kodama, T. (2004). Interaction between hex and GATA transcription factors in vascular endothelial cells inhibits flk-1/KDR-mediated vascular endothelial growth factor signaling. The Journal of biological chemistry *279*, 20626-20635.
Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631-642.
Molls, M., Zamboglou, N., and Streffer, C. (1983). A comparison of the cell kinetics of pre-implantation mouse embryos from two different mouse strains. Cell and tissue kinetics *16*, 277-283.
Morgani, S.M., and Brickman, J.M. (2013). Survival of the fattest. eLife *2*.
Morris, S.A., Graham, S.J., Jedrusik, A., and Zernicka-Goetz, M. (2013). The differential response to Fgf signalling in cells internalized at different times influences lineage segregation in preimplantation mouse embryos. Open biology *3*, 130104.

Morrison, S.J., and Kimble, J. (2006). Asymmetric and symmetric stem-cell divisions in development and cancer. Nature *441*, 1068-1074.
Mummery, C.L., van Achterberg, T.A., van den Eijnden-van Raaij, A.J., van Haaster, L., Willemse, A., de Laat, S.W., and Piersma, A.H. (1991). Visceral-endoderm-like cell lines induce differentiation of murine P19 embryonal carcinoma cells. Differentiation; research in biological diversity *46*, 51-60.
Najm, F.J., Chenoweth, J.G., Anderson, P.D., Nadeau, J.H., Redline, R.W., McKay, R.D., and Tesar, P.J. (2011). Isolation of epiblast stem cells from preimplantation mouse embryos. Cell stem cell *8*, 318-325.
Nakatake, Y., Fukui, N., Iwamatsu, Y., Masui, S., Takahashi, K., Yagi, R., Yagi, K., Miyazaki, J., Matoba, R., Ko, M.S.*, et al.* (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.
Nichols, J., Chambers, I., and Smith, A. (1994). Derivation of germline competent embryonic stem cells with a combination of interleukin-6 and soluble interleukin-6 receptor. Experimental cell research *215*, 237-239.
Nichols, J., Chambers, I., Taga, T., and Smith, A. (2001). Physiological rationale for responsiveness of mouse embryonic stem cells to gp130 cytokines. Development *128*, 2333-2339.
Nichols, J., Davidson, D., Taga, T., Yoshida, K., Chambers, I., and Smith, A. (1996). Complementary tissue-specific expression of LIF and LIF-receptor mRNAs in early mouse embryogenesis. Mech Dev *57*, 123-131.
Nichols, J., Silva, J., Roode, M., and Smith, A. (2009). Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. Development *136*, 3215-3222.
Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell *95*, 379-391.
Nishikawa, S.I., Nishikawa, S., Hirashima, M., Matsuyoshi, N., and Kodama, H. (1998). Progressive lineage analysis by cell sorting and culture identifies FLK1+VE-cadherin+ cells at a diverging point of endothelial and hemopoietic lineages. Development *125*, 1747-1757.
Nishioka, N., Inoue, K., Adachi, K., Kiyonari, H., Ota, M., Ralston, A., Yabuta, N., Hirahara, S., Stephenson, R.O., Ogonuki, N.*, et al.* (2009). The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. Developmental cell *16*, 398-410.
Niwa, H., Burdon, T., Chambers, I., and Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes & development *12*, 2048-2060.
Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nature genetics *24*, 372-376.
Niwa, H., Ogawa, K., Shimosato, D., and Adachi, K. (2009). A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. Nature *460*, 118-122.
Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. Cell *123*, 917-929.
O'Brien, T., and Lis, J.T. (1991). RNA polymerase II pauses at the 5' end of the transcriptionally induced Drosophila hsp70 gene. Molecular and cellular biology *11*, 5285-5290.
Obinata, A., Akimoto, Y., Omoto, Y., and Hirano, H. (2002). Expression of Hex homeobox gene during skin development: Increase in epidermal cell proliferation by transfecting the Hex to the dermis. Development, growth & differentiation *44*, 281-292.
Ogawa, K., Saito, A., Matsui, H., Suzuki, H., Ohtsuka, S., Shimosato, D., Morishita, Y., Watabe, T., Niwa, H., and Miyazono, K. (2007). Activin-Nodal signaling is involved in propagation of mouse embryonic stem cells. Journal of cell science *120*, 55-65.
Ohnishi, Y., Huber, W., Tsumura, A., Kang, M., Xenopoulos, P., Kurimoto, K., Oles, A.K., Arauzo-Bravo, M.J., Saitou, M., Hadjantonakis, A.K.*, et al.* (2014). Cell-to-cell expression variability followed by signal reinforcement progressively segregates early mouse lineages. Nature cell biology *16*, 27-37.
Paca, A., Seguin, C.A., Clements, M., Ryczko, M., Rossant, J., Rodriguez, T.A., and Kunath, T. (2012). BMP signaling induces visceral endoderm differentiation of XEN cells and parietal endoderm. Developmental biology *361*, 90-102.
Paling, N.R., Wheadon, H., Bone, H.K., and Welham, M.J. (2004). Regulation of embryonic stem cell self-renewal by phosphoinositide 3-kinase-dependent signaling. The Journal of biological chemistry *279*, 48063-48070.

Palmieri, S.L., Peter, W., Hess, H., and Scholer, H.R. (1994). Oct-4 transcription factor is differentially expressed in the mouse embryo during establishment of the first two extraembryonic cell lineages involved in implantation. Developmental biology *166*, 259-267.
Parfitt, D.E., and Zernicka-Goetz, M. (2010). Epigenetic modification affecting expression of cell polarity and cell fate genes to regulate lineage specification in the early mouse embryo. Molecular biology of the cell *21*, 2649-2660.
Pauken, C.M., and Capco, D.G. (2000). The expression and stage-specific localization of protein kinase C isotypes during mouse preimplantation development. Developmental biology *223*, 411-421.
Pauklin, S., and Vallier, L. (2013). The cell-cycle state of stem cells determines cell fate propensity. Cell *155*, 135-147.
Pennica, D., Shaw, K.J., Swanson, T.A., Moore, M.W., Shelton, D.L., Zioncheck, K.A., Rosenthal, A., Taga, T., Paoni, N.F., and Wood, W.I. (1995). Cardiotrophin-1. Biological activities and binding to the leukemia inhibitory factor receptor/gp130 signaling complex. J Biol Chem *270*, 10915-10922.
Perea-Gomez, A., Vella, F.D., Shawlot, W., Oulad-Abdelghani, M., Chazaud, C., Meno, C., Pfister, V., Chen, L., Robertson, E., Hamada, H.*, et al.* (2002). Nodal antagonists in the anterior visceral endoderm prevent the formation of multiple primitive streaks. Developmental cell *3*, 745-756.
Pesce, M., and Scholer, H.R. (2001). Oct-4: gatekeeper in the beginnings of mammalian development. Stem Cells *19*, 271-278.
Piliszek, A., Kwon, G.S., and Hadjantonakis, A.K. (2011). Ex utero culture and live imaging of mouse embryos. Methods Mol Biol *770*, 243-257.
Piotrowska, K., Wianny, F., Pedersen, R.A., and Zernicka-Goetz, M. (2001). Blastomeres arising from the first cleavage division have distinguishable fates in normal mouse development. Development *128*, 3739-3748.
Piotrowska-Nitsche, K., Perea-Gomez, A., Haraguchi, S., and Zernicka-Goetz, M. (2005). Four-cell stage mouse blastomeres have different developmental properties. Development *132*, 479-490.
Piotrowska-Nitsche, K., and Zernicka-Goetz, M. (2005). Spatial arrangement of individual 4-cell stage blastomeres and the order in which they are generated correlate with blastocyst pattern in the mouse embryo. Mechanisms of development *122*, 487-500.
Plusa, B., Frankenberg, S., Chalmers, A., Hadjantonakis, A.K., Moore, C.A., Papalopulu, N., Papaioannou, V.E., Glover, D.M., and Zernicka-Goetz, M. (2005). Downregulation of Par3 and aPKC function directs cells towards the ICM in the preimplantation mouse embryo. Journal of cell science *118*, 505-515.
Plusa, B., Piliszek, A., Frankenberg, S., Artus, J., and Hadjantonakis, A.K. (2008). Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. Development *135*, 3081-3091.
Poehlmann, T.G., Fitzgerald, J.S., Meissner, A., Wengenmayer, T., Schleussner, E., Friedrich, K., and Markert, U.R. (2005). Trophoblast invasion: tuning through LIF, signalling via Stat3. Placenta *26 Suppl A*, S37-41.
Poueymirou, W.T., Auerbach, W., Frendewey, D., Hickey, J.F., Escaravage, J.M., Esau, L., Dore, A.T., Stevens, S., Adams, N.C., Dominguez, M.G.*, et al.* (2007). F0 generation mice fully derived from gene-targeted embryonic stem cells allowing immediate phenotypic analyses. Nature biotechnology *25*, 91-99.
Prakash, G.J., Suman, P., Morales Prieto, D.M., Markert, U.R., and Gupta, S.K. (2011). Leukaemia inhibitory factor mediated proliferation of HTR-8/SVneo trophoblast cells is dependent on activation of extracellular signal-regulated kinase 1/2. Reproduction, fertility, and development *23*, 714-724.
Przybyla, L.M., Theunissen, T.W., Jaenisch, R., and Voldman, J. (2013). Matrix remodeling maintains embryonic stem cell self-renewal by activating Stat3. Stem Cells *31*, 1097-1106.
Puppin, C., Presta, I., D'Elia, A.V., Tell, G., Arturi, F., Russo, D., Filetti, S., and Damante, G. (2004). Functional interaction among thyroid-specific transcription factors: Pax8 regulates the activity of Hex promoter. Molecular and cellular endocrinology *214*, 117-125.
Quinn, J., Kunath, T., and Rossant, J. (2006). Mouse trophoblast stem cells. Methods in molecular medicine *121*, 125-148.
Quyn, A.J., Appleton, P.L., Carey, F.A., Steele, R.J., Barker, N., Clevers, H., Ridgway, R.A., Sansom, O.J., and Nathke, I.S. (2010). Spindle orientation bias in gut epithelial stem cell compartments is lost in precancerous tissue. Cell stem cell *6*, 175-181.
Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y., and Tyagi, S. (2006). Stochastic mRNA synthesis in mammalian cells. PLoS biology *4*, e309.

Ralston, A., Cox, B.J., Nishioka, N., Sasaki, H., Chea, E., Rugg-Gunn, P., Guo, G., Robson, P., Draper, J.S., and Rossant, J. (2010). Gata3 regulates trophoblast development downstream of Tead4 and in parallel to Cdx2. Development *137*, 395-403.
Ralston, A., and Rossant, J. (2008). Cdx2 acts downstream of cell polarization to cell-autonomously promote trophectoderm fate in the early mouse embryo. Developmental biology *313*, 614-629.
Rasmussen, E.B., and Lis, J.T. (1993). In vivo transcriptional pausing and cap formation on three Drosophila heat shock genes. Proceedings of the National Academy of Sciences of the United States of America *90*, 7923-7927.
Rathjen, P.D., Toth, S., Willis, A., Heath, J.K., and Smith, A.G. (1990). Differentiation inhibiting activity is produced in matrix-associated and diffusible forms that are generated by alternate promoter usage. Cell *62*, 1105-1114.
Reddington, J.P., Perricone, S.M., Nestor, C.E., Reichmann, J., Youngson, N.A., Suzuki, M., Reinhardt, D., Dunican, D.S., Prendergast, J.G., Mjoseng, H.*, et al.* (2013). Redistribution of H3K27me3 upon DNA hypomethylation results in de-repression of Polycomb target genes. Genome biology *14*, R25.
Riethmacher, D., Brinkmann, V., and Birchmeier, C. (1995). A targeted mutation in the mouse E-cadherin gene results in defective preimplantation development. Proceedings of the National Academy of Sciences of the United States of America *92*, 855-859.
Robertson, M., Chambers, I., Rathjen, P., Nichols, J., and Smith, A. (1993). Expression of alternative forms of differentiation inhibiting activity (DIA/LIF) during murine embryogenesis and in neonatal and adult tissues. Developmental genetics *14*, 165-173.
Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.
Rodda, D.J., Chew, J.L., Lim, L.H., Loh, Y.H., Wang, B., Ng, H.H., and Robson, P. (2005). Transcriptional regulation of nanog by OCT4 and SOX2. The Journal of biological chemistry *280*, 24731-24737.
Rodriguez, A., Allegrucci, C., and Alberio, R. (2012). Modulation of pluripotency in the porcine embryo and iPS cells. PloS one *7*, e49079.
Rodriguez, T.A., Casey, E.S., Harland, R.M., Smith, J.C., and Beddington, R.S. (2001). Distinct enhancer elements control Hex expression during gastrulation and early organogenesis. Developmental biology *234*, 304-316.
Rose, T.M., Weiford, D.M., Gunderson, N.L., and Bruce, A.G. (1994). Oncostatin M (OSM) inhibits the differentiation of pluripotent embryonic stem cells in vitro. Cytokine *6*, 48-54.
Rossant, J., and Frels, W.I. (1980). Interspecific chimeras in mammals: successful production of live chimeras between Mus musculus and Mus caroli. Science *208*, 419-421.
Rossant, J., and Lis, W.T. (1979). Potential of isolated mouse inner cell masses to form trophectoderm derivatives in vivo. Developmental biology *70*, 255-261.
Rossant, J., and Tamura-Lis, W. (1981). Effect of culture conditions on diploid to giant-cell transformation in postimplantation mouse trophoblast. Journal of embryology and experimental morphology *62*, 217-227.
Rossant, J., and Vijh, K.M. (1980). Ability of outside cells from preimplantation mouse embryos to form inner cell mass derivatives. Developmental biology *76*, 475-482.
Rougvie, A.E., and Lis, J.T. (1988). The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged. Cell *54*, 795-804.
Rugg-Gunn, P.J., Cox, B.J., Lanner, F., Sharma, P., Ignatchenko, V., McDonald, A.C., Garner, J., Gramolini, A.O., Rossant, J., and Kislinger, T. (2012). Cell-surface proteomics identifies lineage-specific markers of embryo-derived stem cells. Developmental cell *22*, 887-901.
Saburi, S., Azuma, S., Sato, E., Toyoda, Y., and Tachi, C. (1997). Developmental fate of single embryonic stem cells microinjected into 8-cell-stage mouse embryos. Differentiation; research in biological diversity *62*, 1-11.
Saiz, N., Grabarek, J.B., Sabherwal, N., Papalopulu, N., and Plusa, B. (2013). Atypical protein kinase C couples cell sorting with primitive endoderm maturation in the mouse blastocyst. Development *140*, 4311-4322.
Sanchez-Ripoll, Y., Bone, H.K., Owen, T., Guedes, A.M., Abranches, E., Kumpfmueller, B., Spriggs, R.V., Henrique, D., and Welham, M.J. (2013). Glycogen synthase kinase-3 inhibition enhances translation of pluripotency-associated transcription factors to contribute to maintenance of mouse embryonic stem cell self-renewal. PloS one *8*, e60148.

Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and 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.
Saunders, A., Core, L.J., Sutcliffe, C., Lis, J.T., and Ashe, H.L. (2013). Extensive polymerase pausing during Drosophila axis patterning enables high-level and pliable transcription. Genes & development *27*, 1146-1158.
Sawarkar, R., Sievers, C., and Paro, R. (2012). Hsp90 globally targets paused RNA polymerase to regulate gene expression in response to environmental stimuli. Cell *149*, 807-818.
Schaefer, L.K., Wang, S., and Schaefer, T.S. (2001). Functional interaction of Jun and homeodomain proteins. The Journal of biological chemistry *276*, 43074-43082.
Schroeder, T., Fraser, S.T., Ogawa, M., Nishikawa, S., Oka, C., Bornkamm, G.W., Nishikawa, S., Honjo, T., and Just, U. (2003). Recombination signal sequence-binding protein Jkappa alters mesodermal cell fate decisions by suppressing cardiomyogenesis. Proceedings of the National Academy of Sciences of the United States of America *100*, 4018-4023.
Semoff, S., Hogan, B.L., and Hopkins, C.R. (1982). Localization of fibronectin, laminin-entactin, and entactin in Reichert's membrane by immunoelectron microscopy. The EMBO journal *1*, 1171-1175.
Shen, M.M., and Leder, P. (1992). Leukemia inhibitory factor is expressed by the preimplantation uterus and selectively blocks primitive ectoderm formation in vitro. Proceedings of the National Academy of Sciences of the United States of America *89*, 8240-8244.
Shu, J., Wu, C., Wu, Y., Li, Z., Shao, S., Zhao, W., Tang, X., Yang, H., Shen, L., Zuo, X.*, et al.* (2013). Induction of pluripotency in mouse somatic cells with lineage specifiers. Cell *153*, 963-975.
Singh, A.M., Chappell, J., Trost, R., Lin, L., Wang, T., Tang, J., Wu, H., Zhao, S., Jin, P., and Dalton, S. (2013). Cell-cycle control of developmentally regulated transcription factors accounts for heterogeneity in human pluripotent cells. Stem cell reports *1*, 532-544.
Singh, A.M., Hamazaki, T., Hankowski, K.E., and Terada, N. (2007). A heterogeneous expression pattern for Nanog in embryonic stem cells. Stem Cells *25*, 2534-2542.
Smith, A. (2005). The battlefield of pluripotency. Cell *123*, 757-760.
Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature *336*, 688-690.
Smith, A.G., Nichols, J., Robertson, M., and Rathjen, P.D. (1992). Differentiation inhibiting activity (DIA/LIF) and mouse development. Developmental biology *151*, 339-351.
Smith, K.K., and Strickland, S. (1981). Structural components and characteristics of Reichert's membrane, an extra-embryonic basement membrane. The Journal of biological chemistry *256*, 4654-4661.
Smith, K.N., Singh, A.M., and Dalton, S. (2010). Myc represses primitive endoderm differentiation in pluripotent stem cells. Cell Stem Cell *7*, 343-354.
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 4 1/4 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 from 4 1/4 to 9 1/2 days. Journal of embryology and experimental morphology *89*, 15-35.
Smith, Z.D., Chan, M.M., Mikkelsen, T.S., Gu, H., Gnirke, A., Regev, A., and Meissner, A. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature *484*, 339-344.
Soufi, A., Smith, C., Clarke, A.R., Gaston, K., and Jayaraman, P.S. (2006). Oligomerisation of the developmental regulator proline rich homeodomain (PRH/Hex) is mediated by a novel proline-rich dimerisation domain. Journal of molecular biology *358*, 943-962.
Spindle, A.I. (1978). Trophoblast regeneration by inner cell masses isolated from cultured mouse embryos. The Journal of experimental zoology *203*, 483-489.
Stead, E., White, J., Faast, R., Conn, S., Goldstone, S., Rathjen, J., Dhingra, U., Rathjen, P., Walker, D., and Dalton, S. (2002). Pluripotent cell division cycles are driven by ectopic Cdk2, cyclin A/E and E2F activities. Oncogene *21*, 8320-8333.
Stewart, C.L., Kaspar, P., Brunet, L.J., Bhatt, H., Gadi, I., Kontgen, F., and Abbondanzo, S.J. (1992). Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. Nature *359*, 76-79.

Strumpf, D., Mao, C.A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., and Rossant, J. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. Development *132*, 2093-2102.
Suemori, H., Kadodawa, Y., Goto, K., Araki, I., Kondoh, H., and Nakatsuji, N. (1990). A mouse embryonic stem cell line showing pluripotency of differentiation in early embryos and ubiquitous beta-galactosidase expression. Cell differentiation and development : the official journal of the International Society of Developmental Biologists *29*, 181-186.
Sun, H., Lesche, R., Li, D.M., Liliental, J., Zhang, H., Gao, J., Gavrilova, N., Mueller, B., Liu, X., and Wu, H. (1999). PTEN modulates cell cycle progression and cell survival by regulating phosphatidylinositol 3,4,5,-trisphosphate and Akt/protein kinase B signaling pathway. Proceedings of the National Academy of Sciences of the United States of America *96*, 6199-6204.
Surani, M.A., Hayashi, K., and Hajkova, P. (2007). Genetic and epigenetic regulators of pluripotency. Cell *128*, 747-762.
Swingler, T.E., Bess, K.L., Yao, J., Stifani, S., and Jayaraman, P.S. (2004). The proline-rich homeodomain protein recruits members of the Groucho/Transducin-like enhancer of split protein family to co-repress transcription in hematopoietic cells. The Journal of biological chemistry *279*, 34938-34947.
Tabansky, I., Lenarcic, A., Draft, R.W., Loulier, K., Keskin, D.B., Rosains, J., Rivera-Feliciano, J., Lichtman, J.W., Livet, J., Stern, J.N.*, et al.* (2013). Developmental bias in cleavage-stage mouse blastomeres. Current biology : CB *23*, 21-31.
Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.
Takahashi, Y., Carpino, N., Cross, J.C., Torres, M., Parganas, E., and Ihle, J.N. (2003). SOCS3: an essential regulator of LIF receptor signaling in trophoblast giant cell differentiation. The EMBO journal *22*, 372-384.
Takeda, K., Noguchi, K., Shi, W., Tanaka, T., Matsumoto, M., Yoshida, N., Kishimoto, T., and Akira, S. (1997). Targeted disruption of the mouse Stat3 gene leads to early embryonic lethality. Proc Natl Acad Sci U S A *94*, 3801-3804.
Tam, P.P., and Beddington, R.S. (1987). The formation of mesodermal tissues in the mouse embryo during gastrulation and early organogenesis. Development *99*, 109-126.
Tanaka, S. (2006). Derivation and culture of mouse trophoblast stem cells in vitro. Methods Mol Biol *329*, 35-44.
Tanaka, S., Kunath, T., Hadjantonakis, A.K., Nagy, A., and Rossant, J. (1998). Promotion of trophoblast stem cell proliferation by FGF4. Science *282*, 2072-2075.
Tang, F., Barbacioru, C., Bao, S., Lee, C., Nordman, E., Wang, X., Lao, K., and 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.
Tang, F., Hajkova, P., Barton, S.C., Lao, K., and Surani, M.A. (2006). MicroRNA expression profiling of single whole embryonic stem cells. Nucleic acids research *34*, e9.
Tarkowski, A.K. (1959). Experiments on the development of isolated blastomers of mouse eggs. Nature *184*, 1286-1287.
Tarkowski, A.K. (1961). Mouse chimaeras developed from fused eggs. Nature *190*, 857-860.
Tesar, P.J. (2005). Derivation of germ-line-competent embryonic stem cell lines from preblastocyst mouse embryos. Proceedings of the National Academy of Sciences of the United States of America *102*, 8239-8244.
Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature *448*, 196-199.
Teves, S.S., and Henikoff, S. (2011). Heat shock reduces stalled RNA polymerase II and nucleosome turnover genome-wide. Genes & development *25*, 2387-2397.
Thomas, P., and Beddington, R. (1996). Anterior primitive endoderm may be responsible for patterning the anterior neural plate in the mouse embryo. Current biology : CB *6*, 1487-1496.
Thomas, P.Q., Brown, A., and Beddington, R.S. (1998). Hex: a homeobox gene revealing peri-implantation asymmetry in the mouse embryo and an early transient marker of endothelial cell precursors. Development *125*, 85-94.
Thomson, M., Liu, S.J., Zou, L.N., Smith, Z., Meissner, A., and Ramanathan, S. (2011). Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. Cell *145*, 875-889.

Tomioka, M., Nishimoto, M., Miyagi, S., Katayanagi, T., Fukui, N., Niwa, H., Muramatsu, M., and Okuda, A. (2002). Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex. Nucleic acids research *30*, 3202-3213.
Topcu, Z., Mack, D.L., Hromas, R.A., and Borden, K.L. (1999). The promyelocytic leukemia protein PML interacts with the proline-rich homeodomain protein PRH: a RING may link hematopoiesis and growth control. Oncogene *18*, 7091-7100.
Topisirovic, I., Culjkovic, B., Cohen, N., Perez, J.M., Skrabanek, L., and Borden, K.L. (2003). The proline-rich homeodomain protein, PRH, is a tissue-specific inhibitor of eIF4E-dependent cyclin D1 mRNA transport and growth. The EMBO journal *22*, 689-703.
Torres-Padilla, M.E., Parfitt, D.E., Kouzarides, T., and Zernicka-Goetz, M. (2007). Histone arginine methylation regulates pluripotency in the early mouse embryo. Nature *445*, 214-218.
Toyooka, Y., Shimosato, D., Murakami, K., Takahashi, K., and Niwa, H. (2008). Identification and characterization of subpopulations in undifferentiated ES cell culture. Development *135*, 909-918.
Trott, J., Hayashi, K., Surani, A., Babu, M.M., and Martinez-Arias, A. (2012). Dissecting ensemble networks in ES cell populations reveals micro-heterogeneity underlying pluripotency. Molecular bioSystems *8*, 744-752.
van der Laan, S., Tsanov, N., Crozet, C., and Maiorano, D. (2013). High Dub3 Expression in Mouse ESCs Couples the G1/S Checkpoint to Pluripotency. Molecular cell *52*, 366-379.
Vinot, S., Le, T., Ohno, S., Pawson, T., Maro, B., and Louvet-Vallee, S. (2005). Asymmetric distribution of PAR proteins in the mouse embryo begins at the 8-cell stage during compaction. Developmental biology *282*, 307-319.
Wang, Z., and Jaenisch, R. (2004). At most three ES cells contribute to the somatic lineages of chimeric mice and of mice produced by ES-tetraploid complementation. Developmental biology *275*, 192-201.
Ward Jr., J.H. (1963). Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association *58*, 236-244.
Ware, C.B., Horowitz, M.C., Renshaw, B.R., Hunt, J.S., Liggitt, D., Koblar, S.A., Gliniak, B.C., McKenna, H.J., Papayannopoulou, T., Thoma, B.*, et al.* (1995). Targeted disruption of the low-affinity leukemia inhibitory factor receptor gene causes placental, skeletal, neural and metabolic defects and results in perinatal death. Development *121*, 1283-1299.
Watanabe, S., Umehara, H., Murayama, K., Okabe, M., Kimura, T., and Nakano, T. (2006). Activation of Akt signaling is sufficient to maintain pluripotency in mouse and primate embryonic stem cells. Oncogene *25*, 2697-2707.
Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P.*, et al.* (2002). Initial sequencing and comparative analysis of the mouse genome. Nature *420*, 520-562.
Watson, A.J., and Kidder, G.M. (1988). Immunofluorescence assessment of the timing of appearance and cellular distribution of Na/K-ATPase during mouse embryogenesis. Developmental biology *126*, 80-90.
Weeks, D.L., and Melton, D.A. (1987). A maternal mRNA localized to the vegetal hemisphere in Xenopus eggs codes for a growth factor related to TGF-beta. Cell *51*, 861-867.
Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature *448*, 318-324.
Western, P., Maldonado-Saldivia, J., van den Bergen, J., Hajkova, P., Saitou, M., Barton, S., and Surani, M.A. (2005). Analysis of Esg1 expression in pluripotent cells and the germline reveals similarities with Oct4 and Sox2 and differences between human pluripotent cell lines. Stem Cells *23*, 1436-1442.
Williams, H., Jayaraman, P.S., and Gaston, K. (2008). DNA wrapping and distortion by an oligomeric homeodomain protein. Journal of molecular biology *383*, 10-23.
Williams, R.L., Hilton, D.J., Pease, S., Willson, T.A., Stewart, C.L., Gearing, D.P., Wagner, E.F., Metcalf, D., Nicola, N.A., and Gough, N.M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature *336*, 684-687.
Willson, T.A., Metcalf, D., and Gough, N.M. (1992). Cross-species comparison of the sequence of the leukaemia inhibitory factor gene and its protein. European journal of biochemistry / FEBS *204*, 21-30.
Wray, J., Kalkan, T., Gomez-Lopez, S., Eckardt, D., Cook, A., Kemler, R., and Smith, A. (2011). Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. Nature cell biology *13*, 838-845.

