THE UNIVERSITY *of* EDINBURGH

Characterizing the function of Transcription factor 15 (Tcf15) in pluripotent cells

Chia-Yi Lin

Thesis presented for the degree of Doctor of Philosophy

Institute for Stem Cell Research

MRC Centre for Regenerative Medicine

The University of Edinburgh

2014

Declaration

I hereby declare that this thesis was composed by me, and the work presented in this thesis is my own, unless otherwise stated. This work has not been submitted for any other degree or professional qualification.

Chia-Yi Lin

2014

Acknowledgements

I would like to express my immense gratitude to my supervisor Sally for giving me the opportunity to carry out this interesting research project in her lab, for her guidance, never ending optimism, passion for science and patience for me. She is always available for open discussion, helps me greatly to find clues from chaotic mess of data, provides me environment for doing research independently while keeps me in the main road without loss of track. Her enormous support throughout my PhD studies and, especially, for my final year, made my PhD studies far more than I could have hoped for.

I would like acknowledge the members of my PhD committee: Dr. Tilo Kunath and Dr. Val Wilson, for their constant encouragement, practical advice and comments from related but also distinct fields, which benefits me and this project deeply.

I would like to thank all members in the Lowell lab. It's my pleasure to work with you all. Xhizhi helped me a lot during my first year of PhD study not only by introducing me every essential technique used in the lab but also by sharing life experience with me during our lunch time conversation. She also kept the lab tidy and everything organized for us, which I missed a lot. Paul was an interesting person who kept a lot of secrets including his true age, but I would like to thank him for his patience of sharing my sometimes not so interesting stories. Matt is a great colleague to work with, who is good at critical thinking, willing to help everyone in the lab and who doesn't hesitate to give useful opinions. He also opens a window for me so I can learn how to play touch rugby. Guilluame's arrival made our lab feel different, not only because we finally have a post-doc in the lab, but also because his in-depth knowledge about early embryogenesis and image quantification analysis. Although I couldn't really get to understand how to do programing, Guilluame do share his other broad interests with us, which include star-gazing, piloting and cooking, the latest made my lunchtime in the SCRM colourful. Amy is always willing to help with my embryo works Talking with her has greatly improved my English listening skill. Juraj was a beard Czech people who went to Taiwan more often than me. I would like to thank him for covering my overload cell culture works. Julia and Tulin's coming made me feel old in the lab. Julia is always energetic, and she makes

our bench-working time full of joys. Tulin is continuing the Tcf15-Nanog project and I am sure she has a great future ahead.

I am also very grateful to the past and present members in the institute, especially for members in the Developmental Biology group, for providing me with practical feedback and suggestion throughout my study. I should listen to Ian's advice to knockout Tcf15 directly without attempting to knockdown it using siRNA. Nick never hesitates to share his expertise in Western blot with me, as well as useful tips of how to keep the knee happy. Yali has given me lots of help for embryo works, and it's always a pleasure to see her with a smile on her face. Nicola is always a good listener and she always cheers me up when I feel upset. Florian's extensive instruction on microarray analysis deepens this project greatly. I would like to express my thanks to Suling for giving me a lot of encouragement (and offering free rides home) during my study.

This PhD would not have been possible at all without the invaluable service provided by the core facility staff in the ISCR, and subsequently the SCRM. Simon, Olivia, Claire and Fiona provide me great help for FACS analysis. Audrey, Euan and Robbie keep our daily research function efficiently. Lynsey, Sally and Jennie help me with the morula aggregation, without them I won't be able to get my lovely transgenic Tcf15 mice. I have been living/buried in the tissue culture room for quite a while during my PhD life in Edinburgh. Therefore, I am ever so grateful to Helen, Marilyn, Jonathan and Theresa for making sure that tissue culture run smoothly at all times. Special thanks should go for Helen and Marilyn who make great efforts to maintain the staff-shortage TC and always share me with interesting stories.

My biggest thanks go to my family in Taiwan. Their constant encouragement and support have been invaluable throughout this journey. I'm grateful to Po-Hsun for his love, support and immense understanding that kept me going and looking forward to the future. Thank you so much!

Finally, I would like to thank the Taiwan government scholarship that grant me the opportunity to cross the Eurasian Plate for studying my PhD at the University of Edinburgh, from a tropical island to this rainy, cold, windy but lovely place.

謝誌(Acknowledgements)

終於到了提筆致謝的時候，也代表博士生涯即將邁入完成的階段。回想起過去這段留學時光，愛丁堡從地圖上一個陌生的異鄉到現在堪稱熟門熟路的半個家，經歷了許多，也成長了許多，可說是既辛苦又充實。一路走來，心中充滿無盡的感謝。

首先，我要誠摯的感謝我的指導教授 Dr. Sally Lowell，Sally 的細心指導使我得以一窺胚胎發育與幹細胞這個熱門領域的深奧。同時，她耐心的指點，也讓我從碩班時期主要專攻的生化、分生領域，踏入一竅不通的幹細胞生物學的過渡期間不至於徬徨無措。在這四年中，她從不拒絕我敲門進去辦公室的即興討論，總是充滿熱情並且笑容滿面的與我討論研究中遇到的各種疑難雜症、總是可以在一團混亂的資料中抓出令人意外的重點、總是認真的鼓勵我度過低潮。同時，她提供我自由的研究環境但是又確保我不會偏離研究的主軸、提供我經費的支持讓我在教育部公費期滿的第四年仍然可以沒有後顧之憂地進行學習。在這篇論文寫作的過程中，她也花費了許多時間提供建議以及替我修改。我非常感謝能夠擁有一個這麼好的博士學位指導老師，沒有她就沒有今天這份研究的成果。

接著，我要感謝我的博士學習委員會成員：Dr. Tilo Kunath 以及 Dr. Val Wilson。他們在我求學過程中給我的寶貴意見跟來自不同領域對於研究方向的探討，讓我這篇論文得以更加豐富。

我非常感謝 Lowell lab 的所有成員，我們是一個來自世界各地充滿歡樂的小家庭。博一那年，Xinzhi (新志)帶著我熟悉環境、一起說著中文聊著愛丁堡生活的點點滴滴，讓我可以快速進入狀況展開研究生活。新志離開之後，實驗室就再也沒有人會替我們主動訂購藥品、把東西分門別類建檔，無奈挑起一半你的工作的我無比想念妳啊！Paul 是個神祕的學長，雖然說我也常搞不清楚他到底幾歲以及他到底拐了八百個彎之後想要表達的是什麼東西，但是在實驗室一起聊天的生活還是挺有趣的。Matt 可以說是真 • 強者我同學了！表達能力超強、邏輯超好、還是蘇格蘭橄欖球國家隊選手(我不會說就是因為他帶我去打 touch rugby 所以我摔斷十字韌帶的…)，他總是不藏私地跟我分享彼此的研究內容，並且總是給我很好的建議，這麼一位同學可以說是砥礪我的最佳目標啊！Guilluame 的來臨給實驗室注入了新技術以及讓我們終於正式成為擁有博士後的實驗室！除了不吝惜與我們分享他豐富的胚胎發育學知識以及我通常是有聽沒有懂的強大無比影像量化軟體，讓以往只能定性分析的免疫螢光染色成為可以定量的分析；他還分享了法國的家庭美食(雖然有時會是老婆作弊替他煮)讓我的午餐時光增色不少！Amy 是實驗室第一個英國學生，跟她聊天讓我

的英文聽力進步飛快啊～！Juraj 雖然來自捷克，但是父母卻是在台北工作，所以他居然是比我還常回台灣的人(羨慕無比)。可以跟外國同學在異鄉聊台灣的點點滴滴也是一件非常有趣的事呢！Julia 是個非常開朗的學妹，總是樂觀的面對一切，跟她相處起來非常的愉快，而且他完美代表了數學不太好的外國人讓我每次都好想出一堆濃度換算題給她回答！Tulin 雖然一半時間不在我們實驗室，但是承接了我的研究主題，希望她可以接續這個研究主題並持續發光發熱。

最感謝的是我的家人們，你們的支持與鼓勵伴我度過這四年的留學時光。感謝老爸，當年要不是你提議要我去考公費留學，現在的這一切也不至於會實現。老爸雖然平常很嚴肅，但我知道你無比關心我的學業以及生活，一直為我禱告，謝謝你～！感謝老媽，這四年來風雨無阻的每天下午守在電話面前等著女兒的越洋電話，雖然總是聊一些雞毛蒜皮的生活瑣事，但是老媽是世界上最棒的聆聽者，給我無條件的支持與陪伴，我愛妳～！明年一起來參加畢業典禮，希望天氣還可以跟上次來的時候一樣棒！感謝幸怡，不辭辛勞前往愛丁堡來探親，愛丁堡真的沒有很冷啦～(逃)，感謝妳提供魯泥的陪伴，讓我去年的冬天充滿毛茸茸的愛。還有，安道，你知道全家只剩誰沒來過愛丁堡了嗎？

同時，我要感謝有情有義的柏勳，從當年機場出發時就來送機、到英國的第一個生日給我寄來了台灣的朋友們聯合的祝福生日卡、之後的日子裡給我寄了好多思鄉零食(跟無數無厘頭的禮物)、寫了好多信。很感謝你這些日子以來的陪伴，更感謝你讓生活陷入低潮的我再度找回信心，得以抬起胸膛繼續努力以及用熱情的態度面對生活。遠距戀愛很辛苦，感謝你的包容與體諒，支持我度過最後這段寫論文寫到腦汁榨乾、胡言亂語、不知所云的時光，我想一切都是值得的！

謹將此論文獻給所有關心我的人，以表達我無限的感激。

林佳怡 謹誌

2014

Abstract

Pluripotent embryonic stem (ES) cells are heterogeneous mixtures of naïve and lineage-primed states defined by distinct transcription factor expression profiles. However, the events that prime pluripotent cells for differentiation are not well understood. Id proteins, which are inhibitors of basic helix-loop-helix (bHLH) transcription factors, contribute to pluripotency by blocking differentiation. Using Yeast-Two-Hybrid screening, our lab identified Tcf15 as an Id-regulated transcription factor.

In this study, I first examined the expression of Tcf15 during differentiation *in vitro* and during early development *in vivo* in the mouse. Tcf15 expression is higher in primed pluripotent embryonic stem (ES) cells than in naïve ES cells or epiblast stem cells (EpiSCs). In addition, Tcf15 is expressed heterogeneously in ES cells and is also detected in the inner cell mass (ICM) of E4.5 mouse embryos. Expression of Tcf15 was upregulated during early stages of differentiation and downregulated before cells committed to any specific lineage. Using Tcf15-Venus reporter cells, I found that expression of Tcf15 is specifically associated with a novel subpopulation of ES cells primed for somatic lineages.

Gain of function and loss of function studies were then performed to perturb Tcf15 expression in ES cells in order to assess the function of Tcf15 in self-renewal and during differentiation. An inducible Id-resistant form of Tcf15 accelerates somatic lineage commitment by maturating naïve pluripotent ES cells transit toward primed epiblast and later on epiblast-derived somatic lineages whilst suppressing differentiation towards

extraembryonic endoderm. Preliminary loss of function studies also suggest that down-regulation of Tcf15 may promote a naïve state within pluripotent cells.

I investigated the mechanism by which Tcf15 expression becomes associated with the epiblast-primed state by identifying the upstream regulators and downstream targets of Tcf15. Tcf15 expression is dependent on FGF signalling. Microarray analysis identified that Tcf15 downregulates the naïve pluripotency determinant *Nanog* and upregulates the epiblast determinant *Otx2*. Taken together, our results suggest that Tcf15 acts in opposition to the pluripotency network to prime pluripotent cells towards differentiation.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

bHLH	Basic helix-loop-helix
BMP	Bone morphogenetic protein
bp	Base pairs
cDNA	Complimentary DNA
ChIP	Chromatin immunoprecipitation
ChIP-Seq	Chromatin immunoprecipitation sequencing
cRNA	Complimentary RNA
DAPI	4',6-diamidino-2-phenylindole
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
EB	Embryoid body
EC	Embryonic carcinoma
ECM	Extra-cellular matrix
EMT	Epithelial-mesenchymal-transition
EpiSCs	Epiblast stem cells
FACS	Fluorescence-activated cell sorting
FCS	Fetal calf serum
FGF	Fibroblast growth factor
GFP	Green fluorescent protein
hES	Human embryonic stem
ICC	Immuno-cyto chemistry
ICM	Inner cell mass
iPSCs	Induced pluripotent stem cells
IRES	Internal ribosome entry site

ISH	*In situ* hybridisation
kb	Kilobase pairs
kDa	Kilodalton
KSR	Knockout serum replacement
LB	Lysogeny broth
LIF	Leukaemia inhibitory factor
mES	Mouse embryonic stem
MET	Mesenchymal to epithelial transition
mRNA	Messenger RNA
NLS	Nuclear localisation signal
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
poly(A)	Polyadenylation signal sequence
PrE	Primitive endoderm
RNA	Ribonucleic acid
RT-qPCR	Real-time quantitative PCR
SDS	Sodium dodecyl sulphate
TE	Trophectoderm
TSS	Transcription start site
UPL	Universal probe pibrary
UTR	Untranslated region
Y2H	Yeast-two-hybrid

CHAPTER 1

Introduction

1.1: Embryonic stem cells

Studying the biological properties and clinical potential of embryonic stem (ES) cells has drawn continued scientific, commercial, and public attention in the hope that it will shed new light on the development of regenerative medicine. Mouse embryonic stem (mES) cells are immortal cell lines derived from the inner cell mass (ICM) of pre-implantation blastocysts and represent a clonal pluripotent cell line with two defining properties: the ability to generate identical copy of themselves ("self-renewal") and the capacity to differentiate into all cell types of the developing embryo and adult ("pluripotency") (Evans and Kaufman, 1981; Martin, 1981). A functional proof of pluripotency of ES cells is their ability to form multi-differentiated teratocarcinomas consisting of derivatives from all three germ layers when grafted to adult mice (Stevens, 1970). Their full potency is revealed by blastocyst injection, which yields chimaeric mice with extensive contribution from the injected ES cells progeny to all tissues, including functional colonization of the germline (Bradley et al., 1984). Combined with genetic manipulations techniques such as gene targeting, ES cells serve as an excellent tool for studying gene function, enabling in-depth investigation of the genetic basis of disease and providing a platform for drug screening. For these reasons, the 2007 Nobel Prize in Physiology or Medicine was awarded jointly to Mario R. Capecchi, Sir Martin J. Evans and Oliver Smithies for the discovery of "Principles for introducing specific gene modifications in mice by the use of embryonic stem cells".

In vivo, pluripotent cells only exist transiently in the epiblast of the developing embryo. Therefore, ES cells could represent an *in vitro* model for studying early epiblast development.

1.1.1: Pluripotent populations

The field of pluripotent stem cells biology is a series of historical discoveries which led to the present achievement, like standing on the shoulders of giants. The discovery of embryonic stem cells depended on prior decades of research on mouse embryonal carcinoma (EC) cells. More recently, the derivation of mouse epiblast stem cells (EpiSCs) and induced pluripotent stem cells (iPSCs) relied on prior studies on mouse and human embryonic stem cells.

1.1.1.1: Embryonal carcinoma (EC) cells

The field of study on pluripotent stem cells began in the 1950s with the study of teratocarcinomas and the pluripotent cells that they harbour, the embryonal carcinoma (EC) cells. Teratocarcinomas are malignant germ cell tumours that comprise a differentiated component that can include all three germ layers (ectoderm, mesoderm and endoderm) and an undifferentiated EC component which was able to seed a new tumour when transplanted into a secondary recipient (Kleinsmith and Pierce, 1964; Stevens and Little, 1954). In their key experiment, Kleinsmith and Pierce (1964) demonstrated that a single EC cell is capable of both unlimited self-renewal and multi-lineage differentiation, providing the framework for ES cell biology: the

clonogenicity of single cells, their heterogeneity, and their pluripotency (potential to differentiate into cells in different germ layers).

EC cell lines that could be stably propagated *in vitro* as adherent monolayers in basal medium supplemented with serum, in the presence or absence of a layer of feeder cells were established in the early 1970s (Kahan and Ephrussi, 1970; Rosenthal et al., 1970). Although EC cells form malignant teratocarcinomas when transplanted to ectopic sites, when re-introduced into a blastocyst, they are able to incorporate and contribute to tissues of the developing embryo (Brinster, 1974). This indicates that the proliferation of undifferentiated EC cells can be brought under control in response to the appropriate environmental cues. These mice which harbour tissues derived from different genetic backgrounds (in this case, the EC cells and the original blastocysts) are known as chimaeras. The "ability to generate chimaera upon blastocyst injection" thus becomes a key criteria for evaluating the pluripotency of ES cells, and provides a crucial link between *in vitro* cell lines and *in vivo* development.

The ground-breaking work performed by studying EC cells was pivotal for the subsequent isolation of pluripotent cells from the embryo proper. A critical point was the finding that pluripotent cells heavily benefitted by their co-culture with fibroblasts (Martin and Evans, 1975). The same authors later reported the isolation of pluripotent stem cells lines from mouse blastocyst (Evans and Kaufman, 1981; Martin, 1981). To denote their tissue of origin, these cells were named embryonic stem (ES) cells.

1.1.1.2: Mouse embryonic stem (ES) cells

Mouse ES cells are pluripotent populations isolated from the ICM of mouse blastocysts at embryonic days 3.5-4.5 (E3.5-4.5). ES cells represent the naïve pluripotent epiblast in the pre-implantation embryo. More importantly, unlike EC cells, these karyotypically normal ES cells can contribute at a high frequency to a variety of tissues in chimaeras, including germ cells, thus providing a practical way to introduce modifications to the mouse germline (Robertson et al., 1986) .

The derivation and culture of ES cells has been achieved using combinations of cytokines, growth factors, hormones, serum, serum extracts, conditioned media and feeders. In short, extrinsic factors including leukaemia inhibitory factor (LIF) and bone morphogenetic protein (BMP) can support the propagation of ES cells due to their specific ability to stimulate or protect the gene regulatory network controlling pluripotency. Considerable evidence indicates that ES cell self-renewal and pluripotency is governed by a transcriptional network centred on Oct4, Sox2 and Nanog (Boyer et al., 2005; Chen et al., 2008b; Loh et al., 2006). These core nuclear regulators target genes encoding transcription factors, signal transduction components and chromatin-modifying enzymes that promote self-renewal, while suppressing differentiation (Chambers et al., 2007; Chambers and Tomlinson, 2009; Niwa et al., 2000). The detailed molecular basis for maintaining self-renewing ES cells will be discussed in section **1.1.2**.

The ES cell genome can be manipulated with precision and relative ease (Thomas and Capecchi, 1986). Moreover, ES cells genetically modified with random or targeted transgene integrations could successfully contribute to live animals and their offspring following blastocyst injection (Hooper et al.,

1987; Robertson et al., 1986), thereby enabling the generation of mice with defined genetic modifications. Thus, the *in vitro* transgenesis in ES cells served as useful tool for the *in vivo* study of the phenotype of transgenic mice.

1.1.1.3: Post-implantation mouse epiblast stem cells (EpiSCs)

Epiblast stem cells (EpiSCs) represent the *in vitro* equivalent of the post-implantation pluripotent epiblast. EpiSCs were isolated from E5.5 to E6.5 post-implantation mouse embryos that differ significantly from mouse ES cells but share key features with human ES cells with regard to their molecular properties, growth factor requirements, colony morphology, X-chromosome inactivation status, and culture dynamics (Brons et al., 2007; Tesar et al., 2007). For example, EpiSC derivation failed in the presence of LIF and/or BMP4, the two factors required for the derivation and self-renewal of mouse ES cells (see section **1.1.2.2**). In contrast, similar to human ES cells, fibroblast growth factor (FGF) and Activin/Nodal signalling appear critical for EpiSC derivation. Despite distinct signalling requirements for self-renewal, EpiSCs are capable of differentiating into derivatives of all three primary germ layers when subjecting to embryoid body and teratoma formation assays (Brons et al., 2007; Tesar et al., 2007). Moreover, although EpiSCs are unable to contribute to chimaeras upon blastocyst injection or morula aggregation, it is recently reported that EpiSCs efficiently form chimaeras in gastrulation-stage embryos (Huang et al., 2012). Thus, EpiSCs do share the two key features characteristic of ES cells: prolonged proliferation *in vitro* (self-renewal) and multi-lineage differentiation (pluripotency).

Morphologically, mES cells grow in more three-dimensional, tightly packed colonies whereas mEpiSCs form flattened two-dimensional colonies. Also, like most human ES cell lines, mEpiSCs cannot be passaged as single cells by trypsin digest (Chenoweth et al., 2010), unless adding Rho-associated kinase (ROCK) inhibitor into culture medium to reduce dissociation-induced apoptosis (Watanabe et al., 2007). EpiSCs express the core set of genes that regulate the pluripotent state including *Oct4 (Pou5f1)*, *Nanog*, and *Sox2*. Additionally, EpiSCs express genes known to be involved in early gastrulation and somatic cell specification including *Brachyury* (*T*) and *Fibroblast growth factor 5* (*Fgf5*), suggesting that EpiSCs are poised for differentiation upon receiving instructive signals. Hierarchical clustering of global gene expression data from mES cells and mEpiSCs strongly suggests that these two cell types are in two distinct pluripotent states (Bao et al., 2009; Brons et al., 2007; Tesar et al., 2007). ES cells have been described as the naïve pluripotent state due to their functional similarity to the pre-implantation epiblast, whereas EpiSCs represent a distinct state of pluripotency from ES cells. Their pluripotent state has been described as ''primed'' based on their transcriptional and epigenetic profiles that are most similar to their source, the post-implantation epiblast (Nichols and Smith, 2009) (see section **1.1.5.2**).

1.1.1.4: Induced pluripotent stem cells (iPSCs)

The ground-breaking discovery by Kazutoshi Takahashi and Shinya Yamanaka at 2006 showing that terminally differentiated somatic cells could be reprogrammed to a pluripotent state by means of ectopic overexpression of four transcription factors: *Klf4*, *Myc*, *Oct4* and *Sox2* (Takahashi and Yamanaka, 2006). The type of cell generated by this method was termed

induced pluripotent stem cells (iPSCs). iPSCs are morphologically indistinguishable from ES cells, display a similar gene expression pattern, can contribute to chimaeras upon blastocyst injection and can be transmitted to the germline of the chimaeric embryos (Okita et al., 2007; Takahashi and Yamanaka, 2006; Wernig et al., 2007). The derivation of iPSCs from human cells (Park et al., 2008; Takahashi et al., 2007) provides an incredible tool not only for the study of patient-specific disease phenotypes in culture but also for the generation of patient-specific cell types for transplantation.

iPSCs can be established with a variety of methods. There are several differences among the original cell source, induction method, reprogramming factors usage, and culture conditions including small molecule supplements (Okita and Yamanaka, 2011). Although using iPSCs for research could avoid ethical problems associated with embryonic stem cell generation, iPSCs do present problems of their own. The reprograming process occurs at low frequency (0.1% to 1%) and over a long period of time (a few weeks). Also, the possibility of transgene integration and alteration of the endogenous genomic organization could cause a negative safety issue when considering medical applications. More comprehensive knowledge of the reprogramming process is therefore crucial for future clinical applications of iPSCs (reviewed in Bayart and Cohen-Haguenauer (2013)).

1.1.2: Maintaining pluripotency of ES cells

The popularity of mES cell research has led to a wealth of large and small scale studies on the regulatory mechanisms that act to maintain the pluripotent state *in vitro*. ES cell pluripotency is maintained during self-renewal by the prevention of differentiation and the promotion of

proliferation and it is now clear that a complex interplay between intrinsic and extrinsic factors exist.

1.1.2.1: Intrinsic determinants governing ES cells self-renewal

The transcription factors Oct4, Sox2 and Nanog are among the pluripotency-associated factors that maintain ES cells (Boiani and Scholer, 2005; Chambers et al., 2007; Masui et al., 2007; Niwa et al., 2000). Their targets have been mapped by chromatin immuno-precipitation (ChIP)-based technologies, revealing their extensive co-binding in both mES cells and hES cells. This has led to the proposal that these factors constitute a core transcriptional regulatory network. This core transcriptional network is a key regulator of its own expression as well as regulating various other transcription factors and epigenetic regulators that are involved in self-renewal and pluripotency (Ivanova et al., 2006; Loh et al., 2006; Niwa et al., 2009).

1.1.2.1.1: Oct4

Oct4, which is encoded by *Pou5f1*, is a POU domain-containing transcription factor that is expressed in pluripotent mES cells and the early embryo. It was essential for the maintenance of both *in vivo* and *in vitro* pluripotent populations (Nichols et al., 1998; Yeom et al., 1996). Although *Pou5f1*-null embryos reach blastocyst stage, the inner cell mass of these mutants only produces differentiated cells of trophoblast lineages when cultured *in vitro*. Oct4 has been reported to directly prevent differentiation towards trophectoderm by interacting with Cdx2, a trigger for trophectoderm differentiation, to form a reciprocal repression complex for lineage specification (Niwa et al., 2005).

The expression level of Oct4 is tightly regulated in the developing embryo and mES cells, for the maintenance of pluripotency and specification of fate. A two fold increase in the levels of Oct4 causes differentiation into primitive endoderm, while repression induces dedifferentiation into trophectoderm. This indicates that there are different sets of genes associated with differentiated states/specific lineages expressed at either high or low levels of Oct4 and that pluripotency genes are only expressed when Oct4 is maintained at levels endogenous to self-renewing mES cells (Niwa et al., 2000). A recent study suggested that reduced Oct4 expression resulting from heterozygosity directs a robust pluripotent state with distinct signaling activity and increased enhancer occupancy by Oct4 and Nanog. These *Oct4*-heterozygous ES cells also display delayed differentiation kinetics compared to wild-type ES cells. Therefore, a defined Oct4 level controls the establishment of naïve pluripotency as well as commitment to all embryonic lineages (Karwacki-Neisius et al., 2013). Similar observations were also made in induced pluripotent stem cells (Radzisheuskaya et al., 2013).

In addition to being an essential regulator of pluripotency, Oct4 plays an important role in somatic reprogramming. Oct4 overexpression is sufficient to induce pluripotency when in combination with small molecules (Li et al., 2011; Yuan et al., 2011), or when using somatic cell types endogenously expressing the remaining reprogramming factors (Kim et al., 2009; Tsai et al., 2011; Wu et al., 2011). It was recently demonstrated that reactivation of endogenous *Oct4* expression plays a critical role in somatic cell reprograming, as several factors (Nr5a1/Nr5Aa2, TET1/Nanog) which were shown to induce pluripotency in the absence of Oct4 actually function through regulating *Oct4* regulatory region (Costa et al., 2013; Gao et al., 2013b; Gu et al., 2005; Heng et al., 2010). Moreover, overexpression of an artificial transcription

factor targeting the *Oct4* distal enhancer leads to strong activation of the endogenous *Oct4* locus and efficient reprograming of iPSCs in the absence of exogenous Oct4 (Gao et al., 2013a).

1.1.2.1.2: Sox2

Sox2 belongs to the family of SRY (Sex-related HMG box) transcription factors that is essential for early mouse development (Avilion et al., 2003). Sox2 interacts with Oct4 at both protein and DNA level (Ambrosetti et al., 1997; Ambrosetti et al., 2000) cooperatively to drive the transcription of target genes. These targets include *Fgf4* (Yuan et al., 1995), the pluripotency factors *Nanog* and *Utf1* (Kuroda et al., 2005; Nishimoto et al., 1999), as well as *Oct4* and *Sox2* themselves (Okumura-Nakanishi et al., 2005; Tomioka et al., 2002). Sox2 is also required for ES cell pluripotency, as its loss leads to the downregulation of *Oct4* and the upregulation of markers of the trophectodermal lineage (Ivanova et al., 2006; Masui et al., 2007). Surprisingly, the expression of many Oct4/Sox2 target genes was not greatly affected by the loss of Sox2. Furthermore, restoring Oct4 levels in *Sox2*-depleted cells by introduction of a Oct4 transgene can rescue, at least in part, the expression of Oct4/Sox2 targets, the differentiation phenotype of the cells and capability of the cells to contribute to chimaeric embryos (Masui et al., 2007). Therefore, Sox2 appears to be to maintain *Oct4* expression and other Sox factors such as Sox4, Sox11 and Sox15 can, to a certain extent, substitute for Sox2 in this context with their ability to bind to the Oct4/Sox2 target sites (Maruyama et al., 2005; Masui et al., 2007).

1.1.2.1.3: Nanog

Nanog is a homeobox-containing transcription factor identified by several groups in 2003 using different approaches: in a screen for molecules that can support self-renewal and pluripotency (Chambers et al., 2003), by degenerate PCR for homeobox-containing genes involved in early embryonic development (Wang et al., 2003), and in a bioinformatics screen for mES cell associated transcripts (Mitsui et al., 2003). Overexpression of Nanog can confer cytokine-independent self-renewal to ES cells, and for this reason the gene was named after the mythological Celtic Land of Youth, Tír na nÓg (Chambers et al., 2003).

During mouse development, Nanog expression is first detected in morulae, being most obvious in cells in the interior of the embryo. High levels of Nanog (mRNA/protein) persist in the early blastocyst, but declines prior to implantation. Following implantation, Nanog is expressed in a subset of epiblast cells with highest levels in the proximal posterior region; upon entry into the primitive streak, Nanog is rapidly downregulated (Chambers and Smith, 2004; Malaguti et al., 2013). It is reported to be required for the specification of the pluripotent epiblast *in vivo* (Frankenberg et al., 2011; Mitsui et al., 2003; Silva et al., 2009). *In vitro*, Nanog displays a stronger pluripotency-inducing phenotype than Oct4 and Sox2, and does not induce differentiation when overexpressed, but cannot rescue trophectodermal specification in *Oct4*-null or *Sox2*-depleted cells (Chambers et al., 2003; Ivanova et al., 2006). However, Nanog is dispensable for ES cell self-renewal. *Nanog*-null ES cells can be maintained in culture, retain multi-lineage differentiation capacity and can contribute to chimaeras, although they display lower clonogenicity than wild-type cells and express markers of embryonic and extraembryonic endoderm (Chambers et al., 2007; Mitsui et

al., 2003). These findings indicate that Nanog is an extremely potent pluripotency determinant as transient downregulation of Nanog appears to predispose cells towards differentiation but does not mark commitment, but that it is neither sufficient nor required for ES cell self-renewal and differentiation.

Nanog protein can be divided loosely into 3 domains: an N-terminal domain; a homeobox-containing domain; and a C-terminal domain that contains the well-conserved tryptophan (Trp) repeat domain (Mullin et al., 2008). The Trp-repeat domain is essential for Nanog homodimerisation to bind DNA (Mullin et al., 2008; Wang et al., 2008) and for direct interaction of Nanog with pluripotency factor Sox2 (Gagliardi et al., 2013). Nanog, Oct4 and Sox2 share a significant number of targets, including *Nanog, Oct4* and *Sox2* themselves (Chen et al., 2008a; Costa et al., 2013; Kim et al., 2008; Liang et al., 2008; Wang et al., 2006). Together, Nanog, Oct4 and Sox2 form the core network of transcription factors supporting embryonic stem cell self-renewal.

Nanog expression is subject to a similar level of complex regulation as seen at the *Oct4* locus, with both functional Oct4/Sox2 and germ cell nuclear factor (GCNF) motifs as well as binding sites for STAT3, Tcf3, FoxD3, p53 and T-Brachyury (Pan and Thomson, 2007). Two recent publications have demonstrated an unexpected role for Nanog in the auto-repression of its own transcription. While Oct4 and Sox2 promote the transcription of *Nanog*, Nanog binds its enhancer region independently of Oct4 and Sox2 and drives the downregulation of its expression. This auto-repression exists in the LIF/serum culture as well as in the 2i/LIF medium (Fidalgo et al., 2012; Navarro et al., 2012) (see section **1.1.2.2.5** for 2i culture). In addition, several factors known to regulate the *Nanog* locus, such as Oct4, FoxD3 and p53, can

act as either activators or repressors depending on their co-factor context indicating dynamic control of Nanog expression in mES cells.

1.1.2.1.4: Esrrb

The estrogen related receptor-β (Esrrb) is an orphan nuclear receptor that has been reported to be an important component of the complex pluripotency network (van den Berg et al., 2008; Zhang et al., 2008). Esrrb is involved in ES cell self-renewal and its expression is reduced upon Nanog and Oct4 knockdown and induced by Nanog overexpression (Ivanova et al., 2006; Loh et al., 2006). When ectopically expressed in conjunction with Oct4 and Sox2, Esrrb has been shown to promote reprogramming of mouse embryonic fibroblasts (Feng et al., 2009). Esrrb can interact with Nanog to promote *Oct4* expression and interact with Oct4 to promote *Nanog* expression, and it can also interact with Sox2 (Hutchins et al., 2013; van den Berg et al., 2008; Zhang et al., 2008). Most recently, Esrrb has been reported as the transcript which is upregulated quickest and to the highest levels following tamoxifen-mediated nuclear translocation of Nanog-ERT2 in *Nanog*-null ES cells. Also, Esrrb overexpression can sustain LIF-independent self-renewal even in the absence of Nanog, and *Esrrb*-null cells cannot sustain LIF-independent self-renewal upon Nanog overexpression, suggesting Esrrb is a crucial Nanog target (Festuccia et al., 2012).

1.1.2.1.5: Krüppel-like factors (Klfs)

The Krüppel-like factors (Klfs) are zinc-finger proteins that share homology with the *Drosophila melanogaster* segmentation gene, *Krüppel* (Schuh et al., 1986). *Klf2*, *Klf4* and *Klf5* are expressed in ES cells and have been

shown to promote pluripotency, whereas *Klf2* and *Klf4* are not expressed in EpiSCs (Brons et al., 2007; Tesar et al., 2007). Loss of all three factors drives differentiation of ES cells (Jiang et al., 2008). Expression of *Klf* factors in ES cells is induced by pluripotency factors: *Klf2* transcription is driven by Oct4, *Klf4* and *Klf5* expression is driven by the LIF-JAK/STAT3 pathway (Hall et al., 2009; Niwa et al., 2009). The expression level of Klf factors is critical for maintaining different state of pluripotency: ectopic expression of *Klf4* can convert EpiSCs into mES cell-like cells (Guo et al., 2009), whereas *Klf2* is essential for sustaining ground state pluripotency (Yeo et al., 2014). Klf factors could bind to the enhancer region of *Oct4* and *Nanog*, share many common targets with the core pluripotency factors (Oct4, Nanog and Sox2) (Chen et al., 2008b; Jiang et al., 2008; Parisi et al., 2008) and can substitute each other in somatic cell reprogramming (Nakagawa et al., 2008), suggesting some functional redundancy between the Klf factors.

1.1.2.1.6: Rex1

Rex1 (also referred to as ZFP42) is a zinc-finger transcription factor that was initially showed to be expressed in F9 embryonal carcinoma (EC) cells and is down-regulated after retinoic acid treatment to induce differentiation (Hosler et al., 1989). The expression of Rex1 was later confirmed in the germ cells of the testis and in ES cells (Rogers et al., 1991). Rex1 has been commonly used as a marker of the naïve pluripotent state (Toyooka et al., 2008), with its expression being either activated or repressed by Oct4 (Ben-Shushan et al., 1998), or being activated by Nanog (Shi et al., 2006). Unexpectedly, Rex1 function is likely dispensable for both the maintenance of pluripotency in ES cells and the development of embryos, as neither knockout nor overexpression of *Rex1* in ES cells affect cell pluripotency.

Moreover, *Rex1*-null embryos develop normally until after E17.5 (Masui et al., 2008). Rex1 also has an essential role in the reactivation of the inactive X chromosome by promoting the transcription of *Tsix* and inhibiting *Xist* (Gontan et al., 2012; Navarro et al., 2010). Therefore, the loss of Rex1 expression from pre- to post-implantation epiblast is coupled with random X-chromosome inactivation during this process.

1.1.2.2: Extrinsic factors governing ES cell self-renewal

1.1.2.2.1: LIF-STAT3

The initial derivation of mES cells was dependent on soluble molecules presented in serum, or secreted by a layer of mitotically inactivated embryonic fibroblasts (Evans and Kaufman, 1981; Martin, 1981), or Buffalo rat liver cells-conditioned medium, a culture medium which allowed for the dispensation of the feeder layer (Smith and Hooper, 1987). The subsequent identification of the leukemia inhibitory factor (LIF) as the peptide secreted by feeder cells, necessary for the maintenance of undifferentiated mES cells in culture, opened a new window for investigating the molecular mechanisms regulating pluripotency (Smith et al., 1988; Williams et al., 1988).

LIF belongs to the interleukin family of cytokines and activates a heterodimeric receptor consisting of two related receptors, the LIF receptor (LIFR), and the GP130 receptor (Boulton et al., 1994; Gearing et al., 1992), resulting in the activation of both Janus Kinase/Signal Transducers and Activators of Transcription (JAK/STAT) and MAPK pathways (Stahl et al., 1994). The primary STAT molecule activated by LIF signaling is STAT3 and overexpression of a constitutively active version of this molecule (STAT3-ER) allows mES cells to self-renew in the absence of LIF (Matsuda et al., 1999). In

addition, lack of STAT3 activity results in ES cell differentiation (Burdon et al., 1999; Niwa et al., 1998). Although LIF is required *in vitro* to derive and maintain mES cells, mouse blastocysts lacking *Lif* could implant in wild type female uteri but not in *Lif*-null uteri *in vivo*, suggesting embryonic LIF is not essential (Chen et al., 2000; Stewart et al., 1992).

Although LIF signalling is thought to maintain ES cell pluripotency, it is insufficient to maintain pluripotency in defined culture conditions without the presence of serum (Ying et al., 2003). Culturing ES cells in serum-free N2B27 medium with LIF results in emergence of $Sox1^+$ neural precursors, indicating that there are other soluble factors present in serum that are essential for the maintenance of the undifferentiated state.

1.1.2.2.2: BMP4

The components of serum used for routine cell culture (in this study, FCS for fetal calf serum) are undefined, and sometimes vary from one batch to another. The differentiation-inhibiting factor present in FCS was identified in 2003 as bone morphogenic protein 4 (BMP4) (Ying et al., 2003). ES cells cultured in serum-free N2B27 medium supplemented with LIF and BMP4 can self-renew, retain their ability of multi-lineage differentiation and can contribute to chimaeras. LIF and BMP acts concurrently to preserve pluripotency in ES cells, as ES cells cultured in BMP4 alone differentiate into non-neural fates (Ying et al., 2003).

BMPs are members of the transforming growth factor-β (TGF-β) superfamily of secreted signaling molecules which is known to have important functions in many biological contexts. They bind to specific heterodimeric transmembrane serine/threonine kinase receptors, which

transduce the signal to the nucleus through similar to mothers against decapentaplegic homologue (SMADs) proteins (reviewed in van Bubnoff and Cho, 2001), intracellular signal transduction molecules that fall into three categories: receptor-regulated SMADs (R- SMADs), cooperating SMADs (Co-SMADs) and inhibitory SMADs (I- SMADs). BMP activation of SMAD4, the sole Co- SMAD known in mammals, culminates in the expression of inhibitor of differentiation (Id) proteins, which act to block differentiation in a lineage specific manner, as BMPs have been shown to have powerful anti-neural activity *in vivo* (Di-Gregorio et al., 2007). Overexpression of Id proteins in mES cells can substitute for BMP4 or serum in maintaining self-renewal of ES cells (Ying et al., 2003).

1.1.2.2.3: Wnt/GSK-3β

Wnt signaling has been reported as a complex signalling pathway to be involved in a vast number of functions in tissue homeostasis and development. There are two types of Wnt signaling pathways, canonical and non-canonical. Canonical Wnt signaling involves the presence of β-catenin, while non-canonical Wnt signaling operates without it. In canonical Wnt signaling, Wnt proteins bind the cell surface receptor Frizzled (Fz)/ low density lipoprotein (LDL) receptor-related protein (LRP) complex at the cell surface. These receptors transduce a signal to several intracellular proteins that include Dishevelled (Dsh), glycogen synthase kinase-3β (GSK-3β), Axin, Adenomatous Polyposis Coli (APC), and the transcriptional regulator, β-catenin. Cytoplasmic β-catenin levels are normally kept low through continuous proteasome-mediated degradation, which is controlled by a complex containing GSK-3β/APC/Axin. When cells receive Wnt signals, the

degradation pathway is inhibited, and consequently β-catenin accumulates in the cytoplasm and nucleus. Nuclear β-catenin interacts with transcription factors such as lymphoid enhancer-binding factor 1/T cell-specific transcription factor (LEF/TCF) to affect transcription (Logan and Nusse, 2004).

It has been shown that GSK-3β, a component of the destruction complex, is the kinase responsible for β-catenin phosphorylation. Furthermore, it was also demonstrated that direct inhibition of GSK-3β could mimic Wnt signaling (Finlay et al., 2004). Inhibition of GSK-3β also allows nuclear re-localization of β-catenin and promotes expression of the pluripotency factor *Rex1* (Sato et al., 2004). The inhibition of GSK-3β-induced enhancement of ES cell self-renewal is through the modulation of TCF3 (Wray et al., 2011). Stimulation of ES cells with Wnt3, Wnt3a, Wnt5a and Wnt6 was also shown to promote pluripotency (Hao et al., 2006; Ogawa et al., 2006), and this positive effect on self-renewal could be stimulated further with a combination of Wnt and LIF or FGF inhibition (Singla et al., 2006).

1.1.2.2.4: Fibroblast growth factor (FGF) signaling

FGF signaling plays an important role in regulating pluripotency and lineage segregation in both the early mouse embryo and in pluripotent mammalian stem cells. The importance of FGF signaling in the early mouse embryo is well established, as mutations in *Fgf4*, *Fgfr2*, *Frs2a*, *Grb2* and *Erk2 (Mapk1)* all result in peri-implantation lethality (reviewed in Lanner and Rossant, 2010). FGF4 is the predominantly expressed ligand in the early embryo and has been shown to be under the direct regulation of the pluripotency factors Oct4 and Sox2 (Yuan et al., 1995).

Although FGF4 is produced in an autocrine fashion by undifferentiated ES cells (Ma et al., 1992), it is dispensable for the maintenance of the undifferentiated state of mouse ES cells, as ES cells lacking *Ffg4* are viable and display no obvious proliferative defect under self-renewing conditions (Wilder et al., 1997). FGF signalling appears to negatively affect ES cell self-renewal, and its inhibition promotes and homogenizes the expression of transcription factors associated with pluripotency (Burdon et al., 1999; Kunath et al., 2007; Wilder et al., 1997; Ying et al., 2008). Study on the *Fgf4*-null and *Erk2*-null ES cells demonstrated that FGF/ERK signaling is required to instruct ES cells to exit from the self-renewal program and to begin differentiation (Kunath et al., 2007). A similar phenotype was observed *in vivo* as the ICM of *Fgf4*-null blastocysts become uniformly positive for the epiblast determinant Nanog and fails to correctly segregate primitive endoderm (Chazaud et al., 2006; Nichols et al., 2009b), suggesting FGF4/ERK is also required for ES cells to exit to primitive endoderm (PrE) fate. Canham et al. showed that PrE lineage-primed, *Hex*-expressing population within ES cells can be reduced by inhibiting ERK signaling or promoted by stimulating phosphorylation with sodium vanadate (Canham et al., 2010). Moreover, inhibition of FGF/ERK signaling promotes epiblast maintenance, whereas excess FGF instructs the formation of primitive endoderm (Yamanaka et al., 2010). Therefore, FGF/ERK signaling pathway is critical for early lineage segregation in pre-implantation embryos and plays an important role in modulating the heterogeneity of pluripotent metastable states of ES cells (Figure 1.1).

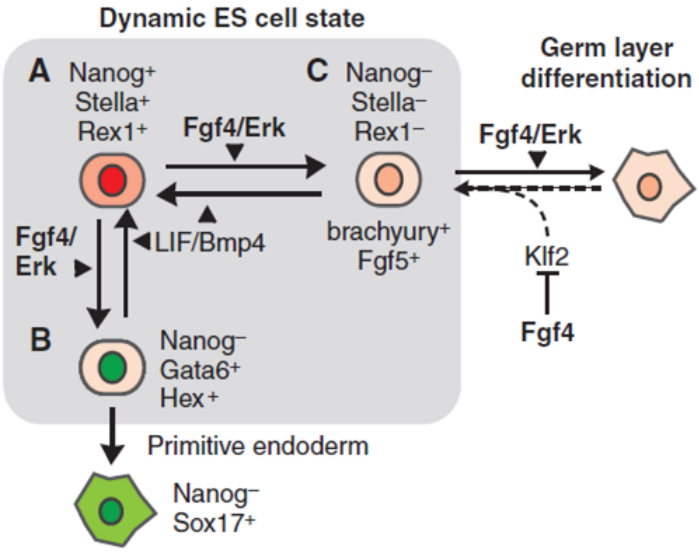


Figure 1.1: FGF signalling is involved in the regulation of dynamic state of mouse ES cells (modified from Lanner and Rossant, 2010)

ES cells exhibit great heterogeneity, whereby some cells appear to be transiently primed for differentiation. At least three states have been suggested: (A) unprimed cells (*Nanog*$^{+}$/*Stella*$^{+}$/*Rex1*$^{+}$); (B) cells primed for primitive endoderm (PrE) (*Gata6*$^{+}$/*Hex*$^{+}$); and (C) cells primed for germ layer differentiation (*Fgf5*$^{+}$/*Brachyury*$^{+}$). FGF4 signalling has been shown to maintain both primed states. LIF and BMP appear to promote ES self-renewal by counterbalancing the pro-differentiation autocrine FGF4/ERK signal. The grey box highlights the cell populations that readily revert back to an unprimed state. Dashed lines indicate suggested function.

1.1.2.2.5: 2i culture and the 'Ground State' hypothesis

As described above, all of these signalling pathways are capable of affecting the self-renewal ability of embryonic stem cells, either positively or negatively, but no single pathway stimulation or inhibition can maintain ES cells in a pluripotent state. Therefore, a combination of these pathways is required for long-term maintenance of pluripotency.

Based on the observation that FGF/ERK was the primary auto-inductive cue that caused mES cells to commit to specific lineages and block β-catenin phosphorylation could enhance ES cell self-renewal, Ying and colleagues developed a self-renewal medium containing only two small molecule inhibitors to block MEK (PD0352901) and GSK-3β (CHIR99021) , a medium they termed "2i" (Ying et al., 2008). ES cells cultured in 2i are shielded from differentiation triggers and ES cells express pluripotency factors more homogeneously (Wray et al., 2011). 2i culture also enabled derivation of germline-competent ES cells from refractory mouse strains (Hanna et al., 2009; Kiyonari et al., 2010; Nichols et al., 2009a). The same culture conditions have been applied to derive ES cell lines from the rat that are capable of contributing to adult chimaeras and passing through the germ line (Buehr et al., 2008; Li et al., 2008).

From this data, the authors propose a model that 2i culture conditions allowed ES cells to remain in a "ground state", and suggested that ES cells are a basal cell state that is intrinsically self-maintaining if shielded effectively from inductive differentiation stimuli including autocrine FGF4 (Ying et al., 2008). Addition of LIF is not required for self-renewal in 2i; however, its addition reinforces the ground state through further upregulation of *Klf4* and *Tcfcp2l1* (Martello et al., 2013). Providing any two of

the three 2i/LIF components is sufficient to support the formation of undifferentiated colonies (Wray et al., 2010), implying complementary inputs to a flexible gene regulatory network.

In summary, maintaining pluripotency within ES cells can be achieved by either activating of inhibiting different combination of signaling pathways. However, no single pathway activation/inhibition is sufficient for long-tern propagation of self-renewing colonies. The flexibility of ES cells to adapt to different culture conditions implied the ability of cells within the ICM, the *in vivo* counterpart of ES cells, to maintain a robust pluripotent state while being exposed to a variety of extrinsic signals.

1.1.3: Cellular heterogeneity

In recent years, using fluorescent reporters and single cell analysis, reports have emerged of heterogeneous expression of transcription factors in ES cells. These include several genes known to have roles in the establishment and/or maintenance of pluripotency, as well as lineage specification: *Nanog* (Chambers et al., 2007), *Rex1* (Toyooka et al., 2008), *Stella* (Hayashi et al., 2008), *Klf4* (Festuccia et al., 2012; Niwa et al., 2009) and *Hex* (Canham et al., 2010). Using fluorescence-activated cell sorting (FACS), sorted fractions can re-establish the parental steady-state heterogeneous distribution when return to culture conditions supporting self-renewal, suggesting that the expression of these factors is dynamic within ES cells.

Heterogeneity of gene expression might have a functional role in cell fate decisions. Rex1-GFP-positive cells plated in serum generate colonies of undifferentiated cells, whereas Rex1-GFP-negative cells produce predominantly small and differentiated colonies (Toyooka et al., 2008).

Nanog-GFP-high cells are associated with a reduced probability of differentiation whereas Nanog-GFP-low cells were shown to be more prone to differentiate (Chambers et al., 2007). In the case of Stella, the ES cells required co-culture with fibroblasts to display heterogeneity. Stella-GFP-positive ES cells are like the pre-implantation inner cell mass (ICM), whereas Stella-GFP-negative cells were like the post-implantation epiblast-derived stem cells (EpiSCs) (Hayashi et al., 2008). Hex-GFP-positive cells have an increased propensity to contribute to extra-embryonic endoderm when reintroduced into blastocysts, with the opposite propensity observed for Hex-GFP-negative cells (Canham et al., 2010). Therefore, the heterogeneous expression of these transcription factors reflects distinct cell states that coexist within the population despite their genetic homogeneity and raise the notion that pluripotency may inherently be a metastable condition (Graf and Stadtfeld, 2008; Hayashi et al., 2008) (Figure 1.1).

Heterogeneity of pluripotent cells may arise via different modes of transcriptional regulation, as well as post-transcriptional events such as protein synthesis and cell cycle dynamics, but the contribution of these different steps to heterogeneity is largely unknown (Torres-Padilla and Chambers, 2014). Recent advances in understanding transcription factor heterogeneity have come from work that focuses on the regulation of Nanog. Nanog protein was shown to decrease the level of transcription of the *Nanog* gene (Navarro et al., 2012), suggesting feedback loops among the gene regulation network components regulate ES cells heterogeneity. Non-transcriptional regulation of heterogeneity might also occur, for example through modulation of signalling pathways. It is shown that ES cells cultured in 2i/LIF expressed Nanog and Rex1 more homogeneously (Wray et al., 2010; Ying et al., 2008), suggesting that modulating signaling pathways could

stabilize ES cells in the naïve ground state which is not intrinsically fluctuating.

Taken together, revealing the molecular basis of the heterogeneity is not only important for understanding the flexible nature of the pluripotent state but might also help to dissect functional implications for the developmental potential of the cell.

1.1.4: Epigenetic regulation of mES cell identity

Recent analysis of specific epigenetic features of human and mouse stem cells has provided important insights into the unique properties of pluripotent and lineage-restricted stem cells. It is considered that transcription factors initiate lineage-specific gene expression programmes and epigenetic regulation contributes to stabilization of expression patterns. Epigenetic regulation of gene expression consists of several layers, including DNA methylation status, histone modifications, nucleosome packaging and rearrangement of higher-order chromatin structures and implementation of RNAi pathways and non-protein-coding RNAs (ncRNA) (Li et al., 2012; Lunyak and Rosenfeld, 2008). Perturbation of epigenetic regulation may cause alterations in the pluripotent stem cells identity, such as their self-renewal and differentiation potentials.

1.1.4.1: Changes in DNA methylation

One well-studied epigenetic process is DNA methylation of cytosine at CpG dinucleotides, which constitutes the first layer of the epigenetic code that is important for lineage specification and reprogramming to

pluripotency. It is shown that approximately 70% of all CpG dinucleotides are methylated and associated with regions of heterochromatin (Bird, 2002). The remaining 30% of un-methylated CpG dinucleotides are generally associated with the gene promoter regions and referred to as CpG islands. Although the global CpG content in ES cells is similar (Meissner et al., 2005), the distribution of the mark is unlike that of any other somatic cell type and also very distinct from the hypomethylated ICM (Meissner et al., 2008).

The DNA methylation pattern of the mammalian genome is maintained by DNA methyltransferases (DNMTs). DNMT3a and DNMT3b, the *de novo* DNA methyltransferase, were involved in re-methylating the genome following implantation (Okano et al., 1999). ES cells can be established and maintained in the absence of DNMTs and DNA methylation (Tsumura et al., 2006). However, *Dnmt*-deficient ES cells (*Dnmt [3a$^{-/-}$, 3b$^{-/-}$]*, *Dnmt1$^{-/-}$*) are markedly deficient in differentiation (Jackson et al., 2004), which is likely due, at least in part, to their inability to completely silence genes encoding Oct4 and Nanog (Li et al., 2007).

DNA demethylation also plays important roles in development and reprogramming. Ten-eleven translocation (TET) family of enzymes convert 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) in DNA. Of the three mammalian TET proteins, TET1 and TET2 are the major regulators of 5hmC levels in mouse embryonic stem cells. TET1 and TET2 are Oct4-regulated enzymes that together sustain 5hmC in mouse ES cells (Koh et al., 2011). During ES cell differentiation, the amount of TET1 and 5hmC decreases (Tahiliani et al., 2009), and TET1 knockdown impairs the self-renewal and maintenance of ES cells (Ito et al., 2010). In somatic reprogramming, demethylation of pluripotency gene promoters is a rate-limiting step towards the fully reprogrammed iPSCs state. Two recent papers

provide a link between the activation of the pluripotency network and changes in DNA hydromethylation status by examining the role of TET proteins (Costa et al., 2013; Gao et al., 2013b).

1.1.4.2: Histone modifications

In ES cells, most differentiation-related genes are regulated by chromatin-mediated mechanisms rather than by DNA methylation. Regulation of gene expression at the chromatin level involves histone modifications and rearrangement of nucleosomes and higher-order chromatin structures. Compared to DNA methylation, these epigenetic modifications are more dynamically regulated during development (Li et al., 2012).

Various histone modifications have been extensively studied in pluripotent stem cells. Polycomb group (PcG) and Trithorax group (TrxG) proteins are major effectors of histone modifications. PcG proteins are transcriptional repressors that have a role in maintaining repression of developmental regulator genes from plants to mammals (Beisel and Paro, 2011). PcG proteins catalyses H3K27 methylation (H3K27me3), a major repressive histone modification, at promoters of many key developmental regulators (Guenther et al., 2010). In pluripotent stem cells, PcG and TrxG proteins are essential for maintaining the balance between self-renewal and differentiation. PcG proteins are not necessary for maintenance of embryonic stem cell self-renewal (Chamberlain et al., 2008). However, cells lacking PRC2 are more prone to differentiation, which might due to increase the transcription of many development-associated PRC2 targets (Boyer et al., 2006). The activity of TrxG proteins opposes the action of PcG proteins

during development, by catalysing H3K4 methylation (H3K4me3), a major activating epigenetic mark (Mikkelsen et al., 2007; Zhao et al., 2007). Genome-wide ChIP analysis performed independently by several groups revealed that H3K27me3 modifications are often found at developmental gene promoters that are also marked with H3K4me3 (Bernstein et al., 2006; Schubeler et al., 2004), termed "bivalency". Bivalency is thought to keep developmental genes silenced yet poised for activation. A significant number of 'bivalent' domains are enriched for binding sites for at least one of the three pluripotency associated factors, Oct4, Nanog and Sox2, which implies that their formation and/or maintenance might be regulated through these 'master regulators'.

Chromatin in pluripotent stem cells is increasingly being recognized as open when compared with somatic cells, implying that its overall structure is less condensed and that the ratio between euchromatin and heterochromatin is higher than in differentiating cells (Efroni et al., 2008; Meshorer et al., 2006). The open ES cell chromatin structure, which is enriched in non-compact euchromatin, allows easy access for transcription factors and the transcriptional machinery and may explain the observed global 'hyper-transcription'. By contrast, lineage commitment is accompanied by the accumulation of regions of highly condensed, transcriptionally inactive heterochromatin (Efroni et al., 2008; Meissner, 2010). The formation of heterochromatin at the pericentric and telomeric regions as well as the variation and propagation of epigenetic states to a dynamic chromatin template have been extensively studied (reviewed in Fodor et al., 2010). In addition, several histone methylases and deacetylases play essential functions in remodelling of heterochromatin in pluripotent stem cells and in development. The histone methyltransferases G9a and Eset catalyse histone

H3 lysine 9 methylation (H3K9me) and contribute to gene regulation in the early embryo (Dodge et al., 2004; Feldman et al., 2006). In ES cells, knocking down G9a increases the efficiency of iPSCs generation (Epsztejn-Litman et al., 2008) whereas Eset-Oct4 interaction regulates pluripotency and represses trophectoderm differentiation (Lohmann et al., 2010; Yeap et al., 2009).

The nucleosome remodelling and deacetylase (NuRD) complex containing both a nucleosome remodelling ATPase and histone deacetylases (HDACs) has also been implicated in the establishment of the epiblast lineage. Loss of NuRD components such as Mbd3 de-represses trophectoderm markers and impairs ES cell differentiation (Fazzio and Panning, 2010; Kaji et al., 2007). Recently, a genome-wide study has shown that TET1 (5hmC) overlaps with Mbd3-binding sites in ES cells, suggesting crosstalk between the Mbd3 and 5hmC pathways (Yildirim et al., 2011). These studies demonstrate the dedicated interplay of different layers of epigenetic regulation in controlling pluripotency and differentiation.

1.1.5: Transition from naïve to primed pluripotency in the embryo and in culture

During early mammalian development, as the pluripotent cells that give rise to all of the tissues of the body proliferate and expand in number, they pass through transition states marked by a stepwise restriction in developmental potential and by changes in the expression of key regulatory genes. *In vitro*, pluripotency can be maintained indefinitely through derivation of stem cell lines derived from different stages of mouse development to mimic these transition states. Different classes of pluripotent stem cell have distinct culture requirements and gene expression programs,

likely reflecting the dynamic development of the epiblast in the embryo (Nichols and Smith, 2012).

1.1.5.1: Early mouse development

Embryogenesis begins from a "totipotent" zygote that will give rise to the entire embryo and to the extra-embryonic lineages followed by progressive loss of developmental capacity. The zygote undergoes several rounds of stereotyped cleavage divisions without increasing in volume but half individual cell size at every division then form blastocyst, with a fluid filled cavity, known as the blastocoel cavity (Rossant and Tam, 2009). Two distinct lineages have emerged from the early blastocyst that exhibit morphological and molecular differences: trophectoderm (TE) and inner cell mass (ICM). The TE, which is characterized by expression of the transcription factor caudal type homeobox 2 (Cdx2) among other specific markers (Beck et al., 1995), and by its epithelial nature, will contribute primarily to the placenta. The ICM further segregates into the pluripotent epiblast (which give rise to ES cells) and the primitive endoderm (PrE). Oct4 is expressed by all cells in the ICM and is essential to establish the distinct identity of the ICM. In the absence of Oct4, ICM cells loss their pluripotency and eventually differentiate along the trophoblast lineage (Nichols et al., 1998). Oct4 expression is negatively correlated with Cdx2 and this appears to be critical for the segregation of trophectoderm and ICM (Niwa et al., 2005). The epiblast and PrE cells express distinct molecular signatures: the epiblast is characterized by the expression of the transcription factor Nanog, and the PrE cells by the expression of the transcription factors Gata6 and Gata4 (Chambers et al., 2003; Chazaud et al., 2006; Plusa et al., 2008). Nanog and Gata6 may act in a mutually antagonistic manner to determine the two lineages (Plusa et al., 2008).

The acquisition of epiblast identity coincides with the reactivation of the inactive X chromosome in female mouse embryos (Mak et al., 2004; Okamoto et al., 2004; Silva et al., 2009). Reactivation occurs transiently in the pluripotent lineage prior to implantation (Heard, 2004), X reactivation is therefore a hallmark of the successful formation of ES cells (Silva and Smith, 2008).

The segregation of the epiblast and primitive endoderm is mediated by FGF signalling. Chemical inhibition of the FFG/ERK pathway entirely suppresses PrE development and concomitantly expands the epiblast (Nichols et al., 2009b; Yamanaka et al., 2010). Conversely, administering high concentrations of FGF4 to early embryos can drive the entire ICM to form PrE cells (Yamanaka et al., 2010). The segregation of the two lineages may be consolidated by reciprocal expression of FGF receptor in the PrE cells and FGF4 in the epiblast (Guo et al., 2010).

Following implantation, which is around E4.5 for mouse embryo, the mouse epiblast grows from a ball of cells into a cup-shaped epithelium, known as the egg cylinder. In addition, random X-inactivation occurs in female embryo throughout the epiblast at E5.5 (Rastan, 1982). This morphological change is accompanied by the reduced expression of naïve pluripotency transcription factors such as *Nanog* and *Rex1* and by the upregulation of *Fgf5* and *T-Brachyury* (Chambers et al., 2003; Pelton et al., 2002). Post-implantation epiblast cells are pluripotent, as they maintain the capacity to give rise to teratocarcinomas and EC cells (Solter et al., 1970; Stevens, 1970). Although post-implantation epiblast cells are unable to contribute to blastocyst chimaeras, it could be they are incompetent to respond to developmental cues in the blastocyst, as lineage tracing using

vital dye labeling of single cells during early gastrulation *ex vivo* reveals progeny in all three germ layers (Lawson et al., 1991).

1.1.5.2: Comparison between naïve ES cells, primed ES cells and post-implantation Epiblast stem cells (EpiSCs)

ES cells are derived from the epiblast of the pre-implantation embryos, as described in section **1.1.1.2**, representing a more naïve pluripotent state when compared with EpiSCs derived from post-implantation epiblast (see section **1.1.1.3**). However, ES cells cultured routinely in the medium containing serum and LIF do display heterogeneous gene expression (see section **1.1.3**). The heterogeneity of gene expression reflects distinct cell states within a pluripotent culture and these metastable states show biases in their differentiation potentials (Graf and Stadtfeld, 2008; Hayashi et al., 2008) (Figure 1.1). Culturing ES cells in two inhibitors (2i), PD0325901 (MEK inhibitor) and CHIR99021 (GSK-3β inhibitor), shield pluripotent cells from differentiation triggers: FGF4 stimulation of the MEK-ERK pathway and endogenous repressor activity of TCF3 (Kunath et al., 2007; Wray et al., 2011). Thus, ES cells cultured in 2i are considered as in "ground state" of pluripotency with rather homogeneous gene expression (see section **1.1.2.2.5**) (Wray et al., 2010; Ying et al., 2008).

The naïve ES cells (ground state, 2i-ESC) differ from primed ES cells (serum-ESC) at both transcriptional and epigenetic levels. 2i-ESC express lower level of lineage affiliated genes (*Pax6, Runx1, T-Brachyury*), reduced repressive histone mark H3K27me3 at promoters, and fewer bivalent domains, which mark genes are inactivated but poised for activation or repression (see section **1.1.4**) (Marks et al., 2012). Therefore, 2i culture is likely to "erase" epigenetic marks generated in serum culture, as cells

cultured in different batch of serum sometimes do display different differentiation potential (unpublished data in the lab).

Despite the difference between naïve ES cells and primed ES cells, there are more studies focusing on the differences between pre-implantation epiblast-derived ES cells and post-implantation epiblast derived EpiSCs (Nichols and Smith, 2009). EpiSCs can also be derived from ES cells by extensive passaging from serum-containing ES cell medium to EpiSC medium which is serum-free medium supplemented with Activin/FGF2 (Guo et al., 2009), or by directly conversion from 2i/LIF culture to EpiSC medium within few days (Hayashi et al., 2011). The differences between naïve and primed states of epiblast are summarized in table 1.1. It is worth noting that there can be some differences in the use of terminology, for example in the table 1.1 the naïve pluripotent state is considered as a general state for pre-implantation epiblast-derived ES cells and the primed state as post-implantation epiblast derived EpiSCs, whereas above ES cells cultured in serum-containing medium were considered to be in a primed state, as distinct from cells cultured in 2i.

In summary, morphologically, EpiSCs share with ES cells a large nuclear-to-cytoplasmic ratio and prominent nucleoli, but their morphology is more two-dimensional and epithelial. The transition from naïve to primed pluripotent state is accompanied with DNA methylation and X-inactivation and loss of E-cadherin expression. The transition from naïve to primed pluripotent state is not irreversible, as it is reported that post-implantation epiblast cells and established EpiSCs can 'overcome' a robust epigenetic barrier and undergo reversion to ES-like cells, either by culturing E5.5-E7.5 pluripotent epiblast in ES cell medium containing LIF and serum on feeder cells (Bao et al., 2009), or by overexpression of either Nanog or Klf4 together

with a change in culture conditions (Guo et al., 2009; Osorno et al., 2012; Silva et al., 2009).

The transition from a naïve to primed pluripotent state, and in more detail, the transition from naïve ES cells to primed ES cells, is the initial important step for the commitment to differentiation. Therefore, understanding the genetic and epigenetic mechanisms controlling these transitions is important, yet these mechanisms are not well understood and are only just beginning to be unravelled.

Table 1.1: Comparison of naïve and primed pluripotent states (adapted from Nichols and Smith, 2009)

Property	Ground state	Primed state
Embryonic tissue	early epiblast	egg cylinder; embryonic disc
Culture stem cell	rodent ES cells	rodent EpiSCs primate "ES cells"
Chimaeras	blastocyst	gastrulation stage embryo
Teratomas	yes	yes
Differentiation bias	none	variable
Pluripotency factors	Oct4, Nanog, Sox2, Klf2, Klf4	Oct4, Sox2, Nanog (low)
Naïve markers	Rex1, NrOb1, Fgf4	absent
Specification markers	absent	Fgf5, T
Response to LIF/STAT3	self-renewal	none
Response to FGF/Erk	differentiation	self-renewal
Clonogenicity	high	low
XX status	XaXa	XaXi
Response to 2i	self-renewal	differentiation/death
DNA methylation	low	high

1.1.6: Embryonic stem cells as a model for lineage specification

The study of differentiating embryonic stem cells represents a convenient experimental tool for the analysis of the molecular events that accompany lineage specification:

• mES cells are easy to grow and can be maintained in chemically defined culture medium, which simplified the complexity of understanding extracellular signalling pathways in their maintenance and differentiation.

• mES cells double rapidly (12-15 hours) which can be useful for biochemical studies where large amounts of materials are needed.

• mES cells exhibit a high frequency of homologous recombination (10^{-5}-10^{-6} per electroporated cell) (Templeton et al., 1997) which allow rapid and straightforward transgenesis experiments for genetic studies.

mES cells can be differentiated into various cell types with high efficiency. Three basic methods have been developed to promote differentiation of ES cells *in vitro*: (1) the formation of three-dimensional aggregates known as embryoid bodies (EBs), (2) the culture of ES cells as monolayers on extracellular matrix proteins, and (3) the culture of ES cells directly on supportive stromal layers.

1.1.6.1: Embryoid body (EB) formation

The most reliable method for generating multi-lineage differentiated cell types is the induction of EB formation, through which ES cells spontaneously develop into three-dimensional, multi-cellular aggregates of differentiated and undifferentiated cells. An EB consists of ectodermal, mesodermal, and endodermal tissues, which recapitulate many aspects of cell differentiation

during early mammalian embryogenesis and differentiate into derivatives of all the three germ layers (Desbaillets et al., 2000). A common technique used is to simply deprive ES cells of contact with feeder cells, or from the presence of LIF, and culture ES cells under the condition in which they are unable to adhere to the surface of the culture dishes. There are several basic methods to induce EB formation from pluripotent ES cells: suspension culture in bacterial-grade dishes, culture in methylcellulose semisolid media, or culture in hanging drops (Hopfl et al., 2004; Kurosawa, 2007).

The features of formed EBs are considered to be not homogeneous embryologically and morphologically. However, EBs could be classified as simple EBs or cystic EBs according to the stage of differentiation (Kurosawa, 2007). For mouse ES cells, spherical ES cell aggregates with morula-like structures formed in 2-4 days in suspension culture are called simple EBs. The outside of the simple EBs is lined by endoderm, and the inside cells contain variable mixtures of mesoderm and ectoderm. In the case of cystic EBs, a central cavity forms in EBs in 4-5 days in suspension culture. Cystic EBs resemble an embryo in the blastula or egg-cylinder stage, consisting of a double-layered structure with an inner ectodermal layer and an outer of endoderm enclosing the cavity. After 8–10 days in suspension culture, cystic EBs expand to larger cystic structures homologous to the visceral yolk sac of post-implantation embryos.

EBs recapitulate many aspects of cell differentiation during early embryogenesis, and play an important role in the differentiation of ES cells into a variety of cell types in *vitro*. However, EBs do present disadvantages. The consistent of mixed cell types within single EB makes the culture too heterogeneous to easily study the specification of particular lineage. Also, the 3-dimensional structure of EBs makes them less practical for imaging *in vivo*.

Thus, although the multi-lineage differentiation potential of EBs served as a great tool for generating non-biased environment for analysing ES cells differentiation, other differentiation methods is needed for accessing cell differentiation for single lineage analysis.

1.1.6.2: Monolayer differentiation in N2B27 medium

Differentiating ES cells in adherent monolayers provides the solution for all of the problems discussed in the previous section. The monolayer differentiation was first established by culturing ES cells on feeder cells to induce specification of different lineages (Kawasaki et al., 2000; Nakano et al., 1994). However, the presence of feeder cells in these protocols made the culture as a non-chemically defined context and thus the differentiation efficiency varied from time to time. Feeder-free ES cell differentiation protocols were then developed by culturing cells on extracellular matrix-coated surface supplemented with different cytokines (Coraux et al., 2003; Nishikawa et al., 1998).

A currently routinely used medium for ES cell differentiation was devised in 2003 by Smith and colleagues by combining N2 and B27-supplemented serum-free basal media and named the new culture medium N2B27 (Ying and Smith, 2003). Culture of ES cells as an adherent monolayer plated on gelatine in N2B27 medium enable the conversion of mouse ES cells to either terminally differentiated neurons and glia or self-renewing but lineage-restricted neural stem cell lines (Pollard et al., 2006; Ying et al., 2003). This N2B27 medium was able to support cell survival and proliferation without providing lineage specifying signals, which support the "default

model of differentiation“ of neural induction in the mouse (Tropepe et al., 2001).

Differentiation of mES cells in N2B27 seems to follow the same molecular transitions that occur *in vivo*: downregulation of naïve pluripotent markers followed by transient upregulation of post-implantation epiblast transcript and subsequent acquisition of neural ectoderm gene expression pattern, indicating that N2B27 culture is a developmentally relevant *in vitro* differentiation system (Aiba et al., 2009; Aiba et al., 2006). Moreover, addition of particular combinations of cytokines in N2B27 medium redirects ES cells to differentiate toward endodermal, mesodermal or epithelial fates (Hansson et al., 2009; Malaguti et al., 2013; Ying et al., 2003), and supplementation of the medium with factors that promote self-renewal enables long-term maintenance of ES cells and EpiSCs (Brons et al., 2007; Tesar et al., 2007; Wray et al., 2010; Ying et al., 2003). Therefore, the N2B27 culture appears to be a well-defined and powerful tool for monitoring the *in vitro* differentiation process.

1.2: Basic helix-loop-helix (bHLH) family of transcription factors

The basic helix-loop-helix (bHLH) family of transcriptional regulatory proteins are key players in a wide array of developmental processes (Massari and Murre, 2000). Members of this family hold a highly conserved functional domain containing a stretch of basic amino acids adjacent to two amphipathic α-helices separated by a loop. The bHLH transcription factors can form either homo- or hetero- dimers depend on their different classes described below. The dimerized bHLH factors bind specifically to a hexa-

nucleotide sequence (CANNTG) called E-box, which is found in the control regions of many lineage-specific genes.

This family is found in a variety of eukaryotic organisms, ranging from the yeast to humans, and in the mouse it comprises over 100 genes, which are involved in a multiplicity of biological processes including cell differentiation, lineage commitment, and sex determination (Massari and Murre, 2000; Skinner et al., 2010). bHLH factors are required for a multitude of important developmental processes, including NeuroD in neurogenesis (Lee et al., 1995), MyoD in myogenesis (Lassar et al., 1989; Sassoon et al., 1989), SCL/TAL-1 in hematopoiesis (Porcher et al., 1996), and NGN3 in pancreatic development (Whelan et al., 1990), with these factors being well conserved between different model organisms.

1.2.1: bHLH superfamily classification

bHLH proteins are classified by tissue distribution, dimerization capabilities, and DNA-binding specificities into 7 classes (I to VII) (Murre et al., 1994).

Class I bHLH proteins (also known as the E proteins) include E12, E47, HEB, E2-2, and Daughterless. E12 and E47 represent spliced variants of *Tcf3* (also called *TcfE2A*) gene (Aronheim et al., 1993). E proteins are ubiquitous expressed in many tissues and capable of forming either homo- or hetero-dimers. Class II bHLH proteins include members such as Twist, MyoD, myogenin, Mash1 and NeuroD, and show a tissue-restricted pattern of expression. With very few exceptions (one exception being Twist (Connerney et al., 2008)), they are not able to form homodimers but preferentially dimerizing with class I proteins (Massari and Murre, 2000).

Class III bHLH proteins include the Myc family of transcription factors. Proteins of this class contain a leucine zipper adjacent to the bHLH motif. Class IV bHLH proteins define a family of molecules, including Mad, Max, and Mxi that are capable of dimerizing with the Myc proteins or with one another.

A group of bHLH proteins that lack a basic trans-activation region, including Id and Emc, define the class V bHLH proteins. Class V members are negative regulators of class I and class II bHLH proteins (Benezra et al., 1990).

Class VI HLH proteins have as their defining feature a proline in their basic region. This group includes the Drosophila proteins Hairy and Enhancer of Split (Hes). Finally, the class VII bHLH proteins are categorized by the presence of the bHLH-PAS domain and include members such as the aromatic hydrocarbon receptor, the AHR nuclear-translocator (ARNT) and hypoxia-inducible factor 1α (HIF-1α) (Massari and Murre, 2000).

1.2.2: Id as a negative regulator of bHLH transcription factors

The inhibitor of DNA binding/inhibitor of differentiation (Id) protein-1 was first identified in a screen for nucleotide sequences homologous to those encoding the second α-helix of the bHLH factors *Myc*, *Myod* and *Myog* (Benezra et al., 1990). There are 4 Ids in mammals with overlapping expression profiles: Id1, Id2 (Sun et al., 1991), Id3 (Christy et al., 1991) and Id4 (Riechmann et al., 1994). Id proteins, lacking the basic domain required for DNA-binding, acts by binding to ubiquitously expressed class I bHLH transcription factor co-factors (E proteins) and blocking their dimerization with tissue-specific class II proteins (Benezra et al., 1990). The Id-E protein

complex lacks DNA binding ability to the tissue-specific target promoters. Therefore, Id is dominant-negative regulator of class II bHLH transcription factors. The bHLH domain of Id1 was shown to be necessary and almost sufficient for the repression of the DNA binding activity of TcfE2A (E12/E47) and MyoD (Pesce and Benezra, 1993).

Id proteins play key roles in the regulation of lineage commitment, cell fate decisions and in the timing of differentiation during neurogenesis, lymphopoiesis and angiogenesis (Norton, 2000). Because of their overlapping expression patterns, single knockout of *Id1* has no effect on the viability of mice and results in no obvious phenotypic defect (Yan et al., 1997). $Id1^{-/-}/Id3^{-/-}$ double-null mice display aberrant neurogenesis with premature withdrawal of neuroblasts from the cell cycle and inappropriate expression of neural-specific markers (Norton, 2000). Knockout of any 4 out of 6 alleles of *Id1*, *Id2* and *Id3* (only one of the three genes is knocked out on both alleles, such as $Id1^{-/-};Id2^{+/-};Id3^{+/-}$) results in embryonic lethality due to defects in cardiogenesis from as early as E9.5 and no triple-knockout embryos survived to E11.5 (Fraidenraich et al., 2004). However, there has been no mouse made with all four *Id* genes knocked out. Therefore, it is possible that an earlier role for *Id* genes is masked by redundancy with *Id4*, although its expression in wild-type embryo is essentially restricted to the developing nervous system, whereas expression of *Id1–Id3* is much more widespread during mouse embryogenesis (Jen et al., 1997; Riechmann and Sablitzky, 1995).

In mouse ES cells, *Id* genes are down-stream targets of BMP signalling (Ying et al., 2003). BMP induced Id-upregulation prevents neural differentiation (Zhang et al., 2010). Forced over-expression of Id enables ES cells to bypass the requirement for either BMP or serum to maintain self-renewal ability (Ying et al., 2003). Expression of Id1 is rapidly downregulated

at the onset of differentiation, both in N2B27 serum-free monolayer differentiation and in serum-free embryoid body differentiation (Aiba et al., 2009; Aiba et al., 2006), consistent with its role in the maintenance of ES cell self-renewal.

Since Id proteins function through antagonism of bHLH transcriptional regulators, it is of interest to study Id-bHLH interactions in the regulation of mammalian cell fate determination.

1.2.3: Identification of Id protein binding partners in mouse ES cells

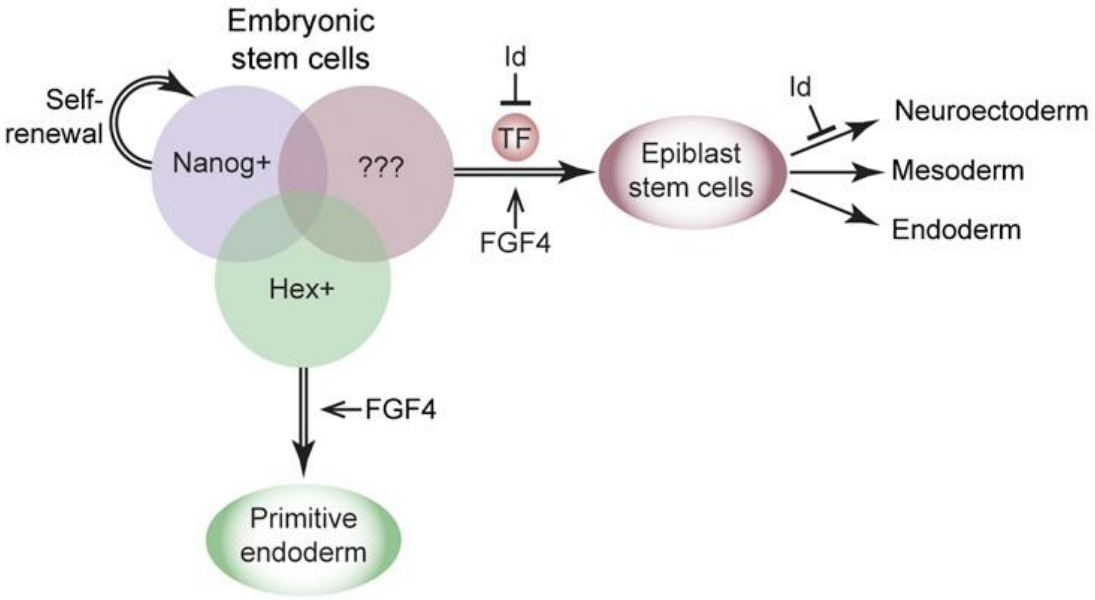
The transcription factors that act downstream of the FGF signalling pathway in order to drive epiblast cells toward differentiation-primed state are not known (Figure 1.2A). A clue to their identity comes from the finding that Id proteins are able to block the transition of ES cells to epiblast stem cells (EpiSCs) (Zhang et al., 2010). Therefore, epiblast priming could be driven by specific bHLH factors that are expressed in pluripotent cells but held in an inactive state through the action of Id proteins (Figure 1.2B). As soon as Id proteins are downregulated, the bHLH activity of these ''primed'' cells would be released from inhibition, allowing epiblast maturation to proceed.

Our previous lab member Dr. Owen Davies set out to identify the targets of Id inhibition by determining the direct binding partners of both Id and E proteins in ES cells. To achieve this, a series of yeast-two-hybrid (Y2H) screens for binding partners of Id1, E47, and E12 within a library generated from the mRNA of pluripotent mouse ES cells were performed. The only validated Id1-binding partners identified were E47 and E12. Id1 is therefore likely to function in ES cells through E protein sequestration, with

downstream bHLH targets constituting direct binding partners of E47/E12. Further Y2H screening identified five validated interacting partners for the HLH domain of E47/E12, two correspond to Id1 and Id3, confirming the results of the Id1 screen; the remaining three correspond to bHLH transcription factors Tcf15, Twist1, and NeuroD1 (Figure 1.3). These bHLH transcription factors are candidates for mediating the effect of Id1 in pluripotent cells.

Functional studies of NeuroD1 and Twist1 in pluripotent stem cells were carried out by other colleagues in the lab (Paul Nistor for Twist1 and Mattias Malaguti for NeuroD1). Therefore, this project will mainly focus on analysing the function of Tcf15.

A)



B)

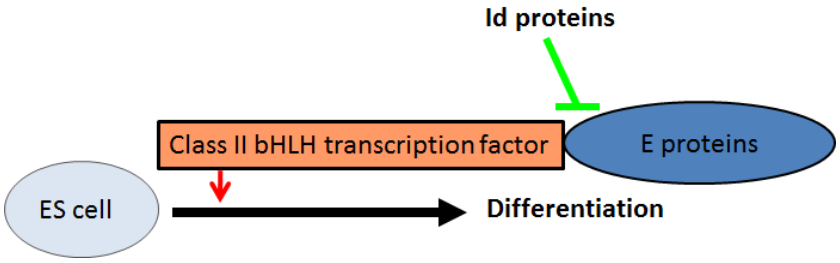


Figure 1.2: Id might act as an intrinsic factor to drive lineage specification

(A) FGF signalling primes pluripotent cells for differentiation. Id inhibits subsequent transitions: first to epiblast and then to neuroectoderm. Thus, unknown Id-regulated transcription factors (TF) might be expressed in epiblast-primed cells to regulate this process.

(B) The function of Id proteins is to block dimerisation between ubiquitously expressed E proteins and class II bHLH transcription factors. Therefore, the binding partners of E proteins in ES cells are likely to play important roles during either epiblast-priming and/or lineage specification.

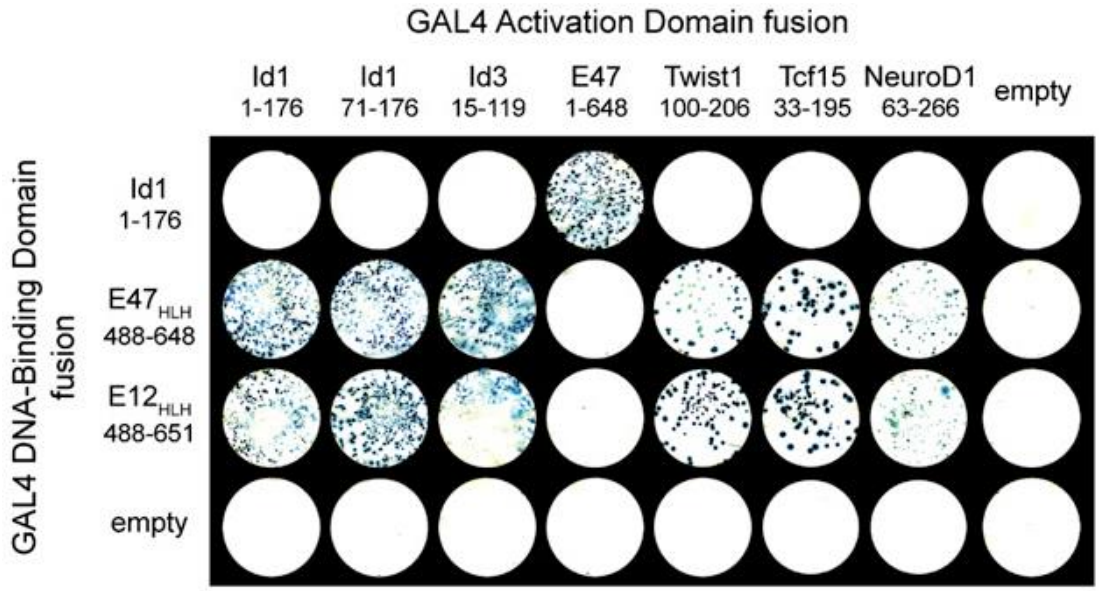


Figure 1.3: Yeast-two-hybrid screen to identify transcription factors regulated by Id1 in ES cells

High-stringency Y2H analysis of interactions among Id1, E47$_{HLH}$, and E12$_{HLH}$, and genes identified through library screens. Positive interactors are determined by growth selection and positive X-Gal reaction.

Y2H screen of Id1 binding partners identified coding sequences of the *TcfE2A* gene products E47. Y2H of the HLH domains of E47/E12 identified five validated interacting partners for E47/E12: Id1, Id3 (confirming the results of the Id1 screen) and bHLH transcription factors Tcf15, Twist1, and NeuroD1.

1.2.4: Tcf15 expression during mouse embryogenesis

Transcription factor 15 (Tcf15), also called Paraxis, is a bHLH transcription factor belongs to the Twist family (Barnes and Firulli, 2009). Prior to this study, the earliest expression of Tcf15 reported in mouse embryo was at E7.5 in a sub-domain of primitive mesoderm detected by *in situ* hybridisation. Between E8.0 to E12.5, the expression proceeds in a rostral to caudal wave in newly formed epithelial somites before the first detectable expression of the myogenic bHLH genes. Later in the mature somite, Tcf15 expression is down-regulated with expression gone after E13.5 (Burgess et al., 1995). The *Tcf15*-null embryo has segmented somites, but the newly formed somites are not fully epithelialized and fail to form a normal epithelial dermomyotome, suggesting that Tcf15 play important role in the morphological mesenchymal to epithelial transition (MET) during somitogenesis (Burgess et al., 1996). Although these *Tcf15* mutant embryos survive to birth, they die shortly afterward as a consequence of muscular-skeletal defects due to the failure in somite morphogenesis earlier in development. Moreover, Tcf15 is necessary for regulating anterior/posterior polarity in somites (Johnson et al., 2001).

Although Tcf15 has been thought as a developing somite marker in the post-implantation embryos, it was recently been reported that its expression was up-regulated during derivation of ES cells from E3.5 blastocysts (Tang et al., 2010). This result suggested that Tcf15 might play another important role in the pre-implantation embryos in addition to its known function in somitogenesis. Since many bHLH transcription factors such as MyoD or NeuroD play important roles in controlling the lineage-commitment of ES cells (Lassar et al., 1989; Lee et al., 1995), the early expression of Tcf15 might act as a lineage-priming pro-differentiation factor within ES cells.

1.3: Aims of this thesis

Despite a known function in controlling somite development (Burgess et al., 1996), a role for Tcf15 at earlier development stage has not been studied. Identification of Tcf15 as a putative Id1 target in ES cells suggests Tcf15 might mediate the effect of Id in pluripotent cells. Id proteins classically function through the inhibition of active bHLH transcription factors and are able to block the transition of ES cells to epiblast stem cells (EpiSCs) (Zhang et al., 2010). Therefore, the function of Tcf15 in ES cells could be driving epiblast priming, however, held in an inactive state through the action of Id proteins. As soon as Id proteins are downregulated, the bHLH activity of these ''primed'' ES cells would be released from inhibition, allowing epiblast maturation to proceed.

Since there is no report addressing the function of Tcf15 in early development, the aim of this thesis is to characterize the function of Tcf15 in pluripotent ES cells and during differentiation.

In this thesis, the expression of Tcf15 during differentiation *in vitro* and during early development *in vivo* will be described. In addition, Tcf15-Venus reporter cells were used to further investigate Tcf15 expression at single cell level. Gain of function and loss of function studies were performed to perturb Tcf15 expression in ES cells in order to assess the function of Tcf15 in self-renewal and during differentiation. Finally, the investigations into the mechanisms that regulate Tcf15 expression and the downstream targets of Tcf15 that mediate its effects in pluripotent cells were presented.

CHAPTER 2

Materials and Methods

2.1: Materials

All restriction enzymes were from either New England Biolabs (NEB) or Roche. All chemicals were obtained from either Sigma or Fisher Scientific. Synthetic oligonucleotides (primers) were synthesized by Eurogentec, with either sePOP desalt or PAGE for cloning primers. RNase/DNase free water from Gibco and Hyclone were used for DNA/RNA suspension and RNA-related solution preparation.

2.1.1: Kits and Reagents

2.1.1.1: Kits

Agilent- Absolutely RNA Purification Kits

Ambion-Illumina TotalPrep RNA Amplification Kit

Amersham- Illustra ProbeQuant™ G-50 Micro Columns

Amersham- MegaPrime™ DNA Labelling System

Invitrogen- Quant-iT™ Protein Assay Kit

Invitrogen- Zero Blunt® TOPO® PCR Cloning Kit

Promega- pGEM®-T and pGEM®-T Easy Vector Systems

Qiagen- QIAprep™ Spin Miniprep Kit

Qiagen- QIAfilter Plasmid Midi Kit

Qiagen- QIAquick Gel Extraction Kit

Qiagen- QIAquick PCR Purification Kit

Qiagen- RNeasy™ Mini Kit

Qiagen- RNeasy™ Micro Kit

Qiagen- DNeasy™ Blood and Tissue Kit

Sigma- Leukocyte Alkaline Phosphatase Assay Kit

2.1.1.2: Reagents

Ambion- tRNA

Ambion- RNase*Zap*® RNase Decontamination Solution

BD- LB agar and broth

Bio-Rad- Protein Assay Dye Reagent Concentrate

Invitrogen- DH5α

Invitrogen- dNTP Mix

Invitrogen- LDS Sample Loading Buffer

Invitrogen- MOPS SDS Running Buffer

Invitrogen- M-MLV Reverse Transcriptase

Invitrogen- NuPAGE® Novex® Bis-Tris Gel

Invitrogen- Platinum® Pfx DNA Polymerase

Invitrogen- Random primers

Invitrogen- RNaseOUT Recombinant Ribonuclease Inhibitor

Invitrogen- SeeBlue™ Prestained Protein Standard

Invitrogen- SuperScript® III Reverse Transcriptase

Invitrogen- SYBR® Safe DNA Gel Stain

Invitrogen- Trizol Reagent

Invitrogen- UltraPure™ Agarose

NEB- 100 bp and 1kb DNA Size Standards

NEB- λ DNA-*Hind*III Digest

NEB-RNase H

NEB- T4 DNA Polymerase

NEB- T4 DNA Ligase

Pierce- SuperSignal™ West Pico/Femto Chemiluminescent Substrate

Promega- DNaseI

Qiagen- Taq DNA Polymerase

Roche- Complete Protease inhibitors (EDTA Free)

Roche- DIG RNA Labelling Mix

Roche- LightCycler® 480 Probes Master Mix

Roche- UPL Set, Mouse

Roche- Proteinase K

Roche- rAPid Alkaline Phosphatase

Roche- SP6 RNA Polymerase

Roche- T7 RNA Polymerase

Sigma- Benzonase® Nuclease

Sigma- Bovine plasma fibronectin

Sigma- Deoxyribonucleic acid from herring sperm

Sigma- Donkey serum

Sigma- GMEM

Sigma- Pefabloc

Sigma- PerfectHyb™ Plus Hybridization Buffer

Sigma- PBS

Sigma- RIPA buffer

Thermo- RiboLock RNase Inhibitor

2.1.2: Solutions

Milli-Q filtered and deionized water (>18 mΩ-cm resistance) was used for making all stock solutions. (Except N2 was made in sterile DMEM-F12 media)

Solution	Formulation
PBS	137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.47 mM KH_2PO_4
1x TAE	40 mM Tris Base, 0.114 % Acetic Acid, 1 mM EDTA. pH 8.0
10x TBE	450 mM Tris Base, 450 mM Boric acid, 10 mM EDTA. pH 8.0
6x DNA loading dye for TBE gel	1x TBE, 0.25 % Orange G, 30 % Glycerol
6x DNA loading dye for TAE gel	1x TAE, 0.25% Bromophenol blue, 15% Ficoll 400
Protein lysis buffer	RIPA buffer, 1x Complete protease inhibitor cocktail, 0.5 mM Pefabloc, 1 mM DTT
Western blot transfer buffer	12 mM Tris pH 7.5, 96 mM Glycine, 10 % methanol
Western blot wash buffer (PBST)	PBS, 0.1 % Tween-20
Western blot blocking buffer	PBST, 5-10 % Skimmed Milk Powder or 5 % BSA (Fraction V)
Western blot stripping buffer	72.5 mM Tris pH 7.5, 125 mM 2-Mercaptoethanol, 2% SDS
ICC blocking buffer	PBS, 0.1 % Triton X-100, 3-5 % species specific serum
ICC wash buffer	PBS, 0.1 % Triton X-100
gDNA lysis buffer	10 mM Tris-HCl pH 7.5, 10 mM EDTA, 10mM NaCl, 1% SDS, 1 mg/mL Protease K
20x SSC (Southern blot)	3 M NaCl, 340 mM Tri-Sodium Citrate pH 7.4
20x SSC (ISH)	3 M NaCl, 340 mM Tri-Sodium Citrate pH 4.5 (in Nuclease-free water)
Southern blot denaturing solution	500 mM NaOH, 1.5 M NaCl

Southern blot neutralization solution	500 mM Tris, 1.5 mM NaCl, pH 8.0
Southern blot wash buffer	2x SSC, 0.1 % SDS and 0.5x SSC, 0.1 % SDS
Southern blot stripping buffer	0.2 M NaOH, 0.1% SDS
N2 supplement	2.5 mg/mL Insulin from bovine pancreas, 10 mg/mL Apo-transferrin, 7.5 mg/mL BSA, 0.002 mg/ml Progesterone, 1.6 mg/ml Putrescine and 3 μM Sodium Selenite
Citrate solution	18 mM Citric Acid, 9 mM Sodium Citrate, 12 mM NaCl
FACS buffer	PBS, 10% FCS
TBS	0.2 M Tris base, 1.5 M NaCl, pH 7.4
TBST	TBS, 0.1 % Tween-20
ISH pre-hybridisation buffer	50 % Ultrapure formamide, 25 % 20x SSC pH 4.5, 50 μg/mL heparin, 0.1 % Tween-20 (in Nuclease-free water)
ISH hybridisation buffer	ISH pre-hybridisation buffer, 100 μg/ml herring sperm DNA, 50 μg/mL Yeast tRNA (in Nuclease-free water)
ISH post-hybridisation wash buffer	50 % Ultrapure formamide, 10 % 20x SSC pH 4.5, 0.1 % Tween-20 (in Nuclease-free water)
ISH alkaline phosphatase buffer	0.1 M NaCl, 0.1 M Tris pH 9.5, 50 mM $MgCl_2$, 0.1 % Tween-20

2.1.3: Plasmids

Name	Antibiotic	Purpose	Reference
pBluescript II SK(+)_Tcf15	Amp	Tcf15 ISH probe	Burgess et al., 1995
AGS-684 (pCMV-GFP)	Amp	Transfection control	From Dr. Andrew Smith
pTcf15_SKO no Poly(A)	Amp	Tcf15 Targeting plasmid missing poly(A)	Made by Aliaksandra Radzisheuskaya
pTcf15_SKO	Amp	For targeting Tcf15 locus	This study
pCAG-mKate2-NLS	Amp	For lineage labelling of cells	From Dr. Guilluame Blin

2.1.4: Primers

2.1.4.1: RT-qPCR primers sequences and UPL probe numbers

Name	Forward	Reverse	UPL probe number
Ccno	TTACGCGTTCACTACCTACACG	GCAGGTCCATCTCGTGCT	27
Cdh1 E-cadherin	ATCCTCGCCCTGCTGATT	ACCACCGTTCTCCTCCGTA	18
Cdh2 N-cadherin	GCCATCATCGCTATCCTTCT	CCGTTTCATCCATACCACAAA	18
Cdh3 P-cadherin	AGGCCCAGCTAACACATGAC	ACAAGGCCACGGTGTCTC	29
Crb2	TCTCTGGCCAGTTCTGTGAA	CTACCTCCAGCAGTGGGAAC	34
DUSP4	GCCTGGCCTACCTGATGAT	GCTGCTTGACGAACTCAAAA	25
DUSP6	GCAGCGACTGGAATGAGAAC	CAGCCCTCGTCTTTGAGTTT	66

EG97	CTAACGTTCTGCACCGTGAC	GATCACAAGTGGTGTTCAGCAG	97
Emp1	TCATCATTACGCCCACAGC	TCCAGGTCAGGATGAAACAA	67
Eomes	ACCGGCACCAAACTGAGA	AAGCTCAAGAAAGGAAACATGC	9
Esrrb	CGATTCATGAAATGCCTCAA	CCTCCTCGAACTCGGTCA	89
Fgf5	AAAACCTGGTGCACCCTAGA	CATCACATTCCCGAATTAAGC	29
FoxA2	TGGCTGCAGACACTTCCTACT	CAACATCAGTACAACCCTCTGGT	32
Gata6	GGTCTCTACAGCAAGATGAATGG	TGGCACAGGACAGTCCAAG	40
GN-TBP	TATCTGCTGGCGGTTTGG	AGGCCTCCAAGCAGAGAAA	104
Hes6	GTTCTGGAGCTGACCGTGA	GATGTAGCCAGCAGCGAAG	32
Hex	CTACACGCACGCCCTACTC	CAGAGGTCGCTGGAGGAA	50
Id1	TCCTGCAGCATGTAATCGAC	GGTCCCGACTTCAGACTCC	78
Id3	CATAGACTACATCCTCGACCTTCA	CACAAGTTCCGGAGTGAGC	53
Irf1	GCACCACTGATCTGTATAACCTACA	CCTCATCCTCGTCTGTTGC	97
Itga3	AGGATATGTGGCTTGGAGTGA	GACCACAGCACCTTGGTGTA	81
Itgb1	ATGCAGGTTGCGGTTTGT	GGGTAAAACAATACCACCAAGTTT	67
Klf2	CTAAAGGCGCATCTGCGTA	TAGTGGCGGGTAAGCTCGT	48
Klf4	CGGGAAGGGAGAAGACACT	GAGTTCCTCACGCCAACG	62
Nanog	CCTCCAGCAGATGCAAGAA	GCTTGCACTTCATCCTTTGG	25

Nanog pre mRNA N26	GGTGATACGTTGGCCTTCTAGT	TTCTCAAATACACACAAGAGCCTTA	SYBG
Nanog pre mRNA N28	AGCCCAGTACTCAGGCTTGT	AGCATCACAACACGCACCT	SYBG
Nanog pre mRNA N29	GCCAGCAGATGGCATAATTT	TGATGGCAATGCTGAGGTTA	SYBG
Nanog pre mRNA N32	GATTCTATTCACCCAGCACCA	CCTTCTGAGTGGAGGTTTATCC	SYBG
Nanog pre mRNA N33	CCATCTCAGCTACTGGAGCA	ATTAGAACCGTGACCGCATC	SYBG
Oct4	GTTGGAGAAGGTGGAACCAA	CTCCTTCTGCAGGGCTTTC	95
Otx2	GACCCGGTACCCAGACATC	GCTCTTCGATTCTTAAACCATACC	103
Pdgfra	GTCGTTGACCTGCAGTGGA	CCAGCATGGTGATACCTTTGT	80
Prr19	CTCCAGATCTTCCTGTGTTGG	TGGCACAATGAGTTGGTACTG	67
Raet1b	TGAAGAGGAAATATTATACCCAAGGA	TTCTTGGAGAGGTCCTGGATT	82
Rex1	CAGCTCCTGCACACAGAAGA	ACTGATCCGCAAACACCTG	16
Scx	ACACCCAGCCCAAACAGAT	TCTGTCACGGTCTTTGCTCA	29
Sox1	GTGACATCTGCCCCCATC	GAGGCCAGTCTGGTGTCAG	60
Sox7	GCGGAGCTCAGCAAGATG	GGGTCTCTTCTGGGACAGTG	97
T-Brachyury	CAGCCCACCTACTGGCTCTA	GAGCCTGGGGTGATGGTA	27
TBP	GGGGAGCTGTGATGTGAAGT	CCAGGAAATAATTCTGGCTCA	97
Tcf15	GTGTAAGGACCGGAGGACAA	GATGGCTAGATGGGTCCTTG	104
Tcf15 transgene	GCTCCATCTGCACCTTCTG	GGCTACACCCCTCACTTTCA	SYBG

Tcf3 (E47)	GTGGGCTCTGACAAGGAACT	ACAGGTAGCGGGAACATCAT	79
Vegfc	CAAGGCTTTTGAAGGCAAAG	GCTCATCTACGCTGGACACA	25
Venus	GAAGCGCGATCACATGGT	CCATGCCGAGAGTGATC	67
Vimentin	CCAACCTTTTCTTCCCTGAA	TGAGTGGGTGTCAACCAGAG	109
Wnt3	GCCAAGAGTGTATTCGCATCTA	GATCCAGCCGCACAATCTAC	48
WT1	CACCAAAGGAGACACACAGG	GGGAAAACTTTCGCTGACAA	47
Zeb1	GCCAGCAGTCATGATGAAAA	TATCACAATACGGGCAGGTG	48
Zeb2	CAAGAGGCGCAAACAAGC	TGCGTCCACTACGTTGTCAT	79
Zfp521	AAGCAAGCGAAACCGAGAT	TTCTGGCCTCTTCTTGCAGT	16
Zyx	CCAGGGAGAAAGTGTGCAGT	GTTCTTGGTCATGTCGTCCA	80

2.1.4.2: Other primers

Name	Sequence	Description
Tcf15_internal_SB_probe_F	AAAGAGGGAGCCCAGAAAGC	Southern blot probe inside the homology arm
Tcf15_internal_SB_probe_R	CCCAGGATGGAGCTATGAACC	Southern blot probe inside the homology arm
Tcf15_3'_SB_Probe_F	GAGCAGGAGTGTCTCGCTCT	Southern blot probe outside the homology arm
Tcf15_3'_SB_Probe_R	GCCTTCCTTCATCTGAGTGC	Southern blot probe outside the homology arm
Venus_SB_Probe_F	CTGCAGTGCTTCGCCCGCTA	Southern blot probe for Venus

Venus_SB_Probe_R	GTCGACCCCCCCGGGTCTAG	Southern blot probe for Venus
CL005_BsRGI-polyA-F	GGACGAGCTGTACAAGTAAGAATTGA	Forward primer for insertion poly(A) into plasmid
CL007_BsrGI-polyA-R2	TCAGTGTACAGTGTTAAGATACATTGATGAGTTTGGAC	Reverse primer for insertion poly(A) into plasmid

2.1.5: siRNAs

siRNAs were reconstituting from lyophilized powder in nuclease-free water according to manufacturer's instruction.

2.1.5.1: siRNA from Qiagen

FlexiTube GeneSolution Tcf15 predesigned siRNA tube 1-4 (GS21407) was obtained from Qiagen.

AllStars Negative Control siRNA conjugated with Alexa-488 (1027292) was used as negative control and for monitoring siRNA transfection efficiency. This RNA had been tested and validated as non-silencing siRNA against any known mammalian gene.

2.1.5.2: siRNA from Dharmacon

Sequences of ON-TARGETplus SMARTpool, Mouse TCF15 were listed as below:

Sequence 1: gaatg tagta tgtga gcaa

Sequence 2: accga aagct gtcta agat

Sequence 3: ggcag ctgct tgaaa gtga

Sequence 4: acgtg aggct gctga gcga

ON-TARGETplus non-targeting negative control siRNA pool contains at least 4 mismatches to any human, mouse, or rat gene.

2.1.6: Antibodies

2.1.6.1: Primary antibodies

Target	Species	Vendor	Dilution	Catalogue number
beta-tubulin	mouse	Sigma	WB 1:4000	T8328
Flag-tag	mouse	Sigma	IHC/WB 1:2000	F9291
Nanog	Rabbit	From Prof. Ian Chambers lab	WB 1:2000	
GFP	Chicken	Molecular probes	IHC 1:1000	A10262
GFP	Rabbit	Molecular probes	IHC/WB 1:1000	A11122
Klf4	Goat	R&D	IHC/WB 1:200	AF3158
Nanog	Rat	eBioscience	IHC 1:200	14-5761-80
Nestin	Rat	DSHB	IHC 1:100	Rat-401
Oct4	Mouse	Santa Cruz	IHC 1:200	Sc-5279
Tcf15	Rabbit	Aviva	WB 1:150	ARP38503_T100
Tuj1(Tubb3)	Mouse	Covence	IHC 1:1500	MMS-435P
PECAM1-APC	Rat	BD	FACS 1:500	551262

2.1.6.1: Secondary antibodies

Alexa Fluor® antibodies (Molecular Probes) conjugated to appropriate fluorophores were used as secondary antibodies for immunostaining experiments. Antibodies were diluted at 1:1000.

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

For detection of primary antibodies in Western blots the following three secondary antibodies were used: ECL Mouse IgG, HRP-linked whole Ab (from sheep) (GE Healthcare, catalogue number NA931-1ML); ECL Rabbit IgG, HRP-linked whole Ab (from donkey) (GE Healthcare, catalogue number NA934-100UL); Donkey Anti-Goat IgG HRP (Promega, catalogue number V8051). All antibodies were used at a 1:4000 dilution.

2.1.7: ES cell lines used in this thesis

Name	Description	Selection	Reference
E14tg2a	Wild type, hypoxanthine phosphoribosyltransferase deficient, mouse embryonic stem cell line 129/Ola background (Clonal)		(Smith and Hooper, 1987)
TTE15	Doxycycline (dox) inducible Flag-Tcf15-E47 expression cell line in AW2 cell background.	puromycin	This study (Davies et al., 2013)
E14tg2a_AW2	E14tg2a derived cell line which contains the coding sequence for the rtTA integrated into the *Rosa 26* (*R26*) and expressed from the *R26* promoter		(Zhou et al., 2013)
216D1 Tcf15-Venus	Gene trapped Tcf15 reporter cell line which a *gtx-IRES-Venus* cassette was incorporated into the *Tcf15* locus	G418	(Tanaka et al., 2008)
Tcf15-Het	*Tcf15* targeted Venus reporter cell line with *Venus* to replace 1st exon in *Tcf15* locus	G418	This study
Tcf15-Het-FRTed	Tcf15-Het with selection cassette being removed		This study
Tcf15-Het-mK2	Tcf15-Het-E6-1 with random integration of *pCAG-mKate-NLS*	puromycin	This study
Tcf15-KO	*Tcf15*-null cell line with *Venus* to replace 1st exon in both *Tcf15* alleles	G418	This study

2.1.8: Cell culture media

Medium	Description
mES cell growth medium	Glasgow Modified Eagle's Medium (GMEM), 10% foetal calf serum, 1 mM sodium pyruvate, 1x non-essential amino acids, 2 mM L-glutamine, 0.1 mM 2-mercapthoethanol and 100 units/mL LIF
LIF withdraw medium	mES cell growth medium without LIF
N2B27 medium	50% Neurobasal, 50% DMEM/F12 (without L-glutamine), 0.1 mM 2-mercapthoethanol, 2 mM L-glutamine, 0.5% modified N2 supplement (made in house) and 1% B27 supplement
Serum-free mES self-renewal medium	N2B27 supplemented with 10 ng/mL BMP4, 100units/mL LIF
2i/LIF medium	N2B27 supplemented with 3 μM CHIR99021, 1 μM PD0325901 and 100 units/ml LIF
EpiSC medium	N2B27 supplemented with 12 ng/mL FGF2 and 20 ng/mL Activin A

2.1.9: Cytokines and inhibitors

Name	Description
Activin A	recombinant human, (R&D)
BMP4	recombinant human (R&D)
Chiron99021	GSK3β inhibitor (Axon Medchem)
Fgf2	FGF2, recombinant human bFGF (R&D)
PD0325901	MEK inhibitor (R&D)
PD173074	FGFR inhibitor (Sigma)
LIF	In house cell culture unit

2.1.10: Cell culture plastics coating

Solution name	concentration	Coating time
Poly-L-ornithine	0.01%	room temperature for 2 hours or 4 degrees overnight
Laminin from EHS murine sarcoma basement membrane	5 μg/mL in PBS	room temperature for 2 hours or 4 degrees overnight
Fibronectin (from bovine plasma)	7.5 μg/mL in PBS	room temperature for 10 minutes
Gelatine	0.1% in PBS	room temperature for 30 minutes

2.1.11: Antibiotics

Name (Manufacturer)	Stock concentration	working concentration
Ampicillin (Calbiochem)	100 mg/ml	100 μg/mL
G418 sulphate (PAA)	200 mg/ml	200 μg/mL
Puromycin (Sigma)	5 mg/ml	12.5 μg/mL
Ganciclovir(Sigma)	10 mg/mL	3 μM
Hygromycin(Sigma)	50 mg/mL	50 μg/mL

2.2: Methods

2.2.1: DNA techniques

2.2.1.1: Plasmid DNA isolation from bacteria

Overnight culture of LB broth, either 5 mL or 100 mL, containing appropriate antibiotics (1:1000 dilution from stock, see materials 2.1.11) was inoculated with a) a single bacterial colony from previous transformed plasmid DNA, or b) glycerol bacterial stock, for mini and midi preparations, correspondingly. After approximately 16 hours culture at 37°C with agitation

at 200 rpm, bacteria were collected by centrifugation at 4000 x g for 10 minutes. Plasmid DNA was isolated using QIAprep™ Spin Miniprep, or QIAfilter Plasmid Midi Kit for mini or midi preparations according to manufacturer's instructions. DNA was eluted in EB buffer for general use or in nuclease-free H_20 for restriction enzyme digestion.

2.2.1.2: DNA isolation from mammalian cells

2.2.1.2.1: Genomic DNA (gDNA) isolation from 24 well plates using home-made lysis buffer

ES cells were grown to confluent in 24 well plates under standard culture condition for gDNA isolation. Culture medium were removed and cells were washed 2 times with PBS and then lysed in 300 μL of gDNA lysis buffer supplemented with fresh added proteinase K to final 0.2 mg/ml. Plates were then sealed completely with tape and incubated at 55°C for overnight in a sealed humidified box to prevent dry-out of lysis buffer. After overnight incubation, cell lysate were transferred to pre-labelled eppendorfs for later procedure. gDNA was precipitated by addition of 300 μL precipitation buffer (nuclease-free isopropanol) directly into tubes followed by vigorous shaking for 15 seconds at room temperature. gDNA was collected by centrifugation at 13,000 rpm for 15 minutes, and excess salt was removed by 2 washing steps with 500 μL of 70% ethanol. DNA pellets were air-dried to remove excess ethanol and resuspended in 50 μL ddH20 at room temperature for 2-3 days. Isolated gDNAs were kept in 4°C.

2.2.1.2.2: gDNA isolation from 6 well plates using commercial kit

For gDNA isolation from wild-type ES cells or smaller sample size, gDNA were isolated using column-binding and spin-elution method. Although this method gives quicker isolation and less protein co-purification, the DNA yield is much lower than using home-made buffer. Therefore, cells must be grown in 6 well plates or larger containers.

ES cells were grown to ~80% confluence in 6 well plates under standard culture condition for gDNA isolation. Cells were trypsinized and cell pellets were collected by centrifugation at 1,300 rpm for 5 minutes. Genomic DNA was isolated using Qiagen DNeasy® Blood & Tissue Kit according to protocol for blood DNA isolation. Genomic DNA was eluted twice in 200 μL elution buffer in order to increase DNA yield.

2.2.1.3: DNA cloning

2.2.1.3.1: Restriction enzyme digestion

For general cloning experiment, 2-5 μg plasmids DNA or PCR product were digested using relevant enzymes in the buffers provided by the manufacturer in 50 μL reaction volume. For validation of ligation reaction using restriction enzyme digestion to check the presence of insert in the extracted plasmids, 100 to 200 ng of DNA were digested in a 10-15 μl reaction. Digestion time was typically 3 hours at a temperature appropriate for the enzyme being used to prevent star activity.

For plasmid linearization used for transfection by electroporation, 100 μg of plasmid was digested in 500 μL reaction volume and incubated overnight. Digested DNA fragment were isolated using gel extraction

method (see section **2.2.1.3.5**) or direct precipitation (see section **2.2.1.3.6**).

2.2.1.3.2: Dephosphorylation of DNA fragment ends

In order to reduce the chance of self-ligation, 5′ phosphates groups of DNA ends from the digested vector DNA were removed by incubating 1-2 μg of digested DNA with 1 unit of alkaline phosphatase according to manufacturer's instructions. Reaction was performed at 37°C for 15 minutes and irreversibly inactivated by heat treatment for 2 minutes at 75°C.

2.2.1.3.3: DNA fragment ligation

DNA fragment with cohesive ends generated by restriction enzyme digestion and follow-up purification were ligated using T4 DNA ligase. Generally, 50 ng of vector DNA was used for a ligation reaction. Molar ratio between vector to insert at 1:3 was used according to manufacturer's instruction. The reaction was performed at room temperature for 2 hours or at 4°C overnight.

2.2.1.3.4: DNA electrophoresis

DNA was separated and analysed by running 0.8 – 2 % (w/v) agarose gel electrophoresis depending on the size of DNA to be resolved. 0.5x TBE buffer were used for making agarose gel and as running buffer. Samples are mixed with 6x DNA loading dye and running at 100 volt for 40 minutes alongside with a 100 bp or 1kb DNA marker for size comparison. Gels are stained with SYBR® Safe dye for visualization by UV illumination.

2.2.1.3.5: DNA extraction from agarose gels

Agarose gel containing desired DNA fragment was examined under UV illumination. Gel slice was excised by clean knife with no longer than 3 seconds of UV exposure to prevent UV-induced mutation. DNA was extracted from gel slices using QIAquick gel extraction kit according to manufacturer's instruction and eluted with 30 μL of elution buffer.

2.2.1.3.6: DNA ethanol precipitation

DNA was precipitation by adding 1/10 volume of 3 M Sodium acetate (NaOAc, pH 5.2) and 2.4 volume of 100 % ethanol into sample then mix thoroughly. This mixture was incubated in -20 °C for ≧ 20 minutes to increase DNA yield. Precipitated DNA was recovered by centrifugation at 13,000 rpm for 15 minutes. Excess salt was removed by washing the DNA pellet twice with ice cold 70% ethanol. DNA was then air-dried and re-suspended in an appropriate volume of nuclease-free H_2O depending on the size of pellet.

2.2.1.3.7: DNA/RNA quantification

Concentration of DNA was quantified by UV spectrometry using a NanoDrop® ND-1000 spectrophotometer or NanoVue Plus spectrophotometer. The Beer-Lambert's law was used to calculate the concentration, with extinction coefficients of 50 for dsDNA, 33 for ssDNA and 40 for RNA.

2.2.1.3.8: Bacterial transformation

For transformations using chemically competent DH5α *E. Coli* cells,

either 5 μL of ligation mix or less than 50 ng of plasmid DNA was used for setup the transformation. DNA was added to 50 μL of freshly thawed competent *E. Coli* and incubated on ice for 5 minutes. Bacteria were heat shocked for 45 seconds at 42°C and immediately placed on ice. 500 μL of LB medium without antibiotics was added to the tube and the culture was allowed to recover at 37°C for ~40 minutes at 250rpm shaker. For plasmid transformation, 20-100 μL of the culture was plated on LB plates with appropriate antibiotics. For ligation product transformation, transformed samples were centrifuged at 4,000 rpm to pellet down the bacteria and re-suspend in 200 μL of LB medium. 100 μL of culture were plated on LB plates with appropriate antibiotics and incubated overnight at 37°C.

For TA ligation product (see section **2.2.1.4.4**), bacteria were plated on LB plate pre-absorb with 100 μL of 100mM IPTG and 20 μL of 50 mg/ml X-Gal over the surface for 30 minutes at 37°C prior to use.

2.2.1.4: Polymerase chain reaction (PCR)

2.2.1.4.1: Generic PCR

All PCR reaction was performed on TProfessional Standard Thermocycler (Biometra). PCR was performed according to manufacturer protocol. A negative control containing no DNA template was used for all reactions.

2.2.1.4.1.1: PCR using Taq

Taq DNA polymerase was used to generate PCR product with 3'-A tail, and when high proofreading activity was not essential, such as colony PCR.

PCR reactions were generally prepared in 50 μL volumes, with 200 μM dNTP. The DNA polymerase was used at 2.5 units/ reaction, primers were used at 0.5 μM and reactions were made-up with the reaction buffer supplied with the polymerase. Depend on the template be used, generally less than 1 ng of template DNA was used for plasmid DNA template and 100-200 ng of DNA was used for cDNA template.

PCR reaction was setup as below, primer annealing temperature were usually between 55°C and 60°C but optimized for each primer pair.

Initial denaturation step		94°C	1 minute
	Denaturation	94°C	30 seconds
30 cycles of	Annealing	45-65°C	30 seconds
	Extension	72°C	1 minutes per kb
Final incubation		72°C	5 minutes
Cool down step		4°C	pause

2.2.1.4.1.1.1: Bacteria colony PCR

Colony PCR was used for quick PCR check of transformed bacteria single colonies in order to determine if ligation is working. Colony PCR provided high-throughput screening of colonies when ligation efficiency is low. In addition, when single restriction enzyme site were used for ligation, colony PCR could be used to check the orientation of insert before plasmid extraction.

Taq PCR master mix was aliquot into 10 μL per reaction. Bacteria single colony was picked and streaked to a new LB plate to create stock. Pick containing the bacteria was then immersed into PCR master mix for 1 minutes and discard afterward. PCR was then performed according to

designated primers.

2.2.1.4.1.2: PCR using Pfx

Pfx Polymerase was used to generate blunt-end DNA fragment for further cloning experiments.

PCR was performed according to protocol provided by the manufacturer. PCR reactions were generally prepared in 50 µL volumes, with 300 µM dNTP. Magnesium sulfate was used at final concentrations of 1 mM. The DNA polymerase was used at 1 units/ reaction, primers were used at 0.3 µM and reactions were made-up with the reaction buffer supplied with the polymerase. Amount of template DNA to be used was depended on type of template as described above in Taq PCR.

PCR reaction was setup as below,

Initial denaturation step		94°C	5 minutes
	Denaturation	94°C	15 seconds
30 cycles of	Annealing	45-65°C	30 seconds
	Extension	68°C	1 minute per kb
Final incubation		68°C	5 minutes
Cool down step		4°C	pause

2.2.1.4.1.3: Long range gDNA PCR using LongAMP

The LongAmp Taq PCR kit (NEB) was used to perform long range PCR reaction on genomic DNA.

PCR was performed according to manufacturer protocol. PCR reactions were generally prepared in 20 µL volumes, with 300 µM dNTP. The DNA polymerase was used at 1 units/ 10 µL reaction, primers were used at 0.4 µM and reactions were made-up with the reaction buffer supplied with the

polymerase. Amount of template DNA to be used was depended on type of template as described above in Taq PCR.

PCR reaction was setup as below,

Initial denaturation step		94°C	30 seconds
30 cycles of	Denaturation	94°C	15 seconds
	Annealing	45-65°C	30 seconds
	Extension	65°C	50 seconds per kb
Final incubation		65°C	10 minutes
Cool down step		4°C	pause

2.2.1.4.2: PCR product purification

When cDNA was used for PCR template and there is only single PCR product after gel electrophoresis, PCR product was purified by QIAquick PCR purification kit according to the manufacturer's instructions. When plasmid DNA was used for PCR template, then PCR product must be gel-extracted to eliminate template plasmid contamination by using QIAquick gel extraction kit.

2.2.1.4.3: A-tailing of blunt-end PCR products

PCR products generated by Pfx polymerase with blunt-end were A-tailed by incubating with Taq polymerase. Purified PCR fragment was mixed with reaction buffer, 0.2 mM dATP and 5 units of Taq DNA polymerase in a total 10 μL volume. A-tailing reaction was performed at 70°C for 30 minutes.

2.2.1.4.4: Direct cloning of PCR products

PCR products with blunt-end were directly cloned using the Invitrogen

Zero Blunt® TOPO® PCR Cloning Kit according to the manufacturer's instructions. PCR products with A-tail were TA-cloned into pGEM-T easy vectors according to the manufacturer's instructions.

2.2.1.4.5: Real-time quantitative PCR (RT-qPCR)

RT-qPCR was performed using the Roche LightCycler® 480 Instrument II in 384 well format. First strand cDNA was used as template. Primers were designed using the Roche Assay Design Center website (http://www.roche-applied-science.com/shop/CategoryDisplay?catalogId=10001&tab=&identifier=Universal+Probe+Library#tab-3) and each primer pair had a specific UPL probe. Reactions were performed using the LightCycler® 480 probes master mix and mouse UPL set (Roche) as outlined by the manufacturer. Primers used for qPCR were listed in **2.1.4.1**. Total reaction volume was 6 µL containing 3 µL of master mix, 0.9 µL of nuclease-free H_2O, 0.54 µL of 5 µM primer mix and 0.06 µL of specific probe. cDNA samples were run as triplicate. Relative gene expression was calculated by normalization against the house-keeping gene TBP or GAPDH. Relative concentrations were calculated from a standard curve from serial diluted plasmids. All experiments have been performed at least three times unless otherwise stated. Statistic analyses were performed using a paired, two tailed Student's T-test and p values below 0.05 were considered as significant. All data are represented as mean ±SD.

2.2.1.5: DNA sequencing

Plasmid DNA was sequenced using the BigDye Terminator Cycle

Sequencing Ready Reaction kit at the Genepool sequencing service center (University of Edinburgh). Reaction samples were submitted in 6 μL volumes containing 500 ng of DNA template with 1 μL of a 3.2 pmole/μL primer solution. Sequencing primers were selected according to different sample demand. Sequencing results were compared by using Multalin website (http://www.sacs.ucsf.edu/cgi-bin/multalin.py).

2.2.1.6: Southern blot analysis

Restriction fragment length polymorphisms were determined by Southern blot analysis (Southern, 1975).

2.2.1.6.1: Sample preparation

Genomic DNA was isolated as outlined in section **2.2.1.2**. 3-5 μg of gDNA was digested in 25 μL volumes with 40 units of restriction enzyme. For Southern blot of *Tcf15* targeting cells, gDNA are digested with high concentration *Eco*RV (Roche) at 37 °C for overnight in an air incubator.

If gDNA concentration is too low to be digested within 25 μL, increase the reaction volume without changing the amount of enzyme and perform ethanol precipitation (**2.2.1.3.6**) of gDNA after digestion.

2.2.1.6.2: DNA electrophoresis and denaturation

Digested gDNA fragments were separated by 0.8% agarose TAE gel electrophoresis alongside a λ DNA-*Hind*III digest size marker for overnight at 20-40 volt in 1x TAE buffer. Gel was stained with SyBR-Safe gel stain (1:20,000) for 20 minutes at room temperature on a 40 rpm shaker protected

from light. The gel was then transferred to a UV trans-illuminator for photographic imaging. The efficiency of digestion was determined by the extent of DNA laddering showing by smear. The gel was then incubated in denaturing solution for 2 x 20 minutes with gentle rocking. After rinsing with ddH_20, the gel was then washed with neutralization solution for 2 x 20 minutes and ready for transfer.

2.2.1.6.3: Sample transfer

DNA was transferred to a nitrocellulose membrane by capillary action (Figure 2.1) for 48 hours. After transfer, the membrane was removed from the transfer apparatus, labelled with well position and date of experiment by pencil and rinsed in 2x SSC buffer. DNA was cross-linked to the membrane by baking at 120°C for 1 hour.