Wray, J., Kalkan, T., and Smith, A.G. (2010). The ground state of pluripotency. Biochemical Society transactions *38*, 1027-1032.
Wu, G., Gentile, L., Fuchikami, T., Sutter, J., Psathaki, K., Esteves, T.C., Arauzo-Bravo, M.J., Ortmeier, C., Verberk, G., Abe, K.*, et al.* (2010). Initiation of trophectoderm lineage specification in mouse embryos is independent of Cdx2. Development *137*, 4159-4169.
Xu, X., Weinstein, M., Li, C., Naski, M., Cohen, R.I., Ornitz, D.M., Leder, P., and Deng, C. (1998). Fibroblast growth factor receptor 2 (FGFR2)-mediated reciprocal regulation loop between FGF8 and FGF10 is essential for limb induction. Development *125*, 753-765.
Yadav, N., Lee, J., Kim, J., Shen, J., Hu, M.C., Aldaz, C.M., and Bedford, M.T. (2003). Specific protein methylation defects and gene expression perturbations in coactivator-associated arginine methyltransferase 1-deficient mice. Proceedings of the National Academy of Sciences of the United States of America *100*, 6464-6468.
Yagi, R., Kohn, M.J., Karavanova, I., Kaneko, K.J., Vullhorst, D., DePamphilis, M.L., and Buonanno, A. (2007). Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development. Development *134*, 3827-3836.
Yagita, K., Horie, K., Koinuma, S., Nakamura, W., Yamanaka, I., Urasaki, A., Shigeyoshi, Y., Kawakami, K., Shimada, S., Takeda, J.*, et al.* (2010). Development of the circadian oscillator during differentiation of mouse embryonic stem cells in vitro. Proceedings of the National Academy of Sciences of the United States of America *107*, 3846-3851.
Yamanaka, Y., Lanner, F., and Rossant, J. (2010). FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. Development *137*, 715-724.
Yang, V.S., Carter, S.A., Hyland, S.J., Tachibana-Konwalski, K., Laskey, R.A., and Gonzalez, M.A. (2011). Geminin escapes degradation in G1 of mouse pluripotent cells and mediates the expression of Oct4, Sox2, and Nanog. Current biology : CB *21*, 692-699.
Ye, S., Li, P., Tong, C., and Ying, Q.L. (2013). Embryonic stem cell self-renewal pathways converge on the transcription factor Tfcp2l1. EMBO J *32*, 2548-2560.
Ying, Q.L., Nichols, J., Chambers, I., and 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., and Smith, A. (2003b). Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. Nat Biotechnol *21*, 183-186.
Ying, Q.L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519-523.
Ying, Y., Liu, X.M., Marble, A., Lawson, K.A., and Zhao, G.Q. (2000). Requirement of Bmp8b for the generation of primordial germ cells in the mouse. Mol Endocrinol *14*, 1053-1063.
Ying, Y., and Zhao, G.Q. (2001). Cooperation of endoderm-derived BMP2 and extraembryonic ectoderm-derived BMP4 in primordial germ cell generation in the mouse. Developmental biology *232*, 484-492.
Zalzman, M., Falco, G., Sharova, L.V., Nishiyama, A., Thomas, M., Lee, S.L., Stagg, C.A., Hoang, H.G., Yang, H.T., Indig, F.E.*, et al.* (2010). Zscan4 regulates telomere elongation and genomic stability in ES cells. Nature *464*, 858-863.
Zamparini, A.L., Watts, T., Gardner, C.E., Tomlinson, S.R., Johnston, G.I., and Brickman, J.M. (2006). Hex acts with beta-catenin to regulate anteroposterior patterning via a Groucho-related co-repressor and Nodal. Development *133*, 3709-3722.
Zappone, M.V., Galli, R., Catena, R., Meani, N., De Biasi, S., Mattei, E., Tiveron, C., Vescovi, A.L., Lovell-Badge, R., Ottolenghi, S.*, et al.* (2000). Sox2 regulatory sequences direct expression of a (beta)-geo transgene to telencephalic neural stem cells and precursors of the mouse embryo, revealing regionalization of gene expression in CNS stem cells. Development *127*, 2367-2382.
Zhang, P., Andrianakos, R., Yang, Y., Liu, C., and Lu, W. (2010). Kruppel-like factor 4 (Klf4) prevents embryonic stem (ES) cell differentiation by regulating Nanog gene expression. The Journal of biological chemistry *285*, 9180-9189.
Zhang, X., Neganova, I., Przyborski, S., Yang, C., Cooke, M., Atkinson, S.P., Anyfantis, G., Fenyk, S., Keith, W.N., Hoare, S.F.*, et al.* (2009). A role for NANOG in G1 to S transition in human embryonic stem cells through direct binding of CDK6 and CDC25A. The Journal of cell biology *184*, 67-82.
Zhao, B., Wei, X., Li, W., Udan, R.S., Yang, Q., Kim, J., Xie, J., Ikenoue, T., Yu, J., Li, L.*, et al.* (2007). Inactivation of YAP oncoprotein by the Hippo pathway is involved in cell contact inhibition and tissue growth control. Genes & development *21*, 2747-2761.

Zwaka, T.P., and Thomson, J.A. (2005). A germ cell origin of embryonic stem cells? Development *132*, 227-233.

SUPPLEMENTARY INFORMATION

Supplementary Table 1. Downregulated GRO-seq gene clusters (related to Chapter 5). Genes were identified as downregulated between the HV^-S^+ and HV^+S^+ ES cell populations and the HV^+S^- differentiated endoderm population using the edgeR statistical package. Genes were then clustered using the Ward methodology using R software. The table lists the RefSeq mRNA ID of all genes within each of the 5 clusters of downregulated genes.

CLUSTER	GENE ID	CLUSTER	GENE ID	CLUSTER	GENE ID
1	NM_008052	3	NM_175093	4	NM_178421
1	NM_001033226	3	NM_001099635	4	NM_144518
1	NM_173862	3	NM_025330	4	NM_007467
1	NM_009262	3	NM_001113478	4	NM_013627
1	NM_001001796	3	NM_175678	4	NM_021408
1	NM_010208	3	NM_027091	4	NM_181401
1	NM_001081283	3	NM_028946	4	NM_028930
1	NM_198656	3	NM_024291	4	NM_001042768
1	NM_144538	3	NM_001134385	4	NM_008934
1	NM_145510	3	NM_001134386	4	NM_001039364
1	NM_175535	3	NM_027965	4	NM_008614
1	NM_053071	3	NM_007715	4	NM_177823
1	NM_130457	3	NM_019703	4	NM_008108
1	NM_016754	3	NM_023598	4	NM_022018
1	NM_015826	3	NM_007905	4	NR_002860
1	NM_008921	3	NM_001042623	4	NM_028387
1	NM_001114662	3	NM_028749	4	NM_029896
1	NM_178936	3	NM_027839	4	NM_053079
1	NM_011771	3	NM_174988	4	NM_016672
1	NM_028034	3	NM_009146	4	NM_026120
1	NM_008261	3	NM_001007573	4	NM_001039644
1	NM_029916	3	NM_001033399	4	NR_004443
1	NM_027921	3	NM_007908	4	NM_172914
1	NM_021530	3	NM_145890	4	NM_001126046
1	NM_008001	3	NM_175503	4	NM_019397
1	NM_172613	3	NM_013584	4	NM_175433
1	NM_008664	3	NM_174991	4	NM_020581
1	NM_183126	3	NM_019501	4	NM_030718
1	NM_020332	3	NM_019971	4	NM_026263
1	NM_001001446	3	NM_146108	4	NM_172436
1	NM_020622	3	NM_007958	4	NM_008884
1	NM_183115	3	NM_012011	4	NM_178087
1	NM_007741	3	NM_001039233	4	NM_007390
1	NM_009593	3	NM_007532	4	NM_153578

1	NM_007506	3	NM_009201	4	NM_133365
1	NM_183023	3	NM_008767	4	NM_146028
1	NM_021511	3	NM_054085	4	NM_172385
1	NM_174854	3	NM_010512	4	NM_016750
1	NM_174853	3	NM_001111275	4	NM_198860
1	NM_029440	3	NM_022016	4	NM_010220
1	NM_178381	3	NM_001004184	4	NM_031388
1	NM_033037	3	NM_199016	4	NM_009176
1	NM_033610	3	NM_001082414	4	NM_011819
1	NM_010826	3	NM_134109	4	NM_028709
1	NM_001037712	3	NM_001033954	4	NM_181541
1	NM_001033304	3	NM_008153	4	NM_178726
1	NM_080575	3	NM_010917	4	NM_178766
1	NM_011973	3	NM_153097	4	NM_172765
1	NM_008811	3	NM_178714	4	NM_001115130
1	NM_001099779	3	NM_010134	4	NM_020577
1	NM_133888	3	NM_022721	4	NM_175407
1	NM_008400	3	NM_001042659	4	NM_011576
1	NM_001001176	3	NM_134092	4	NM_080856
1	NM_172378	3	NM_001039209	4	NM_001024141
1	NM_011876	3	NM_001015046	4	NM_030697
1	NM_145142	3	NM_144855	4	NM_008652
1	NM_010246	3	NM_178224	4	NM_183109
1	NM_001025585	3	NM_001042727	4	NM_028937
1	NM_145394	3	NM_001033419	4	NM_016687
1	NM_011287	3	NM_144795	4	NM_178672
1	NM_010606	3	NM_009337	4	NM_009646
1	NM_001112731	3	NM_029897	4	NM_001081209
1	NM_007592	3	NM_001109688	4	NM_177123
1	NM_001033338	3	NM_029081	4	NM_001030307
1	NM_010062	3	NM_001033344	4	NM_026281
1	NM_001081134	3	NM_177742	4	NM_022316
1	NM_145841	3	NM_001042760	4	NM_016743
1	NM_007378	3	NM_020599	4	NM_172737
1	NM_008616	3	NM_001039198	4	NM_013690
1	NM_001001792	3	NM_177678	4	NM_008803
1	NM_139197	3	NM_001004365	4	NM_172818
1	NM_023816	3	NM_172723	4	NM_001083894
1	NM_028057	3	NM_028012	4	NM_010686
1	NM_001034894	3	NM_008938	4	NM_001033302
1	NM_026480	3	NM_054053	4	NM_153404
1	NM_001029912	3	NM_011308	4	NM_019679
1	NM_001038887	3	NM_031159	4	NM_001077698

1	NM_008586	3	NM_001134391	4	NM_026483
1	NM_029066	3	NM_024474	4	NM_001033488
1	NM_010201	3	NM_148931	4	NM_009624
1	NM_008584	3	NM_028576	4	NM_172872
1	NM_001003893	3	NM_177328	4	NM_016806
1	NM_001039484	3	NM_001114660	4	NM_016762
1	NM_011443	3	NM_031384	4	NM_172923
1	NM_133754	3	NM_009822	4	NM_182808
1	NM_011158	3	NM_001115075	4	NM_182650
1	NM_019945	3	NM_001122733	4	NM_177003
1	NM_001033186	3	NM_021099	4	NM_026524
1	NM_015767	3	NM_001109761	4	NM_138595
1	NM_001039146	3	NM_027308	4	NM_007868
1	NM_001033350	3	NM_017479	4	NM_178269
1	NM_029422	3	NM_019864	4	NM_025639
1	NM_001025584	3	NM_178845	4	NM_145536
1	NM_172510	3	NM_001081139	4	NM_010097
1	NM_011638	3	NM_144953	4	NM_009199
1	NM_008262	3	NM_172575	4	NM_001081127
1	NM_001042612	3	NM_029441	4	NM_001026214
1	NM_183146	3	NM_173737	4	NM_007647
1	NM_031396	3	NM_009814	4	NM_028797
1	NM_010406	3	NM_010827	4	NM_177618
1	NM_001085509	3	NM_010059	4	NM_138304
1	NM_008625	3	NM_001103158	4	NM_001102468
1	NM_029920	3	NM_025861	4	NM_145402
1	NM_175296	3	NM_001080928	4	NM_008866
1	NM_008432	3	NM_181548	4	NM_011143
1	NM_146011	3	NM_001033378	4	NM_172898
1	NM_207523	3	NM_011535	4	NM_145634
1	NM_080470	3	NM_198052	4	NM_008924
1	NM_010924	3	NM_001081220	4	NM_026514
1	NM_027678	3	NM_011465	4	NM_138628
1	NM_010142	3	NM_028610	4	NM_016789
1	NM_007595	3	NM_001018002	4	NM_013486
1	NM_146123	3	NM_013454	4	NM_001080158
1	NM_013495	3	NM_025950	4	NM_025840
1	NM_027504	3	NM_013612	4	NM_175562
1	NM_001033403	3	NM_029044	4	NM_133720
1	NM_016774	3	NM_198022	4	NM_177352
1	NM_008957	3	NM_153179	4	NM_028603
1	NM_172796	3	NM_001011732	4	NM_011762
1	NM_013496	3	NM_011635	4	NM_001083906

1	NM_031260	3	NM_023146	4	NM_009556
1	NM_023871	3	NM_001013802	4	NM_001006676
1	NM_177802	3	NM_009539	4	NM_026668
1	NM_009723	3	NM_153166	4	NM_001033264
1	NM_146256	3	NM_080644	4	NM_001013372
1	NM_001036684	3	NM_008815	4	NM_177882
1	NM_001004154	3	NM_175651	4	NM_198962
1	NM_017397	3	NM_024244	4	NM_001037926
1	NM_010487	3	NM_028804	4	NM_031165
1	NM_199055	3	NM_175488	4	NM_172784
1	NM_153513	3	NM_175328	4	NM_172805
1	NM_007583	3	NM_177293	4	NM_145078
1	NM_029341	3	NM_133997	4	NM_015820
1	NM_023755	3	NM_029494	4	NM_010904
1	NM_031180	3	NM_013465	4	NM_013743
1	NM_024406	3	NM_173030	4	NM_001006677
1	NM_011217	3	NM_001135152	4	NM_027482
1	NM_153127	3	NM_007515	4	NM_178621
1	NM_010830	3	NM_001135151	4	NM_025448
1	NM_001113734	3	NM_144808	4	NM_011517
1	NM_001099674	3	NM_029570	4	NM_007414
1	NM_001081250	3	NM_175164	4	NM_172916
1	NM_173386	3	NM_011138	4	NM_001102404
1	NR_003270	3	NM_178915	4	NM_001085555
1	NM_010296	3	NM_001081346	4	NM_019869
1	NM_152895	3	NM_028230	4	NM_007601
1	NM_001033865	3	NM_018773	4	NM_145448
1	NM_024277	3	NM_032000	4	NM_013687
1	NM_007648	3	NM_023844	4	NM_011926
1	NM_212483	3	NM_008081	4	NM_001039186
1	NM_011010	3	NM_008892	4	NM_001039187
1	NM_134130	3	NM_030889	4	NM_001039185
1	NM_001134646	3	NM_028760	4	NM_018810
1	NM_001099675	3	NM_008382	4	NM_001033321
1	NM_172763	3	NM_001033454	4	NM_008124
1	NM_028602	3	NM_144787	4	NM_009767
1	NM_033571	3	NM_001033452	4	NM_007977
1	NM_030703	3	NM_001024468	4	NM_007388
1	NM_010202	3	NM_009016	4	NM_008846
1	NM_021453	3	NM_021878	4	NM_011934
1	NM_172451	3	NM_007936	4	NM_027334
1	NM_144520	3	NM_001045543	4	NM_011436
1	NM_030052	3	NM_019811	4	NM_007919

1	NM_011027	3	NM_019439	4	NM_146125
1	NM_001038839	3	NM_001007596	4	NM_010021
1	NM_007789	3	NM_013906	4	NM_199257
1	NM_001099296	3	NM_013722	4	NM_177811
1	NM_028182	3	NM_010814	4	NM_145827
1	NM_183024	3	NM_007587	4	NM_001039536
1	NM_021542	3	NM_007701	4	NM_009950
1	NM_026489	3	NM_001033043	4	NM_172513
1	NM_001038845	3	NM_008169	4	NM_172896
1	NM_008234	3	NM_010816	4	NM_010486
1	NM_133204	3	NM_007819	4	NM_207686
1	NR_003546	3	NM_001038609	4	NM_030709
1	NM_001007584	3	NM_010838	4	NM_029784
1	NM_016865	3	NM_146064	4	NM_207685
1	NM_021022	3	NM_198601	4	NM_146074
1	NM_146100	3	NM_008782	4	NM_029163
1	NM_001033542	3	NM_001034013	4	NM_175290
1	NM_009446	3	NM_009765	4	NM_008639
1	NM_001081125	3	NM_001081001	4	NM_026257
1	NM_010324	3	NM_001013774	4	NM_181545
1	NM_001081435	3	NM_015790	4	NM_028945
1	NM_178935	3	NM_010074	4	NM_001114879
1	NM_025995	3	NM_001010973	4	NM_001017525
1	NM_007731	3	NM_007384	4	NM_028964
1	NM_008280	3	NM_025495	4	NM_197985
1	NM_175111	3	NM_008596	4	NM_027957
1	NM_025363	3	NM_001040400	4	NM_027649
1	NM_026630	3	NM_001111145	4	NM_021308
1	NM_028028	3	NM_148958	4	NM_001102471
1	NM_001078167	3	NM_148942	4	NM_033569
1	NM_173374	3	NM_001077354	4	NM_029372
1	NM_001007583	3	NM_009889	4	NM_028231
1	NM_172522	3	NM_010636	4	NM_176844
1	NM_001077499	3	NM_012055	4	NM_001081206
1	NM_001110193	3	NM_001024138	4	NM_033601
1	NM_001110192	3	NM_022563	4	NM_029660
1	NM_010566	3	NM_007804	4	NM_001015039
1	NM_001134399	3	NM_172287	4	NM_172892
1	NM_027426	3	NM_144942	4	NM_019500
1	NM_001008424	3	NM_001081064	4	NM_001081104
1	NM_011242	3	NM_175628	4	NM_009855
1	NM_001003666	3	NM_009559	4	NM_001024837
1	NM_177764	3	NM_011858	4	NR_021486

1	NM_172133	3	NM_201371	4	NM_130895
1	NM_028053	3	NM_011932	4	NR_004429
1	NM_145820	3	NM_172951	4	NM_001033405
1	NM_145155	3	NM_011891	4	NM_178689
1	NM_201353	3	NM_015739	4	NM_001081025
1	NM_023517	3	NM_133900	4	NM_001080793
1	NM_007413	3	NM_011670	4	NM_027749
1	NM_080465	3	NM_023328	4	NM_032002
1	NM_001081178	3	NM_001085505	4	NM_001013784
1	NM_001102411	3	NM_008741	4	NM_146073
1	NM_001102412	3	NM_146198	4	NM_009228
1	NM_024175	3	NM_175003	4	NM_173866
1	NM_011291	3	NM_001099288	4	NM_008855
1	NM_009229	3	NM_001033439	4	NM_172907
1	NM_001081257	3	NM_178098	4	NM_025541
1	NM_001001160	3	NM_145470	4	NM_032610
1	NM_011708	3	NM_145823	4	NM_030706
1	NM_020253	3	NM_019800	4	NM_177753
1	NM_001100394	3	NM_134005	4	NM_033374
1	NM_013669	3	NM_011652	4	NM_001039652
1	NM_001100608	3	NM_028004	4	NM_144936
1	NM_031156	3	NM_024185	4	NM_139148
1	NM_198702	3	NM_011741	4	NM_001033191
1	NM_053261	3	NM_010197	4	NM_001039103
1	NM_023125	3	NM_027238	4	NM_133914
1	NM_001081663	3	NM_207624	4	NM_001001650
1	NM_010953	3	NM_027241	4	NM_030046
1	NM_001081176	3	NM_011518	4	NM_029619
1	NM_009421	3	NM_019677	4	NM_012054
1	NM_172258	3	NM_172578	4	NM_008952
1	NM_029248	3	NM_013456	4	NM_133197
1	NM_001033597	3	NM_001081012	4	NM_008055
1	NM_144823	3	NM_029736	4	NM_133220
1	NM_010600	3	NM_199015	4	NM_177547
1	NM_001038607	3	NM_010262	4	NM_010168
1	NM_001033598	3	NM_198412	4	NM_021876
1	NM_028060	3	NM_001013745	4	NM_015828
1	NM_009663	3	NM_011811	4	NM_022722
1	NM_177261	3	NM_011542	4	NM_198618
1	NM_027261	3	NM_007430	4	NM_027455
1	NM_001008420	3	NM_001014399	4	NM_009258
1	NR_002896	3	NM_001014424	4	NM_026408
1	NM_198306	3	NM_178790	4	NM_007702

1	NM_146201	3	NM_001014422	4	NM_022411
1	NM_001002894	3	NM_001014423	4	NM_017461
1	NM_133982	3	NM_008861	4	NM_172150
1	NM_009084	3	NM_009401	4	NM_023784
1	NM_172517	3	NM_029747	4	NM_011846
1	NM_026940	3	NM_019535	4	NM_139293
1	NM_145356	3	NM_178617	4	NM_177942
1	NM_016964	3	NM_178259	4	NM_008816
1	NM_009292	3	NM_053122	4	NM_001032378
1	NM_144854	3	NM_019430	4	NM_023059
1	NM_177600	3	NM_009533	4	NM_025610
1	NM_172658	3	NM_029942	4	NM_011110
1	NM_011541	3	NM_013762	4	NM_146250
1	NM_001083342	3	NM_199195	4	NM_177857
1	NM_007431	3	NM_010565	4	NR_024051
1	NM_145209	3	NM_178378	4	NM_010189
1	NM_010601	3	NM_007940	4	NM_011374
1	NM_025300	3	NM_011898	4	NM_175236
1	NM_175329	3	NM_026262	4	NM_008247
1	NM_027941	3	NM_001128625	4	NM_001039114
1	NM_001004153	3	NM_001080927	4	NM_023529
1	NM_021463	3	NM_011780	4	NM_175641
1	NM_172307	3	NM_009035	4	NM_001113549
1	NM_011987	3	NM_019926	4	NM_008105
1	NM_177204	3	NM_145838	4	NM_198635
1	NM_001037740	3	NM_025285	4	NM_001005341
1	NM_008722	3	NM_028296	4	NM_007877
1	NM_001104642	3	NM_146086	4	NM_001006674
1	NM_175511	3	NM_001037937	4	NM_001006668
1	NM_201374	3	NM_011157	4	NM_001006669
1	NM_020282	3	NM_177343	4	NM_001003824
1	NM_011323	3	NM_027871	4	NM_001003825
1	NM_172822	3	NM_001038624	4	NM_008903
1	NM_177393	3	NM_198420	4	NM_019388
1	NM_028892	3	NM_013512	4	NM_010898
1	NM_001033599	3	NM_080457	4	NM_008511
1	NM_001034059	3	NM_010137	4	NM_011107
1	NM_001005508	3	NM_177363	4	NM_028829
1	NR_002840	3	NM_028944	4	NM_001033759
1	NM_009308	3	NM_001081383	4	NM_031997
1	NM_008011	3	NM_011139	4	NM_080467
1	NM_011358	3	NM_001126487	4	NM_019459
1	NM_011569	3	NM_021344	4	NM_001122954

1	NM_001039365	3	NM_178780	4	NM_001006675
1	NM_019647	3	NM_026458	4	NM_173446
1	NM_146017	3	NM_145953	4	NM_021495
1	NM_013806	3	NM_026594	4	NM_183161
1	NM_011516	3	NM_019445	4	NM_028199
1	NM_029794	3	NM_177845	4	NM_177167
1	NM_027227	3	NM_008958	4	NM_001037759
1	NM_011295	3	NM_013825	4	NM_008604
1	NM_007749	3	NM_009598	4	NM_023887
1	NM_016723	3	NM_023279	4	NM_183199
1	NM_183278	3	NM_013748	4	NM_010900
1	NM_144804	3	NM_007956	4	NM_025768
1	NM_026682	3	NM_175484	4	NM_026054
1	NM_011587	3	NM_028844	4	NM_007918
1	NM_010356	3	NM_010889	4	NM_011925
1	NM_001077353	3	NM_178440	4	NM_009721
1	NM_175482	3	NM_009484	4	NM_178728
1	NM_201365	3	NM_178415	4	NM_177420
1	NM_145441	3	NM_181316	4	NM_001083121
1	NM_015802	3	NM_029238	4	NM_001083120
1	NM_199024	3	NM_027452	4	NM_008680
1	NM_013731	3	NM_139292	4	NM_010135
1	NM_001033472	3	NM_023716	4	NM_009769
2	NM_001083587	3	NM_198610	4	NM_010357
2	NM_175263	3	NM_030255	4	NM_133999
2	NM_053214	3	NM_031867	4	NM_027268
2	NM_008437	3	NM_201366	4	NM_026763
2	NM_019688	3	NM_020259	4	NM_026233
2	NM_173419	3	NM_001048176	4	NM_019448
2	NM_024271	3	NM_009065	4	NM_001130184
2	NM_029170	3	NM_178256	4	NM_183096
2	NM_016745	3	NM_020026	4	NM_027285
2	NM_011546	3	NM_011372	4	NM_023635
2	NM_175276	3	NM_172830	4	NM_177192
2	NM_177829	3	NM_198300	4	NM_001122639
2	NM_001037906	3	NM_019576	4	NM_009109
2	NM_001033382	3	NM_153589	4	NM_009250
2	NM_019950	3	NM_019641	4	NM_010611
2	NM_008712	3	NM_177216	4	NM_008772
2	NR_015351	3	NM_029525	4	NM_198408
2	NM_011613	3	NM_001033636	4	NM_028838
2	NM_028783	3	NM_008624	4	NM_001085529
2	NM_001113210	3	NM_027288	4	NM_008638