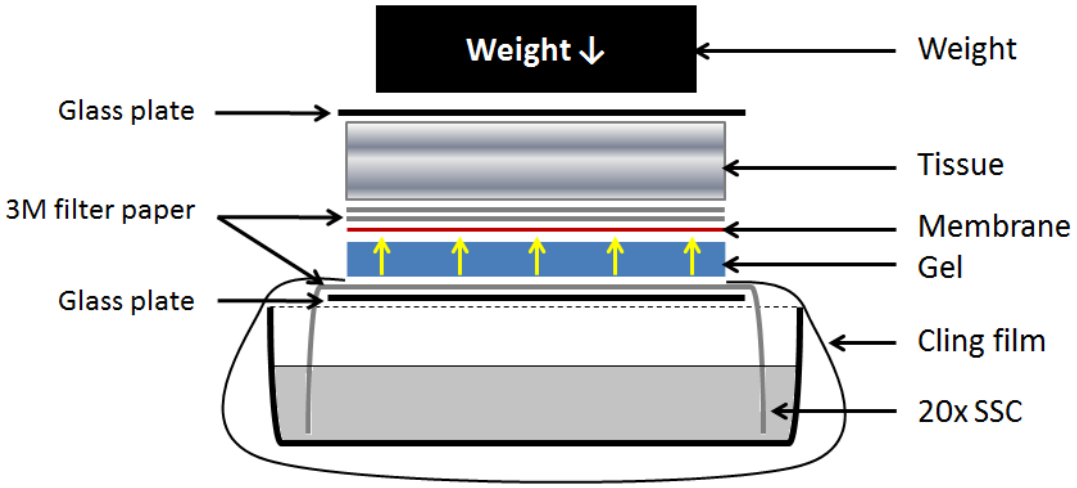


Figure 2.1: Diagram of the Southern blot transfer apparatus

The transfer apparatus was assembled as above. A glass plate was suspended over a buffer tray where a paper wick was setup to draw buffer from the tray. The bottom 3M filter paper was pre-soak with transfer buffer which is 20x SSC. The gel was placed up-side-down on the paper wick and all bubbles need to be removed under the gel. The membrane was pre-soaked in 2x SSC buffer and placed between the gel and a layer of blotting paper. A plastic pipette was used to roll over the membrane and gel to remove air bubbles. A stack of tissue paper was placed over the blotting paper and secured with a glass plate and weight. The bottom of the transfer apparatus was wrapped with cling film to prevent dry out of 20x SSC outside the gel.

2.2.1.6.4: Probe preparation and hybridization

After baking, the membrane could be stored at 4°C or continued to probe hybridization process. Membrane was rinsed with 2x SSC and placed in a glass hybridization bottle. Either 10ml (small membrane, no exceed 12x15 cm) or 20 mL (large membrane) Perfect-Hyb solution (Sigma) containing 10 μg/mL herring sperm DNA were added and incubated in a rolling oven at 65°C for a minimum of 2 hours. During this time the probe was labelled. Probe templates were generally between 200-700 bp in length. The probe template was amplified by either PCR purification or by restriction enzyme digestion from plasmid DNA containing the desired genomic sequence from a wild-type mouse genomic library. λ DNA-*Hind*III digest probe was provided by the labelling kit. The probe was labelled using the Amersham MegaPrime™ DNA Labelling System according to manufacturer's instructions, using [α-^{32}P] dCTP, 3000 Ci/mmol. After probe synthesis, radio-active probes were purified twice by Amersham Illustra ProbeQuant™ G-50 Micro Columns to remove unbound primer and nucleotide. After boiling at 100°C for 5 minutes, the probes (sample and λ DNA-*Hind*III) were then added to the hybridization tube and incubated with the membrane overnight at 65-68°C, according to different probes.

2.2.1.6.5: Membrane washing and exposure

The radio-active probes solution was disposed and the membrane was washed twice in 2x SSC wash buffer at 65°C for 20 minutes. After washes, membrane was examined by Geiger counter. If the signal counts were between 5 and 20, then the membrane is ready to be exposure. If the signal

counts were over 20, then continue the washing steps by using 0.5x SSC wash buffer at 65°C for 20 minutes. The membrane was then wrapped in cling film and exposed to X-ray film (Kodak) at -80°C in a cassette with intensifying screen. For signal counts between 5 and 10, membranes need to be exposed for at least 3 days. For signal counts between 10 and 20, overnight to 2 days of exposure is enough. After exposure, the film was removed and developed in an automated auto-radiographic film processor.

2.2.1.6.6: Stripping of Southern blot membranes

Membranes could be stripped for 2 to 3 times and re-hybridize with other probes. To perform stripping, membranes were incubated with stripping buffer (see **2.1.2**) in hybridization bottles at 37°C for 30 minutes. After stripping, membranes were examined using Geiger counter to make sure no radio-active signal was remained. The stripped membrane was then wrapped in cling film and exposed to X-ray film (Kodak) at -80°C for overnight. The film was developed to make sure no radio-active signal was left on the membrane. The membrane was then ready for hybridization of new probes.

2.2.1.7: DNA sequence analysis

All DNA sequence analysis for labelling genomic markers, identifying restriction enzyme sites, primer design and plasmid map generation were performed using: ApE- A plasmid Editor v2.0.36, copyright © 2003-2009 M. Wayne Davies. Cloning primer sequence was designed and optimized using online Primer3 software (http://bioinfo.ut.ee/primer3-0.4.0/).

2.2.1.8: Generation of specific constructs and plasmids

2.2.1.8.1: Adding poly(A) tail to pTcf15_SKO plasmid

pTcf15_SKO was constructs previously made by Aliaksandra Radzisheuskaya for targeting *Tcf15* locus in mES cells in the lab. pTcf15_SKO construct was designed for introducing *Venus* to replace first intron of *Tcf15*. This targeting vector contained homology arms with *DTA* as negative selection marker outside the homology arms. The *Venus* gene was placed at 5′ of selection cassette included a *PGK_Neomycin* resistant gene and a positive selection marker *MC1_Tk* flanking by *FRT* site. However, after finish generating of these constructs, DNA sequencing result showed that the *poly(A)* tail of *Venus* gene was missing. Therefore, *poly(A)* was re-cloned into the final targeting vector using a restriction enzyme site *Bsr*GI at 3′ end of *Venus* sequence. Briefly, *poly(A)* sequence was PCR amplified using primer pairs with *Bsr*GI site on each side. The PCR product was sub-cloned into pGEM-T-Easy vector as described in section **2.2.1.3**. Sequence varied clone was picked, expanded and digested with *Bsr*GI and gel purified. The pTcf15_SKO plasmids was also digested with *Bsr*GI, dephosphorylated and gel purified. Insert were then ligated with vector and the orientation of insert was verified by colony PCR. Clones with correct orientation of insert were then sent to perform sequence analysis to confirm the *poly(A)* sequence was inserted to the end of the *Venus* gene. The orientation of final targeting construct is shown in figure 2.2.

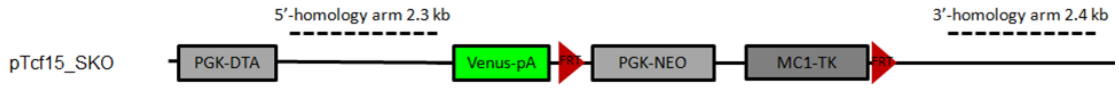


Figure 2.2: Genomic orientation of pTcf15_SKO construct with poly(A) sequence inserted to the end of *Venus* gene

The final construct of pTcf15_SKO for targeting the *Tcf15* locus contains homology arms with *DTA* as negative selection marker outside the homology arms. The *Venus* gene was placed at 5′ of selection cassette included a *PGK_neomycin* resistant gene and a positive selection marker *MC1_Tk* flanking by *FRT* site.

2.2.1.8.2: Generation of plasmids for Southern blot probes

In order to perform southern blot on *Tcf15* targeting cells. Tcf15-3′-probe, Tcf15-internal probe and Venus probe were PCR amplified from wild-type gDNA using the primers as listed in section **2.1.4.2**. PCR products were ligated to pGEM-T-Easy vector. Sequence verified clones were amplified and plasmids were isolated using QIAfilter Plasmid Midi Kit (Qiagen). Plasmids were digested with *Eco*RI to release inserts and probe template DNAs were purified using gel extraction (see section **2.2.1.3.5**). Probe template DNAs were then diluted to 25 ng/μL for probe synthesis.

Tcf15-3′-probe located at 3′ of targeted region and Tcf15-internal probe located inside the homology arms. Tcf15-3′ probe and Venus probes were designed to identify correctly targeted clones and Tcf15-internal probes was designed for detecting the existence of multiple integrations of targeting construct.

2.2.2: RNA methods

For routine RNA work, lab benches were wiped with RNaseZap® RNase decontamination solution before starting experiments. Solutions were made with RNase/DNase free water (Gibco or Hyclone).

2.2.2.1: Total RNA extraction

Total RNA was extracted using either Absolutely RNA Purification Kits (Agilent) or Trizol Reagent (Invitrogen) according to manufacturer's instructions.

2.2.2.1.1: RNA isolation using Absolutely RNA Purification Kits

For RNA isolation using Absolutely RNA Purification Kits, cells were lysed on plate in RNA lysis buffer containing 2-mercaptoethanol. 350 μL lysis buffer with 2.5 μL 2-mercaptoethanol were used for 6 well plates and volume was doubled for 10 cm dishes. Lysates were transferred to pre-filter columns and centrifuged to homogenise. Genomic DNA was removed by on-column purification and the flow-through was loaded onto RNA binding columns for RNA purification, on-column DNase I digestion and washing. RNA was eluted in 30μL of buffer EB.

2.2.2.1.2: RNA isolation using Trizol

For cells isolated by Fluorescence-activated cell sorting (FACS) with concentration less than $5x10^5$ cells/sample, RNA was extracted manually using Trizol reagent according to manufacturer protocol. Cells were pellet down by centrifuged at 2,000 rpm for 5 minutes and re-suspended in 1 ml Trizol. 200 μL of chloroform was added to each sample and vigorous mixing to achieve a homogenous solution. After 5 minutes of incubation at room temperature, the sample was centrifuged at 12,000 x g for 15 minutes at 4°C. Upper aqueous phase was moved carefully to a new tube and 650 μL of nuclease-free isopropanol were added and mixed. The mixture was incubated at -20°C for at least 2 hours with 5 μg/mL carriers DNA to increase RNA yield. RNA was precipitated by centrifugation at 12,000 x g for 15 minutes at 4°C. RNA pellet was then washed twice with ice-cold 70% ethanol to remove excess salt and isopropanol. Pellet was air-dried and re-suspend in nuclease water.

2.2.2.2: First strand cDNA synthesis

First strand cDNA synthesis was performed on TProfessional Standard Thermocycler (Biometra) using 500 ng to 1 μg of total RNA as template. M-MLV reverse transcriptase was used according to manufacturer's instructions. Reverse transcription reaction was primed using random hexamers in the presence of RiboLock RNase Inhibitor (Thermo) in a reaction volume of 20 μL. Following first strand synthesis, the reaction mixture was diluted 1:6 in nuclease-free H_2O, and 1.5 μL was used per qPCR reaction.

No-RT control was performed for each RNA sample by not adding the Reverse Transcriptase to the reaction mixture. This is to detect if there is any significant levels of genomic DNA contamination in the RNA samples by performing qPCR (see section **2.2.1.4.5**) using a non-intron spanning pair of primers (e.g. Tcf15). The samples were only used if the transcript levels were at least 10^3 times higher in the RT samples compared with No-RT controls.

2.2.2.3: Preparing RNA samples for microarray analysis

Illumina TotalPrep™ RNA amplification kit was used for generating cRNA from RNA samples for microarray analysis of gene expression according to manufacturer's instruction. RNA samples were analysed using normal cDNA synthesis followed by qPCR analysis to validate control gene expression and the quality of samples before starting cRNA synthesis procedure.

2.2.2.3.1: First and second strand cDNA synthesis and purification

First strand cDNA synthesis was carried out by preparing 100 ng of RNA sample with 1x first strand buffer, T7 oligo(dT) primer, dNTP mix,

RNase inhibitor and ArrayScript reverse transcriptase in 20 μL nuclease-free H_2O in nuclease-free and sterile PCR tubes. The reaction mix was then placed into PCR machine for 2 hours at 42°C with lid temperature at 50°C according to manufacturer's instruction.

After first strand cDNA synthesis, the lid of the PCR machine was set to 16°C or cooled to room temperature before starting second strand cDNA synthesis. 80 μL of second strand cDNA synthesis master mix consisting with 1x second strand buffer, dNTP mix, DNA polymerase and RNase H was added into samples and mixed well on ice. The reaction was performed at 16°C for 2 hours.

Double-stranded cDNA was then purified by Illumina TotalPrep™ RNA amplification kit and eluted in 20 μL of 55°C nuclease-free water.

2.2.2.3.2: cRNA synthesis and purification

Double-stranded cDNA was *in vitro* transcribed to synthesize cRNA according to manufacturer's instruction. cDNA samples were mixed with 1x reaction buffer, T7 Enzyme mix and biotin-NTP mix for synthesize biotin-labelled cRNAs. The reaction was performed in 37°C for 14 hours with the lid of PCR machine set to 100°C.

cRNA was purified twice after synthesize, the first purification was to removes enzymes, salts and unincorporated nucleotides and the second purification was to concentrate the cRNA for sending to perform microarray analysis. For the first purification, Illumina TotalPrep™ RNA amplification kit was used and cRNA was eluted with 200 μL 55°C nuclease-free water. For the second purification, RNeasy™ micro kit (Qiagen) was used by

following the protocol for RNA clean-up and concentration. The cRNA was eluted in 14 μL of nuclease-free water with total concentration around 2-3 μg.

2.2.2.4: Synthesis of digoxigenin-labelled RNA probe for *in situ* hybridization (ISH)

2.2.2.4.1: Plasmid linearization

Plasmid vectors used as *in vitro* transcription templates should be linearized by restriction enzyme digestion in order to provide defined length of RNA transcripts. For *Tcf15* ISH anti-sense probe, pBluescript II SK(+)_Tcf15 plasmid (see section **2.1.3**) was linearized using *Eco*RI site. Linearized plasmid was ethanol precipitated (see section **2.2.1.3.6**) and validated by gel electrophoresis.

2.2.2.4.2: Synthesis of Digoxigenin- labelled RNA probes

The *in vitro* synthesis of digoxigenin (DIG)-labelled RNA probe was performed by using 1 μg of linearized plasmid, 1x transcription buffer, 1x DIG-UTP nucleotide mix, 150 units of RNase inhibitor and 20 units of T7 RNA polymerase for *Tcf15* probe. The reaction was in 20 μL volume in RNase-free PCR strips. After incubation at 37°C in PCR machine for 2 hours, 20 units of RNase-free DNase I (Promega) was added for another 30 minutes of incubation. The reaction was stopped by adding 1x stopping buffer (Promega) and heat-inactivated the enzyme at 65°C for 10 minutes. Transcribed RNA probe was ethanol precipitated, dissolved in RNase-free H_2O and validated by gel electrophoresis. RNA probes were aliquoted into 1 μg/tube and stored at -80°C for long term preservation.

2.2.3: Protein methods

All protein work was performed on ice to prevent protein degradation. Protease and phosphatase inhibitors were added fresh to all buffers. Dithiothreitol (DTT) was also made and added freshly in to all the buffers.

2.2.3.1: Total protein isolation from mammalian cells using RIPA buffer

Cells were trypsinized, counted and collected in eppendorfs. In general, 60 μL of RIPA lysis buffer were used to lyse 10^6 cells. After mixing cell pellets with RIPA buffer, samples were vortexed for 15 seconds and incubated on ice for 10 minutes. This step was repeated for 4 times. 1 μL of Benzonase nuclease was added and incubated at room temperature for 10 minutes before samples were centrifuged at 13,000 rpm for 15 minutes at 4°C. 10-20 μL lysate (or 30-50 μg of total proteins) was used for Western blot analysis.

2.2.3.2: Protein quantification

Protein concentration was quantified by either traditional Bradford assay or using Quant-iT ™ Protein Assay Kit (Invitrogen).

2.2.3.2.1: Protein quantification using Bradford assay

Protein concentration was quantified by Bradford assay according to reagent guidelines (Bradford, 1976). Briefly, cell lysates were diluted at least 1:10 in ddH_2O to reduce reading background generated by detergents in the lysis buffer. Protein standards ranging from 0.05-2 mg/mL of BSA (0.05 mg, 0.1 mg, 0.25 mg, 0.5 mg, 1 mg, and 2 mg) were prepared in ddH_2O as well. Protein assay dye (Bio-Rad) was diluted 1:5 in ddH_2O and 200 μL/well of

diluted dye was added into flat bottom 96 well plates. 2.5 µL of diluted samples or BSA standards was mixed with assay dye together in triplicate format. Controls including blank well, 10x diluted lysis buffer, and ddH_2O were also included in the plate for comparison. Samples were incubated at room temperature for approximately 10 minutes. The absorbance was read using Glomax Multi Detection System (Promega) at 595 nm and sample concentrations were calculated from the standard curve derived from protein standard measurements.

2.2.3.2.2: Protein quantification using Quant-iT ™ Protein Assay Kit

Where possible, protein concentration was determined using the Quant-iT™ protein assay kit with the Qubit™ fluorometer (Invitrogen) according to manufacturer's instruction. In short, 1 – 20 µL of 10x diluted protein samples and protein standard #1 to #3 were mixed with Quant-iT™ working solution to make 200 µL reaction mixtures in assay tubes. Samples were vortex for 3 seconds and incubated at room temperature for 15 minutes. After calibration of the Qubit™ fluorometer using Standard #1, #2 and #3, samples were read on the fluorometer with excitation and emission filters appropriate for fluorescein or Alexa Fluor 488 dyes. Sample concentration was calculated automatically by the Qubit™ fluorometer.

2.2.3.3: Western blot analysis

2.2.3.3.1: Protein separation by SDS-PAGE

SDS-PAGE electrophoresis was performed using the NuPAGE system (Invitrogen) with 4-12% Bis-Tris gradient gels. Samples with similar amount

of protein concentration were adjusted to same volume (< 20 μL) using RIPA buffer in order to obtain equal salt concentration. Samples were then mixed with 4x LDS loading dye with final 100 mM of DTT and boiled for 5 minutes to obtain denatured and reduced proteins. The electrophoresis apparatus was assembled as outlined by the manufacturer. Gel wells were washed free from storage buffer using 1x NuPAGE MOPS running buffer. Samples were loaded alongside with 8 μL of SeeBlue™ pre-stained protein marker (Invitrogen) for size estimation.

Gels were run at 100 volt for 15 minutes to obtain initial separation followed by 150 volt until the 3 kDa band of the protein marker reached the end of the gel (typically within 1.5 hours). The apparatus was then disassembled. Gel was removed from the casing, washed briefly with ddH_20 and bottom part of gel was cut for preparing for transfer.

2.2.3.3.2: Protein transfer to nitrocellulose membrane

Hybond™ ECL nitrocellulose membrane was used for protein transfer. The transfer apparatus was assembled using the Novex MultiCell gel tank. In short, gel was stacked within 2 pieces of 3M filter paper and sponge at each side, with nitrocellulose membrane on the anode side to capture the anion-containing protein molecules. All blotting papers and sponges must be well soaked in transfer buffer and air bubbles should be avoided. The proteins were transferred to the membrane at 200 mA for 60-90 minutes in transfer buffer depending on desired molecular weight of target protein.

Protein transfer efficiency was determined by staining the membrane with Ponceau S solution. Membrane was incubated in Ponceau S for 2

minutes at room temperature followed by gradually washing with ddH_2O to reveal protein bands.

2.2.3.3.3: Membrane blocking, probing and washing

Membranes were blocked with 5% non-fat dry milk in PBST at room temperature for at least 1 hour or 4°C overnight. After blocking, if necessary, membrane was crosscut into strips in order to probe different primary antibodies from the same samples. Primary antibodies incubation were performed as advised by the manufacturer (see section **2.1.6**), generally overnight at 4°C in blocking buffer. After primary antibody incubations, the membrane was washed with PBST for 15 minutes and repeated for 4 times at room temperature on a shaker. Appropriate secondary HRP-conjugated antibody (Promega or GE) were added at 1:4000 in blocking buffer at room temperature for 1 hour. The membrane was then again washed 4 times for 15 minutes each with PBST at room temperature.

2.2.3.3.4: Signal detection

After washing off secondary antibodies, chemoluminescent signal was detected by using Super Signal West Pico Chemiluminescent Substrate (Pierce). Equal amount of each solution was mixed for 1 minute and applied to membranes for 1 minute. The membrane was then wrapped in cling film and placed into a cassette, ready for exposure. Amersham Hyperfilm ECL was used for developing the signal. The developing time is ranging from 3 seconds to 10 minutes depending on the strength of the antibody signal. Films were developed using automatic film processor.

2.2.4: Tissue culture

All cells were grown at 37°C with 5% CO_2 supply in a humidified incubator. All solutions were either prepared sterile in the SCRM tissue culture facility or obtained directly from manufacturer. All cell manipulations were performed in a sterile laminar flow hood. Media and buffers were stored at 4°C and not used for longer than 2 months. All cells were kept in penicillin/streptomycin-free media except cells after FACS. Culture plastics were coated by different substrates with concentrations and duration as listed in section **2.1.10**. All newly derived lines and clonal lines were karyotyped before used in experiments.

2.2.4.1: Routine culture of pluripotent mouse cells

2.2.4.1.1: Routine culture of mouse embryonic stem cells

Mouse embryonic stem cells were routine cultured on gelatine-coated plastics until 80% confluence in complete mES cell growth medium (see section **2.1.8**). Culture medium was aspirated off and the cell monolayer was rinsed with room temperature Dulbeco's PBS (Gibco). 0.125 % (w/ v) trypsin was used for passage mES cells. In general, 1 mL of trypsin was added per T75 flask, and scaled accordingly for other sized flasks. The flask was incubated at 37°C and then tapped to disassociate the cells. 9 mL of complete culture medium was then added to quench trypsin. Cells suspension was then centrifuged at 1,300 rpm for 3 minutes. Cells were split 1:3 for being passage every second day or 1:5 for every third day. In the case when precise cell number was required, cells were counted using haemocytometer and plated to desired containers. Culture medium is changed every second day.

Foe serum-free pluripotent culture, cells were cultured in N2B27 containing LIF and BMP4 as described in section **2.1.8** with same passaging method as above except N2B27 was used to quench trypsin.

2.2.4.1.2: Routine culture of mES cells at ground state (2i)

For culturing mES cells at ground state, cells were maintained in serum-free N2B27 medium supplied with PD03 and Chiron (see section **2.1.8** for 2i medium) on poly-ornithine and laminin-coated plastics. Accutase (Sigma) was used to dissociate 2i cultured cells. Cells were passaged around 40% confluence every 3-4 days as described in section **2.2.4.1.1**, except N2B27 was used to quench accutase. Culture medium is changed every second day.

2.2.4.1.3: Routine culture of mouse Epiblast stem cells (mEpiSCs)

Mouse Epiblast stem cells (mEpiSCs) were cultured in EpiSC medium as described in section **2.1.8** on fibronectin-coated plastics. mEpiSCs were generally maintained in 6 well plate. Cells were initially plated as 10^5 cells per well and passaged every two days. To passage, cells were briefly washed with non-supplemented N2B27 and detached using accutase. EpiSCs were rinsed with N2B27 gently in order to maintain the cells in small clumps rather than dispersing them into single cells. After centrifugation at 1,300 rpm for 3 minutes, EpiSCs were passaged at 1:4 before passage 7, and 1:10 after passage 7 which was determined empirically.

2.2.4.2: Transfection of mES cells

2.2.4.2.1: Transfection by liposomes

2.2.4.2.1.1: Transfection of plasmid DNA using Lipofectamine 2000

Transient or sometimes stable transfection was performed using Lipofectamine 2000 (Invitrogen) according manufacturer's instruction. In general, 10^5 cells were plated per well in a 6 well plate and cultured overnight in complete media without any antibiotics. 3 μg of plasmid DNA was mixed with 3 μL of transfection reagent in 200 μL of GMEM basal medium (Sigma) and incubated at room temperature for 15 minutes. For transfection controls, DNA was replaced with either water or with a plasmid expressing GFP (AGS-684, see section **2.1.3**) for monitoring transfection efficiency. The liposome-containing medium was then added in drop-wise manner to the culture cells and cells were returned to incubator for culture overnight. The medium was changed to complete mES cell media on the next day and cells were passaged for downstream applications if necessary.

2.2.4.2.1.2: Transfection of siRNA using HiPerfect

Tcf15 siRNA 1-4 and Alexa488-conjugated non-silencing siRNA control were obtained from Qiagen (see **2.1.5.1**). HiPerfect transfection reagent was used to transfect Qiagen siRNA into mES cells according to manufacturer's instruction. Briefly, 1.5 x 10^5 mES cells were plated per well in a 12 well plate in mES cell growth medium 5 hours prior of transfection. 3 μL of 2 μM siRNA hairpins (75 ng) were mixed with 6 μL of HiPerfect transfection reagent in serum-free medium to obtain final 10 nM of siRNA concentration. The mixture was incubated at room temperature for 15 minutes before

adding drop-wise onto the cells. Cells were returned to incubator for culture overnight. The medium was changed to complete mES cell medium on the next day of transfection and cells were collected at different time point after transfection.

2.2.4.2.1.3: Transfection of siRNA using Dharmafect

DharmaFECT transfection reagent (Darmacon) was used to deliver SMARTpool ON-TARGET plus Tcf15 siRNA and non-targeting control siRNA (see section **2.1.5.2**) into cultured mES cells according to manufacturer's instruction. mES cells were plated in 12 well plate at 3 x 10^4 cells per well in complete ES medium without antibiotics for overnight. The next day, 2.5 μL of Dharmafect reagent was mixed with 5 μL of 5 μM siRNA stock in 200 μL of serum-free basal GMEM medium and incubated at room temperature for 20 minutes. During the incubation, culture medium in the 12 well plates was aspirated and changed to 800 μL of fresh ES complete medium without antibiotics. The siRNA mixture was then added in drop-wise manner into designated well and returned to 37°C incubator. All transfection experiments were carried out at duplicate and control wells with siRNA been substituted by H_2O must be included. Culture media were changed after 16 to 24 hours after transfection and cells were collected on day 2 or 3 after transfection for qPCR and/or Western blot analysis.

2.2.4.2.2: Transfection by electroporation

For electroporation of plasmid DNA into mES cells, 100 μg of plasmid DNA was digested with the appropriate enzyme for overnight. Digestion

efficiency was assessed by agarose gel electrophoresis after digestion. Digested DNA was then precipitated by ethanol and re-suspend in 200 μL sterile PBS. 10^7 cells were needed for each electroporation. Cells were collected as general sub-culture procedure. Cell pellets were re-suspended in 600 μL PBS, mixed with plasmid DNA and transferred to electroporation cuvettes (Bio-Rad). A negative control sample was performed by replacing DNA with PBS only. After incubation at room temperature for 5 minutes, cells were electroporated at 0.8kV and 3 μF giving a time constant between 0.1 and 0.2. Electroporated cells were then immediately transferred to pre-warmed complete ES medium and plated at $1.5x10^6$, 10^6 and $0.5x10^6$ cells per 10 cm dish in ES complete medium. After culturing in 37°C incubator for overnight, medium was replaced with antibiotic-containing complete medium. For preventing spontaneous differentiation, 0.5 μM of PD03 was added into culture medium. Cells were under selection until the emergence of single colonies with ~1 mm diameter which were then picked into a 96 well plate.

2.2.4.3: Picking mES cell colonies

For picking mES cells single colonies from 10 cm dished to 96 well plates, 20 μL of trypsin solution was added to each well of a pre-gelatinized 96 well plate before starting. Medium was removed from the culture dish and cells were washed twice with PBS. 1 mL of PBS was added into the dish to cover the colonies. Colony-picking pipette tips fitted with p20 pipette were used to lift up the cells. To pick a colony, 10 cm dish was held in horizontal position. 7 μL of trypsin was taking from the trypsin-containing well and added on

the top of the desired colony. The pipette was then touched the cells and operated as reverse mode to lift the cells. Once picked the colony was placed into the well of the 96 well plates and incubated at room temperature for 5 minutes. 200 μL of selection medium was then added to the well and pipette up-and down to dissociate the cells. Cells were grown in selection medium containing PD03 until it reached 80% confluence. Then, cells were passaged to 24 well plates using multi-channel pipette for further experiments.

2.2.4.4: Freezing mES cells

For routine freezing of mES cells, freezing solution of FCS containing 10% DMSO (v/v) was prepared. Cells were collected as usual by centrifugation, re-suspended in freezing medium and transferred to a cryovial (Nunc). The tube was detailed-labelled with the cell line, passage number, date, owner, and cell concentration. Freezing vials were transferred to -80°C freezer and stored in -80°C for less than 2 months or moved to liquid nitrogen cell bank after 1 day in -80°C.

For freezing mES cells in 24 well plates, cells were grown to 90% confluence prior to freezing. Cells were washed with PBS and then lifted by adding 70 μL trypsin solution for 3 minutes. 430 μL of complete ES medium was added in each well using multi-channel pipette to re-suspend the cells. 250 μL of cell suspension was then transferred to a pre-gelatinized 24 well plate containing 250 μL of complete ES media and the plate was placed back to incubator for further examination such as gDNA isolation. Then 250 μL of ES complete medium containing 50 % FCS and 20 % DMSO (v/v) was added into the old plate and mixed homogeneously. The plate was then sealed with

tape and place into -80°C freezer until required.

2.2.4.5: Thawing mES cells

Cells stocks frozen in cryovials were retrieved from either -80°C freezer or liquid nitrogen cell bank and quickly thawed at 37°C in waterbath. The cell suspension was then added to 9 mL of 37°C pre-warmed complete medium and centrifuged at 1,300 rpm for 3 minutes. Cells were re-suspended in complete medium and transferred to a pre-gelatinized culture plastic according to suitable concentration. The medium was changed on the next day.

For cells frozen in 24 well plates, the entire plate was thawed at 37°C in waterbath taking care of not letting water come into the plate. 500 μL of pre-warmed complete medium was added to each well. Cells were transferred to a universal centrifugation tube for centrifugation at 1,300 rpm for 3 minutes. 500 μL of fresh medium was then used for re-suspension and cells were transferred to a new gelatinised well of a 24 well plate. The media was changed on the second day.

2.2.4.6: Induction of doxycycline-dependent transgene

For doxycycline-dependent transgene induction, TTE15 cells were pre-treat with 1 μg/mL of doxycycline 24 hours before differentiation experiments. The pre-treatment of doxycycline was considered to induce enough amount of transgene protein expression before starting differentiation.

2.2.4.7: Differentiation of mES cells

2.2.4.7.1: Monolayer neural differentiation

Monolayer neural differentiation was performed under monolayer conditions as outlined in Pollard et al., 2006. mES cells were grown to 80 % confluence and passaged 1 day before starting neural differentiation. Cells were washed with PBS twice to remove all traces of serum and then trypsinized followed by quench with N2B27 serum-free medium (see section **2.1.8**). Cells were then plated on gelatinized tissue culture dish as in N2B27 serum-free medium. For neural differentiation in 10 cm dishes, 10^6 cells were plated. For neural differentiation in 6 well plates, 1.5 x 10^5 cells were plated. Culture medium was changed every day.

2.2.4.7.2: LIF withdraw

Non-neural differentiation was performed using complete mES media without supplementation of LIF. Cells were passaged using standard procedure and 3 x 10^4 cells were plated per well in a 6 well plate. Cultures were assayed for non-neural differentiation by morphology examination and gene expression.

2.2.4.7.3: Derivation of Epiblast-like stem cells from mES cells cultured in LIF and serum

mES cells was transformed *in vitro* to Epiblast-like stem cells by culturing in N2B27 supplemented with 20 ng/ml Activin A and 12 ng/ml FGF2 (see section **2.1.8**) (Brons et al., 2007; Tesar et al., 2007; Zhang et al., 2010). Briefly, ESCs were seeded on fibronectin-coated 6 well plate at a density of 4x10^4 per well and grow in ES cell complete media for 24 hours.

The cells were washed with PBS twice to remove serum and media were changed to EpiSC media. The culture was kept in 37°C for 5 days, which corresponded to passage 0. Cells were grown until about 70% confluence, washed twice with PBS and then incubated in 200 μL of accutase for 5 minutes. Cells were passaged until reached 70% confluence. Cells were transformed to Epiblast-like stem cells by examination on morphology and gene expression.

2.2.4.7.4: Derivation of Epiblast-like stem cells from 2i culture

To derive Epiblast-like stem from 2i culture, 2i-cultured ES cells were dissociated by accutase and re-suspend as single cell in N2B27 medium containing Activin A (20 ng/ml), FGF2 (12 ng/ml) and KSR (1 %) (Hayashi et al., 2011). Cells were plated at 10^5 cells per well in a 12-well plate coated with fibronectin. Cells were cultured in EpiSC media and collected at different time point.

2.2.4.7.5: Multi-lineage EB differentiation in hanging drops

mES cells were differentiated into a mixture of mesoderm, endoderm, and ectoderm lineages in hanging drops (Wang and Yang, 2008). Cells were collected as usual and re-suspended in mES medium without LIF at a concentration of 10^5 cells/mL. 30 μL of cell suspension were placed on the inside of the lid of a 15 cm sterile non-cell culture dish using a multi-channel pipette. Each drop was consisted of approximately 1,500 cells and roughly 80 drops were placed on a lid. PBS was added to the bottom of the dish and the lid was inverted, placed back and returned to the incubator. Embryoid-like

bodies were collected using 1 mL pipette tips and spin down for RNA isolation at later time point.

2.2.4.8: Removing *FRT*-containing selection cassette from targeted Tcf15-Het cells by FlpO

To remove the selection cassette, Tcf15-Het cells were transfected with 3 μg of pFlpO by liposome- mediated transfection. FlpO is a mouse codon-optimized flippase which binds to flippase recognition target (*FRT*) sites and trigger homologous recombination. Cells were re-plated to 10 cm dish in complete ES culture medium in the presence of 3 μM ganciclovir for selection. Although the transfection efficiency of FlpO was known to be low, the ganciclovir treatment will kill cells which retain the *MC1-tk* selection cassette. Approximately 30-50 colonies were formed after 7 days of selection from each Tcf15-Het clones and 12-15 sub-clones were picked to 96 well plate. Cells were expanded from 96 well plates to 24 well plates and passaged 1:3 into 3 of 24 well plates once confluent: one for frozen at -80°C as cell stock, one for back selection with G418 and the last one for genomic DNA isolation.

2.2.5: Histological techniques

2.2.5.1: Alkaline phosphatase (AP) staining

Alkaline phosphatase (AP) staining was performed using the Leukocyte Alkaline Phosphatase Assay Kit (Sigma) as outlined in the product manual. Cells were washed with PBS and then fixed in a solution of acetone, citrate solution, and formaldehyde (37% v/v) (Sigma) at a ratio of 8.2 : 2.75 : 1. The fixation was performed in a ventilation hood. After 1 minute of fixation, the

cells were washed with ddH_20 twice. Staining solution was made by mixing ddH_20, sodium nitrate, Napthol AS-B1 and FRV alkaline solution at a ratio of 45:1:1:1. Staining solution was then added to cover each well, and samples were incubated at room temperature in the dark for approximately 15 minutes. The staining solution was aspirated and the wells were washed twice with ddH_20. Samples could be kept in ddH_20 for microscopic imaging or dried for colony counting.

2.2.5.2: Immunocytochemistry staining of cultured cells

mES cells were washed twice with PBS and fixed in 4% paraformaldehyde (w/v) at room temperature for 15 minutes or 4°C for overnight. The fixation was aspirated and the wells were washed 2 times with PBS. Permeabilisation and blocking was performed simultaneously by incubation in ICC blocking buffer (see section **2.1.2**) for at least an hour at room temperature. Primary antibodies were incubated as recommended by the manufacturer at 4°C overnight. For most of primary antibodies, antibodies were re-used by adding of 0.02% (w/v) sodium azide and kept in the 4°C fridge. Cells were washed 4 times for 15 minutes each in PBS. The appropriate Alexa Fluor™ secondary antibody (Invitrogen) was diluted 1:1000 in ICC blocking buffer and incubated with samples at room temperature for at least an hour or 4°C overnight. The antibody solution was aspirated and washed 4 times for 15 minutes each with PBS. DNA was counter stained with DAPI at a concentration of 1 μg/mL and fluorescence was visualized using an Olympus IX51 fluorescent microscope. Images were taken and analysed using the Velocity software onsite.

2.2.5.3: Flow cytometry analysis and fluorescence-activated cell sorting (FACS)

Flow cytometry analysis was performed using a Becton Dickinson FACS Calibur flow cytometer. Fluorescence-activated cell sorting (FACS) was carried out on a Becton Dickinson FACS Aria flow cytometer. For collecting cells for flow cytometry analysis, cultured cells were washed with PBS and dissociated using trypsin. After centrifugation at 1,300 rpm for 3 minutes, cells were washed once with ice-cold PBS and spin down again. For live monitoring of reporter cells, cells were re-suspended in FACS buffer and analysed immediately on FACS Calibur. For staining with PECAM1, cells were resuspended in FACS buffer to achieve concentration of $5x10^6$ cells/mL and stained with 1:500 dilution of PECAM1-APC antibody for 20 minutes on ice protecting for light. Stained cells were centrifuged at 1,300 rpm for 3 minutes and washed twice with FACS buffer and spin down again. All samples were filtered through the Cell-Strainer cap of a 5 ml Polystyrene round-bottom tubes (BD Falcon) to remove any cell clumps. Data was analyzed using FLOWJO flow cytometry analysis software (http://www.flowjo.com/index.php).

For FACS of Venus-positive cells, PECAM1-stained cells were re-suspended in FACS buffer containing 100 ng/mL DAPI to exclude dead cells. For sorting of Tcf15-Venus cells, the top 30% and the bottom 30% of Venus-positive cells were sorted from within a platelet endothelial cell adhesion molecule 1 ($PECAM1^+$) gated population, excluding differentiated cells. Cells were sorted into sterile 15 mL tubes containing 1 mL of complete ES media in a refrigerated collecting device. Sorting quality was assessed by

performing purity check on sorted populations after FACS and 95% purity was achieved every time. After sorting, cells were centrifuged at 1,500 rpm for 5 minutes. For collecting RNA, cells were re-suspended in either Trizol or RNA lysis buffer. For keep culturing or performing differentiations, cells were plated at desired density in 1x penicillin/streptomycin containing medium for preventing contamination.

2.2.5.4: *In situ* hybridization of mouse blastocysts

All steps were carried out in RNase-free conditions, up-until post-hybridisation washed. RNaseZap® RNase decontamination solution was used before starting experiments and RNase-free plastics were used to minimise possible contamination. All solutions were prepared in Nuclease-free H_2O and reactions were carried out in a 12 mm diameter transwell (BD) with 8 μm pore size hanging into a sterile 12-well plate (Corning).

2.2.5.4.1: Blastocysts preparation, fixation and dehydration

MF1 x MF1 F1 morula were flushed on E 2.5 by the SCRM animal facility and cultured in KSOM medium in 37°C incubator to reach E3.5 and E4.5 blastocyst stage. Blastocysts at right developmental stage were transferred to transwell by mouth pipette and then fixed in 4 % paraformaldehyde in PBS at 4°C overnight. After fixation, blastocysts were rinsed 2 times with 0.1% PBST and dehydrated through a series of methanol concentrations (25%, 50%, 75% and 100% (v/v) methanol in PBS). Blastocysts could be stored in 100% methanol at -20°C until required.

2.2.5.4.2: Rehydration, post-fixation and hybridization

Blastocysts were taking out of -20°C, placed on ice and dehydration was performed through same methanol series from 100% methanol to PBS. After rehydration, blastocysts were rinsed 3 times with PBST and 1 time in PBSTX (0.1 % Triton X-100 in PBS). Samples were re-fixed 20 minutes with 4 % paraformaldehyde and 0.2% glutaraldehyde followed by 3 times of washed in PBST. Samples were then immersed to pre-hybridization solution, placed into a humidified container with tissue soaking with 1:1 ddH_2O: formamide lying on the bottom and incubated at 70°C overnight.

After pre-hybridization, samples were immersed to RNA probe-containing hybridization buffer. 1 μg of probe was denatured at 80°C for 10 minutes before adding to the buffer. Samples were placed back to 70°C hybridization oven for overnight.

2.2.5.4.3: Sample wash and antibody labelling

After hybridization, samples were washed 5 minutes in post-hybridization washing buffer at 65°C followed by another 3 washes for 30 minutes. To avoid cross-contamination of probes, use different tips for wells hybridized with different probes until back at room temperature. After post-hybridization washes, samples were cooled to room temperature, rinsed 3 times with TBST and blocked in TBST containing 10% heat-inactivated sheep serum at 4°C overnight. Anti-DIG AP-conjugated Fab fragment in 1% sheep serum in TBST was used at 1:2000 for labelling.

2.2.5.4.4: Staining

Blastocysts samples were washed 3 times in TBST for 5 minutes each

followed by 2 hours washes at room temperature on a shaker. TBST was removed and samples were further washed 3 times for 10 minutes in alkaline phosphatase buffer. Staining solution was prepared by mixing 4.5 μL NBT with 3.5 μL BCIP (Roche) per ml of alkaline phosphatase buffer. The colour reaction was carried out in the dark, assessed several time during the reaction until blue colour was formed. When staining is dark enough, stop the reaction by rinsing in PBT with 1 mM EDTA. For *Tcf15* staining on blastocysts, it took around 36 hours to stain the blastocyst. Blastocysts could be store at 4°C for later image acquisition.

2.2.5.5: Karyotype analysis of mES cells

In general, chromosomal counting was performed with the help from Theresa O'Connor of the SCRM tissue culture core facility. Briefly, metaphase spreads were prepared by adding mES cell medium supplemented with colcemid (10μg/ml) to exponentially growing at ~70% confluence. Cultures were returned to the incubated for approximately 1 hour after which time metaphase arrested cells were collected and incubated in 2 ml hypotonic solution (0.4% KCl, 0.4% Sodium Citrate) for a further 5 minutes. Cells were fixed in 2ml fixation buffer (3:1 methanol: acetic acid) at RT for 30 minutes. Cells were dropped onto a glass slide, flooded with fixative and allowed to dry before staining with Giemsa and chromosome counting.

2.2.6: Microarray analysis of gene expression

TTE15 and parental AW2 cells were plated in 6well plates (IWAKI) one

day before stimulation. Cells were stimulated by 1 μg/mL doxycycline for the indicated time and lysed for RNA. RNA was prepared using Absolutely RNA Purification Kits (see section **2.2.2.1.1**). cRNA was synthesized using an Illumina TotalPrep RNA Amplification Kit (see section **2.2.2.3**). Labelled RNA was submitted to the WTCRF MRC Human Genetics Unit (University of Edinburgh) for further processing. For each sample, 1500 ng of cRNA in 10 μL of elution solution was needed. cRNA quality was checked using a Agilent 2100 Bioanalyser and hybridization was performed on an MouseWG-6 v2 BeadChip (Illumina). Raw data was processed in R using the beadarray (Dunning et al., 2007) and limma (Smyth et al., 2005) packages from the Bioconductor suite (Gentleman et al., 2004), with the help from Dr. Mattias Malaguti in our lab and Dr. Florian Halbritter in Dr. Simon Tomlinson's lab in University of Edinburgh. Briefly, we removed low-quality probes from the input data. The data was subsequently quantile-normalized and $\log_2$-transformed before assessing differential expression with the limma algorithms. We considered genes differentially expressed which had an FDR-adjusted p-value of at most 0.05 and a fold change of 1.5 or more for at least one time point in comparison to the 0-hour baseline.

2.2.7: Quantification of nuclear immunostaining

In order to quantify the fluorescence signal in each individual cell, we generated an automated pipeline for image analysis. Briefly, RGB pictures were registered and pre-processed using a plugin in ImageJ (http://rsbweb.nih.gov/ij/) developed by Dr. Guillaume Blin. The pre-processing step consisted in a background subtraction in each channel as

well as a gamma correction of the blue channel to reveal low intensity nuclei. Then, to detect single cell nuclei, the blue channel (DAPI) was segmented using a previously published algorithm (Li et al., 2007) with the following parameters values: sigma = 0.15, minimum nucleus size = 350 pixels and fusion threshold = 1. This algorithm provides a picture within which each nucleus is labelled with a unique colour in the image. Using a homemade java application that Dr. Guillaume Blin developed with eclipse (http://www.eclipse.org), the signal in the red and green channels of the pre-processed RGB picture was measured in the superimposed area of each nucleus to calculate the average intensity. Finally, the red and green average intensities were plotted using R.

CHAPTER 3

Tcf15 expression during differentiation *in vitro* and during early development *in vivo*

3.1: Aims of this chapter

Tcf15 has recognized functions during somitogenesis and is known to be a somite marker (Burgess et al., 1995). However, there are no previously identified functions in mouse embryonic stem (mES) cells or in pre-implantation embryos. A previous study within our lab using quantitative PCR (qPCR) to analyse *Tcf15* expression in mouse embryos at different developmental stages from E4.5 to E10.5 showed a two-wave temporal expression pattern of *Tcf15* (Figure 3.1). The first wave was in pre-implantation E4.5 stage. Then, the expression of *Tcf15* dramatically declined when embryo implanted until rising again at E7.5 to E8.5, in keeping with the previously reported expression of *Tcf15* in the developing somites. The newly identified expression of *Tcf15* in E4.5 blastocysts raised the possibility that Tcf15 might play a role in the pre-implantation embryos in addition to its known function during somitogenesis.

The primary aim of this chapter was to characterize the expression pattern of Tcf15 *in vivo* and *in vitro*. The *in vivo* study was performed using pre-implantation mouse blastocysts. Mouse ES cells were used as *in vitro* materials for identifying differential expression levels of Tcf15 in pluripotent self-renewing cultures from naïve to primed states as well as during the differentiation process. Also, two independently generated Tcf15-Venus reporter cell lines were characterized and used for monitoring Tcf15 expression at the single cell level during mES cell differentiation and self-renewal.

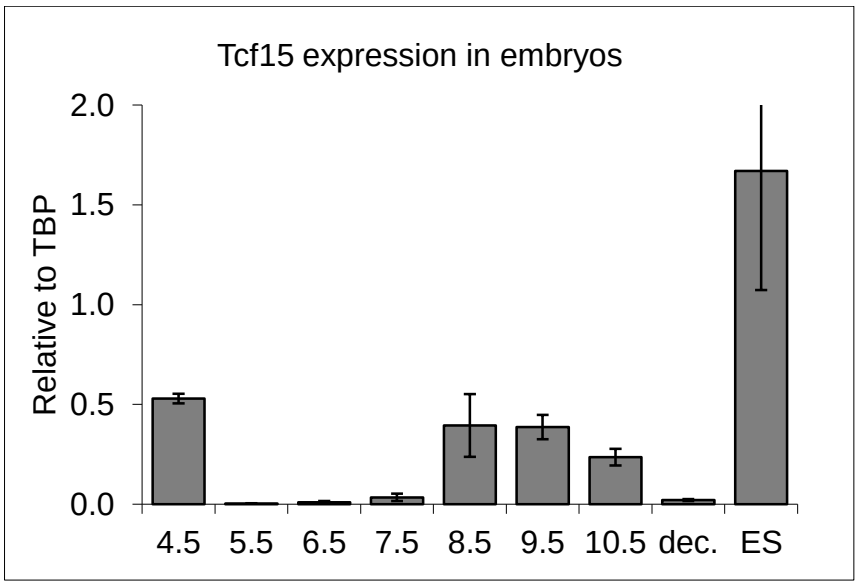


Figure 3.1: qPCR analysis of Tcf15 expression in mouse embryos at different developmental stages

Expression of *Tcf15* during mouse embryogenesis displayed a two-wave expression pattern. The first wave was in pre-implantation E4.5 stage. The second wave of *Tcf15* expression arose at E7.5 to E8.5 stage, in keeping with the previously reported expression of *Tcf15* in the developing somites.

Dec, decidual tissue; ES, embryonic stem cells. (data from Aliaksandra Radzisheuskaya)

3.2: Results

3.2.1: Characterization of Tcf15 expression in mouse blastocysts

3.2.1.1: *In situ* hybridization identified Tcf15 expression in the inner cell mass (ICM) of mouse blastocysts

As shown in figure 3.1, *Tcf15* mRNA was expressed in pre-implantation blastocysts. However, pre-implantation blastocysts consist of outer trophectoderm (TE) and inner cell mass (ICM). It is therefore important to ask where Tcf15 is expressed in pre-implantation blastocyst.

Due to lack of commercially available antibodies which could recognize Tcf15 clearly in mouse embryos by immunostaining, *in situ* hybridization against *Tcf15* mRNA was performed to investigate where *Tcf15* was expressed within the blastocyst. An *Oct4* probe was used as positive control to mark ICM cells within the blastocyst (Figure 3.2A). *Tcf15* was expressed in the E4.5 blastocyst, and was specifically detected in the ICM compartment (Figure 3.2B). This also confirmed that the expression of *Tcf15* we had detected in mouse ES cells (Figure 3.1) is consistent with its expression in the *in vivo* counterpart of ES cells, the ICM.

A) ***Oct4*** **B)** ***Tcf15***

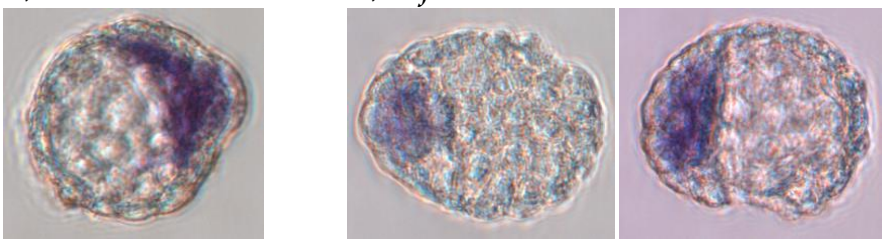


Figure 3.2: Whole mount *in situ* hybridization of *Oct4* and *Tcf15* in early mouse blastocysts

MF1 x MF1 mouse morulae were flushed at E2.5 and cultured in KSOM medium to reach blastocyst stage. Blastocysts were fixed after hatching in 4% paraformaldehyde and transferred to transwells for staining procedures. Digoxigenin-labelled anti-Tcf15 riboprobe was synthesized from *Tcf15* cDNA clones (Burgess et al., 1995) and Oct4 probe was gifted by Dr. Val Wilson's lab.

(A) Blastocyst staining with Oct4 probe as control.

(B) Two blastocysts were shown here where *Tcf15* expression was consistently detected in the ICM. Of 33 blastocysts examined from two independent ISH that had reached a similar stage, 31 had a clear signal for *Tcf15* specifically within the ICM. The remainder of the embryos had no detectable expression. *Tcf15* was never detected in the trophectoderm.

3.2.2: Characterization of Tcf15 expression in mouse embryonic stem cells

After confirming the *in vivo* expression of *Tcf15*, both mRNA and protein expression of Tcf15 were investigated *in vitro* using mouse ES cells. Expression of Tcf15 was analysed in pluripotent culture conditions as well as during differentiation.

3.2.2.1: Expression level of Tcf15 in pluripotent cultures

3.2.2.1.1: Tcf15 expression is higher in primed pluripotent ES cells than in naïve ES cells and EpiSCs

As shown above, *Tcf15* is expressed in the ICM of mouse blastocysts (Figure 3.2), as well as ICM-derived pluripotent mES cells (Figure 3.1). In addition, single-cell qPCR analysis by others (Tang et al., 2010) has revealed that *Tcf15* is up-regulated as ES cells are derived from E3.5 blastocysts and that *Tcf15* is predominantly associated with the *Gata6*-negative subpopulation during the derivation process (Figure 3.3). This suggests that Tcf15 may become specifically up-regulated in cells that are maturing into epiblast-primed ES cells but not into primitive endoderm, although relatively few cells were examined in this study and so firm conclusions cannot be drawn.

In order to understand the relative expression level of *Tcf15* in different self-renewing pluripotent cultures from naïve to primed and post-implantation states, mES cells were established and analysed in three different pluripotent culture conditions. mES cells cultured in the serum-free N2B27 medium containing 2i mimicked the ground state naïve ES cells in

which differentiation cues had been blocked by two pharmacological inhibitors (PD0325901 and Chiron99021) against MEK and GSK-3β (Ying et al., 2008) (see section **1.1.2.2.5**). mES cells cultured in normal culture medium containing LIF and serum (FCS) were considered as a mixture of cells in naïve and differentiation-primed states in which cells had heterogeneous gene expression and some might already be primed for certain lineages (Nichols and Smith, 2009). Post-implantation epiblast stem cells (EpiSCs) culture was established by maintaining cells in serum-free N2B27 medium containing FGF2 and Activin (Tesar et al., 2007). *Tcf15* expression was compared with naïve pluripotency marker *Klf4* and epiblast stem cell marker *Fgf5* (Figure 3.4A). The result showed cells in a primed state expressed higher levels of *Tcf15* than cells in a naïve pluripotent state. In addition, EpiSCs (which are equivalent to post-implantation pluripotent epiblast cells) also expressed lower levels of *Tcf15* than ES cells. These findings are consistent with the qPCR analysis of embryos which indicates that *Tcf15* expression drops after embryos implant, i.e. between E4.5 and E5.5 (Figure 3.1).

3.2.2.1.2: Tcf15 expression is transiently up-regulated during the 2i to EpiSC transition

In order to further investigate how the expression of Tcf15 changes as cells move between pluripotent naïve, primed and epiblast states, a 2i to EpiSC transition experiment was performed as showed in figure 3.4B. ES cells were cultured in 2i conditions supplemented with LIF for two passages before starting the experiment. Cells were then re-plated directly onto fibronectin coated plates in EpiSC medium containing Activin and FGF2. EpiSC medium was also supplemented with 1% knockout serum

replacement (KSR) to reduce cell death and to sustain uniform induction of ES cells into flattened epithelial structures resembling the EpiSCs over the experiment time period (Hayashi et al., 2011). Gene expression was analysed during the 2i to EpiSC transition.

qPCR analysis on marker gene expression was first confirmed that the cells behaved as expected under these conditions. As shown in figure 3.4C, the expression of naïve pluripotency markers *Nanog* and *Klf4* declined dramatically within 24 hours, indicating cells were progressing from a naïve state to primed state as expected. Expression of EpiSC marker *Fgf5* began at day 1, indicating cells start acquiring EpiSC-like gene expression after this time point.

Expression of Tcf15 was then examined under these conditions. *Tcf15* expression was low when culturing in 2i conditions. Expression of *Tcf15* significantly increased rapidly when cells were released from anti-differentiation 2i culture conditions and re-plated in Activin/ FGF2-containing EpiSC medium. After the initial switch from a naïve to primed state, *Tcf15* up-regulation persisted for first 12 hours and then started decreasing between 12 and 24 hours as cells loss naïve pluripotency markers and gradually transited to EpiSC-like state. The epiblast-determinant gene *Otx2* (Acampora et al., 2012) showed a correlation with *Tcf15* during the early part of this transition. However, unlike *Tcf15*, *Otx2* remained highly expressed in the EpiSC-like state. The transient peak of *Tcf15* expression at early time points during 2i to EpiSC transition is consistent with the steady-state expression pattern in LIF and serum culture conditions which represent an intermediate between naïve and post-implantation epiblast states. This raises the possibility that Tcf15 might play a role during the transition from a naïve pluripotent state to primed state.

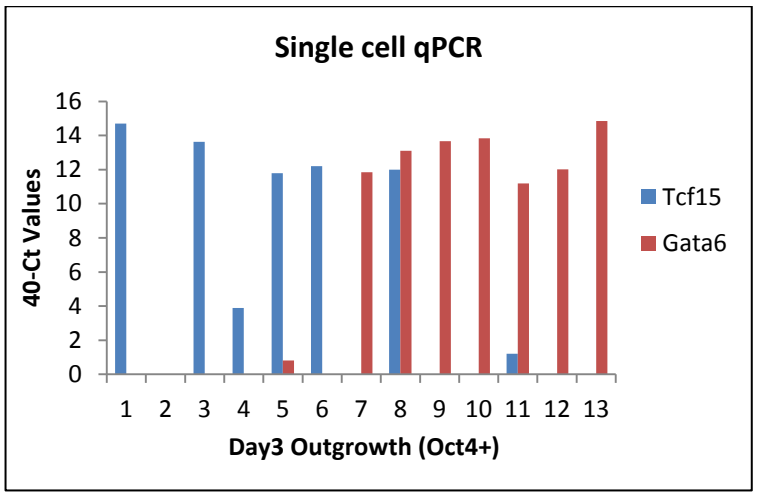


Figure 3.3: Single cell qPCR from day three ICM outgrowth cells

Tcf15 and *Gata6* expression from Oct4-positive cells after 3 days of ICM outgrowth were extracted from the supplementary data from Tang et al., 2010. Expression of *Tcf15* is predominantly accociated with *Gata6*-negative population.

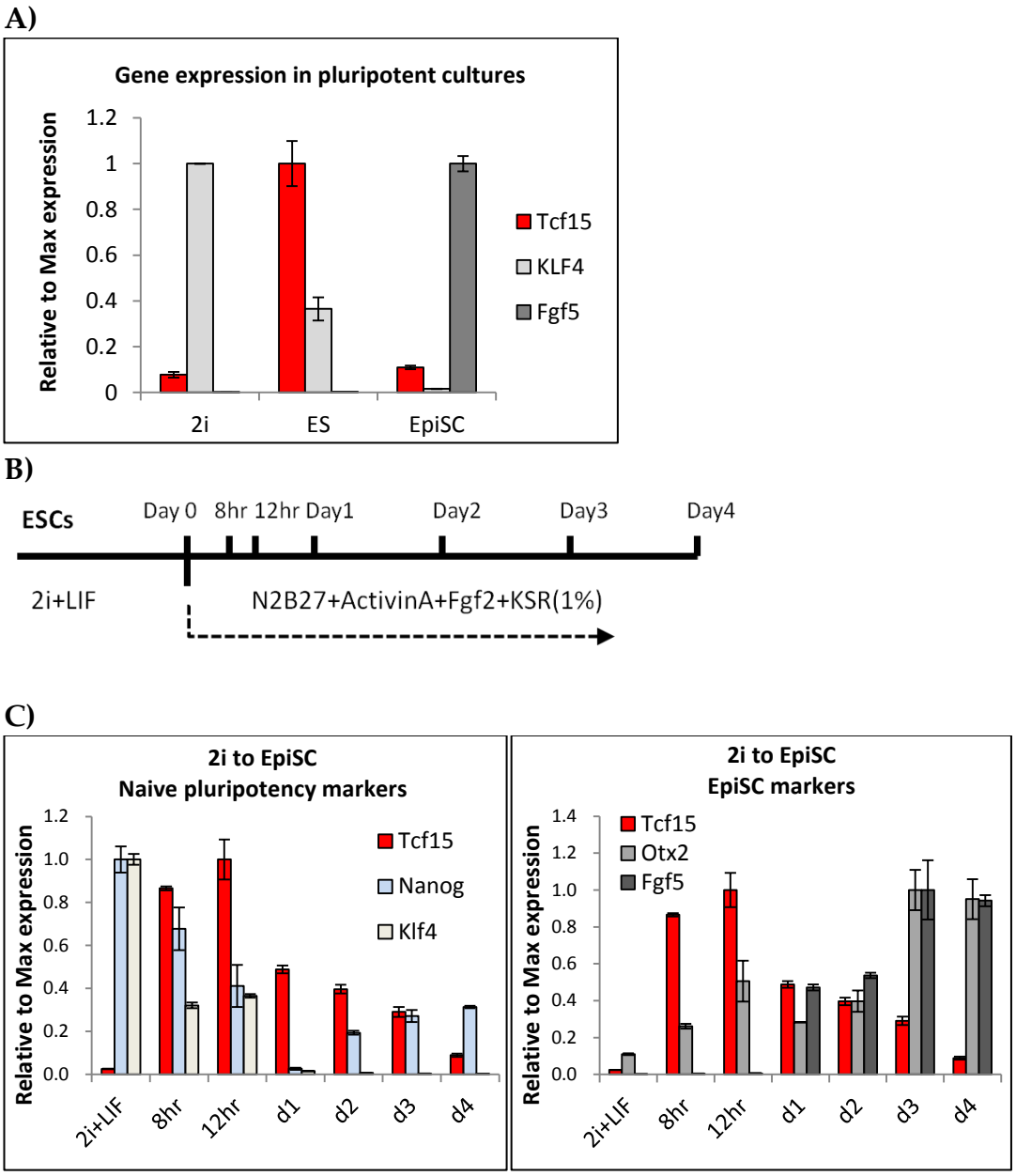


Figure 3.4: Endogenous expression of Tcf15 in pluripotent cultures

(A) Mouse ES cells gene expression in self-renewing cultures. *Klf4* (naïve pluripotency marker) and *Fgf5* (EpiSC marker) were used to monitor cell identity. Expression of each gene was normalized to highest expression. *Tcf15* expression was significantly higher in ES cells than in 2i and EpiSC cultures.

(B) The scheme for 2i to EpiSC induction.

(C) Gene expression during 2i to EpiSC transition. Relative expression was normalized to highest expression. This experiment was performed twice (N=2, n=3).

3.2.2.2: Expression level of Tcf15 during differentiation

After analysing *Tcf15* expression in self-renewing conditions, the expression level of *Tcf15* was examined during differentiation of ES cells to neural cells under monolayer neural differentiation protocol (Ying et al., 2003) (Figure 3.5A) or to non-neural lineages by withdrawing LIF from the culture medium (Figure 3.5B). In both differentiation experiments, *Tcf15* expression persisted after the naïve pluripotency marker *Nanog* and *Rex1* were down-regulated. The expression level of *Tcf15* was higher in pluripotent ES cells than in differentiated cells. *Tcf15* expression then declined at around the time that cells commit to differentiation, as measured by the up-regulation of early neural marker *Sox1* (Wood and Episkopou, 1999) and *N-cadherin* (Riehl et al., 1996) in neural differentiation conditions and *T-Brachyury* in non-neural differentiation conditions (Lanner and Rossant, 2010).

To further study the expression level of *Tcf15* during spontaneous differentiation, ES cells were cultured in suspension in the absence of anti-differentiation factor LIF to form three-dimensional (3D) aggregates called embryoid-like bodies (EBs) (Kurosawa, 2007) (Figure 3.5C). ES cells could differentiate into derivatives of ectodermal, mesodermal, and endodermal tissues which recapitulate many aspects of cell differentiation during early mammalian embryogenesis. *Tcf15* expression during EB formation displayed a similar trend as in the directed differentiation assays described above. The expression of *Tcf15* persisted after naïve gene *Rex1* declined and down-regulated before lineage specific gene *Sox1* (neuroectoderm), *T* (mesoderm) and *FoxA2* (endoderm) began to be expressed.

These results suggested that the transient up-regulation of *Tcf15* prior to differentiation is not associated only with the transition to neural fates but is

associated with differentiation into all epiblast-derived lineages. Moreover, it suggests that any role played by Tcf15 must take place during the early stages of differentiation, i.e. during the transition from a naïve pluripotent state to primed state, before cells commit to any specific lineage.