2	NM_008687	3	NM_011487	4	NM_026301
2	NM_001113209	3	NM_009570	4	NM_145835
2	NM_001033367	3	NM_177652	4	NM_028157
2	NM_145123	3	NM_021470	4	NM_029112
2	NM_001013749	3	NM_173780	4	NM_001111119
2	NM_178738	3	NM_012008	4	NM_177905
2	NM_019413	3	NM_175481	4	NM_028953
2	NM_211138	3	NM_029620	4	NM_027407
2	NM_001085521	3	NM_007545	4	NM_001007221
2	NM_010401	3	NM_172453	4	NM_001007220
2	NM_001102607	3	NM_001083902	4	NM_010243
2	NM_145561	3	NM_027539	4	NM_145467
2	NM_080288	3	NM_007513	4	NM_013827
2	NM_001113238	3	NM_008830	4	NM_001081695
2	NM_009506	3	NM_134251	4	NM_133977
2	NM_027689	3	NM_010425	4	NM_001111026
2	NM_178666	3	NM_020295	4	NM_153514
2	NM_025891	3	NM_173763	4	NM_001081027
2	NM_001024703	3	NR_003188	4	NM_080446
2	NM_177546	3	NM_153386	4	NM_001033669
2	NM_172515	3	NM_153385	4	NM_023508
2	NM_183289	3	NM_153384	4	NM_029458
2	NM_008804	3	NM_172496	4	NM_177086
2	NM_013569	3	NM_001081129	4	NM_175017
2	NR_003492	3	NM_145227	4	NM_001111059
2	NM_009233	3	NM_010279	4	NM_133654
2	NM_177368	3	NM_007765	4	NM_021390
2	NM_008713	3	NM_178656	4	NM_028615
2	NM_011978	3	NM_019779	4	NM_027206
2	NM_178634	3	NM_207649	4	NM_001111027
2	NM_153409	3	NM_172813	4	NM_182928
2	NM_008480	3	NM_172449	4	NM_028913
2	NM_029002	3	NM_019631	4	NM_001006678
2	NM_172563	3	NM_028719	4	NM_001006679
2	NM_001101588	3	NM_013661	4	NM_001006680
2	NM_001004148	3	NM_172870	4	NM_008087
2	NM_178877	3	NM_178662	4	NM_009762
2	NM_001025576	3	NM_001039000	4	NM_023879
2	NM_001024918	3	NM_134090	4	NM_178667
2	NM_011040	3	NM_177597	4	NM_009036
2	NM_178681	3	NM_001081342	4	NM_001098225
2	NM_027902	3	NM_001136058	4	NM_008602
2	NM_009008	3	NM_207255	4	NM_001081074

2	NM_001033538	3	NM_153543	4	NM_028916
2	NM_001080820	3	NM_028679	4	NM_033077
2	NM_001025382	3	NM_016803	4	NM_053171
2	NM_173767	3	NM_028870	4	NM_010881
2	NM_001103156	3	NM_028039	4	NM_001007591
2	NM_001103157	3	NM_025429	4	NM_001034098
2	NM_001033266	3	NM_008447	4	NM_001034097
2	NM_019588	3	NM_173052	4	NM_011076
2	NM_029967	3	NM_001025163	4	NM_009622
2	NM_009122	3	NM_001112805	4	NM_016919
2	NM_013811	3	NM_177888	4	NM_001111274
2	NM_001081351	3	NM_009417	4	NM_001111276
2	NM_148943	3	NM_144819	4	NM_030141
2	NM_001113514	3	NM_177233	4	NM_133219
2	NM_178703	3	NM_032465	4	NM_023873
2	NM_173016	3	NM_021292	4	NM_010657
2	NM_010063	3	NM_028826	5	NM_011795
2	NM_133721	3	NM_001113198	5	NM_019508
2	NM_026138	3	NM_013834	5	NM_024230
2	NM_172885	3	NM_175498	5	NM_153547
2	NM_011855	3	NM_001085522	5	NM_010848
2	NM_013737	3	NM_177200	5	NM_011889
2	NM_009869	3	NM_181595	5	NM_146126
2	NM_021374	3	NM_011491	5	NM_013602
2	NM_010128	3	NR_001463	5	NM_021286
2	NM_010014	3	NR_001570	5	NM_028903
2	NM_153599	3	NM_177193	5	NM_007586
2	NM_177259	3	NM_010187	5	NM_010464
2	NM_008744	3	NM_001077189	5	NM_001033468
3	NM_183171	3	NM_001081208	5	NM_008509
3	NM_009327	3	NM_130886	5	NM_010493
3	NM_001113386	3	NM_148935	5	NM_001005426
3	NM_198967	3	NM_139303	5	NM_175293
3	NM_001034873	3	NM_007897	5	NM_010077
3	NM_001034893	3	NM_172583	5	NM_009581
3	NM_001013779	3	NM_144799	5	NM_023113
3	NM_026183	3	NM_021480	5	NM_029993
3	NM_018872	3	NM_001037915	5	NM_001009948
3	NM_009477	3	NM_175291	5	NM_145570
3	NM_011173	3	NM_008137	5	NM_023258
3	NM_172450	3	NM_029953	5	NM_023063
3	NM_019738	3	NM_001039231	5	NM_009482
3	NM_175238	3	NM_018732	5	NM_001033162

3	NM_001114545	3	NM_008601	5	NM_029755
3	NM_001029868	4	NM_145569	5	NM_001100177
3	NM_175460	4	NM_023868	5	NM_009989
3	NM_009904	4	NM_152810	5	NM_029045
3	NM_013566	4	NM_001014397	5	NM_011148
3	NM_009667	4	NM_022004	5	NM_001008419
3	NM_011614	4	NM_030676	5	NM_025807
3	NM_146010	4	NM_001105252	5	NM_011996
3	NM_001130186	4	NM_001033477	5	NM_001113545
3	NM_021889	4	NM_021788	5	NM_178593
3	NM_027671	4	NM_029444	5	NM_001038846
3	NM_021610	4	NM_198884	5	NM_001033253
3	NM_028236	4	NM_175525	5	NM_001013771
3	NM_183182	4	NM_023680	5	NM_172795
3	NM_001130185	4	NM_001080809	5	NM_001014394
3	NM_010407	4	NM_007981	5	NM_145153
3	NM_013680	4	NM_181853	5	NM_001025590
3	NM_001110780	4	NM_015798	5	NM_133819
3	NM_009263	4	NM_001005423	5	NM_144921
3	NM_022026	4	NM_001081124	5	NM_001029988
3	NM_001039586	4	NM_001033336	5	NM_177081
3	NM_174846	4	NM_133167	5	NM_026147
3	NM_180678	4	NM_146144	5	NM_146069
3	NM_020005	4	NM_001113566	5	NM_207667
3	NM_172303	4	NM_001113565	5	NM_177310
3	NM_001101546	4	NM_001113564	5	NM_008575
3	NM_009752	4	NM_025814	5	NM_023892
3	NR_002844	4	NM_146148	5	NM_001080707
3	NM_007470	4	NM_025693	5	NM_026316
3	NM_026929	4	NM_001042767	5	NM_144805
3	NM_177158	4	NM_026030	5	NM_001008423
3	NM_175213	4	NM_183307	5	NM_001033446
3	NM_008381	4	NM_001098723	5	NM_028025
3	NM_177376				

Supplementary Table 2. Upregulated GRO-seq gene clusters (related to Chapter 5). Genes were identified as upregulated between the $HV^{-}S^{+}$ and $HV^{+}S^{+}$ ES cell populations and the $HV^{+}S^{-}$ differentiated endoderm population using the edgeR statistical package. Genes were then clustered using the Ward methodology using R software. The table lists the RefSeq mRNA ID of all genes within each of the 5 clusters of upregulated genes.

CLUSTER	GENE ID	CLUSTER	GENE ID	CLUSTER	GENE ID
1	NM_001025192	1	NM_013693	1	NM_009565
1	NM_009673	1	NM_001017426	1	NM_054042
1	NM_029586	1	NM_009755	1	NM_173395
1	NM_145688	1	NM_026381	1	NM_145619
1	NM_172267	1	NM_001048207	1	NM_033144
1	NR_003243	1	NM_053099	1	NM_028207
1	NM_027626	1	NM_175512	1	NM_010403
1	NM_027219	1	NM_010871	1	NM_007986
1	NM_029832	1	NM_170758	1	NM_145933
1	NM_030228	1	NM_021545	1	NM_013918
1	NM_054044	1	NM_029614	1	NM_010346
1	NM_152813	1	NM_008391	1	NM_178405
1	NM_007400	1	NM_021521	1	NM_007603
1	NM_008859	1	NM_027998	1	NM_026446
1	NM_001111053	1	NM_011803	1	NM_172524
1	NM_016866	1	NM_008541	1	NM_025282
1	NM_134062	1	NM_176933	1	NM_009395
1	NM_029653	1	NM_173739	1	NM_176942
1	NM_144821	1	NM_001007580	1	NM_001122953
1	NM_001009573	1	NM_001039485	1	NM_010905
1	NM_011416	1	NM_009497	1	NM_028788
1	NM_007752	1	NM_011066	1	NM_009741
1	NM_178214	1	NM_030708	1	NM_008783
1	NM_009794	1	NM_010720	1	NM_144853
1	NM_130888	1	NM_011808	1	NM_198960
1	NM_178653	1	NM_001038642	1	NM_009520
1	NM_130867	1	NM_001045525	1	NM_007873
1	NM_001034892	1	NM_007746	1	NM_011502
1	NM_030728	1	NM_010259	1	NM_001033227
1	NM_183000	1	NM_145978	1	NM_138672
1	NM_008092	1	NM_011845	1	NM_144839
1	NM_053252	1	NM_010870	1	NM_183355
1	NM_001114596	1	NM_028493	1	NM_133789
1	NM_001114595	1	NM_001033339	1	NM_001039878

1	NM_001114597	1	NM_019566	1	NM_026780
1	NM_013870	1	NM_001033531	1	NM_199470
1	NM_178637	1	NM_031170	1	NM_011513
1	NM_145732	1	NM_015775	1	NM_001033908
1	NM_001109764	1	NM_007778	1	NM_010495
1	NM_009819	1	NM_001113529	1	NM_021409
1	NM_001008700	1	NM_001139519	1	NM_172767
1	NM_009425	1	NM_001113530	1	NM_027543
1	NM_144560	1	NM_013640	1	NM_057171
1	NM_172414	1	NM_018884	1	NM_001037928
1	NM_028303	1	NM_001008231	1	NM_029098
1	NM_146119	1	NM_001115154	1	NM_031868
1	NM_019978	1	NM_007811	1	NM_009922
1	NM_026840	1	NM_030711	1	NM_018832
1	NM_016736	1	NM_021452	1	NM_207237
1	NM_001033773	1	NM_027835	1	NM_010728
1	NM_009390	1	NM_008498	1	NM_009021
1	NM_010386	1	NM_013780	1	NM_001122952
1	NM_001024508	1	NM_177694	1	NM_001005784
1	NM_001048229	1	NM_027468	1	NM_009660
1	NM_021050	1	NM_024269	1	NM_029604
1	NM_001083316	1	NM_177306	1	NM_016850
1	NM_145515	1	NM_011783	1	NM_021447
1	NM_178020	1	NM_009197	1	NM_009277
1	NM_001110783	1	NM_001039154	1	NM_001082552
1	NM_010452	1	NM_007667	1	NM_010068
1	NM_021384	1	NM_175483	1	NM_001003961
1	NM_001085373	1	NM_001083912	1	NM_001003960
1	NM_010225	1	NM_138752	1	NM_001003963
1	NM_008872	1	NM_172572	1	NM_001122997
1	NM_028054	1	NM_008923	1	NM_007559
1	NM_033322	1	NM_133807	1	NM_008140
1	NM_134437	1	NM_027910	1	NM_008878
1	NM_001081416	1	NM_001126047	1	NM_001076554
1	NM_153581	1	NM_173007	1	NM_008032
1	NM_001042611	1	NM_139151	1	NM_001081328
1	NM_010500	1	NM_015764	1	NM_007471
1	NM_007542	1	NM_007828	1	NM_007655
1	NM_026951	1	NM_019691	1	NM_025808
1	NM_001141922	1	NM_001113180	1	NM_177576
1	NM_001141924	1	NM_011261	1	NM_029508
1	NM_001141925	1	NM_009527	1	NM_007927
1	NM_173406	1	NM_001035531	1	NM_011901

1	NM_001079844	1	NM_026178	1	NM_177793
1	NM_011821	1	NM_019724	1	NM_008138
1	NM_007498	1	NM_008213	1	NM_007729
1	NM_001105180	1	NM_022315	1	NM_026122
1	NM_008711	1	NM_134131	1	NM_175074
1	NM_175240	1	NM_007713	1	NM_001038990
1	NM_001033320	1	NM_172477	1	NM_001038991
1	NM_001048228	1	NR_015386	1	NM_010498
1	NM_178114	1	NM_026883	1	NM_170671
1	NM_028905	1	NM_009053	1	NM_007755
1	NM_007993	1	NM_007739	1	NR_002853
1	NM_028805	1	NM_021352	1	NM_011334
1	NM_178665	1	NM_001013766	1	NM_008873
1	NM_172868	1	NM_001012667	1	NM_133921
1	NM_009202	1	NM_011838	1	NM_001010836
1	NM_011529	1	NM_172874	1	NM_175459
1	NM_026797	1	NM_021492	1	NM_020028
1	NM_001085374	1	NM_146168	1	NM_145578
1	NM_029789	1	NM_028000	1	NM_029875
1	NM_177364	1	NM_010203	1	NM_029601
1	NM_008508	1	NM_011267	1	NM_178594
1	NM_001080933	1	NM_021883	1	NM_008394
1	NM_001033198	1	NM_133738	1	NM_145129
1	NM_172409	1	NM_027320	1	NM_027874
1	NM_023046	1	NM_027704	1	NM_139059
1	NM_144916	1	NM_170684	1	NM_019990
1	NM_007479	1	NM_010194	1	NM_001081395
1	NM_011058	1	NM_130448	1	NM_008091
1	NM_010833	1	NM_026324	1	NM_010700
1	NM_007521	1	NM_023182	1	NM_001077514
1	NM_199473	1	NM_183261	1	NM_017462
1	NM_010559	1	NM_021362	1	NM_177067
1	NM_008338	1	NM_001039466	1	NM_133687
1	NM_001039959	1	NM_172289	1	NM_133764
1	NM_023122	1	NM_028135	1	NM_011782
1	NM_172475	1	NM_009902	1	NM_020510
1	NM_001093764	1	NM_010228	1	NM_138631
1	NM_173402	1	NM_013655	1	NM_001115085
1	NM_008259	1	NM_001012477	1	NM_020285
1	NM_001013780	1	NM_146118	1	NM_013886
1	NM_134063	1	NM_010928	1	NM_010206
1	NM_008798	1	NM_011909	1	NM_001079908
1	NM_201600	1	NM_008380	1	NM_001079909

1	NM_007668	1	NM_009368	1	NM_175414
1	NM_021451	1	NM_008627	1	NM_012028
1	NM_026669	1	NM_001081009	1	NM_001033415
1	NM_001122765	1	NM_023275	1	NM_008331
1	NM_001122766	1	NM_009188	1	NM_015760
1	NM_022432	1	NM_010907	1	NM_147221
1	NM_144885	1	NM_001085549	1	NM_001081235
1	NM_176930	1	NM_009242	1	NM_033563
1	NM_008115	1	NM_001134733	1	NM_011977
1	NM_181345	1	NM_010724	1	NM_182782
1	NM_001081302	1	NM_147217	1	NM_007570
1	NM_001039126	1	NM_028980	1	NM_001077515
1	NM_019681	1	NM_010205	1	NM_011393
1	NM_080637	1	NM_183187	1	NM_029930
1	NM_134133	1	NM_183294	1	NM_001037764
1	NM_145148	1	NM_011719	1	NM_173770
1	NM_020024	1	NM_008168	1	NM_016890
1	NM_007537	1	NM_001110315	1	NM_133924
1	NR_015387	1	NM_008440	1	NM_009569
1	NM_008217	1	NM_026228	1	NM_145375
1	NM_010234	1	NM_001135149	1	NM_019442
1	NM_001048227	1	NM_001135150	1	NM_010271
1	NM_011402	1	NM_172589	1	NM_008592
1	NM_053245	1	NM_001081335	1	NM_019510
1	NM_133626	1	NM_178660	1	NM_181516
1	NM_008413	1	NM_001037877	1	NM_016798
1	NM_001048177	1	NM_028666	1	NM_011057
1	NM_009112	1	NM_011190	1	NM_138649
1	NM_009643	1	NM_053195	1	NM_178084
1	NM_018867	1	NM_001025577	1	NM_001004143
1	NM_145100	1	NM_011412	1	NM_001039537
1	NM_153156	1	NM_007948	1	NM_011454
1	NM_001033141	1	NM_175473	1	NM_008343
1	NM_145385	1	NM_175149	1	NM_028013
1	NM_008130	1	NM_001033281	1	NM_001081327
1	NM_008216	1	NM_011380	1	NM_026563
1	NM_023483	1	NM_198429	1	NM_001085407
1	NM_008988	1	NM_146030	1	NM_001085408
1	NM_001038614	1	NM_001081215	1	NM_001048005
1	NM_175271	1	NM_007643	1	NM_008019
1	NM_178931	1	NM_023386	1	NM_025629
1	NM_028864	1	NM_146015	1	NM_011841
1	NM_028757	1	NM_010562	1	NM_025725

1	NM_019517	1	NM_021893	1	NM_146182
1	NM_025988	1	NM_009498	1	NM_153392
1	NM_148928	1	NM_028176	1	NM_013682
1	NM_133979	1	NM_007607	1	NM_009543
1	NM_019742	1	NM_025685	1	NM_009964
1	NM_026502	1	NM_008881	1	NM_029528
1	NM_172738	1	NM_172309	1	NM_001113327
1	NM_008176	1	NM_017366	1	NM_028624
1	NM_010860	1	NM_009707	1	NM_007763
1	NM_145158	1	NM_001024139	1	NM_028657
1	NM_001081063	1	NM_133232	1	NM_181273
1	NM_001001309	1	NM_011595	1	NM_026791
1	NM_207246	1	NM_010664	1	NM_010544
1	NM_177769	1	NM_013657	1	NM_029761
1	NM_199222	1	NM_001130529	1	NM_026849
1	NM_172051	1	NM_172941	1	NM_001081393
1	NM_017369	1	NM_019447	1	NM_019538
1	NM_009955	1	NM_001033776	1	NM_144827
1	NM_001024945	1	NM_007866	1	NM_029362
1	NM_023268	1	NM_183027	1	NM_009331
1	NM_026931	1	NM_001110337	1	NM_027950
1	NM_139310	1	NM_172257	1	NM_009936
1	NM_001136056	1	NM_013565	1	NM_009940
1	NM_016673	1	NM_001029855	1	NM_181821
1	NM_029413	1	NM_023056	1	NM_177774
1	NM_172457	1	NM_028266	1	NR_004446
1	NM_175501	1	NM_175363	1	NM_001037761
1	NM_178680	1	NM_146062	1	NM_007857
1	NM_001081224	1	NM_001083114	1	NM_001042557
1	NM_008512	1	NM_175549	1	NM_008482
1	NM_057173	1	NM_025804	1	NM_009685
1	NM_016894	1	NM_194348	1	NM_009798
1	NM_021396	1	NM_008840	1	NM_013717
1	NM_001039162	1	NM_029781	1	NM_145211
1	NM_009990	1	NM_019819	1	NM_134125
1	NM_027225	1	NM_001109992	1	NM_134076
1	NM_177025	1	NM_011202	1	NM_172264
1	NM_011809	1	NM_010736	1	NM_175343
1	NM_177340	1	NM_181593	1	NM_013535
1	NM_126166	1	NM_020557	1	NM_080466
1	NM_013880	1	NM_011366	1	NM_213614
1	NM_001042502	1	NM_028881	1	NM_020496
1	NM_011303	1	NM_029100	1	NM_001085417

1	NM_181820	1	NM_172564	1	NM_017374
1	NM_198168	1	NM_153507	2	NM_172448
1	NM_008350	1	NM_152800	2	NM_139130
1	NM_178598	1	NM_010136	2	NM_139128
1	NM_146019	1	NM_178920	2	NM_139129
1	NM_001085450	1	NM_025611	2	NM_007736
1	NM_007615	1	NM_025399	2	NM_028722
1	NM_001085448	1	NM_001039710	2	NM_001130526
1	NM_001085449	1	NM_026424	2	NM_026738
1	NM_001085453	1	NM_001101502	2	NM_023154
1	NM_001093765	1	NM_001110508	2	NM_013792
1	NM_016969	1	NM_153515	2	NM_018820
1	NM_028882	1	NM_010212	2	NM_144794
1	NM_177740	1	NM_023580	2	NM_016696
1	NM_028627	1	NM_001110338	2	NM_027877
1	NM_027013	1	NM_183148	2	NM_026999
1	NM_022995	1	NM_053014	2	NM_009994
1	NM_001128086	1	NM_013891	2	NM_001085421
1	NM_022435	1	NM_009630	2	NM_024210
1	NM_178189	1	NM_026068	2	NM_028766
1	NM_026753	1	NM_008421	2	NM_026163
1	NM_030614	1	NM_011620	2	NM_001008425
1	NM_001034874	1	NM_001038610	2	NM_001083628
1	NM_001031621	1	NM_007826	2	NM_001136240
1	NM_001042504	1	NM_027823	2	NM_011565
1	NM_011098	1	NM_024289	2	NM_008121
1	NM_146139	1	NM_029786	2	NM_001081135
1	NM_053185	1	NM_021565	2	NM_021440
1	NM_015732	1	NM_001112739	2	NM_021528
1	NM_177583	1	NM_013601	2	NM_008230
1	NM_153088	1	NM_178380	2	NM_007874
1	NM_007504	1	NM_009523	2	NM_009223
1	NM_134080	1	NM_018808	2	NM_001025103
1	NM_175316	1	NM_019946	2	NM_199364
1	NM_015819	1	NM_001037878	2	NM_015772
1	NM_011272	1	NM_008326	2	NM_133214
1	NM_027756	1	NM_021880	2	NM_011227
1	NM_019762	1	NM_008156	2	NM_001014995
1	NM_025480	1	NM_178066	2	NM_024173
1	NM_176860	1	NM_009500	2	NM_007523
1	NM_008478	1	NM_001136059	2	NM_145144
1	NM_011994	1	NM_178410	2	NM_009775
1	NM_023438	1	NM_001029837	2	NM_172947

1	NM_138581	1	NM_024191	2	NM_016794
1	NM_010108	1	NM_021387	2	NM_029662
1	NM_001001979	1	NM_001040085	2	NM_016691
1	NM_008620	1	NM_031394	2	NM_172911
1	NM_001077202	1	NM_026003	2	NM_001114322
1	NM_028878	1	NM_001081098	2	NM_028069
1	NM_172987	1	NM_009152	2	NM_016917
1	NM_001077364	1	NM_001081390	2	NM_030064
1	NM_021476	1	NM_053207	2	NM_172863
1	NM_028883	1	NM_010948	2	NM_175429
1	NM_007496	1	NM_175938	2	NM_138580
1	NM_146090	1	NM_001039387	2	NM_001045863
1	NM_009373	1	NM_020276	2	NM_001045864
1	NM_010753	1	NM_001039386	2	NM_198029
1	NM_013520	1	NM_009058	2	NM_146018
1	NM_007719	1	NM_009273	2	NM_017405
1	NM_001025570	1	NM_172752	2	NM_013724
1	NM_010167	1	NM_001039967	2	NM_001013371
1	NM_010071	1	NM_001039965	2	NM_027828
1	NM_025858	1	NM_008675	2	NM_021354
1	NM_175256	1	NM_020606	2	NM_145473
1	NM_144832	1	NM_001081320	2	NM_145760
1	NM_009271	1	NM_028061	2	NM_053244
1	NM_001025395	1	NM_145383	2	NM_080837
1	NM_013588	1	NM_011785	2	NM_028584
1	NM_009549	1	NM_007416	2	NM_133889
1	NM_001109043	1	NM_001082553	2	NM_212447
1	NM_027951	1	NM_021508	2	NM_194263
1	NM_177157	1	NM_009636	2	NM_021551
1	NM_021896	1	NM_199299	2	NM_177870
1	NM_001101603	1	NM_001081020	2	NM_177092
1	NM_015753	1	NM_010015	2	NM_178060
1	NM_028027	1	NM_007599	2	NM_133350
1	NM_008583	1	NM_028736	2	NM_145934
1	NM_207176	1	NM_130891	2	NM_145489
1	NM_007912	1	NM_133922	2	NM_015749
1	NM_009358	1	NM_019697	2	NM_001130459
1	NM_019401	1	NM_146124	2	NM_001130458
1	NM_001141948	1	NM_001134741	2	NM_012021
1	NM_001141949	1	NM_025898	2	NM_001039242
1	NM_001005608	1	NM_133442	2	NM_019662
1	NM_133663	1	NM_001033126	2	NM_010248
1	NM_026818	1	NM_001042564	2	NM_008281

1	NM_009045	1	NM_018882	2	NM_001110252
1	NM_011104	1	NM_009679	2	NM_026240
1	NM_177226	1	NM_027464	2	NM_201601
1	NM_001081354	1	NM_177073	2	NM_010207
1	NM_001001565	1	NM_172884	2	NM_172145
1	NM_016921	1	NM_011305	2	NM_011198
1	NM_146004	1	NM_028713	2	NM_009170
1	NM_019813	1	NM_027551	2	NM_015759
1	NM_007859	1	NM_007738	2	NM_027294
1	NM_054056	1	NM_028258	2	NM_001033301
1	NM_001113569	1	NM_011459	2	NM_021567
1	NM_009295	1	NM_172691	2	NM_172546
1	NM_001040086	1	NM_177305	2	NM_175730
1	NM_010104	1	NM_001114088	2	NM_133743
1	NM_207655	1	NM_001081142	2	NM_175677
1	NM_009304	1	NM_176958	2	NM_008853
1	NM_009055	1	NM_018774	2	NM_001083110
1	NM_021339	1	NM_025626	2	NM_001110513
1	NM_201519	1	NM_029981	2	NM_016851
1	NM_133201	1	NM_007992	2	NM_001040426
1	NM_178748	1	NM_001081437	2	NM_020589
1	NM_183064	1	NM_001136078	2	NM_175104
1	NM_018781	1	NM_198091	2	NM_007754
1	NM_178218	1	NM_145419	2	NM_025473
1	NM_023732	1	NM_009930	2	NM_001033328
1	NM_001013386	1	NM_145491	2	NM_013664
1	NM_001126490	1	NM_173396	2	NM_007421
1	NM_010199	1	NM_030700	2	NM_133974
1	NM_133222	1	NM_172395	2	NM_172568
1	NM_026455	1	NM_001136260	2	NM_024499
1	NM_008103	1	NM_018760	2	NM_009503
1	NM_133897	1	NM_153107	2	NM_172802
1	NM_009821	1	NM_007485	2	NM_011880
1	NM_001111023	1	NM_023824	2	NM_001085416
1	NM_001038612	1	NM_026651	2	NM_001085415
1	NM_007441	1	NM_011578	2	NM_001081115
1	NM_016846	1	NM_008608	2	NM_153804
1	NM_201352	1	NM_007825	2	NM_001081252
1	NR_024085	1	NM_010518	2	NM_001136222
1	NM_177683	1	NM_018865	2	NM_028817
1	NM_001033416	1	NM_013845	2	NM_001033606
1	NM_011452	1	NM_028472	2	NM_177919
1	NM_011763	1	NM_009230	2	NM_133710

1	NM_177282	1	NM_028325	2	NM_027127
1	NM_010755	1	NM_011063	2	NM_030113
1	NM_001111051	1	NM_013585	2	NM_145542
1	NM_001111052	1	NM_022814	2	NM_011594
1	NM_144783	1	NM_007656	2	NM_019458
1	NM_008595	1	NM_001122754	2	NM_008551
1	NM_020505	1	NM_027419	2	NM_011186
1	NM_175640	1	NM_001042699	2	NM_144914
1	NM_001113471	1	NM_172053	2	NM_025945
1	NM_008882	1	NM_001009947	2	NM_011519
1	NM_026439	1	NM_172710	2	NM_001098636
1	NM_008546	1	NM_133804	2	NM_001033206
1	NM_172736	1	NM_024174	2	NM_009168
1	NM_010231	1	NM_010927	2	NM_133234
1	NM_001111021	1	NM_001111324	2	NM_027770
1	NM_001111022	1	NM_011882	2	NM_009062
1	NM_009825	1	NM_028780	2	NM_177303
1	NM_001111043	1	NM_027552	2	NM_177140
1	NM_001111044	1	NM_080726	2	NM_022424
1	NM_001083929	1	NM_010872	2	NM_053082
1	NM_199305	1	NM_001126182	2	NM_026162
1	NM_146131	1	NM_001033243	2	NM_024263
1	NM_001014390	1	NM_011052	2	NM_008070
1	NM_008285	1	NM_177724	2	NM_001085371
1	NM_177883	1	NM_023476	2	NM_011356
1	NM_177152	1	NM_030127	2	NM_022027
1	NM_026380	1	NM_011254	2	NM_010417
1	NM_001033877	1	NM_001039543	2	NM_133926
1	NM_001081295	1	NM_010801	2	NM_030174
1	NM_001012638	1	NM_173414	2	NM_013770
1	NM_001081298	1	NM_010446	2	NM_010181
1	NM_022009	1	NM_019440	2	NM_021544
1	NM_177469	1	NM_198419	2	NM_001077398
1	NM_001033461	1	NM_030210	2	NM_010698
1	NM_021469	1	NM_001042779	2	NM_030226
1	NM_016976	1	NM_016697	2	NM_133983
1	NM_001114333	1	NM_011449	2	NM_172118
1	NM_177271	1	NM_008937	2	NM_001113415
1	NM_172476	1	NM_010351	2	NM_001113414
1	NM_029802	1	NM_008867	2	NM_010096
1	NM_019457	1	NM_001113248	2	NM_178029
1	NM_183208	1	NM_001083334	2	NM_026179
1	NM_177723	1	NM_009668	2	NM_010273

1	NM_010820	1	NM_153142	2	NM_007795
1	NM_144537	1	NM_001002846	2	NM_001037723
1	NM_172523	1	NM_021899	2	NM_001081456
1	NM_008161	1	NM_009066	2	NM_133225
1	NM_145122	1	NM_027984	2	NM_019456
1	NM_175264	1	NM_012030	2	NM_080793
1	NM_011028	1	NM_001037905	2	NM_010237
1	NM_130449	1	NM_001102400	2	NM_133192
1	NM_182927	1	NM_023118	2	NM_175478
1	NM_001034909	1	NM_001136227	2	NM_207298
1	NM_013586	1	NM_009396	2	NM_013819
1	NM_001141932	1	NM_023854	2	NM_031380
1	NM_001141931	1	NM_029770	2	NM_009655
1	NM_010899	1	NM_027280	2	NM_009576
1	NM_138674	1	NM_177616	2	NM_001079686
1	NM_016886	1	NM_011150	2	NM_021457
1	NM_199198	1	NM_001081286	2	NM_001081236
1	NM_001136073	1	NM_001113379	2	NM_153784
1	NM_144551	1	NM_029053	2	NM_024211
1	NM_007725	1	NM_009416	2	NM_026123
1	NM_008244	1	NM_011571	2	NM_007669
1	NM_026137	1	NM_013505	2	NM_001136076
1	NM_015823	1	NM_198214	2	NM_145434
1	NM_001122998	1	NM_145391	2	NM_018777
1	NM_134117	1	NR_003248	2	NM_178873
1	NM_053197	1	NM_001081212	2	NM_001115015
1	NM_053093	1	NM_009621	2	NM_019487
1	NM_023463	1	NM_026131	2	NM_172861
1	NM_145459	1	NM_001114087	2	NM_011031
1	NM_021359	1	NM_145135	2	NM_019822
1	NM_001039647	1	NM_133363	2	NM_027315
1	NM_009367	1	NM_181274	2	NM_010473
1	NM_008245	1	NM_008876	2	NM_027924
1	NM_009347	1	NM_009177	2	NM_001111099
1	NM_177757	1	NM_001113181	2	NM_001033231
1	NM_001081170	1	NM_177751	2	NM_011103
1	NR_003364	1	NM_145554	2	NM_028419
1	NM_031184	1	NM_001127324	2	NM_011224
1	NM_001077694	1	NM_009926	2	NM_010469
1	NM_172539	1	NM_007913	2	NM_198095
1	NM_010332	1	NM_007880	2	NM_009728
1	NM_025897	1	NM_029609	2	NM_010242
1	NM_030259	1	NM_026042	2	NM_011874

1	NM_010513	1	NM_009290	2	NM_009782
1	NM_021493	1	NM_030218	2	NM_001101488
1	NM_133674	1	NM_013912	2	NM_029466
1	NM_011102	1	NM_001081425	2	NM_001042615
1	NM_145450	1	NM_008174	2	NM_011427
1	NM_013568	1	NM_001081190	2	NM_153574
1	NM_001013753	1	NM_025482	2	NM_019466
1	NM_027859	1	NM_008851	2	NM_138743
1	NM_133968	1	NM_030554	2	NM_173180
1	NM_010394	1	NM_013599	2	NM_023065
1	NM_011348	1	NM_009519	2	NM_001033292
1	NM_021537	1	NM_010110	2	NM_033474
1	NM_001004155	1	NR_024257	2	NM_008445
1	NM_001112735	1	NM_029849	2	NM_022886
1	NM_019770	1	NM_001005740	2	NM_024231
1	NM_172787	1	NM_008945	2	NM_001004367
1	NM_009181	1	NM_153679	2	NM_001025258
1	NM_020296	1	NM_023057	2	NM_001025256
1	NM_027143	1	NM_030725	2	NM_001025259
1	NM_001093766	1	NM_008315	2	NM_001025255
1	NM_001080126	1	NM_175097	2	NM_001025254
1	NM_172682	1	NM_011097	2	NM_001025251
1	NM_008332	1	NM_001042534	2	NM_175491
1	NM_133832	1	NM_011050	2	NM_011756
1	NM_008862	1	NM_011770	2	NM_001077632
1	NM_008393	1	NM_030733	2	NM_001033541
1	NM_027896	1	NM_019467	2	NM_001130408
1	NM_001080775	1	NM_010591	2	NM_001080549
1	NM_054071	1	NM_009473	2	NM_029296
1	NM_010380	1	NM_053242	2	NM_019979
1	NM_010517	1	NM_024427	2	NM_026436
1	NM_133778	1	NM_016769	2	NM_019652
1	NM_177662	1	NM_013460	2	NM_001111141
1	NM_010111	1	NM_011349	2	NM_001013817
1	NM_144796	1	NM_147776	2	NM_007709
1	NM_133955	1	NM_153068	2	NM_018794
1	NM_013497	1	NM_172444	2	NM_080857
1	NM_012049	1	NM_019840	2	NM_173053
1	NM_009183	1	NM_027257	2	NM_025836
1	NM_172894	1	NM_152799	2	NM_007856
1	NM_026473	1	NM_001032414	2	NM_175170
1	NM_027376	1	NM_001032413	2	NM_178283
1	NM_001001602	1	NM_028460	2	NM_007476

1	NM_010734	1	NM_053083	2	NM_025961
1	NM_022025	1	NM_145608	2	NM_008304
1	NM_145575	1	NM_172381	2	NM_178870
1	NM_011193	1	NM_023872	2	NM_134159
1	NM_028769	1	NM_007403	2	NM_011306
1	NM_018811	1	NM_008209	2	NM_178659
1	NM_027886	1	NM_019793	2	NM_134160
1	NM_028730	1	NM_001130191	2	NM_001024619
1	NM_011878	1	NM_001130190	2	NM_007625
1	NM_024217	1	NM_011360	2	NM_008951
1	NM_173749	1	NM_001130188	2	NM_027917
1	NM_016715	1	NM_001130189	2	NM_011313
1	NM_023043	1	NM_007803	2	NR_003627
1	NM_001126338	1	NM_145586	2	NM_001024927
1	NM_001109042	1	NM_008515	2	NM_001081641
1	NM_001109041	1	NM_011373	2	NM_025661
1	NM_001109040	1	NM_019687	2	NM_024208
1	NM_016705	1	NM_008742	2	NM_009740
1	NM_011935	1	NM_029655	2	NM_011639
1	NM_001001980	1	NM_009913	2	NM_008501
1	NM_001042655	1	NM_001081175	2	NM_025573
1	NM_174876	1	NM_007392	2	NM_172398
1	NM_016872	1	NM_177265	2	NM_025944
1	NM_001080742	1	NM_001017427	2	NM_025806
1	NM_001038613	1	NM_213727	2	NM_010312
1	NM_029005	1	NM_017463	2	NM_144883
1	NM_009812	1	NM_001033273	2	NM_001001297
1	NM_001034895	1	NM_133969	2	NM_177588
1	NM_013492	1	NM_175548	2	NM_145987
1	NM_013723	1	NM_173372	2	NM_019719
1	NM_010051	1	NM_015743	2	NM_001111111
1	NM_009044	1	NM_001009979	2	NM_153484
1	NM_172621	1	NM_001009978	2	NM_013610
1	NM_025545	1	NM_011842	2	NM_194066
1	NM_001136091	1	NM_023422	2	NM_194069
1	NM_008235	1	NM_001038708	2	NM_194068
1	NM_001025444	1	NM_024439	2	NM_194067
1	NM_001045516	1	NM_030074	2	NM_026790
1	NM_133719	1	NM_144552	2	NM_207689
1	NM_001111142	1	NM_146120	2	NM_207691
1	NM_177371	1	NM_011141	2	NM_207690
1	NM_011772	1	NM_178111	2	NM_207688
1	NM_172466	1	NM_021437	2	NM_001033500

1	NM_007664	1	NM_001081043	2	NM_183297
1	NM_198657	1	NM_008750	2	NM_177692
1	NM_027531	1	NM_001025245	2	NM_027897
1	NM_053115	1	NM_001033217	2	NM_028434
1	NM_011170	1	NM_013543	2	NM_031880
1	NM_001110149	1	NM_139051	2	NM_178762
1	NM_146007	1	NM_008626	2	NM_025317
1	NM_030257	1	NM_010947	2	NM_011894
1	NM_019980	1	NR_002855	2	NM_144810
1	NM_009518	1	NM_030075	2	NM_010595
1	NM_009954	1	NM_172486	2	NM_019585
1	NM_177620	1	NM_175935	2	NM_007528
1	NM_011877	1	NM_009106	2	NM_178411
1	NM_001040459	1	NM_133641	2	NM_009419
1	NM_001039562	1	NM_207202	2	NM_009357
1	NM_010195	1	NM_019686	2	NM_177717
1	NM_198037	1	NM_028820	2	NM_175023
1	NM_027862	1	NM_001001738	2	NM_177822
1	NM_133768	1	NM_001141982	2	NM_013900
1	NM_019572	1	NM_025635	2	NM_010923
1	NM_019764	1	NM_178696	2	NM_180960
1	NM_001110150	1	NM_181071	2	NM_008640
1	NM_172775	1	NM_016765	2	NM_011980
1	NM_146127	1	NM_019521	2	NM_001081072
1	NM_008426	1	NM_009933	2	NM_007557
1	NM_001033167	1	NM_173751	2	NM_172271
1	NM_009213	1	NM_009256	2	NM_148937
1	NM_144882	1	NM_175234	2	NM_177366
1	NM_010575	1	NM_008858	2	NM_030245
1	NM_173390	1	NM_198711	2	NM_027870
1	NM_008997	1	NM_029838	2	NM_133182
1	NM_008829	1	NM_011310	2	NM_001077638
1	NM_021478	1	NM_001111311	2	NM_026752
1	NM_001080774	1	NM_022032	2	NM_026667
1	NM_183162	1	NM_175485	2	NM_020590
1	NM_026198	1	NM_177384	2	NM_018795
1	NR_002863	1	NM_001081116	2	NM_010055
1	NM_028807	1	NM_026922	2	NM_007406
1	NM_027865	1	NR_004842	2	NM_001039074
1	NM_153410	1	NM_023842	2	NM_021323
1	NM_008659	1	NM_008309	2	NM_001085440
1	NM_026221	1	NM_001113400	2	NM_001039072
1	NM_010118	1	NM_001001491	2	NM_001039073

1	NM_175274	1	NM_011854	2	NM_001039071
1	NM_001110849	1	NM_138606	2	NM_011918
1	NM_133215	1	NM_008976	2	NM_145509
1	NM_001033779	1	NM_011018	2	NM_198959
1	NM_145839	1	NM_175313	2	NM_033270
1	NM_181318	1	NM_001029929	2	NM_028900
1	NM_138956	1	NM_020260	2	NM_008669
1	NM_146162	1	NM_001081241	2	NM_011769
1	NM_001001566	1	NM_013673	2	NM_010070
1	NM_001037178	1	NM_007609	2	NM_153087
1	NM_027117	1	NM_173733	2	NM_028177
1	NM_174857	1	NM_207209	2	NM_021313
1	NM_025730	1	NM_007547	2	NM_009633
1	NM_009806	1	NM_009153	2	NM_177278
1	NM_028450	1	NM_013611	2	NM_007483
1	NM_016849	1	NM_010221	2	NM_023668
1	NM_019704	1	NM_172893	2	NM_025616
1	NM_053095	1	NM_009970	2	NM_009459
1	NM_178900	1	NM_152923	2	NM_026417
1	NM_013470	1	NM_001024136	2	NM_016922
1	NM_001083810	1	NM_026485	2	NM_026676
1	NM_001081331	1	NM_175033	2	NM_001114541
1	NM_019586	1	NM_008495	2	NM_016878
1	NM_146110	1	NM_009714	2	NM_009124
1	NM_001014398	1	NM_019675	2	NM_197986
1	NM_001113331	1	NM_001136055	2	NM_020012
1	NM_175181	1	NM_001113464	2	NM_134150
1	NM_021464	1	NM_001113460	2	NM_146102
1	NM_008455	1	NM_001113461	2	NM_008887
1	NM_016768	1	NM_172500	2	NM_026556
1	NM_153534	1	NM_025381	2	NM_175246
1	NM_008909	1	NM_001113358	2	NM_053106
1	NM_013458	1	NM_023663	2	NM_207105
1	NM_172604	1	NM_177090	2	NM_175541
1	NM_015771	1	NM_130877	2	NM_008494
1	NM_177161	1	NM_001040611	2	NM_026531
1	NM_009896	1	NM_019417	2	NM_025802
1	NM_133943	1	NM_172944	2	NM_172769
1	NM_001040684	1	NM_011609	2	NM_026778
1	NM_007584	1	NM_001039223	2	NM_011906
1	NM_172962	1	NM_021566	2	NM_011498
1	NM_001080814	1	NM_146107	2	NM_177591
1	NM_153537	1	NM_011065	2	NM_177595

1	NM_001003948	1	NM_021274	2	NM_009754
1	NM_023320	1	NM_008192	2	NM_207680
1	NM_001013012	1	NM_133194	2	NM_207681
1	NM_175307	1	NM_011520	2	NM_153417
1	NM_001037128	1	NM_177638	2	NM_178149
1	NM_001037129	1	NM_018784	2	NM_172703
1	NM_001037127	1	NM_029381	2	NM_201244
1	NM_010944	1	NM_027804	2	NM_026373
1	NM_001037130	1	NM_001081182	2	NM_025272
1	NM_010444	1	NM_176922	2	NM_011354
1	NM_001004363	1	NM_008416	2	NM_153533
1	NM_177914	1	NM_017377	2	NM_134050
1	NM_023738	1	NM_145538	2	NM_011032
1	NM_010235	1	NM_001025208	2	NM_008484
1	NM_019581	1	NM_026025	2	NM_027249
1	NM_007555	1	NM_026150	2	NM_023734
1	NM_011766	1	NM_178771	2	NM_080420
1	NM_018746	1	NM_001122830	2	NM_008542
1	NM_001083959	1	NM_013683	2	NM_027512
1	NM_012046	1	NM_148932	2	NM_175270
1	NM_001083960	1	NM_172503	2	NM_175138
1	NM_175437	1	NM_030699	2	NM_133664
1	NM_001024614	1	NM_001039076	2	NM_008360
1	NM_130859	1	NM_001039075	2	NM_010154
1	NM_010276	1	NM_021395	2	NM_013718
1	NM_030262	1	NM_025793	2	NM_009746
1	NM_170701	1	NM_001083925	2	NM_027725
1	NM_170702	1	NR_003507	2	NM_010072
1	NM_170704	1	NM_001048060	2	NM_175937
1	NM_170703	1	NM_198004	2	NM_027088
1	NM_011611	1	NM_029077	2	NM_016974
1	NM_198617	1	NM_024206	2	NM_007562
1	NM_146258	1	NM_026146	2	NM_010786
1	NM_010605	1	NM_183180	2	NM_019542
1	NM_173788	1	NM_033073	2	NM_009778
1	NM_001081102	1	NM_201373	2	NM_001110831
1	NM_008809	1	NM_010893	2	NM_025452
1	NM_011078	1	NM_145403	2	NM_028394
1	NM_031169	1	NM_007484	2	NM_015766
1	NM_001077709	1	NM_153136	2	NM_201359
1	NM_008202	1	NM_011715	2	NM_172647
1	NM_011904	1	NM_026582	2	NM_001039385
1	NM_172530	1	NM_177290	2	NM_028076

1	NM_001114124	1	NM_001112715	2	NM_177386
1	NM_010441	1	NM_173379	2	NM_134083
1	NM_027444	1	NM_173383	2	NM_001035509
1	NM_172275	1	NM_153790	2	NM_001035510
1	NM_019713	1	NM_009397	2	NM_025893
1	NM_009808	1	NM_016752	2	NM_026831
1	NM_010925	1	NM_019749	2	NM_025833
1	NM_008239	1	NM_001033980	2	NM_133694
1	NM_011046	1	NM_001083936	2	NM_011950
1	NM_019651	1	NM_178202	2	NM_026064
1	NM_023564	1	NM_177130	2	NM_177838
1	NM_020047	1	NM_008716	2	NM_017376
1	NM_027665	1	NM_013729	2	NM_177604
1	NM_173745	1	NM_172441	2	NM_001102613
1	NM_175750	1	NM_019569	2	NM_130860
1	NM_029365	1	NM_007742	2	NM_007801
1	NM_023053	1	NM_175260	2	NM_027627
1	NM_177059	1	NM_053248	2	NM_027323
1	NM_011866	1	NM_033444	2	NM_009496
1	NM_010213	1	NM_010107	2	NM_001080557
1	NM_010227	1	NM_178595	2	NM_001037711
1	NM_146191	1	NM_007692	2	NM_011509
1	NM_008036	1	NM_008321	2	NM_025397
1	NM_008136	1	NM_026253	2	NM_025907
1	NM_018830	1	NM_023850	2	NM_009296
1	NM_026346	1	NM_026428	2	NM_013871
1	NM_172393	1	NM_001134384	2	NM_024176
1	NM_177725	1	NM_001111312	2	NM_016677
1	NM_001081300	1	NM_145713	2	NM_001081655
1	NM_025994	1	NM_145133	2	NM_011375
1	NM_175259	1	NM_173863	2	NM_001035228
1	NM_001098226	1	NM_146025	2	NM_138653
1	NM_139206	1	NM_030256	2	NM_145986
1	NM_001037177	1	NM_001081171	2	NM_178804
1	NM_001081263	1	NM_007988	2	NM_010140
1	NM_178408	1	NM_198249	2	NM_019634
1	NM_001039175	1	NM_033608	2	NM_145469
1	NM_019422	1	NM_001083935	2	NM_008404
1	NM_001136088	1	NM_028110	2	NM_011171
1	NM_001077362	1	NM_001001892	2	NM_010725
1	NM_007976	1	NM_175136	2	NM_009528
1	NM_001013826	1	NM_029928	2	NM_007650
1	NR_003623	1	NM_001003911	2	NM_133655

1	NM_023699	1	NM_153778	2	NM_009431
1	NM_009142	1	NM_019965	2	NM_001099298
1	NM_021559	1	NM_178618	2	NM_013899
1	NM_207279	1	NM_010813	2	NM_139001
1	NM_013691	1	NM_008506	2	NM_011324
1	NM_008827	1	NM_016701	2	NM_133731
1	NM_172396	1	NM_026017	2	NM_009907
1	NM_016807	1	NM_007704	2	NM_010579
1	NM_001098227	1	NM_007910	2	NM_019723
1	NM_177078	1	NM_009713	2	NM_172162
1	NM_009948	1	NM_029153	2	NM_001081062
1	NM_001081453	1	NM_009640	2	NM_145532
1	NM_001037859	1	NM_177448	2	NM_153566
1	NM_001081454	1	NM_001130412	2	NM_025605
1	NM_177460	1	NM_015763	2	NM_021486
1	NM_175686	1	NM_172950	2	NM_001043354
1	NM_011127	1	NM_212473	2	NM_013803
1	NM_023622	1	NM_001083937	2	NM_010613
1	NM_024223	1	NM_025276	2	NM_023231
1	NM_011566	1	NM_001081333	2	NM_011322
1	NM_008353	1	NM_175770	2	NM_134020
1	NM_001112701	1	NM_009885	2	NM_026155
1	NM_011181	1	NM_194333	2	NM_033475
1	NM_178630	1	NM_001113548	2	NM_027855
1	NM_016808	1	NM_078484	2	NM_212470
1	NM_015736	1	NM_023048	2	NM_021398
1	NM_025619	1	NM_144879	2	NM_144906
1	NM_007727	1	NM_021439	2	NM_029182
1	NM_029357	1	NM_026186	2	NM_013558
1	NM_008037	1	NM_021506	2	NM_016686
1	NM_026073	1	NM_011577	2	NM_172148
1	NM_023813	1	NM_144525	2	NM_177657
1	NM_001025438	1	NM_001081050	2	NM_011192
1	NM_001025439	1	NM_153287	2	NM_009800
1	NM_134156	1	NM_026441	2	NM_025319
1	NM_011784	1	NM_001081047	2	NM_009279
1	NM_198438	1	NM_010757	2	NM_001083317
1	NM_023672	1	NM_011902	2	NM_026404
1	NM_010835	1	NM_001136083	2	NM_175247
1	NM_001080926	1	NM_010777	2	NM_028787
1	NM_053073	1	NM_029564	2	NM_153512
1	NM_173731	1	NM_001005787	2	NM_008430
1	NM_177699	1	NM_130451	2	NM_011176

1	NM_010999	1	NM_153535	2	NM_177708
1	NM_198411	1	NM_028779	2	NM_133850
1	NR_004843	1	NM_172290	2	NM_025439
1	NM_010440	1	NM_021462	2	NM_030612
1	NM_001079869	1	NM_080435	2	NM_028202
1	NM_009893	1	NM_008090	2	NM_001033208
1	NM_009524	1	NM_025821	2	NM_010708
1	NM_183022	1	NM_026376	2	NM_001037743
1	NM_029881	1	NM_010092	2	NM_198013
1	NM_011526	1	NM_001122780	2	NM_009063
1	NM_028443	1	NM_029652	2	NM_178765
1	NM_023912	1	NM_020493	2	NM_001039077
1	NM_019498	1	NM_023116	2	NM_001039078
1	NM_133816	1	NM_011942	2	NM_007991
1	NM_001033380	1	NM_013509	2	NM_011873
1	NM_198092	1	NM_144797	2	NM_152804
1	NM_007417	1	NM_001130179	2	NM_138607
1	NM_011204	1	NM_001130178	2	NM_213616
1	NM_175386	1	NM_001130177	2	NM_019465
1	NM_175183	1	NM_001130175	2	NM_024216
1	NM_010284	1	NM_001130174	2	NM_022021
1	NM_001042605	1	NM_001130180	2	NM_080510
1	NM_010545	1	NM_001130181	2	NM_018763
1	NM_010794	1	NM_011619	2	NM_021388
1	NM_001110148	1	NM_010019	2	NM_010043
1	NM_181409	1	NM_010117	2	NM_175266
1	NM_008074	1	NM_134068	2	NM_008987
1	NM_001048147	1	NM_153422	2	NM_007450
1	NM_181395	1	NM_011199	2	NM_001081473
1	NM_001038499	1	NM_008663	2	NM_027404
1	NM_017464	1	NM_178407	2	NM_172715
1	NM_133485	1	NM_153528	2	NM_029035
1	NM_011607	1	NM_001122739	2	NM_025649
1	NM_015800	1	NM_145589	2	NM_026552
1	NM_175168	1	NM_020604	2	NR_015349
1	NM_007730	1	NM_019791	2	NM_153565
1	NM_001025313	1	NM_001077361	2	NM_058212
1	NM_009318	1	NM_009154	2	NM_028117
1	NM_008240	1	NM_001130525	2	NM_011627
1	NM_019873	1	NM_001112813	2	NM_016767
1	NM_181397	1	NM_009783	3	NM_009867
1	NM_016978	1	NM_009472	3	NM_001081084
1	NM_212435	1	NM_145503	3	NM_020267