Figure 3.5: Tcf15 expression in different differentiation contexts

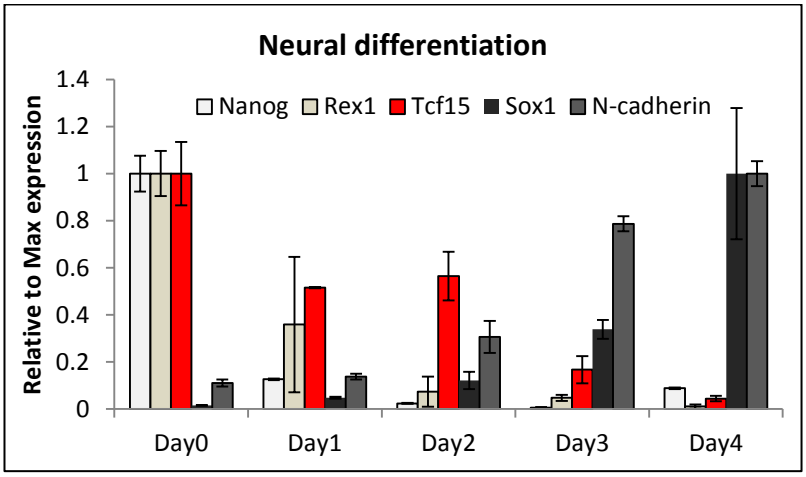
Tcf15 expression persisted after the naïve pluripotency markers *Nanog* or *Rex1* were down-regulated and then declined at around the time cells commit to differentiation, as marked by lineage specific markers *Sox1*, *N-cadherin*, *T-Brachyury* or *FoxA2*.

(A) Expression of *Tcf15* during serum-free monolayer neural induction conditions.

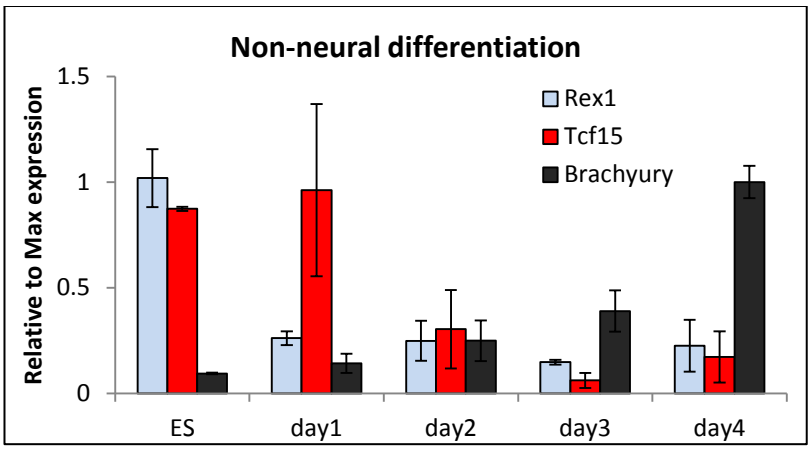
(B) Expression of *Tcf15* during non-neural differentiation condition: LIF withdraw from complete medium. (This experiment is performed by Xinzhi Zhou in the lab)

(C) Expression of *Tcf15* during multi-lineage differentiation: EB formation.

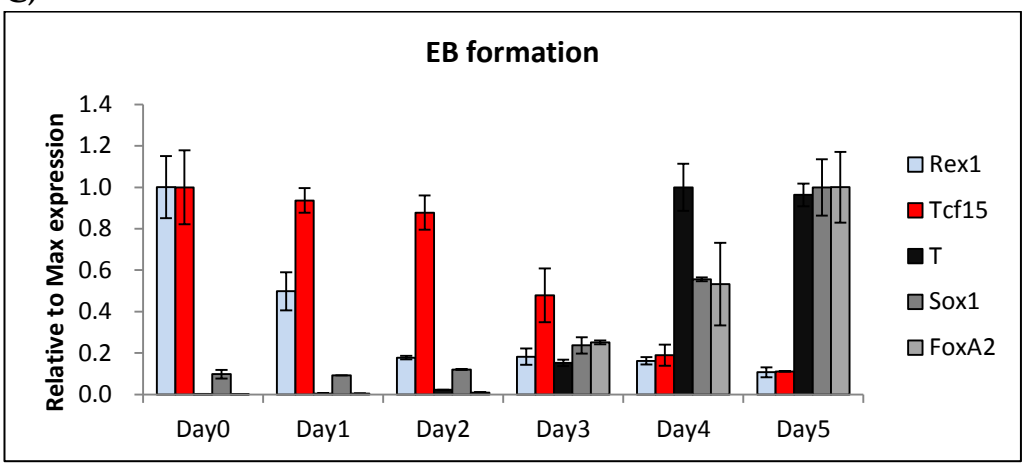
A)


Neural differentiation
Nanog
Rex1
Tcf15
Sox1
N-cadherin
Relative to Max expression
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Day0
Day1
Day2
Day3
Day4

B)


Non-neural differentiation
Rex1
Tcf15
Brachyury
Relative to Max expression
1.5
1
0.5
0
ES
day1
day2
day3
day4

C)


EB formation
Relative to Max expression
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Day0
Day1
Day2
Day3
Day4
Day5
Rex1
Tcf15
T
Sox1
FoxA2

3.2.2.3: Tcf15 reporter cells

After checking the gene expression of *Tcf15* in bulk populations, the expression pattern of Tcf15 within ES cells was investigated at the single cell level using live-fluorescent Tcf15 reporter cell lines.

3.2.2.3.1: Analyzing Tcf15 expression pattern in mES cells by 216D1 Tcf15-Venus reporter cells

3.2.2.3.1.1: Tcf15-Venus is heterogeneously expressed within ES cells and is negatively correlated with Nanog and Klf4, but not Oct4

ES cells cultured in LIF and serum are heterogeneous mixtures of primed cells and naïve cells (Nichols and Smith, 2009). Naïve cells within these cultures are marked by Klf4 and Nanog. In order to find out whether Tcf15 expression was associated with the primed state in these mixed cultures, a sensitive reporting strategy for Tcf15 was used based on translational amplification of a fluorescent protein, Venus. The 216D1 Tcf15-reporter cell line (referred hereafter as Tcf15-Venus cell line) was generated by the International Gene Trap Consortium (Toronto Centre for Phenogenomics) using gene-trapping: a *gtx-IRES-Venus* cassette (Figure 3.6A) was incorporated into the only intron of *Tcf15* (Tanaka et al., 2008) (Figure 3.6B). *Tcf15* transcripts were visualized based on the expression of the enhanced YFP, Venus, coupled to a unique translational amplifier, *gtx-IRES*.

The 216D1 Tcf15-Venus reporter cell line had previously been validated as a faithful reporter of Tcf15 expression (Tanaka et al., 2008). In order to further confirm this, FACS were used to sort Tcf15-Venus cells from the heterogeneous pluripotent LIF and serum cultures through a series of gating strategies as showed in figure 3.7A. In short, Tcf15-Venus cells were first

gated with FSC and SSC for size and granularity, followed by cell doublets exclusion. Then, a live-dead gate was carried out to identify DAPI-negative live cells. Furthermore, Platelet endothelial cell adhesion molecule 1 (PECAM1, CD31) (Furusawa et al., 2004) was used to exclude differentiated cells (Figure 3.7A). The top and bottom 30 % of Venus-positive cells were sorted from PECAM1-positive population and labelled as Venus-high or Venus-low, respectively (Figure 3.7B). Expression of Venus correlates with *Tcf15* mRNA (assessed by qPCR) and with Tcf15 protein (assessed by western blot analysis) (Figure 3.7C and 3.7D). It is worth noting that after screening all the commercial available Tcf15 antibodies, the one used in this thesis (Aviva) was the only one which is working (data not shown). However, the western blot displayed an additional 200kD band along with the predicted 21kD Tcf15 band. Therefore, it is not suitable for performing immunostaining using this antibody.

Next, the distribution of Venus within populations of ES cells cultured in LIF and serum was examined. Expression of Venus protein showed a heterogeneous distribution within this pluripotent cell culture (Figure 3.8A). Moreover, co-immunostaining of pluripotency factors Oct4, Nanog and Klf4 with Tcf15-Venus showed that Tcf15-Venus expression was higher in Nanog-low cells as well as in Klf4-low cells and lower in Nanog-high and Klf4-high cells. However, Oct4 did not seem to be enriched in Tcf15-Venus-high cells, suggested that Tcf15-Venus marks primed ES cells within the pluripotent population rather than marking differentiated cells.

To gain further insight into the relationship between Tcf15-Venus and Nanog, Klf4, or Oct4 in ES cells, an image analysis software developed by Dr. Guillaume Blin in the lab (http://www.eurostemcell.org/MultiCell3D) was used to quantify the immunofluorescence staining (Figure 3.8B and 3.8C).

Fluorescent signals were plotted to a FACS-like 2-dimensional scatter plot (Figure 3.8B) and pie charts were made to illustrate the percentage distribution of populations (Figure 3.8C). Tcf15-Venus was heterogeneously expressed within pluripotent cultures and most of the Venus-positive cells expressed Oct4, whereas the presence of double-positive Tcf15-Venus-high/ Nanog-high and Tcf15-Venus-high/ Klf4-high was low, as seen in the near absence of this population in the scatter plot and less than 10% in the pie chart. This indicates that Tcf15 marks an undifferentiated but differentiation-primed state.

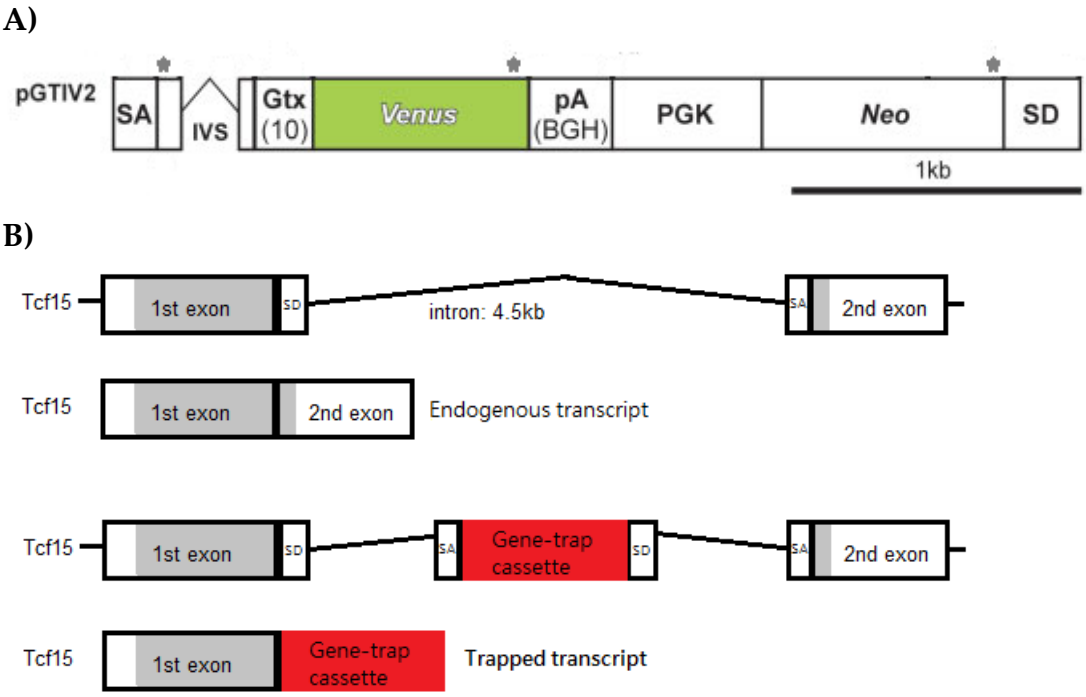
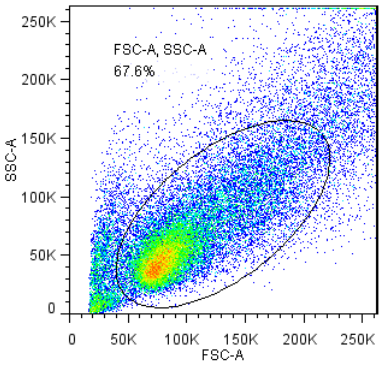
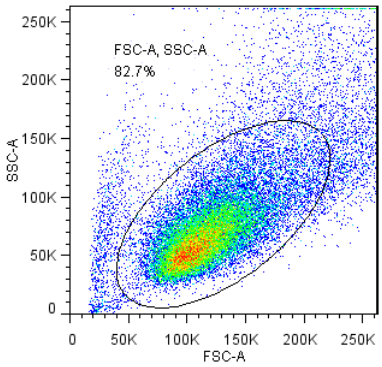
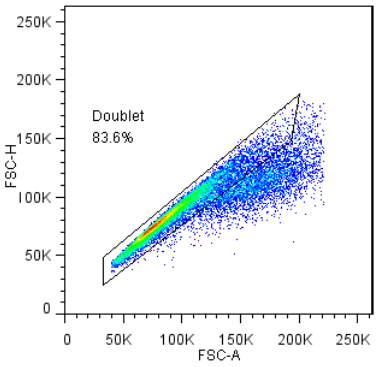
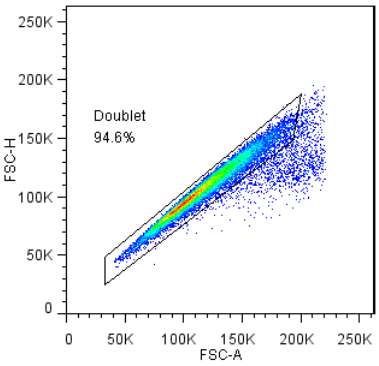
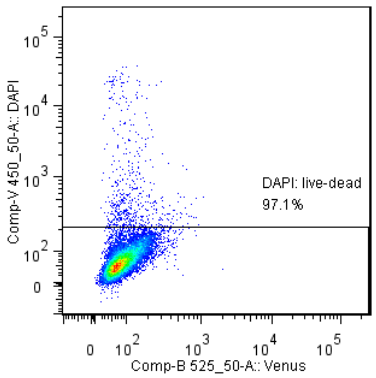
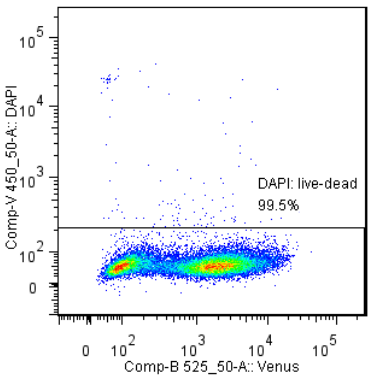
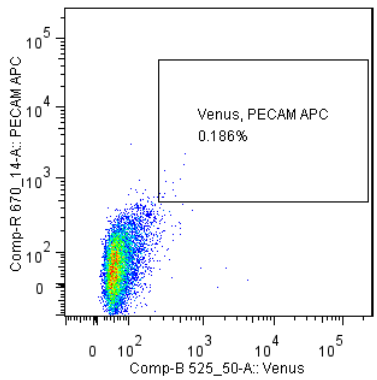
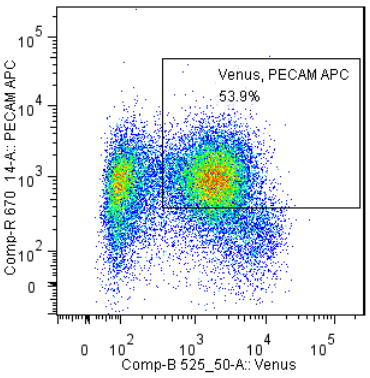


Figure 3.6: Schematic representation of Tcf15-Venus reporter line (Adapted from Tanaka et al., 2008)

(A) Gene-trap cassette used to generate Tcf15 reporter cell line. A bar at the bottom indicates 1 kb. Stop codons are indicated as asterisks (*). IVS, synthetic intervening sequence; pA, polyadenylation signals from bovine growth hormone; SA, adenoviral splice acceptor site; SD, HPRT splice donor site.

(B) Gene-trap strategy for generating Tcf15 reporter line. White boxes: untranslated regions, grey boxes: coding region, diagonal lines: intronic DNA, straight lines: untranscribed intergenic regions. Trapped transcript will generate 3 mRNAs including truncated *Tcf15* with only 1st exon sequence, *Venus*, and *Neo-resistant* gene.

A)

	E14tg2a	216D1 Tcf15-Venus
FSC, SSC gate		
Doublet excluding gate: FSC-A, FSC-H		
Live/Dead gate: DAPI^{-}		
Pluripotency /Venus gate: PECAM1-APC^{+} /Venus^{+}		

(Figure continued to next page)

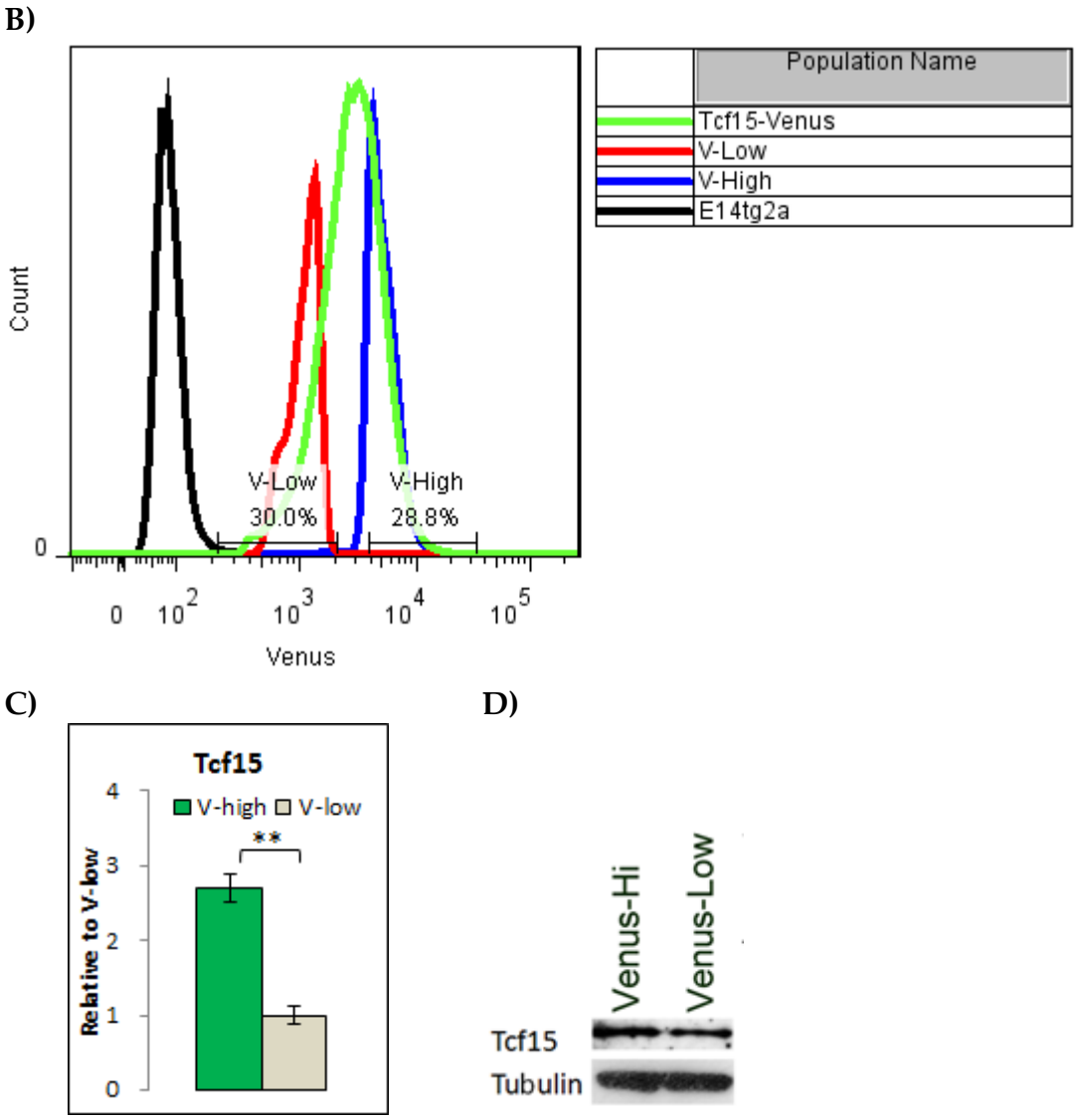


Figure 3.7: Differential Tcf15 expression between FACS sorted 216D1 Tcf15-Venus reporter cells

(A) Sequential FACS gating of 216D1 Tcf15-Venus cells and the control E14tg2a cells.

(B) FACS profile of Tcf15-Venus reporter line. E14tg2a cells were used as negative control (black). Top and bottom 30% of Tcf15-Venus cells (green) were sorted as indicated in the figure. Venus-high (blue) and Venus-low (red) populations were analysed right after sort for purity check.

(C) qPCR analysis of *Tcf15* from FACS sorted Tcf15-Venus reporter cells. **: $p<0.01$

(D) Western blot analysis of Tcf15 from FACS sorted Tcf15-Venus cells. Tubulin was used as internal control.

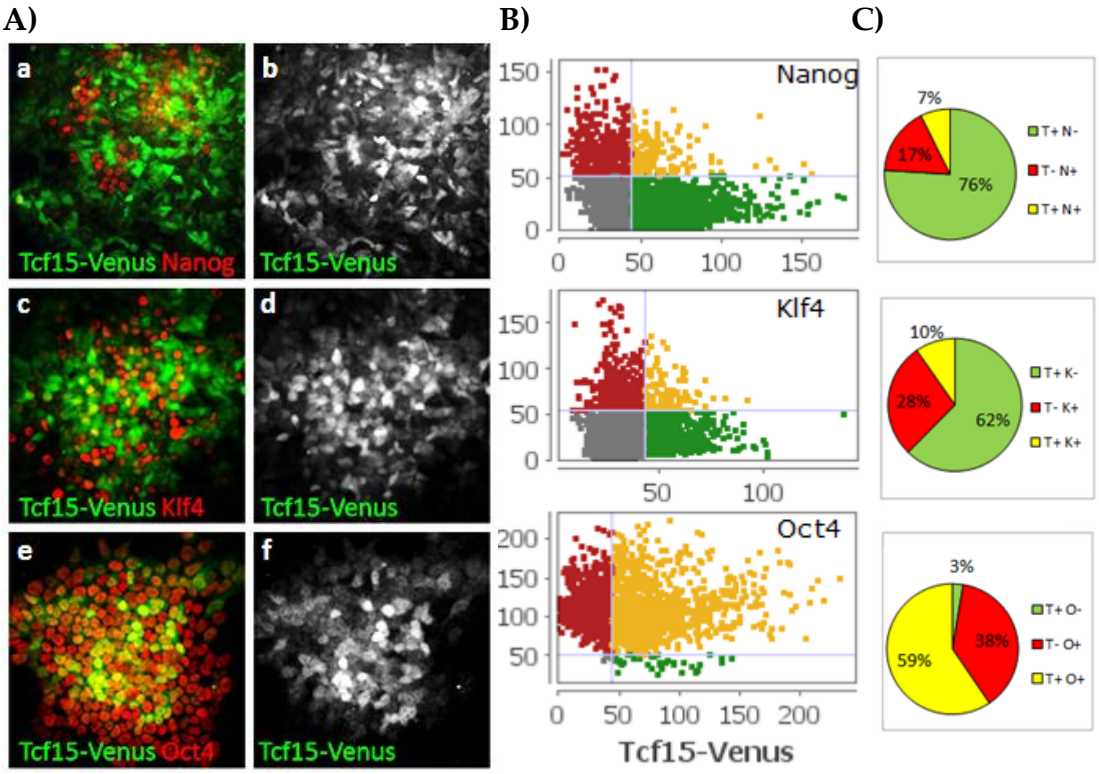


Figure 3.8: Tcf15-Venus is heterogeneously expressed within ES cells and is negatively correlated with Nanog and Klf4, but not Oct4

(A) Immuno-staining of Tcf15-Venus reporter cells cultured five days from a single cell in pluripotent culture. Each row (a and b, c and d, and e and f) was same colony stained with different antibodies. Venus was stained with rabbit anti-GFP antibody to achieve better fluorescent signal. Tcf15-Venus was heterogeneous expressed within colonies (b, d and f). Naïve pluripotency markers Nanog (a, showed in red) and Klf4 (c, showed in red) expressed at higher levels in Venus-low cells, showing a negative correlation with Tcf15-Venus. However, Oct4 was expressed in both Venus-high and Venus-low cells (e, showed in red).

(B) Quantitative immunofluorescence analysis of Tcf15-Venus reporter cells stained for Nanog, Klf4, or Oct4.

(C) Proportion of cells positive for Tcf15-Venus only (green) or Nanog/Klf4/Oct4 (red) or both (yellow). Double negative cells were excluded.

This image analysis is performed with the help from Dr. Guilluame Blin.

3.2.2.3.1.2: Gene expression analysis of FACS sorted Tcf15-Venus reporter cells indicates Tcf15 might mark cells primed for somatic lineages

To further address the population heterogeneity of Tcf15, qPCR and western blot analysis were performed on FACS sorted Tcf15-Venus-high and Tcf15-Venus-low cells. Venus-high cells expressed lower Nanog and Klf4 at both protein and mRNA level, supporting our previous observation from immunostaining (Figure 3.9A and 3.9C). Venus-high cells also expressed lower level of naïve pluripotency markers *Klf2* and *Rex1* compared with Venus-low cells. The expression of pan-pluripotency marker *Oct4* was similar between these two populations (Figure 3.9B), confirming the results of the immunostaining (Figure 3.8B). Moreover, Venus-high cells expressed slightly lower level of the extra-embryonic endoderm markers *Hex* (Canham et al., 2010) and *PdgfRa* (Plusa et al., 2008) (Figure 3.9D) and slightly higher level of the epiblast markers *Fgf5* and *T* (Toyooka et al., 2008) (Figure 3.9E). These differences may not be significant, but they do at least exclude the possibility that Tcf15 specifically marks cells that are primed for extra-embryonic endoderm differentiation (Canham et al., 2010).

In conclusion, based on experiments described above, pluripotent ES cells contained both Tcf15-Venus-high and Tcf15-Venus-low populations. The Venus-high cells were neither in a naïve pluripotent state nor primed for extra-embryonic endoderm, based on marker analysis. This, together with the slightly higher expression of epiblast markers within the Venus-high population, lead to our hypothesis that Tcf15 is marking pluripotent cells which are primed toward somatic lineages (Figure 3.10).

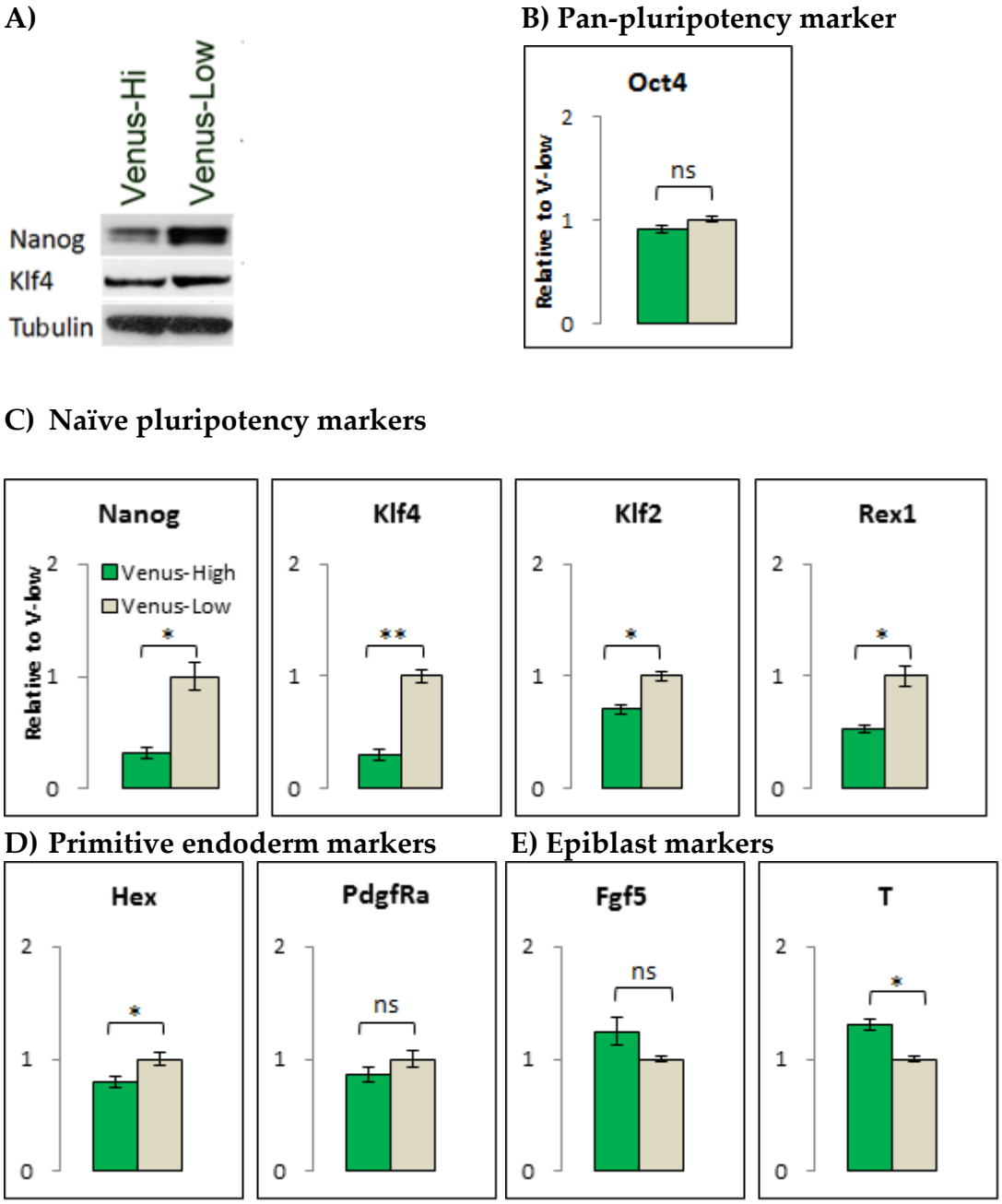


Figure 3.9: Differential expression between two populations of FACS sorted Tcf15-Venus reporter cells

(A) Western blot analysis of Nanog and Klf4 from FACS sorted Tcf15-Venus cells. Tubulin was used as internal control. This blot is the same as that shown in figure 3.5D.
(B) qPCR of pan-pluripotency marker *Oct4*.
(C) Naïve pluripotency markers *Nanog*, *Klf4*, *Klf2* and *Rex1*.
(D) Extra-embryonic marker genes *Hex* and *PdgfRα*.
(E) Epiblast markers *Fgf5* and *T*.
ns: non-significant; *: $p<0.05$; **: $p<0.01$

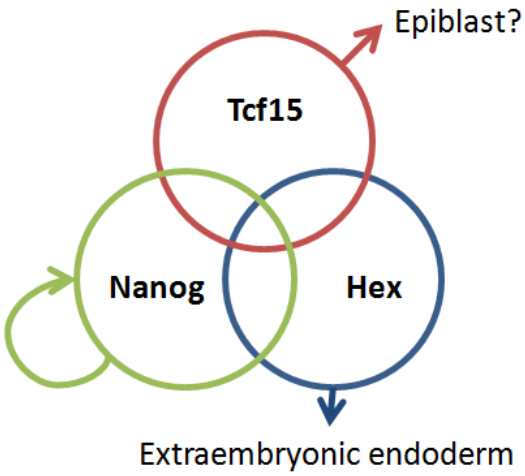


Figure 3.10: Hypothesis: Tcf15 might prime cells for somatic lineages

Differential gene expression between Tcf15-high and Tcf15-low cells indicated Tcf15-high cells were not in a naïve pluripotent state (green circle, marked by Nanog expression) nor primed for extraembryonic endoderm (blue circle, marked by Hex expression). Therefore Tcf15 might mark cells primed for epiblast-derived somatic lineages (red circle).

3.2.2.3.1.3: Differential gene expression between Tcf15-high and Tcf15-low cells under conditions in which Id inhibition is reduced

Experiments above describe changes in gene and protein expression, notably a down-regulation of Nanog and Klf4, which correlate with a high level of Tcf15 expression in pluripotent cells.

However, in some cases Tcf15 may be expressed but held in an inactive state by Id proteins, thus complicating this type of analysis. Id proteins are present in pluripotent ES cell culture and generally act as inhibitors of bHLH transcription factors activity (Ying et al., 2003, Norton, 2000) (see section **1.2.2**). In the presence of Id proteins, Id dimerizes with and sequesters the ubiquitous bHLH transcription factor E12/E47, blocking its interaction with Tcf15. E12/E47 is an essential heterodimerisation partner for Tcf15 in its activation of downstream target genes, and so this effect of Id abolishes Tcf15 activity (Wilson-Rawls et al., 2004) (Figure 3.11A). Hence, expression of Id within pluripotent cells is likely to render Tcf15 inactive.

Therefore, the differential gene expression between Tcf15-Venus-high and Tcf15-Venus-low cells in the absence of Id proteins was being investigated. Id expression is sustained by BMP or serum in culture medium, and is lost within 24 hours of transferring to serum-free medium (Ying et al., 2003). Tcf15-Venus cells were therefore switched to N2B27 basal medium 24 hours before performing FACS analysis. As shown in Figure 3.11B, ES cells cultured in LIF and serum expressed higher *Id1* than N2B27 due to functional BMP signalling machinery within undifferentiated ES cells induced by the presence of serum (Ying et al., 2003), although under these conditions Id1 is only expressed in around 50% of cells (Mattias Malaguti, manuscript in preparation). Higher Id expression is likely to sequester E12/E47 and therefore suppress the transcription activity of Tcf15 in at least some of the

cells within ES cell cultures. *Id1* expression was down-regulated after cells were switched from culturing in LIF and serum to N2B27 (Figure 3.11B), as expected (Ying et al., 2003). Analysis of *Oct4* expression suggested that cells were not committed to differentiation at this point. *E47* expression was maintained after medium switch while *Tcf15* expression was slightly downregulated. The downregulation of *Id1* expression in these conditions is expected to release E12/E47 proteins and to increase Tcf15 activity. Therefore, this condition may reveal additional transcriptional differences between Tcf15-high and Tcf15-low cells (Figure 3.11C). The top and bottom 30 % of Venus-positive cells were sorted from PECAM1-positive population for further analyses.

Gene expression analysis confirmed that the negative correlation between *Tcf15* and *Nanog*/*Klf4*/*Hex* was still observed in cells sorted from 24 hours N2B27 culture (Figure 3.12A), although the difference of *Hex* expression was not significant. *Oct4* did not differ between the two populations, as had been previously observed for LIF and serum cultures (Figure 3.9). Moreover, expression of epiblast markers *Fgf5* and *T* were both enriched in Venus-high cells (Figure 3.12B) compared with cells sorting from culturing in LIF and serum (Figure 3.9E), indicating the epiblast-priming effect of Tcf15 might be stronger when Id inhibition was removed.

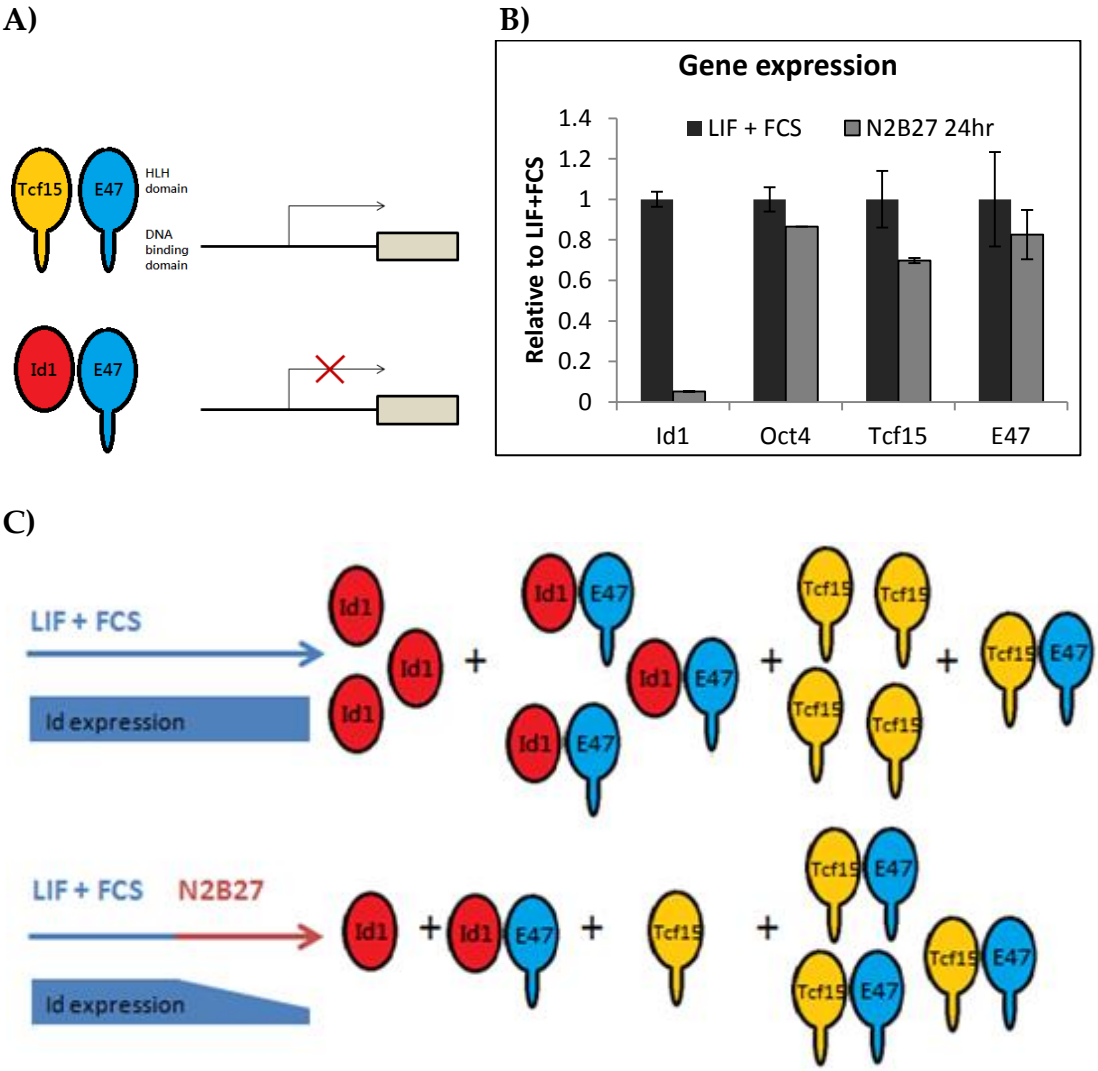


Figure 3.11: Illustration for changing Id expression level in pluripotent ES cells to induce higher Tcf15 activity

(A) The scheme for Id1 protein (showing in red) function in blocking lineage-specific transcription factor such as Tcf15 (showing in yellow) dimerization with ubiquitous E protein, E47 (showing in blue), therefore inhibits downstream gene expression.

(B) qPCR analysis of *Id1, Oct4, Tcf15* and *E47* expression after switching from culturing in LIF and serum to N2B27 for 24 hours.

(C) Graphic view showing different levels of Id expression controlling transcriptional activity of Tcf15. Id expression was decreased when cells were placed in N2B27 medium. Lower Id expression released more E47 protein to dimerize with Tcf15 and induced more Tcf15 downstream gene expression.

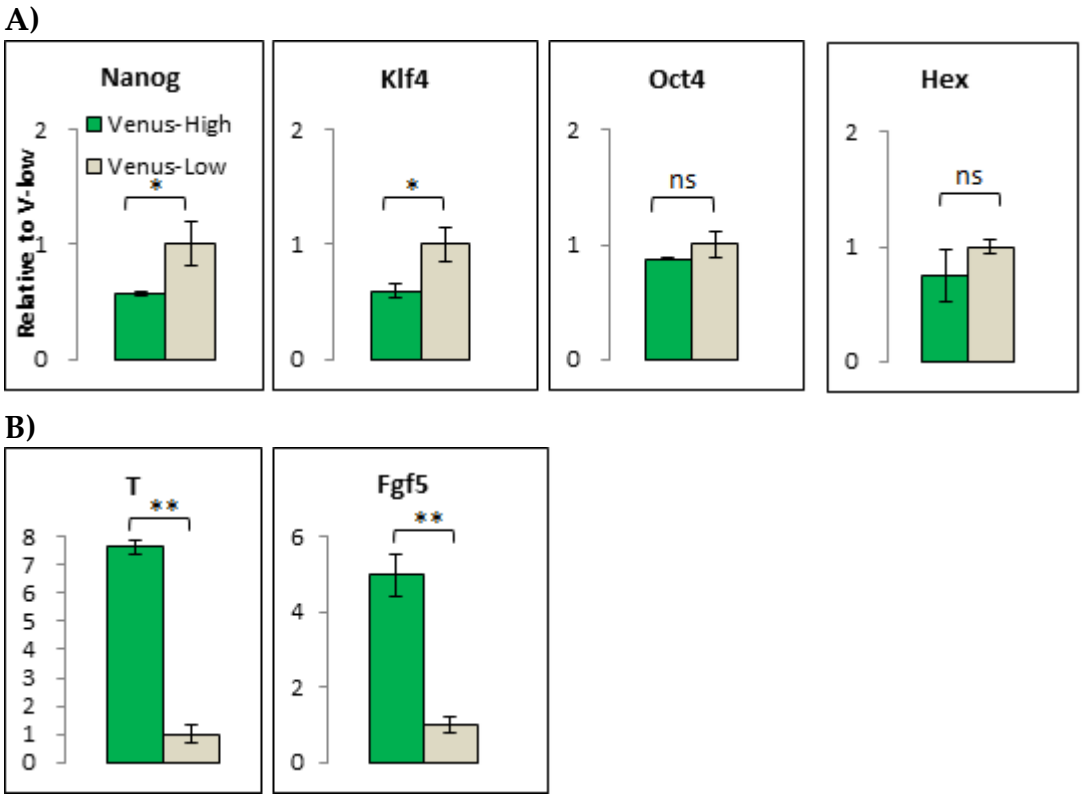


Figure 3.12: qPCR analysis of Tcf15-Venus reporter cells after down-regulation of Id1

216D1 Tcf15-Venus reporter cells were maintained in LIF and serum complete medium and switched to N2B27 basal medium 24 hours before performing FACS. In this condition, Id-induced suppression of Tcf15 transcription activity was reduced due to downregulation of Id1 expression. Gene expression was analysed from sorted Venus-high and Venus-low subpopulations. (ns: non-significant; *: $p<0.05$; **: $p<0.01$)

(A) qPCR analysis of pluripotent and extra-embryonic genes showed similar results as cells were cultured in complete medium containing LIF and serum.

(B) qPCR analysis of epiblast markers *T* and *Fgf5* showed enrichment of expression when Id inhibition was reduced.

3.2.2.3.1.4: Tcf15 expression is dynamic within self-renewing ES cells

Tcf15 was heterogeneously expressed in self-renewing ES cells and Venus-high and Venus-low populations displayed differential gene expression. In order to understand if Tcf15 heterogeneity was dynamic within ES cells, sorted Venus-high and Venus-low cells were analysed daily by flow cytometry for Venus expression until day three after sorting. E14tg2a cells were used as a negative control for non-fluorescent cell and unsorted Tcf15-Venus cells were a positive control and an indicator of the steady state distribution of Tcf15-Venus. As shown in Figure 3.13A, Venus-high and Venus-low cells were sorted from culturing in complete media containing LIF and serum. Cell purity was at least >90% immediately after sorting with no overlapping of the two populations. Cells were plated back to 6 well plates at 10^5 cells per well in complete medium containing LIF and serum plus pen/strep to prevent possible bacterial contamination after the sorting process. Cells from both Venus-high and Venus-low populations were both competent to re-establish a heterogeneous culture within 24 hours, as indicated by overlapping of Venus expression (Figure 3.13B). However, Venus-low cells contributed more Venus-negative cells after 3 days.

An alternative method was used for analysing dynamic expression of Tcf15 to test whether heterogeneous expression of Tcf15-Venus could be rapidly established in a colony of cells that emerged from a single cell. Tcf15-Venus cells were plated at clonal density and cultured in LIF and serum condition for 6 days. Heterogeneous expression of Tcf15-Venus could be regenerated within a colony (Figure 3.13C) in > 90% of the colonies, suggesting that the heterogeneous expression of Tcf15 was in a dynamic rather than a stable state.

These results indicated both Tcf15-high and Tcf15-low existed within

un-differentiated population and were able to interconvert when ES cells were cultured in self-renewing conditions.

3.2.2.3.1.5: Tcf15 does not mark cells that have committed to differentiate

The data presented above suggests that Tcf15 marks a population of cells that express low levels of naïve pluripotency markers but which can revert back to a Tcf15-low state. In order to address the question whether Tcf15 marks cells that have committed to differentiate, experiments were designed to ask whether individual Tcf15-high cells had a reduced ability to form self-renewing colonies compared with Tcf15-low cells.

Self-renewing, un-differentiated PECAM1-positive ES cells were sorted according to Venus expression. Cells were plated at single cell density and cultured in complete medium for 5 days to perform alkaline phosphatase (AP) staining: a commonly used phenotypic assessment for undifferentiated ES cells (Figure 3.14A) (Palmqvist et al., 2005). No obvious difference in colony numbers between unsorted, Venus-high and Venus-low populations could be detected.

These results suggested that ES cells cultured in LIF and serum contain both Tcf15-high and Tcf15-low populations and these two populations were inter-changeable. In addition, Tcf15-high cells were not cells that had committed to differentiation, although based on the relative lower expression of naïve pluripotency markers they are likely to be cells that are primed for differentiation (Figure 3.14B).

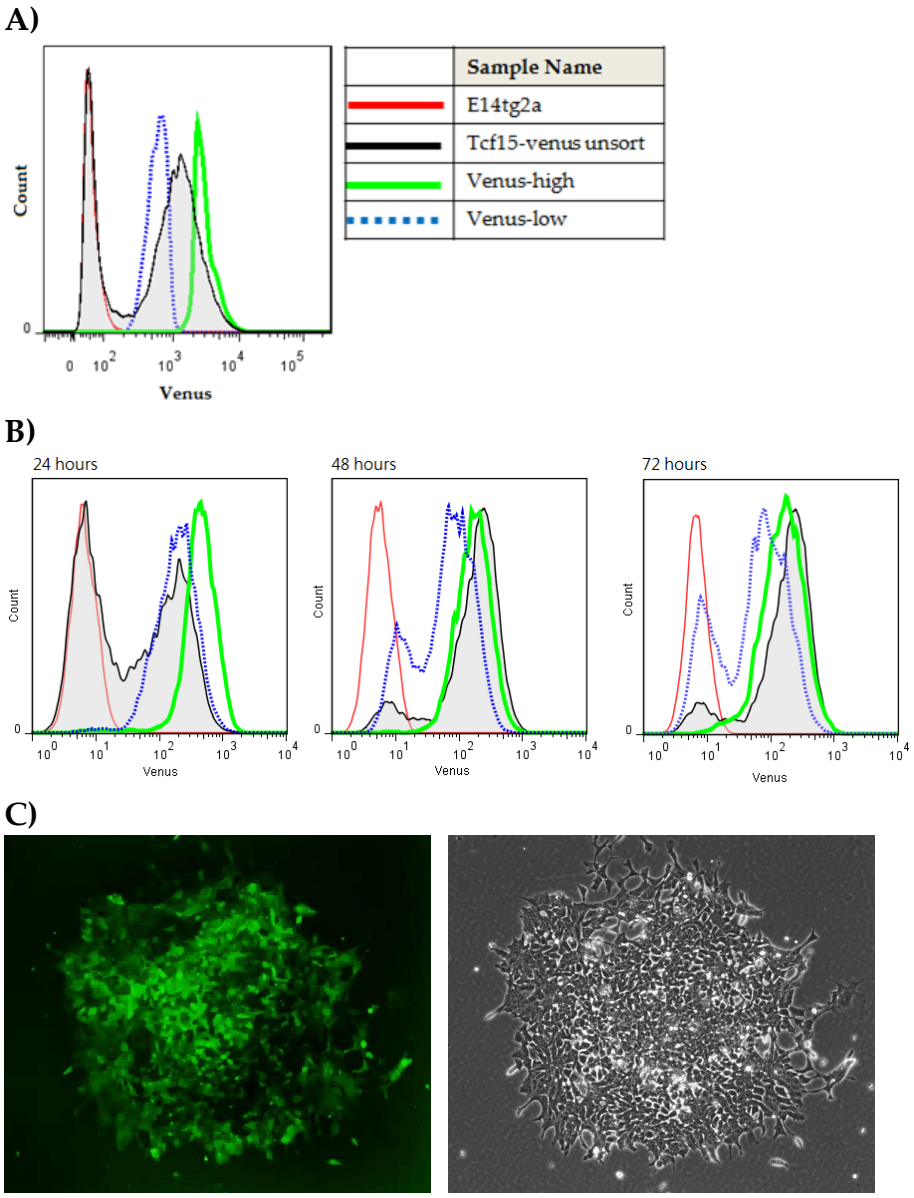


Figure 3.13: Tcf15 expression is dynamic within self-renewing culture conditions

(A) Venus expression of sorted Venus-high (solid green), Venus-low (dashed blue) cells, unsorted Tcf15-Venus control cells (solid black with shadow) and negative control E14tg2a cells (solid red).

(B) Daily flow cytometry analysis of Venus expression on FACS sorted populations. Venus-high and Venus-low cells return to heterogeneous expression within 48-72 hours.

(C) Single colony from 216D1 Tcf15-Venus, cultured from a single cell in LIF and serum for 6 days. Venus was heterogeneously expressed within the colony.

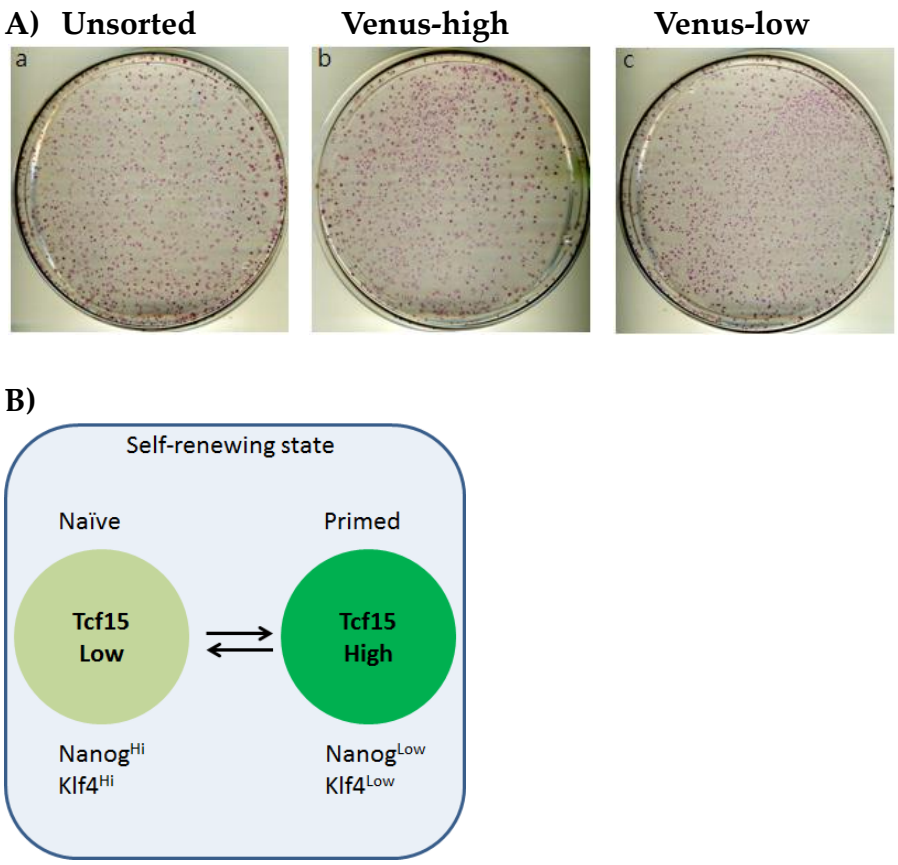


Figure 3.14: Tcf15-high cells were not early-differentiating cells

(A) Alkaline phosphatase staining of unsorted (a), Venus-high (b) and Venus-low (c) populations after 5 days culture in complete medium with LIF and serum.

(B) Model for Tcf15 heterogeneity within self-renewing ES cells. Self-renewing ES cells contain both Tcf15-high and Tcf15-low populations and these two populations were inter-changeable.

3.2.2.3.1.6: Tcf15 marks cells that are primed for rapid differentiation

To test whether Tcf15-Venus-high cells are functionally primed for differentiation, FACS sorted Tcf15-Venus reporter cells were directed to differentiate into neural cells using monolayer neural differentiation protocol (Ying et al., 2003). The extent of differentiation into the neural lineage was then assessed daily by qPCR analysis. As shown in figure 3.15, *Tcf15* expression was down-regulated during neural differentiation and this down-regulation was more rapid when differentiation was initiated from a Venus-high state. *Oct4* expression was also down-regulated during neural differentiation, slightly more efficiently in Venus-high cells, although these differences are not significant. *Nanog* was lost rapidly from both populations but the neural markers *N-Cadherin* and *Sox1* were up-regulated more rapidly when differentiation was initiated from a Venus-high state. These data suggest that the Tcf15-Venus-high subpopulation is primed to undergo neural differentiation more efficiently than the Tcf15-Venus-low subpopulation. The more rapid down-regulation of *Tcf15* from the Venus-high state is therefore likely to reflect the down-regulation of *Tcf15* during differentiation (Figure 3.5A).

Next, to further examine if Tcf15-high cells progress more rapidly toward differentiation into other lineages, 3-dimentional embryoid body (EB) aggregates were generated in hanging drops in order to encourage multi-lineage differentiation. qPCR analysis was performed from day 1 to day 5 to monitor differential gene expression during early EB formation (Figure 3.16). *Tcf15* was downregulated rapidly during EB formation. Venus-high cells also down-regulated *Oct4* expression slightly more efficiently, although these differences are not significant. Expression of epiblast marker *Fgf5* was up-regulated in Venus-high cells more rapidly at

day one. Both Venus-high and Venus-low cells up-regulated *Sox1* (neuroectoderm), *T-Brachyury* (mesoderm) and *FoxA2* (endoderm) expression, indicating there was no differential bias of either population for commitment to any specific lineage. However, expression of *Sox1* and *T-Brachyury* in Venus-low cells were about 24 hours delayed compared with Venus-high cells. The re-activation of *Nanog* during EB formation was also accelerated by around 24 hours: this peak of *Nanog* expression is thought to recapitulate the formation of proximal posterior epiblast prior to gastrulation (Osorno et al., 2012, Malaguti et al., 2013).

Taken together, these data suggest that Tcf15 marks cells that have low expression of naïve markers, are not committed to differentiate, but are primed to differentiate more rapidly.

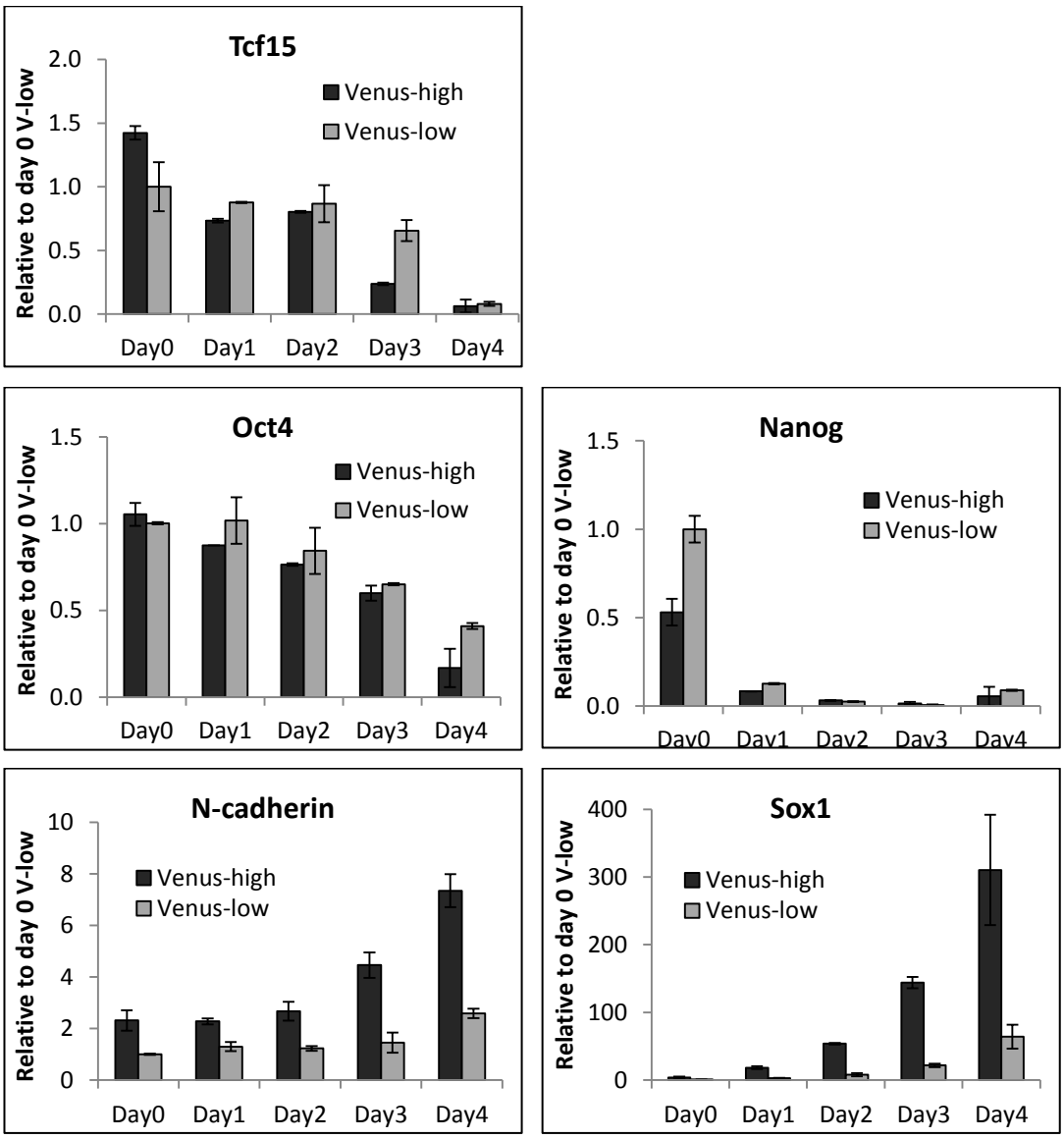


Figure 3.15: qPCR analysis of FACS sorted Venus-high and Venus-low cells during neural differentiation

FACS sorted Tcf15-Venus reporter cells were plated into N2B27 medium containing penicillin/streptomycin after sorting to perform neural differentiation with daily media change. Venus-high subpopulation progresses more rapidly toward neural differentiation.

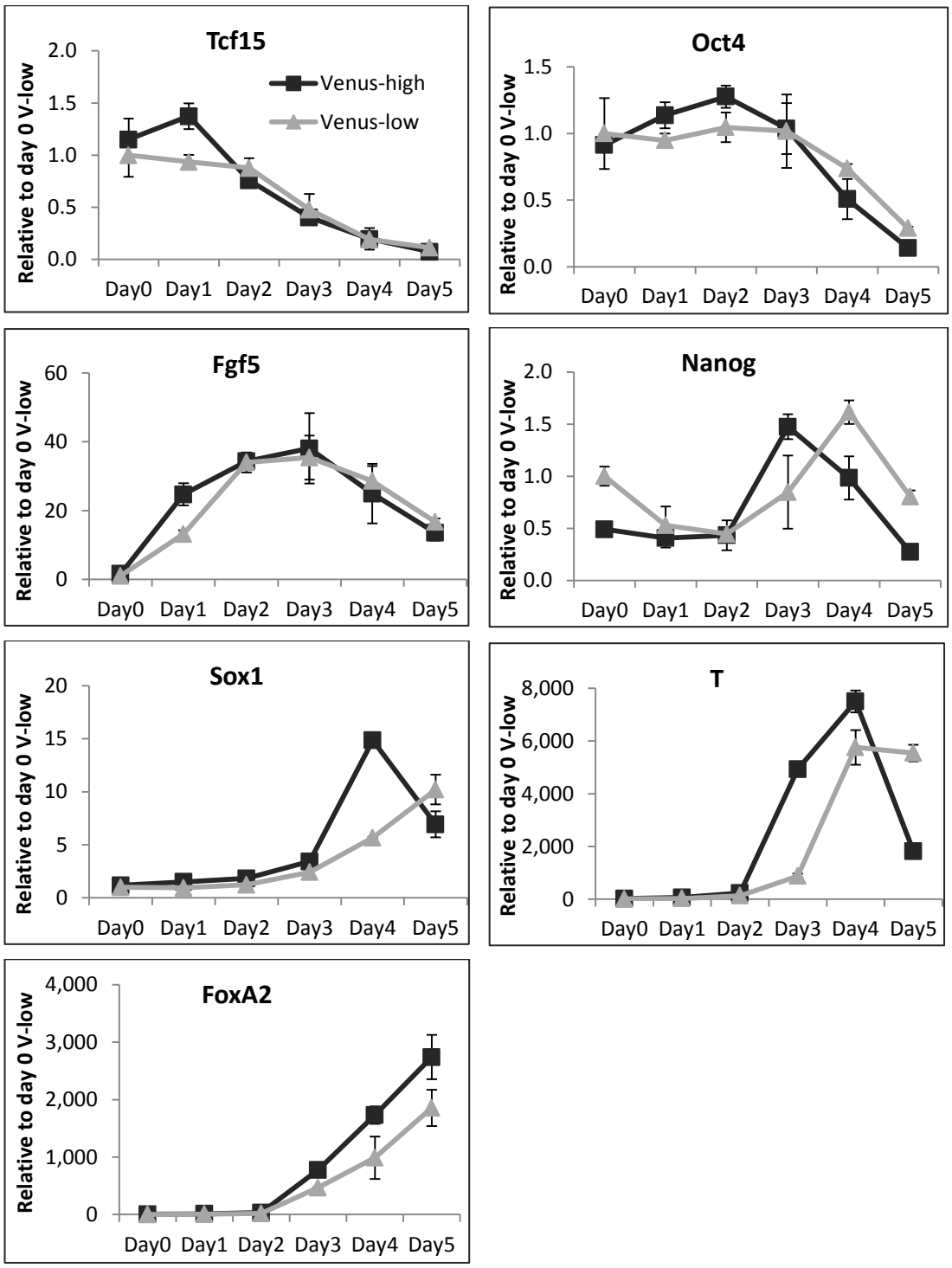


Figure 3.16: qPCR analysis of FACS sorted Venus-high and Venus-low cells during EB formation

qPCR analysis of FACS sorted Tcf15-Venus reporter cells underwent hanging drop EB formation. Venus-high subpopulation displayed rapid differentiation than Venus-low subpopulation without bias to any specific lineage. This experiment was performed twice.