1	NM_009099	1	NM_134127	3	NM_144848
1	NM_013901	1	NM_010211	3	NM_008002
1	NM_011923	1	NM_022417	3	NM_194342
1	NM_172397	1	NM_175268	3	NM_201393
1	NM_015734	1	NM_172777	3	NM_023270
1	NM_011773	1	NM_172588	3	NM_013917
1	NM_207654	1	NM_011214	3	NM_001131054
1	NM_010109	1	NM_001083119	3	NM_001042714
1	NM_030236	1	NM_008781	3	NM_011019
1	NM_008555	1	NM_007405	3	NM_017373
1	NM_001011874	1	NM_029264	3	NM_011701
1	NM_001048178	1	NM_029555	3	NM_175314
1	NM_009525	1	NM_139298	3	NM_201392
1	NM_007966	1	NM_010320	3	NM_201391
1	NM_199239	1	NM_019391	3	NM_201390
1	NM_199241	1	NM_026820	3	NM_201389
1	NM_199240	1	NM_020265	3	NM_010258
1	NM_199238	1	NM_009732	3	NM_201388
1	NM_007903	1	NM_177145	3	NM_016753
1	NM_172913	1	NM_182807	3	NM_001037713
1	NM_175096	1	NM_007408	3	NM_001002011
1	NM_153412	1	NM_153393	3	NM_029575
1	NM_199449	1	NM_001085370	3	NM_009371
1	NM_011280	1	NM_016887	3	NM_001111102
1	NM_011074	1	NM_134082	3	NM_009451
1	NM_022019	1	NM_138596	3	NM_201387
1	NM_133727	1	NM_023524	3	NM_021509
1	NM_021394	1	NM_178890	3	NM_033327
1	NM_025368	1	NM_009735	3	NM_018789
1	NM_013502	1	NM_198247	3	NM_015744
1	NM_021716	1	NM_010153	3	NM_001136077
1	NM_009866	1	NM_008756	3	NM_001025568
1	NM_029935	1	NM_010908	3	NM_175013
1	NM_001039151	1	NM_013539	3	NM_201386
1	NM_001039150	1	NM_007759	3	NM_011054
1	NM_009851	1	NM_008390	3	NM_145492
1	NM_010876	1	NM_175398	3	NM_026058
1	NM_019661	1	NM_011448	3	NM_010100
1	NM_010762	1	NM_018880	3	NM_001081198
1	NM_147220	1	NM_009992	3	NM_183308
1	NM_001081044	1	NM_001080708	3	NM_010723
1	NM_019408	1	NM_027996	3	NM_001013381
1	NM_130863	1	NM_009365	3	NM_176837

1	NM_009101	1	NM_177574	3	NR_024329
1	NM_029862	1	NM_001081120	3	NM_201385
1	NM_018744	1	NM_021474	3	NM_010892
1	NM_172932	1	NM_001037957	3	NM_001122730
1	NM_001122949	1	NM_009820	3	NM_009931
1	NM_010823	1	NM_010567	3	NM_001109661
1	NM_172411	1	NM_001081244	3	NM_009805
1	NM_023047	1	NM_145445	3	NM_015773
1	NM_177866	1	NM_016845	3	NM_013626
1	NM_001130030	1	NM_001136071	3	NM_007585
1	NM_008479	1	NM_001008533	3	NM_009781
1	NM_027728	1	NM_001039510	3	NM_016928
1	NM_008935	1	NM_001003817	3	NM_207277
1	NM_011129	1	NM_133351	3	NM_001114079
1	NM_001037221	1	NM_145940	3	NM_029007
1	NM_028966	1	NM_001083315	3	NM_011117
1	NM_019449	1	NM_022332	3	NM_025671
1	NM_011777	1	NM_001033040	3	NM_001039156
1	NM_001122737	1	NM_146142	3	NM_138579
1	NM_172537	1	NM_011503	3	NM_010681
1	NM_153319	1	NM_026832	3	NM_001039038
1	NM_178754	1	NM_133665	3	NM_177698
1	NM_183220	1	NM_028479	3	NM_178242
1	NM_018764	1	NM_013613	3	NM_027905
1	NM_001040088	1	NM_173008	3	NM_031158
1	NM_001040087	1	NM_010586	3	NM_026481
1	NM_183037	1	NM_019923	3	NR_003258
1	NM_133949	1	NM_018738	3	NM_026988
1	NM_016740	1	NM_011155	3	NM_008109
1	NM_139307	1	NM_001037955	3	NM_145456
1	NM_001005342	1	NM_001080813	3	NM_025654
1	NM_011829	1	NR_003946	3	NM_134257
1	NM_144941	1	NM_026601	3	NM_013607
1	NM_025831	1	NM_152220	3	NM_018779
1	NM_013606	1	NM_144955	3	NM_011544
1	NM_146226	1	NM_009846	3	NM_054099
1	NR_003508	1	NM_148950	3	NM_001112698
1	NM_009059	1	NM_028416	3	NM_013609
1	NM_175561	1	NM_001024720	3	NM_013815
1	NM_011530	1	NM_146085	3	NM_011851
1	NM_026145	1	NM_010112	3	NM_030263
1	NM_008737	1	NM_007737	3	NM_172728
1	NM_030137	1	NM_026977	3	NM_008969

1	NM_017467	1	NM_172696	3	NM_008150
1	NM_008471	1	NM_144872	3	NM_028810
1	NM_183221	1	NM_001081633	3	NM_008986
1	NM_009801	1	NM_022408	3	NM_017379
1	NM_025735	1	NM_147201	3	NM_133980
1	NM_001081401	1	NM_019777	3	NM_175668
1	NM_007442	1	NM_133794	3	NM_010730
1	NM_001004176	1	NM_010516	3	NM_033602
1	NM_007673	1	NM_010942	3	NM_001024716
1	NM_008697	1	NM_134021	3	NM_001081424
1	NM_001093754	1	NM_013781	3	NM_009037
1	NM_009697	1	NM_145629	3	NM_008690
1	NM_009309	1	NM_015787	3	NM_007929
1	NM_010437	1	NM_001113527	3	NM_026599
1	NM_001024602	1	NM_025829	3	NM_009526
1	NM_011189	1	NM_013642	3	NM_010468
1	NM_011893	1	NM_025779	3	NM_009333
1	NM_018734	1	NM_001081344	3	NM_009932
1	NM_001045527	1	NM_020043	3	NM_001081088
1	NM_025289	1	NM_001029838	3	NM_144791
1	NM_019989	1	NM_013749	3	NM_010576
1	NM_010514	1	NM_198109	3	NM_009988
1	NM_025744	1	NM_172708	3	NM_028146
1	NM_008748	1	NM_009502	3	NM_144549
1	NM_023131	1	NM_001134383	3	NM_015817
1	NM_017378	1	NM_024189	3	NM_178884
1	NM_007743	1	NM_011800	3	NM_008046
1	NM_133994	1	NM_133950	3	NM_145216
1	NM_031373	1	NM_031393	3	NM_001111267
1	NM_001077696	1	NM_172994	3	NM_001009950
1	NM_010412	1	NM_007791	3	NM_016719
1	NM_172685	1	NM_009465	3	NM_009811
1	NM_027864	1	NM_020583	3	NM_027154
1	NM_021715	1	NM_007463	3	NM_001039530
1	NM_001009819	1	NM_013863	3	NM_009393
1	NM_001101483	1	NM_018854	3	NM_001033322
1	NM_181328	1	NM_023230	3	NM_019750
1	NM_023158	1	NM_011665	4	NM_001045523
1	NM_001008233	1	NM_025516	4	NM_178791
1	NM_027533	1	NM_001008702	4	NM_009644
1	NM_001001881	1	NM_177801	4	NM_010496
1	NM_001038602	1	NM_001081279	4	NM_019772
1	NM_007896	1	NM_027373	4	NM_022327

1	NM_007553	1	NM_008483	4	NM_053253
1	NM_028351	1	NM_011581	4	NM_001083916
1	NM_029999	1	NM_010008	4	NM_133707
1	NM_021422	1	NM_027893	4	NM_053250
1	NM_183390	1	NM_001122680	4	NM_181664
1	NM_178227	1	NM_007588	4	NM_134024
1	NM_001083917	1	NM_025401	4	NM_011537
1	NM_153522	1	NM_053108	4	NM_001081123
1	NM_027990	1	NM_001085390	4	NM_172471
1	NM_030243	1	NM_178220	4	NM_176838
1	NM_001141981	1	NM_177231	4	NM_023211
1	NM_001122736	1	NM_178782	4	NM_008255
1	NM_007855	1	NM_001043335	4	NM_172705
1	NM_013869	1	NM_029688	4	NM_025658
1	NM_011281	1	NM_019829	4	NM_194055
1	NM_007653	1	NM_008817	4	NM_007923
1	NM_029509	1	NM_008871	4	NM_144731
1	NM_001109752	1	NM_008187	4	NM_025427
1	NM_007864	1	NM_001130176	4	NM_134252
1	NM_001081185	1	NM_026545	4	NM_013755
1	NM_009553	1	NM_177876	4	NM_178802
1	NM_008994	1	NM_011317	4	NM_013710
1	NM_170599	1	NM_011602	4	NM_007921
1	NM_001122758	1	NM_026921	4	NM_001032727
1	NM_008358	1	NM_009423	4	NM_176998
1	NM_001008422	1	NM_011368	4	NM_025313
1	NM_011580	1	NM_130456	4	NM_025367
1	NM_015748	1	NM_175534	4	NM_001081172
1	NM_009652	1	NM_010570	4	NM_007793
1	NM_145711	1	NM_026139	4	NM_013667
1	NM_019958	1	NM_145857	4	NM_001141950
1	NM_027952	1	NM_001024645	4	NM_027797
1	NM_025638	1	NM_001025307	4	NM_013628
1	NM_008984	1	NM_028732	4	NM_008552
1	NM_010768	1	NM_001012310	4	NM_134081
1	NM_133734	1	NM_173785	4	NM_008869
1	NM_133653	1	NM_001043336	4	NM_139198
1	NM_145136	1	NM_007987	4	NM_007685
1	NM_146386	1	NM_177630	4	NM_009061
1	NM_008885	1	NM_001112744	4	NM_172781
1	NM_199027	1	NM_172399	4	NM_007670
1	NM_144817	1	NM_001013022	4	NM_001113515
1	NM_007616	1	NM_011944	4	NM_031163

1	NM_009521	1	NM_021704	4	NM_198604
1	NM_153103	1	NM_010658	4	NM_133902
1	NM_133836	1	NM_153546	4	NM_001013376
1	NM_030713	1	NM_028071	4	NM_133191
1	NM_016791	1	NM_011512	4	NM_011799
1	NM_175133	1	NM_021557	4	NM_001025779
1	NM_001033246	1	NM_007435	4	NM_172152
1	NM_133805	1	NM_001042725	4	NM_178934
1	NM_026432	1	NM_029270	4	NM_013750
1	NM_001136471	1	NM_146161	4	NM_026823
1	NM_001083312	1	NM_019650	4	NM_019877
1	NM_145545	1	NM_013526	4	NM_010094
1	NM_010260	1	NM_013860	4	NR_003292
1	NM_028235	1	NM_133229	4	NM_007480
1	NM_172611	1	NM_011563	4	NM_008655
1	NM_177139	1	NM_021515	5	NM_010479
1	NM_175751			5	NM_019390

PUBLICATIONS

Cell Reports
Article

Totipotent Embryonic Stem Cells Arise in Ground-State Culture Conditions

Sophie M. Morgani,[1,2] Maurice A. Canham,[1] Jennifer Nichols,[3] Alexei A. Sharov,[4] Rosa Portero Migueles,[1] Minoru S.H. Ko,[4,5] and Joshua M. Brickman[1,2,*]

[1]MRC Centre for Regenerative Medicine, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, 5 Little France Drive, EH16 4UU Edinburgh, UK
[2]The Danish Stem Cell Centre, DanStem, University of Copenhagen, 3B Blegdamsvej, DK-2200 Copenhagen N, Denmark
[3]Wellcome Trust Centre for Stem Cell Research, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK
[4]Laboratory of Genetics, National Institute on Aging, NIH Biomedical Research Centre, 251 Bayview Boulevard, Suite 100, Baltimore, MD 21224, USA
[5]Department of Systems Medicine, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160, Japan
*Correspondence: joshua.brickman@sund.ku.dk
http://dx.doi.org/10.1016/j.celrep.2013.04.034

SUMMARY

Embryonic stem cells (ESCs) are derived from mammalian embryos during the transition from totipotency, when individual blastomeres can make all lineages, to pluripotency, when they are competent to make only embryonic lineages. ESCs maintained with inhibitors of MEK and GSK3 (2i) are thought to represent an embryonically restricted ground state. However, we observed heterogeneous expression of the extraembryonic endoderm marker *Hex* in 2i-cultured embryos, suggesting that 2i blocked development prior to epiblast commitment. Similarly, 2i ESC cultures were heterogeneous and contained a *Hex*-positive fraction primed to differentiate into trophoblast and extraembryonic endoderm. Single *Hex*-positive ESCs coexpressed epiblast and extraembryonic genes and contributed to all lineages in chimeras. The cytokine LIF, necessary for ESC self-renewal, supported the expansion of this population but did not directly support *Nanog*-positive epiblast-like ESCs. Thus, 2i and LIF support a totipotent state comparable to early embryonic cells that coexpress embryonic and extraembryonic determinants.

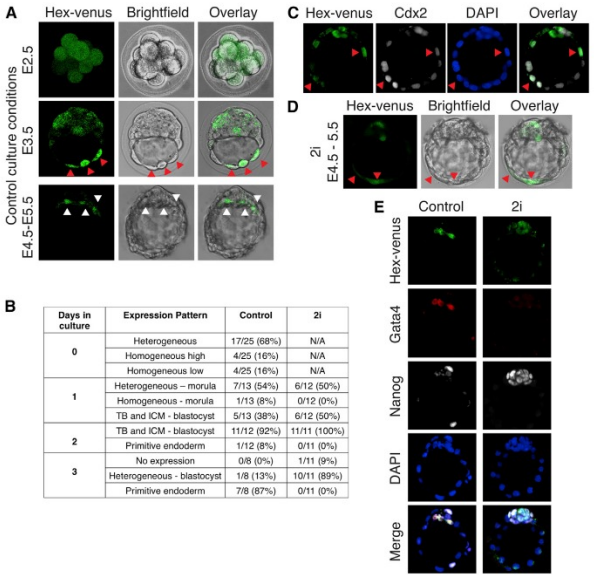
INTRODUCTION

Embryonic stem cells (ESCs) are karyotypically normal cells derived from the inner cell mass (ICM) or epiblast of peri-implantation embryos. They are classically defined as pluripotent. In mouse, this is assessed by their ability to differentiate into all embryonic but not extraembryonic lineages when reintroduced into morulae or blastocysts. However, because this definition is based on retrospective function, the precise cellular phenotype of pluripotent cells is unknown, especially as there has been little characterization of the functional potential of single cells. Additionally, although ESCs have been shown to be pluripotent only in standard injections, they can generate extraembryonic primitive endoderm (PE) in vitro (Hayashi et al., 2010; Xu et al., 2002) and, in rare events, contribute to both embryonic and extraembryonic lineages in vivo (Beddington and Robertson, 1989; Canham et al., 2010; Lallemand and Brûlet, 1990; Macfarlan et al., 2012; Suemori et al., 1990). This suggests that ESC cultures contain a mixture of cells resembling precursors of embryonic epiblast and extraembryonic tissues but that epiblast-like precursors have a competitive advantage when reintroduced into chimeras.

ESCs can be maintained in the presence of LIF and either BMP4 or serum (Ying et al., 2003a). Under these conditions, numerous genes are expressed in a heterogeneous manner (Canham et al., 2010; Chambers et al., 2007; Hayashi et al., 2008; Kobayashi et al., 2009; Singh et al., 2007; Toyooka et al., 2008), implying that ESC cultures may harbor cells with distinct functional potentials. ESCs can also be cultured in minimal medium with MEK (mitogen-activated protein or extracellular signal-regulated kinase) and GSK3 (glycogen synthase kinase 3) inhibitors (2i) (Ying et al., 2008). These inhibitors shield ESCs from differentiation-inducing signals and are thought to generate a homogeneous early epiblast-like ground state for embryonic but not extraembryonic development (Nichols et al., 2009; Wray et al., 2010, 2011). Culture in 2i is often supplemented with LIF, which not only supports self-renewal of ESCs but also has a function in extraembryonic development, where it promotes trophoblast proliferation, differentiation, and invasion (Poehlmann et al., 2005; Prakash et al., 2011; Takahashi et al., 2003).

We previously described a sensitive reporter for the endoderm marker *Hex* utilizing a reiterated IRES element to translationally amplify expression of the fluorescent protein Venus, encoded downstream of *Hex* in the endogenous locus (Canham et al., 2010). Here, we utilize ESCs containing this reporter, and a transgenic reporter mouse derived from them, to explore the nature of the ground state and investigate the cell-intrinsic role of LIF in this defined context. We show that embryos and ESCs cultured in 2i are heterogeneous and contain a fraction of cells coexpressing markers of both embryonic and extraembryonic lineages. This population demonstrated an enhanced capacity to generate extraembryonic cell types, including trophoblast, in vitro, and single cells from this fraction were totipotent when assessed by morula aggregation in vivo. Thus,

Cell PRESS
Open ACCESS



B

Days in culture	Expression Pattern	Control	2i
0	Heterogeneous	17/25 (68%)	N/A
	Homogeneous high	4/25 (16%)	N/A
	Homogeneous low	4/25 (16%)	N/A
1	Heterogeneous – morula	7/13 (54%)	6/12 (50%)
	Homogeneous - morula	1/13 (8%)	0/12 (0%)
	TB and ICM - blastocyst	5/13 (38%)	6/12 (50%)
2	TB and ICM - blastocyst	11/12 (92%)	11/11 (100%)
	Primitive endoderm	1/12 (8%)	0/11 (0%)
3	No expression	0/8 (0%)	1/11 (9%)
	Heterogeneous - blastocyst	1/8 (13%)	10/11 (89%)
	Primitive endoderm	7/8 (87%)	0/11 (0%)

Figure 1. Culture in 2i Does Not Eliminate Endoderm Marker Expression in the Early Embryo

(A) HV transgenic embryos were collected at E2.5 and cultured in vitro for different time periods before live confocal microscopy imaging. Representative images are shown for each stage. At E2.5, an extended focus image is shown. E3.5 and E4.5 images show confocal optical sections.

(B) Scoring of the HV expression pattern in embryos collected at E2.5 and cultured in either control or 2i medium. Day 0 = E2.5, day 1 = E3.5 (early blastocyst), day 2 = E4.5 (midblastocyst), and day 3 = E5.5 (late blastocyst, equivalent to E4.5 in vivo due to culture induced delay). Only HV-expressing transgenic embryos were scored. TB, trophoblast; ICM, inner cell mass.

(C) Confocal optical section of an E2.5 HV embryo cultured for 3 days immunostained for the trophoblast marker CDX2. Red arrowheads indicate trophoblast cells coexpressing HV and CDX2. Although in vivo PE is specified by E4.5, this segregation is delayed in vitro, and, hence, embryos were cultured until E5.5 to observe differentiated PE.

(D) HV embryos were flushed at E2.5 and cultured in 2i media for 3 days before live confocal microscopy imaging. Red arrowheads indicate HV-expressing trophoblast cells, whereas white arrowheads indicate HV^+ PE cells.

(E) Confocal optical sections of immunostained transgenic HV embryos after flushing at E2.5 and 3 days in culture either in control or 2i medium.

the combination of 2i and LIF promoted the expansion of individual totipotent cells reminiscent of the morula or early blastocyst stage, before lineage restrictions have occurred.

RESULTS

Preimplantation Embryo Culture in 2i Captures an Early Blastocyst Stage of Development

We generated a transgenic mouse line from our *Hex*-Venus (HV) reporter ESCs and assessed HV expression in preimplantation embryos. HV expression was observed at embryonic day (E) 2.5 (Figures 1A and 1B) and, in the majority of cases, was heterogeneous (Figures 1A and 1B). On average, 50% of blastomeres expressed HV at E2.5, although there was variation in the proportions between individual embryos. At E3.5, HV was expressed heterogeneously in both the ICM and trophoblast, being coexpressed with the trophoblast marker CDX2 in a number of cells (Figures 1A–1C). At E3.5, around one-third of the cells of the embryo expressed HV. By the late blastocyst stage (E4.5–E5.5), HV expression became restricted to the PE (Figures 1A, 1B, and 1E).

Embryo culture in 2i has been reported to support the expansion of a homogeneous epiblast-like state at the expense of the expression of the late PE marker GATA4 (Nichols et al., 2009). We cultured transgenic HV embryos in 2i to determine whether the ICM differentiates into homogeneous epiblast. Embryos cultured in 2i, from E2.5 for 3 days, maintained heterogeneous HV expression (Figures 1B, 1D, and 1E) but did not segregate the PE and epiblast lineages, demonstrated by the absence of the late PE marker GATA4 (Figure 1E). Thus, 2i appears to block the progression of cultured embryos toward later blastocyst stages but does not eliminate the early endodermal precursor population.

Ground State In Vitro ESC Cultures Are Heterogeneous and Contain an Extraembryonically Primed Subpopulation

We then asked whether the in vitro ground state observed in ESCs is comparable to an early blastocyst stage in vivo, containing cells heterogeneously expressing low-level extraembryonic and pre-epiblast determinants (Chazaud et al., 2006). Previous observations suggested that ESCs in 2i maintained the expression of some endoderm markers (Canham et al., 2010; Marks et al., 2012). To assess the degree to which this represented heterogeneous gene expression, we clonally rederived HV ESC lines in 2i. We confirmed that ESCs in 2i demonstrated a block to extracellular signal-regulated kinase (ERK) signaling (Figure S1A) and found that at least nine independent clones showed heterogeneous expression of HV as judged by flow cytometry (gating methodology shown in Figures 2A and S1B–S1F). This was assessed for all three conditions (serum/LIF, 2i, and 2i/LIF). When single HV low (HV^-) or HV high (HV^+) cells were cultured in 2i or 2i/LIF, they regenerated heterogeneous cultures, although a proportion of cells maintained a bias toward

regenerating mixed populations that contained a higher fraction of the seeding cell type (Figure 2B).

We asked whether HV^+ and HV^- cells derived from serum/LIF or 2i/LIF culture conditions represented equivalent cell types by carrying out RNA sequencing (RNA-seq) genome-wide expression analysis on each population (Gene Expression Omnibus accession number GSE45182). Hierarchical clustering of gene expression based on an alternative pairwise comparison of the different populations (Figure 2C) suggested that HV^+ and HV^- states were more distinct in 2i/LIF than in serum/LIF (>2-fold change, false discovery rate [FDR] 0.05; Figure 2C; Tables S1 and S2).

While in this data set there was little change in early epiblast or pluripotency genes between the sorted 2i populations, we observed more than 20 imprinted genes that were enriched in the 2i/LIF HV^+ compared to 2i/LIF HV^- population; for example, the *Dlk1-Dio3* cluster (Figures S2A and S2B) associated with efficient reprogramming (Liu et al., 2010). We also observed increased levels of trophoblast gene expression in the 2i/LIF HV^+ population, including markers specifically expressed in trophoblast stem cells (Rugg-Gunn et al., 2012) (Figure S2C). In addition, endogenous retroviral (ERV) genes, enriched in an ESC population comparable to the two-cell-stage embryo (Macfarlan et al., 2012), such as *Abcb5*, *AD067063*, and *Gm10696*, were upregulated in the 2i/LIF HV^+ population (Figure S2D). However, unlike the population of ERV-marked cells, 2i/LIF ESCs also continued to express pluripotency markers.

This coexpression of pluripotency genes and trophoblast determinants is reminiscent of the stages of preimplantation development when blastomeres are competent to make all lineages. As ESCs are not thought to be able to generate trophoblast, we asked if 2i/LIF HV^+ cells could differentiate into trophoblast in vitro. Figures 2D and 2E show that HV^+ cells generated 40-fold more $CDX2^+$ cells than HV^- cells in trophoblast stem cell conditions (Quinn et al., 2006). $CDX2^+$ cells appeared to be trophoblast-like, coexpressing neither the endoderm marker GATA6 nor the mesoderm marker BRACHYURY (Figure S2E; data not shown). We also observed that, upon differentiation by LIF withdrawal, only HV^+ cells from 2i produced robust levels of trophoblast gene expression (Figure S2F). These observations revealed that HV^+ ESCs in serum/LIF and 2i/LIF are fundamentally different from each other in both gene expression and functional capabilities, with cells from 2i/LIF demonstrating the additional capacity to generate trophoblast in vitro.

To determine whether 2i/LIF HV^+ cells were restricted to the trophoblast lineage, we assessed their capacity to differentiate into endoderm and the epiblast-derived neural lineage. We observed a marked bias of the HV^+ population to form endoderm, whereas the HV^- population was biased toward a neural fate, even after prior culture in 2i (Figures 2F, 2G, and S3A–S3G; $p < 0.001$). Levels of differentiation were scored based on the number of $GATA6^+$ cells (Figure S3B), $GATA6^+$ colonies (Figure 2F), gene expression (Figures S3E and S3F), and flow cytometry to quantify the expression of an endodermal cell surface marker (Figure S3G). Absolute levels of differentiation were also higher in cells differentiated from 2i (Figures 2E–2G; $p < 0.001$).

ESCs Cultured in 2i Can Contribute to Both Embryonic and Extraembryonic Lineages

We tested the capacity of 2i-cultured HV^- and HV^+ ESCs to colonize embryonic and extraembryonic lineages in aggregation chimeras using HV ESCs constitutively expressing LACZ (Figures 3A and 3B). We previously found that HV^- cells in serum/LIF contributed efficiently to the epiblast, whereas HV^+ cells contributed to the extraembryonic endoderm and only weakly to epiblast (Canham et al., 2010). While HV^- cells in 2i displayed high-level epiblast contribution (Figure 3A), HV^+ cells from 2i contributed efficiently to the epiblast and to all extraembryonic lineages (i.e., exhibited totipotent properties) (Figures 3A, 3B, and S4A). In a large proportion of chimeras generated from the HV^+ population, cells were detected in the trophoblast as well as visceral and parietal endoderm (n = 10/26; Figures 3A, 3B, and S4A), while the HV^- population from 2i contributed only to epiblast (n = 18; Figure 3A). We assessed cells cultured in 2i as well as in 2i/LIF and found enhanced extraembryonic contribution from cells grown in 2i/LIF (n = 30/55) compared to those in 2i alone (n = 10/26). In 2i/LIF, a proportion of the HV^- cells also contributed to extraembryonic lineages, although less efficiently than HV^+ cells (n = 18/60; Figure 3B).

We additionally generated chimeras using HV cells expressing a constitutive H2B-Tomato fluorescent protein (Figures 3C–3F and S4B–S4D). We sorted HV^+ cells from 2i/LIF conditions, aggregated these cells with wild-type morulae, and assessed chimera contribution by fluorescence and immunohistochemistry. Whole-mount immunostaining for the extraembryonic endoderm marker GATA6 and trophoblast marker KRT7 showed that H2B-Tomato ESCs in the extraembryonic region were found to express either GATA6 or KRT7 (Figures S4C and S4D). In late blastocysts, H2B-Tomato ESCs were found both in the epiblast and extraembryonic lineages (Figure 3C), and in E9.5 embryos, H2B-Tomato cells were integrated into the placenta and the yolk sac (Figures 3D–3F).

Single Cells Cultured in 21 Are Totipotent

Although we observed contribution of 2i-cultured ESCs to both embryonic and extraembryonic lineages, the definition of totipotency is based on the capacity of a single cell to contribute to all lineages. To distinguish between the presence of individual totipotent cells or a population-based explanation for the totipotent activity of 2i ESC cultures, we assessed gene expression and functional properties of single ESCs. As totipotent cells in the early embryo coexpress both embryonic and extraembryonic determinants, we examined gene expression in single HV^- and HV^+ sorted ESCs from serum/LIF or 2i to ask if this is also the case in our ESC cultures. For this experiment, a BioMark HD System (Fluidigm) was used with custom designed DELTAgene Assay primer pairs (Fluidigm; sequences are provided in Extended Experimental Procedures). Consistent with our RNA-seq data, HV^+ and HV^- populations were more distinct in 2i than in serum conditions (Figure 4A). The extraembryonic markers *Eomes*, *Gata3*, *Serpine2*, and *Tcfap2a* showed considerable variability across all four populations, with low or no expression in HV^- and HV^+ cells from serum as well as HV^- cells in 2i but dramatically enhanced expression in HV^+ cells from 2i (Figure 4A). We also observed single-cell enrichment in the

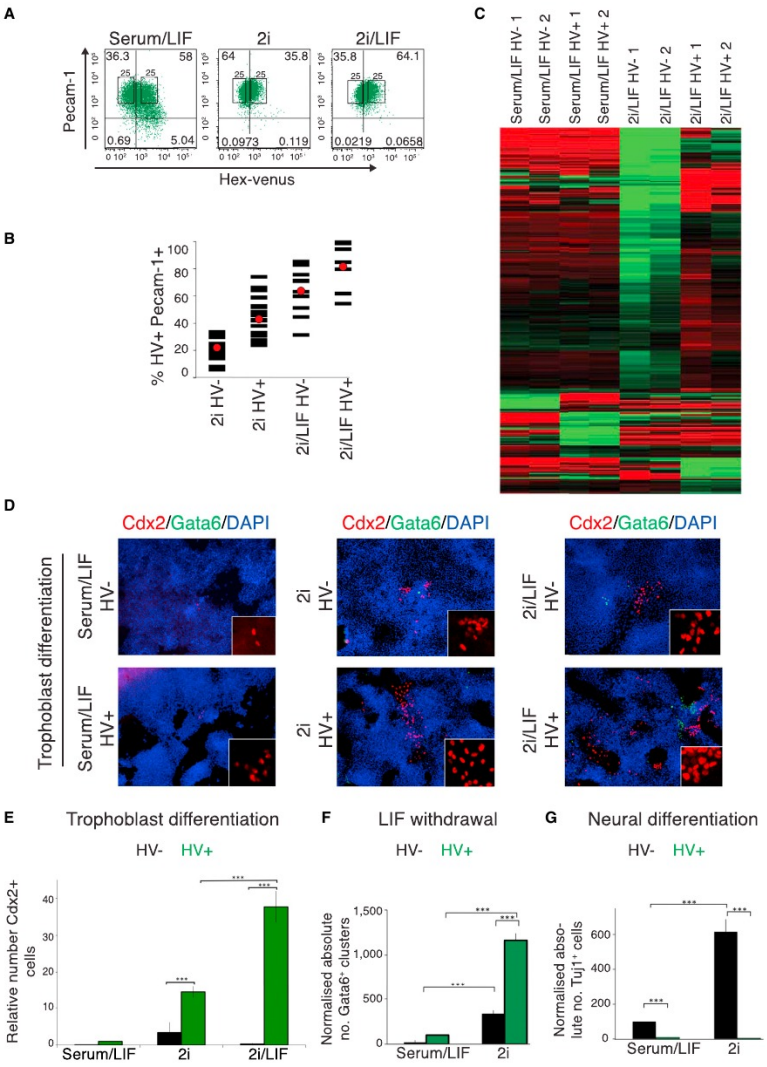


Figure 2. ESCs Cultured in 2i Are Heterogeneous

(A) Flow cytometry of HV and ESC marker PECAM-1 in cells cultured in serum/LIF, 2i, and 2i/LIF. Gates set using unstained E14 ESCs (Figure S1B). Black boxes indicate sorting gates for the upper and lower 25% of HV expression used to separate HV^- and HV^+ populations.

(B) Flow cytometry of clones (represented by individual bars) after expansion from single HV^- or HV^+ sorted cells from 2i or 2i/LIF. Red circles indicate mean.

(legend continued on next page)

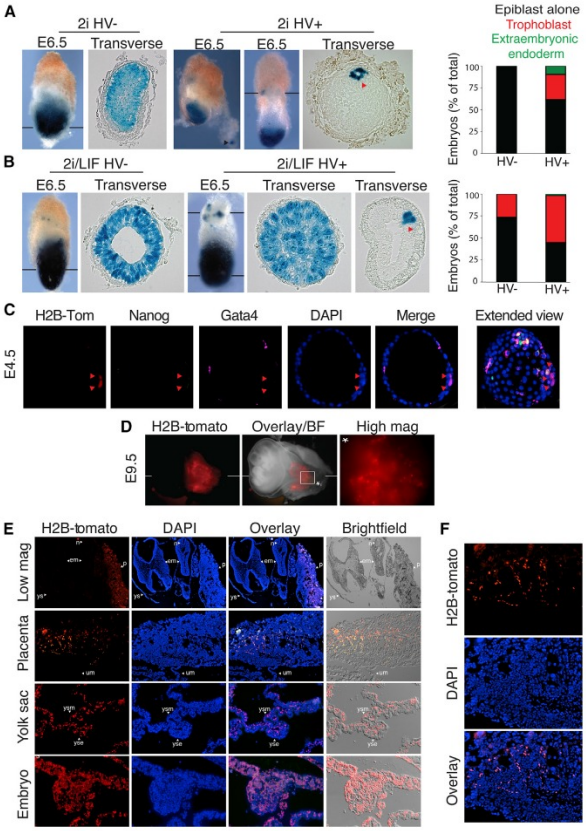


Figure 3. Subpopulations of ESCs Cultured in 2i Demonstrate Extraembryonic Potential In Vivo

(A and B) HV cells, expressing a constitutive LACZ marker, cultured in 2i (A) or 2i/LIF (B), were sorted via FACS into HV^- (A: n = 18; B: n = 60) and HV^+ (A: n = 26; B: n = 55) PECAM-1^+ cells and aggregated with wild-type F1 morulae. Embryos were collected at E6.5 and X-gal stained. Black bars indicate the plane of section. Arrowheads mark extraembryonic endoderm (black) and trophoblast (red).

(C–F) Sorted HV^+ cells expressing a constitutive H2B-Tomato fusion protein, cultured in 2i/LIF, were used in morula aggregations. Embryos were cultured until E4.5 (C) or dissected at E9.5 (D). (C) Confocal optical section of an immunostained chimera after 3 days of culture (equivalent to E4.5 in vivo). Red arrowheads indicate H2B-Tomato cells within the trophoblast. Extended view throughout the embryo also shown for orientation. (D) E9.5 chimera. White line indicates plane of the low-magnification section in (E). White box indicates area of placenta shown in the high-magnification image marked with an asterisk (*). (E) Wax sections of an E9.5 chimeric embryo showing HV H2B-Tomato cells in the placenta, yolk sac, and embryo. The uppermost panels show a low-magnification section of the whole embryo (shown in D) within the yolk sac and attached to the placenta. ys, yolk sac; yse, yolk sac endoderm; ysm, yolk sac mesoderm; em, embryo; n, notochord; p, placenta; um, umbilical cord. (F) High-magnification image of placental section.

See also Figure S4.

imprinted gene *Dlk1*, identified based on our global expression analysis, specifically in 2i HV^+ cells (Figures 4A and S5A). Taken together, our single-cell data showed that individual cells in the 2i HV^+ population simultaneously coexpressed epiblast and extraembryonic genes (Figures 4A and S5A). This reinforces the notion that HV^+ cells in 2i are distinct from the HV^+ population in serum and suggests that in 2i these single cells may possess totipotent properties.

Oct4 and *Sox2* showed relatively uniform expression across all populations, while epiblast markers shown to be heterogeneously expressed, such as *Nanog* and *Stella*, exhibited more variable expression (Figure 4A). However, when 2i was compared to serum, we observed that *Nanog* expression became more homogenous. A total of 3 out of 30

(C) Heat map of sorted HV^- and HV^+ populations from serum/LIF and 2i/LIF culture, based on gene expression data from RNA-seq. Heat map shows differentially expressed genes identified by pairwise comparison of all sorted fractions. Data were normalized by subtracting the average log expression from all samples. Genes are hierarchically clustered by average Euclidean distance. Two biological replicates are shown per sample. Red represents upregulation and green represents downregulation of expression.

(D) Immunostaining displaying representative images of CDX2-positive cells in trophoblast differentiation. High-magnification images are shown as insets.

(E and F) Quantification of in vitro differentiation of sorted HV^- and HV^+ ESCs after differentiation in trophoblast stem cell conditions for 7 days (E), LIF withdrawal (F), or neural differentiation (G). (E) Trophoblast differentiation was quantified by counting CDX2-positive cells, which were negative for GATA6 and BRACHYURY (n = 3). Values are shown relative to HV^+. (F) Endoderm differentiation was quantified by counting the number of GATA6-positive endodermal clusters (see also Figures S3A and S3C; n = 3). Values are shown relative to HV^-.

(G) Neural differentiation was quantified by counting elongated TUJ1-positive neurons (n = 5). Values are shown relative to HV^-. Error bars indicate mean ± SD of biological replicates.

See also Figures S1, S2, and S3.

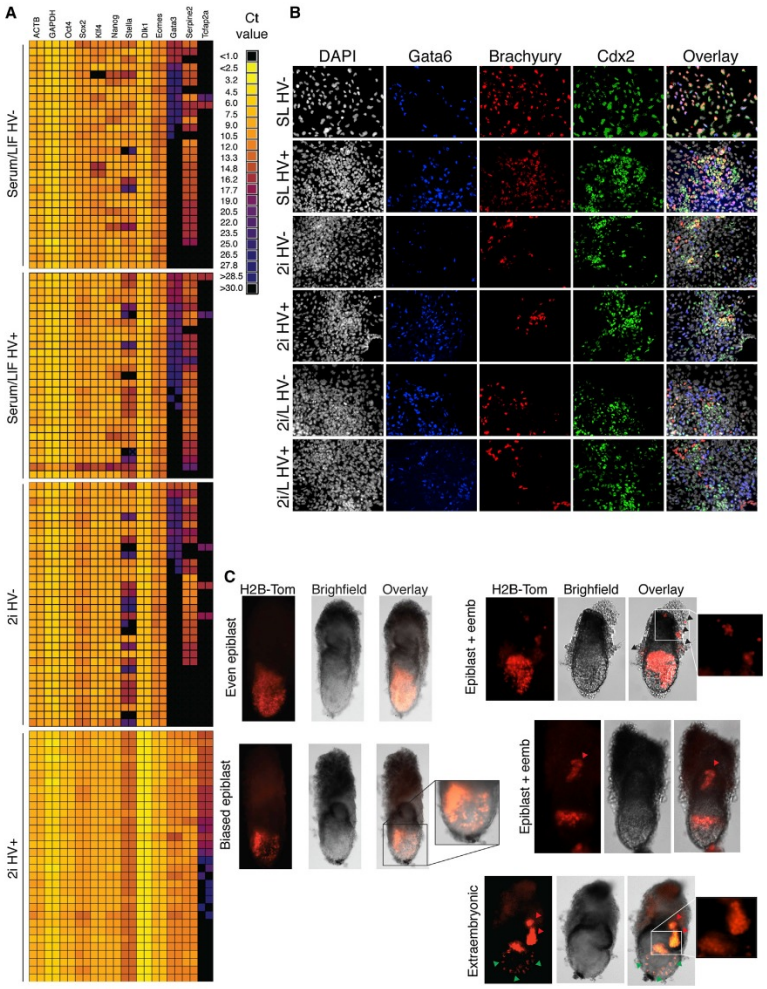


Figure 4. Single HV$^+$ Cells in 2i Can Contribute to Both Epiblast and Extraembryonic Lineages

Single HV$^-$ or HV$^+$ cells from serum/LIF or 2i were sorted by FACS and analyzed by qRT-PCR using the Biomark HD system (Fluidigm).

(A) Heat map for each condition and cell type showing the expression of genes with reliable primer melting curves. Adjacent squares show technical replicates for each cell.

(B) Single cells were sorted by flow cytometry into LIF withdrawal conditions. After 7 days, differentiated colonies were immunostained for GATA6, CDX2, and BRACHYURY.

(C) Single-sorted HV$^+$ cells (sorted from the top 10% of HV expression), previously cultured in 2i/LIF, constitutively expressing H2B-Tomato were injected into morulae and dissected at E6.5 and chimeras (n = 23/64) were assessed for lineage contribution.

(legend continued on next page)

ESCs in the serum/LIF HV^+ population showed very low or no *Nanog* expression (data not shown). This was not seen in any other population analyzed. However, these *Nanog*-negative cells, although expressing genes including *Oct4*, showed lower levels of housekeeping genes and hence were excluded from further analysis.

Based on the coexpression of epiblast and extraembryonic genes in single cells, we asked whether individual cells grown in 2i could differentiate into both embryonic and extraembryonic cells in vitro. Single cells that had previously been cultured in serum/LIF, 2i, or 2i/LIF were sorted into medium without LIF and differentiated under clonal conditions. Multilineage differentiation was quantified based on the expression pattern of specific lineage markers (GATA6, CDX2, and BRACHYURY) within differentiated colonies (Figures 4B and S5B). Coexpression of these markers in the same cell is likely to represent mesendoderm, while the presence of all three in separate cells would suggest the presence of PE, trophoblast, and epiblast/mesoderm. Cells that had previously been cultured in serum/LIF generated colonies either exclusively expressing one or two markers or coexpressing multiple markers within the same cells. Conversely, cells cultured in 2i or 2i/LIF showed a decrease in single-marker expression and coexpression in favor of mutually exclusive expression of all three lineage markers within the same colony (Figures 4B and S5B). Thus, individual 2i-cultured and, in particular, HV^+ cells have the capacity to generate all three lineages: epiblast, trophoblast, and PE. Additionally, cells cultured in 2i before differentiation showed an increase in the expression of CDX2 alone, indicative of trophoblast differentiation, compared to cells previously cultured in serum (data not shown). While this in vitro assay gives a crude estimate of the potency of individual ESCs, it is difficult to assign totipotency based solely on marker expression in the absence of embryonic context.

We therefore also assessed totipotency in vivo by asking whether single ESCs could contribute to all lineages when reintroduced into embryos. Single-cell morula injections of 2i/LIF HV^+ cells were carried out and resulting chimeric E6.5 embryos were scored for contribution to different lineages (Figure 4C). More than half of the chimeras showed evidence of extraembryonic contribution (n = 23). One embryo showed contribution only to extraembryonic tissue, both endoderm and trophoblast, while 13 showed contribution both to extraembryonic tissue and epiblast (Figures 4C and S5C), indicating that this fraction contained a significant proportion of cells with extraembryonic potential. Of these 13, two showed contribution to the endoderm, eight showed trophoblast contribution, and three showed contribution to both the endoderm and trophoblast. Single HV^- cells cultured in 2i/LIF were also injected as a control and, as observed for the multiple cell injections, only a small fraction of cells were able to contribute to the extraembryonic region (Figure S5D; n = 13).

LIF Supports Extraembryonic Priming

In our assessment of clonal ESC potential, we observed a significant enhancement of the potential of cells grown in LIF to generate the extraembryonic lineages (Figures 3A, 3B, and 4B). When either 2i or serum-containing medium was supplemented with increasing doses of LIF, the expression of the HV transgene and other extraembryonic markers increased (Figures 2A, 5A–5D, and S6A–S6C). However, a similar dose-response experiment using a *Nanog*-GFP ESC reporter line (Chambers et al., 2007) showed no change in GFP levels, even when LIF was present at a 5-fold excess of its normal saturating dose (Figures 5A and S6B). In both 2i and 2i/LIF, the majority of HV^+ cells expressed NANOG, with this coexpressing cell population representing approximately 15%–20% of the culture (Figures 5C and S6D). In serum/LIF, NANOG and HV expression were predominantly mutually exclusive (Figure 5C).

As the ability of LIF to mediate self-renewal is thought to be dependent on STAT3, we examined gene expression in $Stat3^{-/-}$ ESCs (Ying et al., 2008) compared to wild-type cells derived at the same time. $Stat3^{-/-}$ ESCs exhibited decreased levels of extraembryonic gene expression (endodermal and trophoblast) (Figure 5D), but, other than the STAT3 target gene *Klf4*, pluripotency markers were not affected.

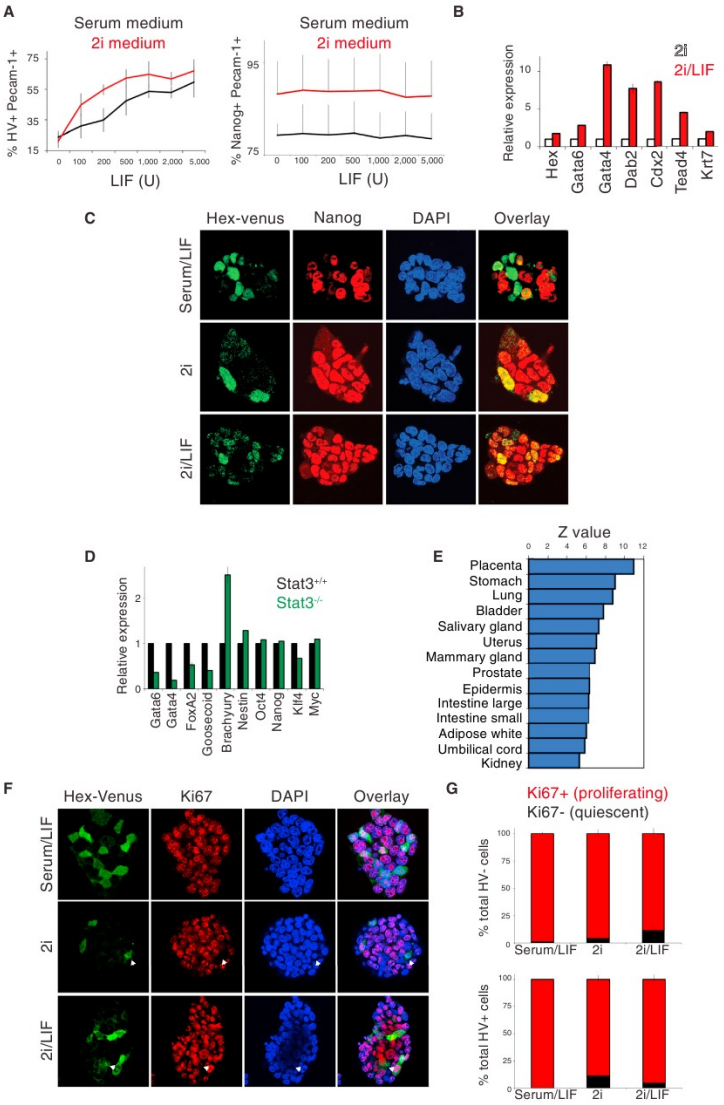
Additionally, preculture of ESCs in 2i/LIF rather than 2i alone before differentiation by LIF withdrawal caused a significant increase in the levels of endoderm generated (Figures S6E and S6F).

We assessed the degree to which LIF induced extraembryonic gene expression in ESC culture by whole genome expression analysis, using RNA-seq, on unsorted ESCs cultured in 2i or 2i/LIF. Extraembryonic endoderm and trophoblast genes were both upregulated upon the addition of LIF, while neuroectoderm and mesoderm markers were reduced (Figure S6G). Moreover, genes upregulated by LIF in 2i demonstrated the strongest correlation with placental gene expression (Genomics Institute of the Novartis Research Foundation [GNF] Gene Expression Database; Figure 5E). In 2i, LIF increased the proportion of proliferating HV^+ cells while decreasing the proportion of HV^- proliferating cells (Figures 5F and 5G), suggesting that LIF promoted the expansion of the HV^+ totipotent population.

DISCUSSION

In this paper, we demonstrated that ESCs and embryos grown in 2i are heterogeneous with respect to extraembryonic gene expression. We identified a LIF-promoted population containing single cells that could give rise to trophoblast and PE as well as epiblast, a behavior characteristic of totipotent cells. This suggests that HV^+ cells cultured in 2i may reflect an earlier developmental stage than widely believed to exist in ESC culture and that the role of LIF in supporting self-renewal may be a consequence

Examples of different contribution patterns of single cells are shown. Some cells showed even contribution across the whole epiblast while others showed a contribution that was biased toward a particular area of the epiblast. Red arrowheads indicate trophoblast contribution, green arrowheads indicate visceral endoderm contribution, and black arrowheads indicate parietal endoderm contribution. Epiblast contribution alone was observed in nine chimeras, extraembryonic contribution alone in one chimera, and 13 chimeras showed both epiblast and extraembryonic contribution.
See also Figure S5.



(legend on next page)

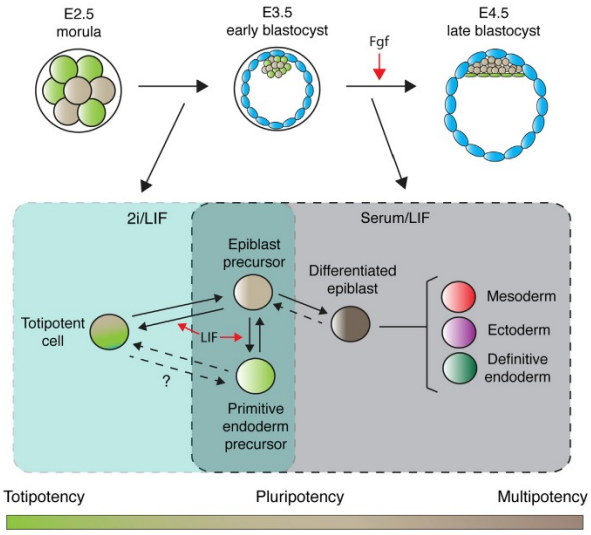


Figure 6. Model Comparing Standard Serum/LIF ESC Culture to 2i ESC Culture

The model relates serum/LIF and 2i ESC cultures to embryonic development. 2i/LIF cultures contain totipotent single cells that are comparable to early preimplantation stages when embryonic cells are still totipotent. As Fgf is needed for later lineage segregation, inhibiting Fgf signaling in 2i seems to block embryos and ESCs at a point prior to this commitment event. Serum/LIF ESC cultures are reminiscent of later embryonic stages when cells are pluripotent and contain more restricted cell types. Although partially differentiated mesoderm and ectoderm appear to be lost in 2i/LIF, endoderm is maintained. Based on our data, totipotent ESCs are able to dynamically interconvert to a pluripotent epiblast precursor state and may also exist in equilibrium with a similar primitive endoderm precursor, although we have not been able to distinguish this population in this study. LIF supports the expansion of the totipotent ESC state and primitive endoderm precursors in serum.

of its capacity to support this population. A model illustrating these findings is shown in Figure 6.

Culture in 2i is purported to maintain a "naïve" pluripotent state by shielding ESCs from differentiation-promoting signals (Nichols et al., 2009; Wray et al., 2011), but despite more homogeneous *Nanog* expression, there is little evidence that cells in 2i represent a single cell type. Moreover, 2i culture is regularly supplemented with LIF (Nichols et al., 2009; Ying et al., 2008) and, while LIF is known to upregulate pluripotency markers (Hall et al., 2009), its role in vivo is to support extraembryonic development (Stewart et al., 1992; Takahashi et al., 2008). In embryoid body differentiation, LIF selectively blocks primitive ectoderm differentiation while permitting PE differentiation (Shen and Leder, 1992). Additionally, the downstream effector of LIF, STAT3, binds directly to extraembryonic gene promoters such as *Gata6*, *Gata3*, and *Eomes* (Kidder et al., 2008). Thus, the role of LIF in supporting ESC culture may be mediated via a totipotent, extraembryonically primed cell type capable of effectively expanding and dynamically generating the heterogeneous distribution of cells normally observed in ESC culture.

Preimplantation embryos cultured in 2i retained the early heterogeneous expression of *Hex*, and ground state ESCs contained a population capable of generating PE and trophoblast. This suggests that 2i blocks the commitment of totipotent cells to embryonic or extraembryonic lineages. Cells and embryos maintain the coexpression of epiblast and extraembryonic markers that occurs in early preimplantation development and has been observed for the embryonic markers NANOG and OCT4 and the extraembryonic marker CDX2 (Dietrich and Hiiragi, 2007). The 2i MEK inhibitor PD0325901 blocks fibroblast growth factor signaling, which is important for the

Figure 5. LIF Supports an Extraembryonically Primed Population of Cells

(A) Flow cytometry of HV or TNGB *Nanog*-GFP cells cultured in serum or 2i medium, with increasing LIF concentrations (n = 3). Raw data and controls are in Figures S6A–S6C. Error bars indicate mean ± SD of two biological replicates.