3.2.2.3.2: Silencing effect within 216D1 Tcf15-Venus cells

The 216D1 Tcf15-Venus cells served as a good tool for analysing Tcf15 expression. However, there were some concerns after extensive studies on this reporter cell line. First, a Venus-negative population within pluripotent culture aroused after long-term (>15 passages) culture (Figure 3.17A). The sorted Venus-negative cells expressed intermediate level of *Tcf15* between Venus-high and Venus-low cells and almost undetectable *Venus* expression (Figure 3.17B). This was first considered to be due to loss of the gene-trap cassette. However, genomic DNA qPCR analysis showed that Venus-negative cells still retained the Venus selection cassette (Figure 3.17C) but somehow lost mRNA and protein expression. This indicated that the Venus-negative population is likely to contain cells in which the gene-trapped allele has been silenced by an unknown epigenetic regulation.

Furthermore, single cell sorting of Venus-positive cells was performed in order to re-clone the pure cells. However, after 15-20 passages culturing in the presence of G418 selection, the Venus-negative cells arise again. In addition, as shown in figure 3.13A above, the Venus-positive population was dynamic which Venus-high cells could regenerate Venus-low cells and vice versa. However, the Venus-negative population never switch-on Venus expression after sorting (Figure 3.17D) and could be maintained in pluripotent LIF and serum cultures for at least 10 passages. We concluded that the gene trap allele is reproducibly prone to irreversible silencing at a low but significant frequency.

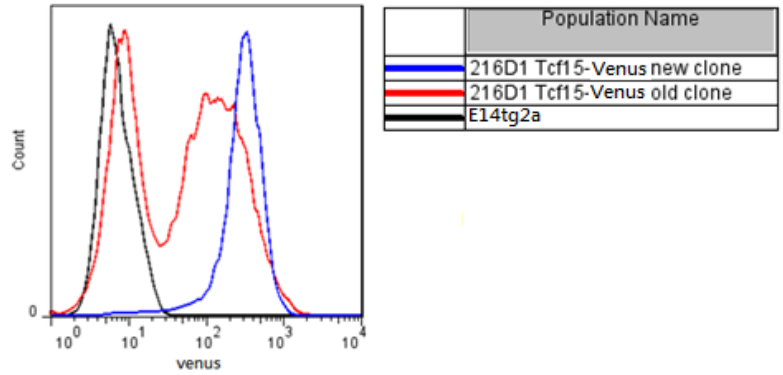
A second concern about this cell line came from the use of *gtx-IRES-Venus* cassette as a translational amplifier. This amplifier generates high levels of fluorescence, which is convenient for monitoring Venus expression in live cells. However, the strong amplification of Venus protein

expression under this amplifier might mask some subtle changes within cells and raises the concern that this reporting strategy may not be as strictly faithful as a strategy that relies entirely on an endogenous promoter in the absence of translational amplification. Therefore, I decided to make another Tcf15-Venus reporter by targeting Venus directly into the *Tcf15* locus.

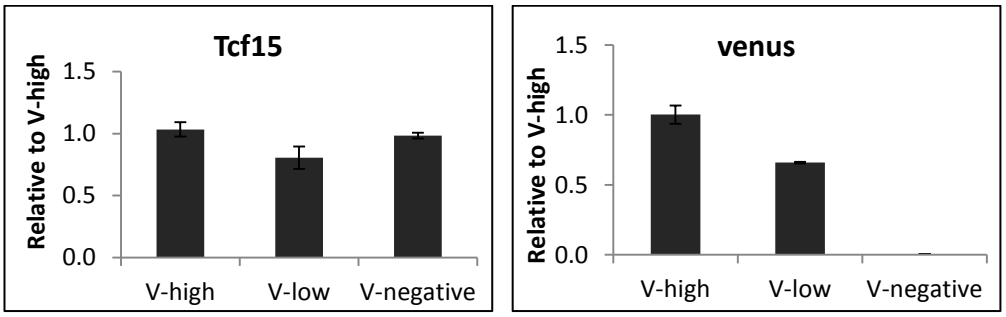
Figure 3.17: Analysis of Venus-negative population within 216D1 cells

(A) Flow cytometry analysis of Venus expression in 216D1 Tcf15-reporter cells. "Old clone' refers to cells cultured for more than 20 passages. "New clone" refers to cells re-cloned from single cell sorting from Venus-positive population.
(B) qPCR analysis of sorted Venus-high, Venus-low and Venus-negative cells from old 216D1 Tcf15 reporter cells. (N=2, n=2)
(C) qPCR analysis of gDNA isolated from sorted Venus-positive and Venus-negative population. Two different endogenous housekeeping controls were used to normalize genomic *Venus* expression. (N=2, n=2)
(D) Flow cytometry analysis of Venus-negative cells after FACS. Sorted cells maintain negative for Venus expression after 72 hours of culture.

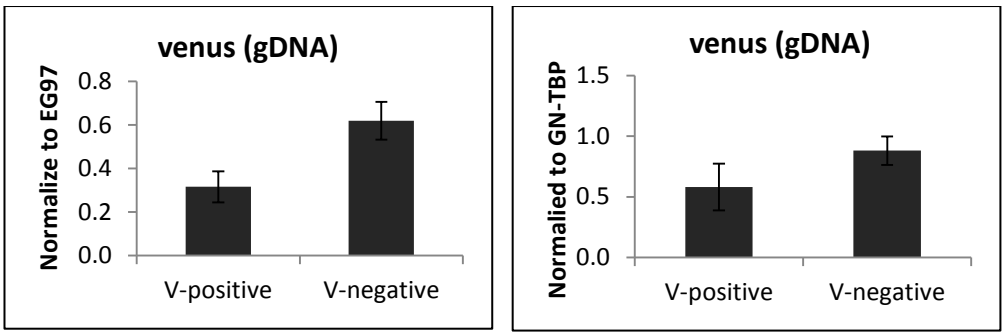
A)


Population Name
216D1 Tcf15-Venus new clone
216D1 Tcf15-Venus old clone
E14tg2a
Count
venus

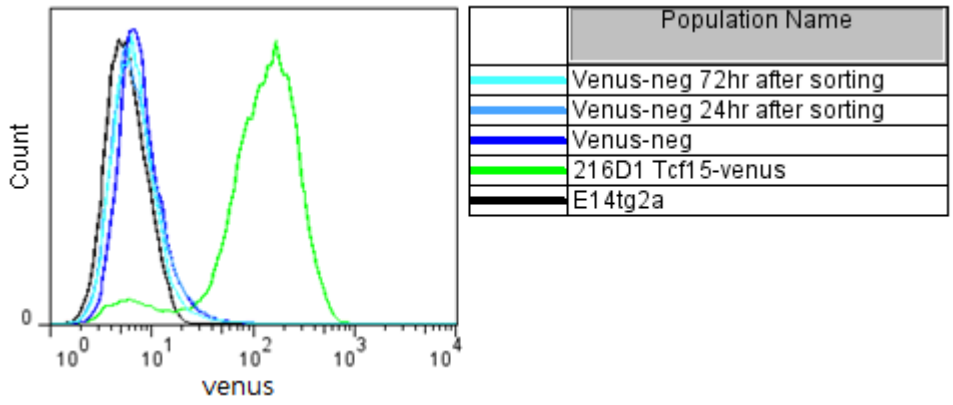
B)


Tcf15
Relative to V-high
1.5
1.0
0.5
0.0
V-high
V-low
V-negative
venus
Relative to V-high
1.5
1.0
0.5
0.0
V-high
V-low
V-negative

C)


venus (gDNA)
Normalize to EG97
0.8
0.6
0.4
0.2
0.0
V-positive
V-negative
venus (gDNA)
Normalied to GN-TBP
1.5
1.0
0.5
0.0
V-positive
V-negative

D)


Population Name
Venus-neg 72hr after sorting
Venus-neg 24hr after sorting
Venus-neg
216D1 Tcf15-venus
E14tg2a
Count
venus

3.2.2.3.3: Analyzing Tcf15 expression pattern in mES cells by Tcf15-Het (*Tcf15*$^{+/Venus}$) reporter cells

3.2.2.3.3.1: Generating *Tcf15* heterozygous Venus knock-in cells as a new Tcf15-reporter cell line

A targeting vector for generating Tcf15-Venus reporter cells was previously constructed in our lab by Aliaksandra Radzisheuskaya. Briefly, the targeting strategy was to remove exon 1 of *Tcf15* allele and replace it with the fluorescence protein Venus followed by a *FRT*-flanked double selection cassettes containing a *PGK-neomycin* resistant gene for positive selection with G418 and a *MC1-TK* (*thymidine kinase*) for negative selection with ganciclovir (Figure 3.18). The targeting construct also contained a negative selection marker *diphtheria toxin A* (*DTA*) outside of the homology region to perform negative selection of non-homologous integrations. However, this targeting vector was found to inadvertently lack the coding sequence for placing a *poly(A)* tail after the *Venus* gene and was therefore likely to result in an unstable *Venus* mRNA and low protein expression. Therefore, the *poly(A)* sequence was inserted after the *Venus* gene through *Bsr*GI site to obtain the final targeting construct, which been designated pSKO-Tcf15 (SKO stands for Straight Knock Out to indicate that this construct knocked out the first exon of *Tcf15* allele).

*Sca*I-linearized pSKO-Tcf15 (100 μg) was then electroporated into wild-type 129/Ola-derived mES cells (E14tg2a), and transfectants were selected with G418 (300 μg/ml) yielding over 10^3 colonies by day 8 of selection. Approximately 190 clones were picked and expanded to 96 well plates. Clones were expanded to 48 well plates whereupon replica plates were made and frozen at -80°C. Cells were pre-screened for Venus

expression before lysing for genomic DNA isolation. Genomic DNA of 83 clones with positive Venus expression under microscopic examination were first isolated with the others frozen at -80°C. Southern blot was used for identifying corrected targeting clones. Correct targeting introduced an exogenous *Eco*RV site into the region of exon 1 replaced by the selection cassette, which allowed both wild-type and targeted alleles to be distinguished from each other by southern blot screening. A Tcf15-3' external probe identified 2 bands, 9.5 kb and 8.3 kb, corresponding to wild-type and targeted alleles, respectively. An internal probe was also designed for identifying potential clones with random-integrated targeting construct. Random-integrated clones will contain additional bands other than the known 8.3 kb targeted allele.

Southern blot analysis identified 10 heterogeneous clones (*Tcf15*$^{+/Venus}$) with both wild-type and correct targeted band by Tcf15-3' external probe (Figure 3.19A), giving a targeting efficiency of approximately 5%. Among these clones, clone E14 had an additional random-integrated fragment revealed by a 4.4 kb band after blotting with internal probe (Figure 3.19B). Therefore, 9 independent *Tcf15* heterozygous clones were obtained, which were clones A29, D4, D22, D34, D43, E5, E6, E39 and E48 (now referred to as Tcf15-Het-A29, Het-D4, Het-D22, Het-D34, Het-D43, Het-E5, Het-E6, Het-E39 and Het-E48, respectively). More detailed characterisation and phenotypic analysis of Tcf15-Het cells will be described in section **4.2.2.2.1**.

Tcf15-Het-E6-1 cells, in which the selection cassette had been successfully removed, displayed a normal complement of 40 chromosomes after karyotyping analysis (Figure 3.20). Therefore, this clone was selected as new Tcf15 reporter for further studies.

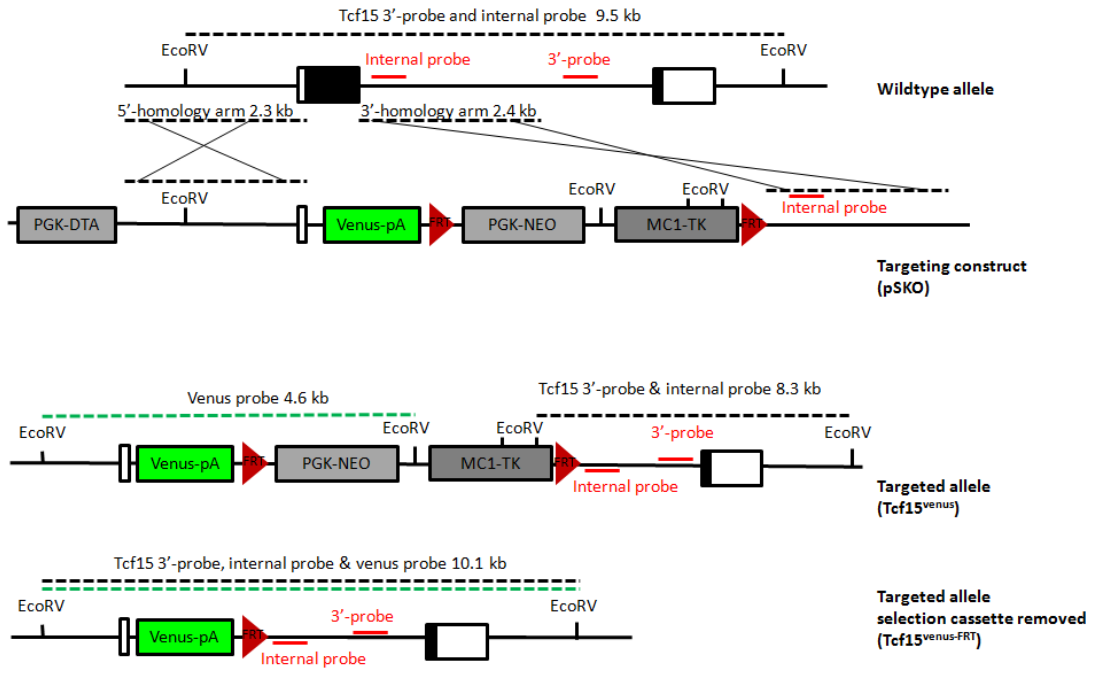
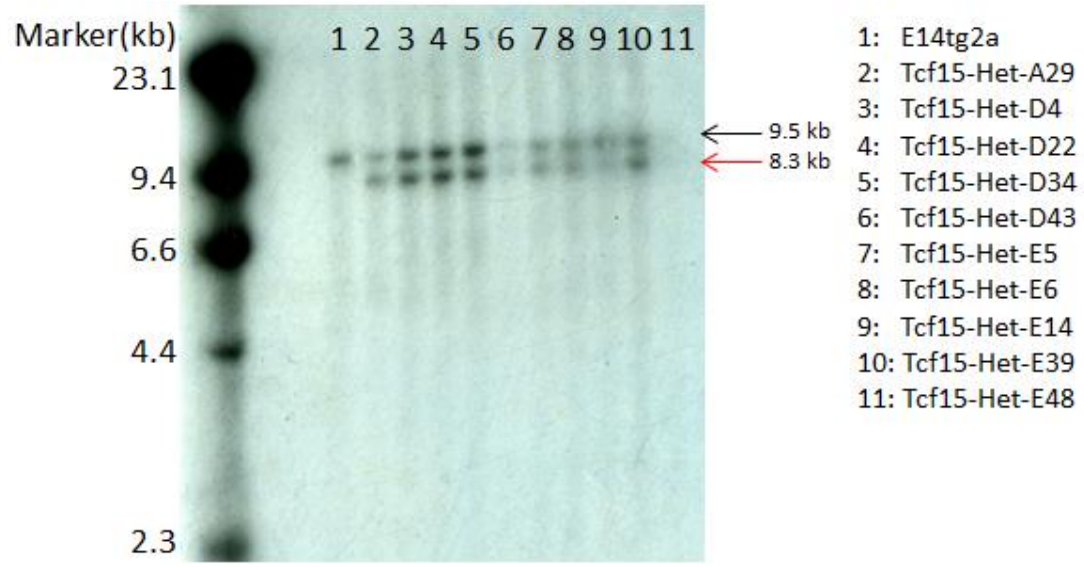


Figure 3.18: Schematic view of the wild type allele (*Tcf15*), *Tcf15* gene-targeting vector (pSKO), targeted allele (*Tcf15*^{Venus}) and targeted allele after removing the selection cassette (*Tcf15*^{VenusFRT})

pSKO *Tcf15* gene-targeting vector contained *Venus* (shown in green), a *PGK-Neo* selectable marker and a negative selection marker *MC1-TK* flanked by *FRT* sties. Exon 1 in the wild-type *Tcf15* allele was replaced to generate the *Tcf15*^{Venus} allele. After FlpO recombinase mediated recombination, the selection cassette could be excised, leaving a *FRT* site at the allele *Tcf15*^{VenusFRT}. Genomic DNA digestion with *Eco*RV followed by Southern blotting, using the Tcf15-3' external probe and internal probe depicted in the diagram, resulted in a 9.5 kb band for the wild-type allele, and a 8.3 kb band for the *Tcf15*^{Venus} targeted allele. After removing the selection cassette, *Eco*RV digestion of *Tcf15*^{VenusFRT} allele will give rise to a 10.1 kb band by Southern blotting. The *Tcf15*^{VenusFRT} allele could also be easily distinguished from the *Tcf15*^{Venus} allele by probing with Venus probe, which gave rise to 10.1 kb and 4.6 kb band, respectively.

A) Tcf15-3′ external probe



B) Tcf15 internal probe

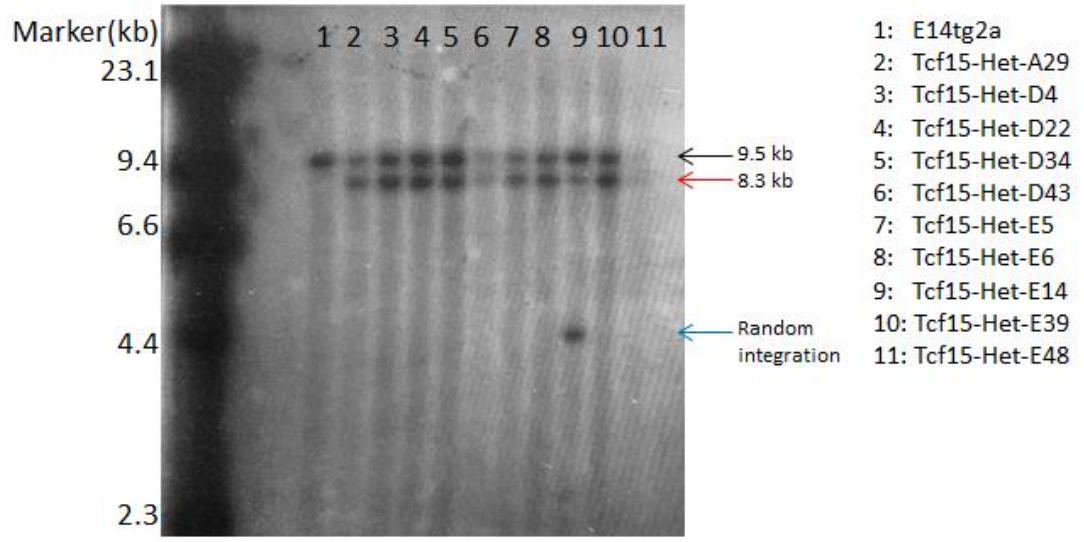


Figure 3.19: Southern blot analysis of *Tcf15* targeted cells

E14tg2a ES cells were electroporated with the pSKO targeting construct. Genomic DNA was isolated and digested with *Eco*RV to perform a Southern blot using the targeting probes depicted in figure 3.18 on G418-resistant, Venus-positive clones.

(A) Clones A29, D4, D22, D34 ,D43, E5, E6, E14, E39 and E48 appeared to be correctly targeted by pSKO as they had the wild-type band (black arrow) and a second band from targeted allele (red arrow).

(B) Tcf15 internal probe identified clone E14 had an additional random integrated fragment of targeting vector (blue arrow), therefore was excluded from further analysis.

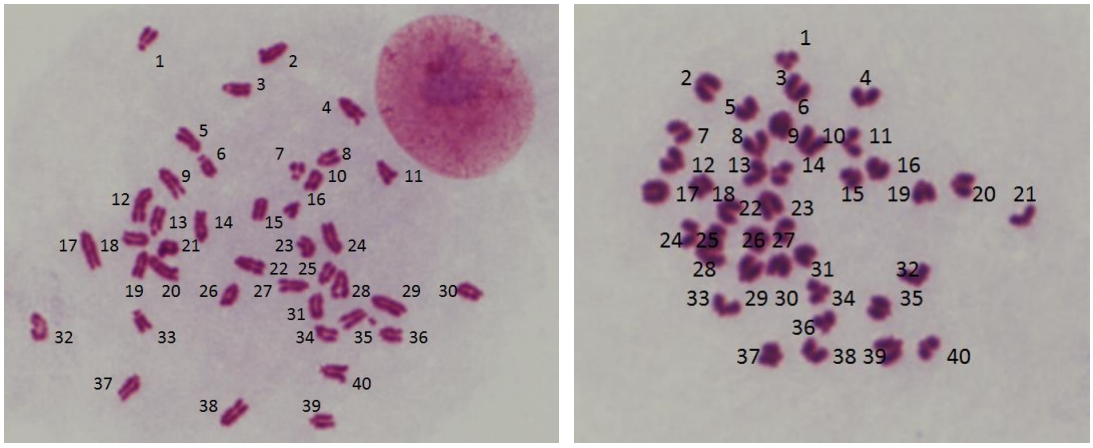


Figure 3.20: Karyotyping of Tcf15-Het-E6-1 cells

Karyotyping of Tcf15-Het-E6-1 cells was carried out by treating cells with colcemid for 2 hours to arrest cells in metaphase. Cells were then dissociated into single cell suspension, fixed, and stained with Giemsa to reveal chromosomes. This figure showed two representative images for chromosome counting. Mouse ES cells should have 40 chromosomes.

3.2.2.3.3.2: Expression pattern of Tcf15-Het cell lines

After obtaining the new line of Tcf15-Het reporter cells, *in vitro* and *in vivo* studies were performed to verify if Tcf15-Hets were faithful Tcf15 reporters. Clone E6-1 (now referred to as Tcf15-Het cells) was selected for further analyses.

3.2.2.3.2.2.1: *In vitro* expression pattern of Tcf15-Het cells is similar to that of the 216D1 Tcf15-Venus cells

Tcf15-Het cells express Venus protein driven by the endogenous *Tcf15* promoter with no translational amplification, and so as expected the Venus expression was lower than observed in the 216D1 Tcf15-Venus reporter which Venus protein expression was amplified by the *gtx-IRES* element (Figure 3.21A). Immuno-staining of Venus expression on Tcf15-Het cells plating at clonal density in complete medium containing LIF and serum showed Venus was heterogeneously expressed (Figure 3.21B), as expected from analysis of 216D1 Tcf15-Venus reporters (Figure 3.7A). Expression of Venus was gradually lost when cells spontaneous differentiated towards the edge of the colony, as expected. Also, when culturing in 2i medium for 48 hours, expression of Venus dramatically decreased (Figure 3.21C), consistent with the down-regulation of *Tcf15* mRNA under 2i conditions (Figure 3.4A). Furthermore, confocal microscopic analysis followed by quantification demonstrated that Venus expression was negatively correlated with Nanog expression but not with Oct4 (Figure 3.22), as had been previously observed in the 216D1 Tcf15-Venus cells (Figure 3.8). This confocal microscopic analysis was done by undergraduate project student Julia Oh under my supervision.

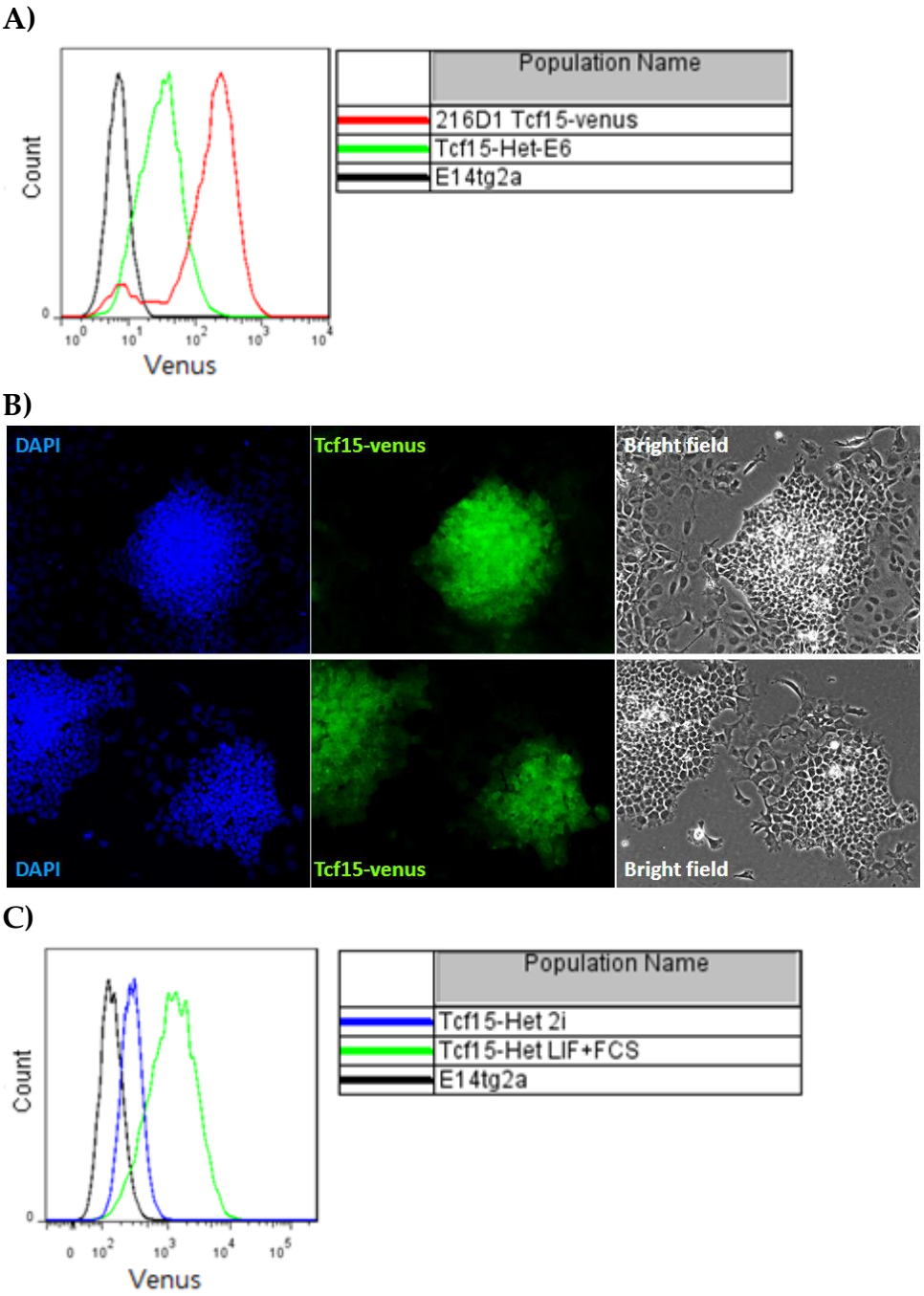


Figure 3.21: Venus expression pattern of Tcf15-Het-E6-1 cells

(A) Flow cytometry analysis was used to compare Venus intensity between 216D1 Tcf15-Venus cells and Tcf15-Het cells.

(B) Immunostaining of Tcf15-Het cell plating at clonal density for 5 days. Venus was heterogenously expressed within colony and expression was loss when cell spontaneously differentiated outward the outside of the pluripotent colony (differentiation is assessed here by morphology).

(C) Flow cytometry analysis of Tcf15-Het cells culturing in LIF + FCS or 2i medium.

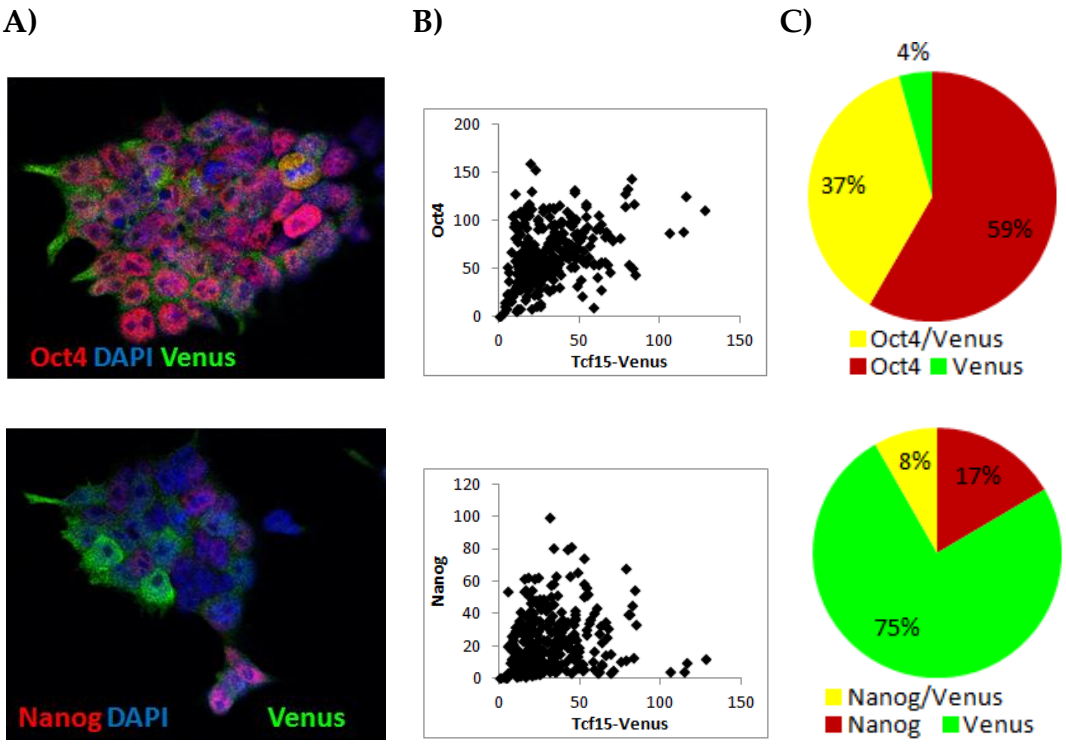


Figure 3.22: Venus expression in Tcf15-Het cells was negatively correlated with Nanog, but not with Oct4 expression

(A) Confocal microscopic analysis of Tcf15-Het cells culturing in LIF and serum, staining with Oct4 (top) and Nanog (bottom).

(B) Quantitative immunofluorescence analysis of Tcf15-Het reporter cells stained for Oct4 and Nanog.

(C) Proportion of cells positive for Tcf15-Het Venus only (green), Oct4 or Nanog only (red), or double positive (yellow).

To further analyse the population heterogeneity, Tcf15-Het cells were sorted by similar gating strategy as described in section **3.2.2.3.1.1**. Briefly, cells were first gated with FSC and SSC for size and granularity, followed by cell doublets exclusion. Then, a live-dead gate was carried out to identify DAPI-negative live cells and PECAM1-satining to exclude differentiated cells (Figure 3.23A). Tcf15-Het cells were sorted into Venus-high and Venus-low populations by the top and bottom 30% of cells (Figure 3.23B). qPCR analysis of FACS sorted Venus-high and Venus-low cells showed a clear correlation between *Venus* and *Tcf15* (Figure 3.23C). Furthermore, expression of pluripotency and differentiation markers showed a similar distribution between the two populations as had been previously observed for the 216D1 Tcf15-Venus cells (Figure 3.7 and 3.9). For example, expression of naïve pluripotency genes *Nanog* and *Klf4* and extra-embryonic endoderm marker *Hex* were lower in Venus-high cells and epiblast marker *Fgf5* was higher in Venus-high cells, but there was no difference between the two populations in expression of *Oct4* (Figure 3.23C).

Venus expression within Tcf15-Het cells also displayed dynamic pattern. FACS sorted Venus-high and Venus-low cells re-establish heterogeneous population within 24 hours after sorting when culturing in self-renewing medium containing LIF and serum (Figure 3.23D). Furthermore, FACS sorted cells were differentiated under monolayer neural differentiation. qPCR analysis showed that Venus-high population in the Tcf15-Het cells up-regulated *Sox1* expression more quickly than Venus-low population (Figure 3.23E), suggesting Venus-high cells in Tcf15-Het were also primed for differentiation more rapidly.

Taken together, the *in vitro* expression of Tcf15-Het cells is similar in all the above-mentioned respects to that of the 216D1 Tcf15-Venus reporter cells,

suggesting that Tcf15-Het is a faithful reporter. Furthermore, although slightly less sensitive, the use of endogenous promoter to drive Venus expression in Tcf15-Het cells made it less subject to gene silencing. Also, the consistent results between Tcf15-Het and 216D1 Tcf15-Venus cells also confirmed the data presented earlier in this chapter.

Figure 3.23: Venus expression in Tcf15-Het cells compared to 216D1 cells

(A) Sequential FACS gating of Tcf15-Het cells and the control E14tg2a cells.

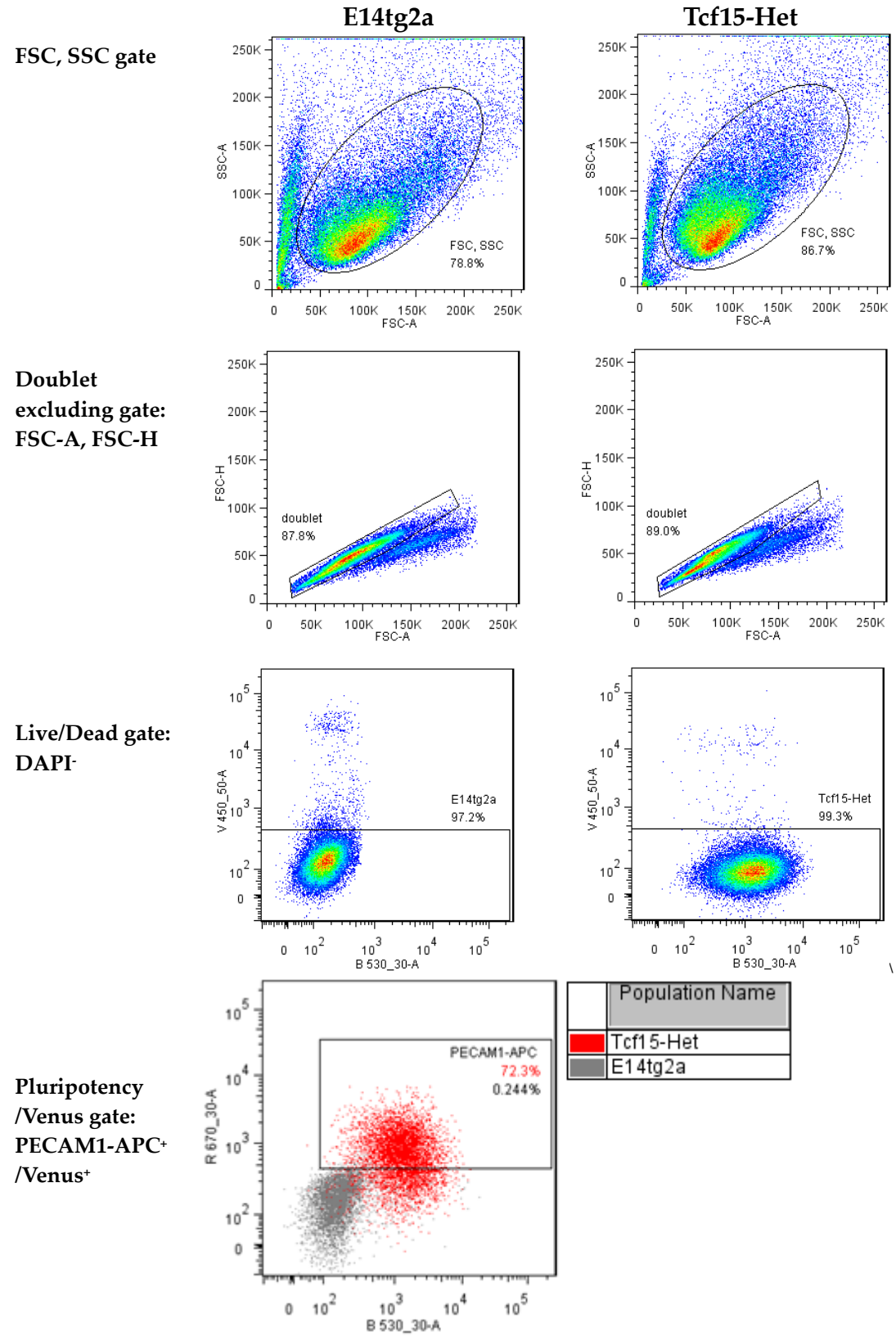
(B) FACS profile of Tcf15-Het reporter line after PECAM1^{+} gating. E14tg2a cells were used as negative control (black). Top and bottom 30% of Tcf15-Het cells (green) were sorted as indicated in the figure. Venus-high (blue) and Venus-low (red) populations were analysed right after sort for purity check.

(C) qPCR analysis of FACS sorted Tcf15-Het cells. Venus-high cells displayed lower *Nanog* and *Klf4* expression and higher *Fgf5* expression with no significant difference of *Oct4* expression when compared with Venus-low cells.

(D) Flow cytometry analysis of Tcf15-Het cells right after sorting (left) and 24 hours culturing in LIF and serum after sorting (right). Both population re-generate overlapping expression of Venus with 24 hours after sorting.

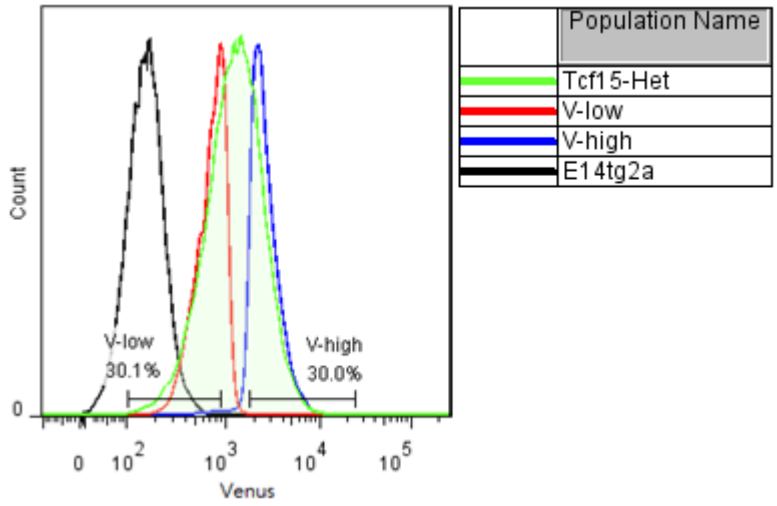
(E) qPCR analysis of FACS sorted Tcf15-Het cells under neural differentiation. Venus-high cells differentiated to neuronal lineage more rapidly, as marked by *Sox1* expression.

A)

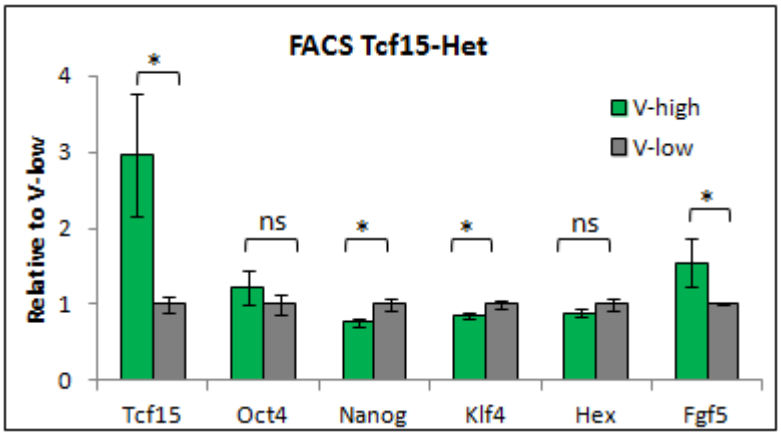


(Figure continued to next page)

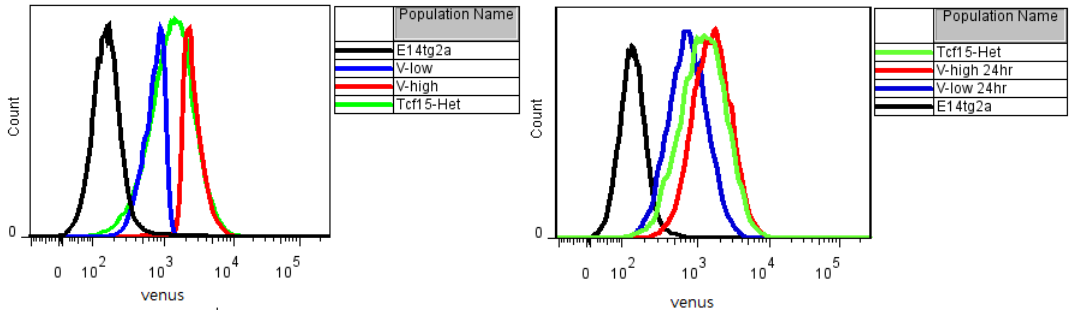
B)


Population Name
Tcf15-Het
V-low
V-high
E14tg2a
Count
V-low
30.1%
V-high
30.0%
Venus

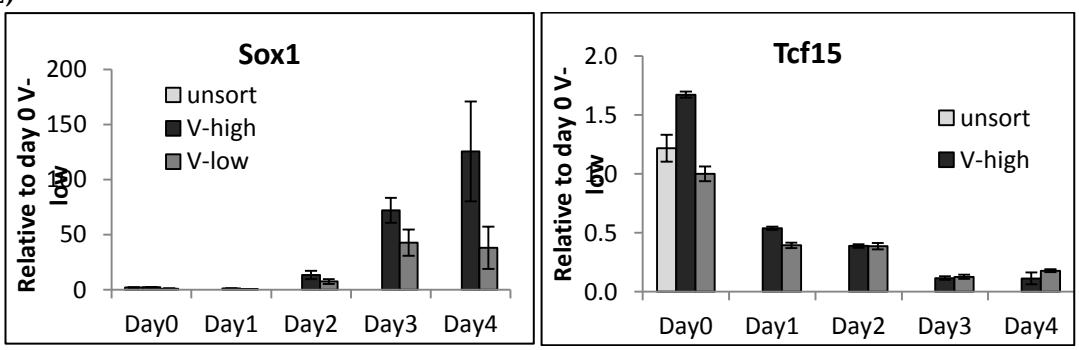
C)


FACS Tcf15-Het
Relative to V-low
V-high
V-low
ns
Tcf15
Oct4
Nanog
Klf4
Hex
Fgf5

D)


Population Name
E14tg2a
V-low
V-high
Tcf15-Het
Count
venus
Population Name
Tcf15-Het
V-high 24hr
V-low 24hr
E14tg2a
Count
venus

E)


Sox1
Relative to day 0 V-low
unsort
V-high
V-low
Day0
Day1
Day2
Day3
Day4
Tcf15
Relative to day 0 V-low
unsort
V-high
Day0
Day1
Day2
Day3
Day4

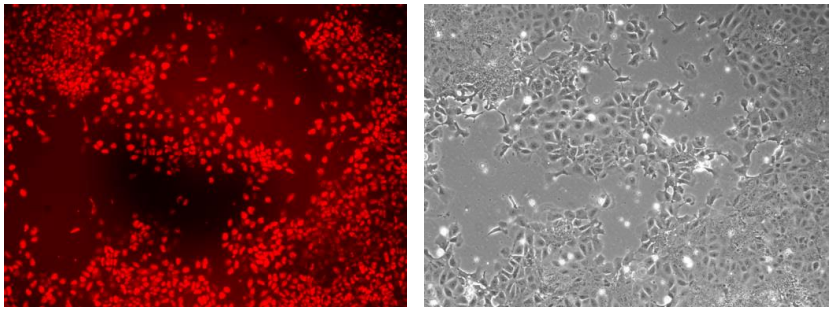
3.2.2.3.2.2.2: Lineage-labelled Tcf15-Het cells recapitulate somitic expression of Tcf15 *in vivo*

In order to further validate if Tcf15-Het cells were faithful reporters of Tcf15 expression, an *in vivo* expression analysis was carried out by using the Tcf15-Het cells to generate high-contribution chimaeric embryos. In order to identify the Tcf15-Het cells within chimaeras, Tcf15-Het cells were transfected with pCAG-mKate2-NLS (NLS stands for Nucleus Localization Signal) in order to obtain nuclear mKate2 protein expression as a lineage label. mKate2 is a high-brightness, low toxicity far-red fluorescent protein suitable for imaging in living tissues (Shcherbo et al., 2009). Transfectants were selected with puromycin and 12 clones were picked and amplified. These cells lines were designated Tcf15-Het-mK2. In order to test whether the lineage label persisted in differentiated cells, the clones were directed to differentiate by LIF withdrawal condition for 4 days. Clone #2 was selected for morula aggregation (Figure 3.24A) based on its maintenance of high mKate2 expression within differentiated cells.

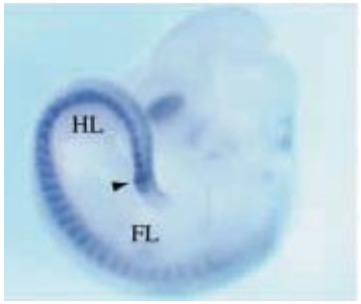
Morula aggregation of Tcf15-Het-mK2 cells was performed with the help of the in house transgenic unit. Chimaeric embryos were dissected at somitogenesis stage on E10.5, when *Tcf15* was known to be expressed specifically in the somites (Kokubu et al., 2004) (Figure 3.24B). The Tcf15-Het-mK2 cells contribute widely within the transgenic embryo, as shown by mKate2 expression (Figure 3.24C). Moreover, the expression of Venus was specifically located in the developing somites (Figure 3.24D), perfectly recapitulating the somitic expression of Tcf15 as reported by *in situ* hybridization.

These results, observed with both *in vitro* and *in vivo* analyses, suggest that Tcf15-Het cells are faithful reporters for monitoring Tcf15 expression.

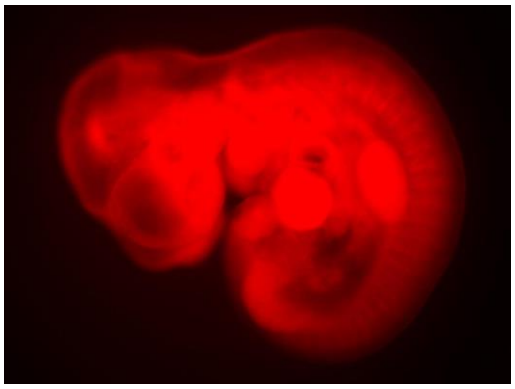
A)



B)



C)



D)

Figure 3.24: Tcf15-Het cells recapitulate the somitic expression of Tcf15 in mouse embryo

(A) Lineage labelled Tcf15-Het-mKate2 cells after 4 days culture in complete medium without LIF. mKate2 remained highly expressed within nucleus of differentiated cells, as indicated by cell morphology in bright field.
(B) ISH of E10.5 mouse embryo, taken from Kokubu et al., 2004.
(C) Transgenic embryo at E10.5, mKate2 expression showed that Tcf15-Het cells contribute widely to the whole embryo.
(D) Transgenic embryo at E10.5 displayed Tcf15-Venus expression in the developing somites.

3.3: Discussion

3.3.1: Tcf15 expression in mouse blastocysts

In this chapter, the *in vivo* expression of *Tcf15* in E4.5 pre-implantation mouse blastocysts was identified. Using *in situ* hybridization, *Tcf15* expression was specifically detected in the ICM compartment, consistent with its expression in the *in vitro* counterpart, ES cells. However, ICM consists of both epiblast and primitive endoderm (PrE) cells, which were distinct cell types but are not readily distinguishable when using ISH analysis. It would be interesting to verify whether *Tcf15* is specifically expressed in the epiblast by generating reporter mouse line using the new Tcf15-Het reporter cells which faithfully reported the somitic expression of Tcf15 *in vivo* (Figure 3.24D). Generation of these reporter mouse lines is currently in progress. Our preliminary data indicate that our Tcf15-Het (clone E6-1) cells have contributed well to chimaeric mice (10/14 upon blastocyst injection, data not shown) and we are currently awaiting confirmation of germ line transmission.

Once the Tcf15 reporter mice become available, immunostaining on the reporter blastocysts could be performed, and it would be possible to find out whether Tcf15 expression is restricted to a particular compartment of the ICM. In addition, although it had been described by single cell qPCR analysis that *Tcf15* is predominantly associated with the *Gata6*-negative subpopulation during the E3.5 ICM outgrowth (Figure 3.3), only very few cells were analysed in the study (Tang et al., 2010). Therefore, the previous RNA expression data could be further confirmed by perform protein-level analysis to investigate if Tcf15-Venus expression is associated with Nanog-low and Gata6-low population *in vivo*.

Two pieces of data presented in this chapter suggested that Tcf15 is expressed at significantly lower level in post-implantation embryos then in pre-implantation embryos: qPCR analysis *in vivo* showed expression of *Tcf15* declined dramatically after embryo implanted around E5.5 (Figure 3.1) and post-implantation EpiSCs expressed lower levels of *Tcf15 in vitro* (Figure 3.4). However, it could not be ruled out the possibility that Tcf15 is only expressed in a small subset of cells in the post-implantation embryos (e.g. epiblasts). It would be therefore informative to perform detailed investigation on expression of Tcf15 within post- implantation embryos at the single cell level using Tcf15-Venus chimaera. Two ES cell lines could be used for the morula aggregation experiments to analyse the expression of Tcf15 *in vivo*: the lineage-labelled line (Tcf15-Het-mKate2) could be used to assess the status of Tcf15 expression both in Tcf15-Venus expressing and non-expressing labelled cells throughout the embryo; the non-lineage labelled line (Tcf15-Het) could be used to assess whether Tcf15 is co-expressed with other transcription factors in the early post-implantation embryos. Any results from these analyses could then be confirmed in the Tcf15-Venus reporter mouse lines, once available.

3.3.2: Tcf15 expression in pluripotent cultures

Pluripotent cells can exist in a "naïve" state, possibly representing cells from the early blastocyst or in a "differentiation primed state" that may correspond to cells within the mature blastocyst or the early post-implantation epiblast (Nichols and Smith, 2009). Pharmacological inhibitors of MEK and GSK-3β can be used to maintain ES cells in a naïve state (Ying et al., 2008), whereas culture in FGF2 plus Activin A drives them

into differentiation-primed EpiSCs (Tesar et al., 2007, Brons et al., 2007). *Tcf15* expression was analysed in these different self-renewing pluripotent cultures from naïve to primed and post-implantation states (Figure 3.4). The results showed that cells in a primed state expressed higher levels of *Tcf15* than cells in naïve pluripotent state or in EpiSC cultures. Time-course 2i to EpiSC transition in serum-free media was performed in order to monitor changes of *Tcf15* as cells move between pluripotent naïve, primed and epiblast states whilst avoiding artifactual effects on *Tcf15* expression that could result from exposure to growth factors present in serum**.** The transient peak of *Tcf15* expression at early time points during 2i to EpiSC transition raises the possibility that Tcf15 might play a role during the transition from a naïve pluripotent state to primed state.

3.3.3: Expression level of Tcf15 during differentiation

Tcf15 expression during cell differentiation including neural, non-neural and EB formation was analysed. The expression of *Tcf15* persisted after naïve genes declined and was downregulated before lineage specific genes began to be expressed (Figure 3.5), suggesting any role played by Tcf15 must take place during the early stages of differentiation. It is worth noting that *Tcf15* expression slight decreased at day 1 of neural differentiation but was sustained at levels similar to ES cells in non-neural differentiation and EB formation at day 1. This difference might due to the presence of serum, which possibly contained some growth factors to sustain *Tcf15* expression, in non-neural differentiation (LIF withdraw) and EB formation.

3.3.4: Tcf15 reporter cells

Two Tcf15-Venus reporter cell lines were used in this study: the highly sensitive gene-trapped 216D1 Tcf15-Venus reporter, which *Tcf15* transcripts were visualized based on the expression of Venus coupled to a unique translational amplifier, *gtx-IRES*, incorporated into the only intron of *Tcf15* (Tanaka et al., 2008) (Figure 3.6B); and the less-sensitive but directly gene-targeted Tcf15-Het Venus reporter, which the 1st exon of *Tcf15* allele was replace by *Venus* (Figure 3.18). Both Tcf15 reporters exhibit heterogeneous Venus expression (Figure 3.8 and 3.22), with negative correlation between Tcf15-Venus and naïve pluripotency markers Nanog/Klf4, but no negative correlation with the pan-pluripotency marker Oct4 expression. The difference in Nanog and Klf4 expression between FACS sorted Tcf15-Het cells is smaller than 216D1 Tcf15-Venus cells (Figure 3.9C and 3.23). This might due to the experimental design, which was based on sorting cells within the top and bottom 30 % of Venus expression. The Venus expression in 216D1 cells was amplified by the *gtx-IRES* cassette where Venus expression in Tcf15-Het cells was driven by endogenous *Tcf15* promoter. Therefore, the same 30 % gate in 216D1 cells might be more stringent than in Tcf15-Het cells. Therefore, a tighter gating strategy could be tried for optimizing the sorting process in the Tcf15-Het reporter cells.

Even in the continued presence of LIF, a subset of ES cells downregulate markers of naïve pluripotency. This subset of ES cells is consequently primed toward differentiation (Toyooka et al., 2008, Chambers et al., 2007, Kalmar et al., 2009). Recent work shows that some of these Nanog-low cells, marked by the endodermal transcription factor Hex, are primed for primitive endoderm, but a third population expresses neither Hex nor naïve pluripotency markers (Canham et al., 2010). It has been speculated (Canham et al., 2010, Lanner

and Rossant, 2010) that this third population is primed for differentiation into mature epiblast and therefore ultimately for differentiation into somatic cell types, but a marker of this early primed state has proved elusive. Gene expression analysis of FACS sorted Tcf15-Venus reporter cells indicate Tcf15 might mark cells primed for somatic lineages, as Tcf15-high cells expressed lower level of naïve marker genes *Nanog* and *Klf4*, slightly higher level of the epiblast markers *Fgf5* and *T* and slightly lower level of primitive endoderm gene *Hex*.

Tcf15 activity is likely to be negatively regulated by Id proteins (Figure 3.11A). The differential gene expression between Venus-high and Venus-low cells was intensified when cells were sorted from 24 hours N2B27 culture (Figure 3.12), in which Id inhibition is reduced, indicating the epiblast-priming effect of Tcf15 might be stronger when Id inhibition was removed. It is known that Id1 overexpression delays the exit from a pluripotent epiblast-like state (Malaguti et al., 2013). Therefore, it would be very informative to generate a Tcf15/Id1 double reporter cell line to analyse if the truly active Tcf15 (Tcf15-high and Id-low population) is more tightly associated with the epiblast-primed state.

Tcf15 heterogeneity was dynamic within pluripotent ES cells, as FACS sorted Venus-high and Venus-low cells could re-establish a heterogeneous culture within 24 hours in self-renewing conditions (Figure 3.13B and 3.23D). In pluripotent cultures containing LIF and serum, Tcf15-high cells were not cells that had committed to differentiation, as Venus-high and Venus-low cells were both capable to generate AP-positive colonies with no obvious difference in colony numbers (Figure 3.14A). However, Tcf15-Venus-high cells were functionally primed for differentiation, as cells differentiated more efficiently toward neural lineage than the Tcf15-Venus-low subpopulation

(Figure 3.15 and 3.23E). It would be therefore interesting to assess if Tcf15-Venus-high cells committed to differentiation more quickly when cells were released from pluripotent cultures. Commitment assays could be setup to challenge cells sorting from LIF and serum: any cell that is more primed to differentiation during the "challenge" phase (N2B27 culture for certain time, for example: 12, 24, or 48 hours) should be less able to form an alkaline-phosphatase-positive colony when plated at clonal density under ES cell self-renewal culture conditions (LIF + serum) (Zhou et al., 2013). If Tcf15-high cells are marking cells primed for differentiation, they were likely to commit to differentiation when releasing from pluripotent culture.

Another experiment to test if Tcf15 primed ES cells to somatic lineages rather than extra-embryonic lineage is to test whether Tcf15-Venus-high cells are less efficient at primitive endoderm differentiation. Chimaeric EBs could be generated to test the extra-embryonic lineage differentiation ability of different Tcf15-Venus populations. Briefly, chimaeric EBs are made by mixing wild-type ES cells with FACS sorted mKate2-lineage-labelled Tcf15-reporter cells (Tcf15-Het-mK2). EBs are examined at day 4 for analysing the contribution of Tcf15-Het cells by assessing the mKate2 expression. This type of approach was used to show that Hex-high ES cells move preferentially to the outside of EBs to form the presumptive visceral endoderm (Canham et al., 2010). Therefore, it would be predictable that Tcf15-Venuus-high cells should exhibit opposite phenotype compare to Hex-reporters, by localizing in the centre of EBs.

The results of the EB differentiation experiments (Figure 3.16) suggest that differentiation is modestly accelerated when initiated from a Tcf15-Venus-high state. It is notable that the priming effect of Tcf15 is less striking during EB differentiation than in serum-free neural monolayer

differentiation (Figure 3.16). However, EB differentiation culture contains serum and so it is likely that Tcf15 will be less active in these conditions than in serum-free conditions due to presence of Id (see section **3.2.2.3.1.3** and figure 3.11C). A future experiment would be to try to identify a non-neural directed differentiation protocol where Id expression will be low (e.g. endoderm differentiation: N2B27 media containing Activin but without BMP).

Taken together, in this chapter, the expression of *Tcf15* during differentiation *in vitro* and during early mouse development *in vivo* were examined. Expression of *Tcf15* during early stage of differentiation suggested that it might act as a pro-differentiation factor. Tcf15-high cells expressed lower level of naïve pluripotency and extraembryonic endoderm markers, indicating it might mark the epiblast-primed subpopulation of ES cells and identify cells taking the first step toward embryonic rather than extraembryonic lineages. To further validate if this differential Tcf15 expression correlation to any functional difference, I would like to perturb Tcf15 expression in ES cells in order to assess the function of Tcf15.

CHAPTER 4

Functional characterization of Tcf15 in mouse embryonic stem (mES) cells

4.1: Hypothesis: Tcf15 acts as a pro-differentiation factor priming mES cells toward somatic lineages

In chapter 3, the expression pattern of endogenous Tcf15 was analyzed. At the single cell level, Tcf15 was expressed heterogeneously within colonies of undifferentiated ES cells. Expression of Tcf15 negatively correlates with expression of naïve pluripotency markers and extraembryonic markers, indicating Tcf15-high cells might represent a new subpopulation within pluripotent cell culture, distinct from both the naïve subpopulation (Chambers et al., 2003) and the primitive endoderm-primed subpopulation (Canham et al., 2010). Moreover, a transient expression window of *Tcf15* as cells exit a naïve self-renewing state and before they start to acquire lineage-specific gene expression may suggest that Tcf15 might play a role in the early phase of differentiation. Furthermore, functional analysis of differentiation speed between Tcf15-Venus-high cells and Tcf15-Venus-low cells confirmed that Tcf15 expression marks a functionally differentiation-primed state.

These observations raise the question of whether Tcf15 is simply a marker of primed cells or whether it plays a functional role in driving cells towards differentiation. In this chapter, the hypothesis that Tcf15 acts as a pro-differentiation factor within pluripotent population, priming mES cells toward somatic but not extraembryonic lineages will be tested (Figure 4.1).

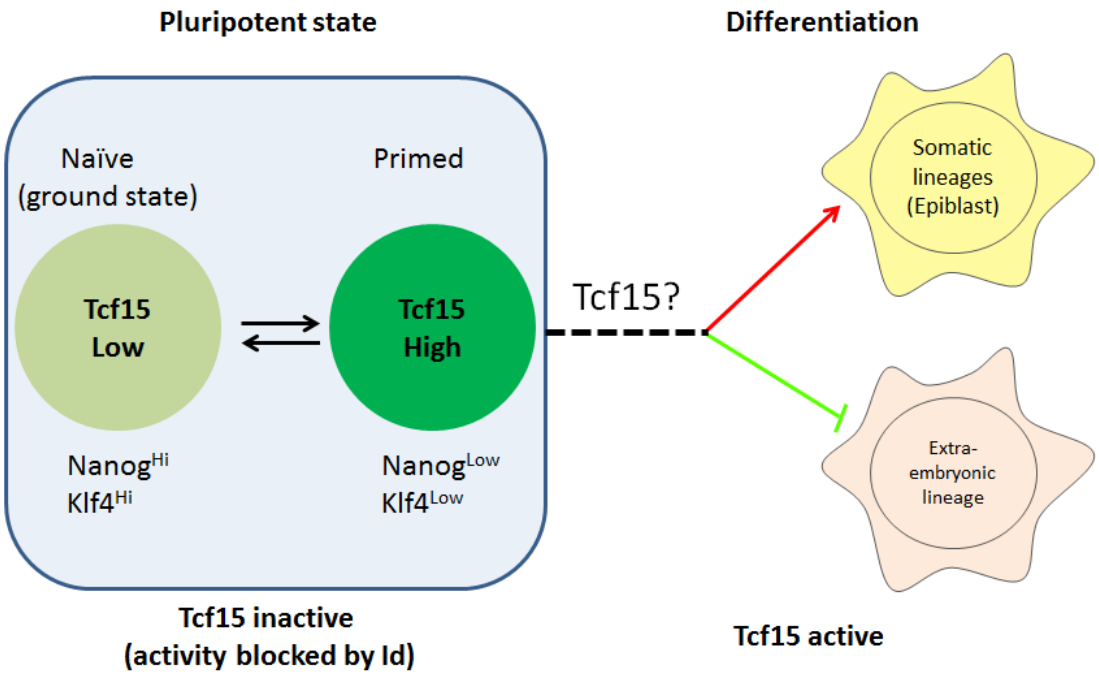


Figure 4.1: Schematic diagram illustrating a hypothetical model of possible function of Tcf15 in priming cells toward somatic lineages

ES cells contained Tcf15-high (dark green) and Tcf15-low (light green) populations and these two populations are dynamic in pluripotent culture (light blue box). However, Tcf15 activity within pluripotent culture was blocked by the presence of Id protein. Therefore, Tcf15 might act as a pro-differentiation factor within pluripotent population, without driving cells to exit pluripotency network. During cell differentiation, expression of Id decreased and Tcf15 was able to prime mES cells toward somatic but not extraembryonic lineages.