(B) qRT-PCR showing relative gene expression of extraembryonic markers in ESCs cultured in 2i or 2i/LIF. Data were normalized to the housekeeping gene TBP and are shown relative to 2i samples. Error bars indicate mean ± SD of biological replicates.

(C) Confocal optical section showing immunostaining of HV cells for NANOG after culture of cells in serum/LIF, 2i, or 2i/LIF.

(D) qRT-PCR showing relative gene expression in Stat3$^{+/+}$ or $^{-/-}$ ESCs. Data are shown relative to wild-type expression levels. All qRT-PCR transcript levels in this figure were normalized to TBP and are shown as mean ± SD.

(E) Statistical significance of correlation of gene expression differences between ESCs cultured in the presence and absence of LIF and tissue-specific gene expression in the GNF Ver.3 database. Correlation was estimated from data on 5,085 genes, which represent the overlap between 10,000 most informative genes in our RNA-seq data and 10,000 most informative genes in the GNF database. Organs/tissues sorted by decreasing correlation.

(F) Confocal optical section of HV cells immunostained for the proliferation marker KI67. White arrowheads indicate KI67-negative cells.

(G) Quantification of KI67 immunostaining in HV^{+} and HV^{-} ESCs. The absolute number of proliferating cells (expressing KI67, red bar) and quiescent cells (not expressing KI67, black bar) were counted. Error bars indicate SD of the mean from five colonies in each condition, analyzed by confocal microscopy.

See also Figure S5.

segregation of epiblast and PE (Hamazaki et al., 2006; Yamanaka et al., 2010; Chazaud et al., 2006; Nichols et al., 2009; Yamanaka et al., 2010). It is also important for the support and expansion of trophoblast (Quinn et al., 2006); thus, blocking this signal could halt embryonic-extraembryonic lineage segregation. Similar observations have been made for another PE reporter, *PDGFR-alpha*-GFP, where embryos cultured in 2i continued to express this marker heterogeneously within the ICM, although this expression was attributed to the persistence of the stable GFP (Yamanaka et al., 2010). We found that this low-level expression of extraembryonic markers correlated with totipotent function in ESCs, reminiscent of the activity of primed blastomeres expressing low-levels of *PDGFR-alpha*, shown to have greater lineage flexibility than their epiblast-primed counterparts (Grabarek et al., 2012).

Although it has been shown previously that ESCs can contribute to extraembryonic lineages (Beddington and Robertson, 1989; Canham et al., 2010; Lallemand and Brûlet, 1990; Macfarlan et al., 2012; Suemori et al., 1990), these events occurred at a low frequency. Thus, totipotent cells could exist in most ESC cultures, but because they are relatively rare, they are frequently not acknowledged in literature. Although relatively rare in standard culture conditions, our data suggest that these remarkable individual ground state ESCs have the capacity to generate the majority of the epiblast. While the extent of extraembryonic contribution that we observed was never as great as that which we observed in the epiblast, the contribution to trophoblast was at least comparable to that observed in chimeras generated from trophoblast stem (TS) cells (Quinn et al., 2006) and the extent of visceral endoderm contribution to that observed in chimeras generated from undifferentiated extraembryonic endoderm cells (Kunath et al., 2005).

It has been noted that some cells within standard ESC cultures express embryonic two-cell-stage transcripts and can contribute to both embryonic and extraembryonic lineages (Macfarlan et al., 2012). These cells lack expression of epiblast markers such as OCT4, NANOG, and SOX2, although these genes are expressed from early in mammalian development and throughout the period that mammalian embryos are considered totipotent. Rossant et al. demonstrated that, up until the early blastocyst stage, embryonic cells posses a similar level of flexibility to that demonstrated here for HV^+ ground state ESCs. ICM cells contribute to the trophoblast in morula aggregations and isolated early ICMs implant into the uterus (Rossant and Lis, 1979). Additionally, the majority of outer trophoblast progenitor cells can contribute to the ICM and epiblast in blastocyst injections (Rossant and Vijh, 1980). We also observed that, at this early stage of blastocyst formation, a subset of trophoblast cells expressed high levels of NANOG in conjunction with CDX2 (Figure S4E). Similarly, NANOG, OCT4, and CDX2 have been observed to be coexpressed throughout early preimplantation development in both ICM and trophoblast cells (Dietrich and Hiiragi, 2007). Thus, at these stages, the ICM and trophoblast appear to retain plasticity, which is lost as embryos progress to the late blastocyst (Handyside, 1978; Hogan and Tilly, 1978; Rossant and Lis, 1979; Spindle, 1978), where certain ICM cells only exhibit pluripotency (Gardner and Rossant, 1979). Our findings suggest that 2i/LIF promotes the expansion of a totipotent population of cells, reminiscent of these early developmental stages in vivo.

EXPERIMENTAL PROCEDURES

Cell Culture and Differentiation Assays

ESCs were cultured in serum/LIF (Canham et al., 2010), 2i (Ying et al., 2008), or 2i/LIF. $Stat3^{-/-}$ ESCs were maintained in 2i/LIF. H2B-Tomato cell lines were generated by introducing a H2B-Tomato vector under the control of a CAG promoter and upstream of IRES Puro cassette (a kind gift from H. Lickert) into the HV cell line.

LIF withdrawal and neural differentiation in monolayer culture were carried out as described previously (Fujikura et al., 2002; Ying et al., 2003b). For LIF withdrawal, 3×10^3 cells were plated per well of a six-well plate. Cells were cultured in standard serum-containing medium without LIF for 5 days. For neural differentiation, 10^4 ESCs were plated per well of a six-well plate in N2B27 medium for 9 days. For trophoblast stem cell differentiation assays, 10^4 cells were plated per well of a six-well plate. Cells were plated into trophoblast stem cell medium (70% mouse embryonic fibroblast-conditioned medium [R&D] and 30% TS cell medium [RPMI, GIBCO; glutamine and sodium pyruvate, GIBCO; 0.1 mM β-mercaptoethanol, Sigma; 20% fetal calf serum]). A total of 25 ng/ml Fgf4 (Peprotech) was added to medium along with 1 μg/ml heparin sulfate (Sigma). Cells were cultured for 7 days before analysis of differentiation levels by immunostaining.

Flow Cytometry

Cells were collected using Accutase dissociation buffer (A6964, Sigma) and antibody staining for PECAM-1 was carried out (Canham et al., 2010). Cells were sorted with a BD FACS Aria II cell sorter SORP or a BD FACS Aria III. DAPI-low, PECAM-1^+ cells were sorted from the upper and lower 10% or 25% of HV expression. In flow cytometry purity checks, these sorted populations showed a good separation and demonstrated gene expression and functional differences upon further analysis (Figure S1B). To assess population interconversion, single cells were sorted twice, taking the top and bottom 1% of HV-expressing cells to ensure clear separation. Flow cytometry and PDGFR-alpha antibody staining after LIF withdrawal was carried out as previously reported (Rugg-Gunn et al., 2012). Analysis was performed using FlowJo software (Tree Star). Gating methodology is shown in Figure S1B and an example purity check is shown in Figure S1C.

Immunostaining and Quantification

Immunostaining was carried out as previously reported (Canham et al., 2010) utilizing antibody concentrations described in Table S1. Endoderm levels were quantified by counting defined clusters (Figure 3F; example shown in Figure S3A) or absolute numbers of $GATA6^+$ cells (Figure S3B) and neural differentiation by counting absolute numbers of $TUJ1^+$ elongated neurons. Absolute numbers of $CDX2^+$ cells were counted. Two wells of a 12-well plate were analyzed per biological replicate and three biological replicates completed for each experiment. p values were calculated using one-way ANOVA tests. Images were acquired at 10× or 20× magnification. Colocalization studies with NANOG and KI67 were analyzed by confocal microscopy using an antibody against GFP to detect HV expression.

qRT-PCR

RNA was isolated from cells (RNeasy, 74104; QIAGEN). Complementary DNA (cDNA) was synthesized from 1 μg RNA using Superscript III according to the manufacturer's guidelines (18080, Invitrogen). Quantitative RT-PCR (qRT-PCR) was carried out on a Lightcycler480 (Roche) using either SYBR green or UPL technologies (Roche) utilizing primer sequences specified in Table S2. Primers were used at a concentration of 1 μM. The housekeeping gene TBP was used for normalization of Ct values detected for each sample.

RNA-Seq

Approximately 10 μg of total RNA underwent two rounds of mRNA enrichment with Dynalbeads $Oligo(dT)_{25}$ (61005, Life Technologies). Solid whole transcriptome libraries were made according to the Solid Total RNA-seq protocol with

the exception of RNA fragmentation (4445374, Life Technologies). RNA was fragmented by chemical hydrolysis; heating to 95°C, 10 min in 1× RNase III buffer (AM2290, Life Technologies), and snap cooled on ice. ATP (0.83 mM, 11140965001, Roche) and 10 U of T4 PNK (M0201L, NEB) were added and incubated at 37°C for 30 min. RNA was purified using Purelink RNA Micro Kit (12183-016, Life Technologies). Equimolar pools of RNA-seq libraries were made following quantitative PCR quantification using a Kapa Library Quantification kit (KK4823, Kapa Biosystems). Emulsion PCR and templated bead enrichment was carried out with Solid EZ bead system according to the manufacturer's guidelines. Enriched beads were sequenced on an ABi SOLiD 4 analyzer according to the manufacturer's instructions to generate 50 bp reads in color space.

Ethics Statement

All animal work was carried in accordance with UK and European legislation and in particular according to the regulations described in the Animals (Scientific Procedures) Act of 1986 (UK). All work in this manuscript was authorized by and carried out under Project License 60/3715 issued by the UK Home Office. Genetic modification for the generation of mouse and mouse ESC lines were approved by the ethics committees of the University of Edinburgh and the University of Copenhagen.

HV Transgenic Mouse Line Generation

The HV mouse line was generated by blastocyst injection of E14 Ju09 HV ESCs into wild-type F1 blastocysts. E14 Ju09 HV cells were generated by homologous recombination using the HV construct described previously (Canham et al., 2010) into E14 Ju09 ESCs. Mice were backcrossed onto the C57BL/6 background and maintained as heterozygotes. PCR genotyping was carried out on tail biopsies using the following primers: 5′-CGGAGGC GAATCTGAAGCCAGC-3′ (forward), 5′-GCATACAGCGGGACTCCCACG-3′ (reverse).

Early Embryo Imaging and Culture

Wild-type C57BL/6 or transgenic HV mouse lines were used for all experiments. Mice were checked for copulation plugs each morning and embryos were considered E0.5 on the day of plug detection. Embryos were flushed either from oviducts at E2.5 or from the uterus at E3.5 in PB1 medium. Embryos were cultured in control KSOM medium (Millipore) or KSOM medium containing 1 μM PD032 and 3 μM CHIR99021 (2i) (Nichols et al., 2009) and imaged by confocal microscopy (Leica, TCS SP2, or TCS SP8) at 20X magnification. Embryos were scored (Figure 1D) as showing homogeneously high or low HV expression relative to one another or as demonstrating heterogeneous HV expression when a mix of high and low HV-expressing cells was present.

Chimera Generation

Chimeric mouse generation was performed by morula aggregation or morula injection of single cells. HV cells constitutively expressing LacZ-IRES-Puro from a CAG promoter or an H2B-Tomato fusion protein were sorted by flow cytometry into HV^+ or HV^- populations. Clusters of eight to ten cells were aggregated with wild-type F1 morulae or single cells were injected. For single cell injections, morulae were incubated in PB1 medium without calcium and magnesium for 15 min at room temperature to facilitate decompaction for ease of injection. Resultant embryos were cultured for 3 days in vitro (equivalent to E4.5 stage in vivo) or transferred to pseudopregnant female mice and harvested at E6.5 or E9.5 and subjected to X-gal staining (Canham et al., 2010) or immunostaining and fluorescence imaging. X-gal-stained embryos were cryosectioned and fluorescent embryos were wax sectioned.

Whole-Mount Immunostaining

E6.5 embryos were dissected and fixed in 4% PFA for 1–2 hr at 4°C. Embryos were washed twice with PBS for 10 min followed by two 1 hr washes in PBST (PBS with 5% serum and 0.1% Triton X-100) at 4°C. Primary antibodies were diluted to 1:100 in PBST and incubated overnight at 4°C. Embryos were washed twice in PBST for 15 min, at room temperature, followed by five 1 hr washes. They were incubated in the secondary antibody for 2 hr at room temperature followed by two 15 min washes and a further five 1 hr washes.

Single-Cell qRT-PCR

Single cells were sorted into 96-well PCR plates containing 5 μl CellsDirect reaction mix, 0.2 μl SuperScriptIII/Platinum Taq mix (CellsDirect One-Step qRT-PCR kit, Invitrogen), 2.8 μl DNA suspension buffer (TEKnova), and 1 μl 500 nM primer mix containing a mix of 48 DELTAgene Assays (Fluidigm) (sequences in Table S3). Controls of 100 and 1,000 cells were included. RT reaction conditions were 50°C, 15 min; 95°C, 2 min; and 22× (95°C, 15 s; 60°C, 4 min). An exonuclease step was performed to remove unincorporated primers at 37°C, 30 min and 80°C, 15 min. Amplification products were then diluted 5-fold in TE buffer. Amplified cDNA was mixed with SsoFast EvaGreen SuperMix with Low ROX (Bio-Rad). The same DELTAgene assays were used in qRT-PCR. Samples and assays were mixed with appropriate loading reagents and loaded onto a 96.96 gene expression Dynamic Array (Fluidigm). Samples were loaded in technical replicates. Arrays were read using a BioMark HD genetic analysis system (Fluidigm). Downstream analysis was completed in Microsoft Excel. Cells that expressed no or low levels of ACTB and GAPDH housekeeping genes, or with a Ct over 30, were excluded from further analysis. DELTAgene assays were custom designed by Fluidigm to cross introns and avoid amplifying genomic DNA. Assays showing poor melting curves were also excluded from analysis. Data were analyzed without normalization or also normalized to the median expression of all genes across the array. No significant difference was observed in the data generated from either analysis method.

ACCESSION NUMBERS

The Gene Expression Omnibus accession number for the RNA-seq data reported in this paper is GSE45182.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, six figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2013.04.034.

LICENSING INFORMATION

ACKNOWLEDGMENTS

We thank Lynsey Robertson, Sally Inverarity, Javier Martin Gonzalez, Kasper Bonderup, Ron Wilkie, Simon Monard, Gelo de la Cruz, and Valerie Wilson for technical assistance; Austin Smith, Heiko Lickert, and Ian Chambers for reagents; and Sophie Astrof, Sally Lowell, and the entire Brickman lab for critical discussion of this manuscript. This work was supported by grants from the Medical Research Council (MRC) (G0701428) and the Novo Nordisk Foundation (to J.M.B.), the Intramural Research Program of the National Institutes of Health, and the National Institute on Aging (Z01AG AG000656 and Z01AG000662 to M.S.H.K.). S.M.M. is supported by an MRC studentship, and J.M.B. is an MRC senior nonclinical fellow. The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript. S.M.M. conceived of and executed experiments and wrote the paper with J.M.B.; M.A.C. generated the HV mice and ESCs; J.N. designed and executed the experiments and helped with embryo culture; A.A.S. analyzed RNA-seq data sets; R.P.M. executed experiments; M.S.H.K. supervised data analysis; and J.M.B. designed and supervised experiments and wrote the paper with S.M.M.

Received: March 13, 2013
Revised: April 26, 2013
Accepted: April 30, 2013
Published: June 6, 2013

REFERENCES

Beddington, R.S., and Robertson, E.J. (1989). An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. Development *105*, 733–737.

Canham, M.A., Sharov, A.A., Ko, M.S., and Brickman, J.M. (2010). Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcript. PLoS Biol. *8*, e1000379.

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

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

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

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

Gardner, R.L., and Rossant, J. (1979). Investigation of the fate of 4-5 day post-coitum mouse inner cell mass cells by blastocyst injection. J. Embryol. Exp. Morphol. *52*, 141–152.

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

Hall, J., Guo, G., Wray, J., Eyres, I., Nichols, J., Grotewold, L., Morfopoulou, S., Humphreys, P., Mansfield, W., Walker, R., et al. (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., and Terada, N. (2006). The Grb2/Mek pathway represses Nanog in murine embryonic stem cells. Mol. Cell. Biol. *26*, 7539–7549.

Handyside, A.H. (1978). Time of commitment of inside cells isolated from preimplantation mouse embryos. J. Embryol. Exp. Morphol. *45*, 37–53.

Hayashi, K., Lopes, S.M., Tang, F., and Surani, M.A. (2008). Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. Cell Stem Cell *3*, 391–401.

Hayashi, Y., Furue, M.K., Tanaka, S., Hirose, M., Wakisaka, N., Danno, H., Ohnuma, K., Oeda, S., Aihara, Y., Shiota, K., et al. (2010). BMP4 induction of trophoblast from mouse embryonic stem cells in defined culture conditions on laminin. In Vitro Cell. Dev. Biol. Anim. *46*, 416–430.

Hogan, B., and Tilly, R. (1978). In vitro development of inner cell masses isolated immunosurgically from mouse blastocysts. I. Inner cell masses from 3.5-day p.c. blastocysts incubated for 24 h before immunosurgery. J. Embryol. Exp. Morphol. *45*, 93–105.

Kidder, B.L., Yang, J., and Palmer, S. (2008). Stat3 and c-Myc genome-wide promoter occupancy in embryonic stem cells. PLoS ONE *3*, e3932.

Kobayashi, T., Mizuno, H., Imayoshi, I., Furusawa, C., Shirahige, K., and Kageyama, R. (2009). The cyclic gene Hes1 contributes to diverse differentiation responses of embryonic stem cells. Genes Dev. *23*, 1870–1875.

Kunath, T., Arnaud, D., Uy, G.D., Okamoto, I., Chureau, C., Yamanaka, Y., Heard, E., Gardner, R.L., Avner, P., and Rossant, J. (2005). Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. Development *132*, 1649–1661.

Lallemand, Y., and Brûlet, P. (1990). An in situ assessment of the routes and extents of colonisation of the mouse embryo by embryonic stem cells and their descendants. Development *110*, 1241–1248.

Liu, L., Luo, G.Z., Yang, W., Zhao, X., Zheng, Q., Lv, Z., Li, W., Wu, H.J., Wang, L., Wang, X.J., and Zhou, Q. (2010). Activation of the imprinted Dlk1-Dio3 region correlates with pluripotency levels of mouse stem cells. J. Biol. Chem. *285*, 19483–19490.

Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O., Trono, D., and Pfaff, S.L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature *487*, 57–63.

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

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

Poehlmann, T.G., Fitzgerald, J.S., Meissner, A., Wengenmayer, T., Schleussner, E., Friedrich, K., and Markert, U.R. (2005). Trophoblast invasion: tuning through LIF, signalling via Stat3. Placenta *26*(*Suppl A*), S37–S41.

Prakash, G.J., Suman, P., Morales Prieto, D.M., Markert, U.R., and Gupta, S.K. (2011). Leukaemia inhibitory factor mediated proliferation of HTR-8/SVneo trophoblast cells is dependent on activation of extracellular signal-regulated kinase 1/2. Reprod. Fertil. Dev. *23*, 714–724.

Quinn, J., Kunath, T., and Rossant, J. (2006). Mouse trophoblast stem cells. Methods Mol. Med. *121*, 125–148.

Rossant, J., and Lis, W.T. (1979). Potential of isolated mouse inner cell masses to form trophectoderm derivatives in vivo. Dev. Biol. *70*, 255–261.

Rossant, J., and Vijh, K.M. (1980). Ability of outside cells from preimplantation mouse embryos to form inner cell mass derivatives. Dev. Biol. *76*, 475–482.

Rugg-Gunn, P.J., Cox, B.J., Lanner, F., Sharma, P., Ignatchenko, V., McDonald, A.C., Garner, J., Gramolini, A.O., Rossant, J., and Kislinger, T. (2012). Cell-surface proteomics identifies lineage-specific markers of embryo-derived stem cells. Dev. Cell *22*, 887–901.

Shen, M.M., and Leder, P. (1992). Leukemia inhibitory factor is expressed by the preimplantation uterus and selectively blocks primitive ectoderm formation in vitro. Proc. Natl. Acad. Sci. USA *89*, 8240–8244.

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

Spindle, A.I. (1978). Trophoblast regeneration by inner cell masses isolated from cultured mouse embryos. J. Exp. Zool. *203*, 483–489.

Stewart, C.L., Kaspar, P., Brunet, L.J., Bhatt, H., Gadi, I., Köntgen, F., and Abbondanzo, S.J. (1992). Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. Nature *359*, 76–79.

Suemori, H., Kadodawa, Y., Goto, K., Araki, I., Kondoh, H., and Nakatsuji, N. (1990). A mouse embryonic stem cell line showing pluripotency of differentiation in early embryos and ubiquitous beta-galactosidase expression. Cell Differ. Dev. *29*, 181–186.

Takahashi, Y., Carpino, N., Cross, J.C., Torres, M., Parganas, E., and Ihle, J.N. (2003). SOCS3: an essential regulator of LIF receptor signaling in trophoblast giant cell differentiation. EMBO J. *22*, 372–384.

Takahashi, Y., Takahashi, M., Carpino, N., Jou, S.T., Chao, J.R., Tanaka, S., Shigeyoshi, Y., Parganas, E., and Ihle, J.N. (2008). Leukemia inhibitory factor regulates trophoblast giant cell differentiation via Janus kinase 1-signal transducer and activator of transcription 3-suppressor of cytokine signaling 3 pathway. Mol. Endocrinol. *22*, 1673–1681.

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

Wray, J., Kalkan, T., and Smith, A.G. (2010). The ground state of pluripotency. Biochem. Soc. Trans. *38*, 1027–1032.

Wray, J., Kalkan, T., Gomez-Lopez, S., Eckardt, D., Cook, A., Kemler, R., and Smith, A. (2011). Inhibition of glycogen synthase kinase-3 alleviates Tcf3

repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. Nat. Cell Biol. *13*, 838–845.

Xu, R.H., Chen, X., Li, D.S., Li, R., Addicks, G.C., Glennon, C., Zwaka, T.P., and Thomson, J.A. (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. Nat. Biotechnol. *20*, 1261–1264.

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

Ying, Q.L., Nichols, J., Chambers, I., and 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., and Smith, A. (2003b). Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. Nat. Biotechnol. *21*, 183–186.

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

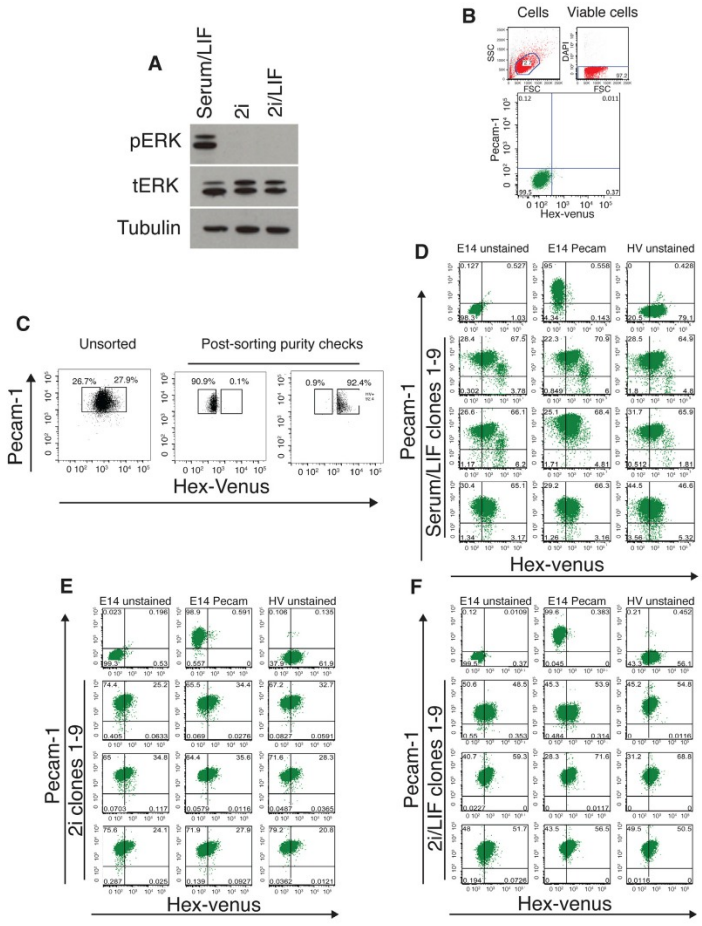


Figure S1. HV Heterogeneity Can Be Observed after Clonal Rederivation of ESC Lines in 2i, Related to Figure 2

(A) Phospho ERK western blots of cells cultured in serum/LIF, 2i or 2i/LIF. Total ERK and Tubulin are shown as loading controls.

(B) Plots showing gating for all flow cytometry in this study. Gating for HV and PECAM-1 expression was based on a control E14 cell line without the HV reporter and unstained for PECAM-1. Cells were separated from debris by gating on forward (FSC) and side scatter (SSC). Dead cells, showing high DAPI levels, were eliminated.

(C) Purity checks after sorting by FACS of upper and lower 25% of HV expressing ESCs in 2i.

(D–F) Flow cytometry of individual clones in serum/LIF (D), in 2i (E) and 2i/LIF (F).

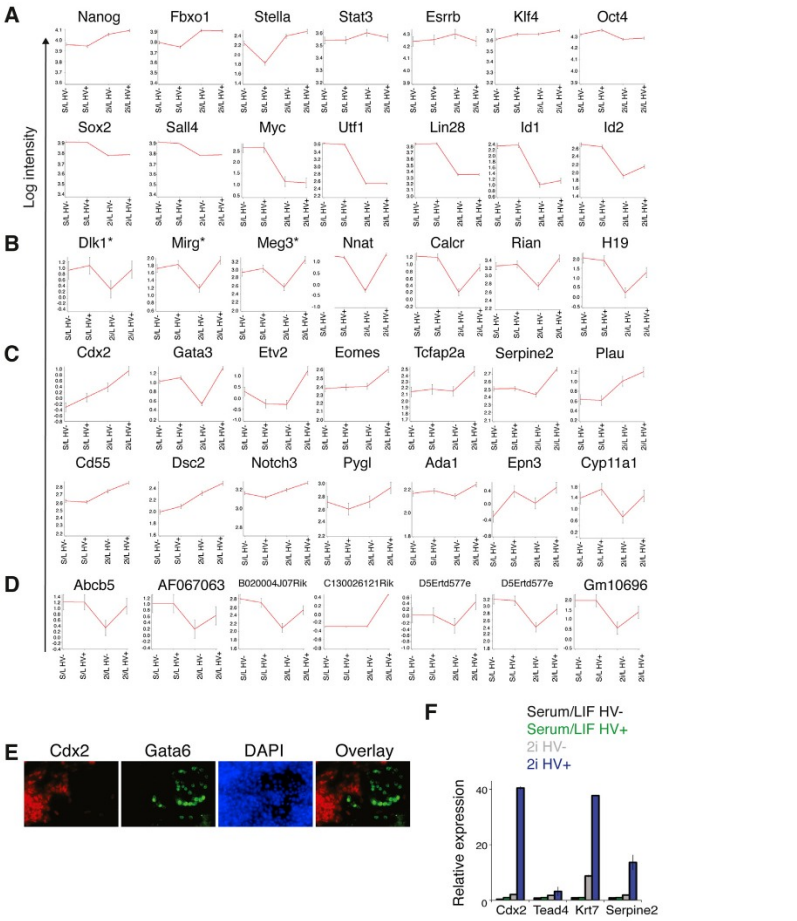


Figure S2. Transcriptomics Suggest the Presence of a Population Primed toward Trophoblast as well as Endoderm, Related to Figure 2

(A–D) Common RNA-seq signatures in specific gene classes. The behavior of the individual genes belonging to the classes described in the text is shown. Plots compare mean log intensity values for individual genes among four sorted populations; serum/LIF HV-, serum/LIF HV+, 2*i*/LIF HV-, 2*i*/LIF HV+. Error bars represent s.d. between 2 biological replicates. (A) Markers of undifferentiated ESCs, (B) imprinted genes (* = member of *Dlk1-Dio3* cluster), (C) trophoblast markers, (D) 2-cell stage embryo retroviral genes.

(E) High-magnification image of GATA6 and CDX2 immunostaining of ESCs, after 7 days differentiation in TSC conditions, demonstrating that there was no coexpression of these 2 markers.

(F) qRT-PCR of HV- and HV+ populations cultured in serum/LIF or 2i after LIF withdrawal for 4 days. Values were normalized to the housekeeping gene TBP and are shown relative to serum/LIF HV+. Error bars indicate mean ± SD.

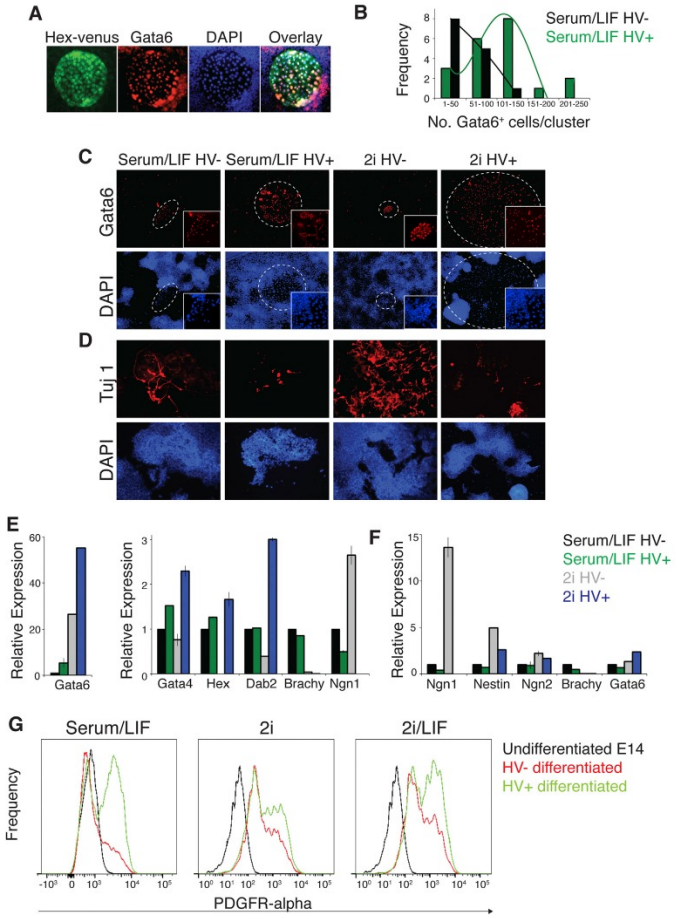


Figure S3. Quantification of Lineage Priming In Vitro, Related to Figure 2

(A) A typical GATA6+ endodermal colony, also expressing HV, as scored in differentiation assays.

(B) Quantification of number of GATA6+ cells per endodermal cluster.

(C and D) Immunostaining of ESCs differentiated by LIF withdrawal (C) or neural differentiation (D) after previous culture in serum/LIF or 2i and then sorting into HV- and HV+, PECAM-1+ ESCs. Cells were plated immediately into differentiation conditions after sorting. Endoderm levels were quantified by GATA6 immunostaining (n = 3) (values relative to serum/LIF HV+), (D) and neural differentiation by TUJ1 immunostaining (n = 5) (Values relative to serum/LIF HV-)

(E and F) qRT-PCR showing relative gene expression of sorted cells differentiated by LIF withdrawal (E) or neural differentiation (F). Data were normalized to the housekeeping gene TBP and is shown relative to serum/LIF HV sample. Values indicate mean ± SD of biological replicates.

(G) Flow cytometry of sorted HV- and HV+ populations following LIF withdrawal and staining for the endoderm marker PDGFR-alpha.

Cell PRESS
Open ACCESS

A

STAGE	CONDITION	EMBRYONIC	EMBRYONIC + EEMB
E4.5	2i/LIF HV+	4/7 (57%)	3/7 (43%)
E6.5	2i HV-	18/18 (100%)	0/18 (0%)
	2i HV+	16/26 (62%)	10/26 (38%)
	2i/LIF HV-	42/60 (70%)	18/60 (30%)
	2i/LIF HV+	25/55 (45%)	30/55 (55%)
E9.5	2i/LIF HV+	5/8 (63%)	3/8 (37%)

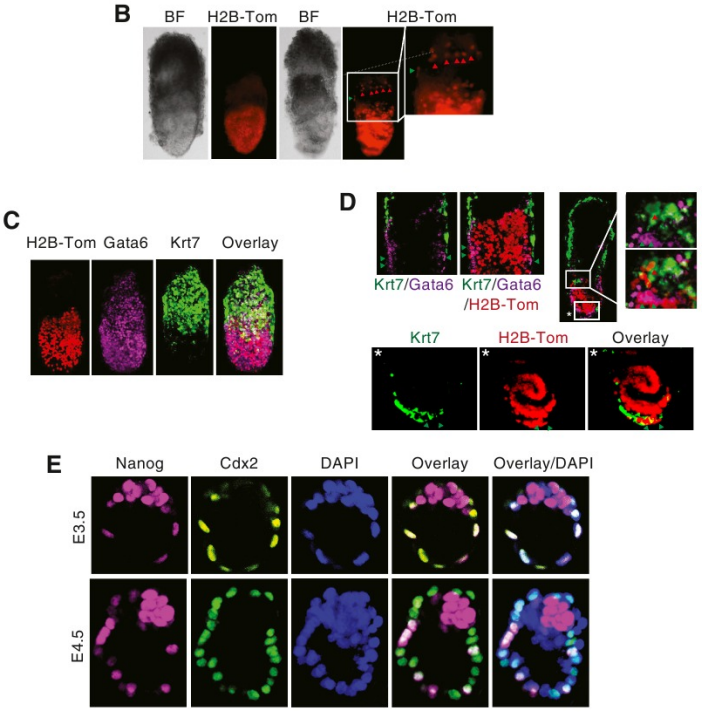


Figure S4. HV+ ESCs Demonstrate Extraembryonic Potential in Morula Aggregations, Related to Figure 3

(A) Table containing a summary of all chimeras produced by morula aggregation in this study.

(B–D) HV ESCs were cultured in 2i/LIF. HV+ cells, expressing a constitutive H2B-Tomato fusion protein, were sorted and used in morula aggregations. Embryos were dissected at E6.5 (B-D) and imaged by confocal microscopy. (B) Example chimeras, one showing only epiblast contribution, the other both epiblast and extraembryonic contribution. Red arrowheads indicate H2B-Tomato cells in the trophoblast, green arrowheads indicate H2B-Tomato cells in the extraembryonic endoderm. (C) Whole-mount immunostaining of chimeras for an endoderm (GATA6) and trophoblast marker (KRT7). (D) Optical sections through whole-mount immunostained embryos. Upper panels show both visceral endoderm and trophoblast contribution at the epiblast/extraembryonic border. Lower panels indicate trophoblast giant cell contribution outside of the distal epiblast. White box marked with * indicates area of high magnification in lower panel.

(E) Immunostaining of E3.5 and E4.5 wild-type blastocysts. Images show confocal optical sections.

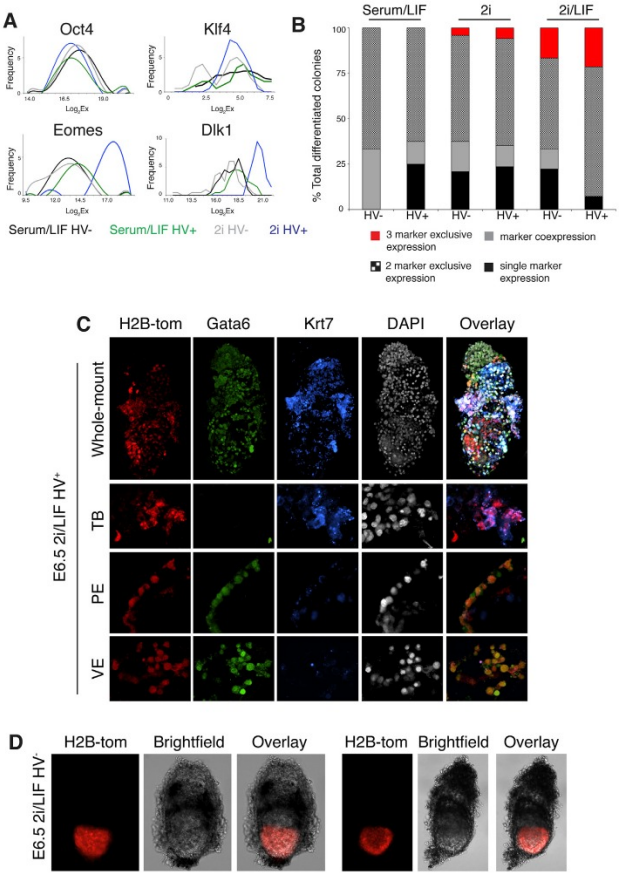


Figure S5. Single HV$^+$ Cells Are Totipotent in the Ground State, Related to Figure 4

Single cell qRT-PCR was carried out on a BioMark HD System (Fluidigm) using a 96 × 96 chip. DELTAgene Assay custom designed primers (Fluidigm) were used for amplification. Technical replicates were generated for each single cell.

(A) Histograms showing the distribution of expression of 4 genes in each population. Data shown unnormalized.

(B) Single cells were sorted by flow cytometry into LIF withdrawal conditions. After 7 days differentiated colonies were immunostained for GATA6, CDX2 and BRACHYURY. Quantification of marker expression patterns. 2 and 3 marker exclusive expression indicates colonies that contain independent cells positive for different lineage markers.

(C) Whole-mount immunostaining of an E6.5 chimeric embryo generated by injection of a single HV+ ES cell constitutively expressed H2B-Tomato, previously cultured in 2i/LIF, into a morula stage wild-type embryo. Upper whole-mount image shows an extended focus view of the entire embryo. Panels below show optical sections, generated by confocal microscopy, with ES cells contributing to trophoblast and visceral and parietal endoderm (VE and PE, respectively).

(D) E6.5 chimeras generated by injection of a single HV- cell that had been cultured in 2i/LIF. The majority of single cells contributed only to epiblast (11/13), with only 2 showing extraembryonic contribution.

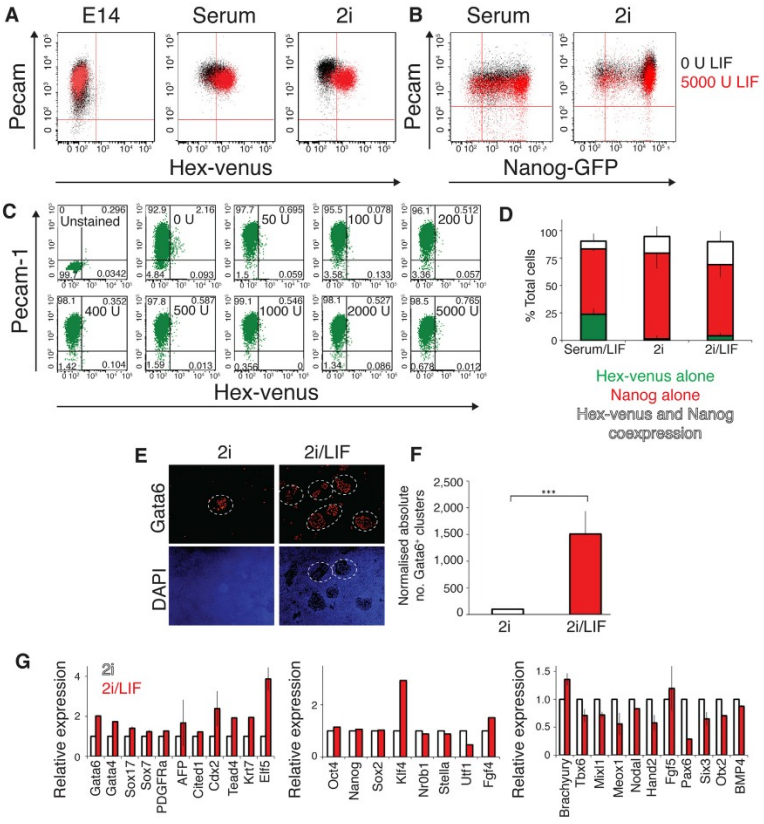


Figure S6. LIF Increases HV Expression within an Undifferentiated Population, Related to Figure 5

(A and B) Flow cytometry of HV cells (A) and Nanog-GFP TNG cells (B) in the presence and absence of LIF. Control E14 cells are also shown. Cells were cultured at high density to prevent differentiation.

(C) While LIF stimulates the HV reporter (Figure S1) it does not enhance levels of autofluorescence in wild-type E14 ESCs.

(D) Quantification of colocalization of HV and NANOG in serum/LIF, 2i and 2i/LIF. Data, based on confocal optical sections of 5 individual colonies of cells per condition, shown as mean ± SD.

(E) Differentiation by LIF withdrawal from unsorted cultures in 2i or 2i/LIF. Cells were immunostained for GATA6 (n = 3).

(F) Quantification of GATA6 immunostaining. Values relative to 2i. Error bars indicate mean ± SD of biological replicates (***$p < 0.001$). White dotted lines highlight GATA6+ endoderm clusters. See also Figure S3 and methods section.

(G) Summary of RNA-seq expression data for endoderm and trophoblast (left), canonical ESC (middle) and mesoderm and neurectoderm markers (right) for unsorted samples cultured in 2i or 2i/LIF. Error bars indicate mean ± SD of two biological replicates. Values are relative to 2i sample.

eLIFE
elife.elifesciences.org

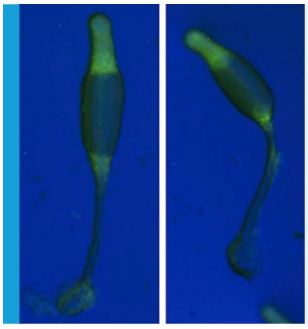
INSIGHT

CELL FATE CHOICE

Survival of the fattest

Experiments on the social amoeba *Dictyostelium discoideum* show that the origins of lineage bias in this system lie in the nutritional history of individual cells. Clues to the molecular basis for this process suggest similar forces may be at work in early mammalian development.

SOPHIE M MORGANI AND JOSHUA M BRICKMAN



Related research article Chattwood A, Nagayama K, Bolourani P, Harkin L, Kamjoo M, Weeks G, Thompson CRL. 2013. Developmental lineage priming in *Dictyostelium* by heterogeneous Ras activation. ***eLife*** **2**:e01067. doi: 10.7554/eLife.01067

Image Wild-type (left) and mutant Dictyostelium discoideum showing the spore (top) and stalk

One of the fundamental questions in developmental biology is how a pattern can be generated within a group of seemingly identical cells. In many animals, including non-mammalian vertebrates, pattern is determined by the position of specific factors inside the mother's egg. However, the initiation of patterns in many other species, including mammals and social amoeba, is not understood. When cells of *Dictyostelium discoideum*, a widely studied amoeba, are deprived of nutrients, they aggregate and differentiate to form a multicellular fruiting body comprised of a stalk and a number of spores. As in the early mammal, differences between the cells emerge seemingly at random. However, these initial differences are not fixed, and although the cells exhibit a bias towards a particular fate, a phenomenon known as 'lineage priming', they still retain the ability to become stalk or spore cells.

Lineage priming could arise within a population of cells as a result of stochastic bursts of gene expression (***Elowitz et al., 2002***) or through a deterministic element such as a molecular oscillator, the cell cycle or physical location. Feedback pathways, both negative and positive, could then amplify these early variations. However, there is another possibility that lies somewhere between these two alternatives. In *Dictyostelium*, the apparently 'random' fate of individual cells is influenced by their nutritional history. Cells grown without glucose are more likely to become stalk cells because they are more responsive to DIF, an extracellular signalling molecule that encourages the development of stalk cells, whereas cells grown with glucose tend to become spore cells (***Leach et al., 1973***; ***Thompson and Kay, 2000***). This bias is only observed when cells with different nutrient histories are mixed together.

Now in eLife, Christopher Thompson and colleagues at Manchester University and the University of British Columbia—including Alex Chattwood and Koki Nagayama as joint first authors—provide a mechanism linking the nutritional history of cells to lineage priming (***Chattwood et al., 2013***). They identified the protein GefE, a guanine nucleotide exchange factor, as being critical in regulating this process. When mutant cells that contain very little GefE were allowed to develop on their own, they still showed a normal bias based on their exposure to glucose. However, when these same mutant cells were grown without glucose and mixed with normal cells grown in glucose, they were no longer biased towards a stalk cell fate.

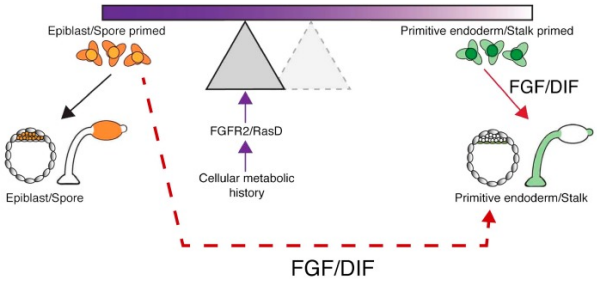


Figure 1. Potential mechanisms for introducing bias into lineage priming. When the amoeba *D. discoideum* is deprived of nutrients, it responds by forming a multicellular fruiting body made up of stalk and spore cells. The decision to become a stalk cell is induced by the signalling molecule DIF, and the nutrient history of each cell influences the likelihood that it will respond to DIF and adopt a particular fate. A similar situation arises in the early mouse embryo where cells can choose between becoming epiblast or primitive endoderm (PrE) cells, with the growth factor FGF promoting a PrE fate. The process can be visualized as a seesaw, with the nutrient history influencing the position of the fulcrum and, therefore, the likelihood that a cell will tipped towards a particular fate when exposed to DIF or FGF. In this way, the nutrient history of individual cells can establish precursor populations that are likely to become a stalk cell or a spore cell. Chattwood, Nagayama and co-workers now show that nutrient history affects the level of a protein called RasD in *Dictyostelium*, which in turn influences the ability of a cell to respond to DIF: high levels of RasD correspond to the fulcrum being closer to the left end of the seesaw (shaded triangle), so that the cells are primed to become stalk cells and exposure to DIF (solid red arrow) is more likely to lead to them becoming stalk cells. However, the primed state of these cells is not fixed because RasD expression can change based on nutrient history and the fulcrum can return to the right (dashed triangle), where cells are primed to become spore cells. When the fulcrum is positioned on the right, the cells are primed, but their fate is not determined because they can still become stalk cells if treated with high levels of DIF (dashed red arrow). In mice, a receptor called FGFR2 and a signalling molecule called FGF influence cell fate in a similar way to RasD and DIF in Dictyostelium. This suggests the intriguing possibility that nutrient history could influence the choice between epiblast and PrE lineages.

The explanation for this competitive loss of lineage bias was that GefE mutants were shown to be less likely to respond to the DIF signalling molecule. Chattwood, Nagayama and colleagues found that GefE activates a protein called RasD that in turn dictates responsiveness to DIF. They showed that RasD expression differed between cells, changed over time, and was increased by a lack of glucose. When cells with different levels of RasD were mixed, those with low levels of RasD mostly became spore cells, whereas cells with high RasD levels exhibited a preference for becoming stalk cells.

The role of DIF in *Dictyostelium*, promoting one cell fate over another, is comparable to the role of the growth factor FGF in mice. In early mouse development some cells will become primitive endoderm (PrE) cells—cells that support the embryo—and others become epiblast (i.e., they become part of the embryo). FGF encourages cells to become PrE (***Yamanaka et al., 2010***). However, before the cells segregate into these distinct lineages, precursor cells already exist that exhibit bias when challenged to differentiate in competition with other cells (***Canham et al., 2010***; ***Grabarek et al., 2012***). PrE precursors have been shown to express higher levels of the FGFR2, the receptor for factor FGF, which renders them more likely to respond to this growth factor.

In *Dictyostelium*, DIF activates the evolutionarily conserved JAK/STAT pathway, while in mice the homolog of RasD is a component of the Fgf pathway, positioned just downstream of the FGF receptor. Interestingly, both of these pathways are involved in the formation of extraembryonic lineages in early mouse development (***Stewart et al., 1992***; ***Yamanaka et al., 2010***) and embryonic stem cells (***Morgani et al., 2013***). Is it a coincidence that species as distantly related as *Dictyostelium* and mammals use the same pathways to modulate probabilistic cell fate choice? It is also noteworthy that the stalk cells in amoeba and PrE cells in mice both give rise to supportive structures that aid in the future development of the organism.

Perhaps the answer to this question lies in thinking about what advantages a historical lineage bias offers over either stochastic or traditional deterministic models. In the case of *Dictyostelium*, bequeathing the most nutritionally healthy cells to the next generation through the spore increases survival odds. A model incorporating historical bias into cell fate choice (***Figure 1***) would ensure that spores could be formed in all cases but, when fitter cells were present, they would be favoured for spore generation. Does the conservation of signalling pathways suggest a similar competitive mechanism is at work in the early mammalian embryo?

In essence, lineage priming is an experimental measurement of the ability of cells to compete for a particular lineage choice. While there is no

evidence of direct competition for a particular lineage, recent work suggests that epiblast cells can sense their defective neighbours and outcompete them (***Sancho et al., 2013***). In this case, cells sense a genetically defective neighbour but thus far there is no evidence that mammalian cells can judge the metabolic fitness of a genetically normal neighbour. However, the first clear example of cell competition, observed in *Drosophila*, was based upon the relative metabolic activity of cells (***Morata and Ripoll, 1975***).

Intriguingly, the first lineage decisions, in the earliest stages of mouse embryonic development, coincide with a switch to glucose metabolism (***Wales et al., 1995***). Perhaps probabilistic lineage bias ensures that the metabolically fittest cells adopt an embryonic fate. How then does RasD impact on the probability of cells to respond to DIF? This is an important, unanswered question that could shed light on the evolution of competitive mechanisms to register metabolic fitness through signalling pathways that might also be conserved in early mammalian development.

Sophie M Morgani is in the Danish Stem Cell Center (DanStem), University of Copenhagen, Copenhagen, Denmark
sophie.morgani@sund.ku.dk
Joshua M. Brickman is in the Danish Stem Cell Center (DanStem), University of Copenhagen, Copenhagen, Denmark
joshua.brickman@sund.ku.dk
Competing interests: The authors declare that no competing interests exist.
Published 26 November 2013

References

Canham MA, Sharov AA, Ko MS, Brickman JM. 2010. Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcript. *PLOS Biology* **8**:e1000379. doi: 10.1371/journal.pbio.1000379.
Chattwood A, Nagayama K, Bolourani P, Harkin L, Kamjoo M, Weeks G, Thompson CRL. 2013. Developmental lineage priming in *Dictyostelium* by heterogeneous Ras activation. *eLife* **2**:e01067. doi: 10.7554/eLife.01067.
Elowitz MB, Levine AJ, Siggia ED, Swain PS. 2002. Stochastic gene expression in a single cell. *Science* **297**:1183–1186. doi: 10.1126/science.1070919.
Grabarek JB, Zyzynska K, Saiz N, Piliszek A, Frankenberg S, Nichols J, Hadjantonakis AK, Plusa B. 2012. Differential plasticity of epiblast and primitive endoderm precursors within the ICM of the early mouse embryo. *Development* **139**:129–139. doi: 10.1242/dev.067702.
Leach CK, Ashworth JM, Garrod DR. 1973. Cell sorting out during the differentiation of mixtures of metabolically distinct populations of *Dictyostelium discoideum*. *Journal of Embryology and Experimental Morphology* **29**:647–661.
Morata G, Ripoll P. 1975. Minutes: mutants of drosophila autonomously affecting cell division rate. *Developmental Biology* **42**:211–221. doi: 10.1016/0012-1606(75)90330-9.
Morgani SM, Canham MA, Nichols J, Sharov AA, Migueles RP, Ko MS, Brickman JM. 2013. Totipotent embryonic stem cells arise in ground-state culture conditions. *Cell Reports* **3**:1945–1957. doi: 10.1016/j.celrep.2013.04.034.
Sancho M, Di-Gregorio A, George N, Pozzi S, Sanchez JM, Pernaute B, Rodríguez TA. 2013. Competitive interactions eliminate unfit embryonic stem cells at the onset of differentiation. *Developmental Cell* **26**:19–30. doi: 10.1016/j.devcel.2013.06.012.
Stewart CL, Kaspar P, Brunet LJ, Bhatt H, Gadi I, Köntgen F, Abbondanzo SJ. 1992. Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. *Nature* **359**:76–79. doi: 10.1038/359076a0.
Thompson CRL, Kay RR. 2000. Cell-fate choice in Dictyostelium: intrinsic biases modulate sensitivity to DIF signaling. *Developmental Biology* **227**:56–64. doi: 10.1006/dbio.2000.9877.
Wales RG, Martin KL, Leese HJ. 1995. Glucose utilization by components of the mouse conceptus during early embryogenesis. *Journal of Reproduction and Fertility* **104**:125–132. doi: 10.1530/jrf.0.1040125.
Yamanaka Y, Lanner F, Rossant J. 2010. FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. *Development* **137**:715–724. doi: 10.1242/dev.043471.