4.2: Aims of this chapter

The primary aim of this chapter was to use gain of function and loss of function approaches to perturb Tcf15 expression in mES cells in order to assess the function of Tcf15 in self-renewal and during differentiation. For gain of function studies, Tcf15 inducible cell lines were generated and validated. For loss of function studies, siRNA knockdown of Tcf15 expression was attempted and a strategy to delete both alleles of *Tcf15* from mouse ES cells was generated.

4.3: Results

4.3.1: Gain of function studies

To investigate whether increasing the activity of Tcf15 within pluripotent mES cells was sufficient to induce cells toward somatic differentiation, gain of function studies were performed. An engineered form of Tcf15 that is tethered by a flexible linker of sequence TGSTGSKTGSTGS to its hetero-dimerization partner E47 was cloned into 3'-end of a Flag tag. As described in section **3.2.2.3.1.3**, in the presence of Id proteins, Id dimerizes with and sequesters the ubiquitous bHLH transcription factor E12/E47, blocking its interaction with Tcf15 (Figure 3.11A). Therefore, this forced-dimerization of Tcf15-E47 strategy renders Tcf15 more resistant to inhibition by Id (Neuhold and Wold, 1993) and also avoids disrupting the balance of endogenous bHLH factors in the cell (Figure 4.2A).

The *Flag-Tcf15-E47* DNA fragment was then sub-cloned into recipient

vectors as below. This part of the cloning work was performed by Aliaksandra Radzisheuskaya and Paul Nistor in the lab.

Two different strategies were selected to generate inducible mES cell lines over-expressing Tcf15.

1) FKBP12-derived inducible stabilization system: This strategy is to induce Tcf15 expression at post-translational level, by Shld1-induced protein stabilization.

   Briefly, the *CMV early enhancer/chicken β-actin* (*CAG*) promoter was selected for maintaining long term expression of the transgene during stem cell differentiation (Alexopoulou et al., 2008). Flag-Tcf15-E47 was fused to the C-terminal of FKBP protein containing a destabilizing domain (DD) followed by a bicistronic hygromycin resistance gene to obtain effective expression. The FKBP-fusion protein was effectively targeted for proteasomal degradation unless a stabilizing ligand "Shld1" was added to bind and mask the destabilizing domain (Banaszynski et al., 2006) (Figure 4.2B).

2) Doxycycline inducible expression system: This strategy is to induce Tcf15 gene transcription followed by protein expression upon addition of doxycycline.

   In short, *Flag-Tcf15-E47* was placed under a Tet-responsive promoter *TetO* within a cell line (AW2) (Zhou et al., 2013) that has constitutive expression of a reverse tetracycline transactivator (rtTA). In the presence of doxycycline, a tetracycline analogue, doxycycline forms a complex

with the rtTA protein and this complex was then capable of binding to *TetO* and activating downstream gene expression (Figure 4.2C).

4.3.1.1: Generation of FKBP12-Flag-Tcf15-E47 inducible cell lines using Oct4-GiP background

To generate the FKBP12-Flag-Tcf15-E47 inducible mES cell line, Oct4-GiP cells were used. The Oct4-GiP cells expressed bicistronic green fluorescent protein (GFP) and puromycin resistance exclusively in pluripotent cells under direction of regulatory sequences of the mouse *Oct4* gene (Ying et al., 2002). Puromycin selection of Oct4-GiP cells eliminated differentiated cells and thus gave purer starting population when performing experiments on heterogeneous mES cells containing a number of spontaneous differentiating cells.

The pCAG-FKBP-Flag-Tcf15-E47-IRES-Hygromycin plasmid was transfected into Oct4-GiP cells by liposome-mediated delivery. Cell transfectants were selected with hygromycin (50 μg/mL) for 10 days. 91 hygromycin-resistant single colonies were picked into 96 well plates, expanded, and screened for *Flag-Tcf15-E47* transgene expression. Clones were triplicated plated in 48 well plates: 1 plate was frozen down in -80°C for maintaining cell stock, 1 plate for non-stimulation control and the last plate was stimulated with 1 μM of Shld1.

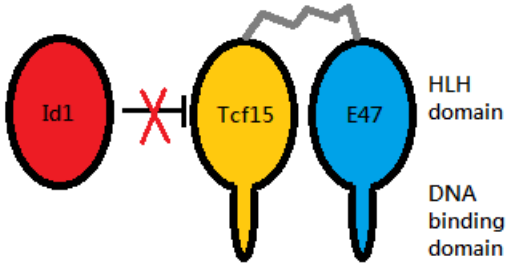
Screening of induction was not possible at mRNA level due to the fact that degradation of transgene occurred only at protein level. Therefore, both control and Shld1-stimulated plates were fixed with 4% paraformaldehyde

after 24 hours of Shld1 stimulation for immunocytochemistry-staining. Anti-Flag antibody was used to detect transgene expression. There should be no Flag staining in the absence of Shld1 stimulation but positive staining after Shld1-inducible stabilization of FKBP-fusion Flag-Tcf15-E47 fusion protein. A stable cell line expressing FKBP-Flag-Hes1 served as a positive control (Figure 4.3A).

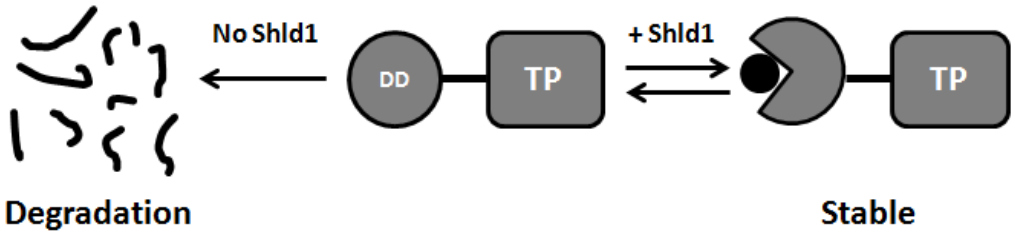
After screening all of the 91 clones, around 80% of those clones showed no induction of transgene expression as no Flag-positive staining could be detected after Shld-stimulation under microscopic examination. Very few clones, such as clone 25, displayed Flag-positive staining, although at relatively lower level when compared with FKBP-Flag-Hes1 cells. The nuclear localization of Flag staining suggested that Flag-Tcf15-E47 protein located within the nucleus, where it could function as a transcription factor. However, the Shld1 induced protein stabilization in clone 25 shown some expression of Flag even in the absence of Shld induction (Figure 4.3B), indicating some 'leakiness' of expression in un-induced cells. Due to not able to pick promising clones in 48 well plate format, the toxicity of transgene expression was not further analysed in this screening.

Therefore, a doxycycline-induction method was selected to over-express Tcf15 within ES cells.

A)



B)



C)

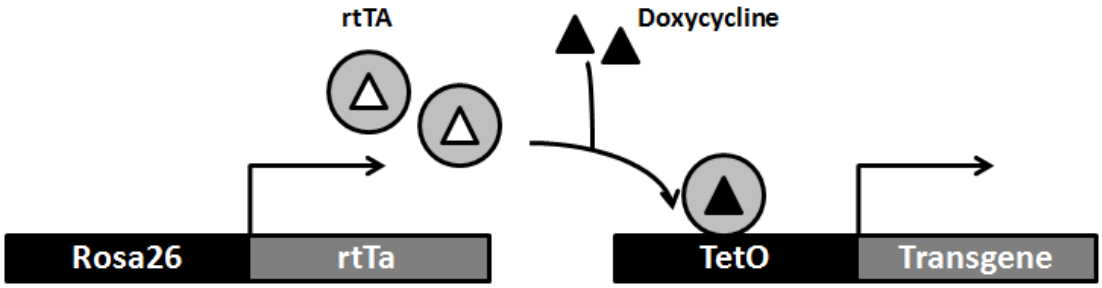


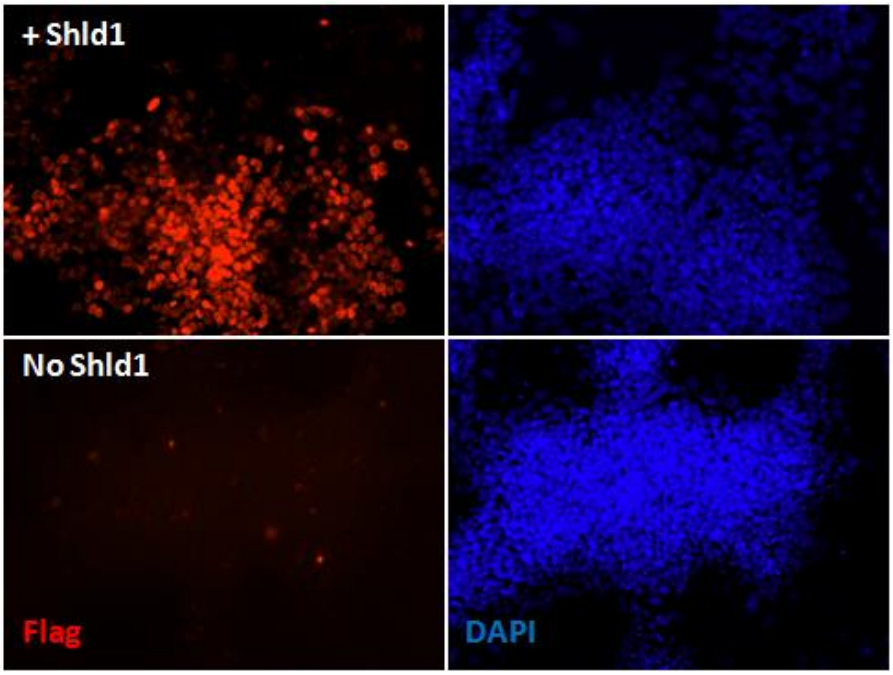
Figure 4.2: Strategies for generating inducible mES cell lines over-expressing Tcf15

(A) Dox-inducible, Id resistant dominant active Tcf15-E47.

(B) FKBP12-derived inducible protein stabilization system: TP fused with DD undergo degradation in the absence of Shld1. Upon addition of Shld1, TP is then maintained in a stable state. DD: destabilizing domain, here FKBP1A-L106P, TP: target protein, here Flag-Tcf15-E47.

(C) Doxycycline inducible gene expression system: Upon addition of doxycycline, doxycycline forms complex with rtTA protein and this complex is then capable for binding to *TetO* promoter for activating downstream gene (*Flag-Tcf15-E47*) expression.

A) FKBP-Flag-Hes1



B) FKBP-Flag-Tcf15-E47: clone 25

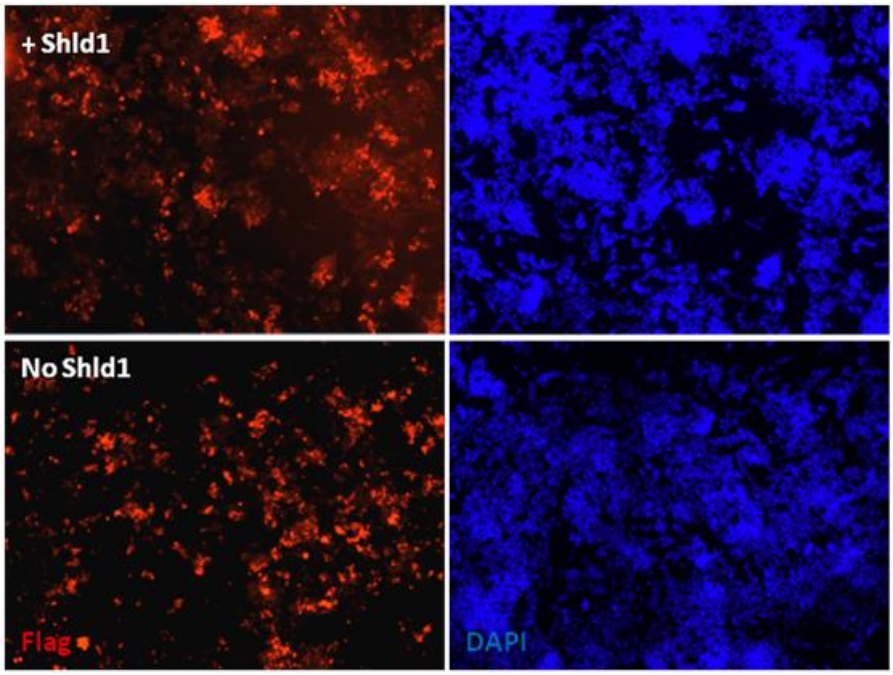


Figure 4.3: Immunostaining on FKBP inducible cell lines

(A) Flag staining (red) and nuclear DAPI staining (blue) of FKBP-Flag-Hes1 cells +/- Shld1 stimulation. 10x objective lens.

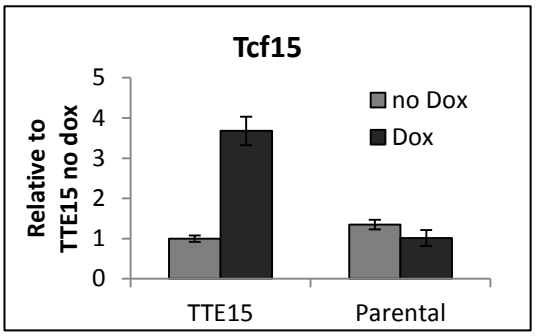
(B) Staining of FKBP-Flag-Tcf15-E47 clone 25, Flag staining (red) was detected in the presence and absence of Shld1 stimulation. 4x objective lens.

4.3.1.2: Gain of function studies using doxycycline-inducible Flag-Tcf15-E47 cell line

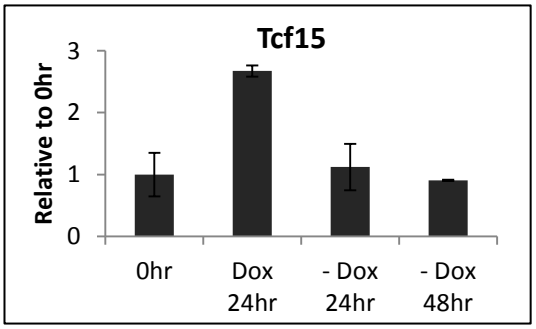
A doxycycline-inducible Flag-Tcf15-E47 cell line was previously made in the lab by introducing *TetO-Flag-Tcf15-E47* construct into the E14tg2a_AW2 cell line, which contains the coding sequence for the reverse tetracycline transactivator (rtTA) integrated into the *Rosa26* locus and expressed from the *Rosa26* promoter (Zhou et al., 2013). Clone TTE15 (TTE stands for Tet-inducible Tcf15-E47) was pre-selected after Flag-staining showed no leakiness in the absence of doxycycline induction. Expression of *Tcf15* mRNA was then examined and compared with the parental E14tg2a_AW2 cell line. When compared with parental cells, TTE15 cells displayed negligible leakiness of *Tcf15* expression without doxycycline induction, and moderate (2-3 times higher) overexpression in the presence of 1 μg/mL doxycycline (Figure 4.4A). All the following experiments were using 1 μg/mL doxycycline to induce transgene expression.

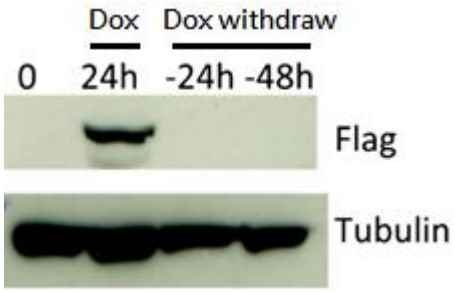
The expression of Flag-Tcf15-E47 transgene was lost within 24 hours after removing doxycycline at both mRNA and protein level (Figure 4.4B). Furthermore, *Tcf15* expression could be artificially maintained under neural differentiation conditions at levels similar to those in ES cells whilst the endogenous *Tcf15* dramatically decreased after 3 days under the same differentiation conditions (Figure 4.4C). Therefore, the clone TTE15 was selected for Tcf15 gain of function studies in which *Tcf15* expression could be artificially sustained at close to physiological levels under differentiation conditions.

A)



B)





C)

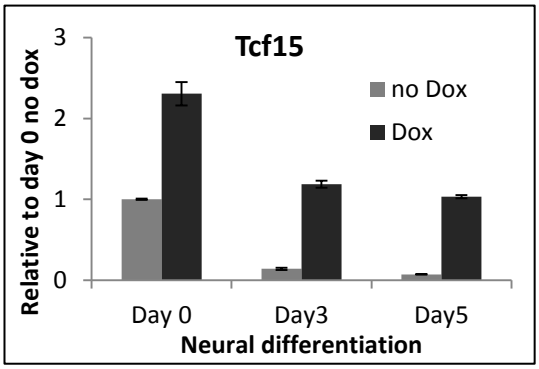


Figure 4.4: Characterization the expression of doxycycline-inducible TTE15 cells

(A) qPCR analysis for *Tcf15* in dox-inducible TTE15 cells or parental control ES cells: 1 day treatment with doxycycline in culture medium containing LIF and serum.

(B) qPCR and western blot to detect Flag-Tcf15-E47 expression in TTE15 ES cells after 24 hours induction with doxycycline and 24 or 48 hours after removal of doxycycline.

(C) qPCR analysis for *Tcf15* in TTE15 cells during N2B27 monolayer neural differentiation. Expression level of *Tcf15* was maintained in dox-treated cells.

4.3.1.2.1: Artificially sustained of Tcf15-E47 fusion protein accelerates ES cells differentiation and suppresses primitive endoderm lineage in neural induction conditions

TTE 15 cells were used to address the hypothesis that Tcf15 might act as a pro-differentiation factor priming ES cells for somatic differentiation. TTE15 cells were pre-treated with doxycycline in complete medium containing LIF and serum for 24 hours before starting the N2B27 monolayer neural differentiation and keep on for entire differentiation. In the presence of doxycycline, Tcf15-E47 promotes neural differentiation, as revealed by the presence of more neurites in the dox-treated culture compared with untreated control culture (Figure 4.5A). Immuno-staining of neuroepithelial progenitor cell marker nestin, an intermediate filament protein (Lendahl et al., 1990), and Tuj1 (Lee et al., 1990), a neuron-specific class III beta-tubulin, revealed that doxycycline-treated TTE15 cells expressed strikingly higher level of these two neural markers (Figure 4.5B).

Tcf15 expression was maintained in dox-treated TTE15 cells at levels similar to those in untreated ES cells throughout the differentiation, whist untreated TTE15 cells down-regulated *Tcf15* to a similar extent as parental cells (Figure 4.6A). In addition, up-regulation of neuronal markers *Sox1* and *N-cadherin* could be detected from day 2 upon induction while in no-dox and parental cells the expression did not begin until one day later (Figure 4.6B). *Sox1* and *N-cadherin* expression persisted at higher level until day 5. These data further confirm that Tcf15-E47 accelerates neural differentiation.

Under neural differentiation conditions, a significant subpopulation of

cells usually differentiates into non-neural cells that morphologically resemble extraembryonic endoderm. To ask whether neural differentiation becomes more uniform in response to Tcf15 activation, TTE15 cells were induced with doxycycline throughout the neural differentiation process and compared with un-induced control cells. Tcf15-E47 therefore might maintain the robustness of neural differentiation by suppressing differentiation toward primitive endoderm fate, as assessed by morphology (Figure 4.7A) and qPCR analysis of primitive endoderm marker *Sox7* (Futaki et al., 2004) expression (Figure 4.7B).

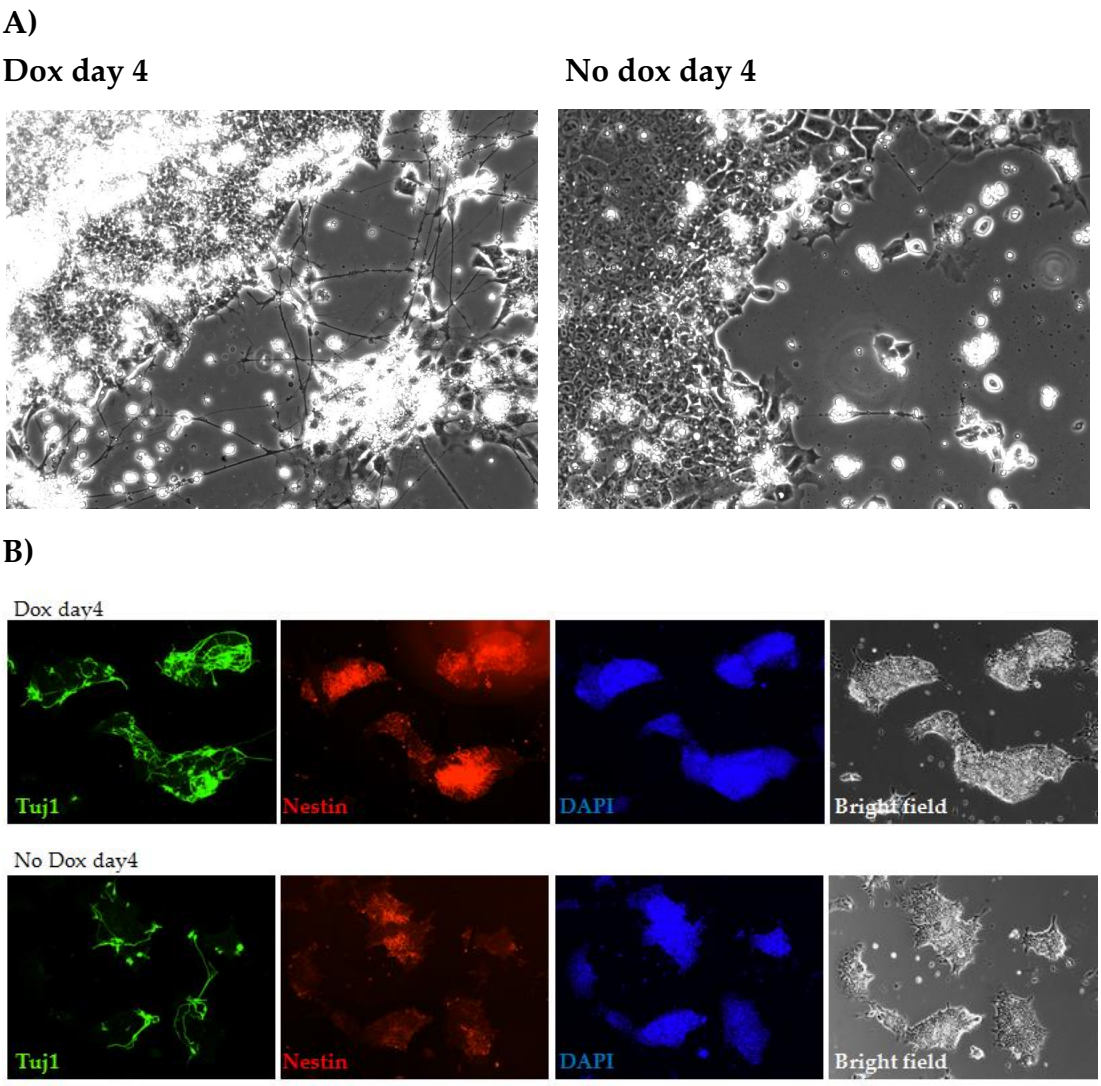
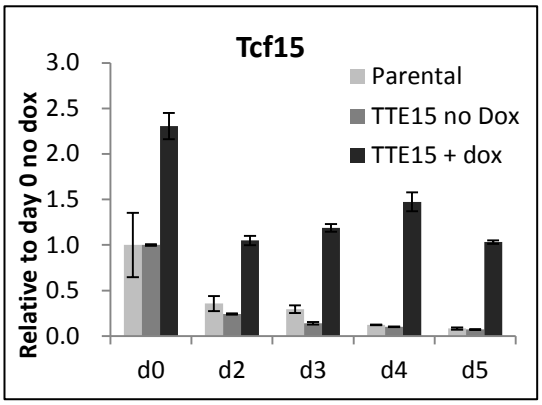


Figure 4.5: Forced over-expression of Tcf15-E47 fusion protein promotes neuronal differentiation

(A) Morphology of dox-inducible TTE15 cells under 4 days of differentiation in monolayer N2B27 neural induction condition. In the presence of doxycycline, there was more neurites outgrowth compared with no-dox control cells.

(B) Tuj1 (green) and nestin (red) staining to detect neurons and neural progenitors during neural differentiation of dox-inducible TTE15 cells.

A)



B)

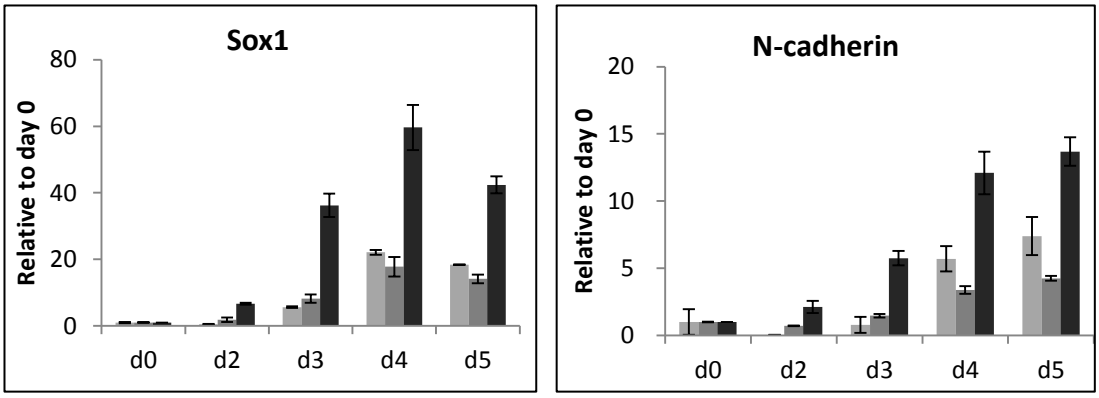


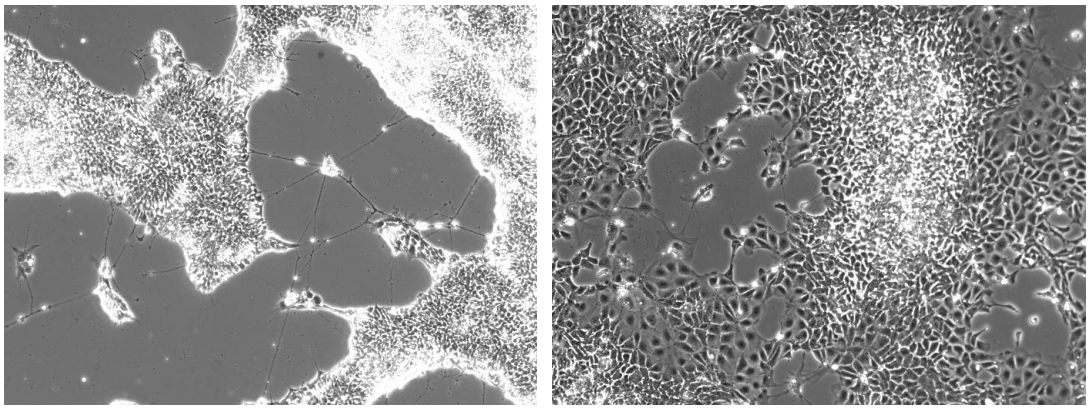
Figure 4.6: Forced over-expression of Tcf15-E47 fusion protein accelerates ES cells differentiation in neural induction condition

(A) qPCR of parental cells (light grey bars), TTE15 cells no-dox control (dark grey bars) and TTE15 cells induced with dox to over-express Tcf15-E47 (black bars). Endogenous expression of *Tcf15* decreased during the differentiation process. In the presence of doxycycline, *Tcf15* expression was maintained at the same level as in the ES cell culture.

(B) *Sox1* and *N-cadherin* were up-regulated from day 2 in Tcf15-E47 over-expressing cells at the similar level as day 3 controls.

A)

Dox day4 **No dox day4**



B)

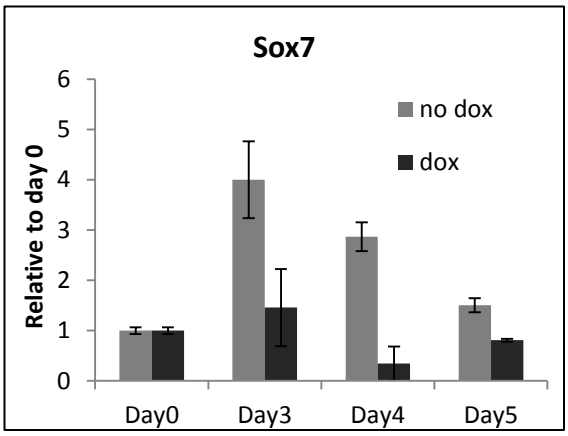


Figure 4.7: Forced over-expression of Tcf15-E47 suppresses ES cells differentiation to primitive endoderm

(A) Morphology of dox-inducible cells underwent N2B27 neural differentiation at day 4. Dox-treated TTE15 cells showed robust neural differentiation with neurites outgrowth from neural epithelium. In no-dox control cells, some cells differentiated toward triangular primitive endoderm-like lineage.

(B) qPCR analysis of primitive endoderm marker gene *Sox7* expression.

4.3.1.2.2: Transient induction of Tcf15-E47 was sufficient to drive differentiation

Endogenous *Tcf15* expression peaks during the first day of differentiation and is subsequently downregulated (Figure 3.4). In addition, *Tcf15* is downregulated *in vivo* after implantation (Figure 3.1). This led us to propose that Tcf15 acts at the earliest stage of differentiation (Figure 4.1). In order to further test this hypothesis, it is worth asking whether exogenous Tcf15 was acting during the first day of differentiation.

To answer this question, *Tcf15-E47* transgene was induced only for the first 24 hours of neural differentiation and compared with induction throughout and control un-induced cells (Figure 4.8A). *Tcf15* expression was sustained in dox-throughout cells, whereas in dox-day 1 withdraw cells induction was reduced to similar level to control cells at day 2 (Figure 4.8B). More Tuj1^{+} neurites in dox-treated cultures could be detected from day 2 and abundant by the third day in dox-treated, but not control cultures (Figure 4.8C). This was consistent with previous qPCR analysis (Figure 4.6B), suggesting that induction of Tcf15-E47 accelerate neural induction. In addition, there were no significant difference between the numbers of neurites presenting in dox-throughout and dox treating in the first 24 hours cultures. Dox-induced cells also down-regulated *Oct4* and *E-cadherin* expression more quickly than un-treated cells, although the speed was slightly slower in day 1 dox-withdraw cells. qPCR analysis of *Sox1* and another neural marker *Zfp521* (Kamiya et al., 2011) expression confirmed that neural differentiation occurred more rapidly in dox-treated cultures, even

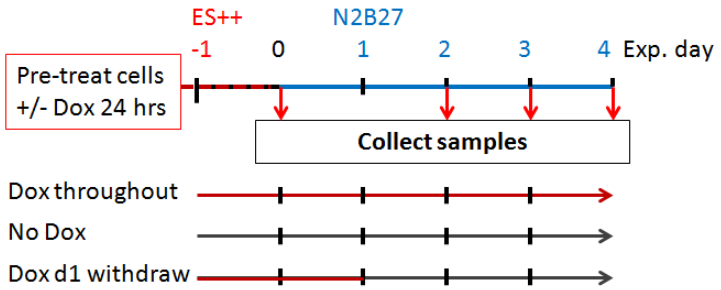
when doxycycline was removed after the first 24 hours of differentiation (Figure 4.8D). Therefore, Tcf15-E47 is effective at accelerating differentiation even if expressed only during the first 24 hours of differentiation. Taken together, these results suggest that Tcf15 activity may be a limiting factor at an early stage of differentiation.

4.3.1.2.3: Somatic lineage choice was not biased by activation of Tcf15

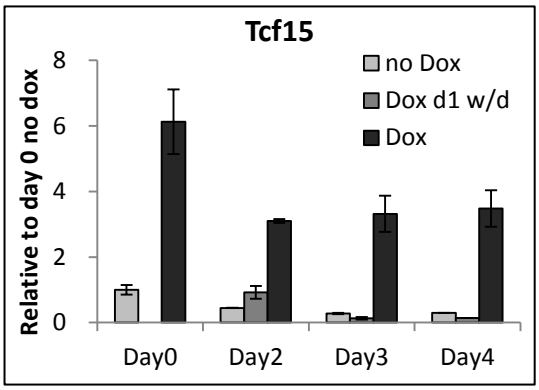
The data above suggest that Tcf15 accelerates somatic differentiation and suppresses extraembryonic endoderm differentiation in neural inducing condition. The next question would be whether Tcf15 was specifically promoting the neural lineage or acting in a more general mode to promote differentiation into other somatic lineages. The working hypothesis (Figure 4.1) that Tcf15 drives the initial transition towards somatic fates, would suggest that Tcf15 would favour differentiation into epiblast-derived somatic lineages without biasing the particular direction of differentiation. In order to test whether or not somatic lineage choice was biased by Tcf15, TTE15 cells were placed into N2B27 basal media supplied with BMP4 to investigate the function of Tcf15 during non-neural differentiation. BMP4 was known to suppress neural differentiation (marked by *Sox1*) in dose-dependent manner and instead promote non-neural differentiation (marked by *T-Brachyury*) (Finley et al., 1999). As shown in figure 4.9A, in the absence of BMP4, forced-overexpression of Tcf15 promotes neural differentiation whereas in the presence of 10 ng/mL of BMP4, Tcf15 promotes non-neural differentiation. Tcf15 promotes both neural and non-neural differentiation in the presence of

low dose (1 ng/mL) of BMP4, suggesting induction of Tcf15-E47 didn't bias ES cells toward either lineage significantly. Tcf15 therefore does not appear to mediate the ability of BMP/Id to suppress neural induction (Ying et al., 2003) but may mediate the earlier role of Id to inhibit the transition of ES cells to differentiation-primed epiblast (Zhang et al., 2010) (see these two transitions illustrated in figure 4.9B).

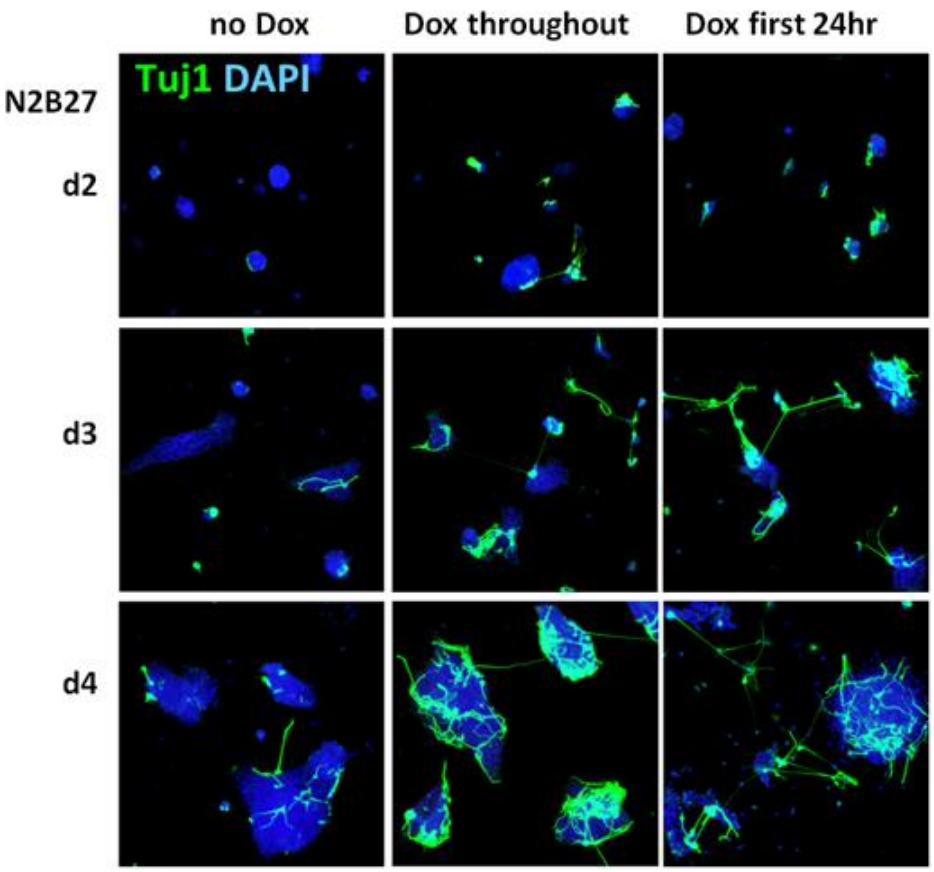
A)


ES++
N2B27
-1
0
1
2
3
4 Exp. day
Pre-treat cells
+/- Dox 24 hrs
Collect samples
Dox throughout
No Dox
Dox d1 withdraw

B)


Tcf15
Relative to day 0 no dox
8
6
4
2
0
no Dox
Dox d1 w/d
Dox
Day0
Day2
Day3
Day4

C)


no Dox
Dox throughout
Dox first 24hr
N2B27
Tuj1 DAPI
d2
d3
d4

D)

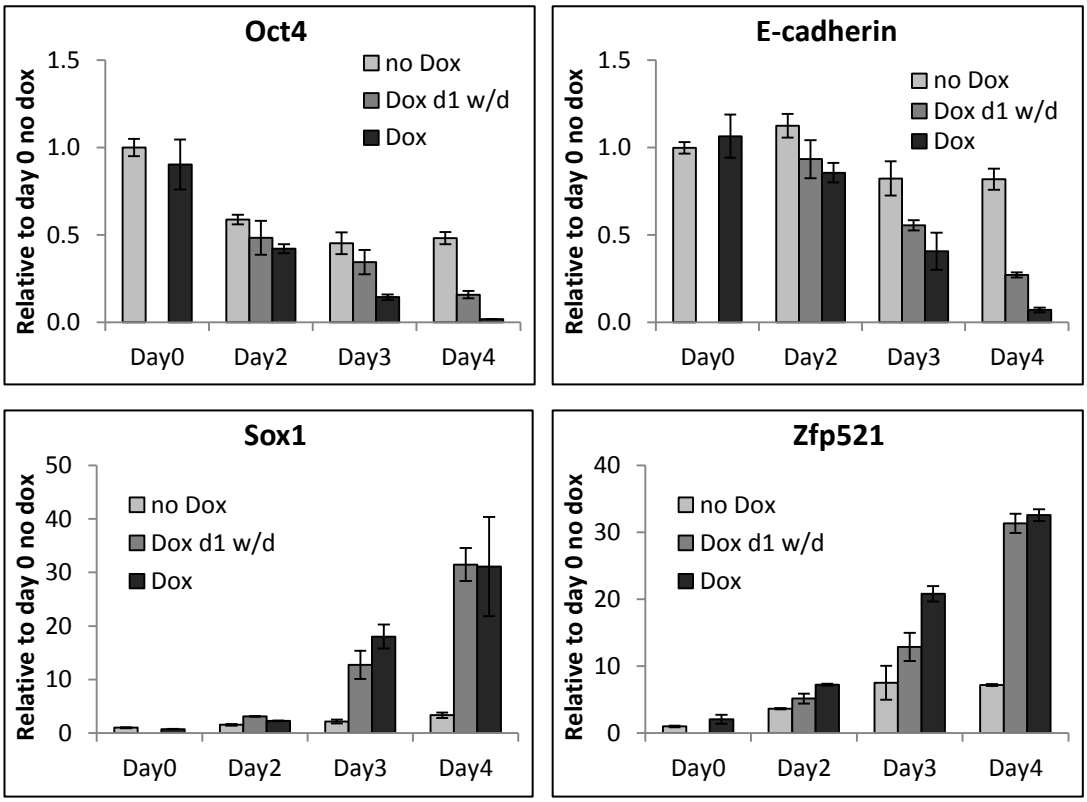


Figure 4.8: Transient induction of Tcf15-E47 is sufficient to drive differentiation

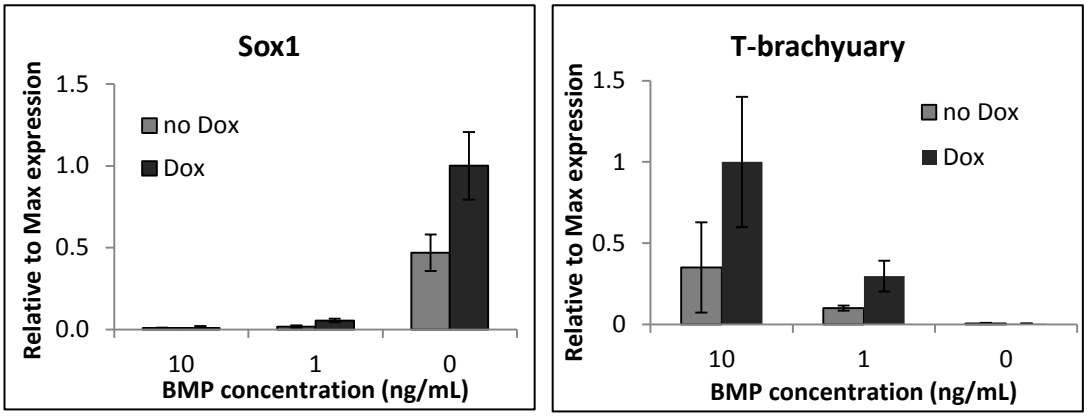
(A) Experimental design to test how transient induction of Tcf15-E47 expression affects cell differentiation.

(B) qPCR of *Tcf15* expression from day 2 to day 4 during neural differentiation.

(C) Immunostaining of Tcf15-E47 dox-inducible cell line cultured in N2B27 medium, showing neurite marker Tuj1 in green and nucleus stained by DAPI in blue.

(D) qPCR analysis from day 2 to day 4 of neural differentiation. Neural differentiation was accelerated in dox-treated cells with up-regulation of *Sox1* and *Zfp521* and more rapidly loss of *E-cadherin* and pluripotent gene *Oct4*. Withdraw dox-treatment from day 1 of differentiation still promoted neural induction but the down-regulation of *Oct4* was slower than cells with doxycycline treatment all the time.

A)



B)

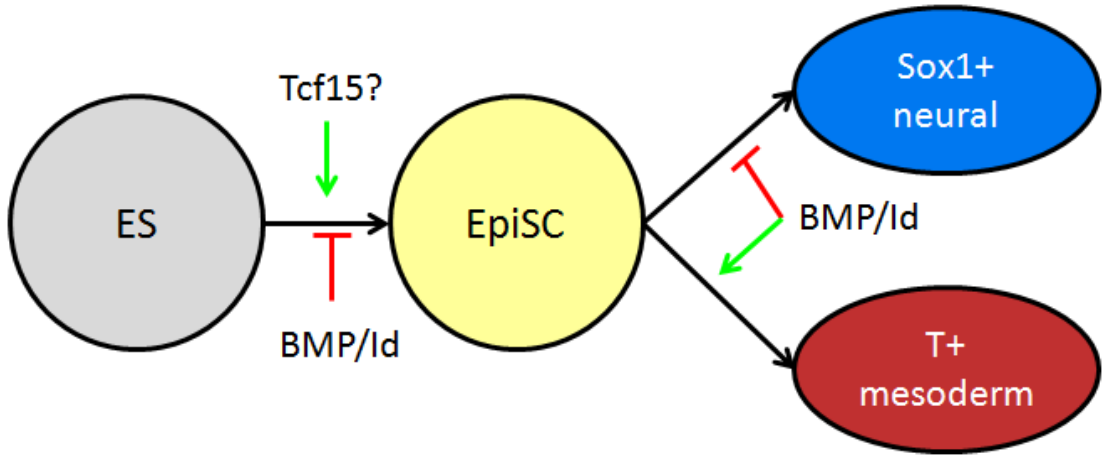


Figure 4.9: Somatic lineage choice was not biased by activation of Tcf15

(A) qPCR analysis of Tcf15-E47 inducible cells cultured in N2B27 medium containing different dose of BMP to block neural differentiation and promote non-neural differentiation. (N=2, n=2)

(B) Tcf15 was acting at early stage of ES cells differentiation, possibly by promoting transition to EpiSCs without bias later on lineage decision.

4.3.1.2.4: Tcf15 drives the transition to primed epiblast

In order to ask whether Tcf15 favours the transition to a differentiation-primed epiblast-like state, more directly experiments were setup to test this hypothesis. ES cells cultured in LIF and serum (FCS) or LIF and BMP are restrained from progressing to overt differentiation but free to explore naïve and primed states of pluripotency (Nichols and Smith, 2009). TTE15 cells were cultured in LIF and serum with induction of Tcf15-E47 for 3 days to ask whether Tcf15-E47 can drive the transition to a primed state. Under this long term persistent activation of Tcf15, Tcf15-E47 was able to significantly down-regulate markers of naïve ES cells and up-regulate the epiblast marker *Fgf5* (Figure 4.10). Tcf15-E47 also down-regulated extraembryonic endoderm marker *Sox7* expression in the pluripotent culture. On the other hand, long-term induction of Tcf15-E47 expression in LIF and serum didn't drive cells irreversibly toward differentiation, as cells could be passaged in the presence of doxycycline for several passages without losing pluripotency. This suggested that in this culture condition, Tcf15 only maturated cells toward primed state but did not drive cells toward differentiation.

To further understand if Tcf15 drives pluripotent cells from naïve to primed epiblast, ES cells were converted into Epiblast stem cells (EpiSCs) by culturing cells from 2i/LIF to N2B27 + FGF2 + Activin (Hayashi et al., 2011). Endogenous *Tcf15* expression was low in 2i culture and was up-regulated after cells were plating in the EpiSC medium (Figure 4.11A). Epiblast-like cells were achieved by monitoring the expression of epiblast determinant

gene *Otx2* and epiblast marker *Fgf5*. *Otx2* and *Fgf5* were up-regulated during 2i to EpiSC transition and induction of Tcf15-E47 promoted *Fgf5* expression significantly at day 1 while slightly increased *Otx2* expression. Tcf15-E47 did not appear to accelerate the loss of naïve pluripotency marker *Klf4*, but this is likely because *Klf4* was down-regulated very rapidly under this EpiSC transition even in the absence of exogenous Tcf15. *Nanog* was down-regulated within 1 day of EpiSC transition, in consistent with published data (Hayashi et al., 2011). The expression was up-regulated again, which might indicate cells were maturating to more proximal posterior pre-gastrulating epiblast-like identity (Malaguti et al., 2013; Osorno et al., 2012), as indicated by *Wnt3* and *Eomes* expression (Brennan et al., 2001).

Induction of Tcf15-E47 increased *Otx2* and *Fgf5* expression during the 2i to EpiSC transition suggested that Tcf15 might have an early function during the EpiSC transition. However, the rapid loss of naïve pluripotency marker and acquisition of epiblast marker genes expression during 2i to EpiSC transition indicated that the exit from a naïve state might occurs at an even earlier time point. Therefore, the 8 hour and 12 hour time points during the 2i to EpiSC transition were selected for further investigation. Tcf15-E47 significantly accelerates up-regulation of epiblast marker *Fgf5* and suppresses expression of primitive endoderm marker *Sox7* (Figure 4.11B), for which expression was only up-regulated transiently during these early time points. The transient up-regulation of *Fgf5* and down-regulation of *Sox7* further suggest that Tcf15 acts during early stage of ES cells transition to differentiation-primed epiblast but not extraembryonic endoderm.

Unlike culturing in LIF and serum, induction of Tcf15-E47 expression in N2B27 + LIF + BMP drove irreversible commitment to differentiation in a significant proportion of cells even in the continued presence of LIF and BMP (Figure 4.12A). A 24 hour pulse of doxycycline induction also displayed similar phenotype. In 24 hours pulse-dox cells, cells differentiated into mixed cultures, these cells were morphologically similar to cells from which LIF had been withdrawn. This suggests that Tcf15-E47 may prime cells toward differentiation by driving the loss of LIF responsiveness, which is a characteristic feature of the transition toward post-implantation EpiSC (Brons et al., 2007; Tesar et al., 2007). Moreover, cells exhibit more mesenchymal-like phenotype when continuous inducing Tcf15-E47 expression for 4 days, indicating Tcf15 might also promote EpiSC to further differentiation depend on environmental cues, although more experiments need to be performed to validate this hypothesis.

Taken together, the gain of function studies confirmed that Tcf15 mature naïve pluripotent ES cells transit toward primed epiblast and later on epiblast-derived somatic lineages whilst suppressing differentiation towards extraembryonic endoderm (Figure 4.12B).

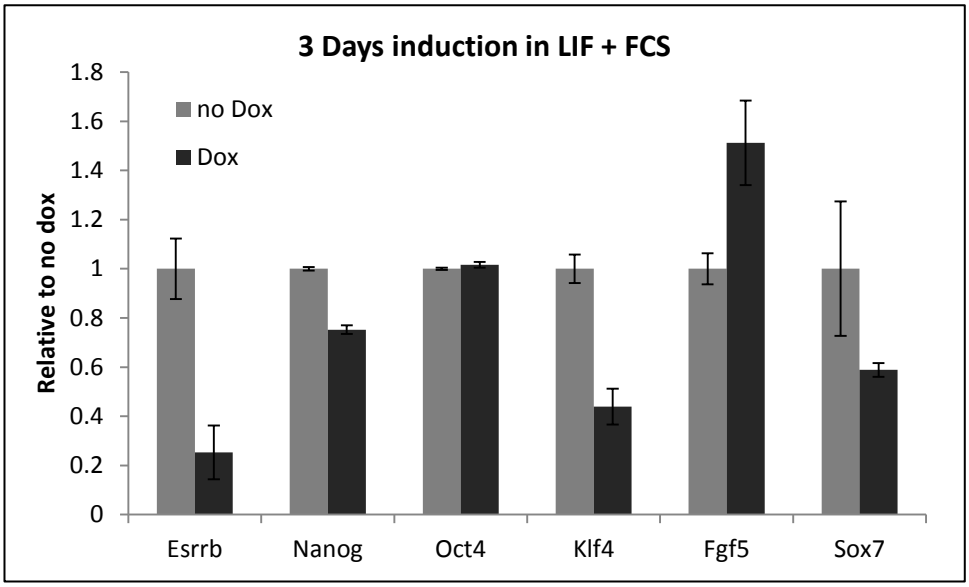
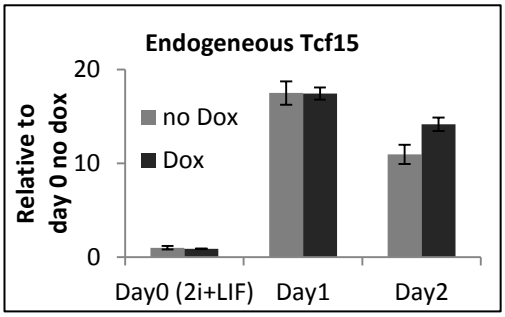
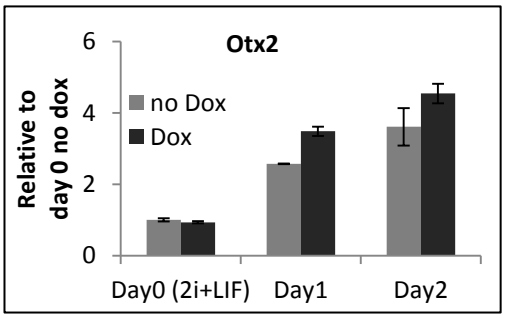


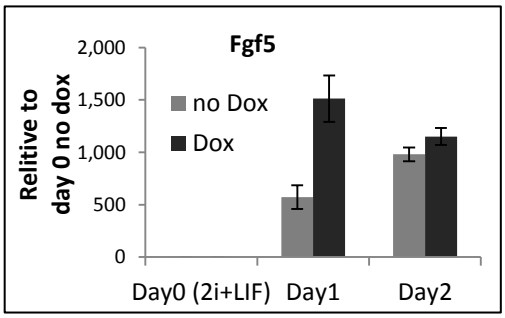
Figure 4.10: qPCR analysis of dox-inducible TTE15 ES cells after 3 days in LIF and serum culture

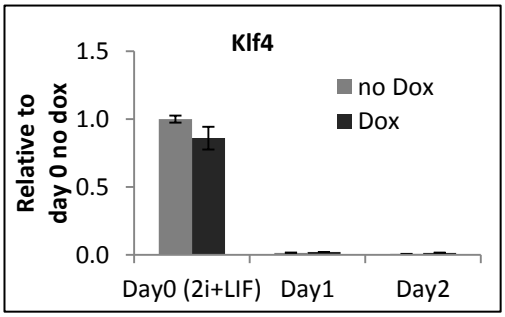
Persistent induction of Tcf15-E47 within ES cells down-regulated naïve pluripotency genes *Nanog, Klf4* and *Esrrb* without changing *Oct4* expression. Tcf15-E47 induction also up-regulated epiblast marker *Fgf5* and reduced extraembryonic endoderm marker *Sox7*. (N=2, n=2)

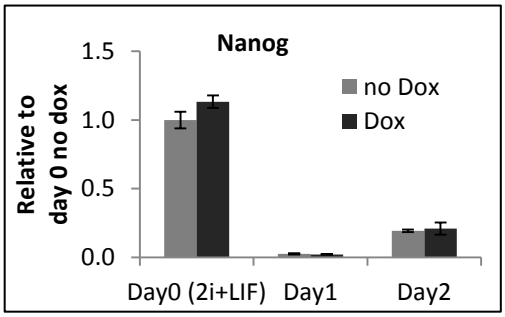
A)

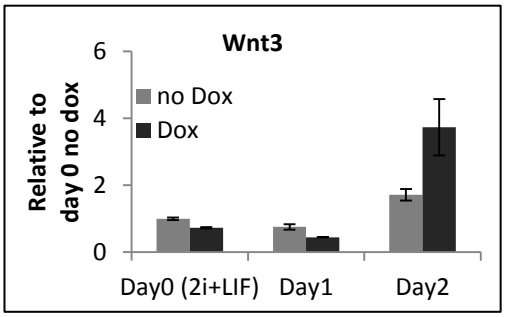

Endogeneous Tcf15
Relative to day 0 no dox
20
10
0
no Dox
Dox
Day0 (2i+LIF)
Day1
Day2

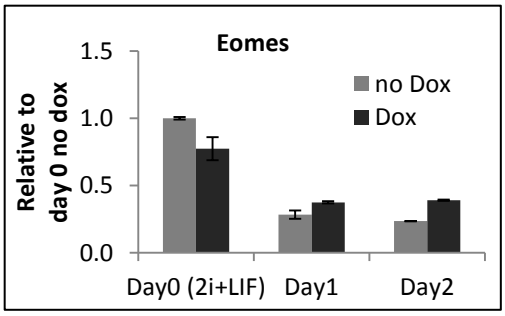

Otx2
Relative to day 0 no dox
6
4
2
0
no Dox
Dox
Day0 (2i+LIF)
Day1
Day2


Fgf5
Relitive to day 0 no dox
2,000
1,500
1,000
500
0
no Dox
Dox
Day0 (2i+LIF)
Day1
Day2

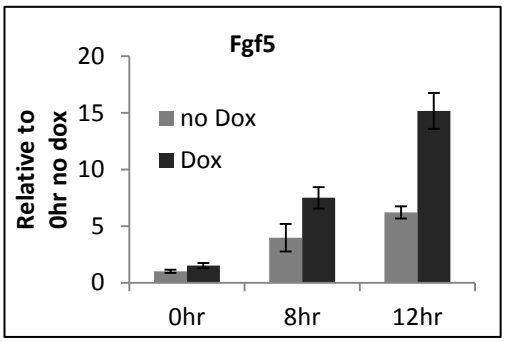

Klf4
Relative to day 0 no dox
1.5
1.0
0.5
0.0
no Dox
Dox
Day0 (2i+LIF)
Day1
Day2


Nanog
Relative to day 0 no dox
1.5
1.0
0.5
0.0
no Dox
Dox
Day0 (2i+LIF)
Day1
Day2


Wnt3
Relative to day 0 no dox
6
4
2
0
no Dox
Dox
Day0 (2i+LIF)
Day1
Day2


Eomes
Relative to day 0 no dox
1.5
1.0
0.5
0.0
no Dox
Dox
Day0 (2i+LIF)
Day1
Day2

B)


Fgf5
Relative to 0hr no dox
20
15
10
5
0
no Dox
Dox
0hr
8hr
12hr

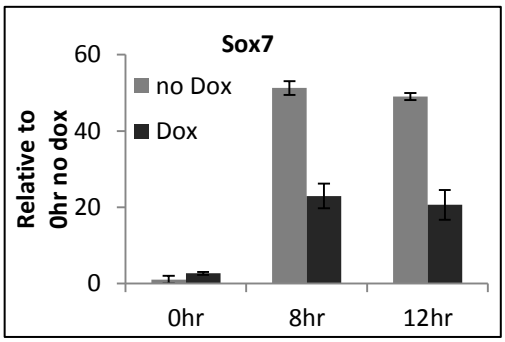
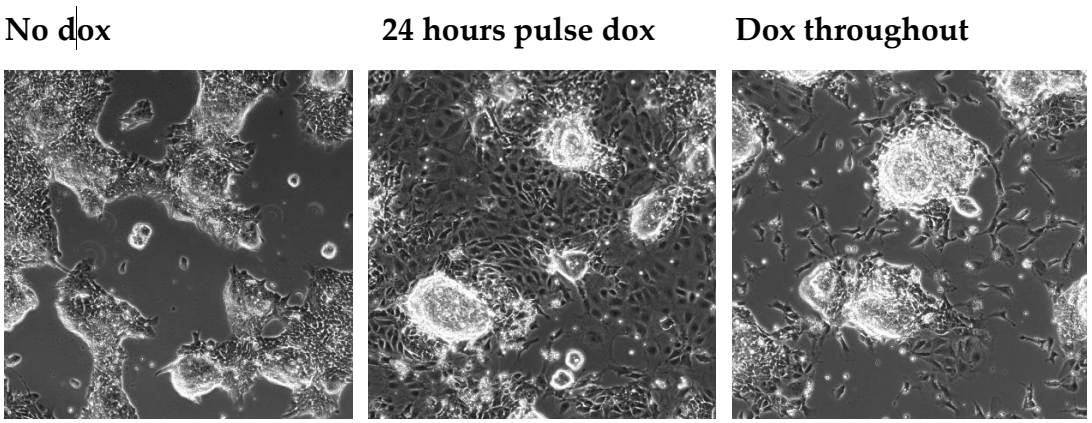

Sox7
Relative to 0hr no dox
60
40
20
0
no Dox
Dox
0hr
8hr
12hr

Figure 4.11: Tcf15 drives the transition to primed epiblast and suppresses cells toward extraembryonic lineage

(A) Day 0 to day 2 qPCR analysis of TTE15 cells in 2i-to-EpiSC transition. (N=2, n=2).

(B) qPCR analysis of TTE15 cells in 2i-to-EpiSC transition at 8 and 12 hours. Induction of Tcf15-E47 at early time point of EpiSC transition up-regulated epiblast marker *Fgf5* expression and suppressed extraembryonic endoderm marker *Sox7* expression. (N=2, n=2).

A)



B)

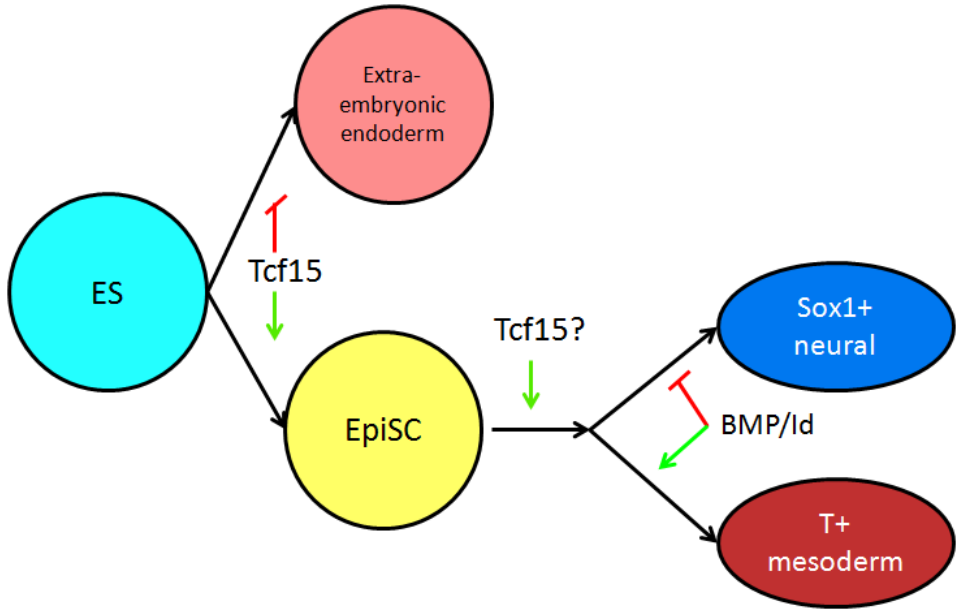


Figure 4.12: Tcf15 drives the transition to primed epiblast

(A) Dox-inducible Tcf15-E47 ES cells after 4 days in N2B27 + LIF + BMP4 only, with 24 hours pulse of dox-treatment during the first day, and in N2B27 + LIF + BMP4 + Dox.

(B) Possible function of Tcf15: Tcf15 promote ES cells transition to EpiSCs by suppression cells toward extraembryonic lineage. Tcf15 might also promote EpiSC to further differentiation, although more experiments need to be performed to validate this hypothesis

4.3.2: Loss of Function studies

4.3.2.1: siRNA knockdown of Tcf15 in mES cells

Loss of function studies were perform to further test the hypothesis that Tcf15 primes pluripotent cells towards epiblast. siRNA was first used to knockdown Tcf15 expression in mES cells. Two strategies of liposome-mediated siRNA knockdown were used, with relative genomic orientation displayed in Figure 4.13:

1) Single hairpin knockdown with 4 hairpins against mouse *Tcf15* from Qiagen, with negative control siRNA conjugated with Alexa-488.
2) Pooled hairpins against mouse *Tcf15* from Dharmacon to achieve reduced off-target effect, with non-targeting negative control siRNA pool.

4.3.2.1.1: Qiagen siRNA knockdown of Tcf15 in mES cells gives poor knockdown efficiency

To test the knockdown efficiency of Qiagen siTcf15 hairpins, E14tg2a cells cultured in complete medium containing LIF and serum were transfected with siTcf15 hairpin 1-4 and negative control hairpin in different combination of siRNA concentrations and transfection reagents according to manufacturer's instruction. A mock transfection was also included in the experiments. Transfection efficiency was observed by the cytosolic expression of Alexa-488 conjugated control siRNA (Figure 4.14A). Cells were collected after 24 hours of transfection for qPCR analysis of *Tcf15* expression. As shown in figure 4.14B, among different combination of siRNA and transfection

reagent HiPerfect (HP), *Tcf15* only had been knockdown in 10 nM siRNA with 6 μL HP treatments. Hairpin 1, 3 and 4 showed 30% of knockdown efficiency compared with Alexa-488 negative control. Therefore, this 10 nM siRNA with 6 μL HP combination was selected to further analyze the knockdown efficiency between 4 hairpins for longer time. However, the knockdown of *Tcf15* was gradually lost after day 2 of transfection, with no significant difference with control and mock cells at day 3 (Figure 4.14C). In order to improve knockdown efficiency, cells were transfected with siRNA in different culture conditions from serum-containing ES complete medium to serum-free N2B27 based medium including LIF + BMP (self-renewing) or FGF2 + Activin (EpiSC) medium (Figure 4.15A). Nevertheless, the knockdown efficiency was poor and always lost after day 2 (Figure 4.15B and 4.15C). Taken together, the knockdown efficiency of Qiagen hairpins were too low to perform any downstream differentiation experiments.

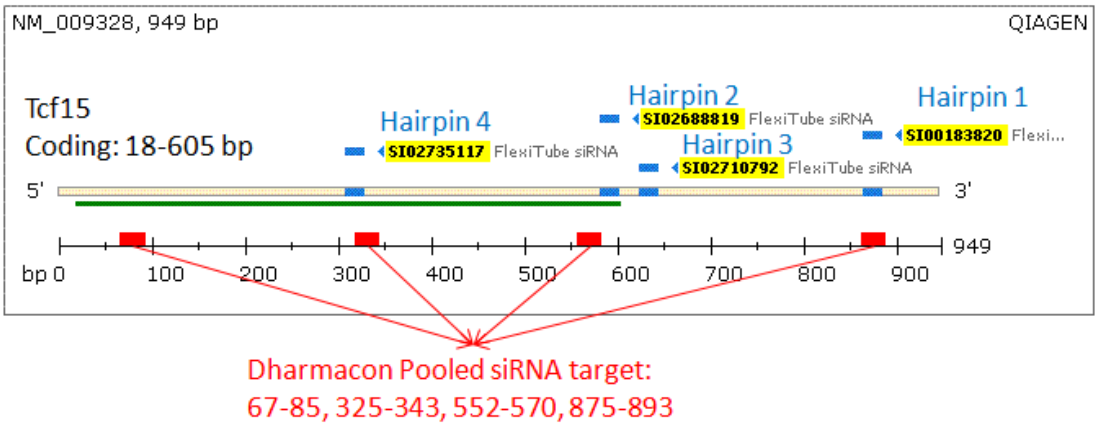
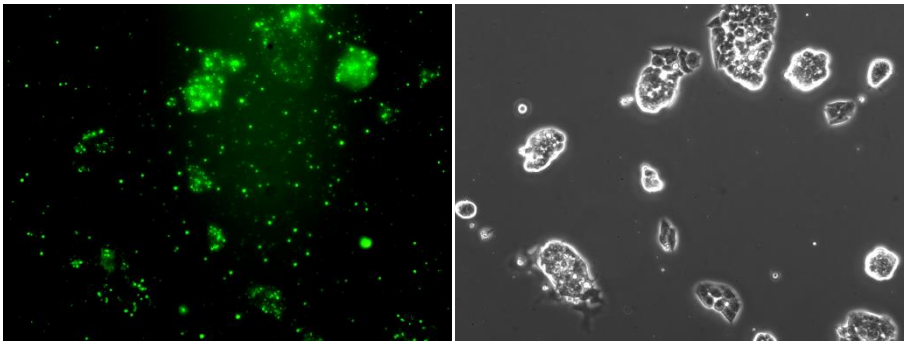


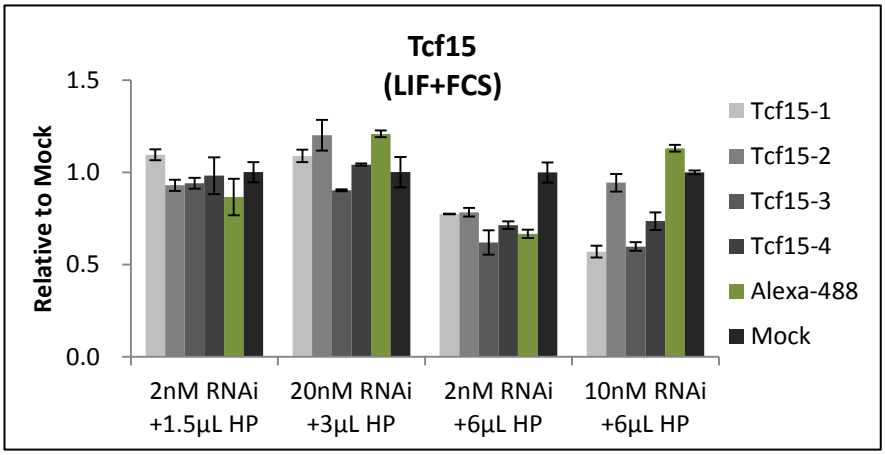
Figure 4.13: Genomic position of Tcf15 siRNAs on *Tcf15* coding region

Mouse *Tcf15* mRNA transcript (NM_009328) consisted of 949 nucleotides. The mature mRNA was shown in green in this figure. GeneSolution *Tcf15* siRNAs 1-4 from Qiagen (sequence unknown) was labelled with blue bars with hairpin numbers indicated by the manufacturer. Dharmacon pooled *Tcf15* siRNAs were labelled as red bars in this figure.

A)



B)



C)

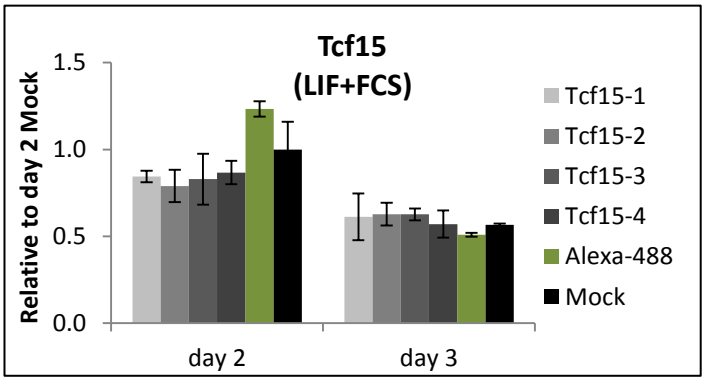


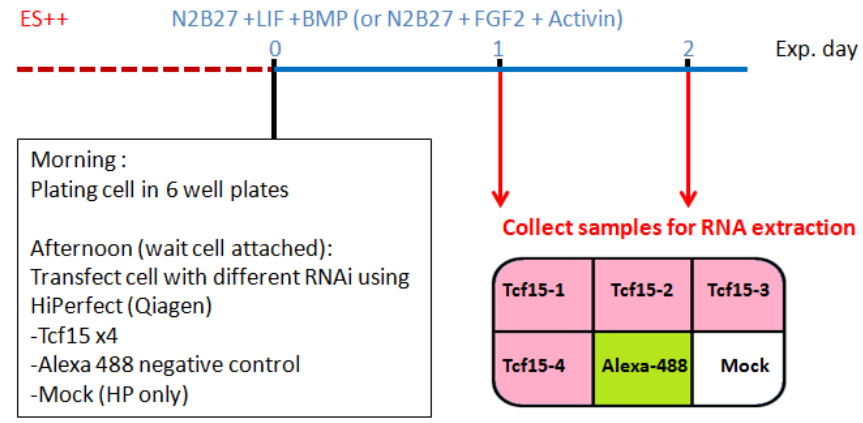
Figure 4.14: Qiagen siRNA knockdown of Tcf15 in serum-containing cultures

(A) Transfection efficiency by Alexa-488 expression within cells, showing here 10nM RNAi with 6µL HP.

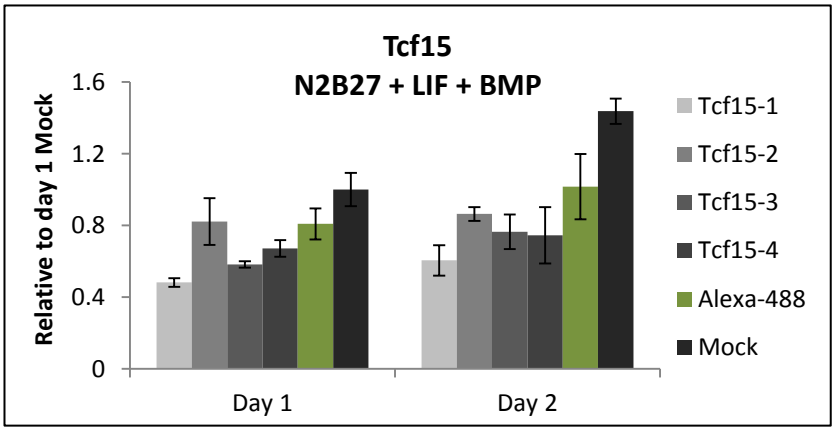
(B) qPCR analysis of *Tcf15* expression after 24 hours of transfection in different combination of siRNA hairpins and transfection reagent HiPerfect (HP).

(C) qPCR analysis of *Tcf15* expression from ES cells transfected with 10 nM siRNA and 6µL HP after 2 and 3 days.

A)



B)



C)

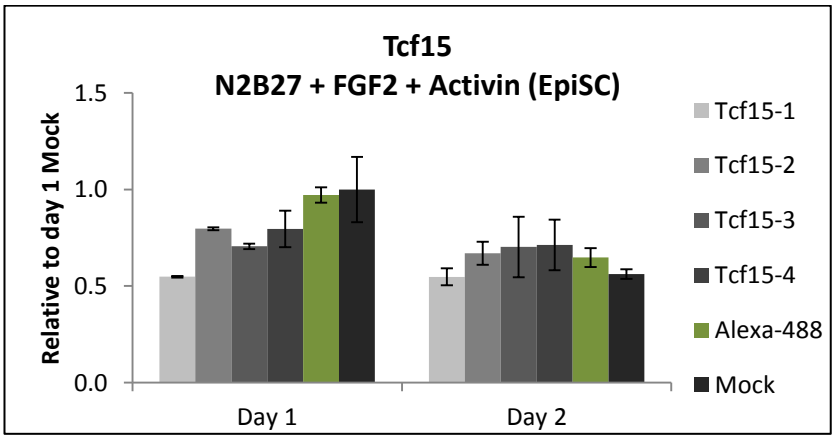


Figure 4.15: Qiagen siRNA knockdown of Tcf15 in serum-free cultures

(A) Experimental design of transfecting cells in serum-free media.

(B) qPCR analysis of *Tcf15* expression on cells transfected with siRNA in N2B27 medium containing LIF and BMP. (N=2, n=2)

(C) qPCR analysis of *Tcf15* expression on cells transfected with siRNA in N2B27 medium containing FGF2 and Activin (EpiSC medium). (N=2, n=2)

4.3.2.1.2: Dharmacon siRNA knockdown of Tcf15 in mES cells gives good knockdown efficiency at mRNA level but not at protein level

Dharmacon smart-pool hairpins against *Tcf15* were purchased after I failed to knockdown *Tcf15* expression using single siRNA hairpins from Qiagen. Pooled siRNA mixture was selected to decrease off-targets effect and enhance siRNA effectiveness, according to manufacturer's instructions. E14tg2a cells were transfected with 25 nM pooled anti-*Tcf15* siRNA or control non-targeting siRNA together with mock transfected cells in complete medium containing LIF and serum according to manufacturer's instruction. Cells were collected after 2 and 3 days of transfection for qPCR analysis. *Tcf15* mRNA was reduced by 75% at 2 days and 30% at day 3 after introducing siRNA into ES cells (Figure 4.16A). The transient knockdown of *Tcf15* was lost after day 2 and expression of *Tcf15* was restored at day 3 after transfection. Tcf15 protein expression was only slightly lower in knockdown cells at day 3 after siRNA transfection (Figure 4.16B). Despite this very moderate reduction of Tcf15 protein expression, I decided to go ahead and test whether this slight reduction in Tcf15 expression would affect the ES to EpiSC transition. Cells were re-plated from LIF + FCS into EpiSC medium after 2 days of siRNA transfection. Although knockdown of *Tcf15* mRNA persisted until day 2 of EpiSC transition (4 days after transfection), *Fgf5* expression only had been slightly down-regulated at day 1 of EpiSC transition. In addition, no significant difference of *Nanog* and *Klf4* expression could be observed (Figure 4.16C). This suggested that although the knockdown on mRNA level was significant, the small and transient

suppression of Tcf15 at protein level might not be enough to generate any significant phenotype. Since functional analysis of Tcf15 was likely to require a stronger and more sustained knockdown during the differentiation procedure, I decided to generate *Tcf15* knockout cell lines.

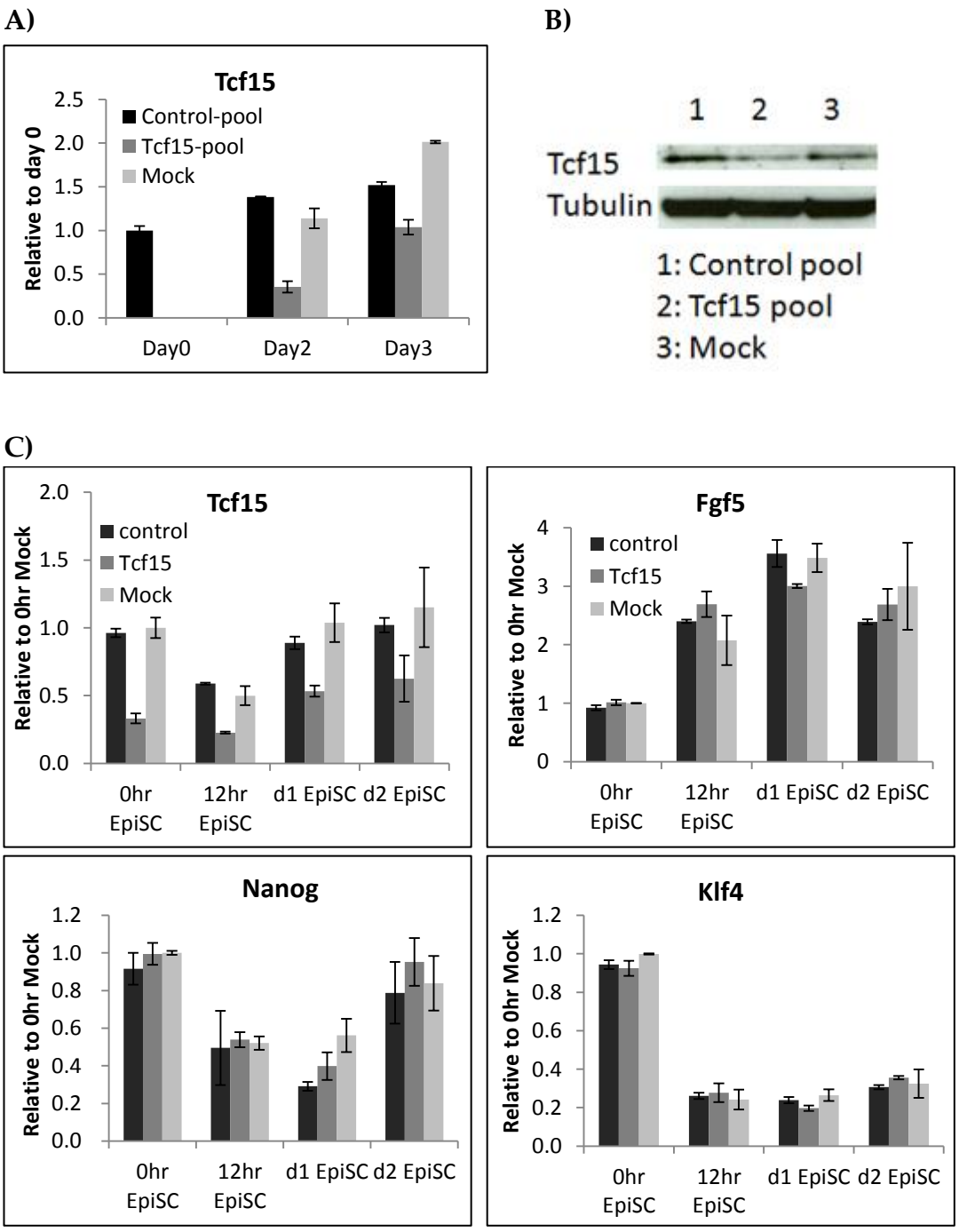


Figure 4.16: Dharmacon pooled siRNA knockdown of Tcf15

(A) qPCR analysis of *Tcf15* expression on E14tg2a cells transfected with 25 nM pooled anti-*Tcf15* siRNA or non-targeting control siRNA.

(B) Western blot analysis of Tcf15 expression after 2 days of siRNA transfection, using tubulin as loading control.

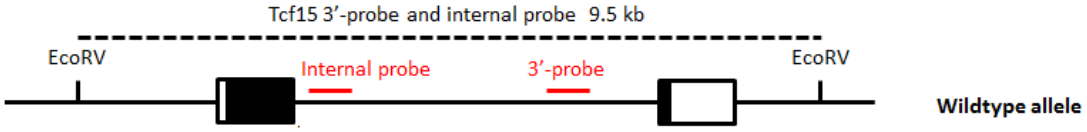
(C) Gene expression from EpiSC transition of siRNA-transfected cells. 0 hour EpiSC time point equaled to 2 days after siRNA transfection in ES complete medium.

4.3.2.2: Generating *Tcf15* knockout cell lines

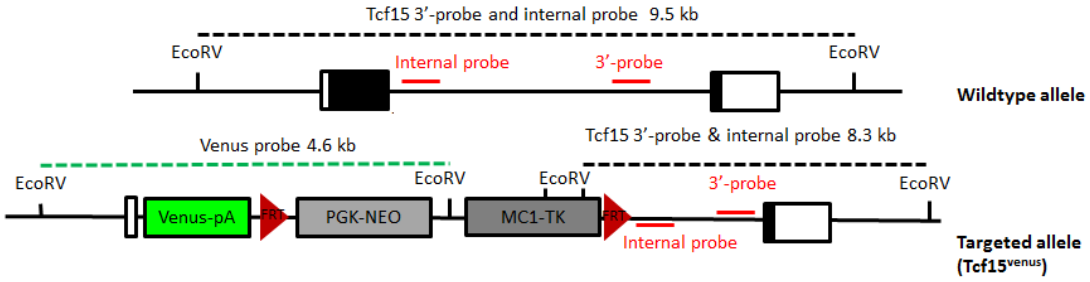
Chapter 3 described Tcf15-Het reporter cells which were generated by introducing the targeting construct pSKO to remove exon 1 of *Tcf15* allele and replace it with the coding sequence for the fluorescence protein *Venus* followed by *FRT*-flanked double selection cassettes including a *PGK-neomycin* resistant gene for positive selection with G418 and a *MC1-TK* (thymidine kinase) for negative selection with ganciclovir (Figure 3.18). This targeting strategy created a heterozygous *Tcf15* allele. Therefore, the same pSKO targeting construct could be used to target the second *Tcf15* allele within Tcf15-Het cells and make *Tcf15* knockout cells.

In this section, gene expression in Tcf15-Het clones was analysed to see whether loss one allele of *Tcf15* has any phenotypic consequence on pluripotent cells. Next, a strategy to remove the selection cassette of Tcf15-Het clones and perform new round of gene targeting was described. All the different targeted *Tcf15* alleles with or without selection cassette are illustrated in figure 4.17.

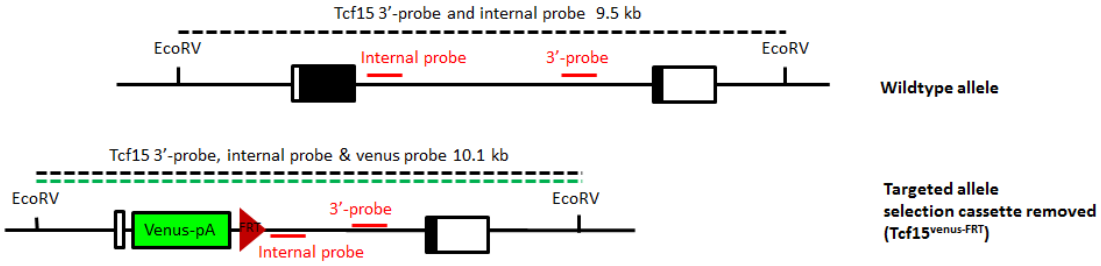
A) Wild-type: *Tcf15*$^{+/+}$



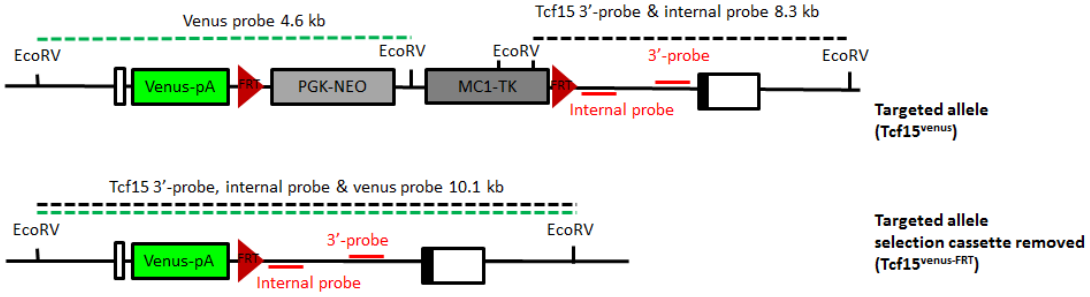
B) Tcf15-Het: *Tcf15*$^{+/Venus}$



C) Tcf15-Het-Frted: *Tcf15*$^{+/VenusFRT}$



D) Tcf15-KO: *Tcf15*$^{Venus/VenusFRT}$



E) Tcf15-KO-Frted: *Tcf15*$^{VenusFRT/VenusFRT}$

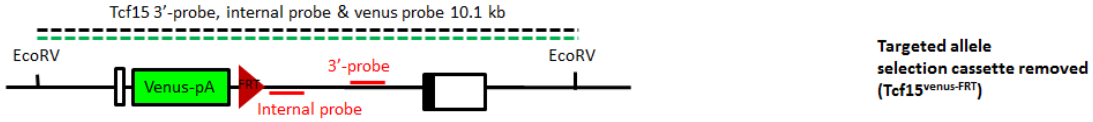


Figure 4.17: Illustration of all the different targeted Tcf15 alleles in ES cells

Genomic orientation of wild-type and *Tcf15*-targeted cells with specific location of southern blot probes: Tcf15-3′, Tcf15-internal and Venus. Venus and Tcf15-internal probe recognized sequence within the targeting construct (showing in figure 3.18). Tcf15-3′ probe recognized sequence outside the targeted region. *Eco*RV: *Eco*RV cutting sites.

(A) Wild-type *Tcf15* allele.

(B) Tcf15-Het cells before removing selection cassette.

(C) Tcf15-Het-FRTed cells are Tcf15-Het cells after removing the selection cassette. This is the Tcf15-Het-Venus reporter cells been analysed in this thesis (clone E6-1).

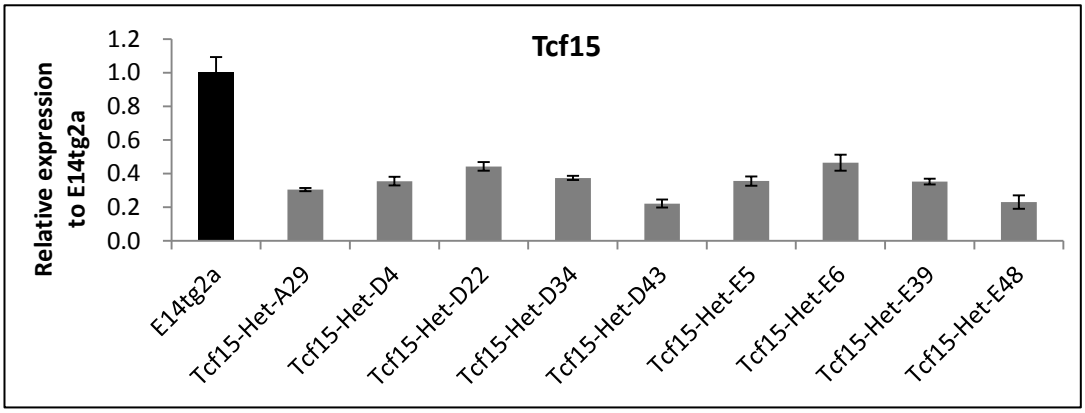
(D) Tcf15 knockout (Tcf15-KO) cells before removing selection cassette.

(E) Tcf15-KO-FRTed cells are Tcf15-KO cells after removing the selection cassette.

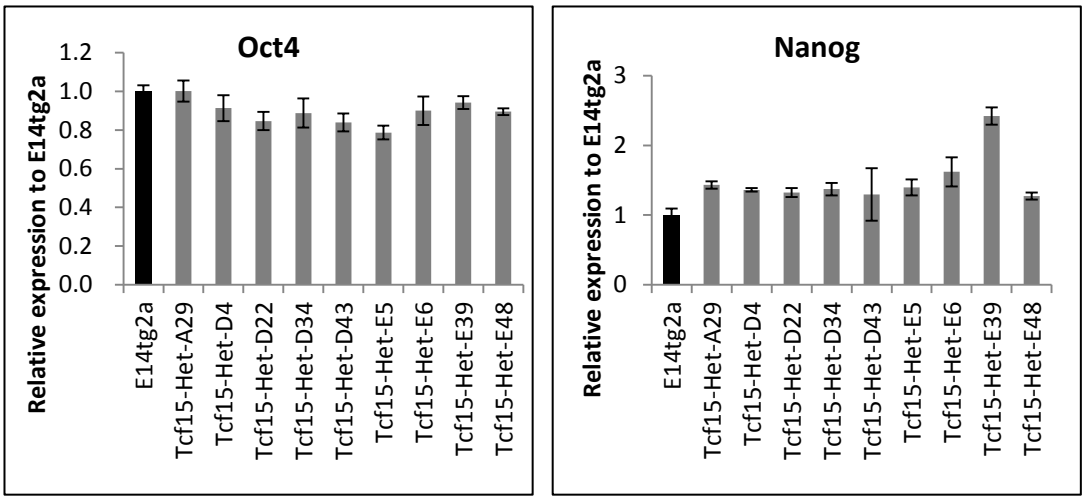
4.2.2.2.1: Expression of Tcf15 in heterozygous Tcf15-Het cell lines

As described in chapter 3, 9 independent *Tcf15* heterozygous clones (Tcf15-Het clones, Figure 4.17B) were obtained after a first round of gene targeting, which were clone A29, D4, D22, D34, D43, E5, E6, E39 and E48 (now referred to as Het-A29, Het-D4, Het-D22, Het-D34, Het-D43, Het-E5, Het-E6, Het-E39 and Het-E48, respectively). Gene expression of Tcf15-Het cells ($Tcf15^{+/Venus}$) was analyzed in order to see if there was any difference between Tcf15-Het cells and wild-type cells. Tcf15-Het cells expressed 30-40% of *Tcf15* transcript compared with parental E14tg2a cells (Figure 4.18A). This result, together with the heterogeneous Venus expression as shown in figure 3.20, indicated that both *Tcf15* alleles were actively transcribed within pluripotent LIF and serum culture and loss of one allele reduced half of *Tcf15* expression compared with parental wild-type control. *Nanog* expression was increased by 30% in Tcf15-Het clones whist *Oct4* expression was not affected (Figure 4.18B). Expression of naïve pluripotency maker *Nanog* and *Klf4* were increased in Tcf15-Het clone E6 and expression of the epiblast marker *Fgf5* was lower (Figure 4.18C). These data are in keeping with the decrease in *Nanog* and *Klf4* and the increase in *Fgf5* observed after overexpression of Tcf15 in ES cells (Figure 4.10A). These data indicate that loss one allele of *Tcf15* expression within pluripotent population may promote a naïve pluripotent state.

A)



B)



C)

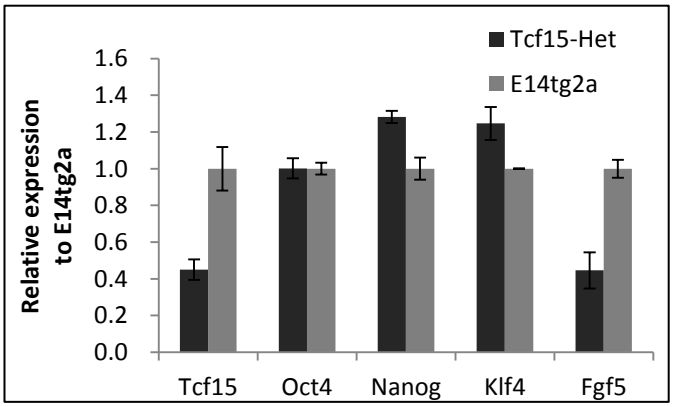


Figure 4.18: Gene expression of Tcf15-Het cells

(A) and (B) qPCR analysis of gene expression on Tcf15-Het cells culturing in LIF and serum. (N=2, n=2)

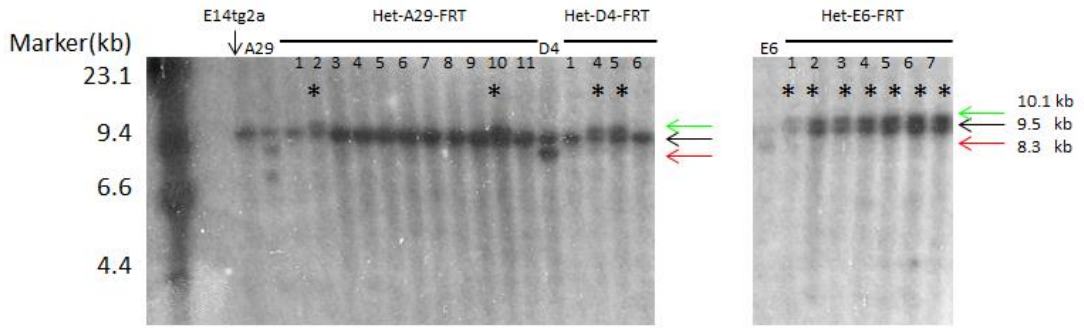
(C) Comparison of gene expression between Tcf15-Het-E6 cells and parental E14tg2a cells culture in LIF and serum.

4.2.3.2.2: Using FlpO/FRT to remove the selection cassette from Tcf15-Het cells

To carry on generating *Tcf15* knockout cell line, 3 independent Tcf15-Het clones A29, D4 and E6 were selected for deletion of the second *Tcf15* allele. The selection cassette was first excised from Tcf15$^{+/Venus}$ cells by FlpO-mediated recombination (details described in section **2.2.4.8**) so that the same targeting vector could be reused for targeting the second allele of the *Tcf15* gene. Only clones which were G418-sensitive and ganciclovir-resistant were selected for further analysis.

Genomic DNA was digested with *Eco*RV to perform Southern blot. Cells losing the selection cassette displayed a 10.1 kb band whilst cells retaining the selection cassette display an 8.3 kb band when probing with Tcf15-internal probe (Figure 4.17 and Figure 4.19A). The blot was stripped and re-probed with Venus probe to better distinguish the *Tcf15^{Venus}* and *Tcf15^{VenusFRT}* alleles, which were 4.6 kb or 10.1 kb respectively (Figure 4.19B). Southern blot analysis of Tcf15-Het-Frted cells (*Tcf15$^{+/VenusFRT}$*, Figure 4.17C) identified 2 clones from Tcf15-Het-A29 (Het-A29-2 and Het-A29-10), 2 clones from Tcf15-Het-D4 (Het-D4-4 and Het-D4-5) and 7 clones from Tcf15-Het-E6 (Het-E6-1 to Het-E6-7) with correct 10.1 kb band for Tcf15-internal probe and Venus probe after *Eco*RV digestion.

A) Tcf15 internal probe



B) Venus probe

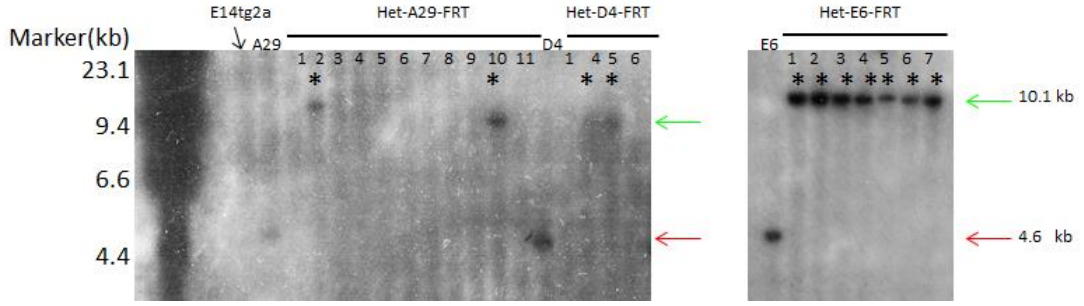


Figure 4.19: Southern blot analysis of Tcf15-Het cells after FlpO transfection to remove the selection cassette (*Tcf15*$^{VenusFRT/+}$ cells)

(A) gDNA from Tcf15-Het-Frt cells, wild-type E14tg2a cells and Tcf15-Het cells (clone A29, D4 and E6) probing for Tcf15 internal probe. In addition to 9.5 kb band for wild-type *Tcf15* allele (black arrow), Tcf15-Het-Frt cells with selection cassette correctly removed will display a 10.1 kb band (green arrow) for *Tcf15*^{VenusFRT} allele instead of an 8.3 kb band (red arrow) for *Tcf15*^{Venus} allele. Clones which had 10.1 kb *Tcf15*^{VenusFRT} allele were marked with *.

(B) gDNA from Tcf15-Het-Frt cells, wild-type E14tg2a cells and Tcf15-Het cells (clone A29, D4 and E6) probing for Venus probe. Wild-type *Tcf15* allele showed no Venus band, Tcf15-Het-Frt cells displayed a 10.1 kb band (green arrow) for *Tcf15*^{VenusFRT} allele instead of a 4.6 kb band (red arrow) for *Tcf15*^{Venus} allele. Clones which had 10.1 kb *Tcf15*^{VenusFRT} allele were marked with *.

Both Tcf15-internal probe and Venus probe identified same Tcf15-Het-Frt clones, which are A29-2, A29-10, D4-4, D4-5 and E6-1-7.

4.2.3.2.3: Generating *Tcf15*-null cells (*Tcf15*$^{Venus/VenusFRT}$)

Tcf15-Het-A29-2, Het-D4-4 and Het-E6-1 cells were electroporated with 100 μg of *Sca*I-linearized pSKO to perform a second round of *Tcf15* gene targeting. Electroporated cells were selected with G418 (300 μg/ml) and over 10^3 colonies were formed by day 8 of selection. Owing to the same targeting construct being used for targeting, there were longer homology arms in the targeted allele than the wild-type allele. Therefore, a lower targeting efficiency of the wild type allele was expected. As a result, approximately 300 clones were picked from each Tcf15-Het-Frted cells and overall 900 clones were expanded from 96 well formats. Clones were passaged to 24 well plates whereupon replica plates were made and frozen at -80°C. Genomic DNA of these Tcf15-KO (KO stands for knockout) cells was isolated and digested with *Eco*RV.

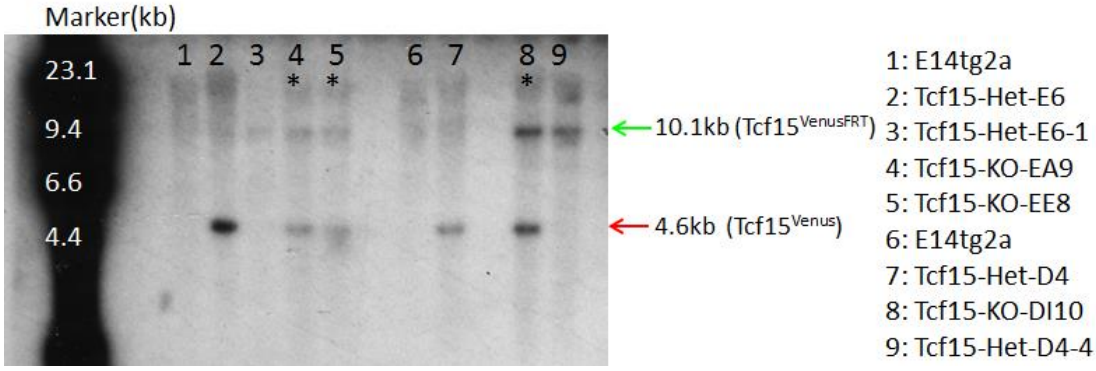
In order to save time and perform high efficiency Southern blot screening of 900 Tcf15-KO clones, 90 digested samples were run in the same agarose gel simultaneously and transferred on a membrane. However, the DNA resolution of agarose gel was reduced due to shortened separation distance. Therefore, instead of probing with Tcf15 probes, membranes were pre-screened by hybridization with the Venus probe which gave a 4.6 kb band for *Tcf15*^{Venus} allele and 10.1 kb band for *Tcf15*^{VenusFRT} allele (Figure 4.17D). A significant proportion of KO cells (around 25%) gave double bands which might indicate correct targeting of the second allele with the pSKO construct (data not shown). However, after probing with Tcf15 3'-probe, only 3 clones displayed correct 8.3 kb band for *Tcf15*^{Venus} allele and 10.1 kb band for

Tcf15^{VenusFRT} allele (Figure 4.20), indicating all the other non-targeted clones with correct Venus bands might be cells with random-integration of targeting construct. Taken together, the targeting efficiency decreased from 5% of Tcf15-Het cells (9/190) to about 0.33% of Tcf15-KO cells (3/900).

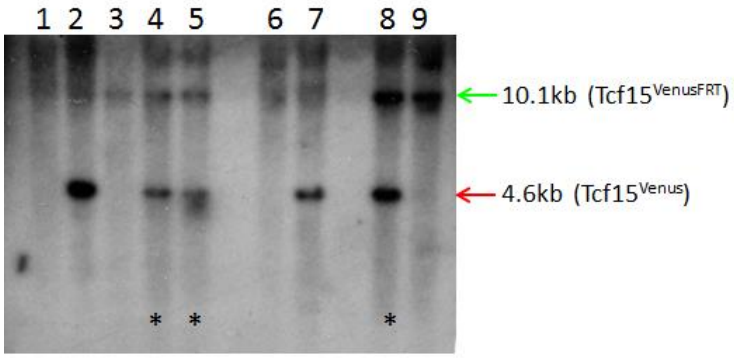
The three Tcf15-KO clones were named as shown in figure 4.21. Clone I10 was made from Tcf15-Het-D4-4 (Now refer to Tcf15-KO-DI10 to indicate the parental Tcf15-Het-D4 line) and clone A9 and E8 were made from Tcf15-Het-E6-1 (Now refer to Tcf15-KO-EA9 and Tcf15-KO-EE8 to indicate the parental Tcf15-Het-E6 line). After further validation (e.g. characterizing if clone DI10 is normal or not, as additional band was identified by Southern blot) of these Tcf15-KO clones, these cells will be analysed in future studies to provide more loss-of-function information of Tcf15 in regulating naïve to primed-epiblast transition.

A) Venus probe

Short exposure

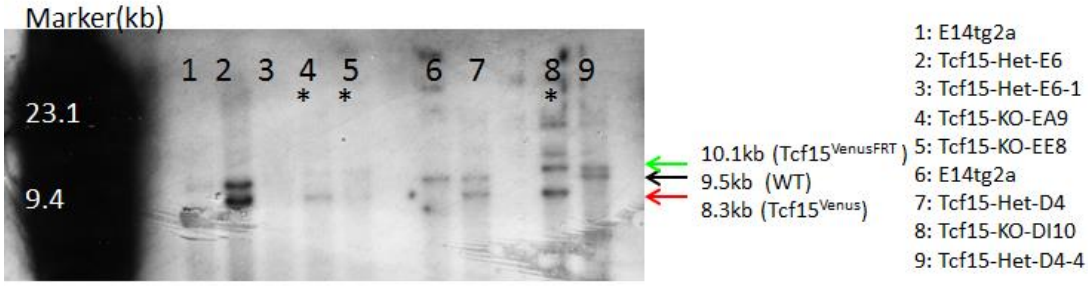


Long exposure



B) Tcf15 3′-probe

Short exposure



Long exposure

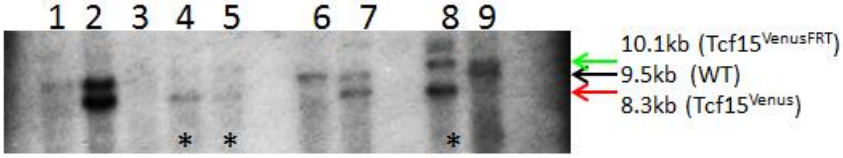


Figure 4.20: Southern blot analyses identified Tcf15-KO cells

(A) gDNA from wild-type E14tg2a cells, Tcf15-Het cells (Het-E6 and Het-D4), Tcf15-Het-Frt cells (Het-E6-1 and Het-D4-4) and Tcf15-KO cells were hybridized with Venus probe. Targeted *Tcf15*^{Venus} allele which retained the selection cassette displayed 4.6kb band (red arrow). *Tcf15*^{VenusFRT} allele which selection cassette had been correctly removed displayed a 10.1 kb band (green arrow). The correctly targeted Tcf15-KO cells (KO-EA9, KO-EE8 and KO-DI10) contained a 10.1 kb band (green arrow) for *Tcf15*^{VenusFRT} allele and a 4.6 kb band (red arrow) for *Tcf15*^{Venus} allele. Corrected Tcf15-KO clones were marked with *.

(B) gDNA from wild-type E14tg2a cells, Tcf15-Het cells, Tcf15-Het-Frt cells and Tcf15-KO cells were hybridized with Tcf15-3' probe. Wild-type *Tcf15* allele displayed a 9.5 kb band (black arrow). Tcf15-Het cells (Het-E6 and Het- D4) which retain the selection cassette displayed 8.3 kb band (red arrow) for *Tcf15*^{Venus} allele and a wild-type *Tcf15* allele. Tcf15-Het-Frt cells (Het-E6-1 and Het- D4-4) with selection cassette correctly removed will display a 10.1 kb band (green arrow) for *Tcf15*^{VenusFRT} allele instead of *Tcf15*^{Venus} allele. The correctly targeted Tcf15-KO cells contained a 10.1 kb band (green arrow) for *Tcf15*^{VenusFRT} allele and an 8.3 kb band (red arrow) for *Tcf15*^{Venus} allele. Corrected Tcf15-KO clones were marked with *.

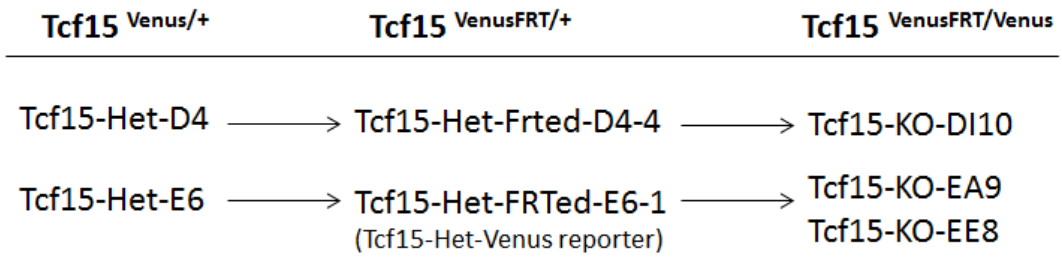


Figure 4.21: Three Tcf15-KO cells from two independent parental Tcf15-Het lines

The three Tcf15-KO cell lines identified by southern blot are named according to their parental Tcf15-Het clones. D stands for Tcf15-Het-D4 and E stands for Tcf15-Het-E6, respectively.

4.4: Discussion:

In this chapter, gain of function studies and preliminary loss of function studies were used to validate the hypothesis that Tcf15 acts as a pro-differentiation factor within pluripotent population, priming mES cells toward somatic but not extraembryonic lineages (Figure 4.1).

4.4.1: Generation of Tcf15-E47 inducible cell lines

To perform gain of function studies, two different inducible systems were selected for investigating the function of Tcf15 in a regulated and reversible manner. However, the FKBP12-Flag-Tcf15-E47 inducible cell lines exhibited some extent of 'leakiness' in the absence of stabilizer Shld1 (Figure 4.3B). This insufficient target protein degradation might due to steric hindrance generated within the fusion FKBP12-Flag-Tcf15-E47 protein. Also, the protein expression was not homogeneous, as seen in figure 4.3B. Therefore, a doxycycline-induction method was selected to over-express Tcf15 within ES cells.

The doxycycline inducible TTE15 cell line displayed negligible leakiness of Tcf15 expression and moderate overexpression in the presence of doxycycline (Figure 4.4A). The addition of doxycycline induced Tcf15 mRNA and protein expression within 24 hours (Figure 4.4B). To obtain sufficient induction of Tcf15 expression at the onset of differentiation, TTE15 cells were pre-treated with doxycycline in complete medium containing LIF and serum for 24 hours before initiating differentiation. Without pre-treatment of

doxycycline to induce Tcf15-E47 expression (i.e. when doxycycline was only added at the time that differentiation was initiated), the promotion of neural differentiation was reduced (data not shown). This is likely because doxycycline takes time to induce protein expression and the up-regulation of Tcf15-E47 after cells proceed to differentiation is less effective compared to pre-treated cells. This is consistent with the idea that Tcf15-E47 is acting at early stage of neural differentiation and serves as a pre-deposited pro-differentiation factor.

4.4.2: Tcf15-E47 accelerates somatic differentiation and suppresses the primitive endoderm lineage

4.4.2.1: Tcf15-E47 fusion protein accelerates ES cells differentiation in neural induction conditions

Forced overexpression of the Id-resistant Tcf15-E47 fusion protein promoted ES cells to differentiate into the neural lineage, as shown by cell morphology and marker gene/protein expression (Figure 4.5 and 4.6). It has been proposed that ES cells neural differentiation occurs in two stages: first from ES cells to differentiation-primed EpiSCs-like stage and then from primed EpiSCs to neural precursor cells (NPCs) (Zhang et al., 2010). These two transitions correspond to the ICM to egg cylinder-epiblast transition and the egg cylinder-epiblast to neural ectoderm transition, respectively.

Tcf15-E47 is effective at accelerating differentiation even if expressed only during the first 24 hours of differentiation (Figure 4.8), suggesting the

function of Tcf15 in promoting ES cells to differentiation-primed EpiSCs-like stage during neural differentiation. It might also be interesting to analyse if Tcf15 could promote EpiSCs to NPCs differentiation, although endogenous expression of *Tcf15* in EpiSCs was much lower than in ES cells cultured in LIF and serum (Figure 3.4A). However, the window of EpiSC to NPC differentiation was hard to capture in the direct ES cells neural differentiation protocol. Therefore, it would be better, in future studies, to generate doxycycline-inducible Tcf15-E47 EpiSC lines *in vitro* and perform direct EpiSC to neural differentiation (Najm et al., 2011).

4.4.2.2: Tcf15-E47 fusion protein suppresses primitive endoderm gene expression

In vitro, ES cells can generate primitive endoderm (PrE)-like cells either in response to LIF withdrawal (Niwa et al., 2000) or through forced expression of the transcription factors Gata4 or Gata6 (Fujikura et al., 2002; Shimosato et al., 2007). These background levels of PrE gene expression might lead to subpopulation of cells differentiating into non-neural cells that morphologically resemble extraembryonic endoderm in direct N2B27 neural differentiation condition. Induction of Tcf15-E47 maintained the robustness of neural differentiation by suppressing differentiation toward primitive endoderm fate (Figure 4.7). Moreover, persistent induction of Tcf15-E47 within ES cells cultured in LIF and serum also down-regulated extraembryonic endoderm marker *Sox7* expression, suggesting Tcf15-E47 push pluripotent cells away from extraembryonic endoderm-primed

population (Figure 4.1).

4.4.2.3: Somatic lineage choice was not biased by activation of Tcf15

The ability of Tcf15 to accelerate differentiation was not restricted to the neural lineage. Somatic lineage choice did not appear to be biased significantly by Tcf15 as measured by the dose-dependent ability of BMP to suppress neural differentiation and instead promote non-neural differentiation (Figure 4.9). Induction of Tcf15-E47 expression promotes neural differentiation in the absence of BMP and promotes non-neural differentiation in the presence of BMP. However, BMP is known to induce Id expression and Id protein acts as negative regulator of Tcf15 activity. Therefore, this doxycycline-induced Tcf15-high and BMP-induced Id-high situation is likely unable to exist *in vivo*. Thus, in future studies, it would be informative to use a BMP-free non-neural differentiation condition such as endoderm differentiation using serum free medium supplemented with Activin in order to provide an Id-low condition to further test if Tcf15 promotes ES cells differentiation in this condition as well.

4.4.2.4: Induction of Tcf15-E47 expression in different pluripotent cultures

The effect of induction Tcf15-E47 expression in pluripotent ES cells cultures, either in LIF + serum condition or in serum-free N2B27 + LIF + BMP condition was investigated. Persistent activation of Tcf15-E47 in LIF + serum

culture didn't drive cells toward differentiation, but significantly down-regulates naïve pluripotency markers expression and up-regulates epiblast marker in ES cells (Figure 4.10). However, induction of Tcf15-E47 expression in N2B27 + LIF + BMP drove irreversible commitment to differentiation in a significant proportion of cells (Figure 4.12A). This indicates that Tcf15-E47 could functionally antagonize the extrinsic anti-differentiation signals. I propose that Tcf15-E47 may drive the loss of LIF responsiveness, which is a characteristic feature of the transition toward post-implantation EpiSCs (Brons et al., 2007; Tesar et al., 2007). Future experiment would be to verify if induction of Tcf15-E47 expression affects phospho-STAT3 activation after LIF stimulation with or without JAK inhibitor pre-treatment (Tesar et al., 2007).

It was also observed that pulse-induction of Tcf15-E47 in ES cells cultured in N2B27 + LIF + BMP displayed distinct phenotype when compared with induction throughout (Figure 4.12A). When continuous inducing Tcf15-E47 expression for 4 days, some cells exited pluripotency and displayed mesenchymal-like morphology. This observation gave the possibility that in addition to triggering the loss of LIF responsiveness, Tcf15 might also promote cells to further differentiation depend on environmental cues. More experiments would be needed to test this hypothesis (Figure 4.12B), for example, as described in section **4.4.2.1**, *in vitro* derivation of EpiSC from TTE15 and induce epiblast-derived lineage differentiation.

4.4.2.5: Tcf15 drives the transition to primed epiblast

It is demonstrated that induction of Tcf15-E47 during 2i to EpiSC transition promoted epiblast marker gene *Otx2* and *Fgf5* expression at day 1 (Figure 4.11A). Also, during early time point of 2i to EpiSC transition, Tcf15-E47 up-regulated *Fgf5* expression and suppressed expression of primitive endoderm marker *Sox7* (Figure 4.11B). The re-activation of *Nanog* and up-regulation of *Wnt3* and *Eomes* expression at later time points indicated that cells were maturating to more proximal posterior pre-gastrulating epiblast-like identity. Induction of Tcf15-E47 did not up-regulated *Nanog* expression at day 2 but dox-treated cells expressed higher levels of *Wnt3* and *Eomes*.

The preliminary loss of function data displayed reciprocal results to the overexpression data: loss of one allele of *Tcf15* expression in Tcf15-Het cells within pluripotent population seems to promote at least some aspects of a naïve pluripotent gene expression signature (Figure 4.18). Based on our results from generation of Tcf15-Het chimaeric mice (Section **3.3.1**), ES cells that have lost one allele of *Tcf15* are capable of proceeding normal developmental process. However, minor effects on the kinetics or efficiency of development in these cells would not have been picked up in these experiments.

Tcf15-null ES cells have not yet been examined. *Tcf15*-null embryos do survive to birth, but die shortly afterward as a consequence of muscular-skeletal defects due to the failure in somite morphogenesis earlier in development (Burgess et al., 1996). Is it possible that there is an earlier

phenotype that has been overlooked? According to my model, we may predict that *Tcf15*-null embryos have a delay in epiblast formation and consequently some disruption in the kinetics of subsequent lineage allocation. However, the early embryo is highly regulative, and such phenotypes may therefore not be readily apparent when examining mutant embryos at late stages *in vivo*. Or, there might be some functional redundancy existed during development between Tcf15 and other HLH transcription factors. This will be further discussed in chapter 6.

The *Tcf15*-KO cells will provide a sensitive tool for studying early function of Tcf15 during peri-implantation stages. *Tcf15*-KO cells express Venus under endogenous *Tcf15* promoter to map the region where *Tcf15* was expressed during development without functional protein expression. Chimaeric embryos generated from lineage-labelled *Tcf15*-KO cells will serve as good tool for future studies to examine whether *Tcf15*-null cells are less capable to transit to post-implantation epiblast or displayed a delay of epiblast formation compared with wild-type cells.

Taken together, in this chapter, it is shown that Tcf15 can accelerate ES cells somatic differentiation by driving the transition to primed epiblast and suppressing primitive endoderm differentiation. Moreover, Tcf15 is acting at early stage of cell differentiation. How Tcf15 can drive pluripotent cells away from naïve pluripotency and toward epiblast? To understand this, the next step would be to investigate the mechanism by which Tcf15 expression becomes associated with this epiblast-primed state, by identifying upstream regulators and downstream targets of Tcf15. These experiments are described in the next chapter.

CHAPTER 5

Mechanism controlling Tcf15 expression and possible downstream targets

5.1: Aims of this chapter

In chapter 4, gain of function and preliminary loss of function studies suggest that Tcf15 is able to prime cells for differentiation by driving the transition to a mature epiblast-like state as well as suppressing differentiation of cells toward extra-embryonic endoderm fate. Since the transcriptional activity of Tcf15 is negatively regulated by Id proteins (Wilson-Rawls et al., 2004), my data may explain why Id is able to suppress the conversion of ES cells into EpiSCs (Zhang et al., 2010). In the experiments described in this chapter, the mechanism by which Tcf15 expression becomes associated with this primed state will be investigated, by identifying upstream regulators and downstream targets of Tcf15.

5. 2: Results

5.2.1: Upstream regulators of Tcf15

5.2.1.1: Tcf15 expression is down-regulated by a MEK inhibitor and a GSK-3β inhibitor

As shown in figure 3.3A, *Tcf15* expression was suppressed when ES cells were cultured in 2i medium, which contained PD0325901 (a MEK inhibitor, referred to hereafter as PD03) and CHIR99021 (a GSK-3β inhibitor, referred to hereafter as Chiron) (Ying et al., 2008). Tcf15-Het reporter cells were used to investigate which of the components of 2i culture is responsible for suppressing *Tcf15* expression. These reporter cells were cultured in complete medium containing LIF and serum in the presence of 1 μM PD03 or 3 μM Chiron for 48 hours before harvesting them for analysis. Flow cytometry analysis revealed that both PD03 and Chiron treatment down-regulated Venus expression (Figure 5.1A), suggesting that *Tcf15* promoter activity was reduced after culturing in these inhibitors. In support of this, qPCR analysis of *Tcf15* mRNA expression showed down-regulation to around 30% under PD03 and Chiron treatment compared with control cells. This down-regulation of *Tcf15* mRNA expression was positively correlated with *Venus* mRNA expression (Figure 5.1B) as expected.

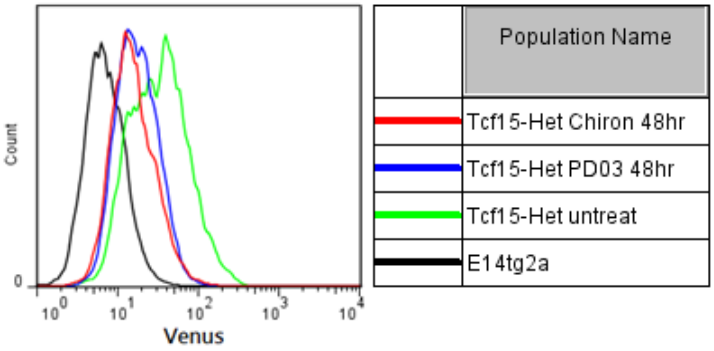
In order to further confirm these results, immunocytochemistry staining of Venus protein and the pluripotent cell marker Oct4 on cultured cells (Figure 5.2A) was performed. Venus expression was relatively weak in PD03-treated cells compared with control cells cultured in LIF and serum. However, Venus expression in Chiron-treated cells appeared to remain at

similar levels to control untreated cells. The much stronger signal in DAPI and Oct4 staining within Chiron-treated cells compared to control untreated cells suggested that this might due to a general accumulation of fluorescent signal from fluorescence microscope in this condition, perhaps due to the change in colony morphology that is induced by Chiron. As seen in the bright field image, cells were very densely packed under Chiron treatment and formed "dome-shaped" colonies. Therefore, the increase of local cell density together with cell stacking is likely to result in the high Venus intensity seen in Chiron-treated cultures.

To further validate Venus expression within cultured cells and to eliminate any artificial increase in fluorescence due to cell stacking, confocal microscopic analysis was performed by Julia Oh, an undergraduate project student within Dr. Sally Lowell lab, under my supervision. Tcf15-Het cells were cultured on glass coverslips in complete medium containing PD03 or Chiron for 48 hours before fixed for immunocytochemistry staining. Cells were stained with Venus and co-stained with Oct4 to identify pluripotent cells. Venus was heterogeneously expressed in cells culturing in LIF and serum (Figure 5.2B). Using this detection method, it became clear that transfer to medium containing PD03 or Chiron both reduced Venus to very low levels. This highlights the importance of using confocal analysis when assessing changes in fluorescence intensity under conditions that may also induce changes in colony morphology.

In conclusion, MEK and GSK-3β signalling are both required for maintaining high levels of Tcf15 expression in ES cells. Due to time constraint, only FGF signalling is further characterized.

A)



B)

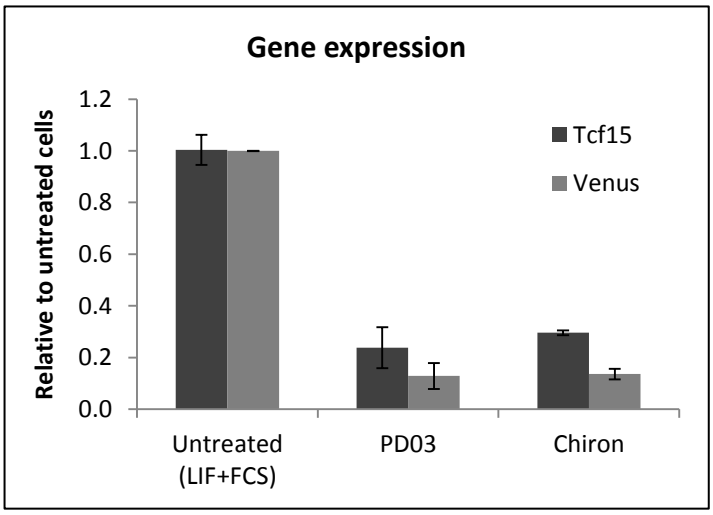


Figure 5.1: Tcf15 expression was down-regulated in PD03 and Chiron treated cultures

(A) Flow cytometry analysis of E14tg2a cells (Venus negative control) and Tcf15-Het cells culturing in LIF and serum (untreated) in the presence of 1 μM PD03 or 3 μM Chiron for 48 hours.

(B) qPCR analysis of *Tcf15* and *Venus* mRNA expression after PD03 and Chiron treatments for 48 hours. Expression of *Tcf15* and *Venus* were down-regulated to a similar extent.

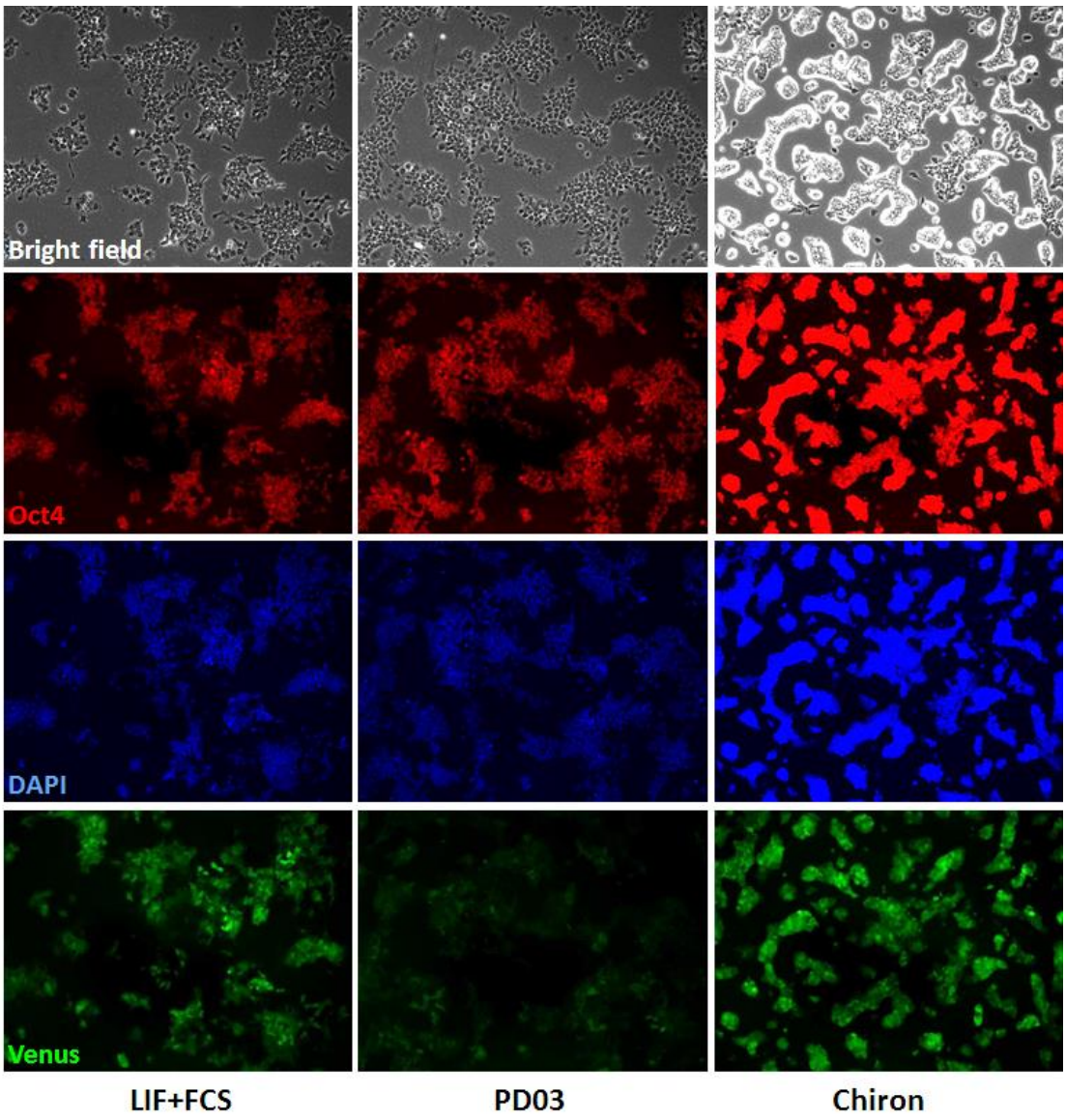


Figure 5.2: Immunofluorescence analysis of Tcf15-Het cells under indicated inhibitor treatments

Tcf15-Het cells were cultured in LIF and serum only, 1 μM PD03 or 3 μM Chiron for 48 hours. Cells were fixed with 4% paraformaldehyde and stained for Oct4 (red), DAPI (blue) and Venus (green). Antibody staining for each condition was performed side by side, and images were taken with the same exposure and identical microscope settings.

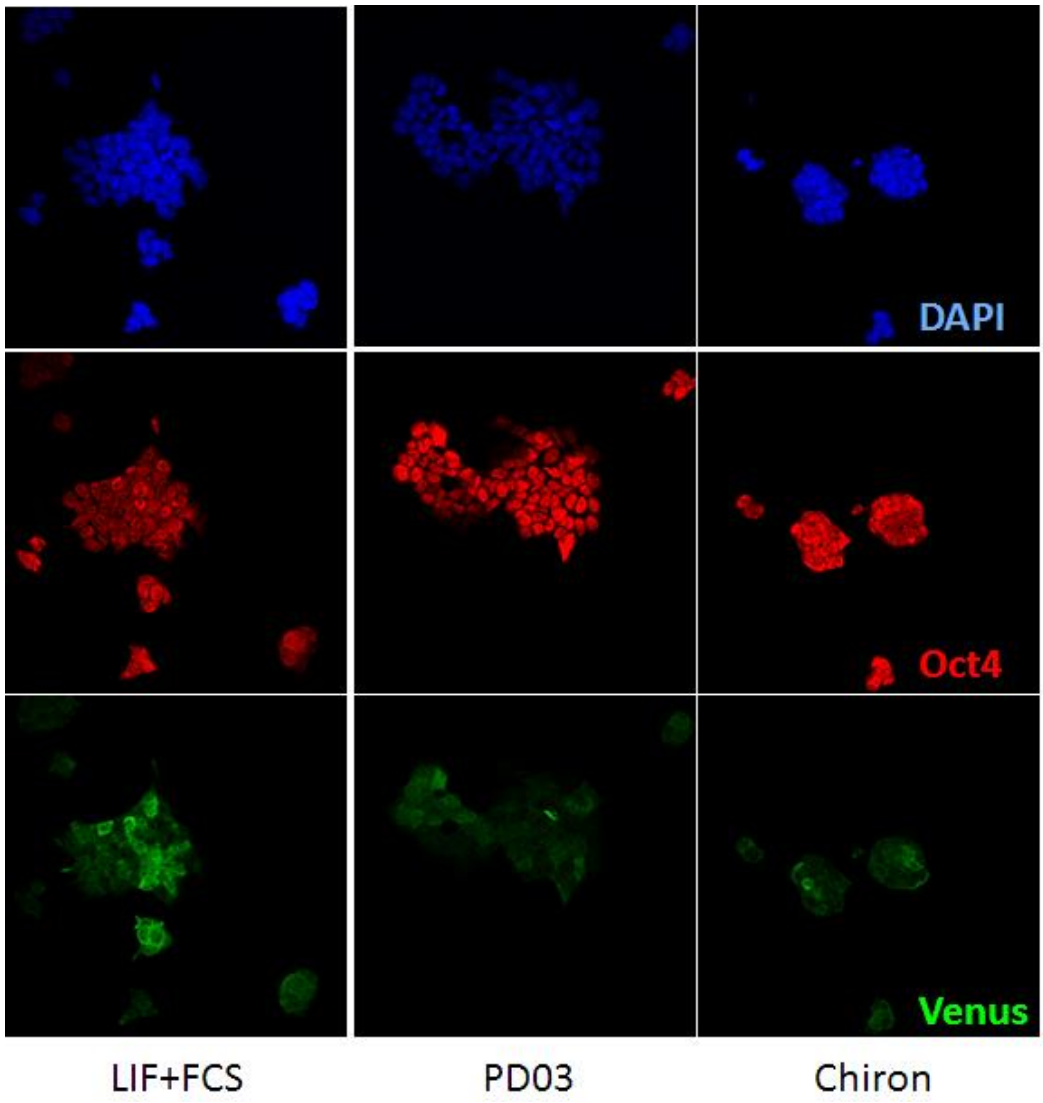


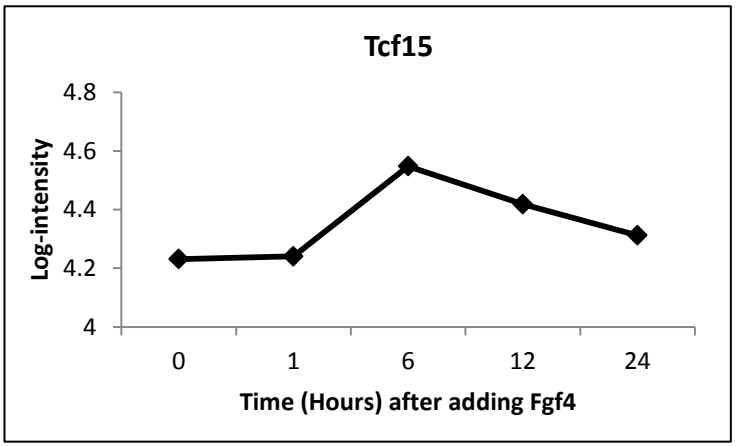
Figure 5.3: Confocal immunofluorescence analysis of Tcf15-Het cells under stated inhibitor treatments

Tcf15-Het cells were cultured in LIF and serum only, 1 μM PD03 or 3 μM Chiron for 48 hours on glass coverslips. Cells were fixed with 4% paraformaldehyde and stained for Oct4 (red), DAPI (blue) and Venus (green).

5.2.1.2: Tcf15 expression is dependent on FGF signalling

FGF signalling is required not only for self-renewing mES cells to differentiate into somatic cell types (Kunath et al., 2007) but also for primitive endoderm differentiation (Nichols and Smith, 2009; Yamanaka et al., 2010). Fibroblast growth factor 4 (FGF4) stimulation of ERK1/2 is an auto-inductive stimulus for naïve mES cells to exit the self-renewal programme and progress to a post-implantation epiblast-like state (Kunath et al., 2007). Previous unpublished microarray data from Dr. Tilo Kunath's lab in University of Edinburgh showed that expression of *Tcf15* increased at 6 hours after adding exogenous FGF4 into *Fgf4*-null ES cell culture (Figure 5.3A). To further validate this result, ES cells were cultured in the presence of 1 μM PD03 or 100 ng/ mL PD173074 (a Fibroblast growth factor receptor inhibitor, referred to hereafter as PD17) for 48 hours, the inhibitors were washed away and then inhibitor-free medium was added back to cultures. The fibroblast growth factor receptor (FGFR) inhibitor PD17 was also able to suppress *Tcf15* expression (Figure 5.3B). The response of Tcf15 to the restoration of autocrine FGF signalling is quite rapid: *Tcf15* is strongly up-regulated within 6 hours of removing FGFR or MEK inhibitors. Taken together, *Tcf15* expression is likely dependent on auto-inductive FGF signalling within pluripotent culture and that it responds rapidly to increases in FGF activity. FGF could therefore prime cells for differentiation by up-regulating the pro-differentiation factor Tcf15, whereas BMP/Id may restrain progression to overt differentiation from this primed state by inhibiting Tcf15 activity (Figure 4.9B).

A)



B)

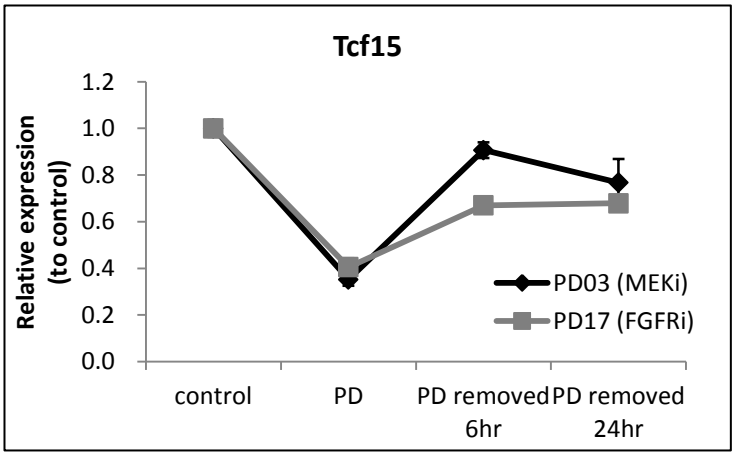


Figure 5.4: Tcf15 expression is under FGF signaling

(A) Microarray data from Dr. Tilo Kunath's lab (unpublished), showing *Tcf15* expression was up-regulated after adding FGF4 into *Fgf4*-null ES cells.

(B) qPCR analysis of ES cells pre-treated 48 hours in PD03- or PD17-containing medium showed that *Tcf15* expression was down-regulated when MEK or FGFR was inhibited, but restored rapidly after 6 hours when inhibitors were removed. (N=2,N=2)

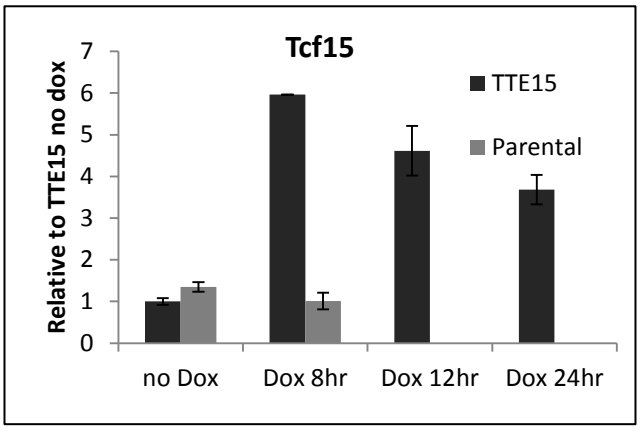
5.2.2: Downstream targets of Tcf15

5.2.2.1: Microarray analysis of possible Tcf15 down-stream targets

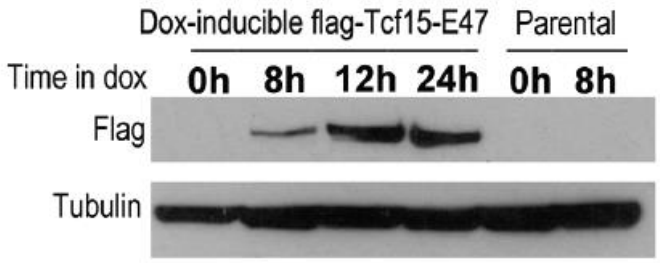
To identify downstream targets of Tcf15, we carried out a time-course microarray analysis aimed at capturing the earliest changes in gene expression in response to Tcf15 activity. Using the TTE15 dox-inducible Tcf15-E47 ES cell line (described in chapter 4, section **4.3.1.2**), it is showed that *Tcf15* mRNA is up-regulated within 8 hours after addition of doxycycline (Figure 5.5A). Tcf15-E47 protein is weakly detectable at around 8 hours after addition of doxycycline and reaches robust levels after 12 hours (Figure 5.5B). Therefore, 8 hours and 12 hours were selected as suitable time points for identifying Tcf15-responsive genes by microarray analysis. A 24 hours' time point was also included in order to profile the longer term response of Tcf15 activation.

An Illumina TotalPrep™ RNA amplification kit was used for generating cRNA from RNA samples for microarray analysis of gene expression according to manufacturer's instruction. cRNA hybridization was performed on a MouseWG-6 v2 BeadChip (Illumina) by the WTCRF MRC Human Genetics Unit (University of Edinburgh). Genes which had an FDR-adjusted p-value of at most 0.05 and a fold change of 1.5 or more for at least one time point in comparison to the 0-hour baseline were considered as differentially expressed. According to these criteria, a total of only 89 genes changed significantly over these early time points, with global transcriptional changes after TTE15 cells stimulation with doxycycline as indicated in figure 5.5C. Figure 5.5D shows the top and bottom 20 genes ranked by fold-change at 12 hours after inducing Tcf15-E47 expression.

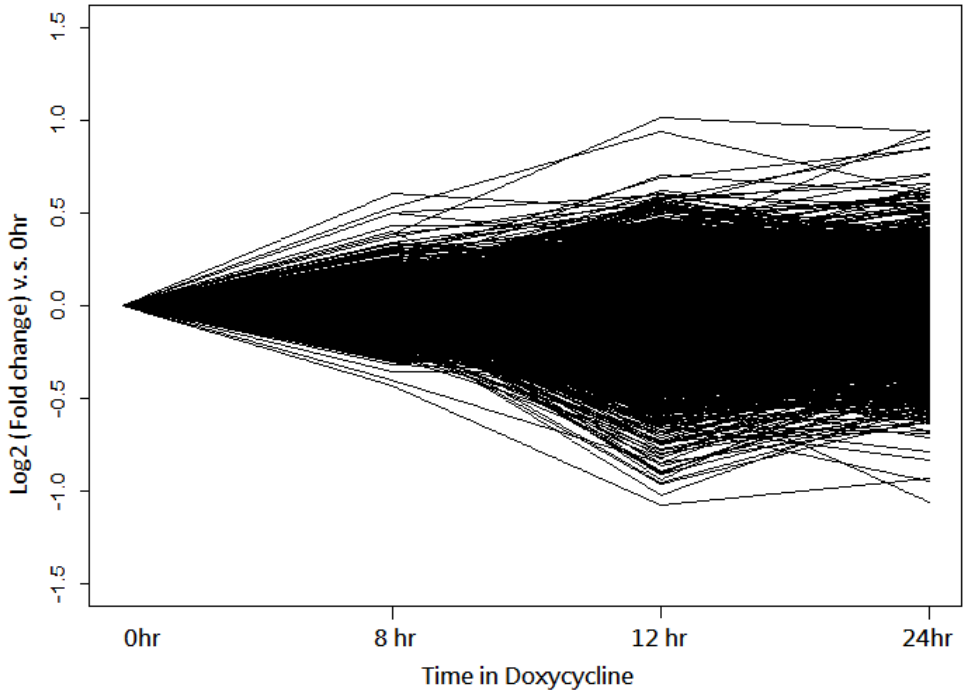
A)



B)



C)



(Figure continued to next page)

D)

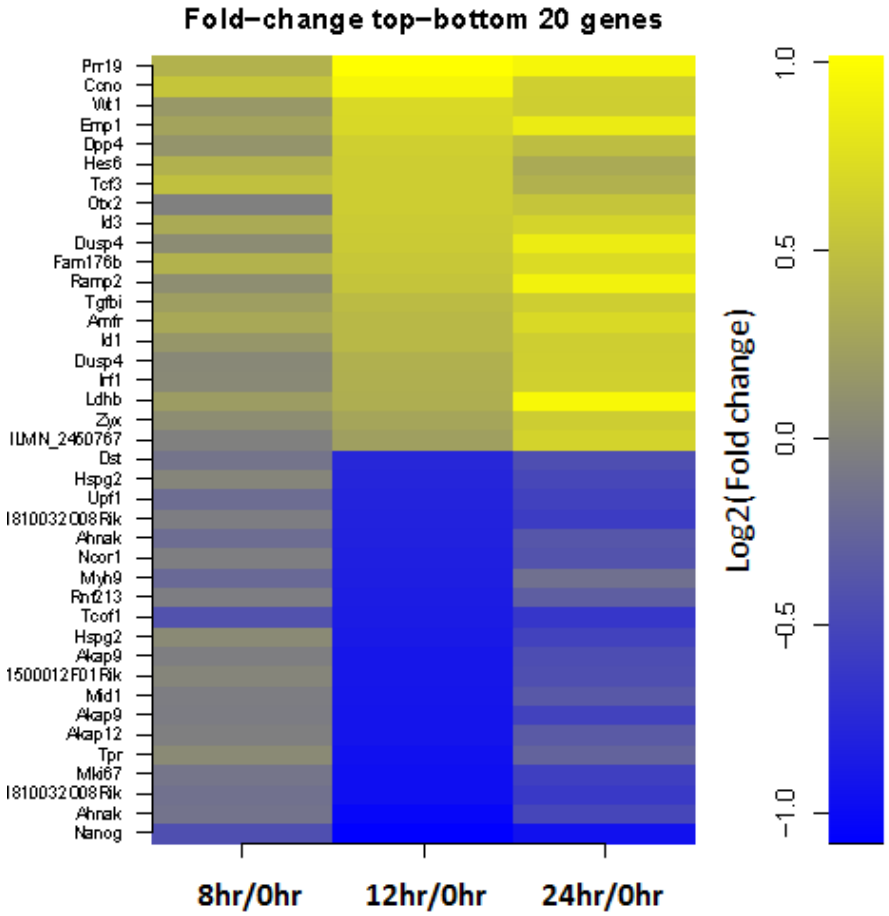


Figure 5.5: Microarray analysis to identify Tcf15-E47 responsive genes

(A) (B) TTE15 and parental control AW2 cells were stimulated by 1 μg/mL doxycycline for the indicated time, Tcf15-E47 induction were confirmed at mRNA (A) and protein level (B).

(C) Microarray analysis was performed on a MouseWG-6 v2 BeadChip (Illumina). Low-quality probes were removed from the input data and data was subsequently quantile-normalized and $\log_2$-transformed before assessing differential expression. Fold change in all genes within the microarray probe set were plot in (C).

(D) Heat-map showing the top and bottom 20 candidate genes which were differentially expressed upon doxycycline-treatment. Genes are ranked by fold change at 12 hours of doxycycline treatment.

The GEO accession number for the microarray data reported in this thesis is GSE42539.

5.2.2.1.1: Up-regulated genes after Tcf15-E47 induction

Table 5.1 displays a list of genes which were up-regulated after inducing Tcf15-E47 expression. qPCR analysis was performed on *Tcf15* and *Tcf3* (*Tcfe2a*) to validate successful induction of *Tcf15-E47* transgene expression (Figure 5.6A). The up-regulation of *Tcf3* (*Tcfe2a*) gene, which generates E47 and E12 proteins through alternative pre-mRNA splicing, was one of the top up-regulated genes, indicating that doxycycline induced *Tcf15-E47* transgene expression as expected. *Tcf15* was, however, not identified as an up-regulated gene in the microarray. This is because the Illumina probe against *Tcf15* (ILMN_1231030) is partially located at the 3'-UTR of *Tcf15* gene, which is absent in the *Tcf15-E47* transgene construct.

To verify whether any of the genes that respond to Tcf15 activation are also enriched in the Tcf15-high subpopulation of ES cells, qPCR analysis was performed on Tcf15-Venus reporter cells sorted by differential Venus expression. The top and bottom 30% of Venus-expression cells were labeled as Venus-high or Venus-low, respectively. Expression of *Tcf15* was first confirmed at higher level in Venus-high cells, whereas the expression of *Tcf3* (*E47*) was similar between Venus-high and Venus-low cells (Figure 5.6B).

To ask whether any of the genes that respond to Tcf15 activation show a similar expression profile to *Tcf15* during differentiation, ES cells were differentiate in monolayer neural conditions. It is first confirmed that the neural marker *Sox1* was up-regulated as expected during the differentiation process (Figure 5.7A), and that *Tcf15* was downregulated after day 1, as previously observed (Figure 5.7B and Figure 3.5A). Since Tcf15 activity is

inhibited by Id1, the expression of *Id1* was also checked in this differentiation time course (Figure 5.7C). *Id1* expression was strongly reduced on day 1, whilst *Tcf15* expression remains high. Therefore, it would be expected that any Tcf15-responsive genes to show a peak of expression at day 1 of neural differentiation.

In order to further validate candidate Tcf15-responsive genes by qPCR and ask whether they are enriched in Tcf15-high ES cells or on day 1 of neural differentiation, genes which were up-regulated by induction of Tcf15-E47 were first manually categorized into 4 groups. The first group encodes transcription factors/regulators, which included *Hes6*, *Id1*, *Id3*, *Irf1*, *Otx2* and *WT1*. The second group encodes signaling modulators, including *Dusp4* and *Dusp6*. The third group encodes adhesion molecules, which includes *Cdh3*, *Itga3*, *Itgb1* and *Zyx*. And the fourth group is genes cannot be categorized as any of the above, including *Ccno*, *Emp1*, *Prr19*, *Raet1b* and *Vegfc*.

Table 5.1: Genes upregulated in response to Tcf15 activation

Illumina Probe ID	Gene Names	Description	log_2FC 8/0hr	log_2FC 12/0hr	log_2FC 24/0hr	qPCR verified Dox-inducible cells	qPCR verified FACS sorted Tcf15-reporter
ILMN_1216541	Prr19	Proline rich 19	0.3851*	1.0142*	0.9380*	Yes	=
ILMN_2736471	Ccno	Cyclin O, Ung2	0.5332*	0.9383*	0.6154*	Yes	−
ILMN_2470646	Wt1	Wilm's tumor 1 homolog	0.1717	0.7024*	0.6061*	Yes	+
ILMN_2642913	Emp1	Epithelial membrane protein 1	0.2593	0.6907*	0.8447*	Yes	=
ILMN_2615096	Dpp4	Dipeptidylpeptidase 4	0.1393	0.6189*	0.4734*		
ILMN_2791952	Hes6	Hairy and enhancer of split 6 (Drosophila)	0.3703	0.6025*	0.3163	Yes	+
ILMN_2596278	Tcf3	Transcription factor E2a	0.4973	0.6016*	0.3698	Yes	=
ILMN_2691752	Otx2	Orthodenticle homolog 2 (Drosophila)	-0.0244	0.5989*	0.5397*	Yes	−
ILMN_2687169	Id3	Inhibitor of DNA binding (differentiation) 3	0.3137	0.5848*	0.6617*	Yes	=
ILMN_2689731	Dusp4	Dual specificity phosphatase 4	0.0737	0.5764*	0.8564*	Yes	+
ILMN_2890357	Fam176b	Family with sequence similarity 176, member B	0.3848	0.5559*	0.7123*		
ILMN_3079461	Raet1b	Retinoic acid early transcript 1 beta	0.2525	0.5453*	0.0667*	Yes	+
ILMN_2661422	Ramp2	Receptor (calcitonin) activity modifying protein 2	0.0884	0.5283*	0.9055*	Yes	
ILMN_2486573	Vegfc	Vascular endothelial growth factor C	0.3180	0.5200*	0.3912*	Yes	+
ILMN_2834379	Tgfbi	Transforming growth factor, beta induced	0.2138	0.4519*	0.6093*		
ILMN_2722902	Amfr	Autocrine motility factor receptor	0.3017	0.4308*	0.7006*		
ILMN_2672190	Id1	Inhibitor of DNA binding (differentiation) 1	0.1520	0.4281*	0.6107*	Yes	+
ILMN_2870965	Dusp4	Dual specificity phosphatase 4	0.0354	0.3623*	0.6152*	Yes	+

***: p-value < 0.05**

(+): positive correlation, (-): negative correlation, (=): no difference, between Tcf15-Het-Venus and target gene expression

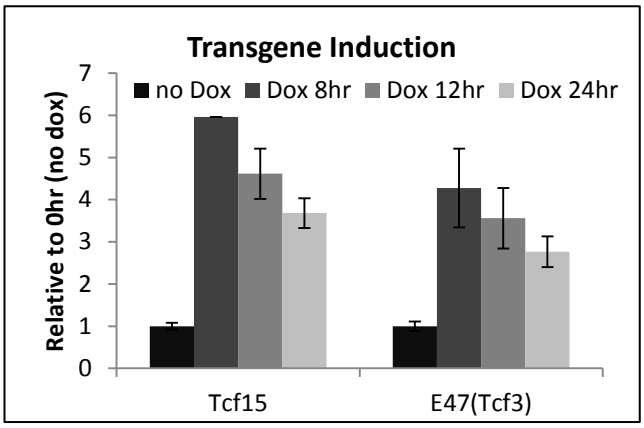
Table 5.1: Genes upregulated in response to Tcf15 activation (continued)

Illumina Probe ID	Gene Names	Description	log_2FC 8/0hr	log_2FC 12/0hr	log_2FC 24/0hr	qPCR verified Dox-inducible cells	qPCR verified FACS sorted Tcf15-reporter
ILMN_2834777	Irf1	Interferon regulatory factor 1	0.0476	0.3541*	0.6280*	Yes	=
ILMN_2660466	Ldhb	Lactate dehydrogenase B	0.2030	0.3451*	0.9496*		
ILMN_2616164	Itga3	Integrin, alpha 3 (antigen CD49C)	0.0384	0.3413	0.2416*	Yes	+
ILMN_2514292	Zyx	Zyxin	0.0759	0.2767	0.6047*	Yes	+
ILMN_3156246	Cdh3	Cadherin 3	0.3109	0.2743	0.5327*	Yes	−
ILMN_2450767		ILMN_2450767	-0.0295	0.2300	0.6516*		
ILMN_1226546	Iffo2	Intermediate filament family orphan 2	0.1218	0.1582	0.5851*		
ILMN_2708477	Spink3	Serine peptidase inhibitor, Kazal type 3	0.1111	0.1577	-0.6103*		
ILMN_1248537	Dusp6	Dual specificity phosphatase 6	-0.0990	0.1045	0.4921*	Yes	+
ILMN_2789077	Itgb1	Integrin, beta 1 (antigen CD29)	-0.0141	0.0066	0.1443*	Yes	+

***: p-value < 0.05**

(+): positive correlation, (-): negative correlation, (=): no difference, between Tcf15-Het-Venus and target gene expression

A) Expression in doxycycline-treated inducible TTE15 cells



B) Expression in FACS sorted Tcf15-Het reporter cell line

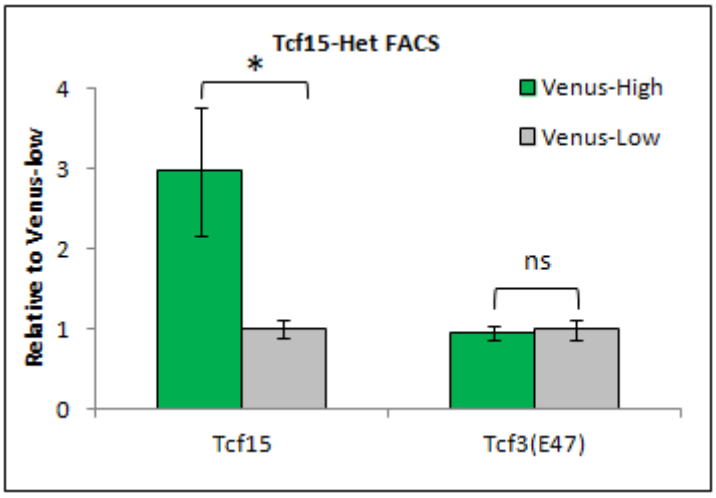


Figure 5.6: qPCR validation of Tcf15 and E47 expression

(A) Expression of *Tcf15* and *E47* (showing in Table 5.1 as *Tcf3*) upon doxycycline induction of Tcf15-E47 expression in TTE15 cells. *Tcf15* and *E47* were up-regulated at similar level.

(B) qPCR analysis of *Tcf15* and *E47* on FACS sorted Tcf15-Venus-high and Tcf15-Venus-low cells. *Tcf15* was differentially expression between V-high and V-low cells and *E47* was not enriched in Tcf15-Venus-high population. ns: non-significant; *: $p<0.05$.

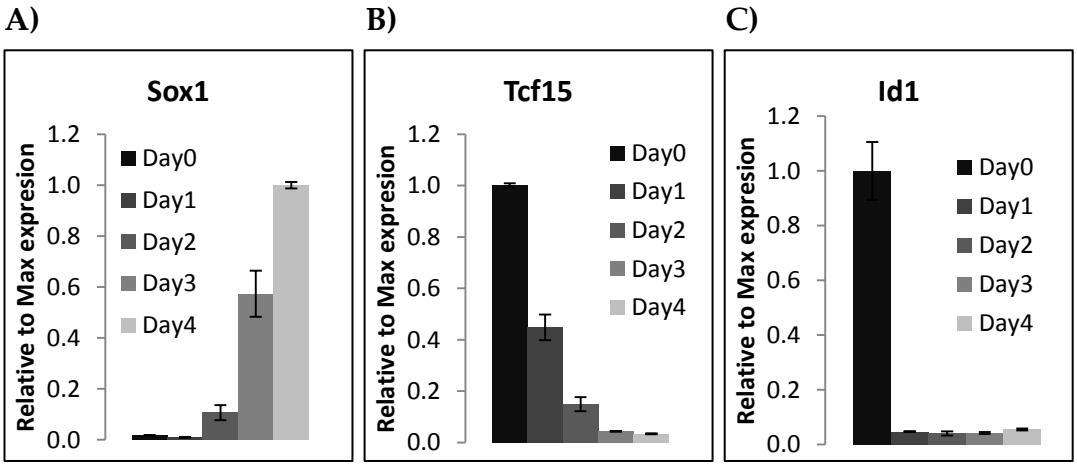


Figure 5.7: Gene expression during neural differentiation

qPCR analysis of *Sox1, Tcf15* and *Id1* during serum-free monolayer neural differentiation. *Id1* expression was strongly reduced on day 1, whilst *Tcf15* expression remains high. Expression of *Tcf15* was down-regulated after day 2 of neural differentiation, before up-regulation of *Sox1*.

5.2.2.1.1.1: Transcription factors/regulators up-regulated after Tcf15-E47 induction

Hes6 (hairy and enhancer of split 6), Id1 (Inhibitor of differentiation/DNA-binding 1), Id3 (Inhibitor of differentiation/DNA-binding 3), Irf1 (Interferon regulatory factor 1), Otx2 (Orthodenticle homeobox 2) and *WT1 (Wilms' tumor 1)* were transcription factors/regulators up-regulated after inducing Tcf15-E47 expression (Table 5.1).

qPCR analysis on doxycycline-treated TTE15 cells was performed to validate the response of these genes to Tcf15 activation (Figure 5.8A). All of these transcription factors/regulators could be detected by qPCR in ES cells. *Hes6* responded rapidly upon Tcf15 induction from 8 hours and persisted until 24 hours. *Id1, Id3, Irf1, Otx2* and *WT1* expression were gradually up-regulated from 8 hours to 12 hours after adding doxycycline.

Expression of these genes in FACS sorted Tcf15-high and Tcf15-low subpopulations of ES cells were then examined. *Id3, Irf1* and *Otx2* were almost equally expressed between Venus-high and Venus-low cells. Expression of *Hes6, Id1* and *WT1* were enriched in Venus-high cells (Figure 5.8B).

Gene expression during N2B27 monolayer neural differentiation was then analyzed in order to ask which if these genes shows a similar expression profile to *Tcf15* (Figure 5.8C). *Id1* was down-regulated immediately when cells began differentiation, as shown in figure 5.7C. *Irf1* expression was also down-regulated during neural differentiation. *Id3* was down-regulated at day 1 and day 2 of neural differentiation and expression was restored to

levels similar to ES cells after day 3. Expression of *Hes6*, *Otx2* and *WT1* were up-regulated at day 1, correlating with the up-regulation of *Tcf15* during early stages of neural differentiation. *Hes6* and *WT1* expression were down-regulated when cells started express neural marker *Sox1* (Figure 5.7A). Expression of *WT1* was lost when cells committed to differentiation, with similar pattern as *Tcf15* (Figure 5.7B). *Hes6* and *Otx2* were still expressed at day 4, consistent with their reported roles in regulating neural development (Acampora et al., 1995; Bae et al., 2000).

Table 5.2 summarized genes of interest which fulfill at least two of our analyses: responds to Tcf15 activation (gene expression changed for 1.5 fold or more for at least one time point in comparison to the un-induced cells), enriched in Tcf15-high subpopulation or displayed similar expression profile to Tcf15 during neural differentiation. Therefore, *Id1*, *Hes6*, *Otx2* and *WT1* are considered as possible Tcf15 targets for future study.

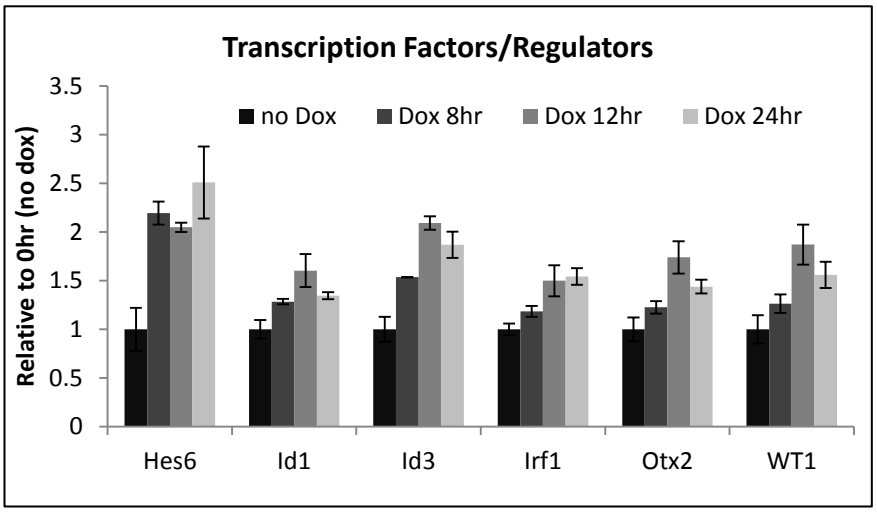
Figure 5.8: Gene expression of transcription factors/regulators up-regulated after induction of Tcf15-E47 expression

(A) Gene expression after doxycycline induction of Tcf15-E47 expression in TTE15 cells.

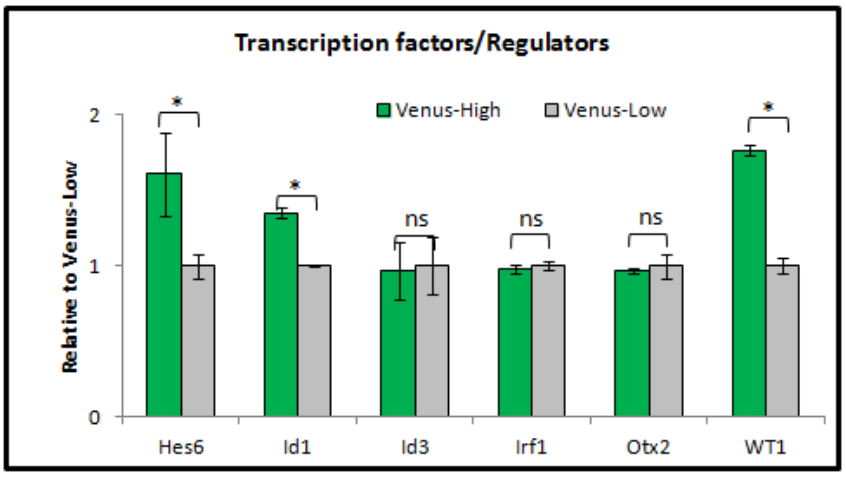
(B) Gene expression of FACS sorted Tcf15-Venus-high and Tcf15-Venus-low cells from Tcf15-Het cells. ns: non-significant; *: $p<0.05$.

(C) Gene expression during serum-free monolayer neural differentiation. *Id1* expression is the same data as displayed in figure 5.7C.

A) Expression of Tcf15-responsive genes in doxycycline-treated inducible TTE15 cells



B) Expression of Tcf15-responsive genes in FACS sorted Tcf15-Het reporter cell line



C) Expression of Tcf15-responsive genes during neural differentiation

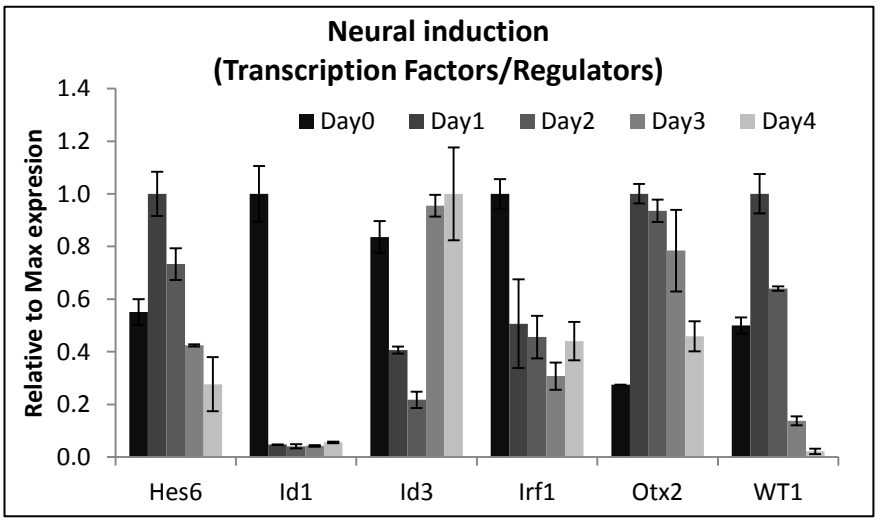


Table 5.2: Summarizing results from transcription factors/regulators of interest

Gene Of Interest	Id1	Hes6	Otx2	WT1
Responds to Tcf15 activation	Yes	Yes	Yes	Yes
Enriched in Tcf15-high subpopulation	Yes	Yes	No	Yes
Similar profile to Tcf15 during neural differentiation	No	Yes	Yes	Yes

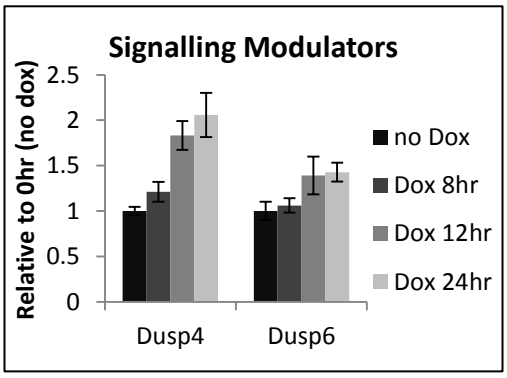
5.2.2.1.1.2: Signaling modulators up-regulated after Tcf15-E47 induction

Dual specificity phosphatases 4 (*DUSP4*) and *DUSP6* were two signaling molecules up-regulated after inducing Tcf15-E47 expression (Table 5.1). DUSPs are members of the mitogen-activated protein kinase phosphatase (MKP) family and potential tumor suppressor. DUSPs inhibit MAPK (mitogen-activated protein kinases) signaling by dephosphorylating activated MAPKs, thus negatively regulate the MAPKs including ERK, p38 and JNK (Keyse, 2008).

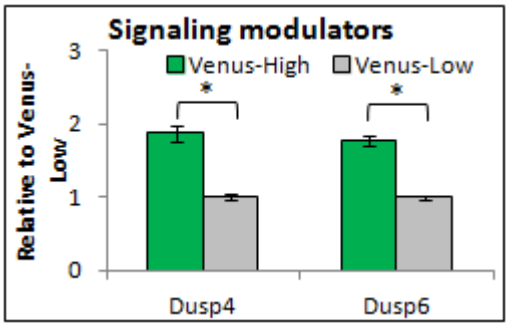
DUSP4 and *DUSP6* were both up-regulated upon Tcf15-E47 induction (Figure 5.9A) and preferentially expressed in FACS sorted Tcf15-Venus-high population (Figure 5.9B). Expression of *DUSP4* was up-regulated at day 1 of neural differentiation and remained at levels similar to undifferentiated cells from day 2 onward (Figure 5.9C). *DUSP6* was down-regulated during neural differentiation.

Table 5.3 summarized *DUSP4* and *DUSP6* expression according to the analyses performed above. *Dusp4* and *Dusp6* are considered as possible Tcf15 responsive genes.

A) Expression of Tcf15-responsive genes in doxycycline-treated inducible TTE15 cells



B) Expression of Tcf15-responsive genes in FACS sorted Tcf15-Venus reporter cell line



C) Expression of Tcf15-responsive genes during neural differentiation

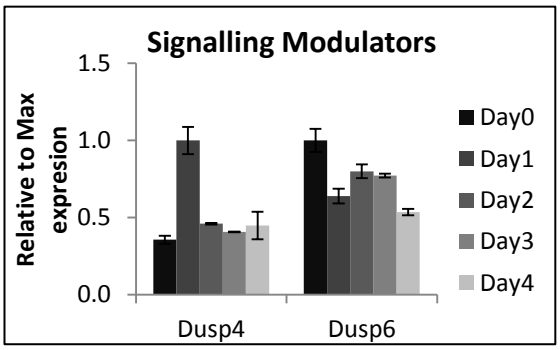


Figure 5.9: Gene expression of signalling modulators up-regulated after induction of Tcf15-E47 expression

(A) Gene expression of *DUSP4* and *DUSP6* after doxycycline induction of Tcf15-E47 expression in TTE15 cells.

(B) Gene expression of FACS sorted Tcf15-Venus-high and Tcf15-Venus-low cells, showing both *DUSP4* and *DUSP6* were enriched in Tcf15-Venus-high population in Tcf15-Venus reporter cells. *: $p<0.05$.

(C) Gene expression during serum-free monolayer neural differentiation.

Table 5.3: Summarizing results from signaling modulators of interest

Gene Of Interest	Dusp4	Dusp6
Responds to Tcf15 activation	Yes	Maybe yes (nearly 1.5 fold increase)
Enriched in Tcf15-high subpopulation	Yes	Yes
Similar profile to Tcf15 during neural differentiation	Yes (only at day 1)	No

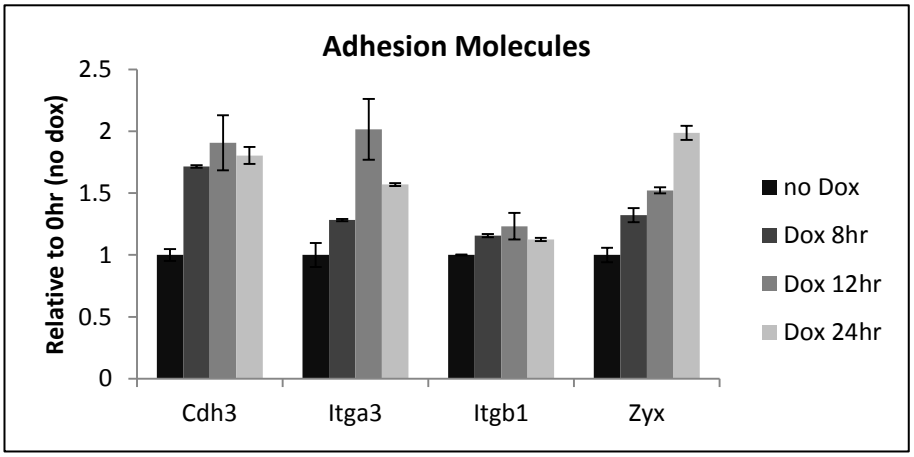
5.2.2.1.1.3: Adhesion molecules up-regulated after Tcf15-E47 induction

Cdh3 (Cadherin 3, P-cadherin), *Itga3 (Integrin alpha3), Itgb1 (Integrin beta1)* and *Zyx (Zyxin)* were adhesion molecules up-regulated after inducing Tcf15-E47 expression.

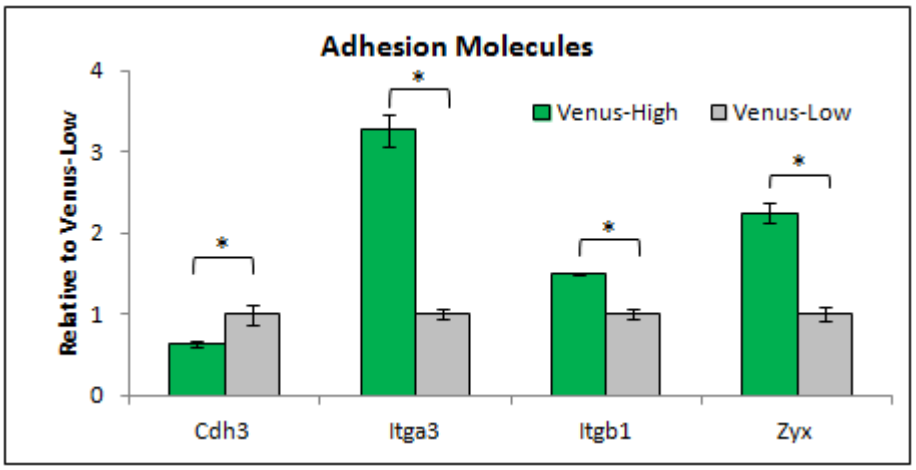
qPCR analysis showed that *Cdh3* and *Zyx* were up-regulated upon doxycycline induction of Tcf15-E47 expression (Figure 5.10A), whilst a modest change in *Itgb1* was seen. *Itga3*, *Itgb1* and *Zyx* were enriched in Tcf15-Venus-high cells. *Cdh3* expression was, however, lower in Tcf15-Venus-high cells, showing negative correlation with *Tcf15* expression (Figure 5.10B).

All the genes examined in this section showed dynamic changes during neural differentiation (Figure 5.10C), with cells progressively switching from an *E-Cadherin (Cdh1)*-high, *P-Cadherin (Cdh3)*-high state to an *N-Cadherin (Cdh2)*-high state (Figure 5.10D). Of these genes, none showed a peak of expression at day 1, although *P-Cadherin* was downregulated with similar kinetics to *Tcf15* (Figure 5.7B).

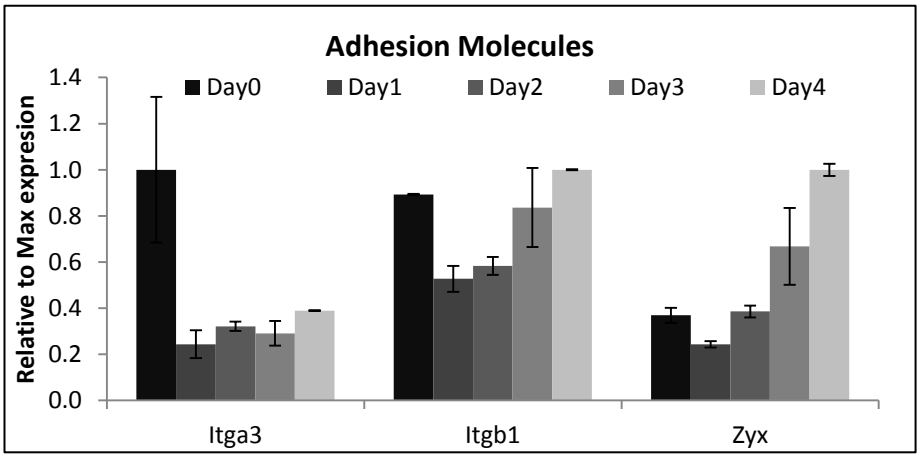
A) Expression of Tcf15-responsive genes in doxycycline-treated inducible TTE15 cells



B) Expression of Tcf15-responsive genes in FACS sorted Tcf15-Venus reporter cell line



C) Expression of Tcf15-responsive genes during neural differentiation



(Figure continued to next page)

D) Expression of cadherins during neural differentiation

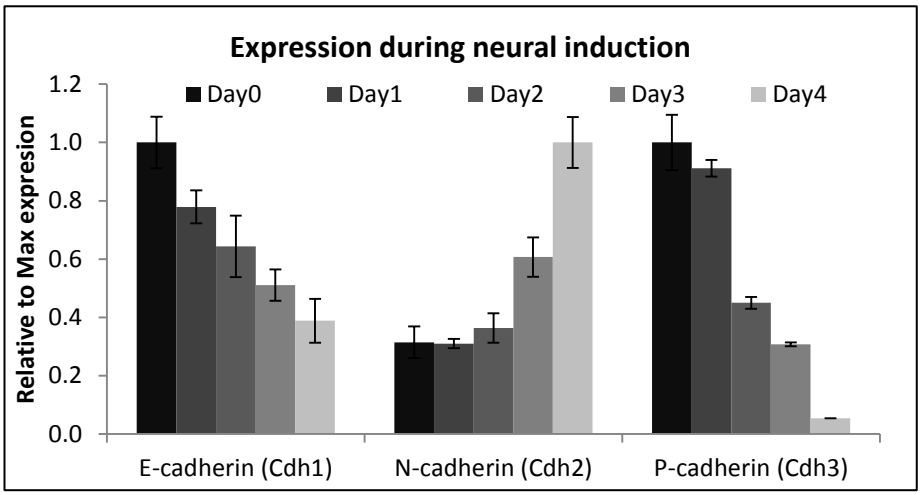


Figure 5.10: Gene expression of adhesion molecules up-regulated after induction of Tcf15-E47 expression

(A) Gene expression of adhesion molecules after doxycycline induction of Tcf15-E47 expression in TTE15 cells.

(B) Gene expression of FACS sorted Tcf15-Venus-high and Tcf15-Venus-low populations from Tcf15-Venus cells. *: $p<0.05$.

(C) Gene expression of Tcf15-responsive adhesion molecules during serum-free monolayer neural differentiation.

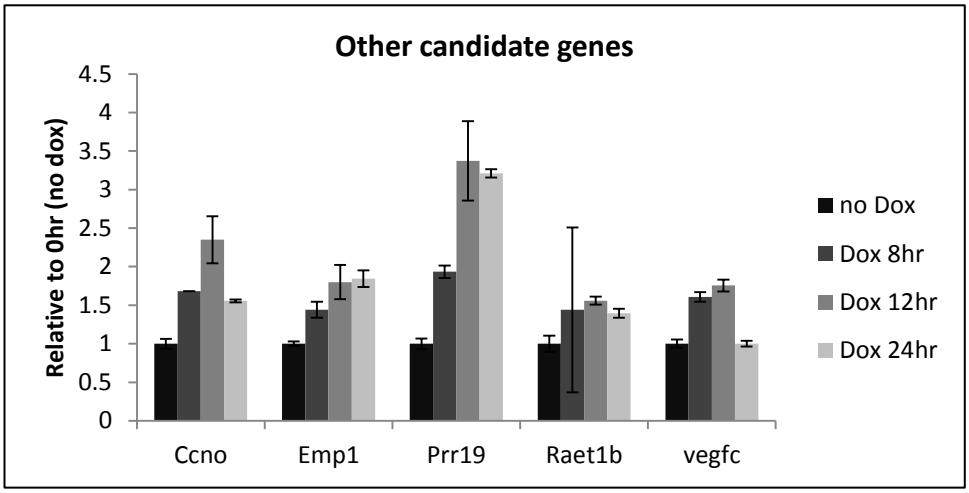
(D) Expression of *Cdh1 (E-cadherin), Cdh2 (N-cadherin)* and *Cdh3 (P-cadherin)* during neural differentiation. *Cdh3* was down-regulated gradually during neural differentiation.

5.2.2.1.1.4: Other genes up-regulated after Tcf15-E47 induction

Ccno (Cyclin O), Emp1 (Epithelial membrane protein 1), Prr19 (Proline rich 19), Raet1b (retinoic acid early transcript 1 beta) and *Vegfc (vascular endothelial growth factor C)* are up-regulated after inducing Tcf15-E47 expression (Table 5.1) that do not fall into any of the above categories.

qPCR analysis confirmed that *Ccno, Emp1, Prr19, Raet1b* and *Vegf*c were up-regulated after doxycycline induction (Figure 5.11A). Expression of *Raet1b* was however very low in ES cells cultured in LIF and serum, therefore it was difficult to obtain reproducible data on gene expression (see large error bars in figure 5.11A). None of these genes showed strong enrichment in the FACS sorted Tcf15-Venus-high subpopulation of ES cells (Figure 5.11B). For this reason, the expression analysis of these genes during neural differentiation was not preceded.

A) Expression of Tcf15-responsive genes in doxycycline-treated inducible TTE15 cells



B) Expression of Tcf15-responsive genes in FACS sorted Tcf15-Het reporter cell line

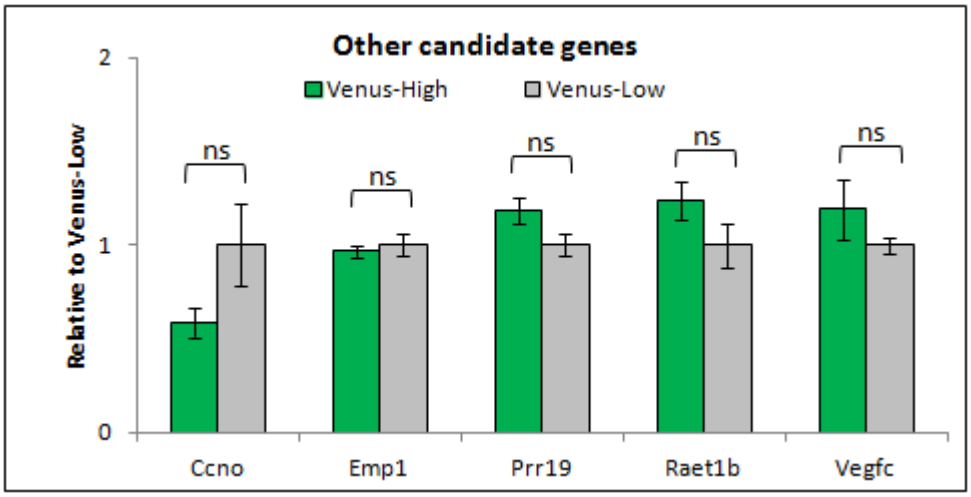


Figure 5.11: Expression of other genes up-regulated after induction of Tcf15-E47 expression

(A) Gene expression after doxycycline induction of Tcf15-E47 expression in TTE15 cells.

(B) Gene expression of FACS sorted Tcf15-Venus-high and Tcf15-Venus-low cells from Tcf15-Het cells.

5.2.2.1.2: Down-regulated genes after Tcf15-E47 induction

5.2.2.1.2.1: Naïve pluripotency marker Nanog was down-regulated by Tcf15-E47

Table 5.4 displays the list of the 18 genes that are most strongly down-regulated after inducing Tcf15-E47 expression and ranked by fold change at 12 hours of doxycycline treatment. Among these, the most strongly down-regulated gene at the two earliest time points is *Nanog*. This further validates the hypothesis that Tcf15 activation represents the first step toward differentiation (Figure 4.12B) and suggests a mechanism by which it drives cells away from a naïve pluripotent state.

qPCR analysis of mature *Nanog* mRNA confirms that *Nanog* expression was down-regulated after inducing Tcf15-E47 expression by doxycycline (Figure 5.12A). To further assess the transcriptional activity of the *Nanog* locus, qPCR analysis was performed by using five primer pairs located within a region of *Nanog* intron 1 to determine the level of pre-mRNA produced by the *Nanog* locus. *Nanog* pre-mRNA expression was also downregulated, confirming the level of transcription of the endogenous *Nanog* locus was reduced after Tcf15-E47 induction. In addition, consistent with data shown in figure 3.9C, *Nanog* expression was negatively correlated with Tcf15-Venus expression (Figure 5.12B).

Taken together, data presented in this chapter support a model in which FGF signalling pathway upregulates *Tcf15* expression and drives the transition to epiblast-primed state by activation of *Otx2* and suppression of *Nanog* expression (Figure 5.13).

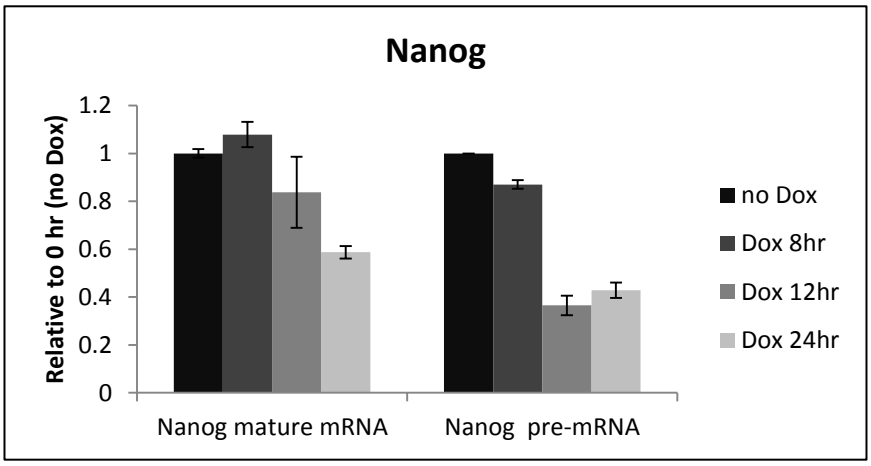
Table 5.4: Genes downregulated in response to Tcf15 activation

Illumina Probe ID	Gene Names	Description	log_2FC 8/0hr	log_2FC 12/0hr	log_2FC 24/0hr	qPCR verified Dox-inducible cells	qPCR verified FACS sorted Tcf15-reporter
ILMN_2706623	Nanog	Nanog homeobox	-0.4298*	-1.0768*	-0.9348*	Yes	—
ILMN_2458765	Ahnak	AHNAK nucleoprotein (desmoyokin)	-0.1327	-1.0251*	-0.4996*		
ILMN_1230176		RIKEN cDNA 1810032O08 gene	-0.1514	-0.9603*	-0.6039*		
ILMN_2740902	Mki67	Antigen identified by monoclonal antibody Ki 67	-0.1190	-0.9543*	-0.5564*		
ILMN_2502996	Tpr	Translocated promoter region	0.0516	-0.9380*	-0.2619*		
ILMN_2624622	Akap12	A kinase (PRKA) anchor protein (gravin) 12	-0.0333	-0.9112*	-0.3344		
ILMN_2627299	Akap9	A kinase (PRKA) anchor protein (yotiao) 9	-0.0562	-0.9077*	-0.5355*		
ILMN_3159435	Mid1	Midline 1	-0.0533	-0.9029*	-0.3571		
ILMN_2939012		RIKEN cDNA 1500012F01 gene	0.0115	-0.8966*	-0.4293*		
ILMN_2520249	Hspg2	Perlecan (heparan sulfate proteoglycan 2)	0.0531	-0.8645*	-0.5365*		
ILMN_2608933	Tcof1	Treacher Collins Franceschetti syndrome 1, homolog	-0.4026	-0.8531*	-0.6286		
ILMN_1229455	Rnf213	Ring finger protein 213	-0.0531	-0.8470*	-0.3064*		
ILMN_2672626	Myh9	Myosin, heavy polypeptide 9, non-muscle	-0.2157	-0.8309*	-0.1547		
ILMN_2599997	Ncor1	Nuclear receptor co-repressor 1	-0.0410	-0.8180*	-0.3944		
ILMN_2933692	Upf1	UPF1 regulator of nonsense transcripts homolog	-0.1810	-0.7833*	-0.5433*		
ILMN_2721385	Dst	Dystonin	-0.1228	-0.7507*	-0.4336*		
ILMN_1231490		2410006H16Rik	-0.1091	-0.7461*	-0.4270*		
ILMN_2678373	Myst4	MYST histone acetyltransferase monocytic leukemia 4	-0.0597	-0.7382*	-0.6337*		

***: p-value < 0.05**

(+): positive correlation, (-): negative correlation, (=): no difference, between Tcf15-Het-Venus and target gene expression

A) Expression of Tcf15-responsive genes in doxycycline-treated inducible TTE15 cells



B) Expression of Tcf15-responsive genes in FACS sorted Tcf15-Venus (Tcf15-Het) reporter cell line

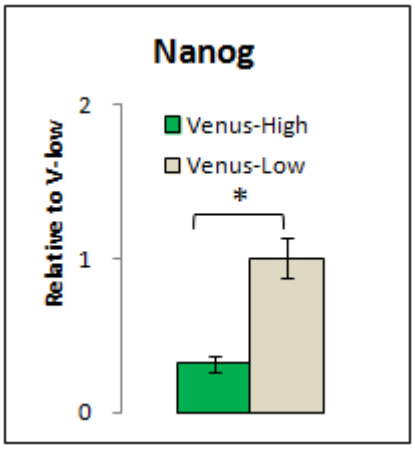


Figure 5.12: Nanog expression after induction of Tcf15-E47 expression

(A) Gene expression of *Nanog* and *Nanog* pre-mRNA after doxycycline induction of Tcf15-E47 expression in TTE15 cells. *Nanog* pre-RNA expression was analysed by averaging 5 pairs of primers within 1st exon of *Nanog* locus.

(B) *Nanog* expression of FACS sorted Tcf15-Venus-high and Tcf15-Venus-low cells. *: $p<0.05$.

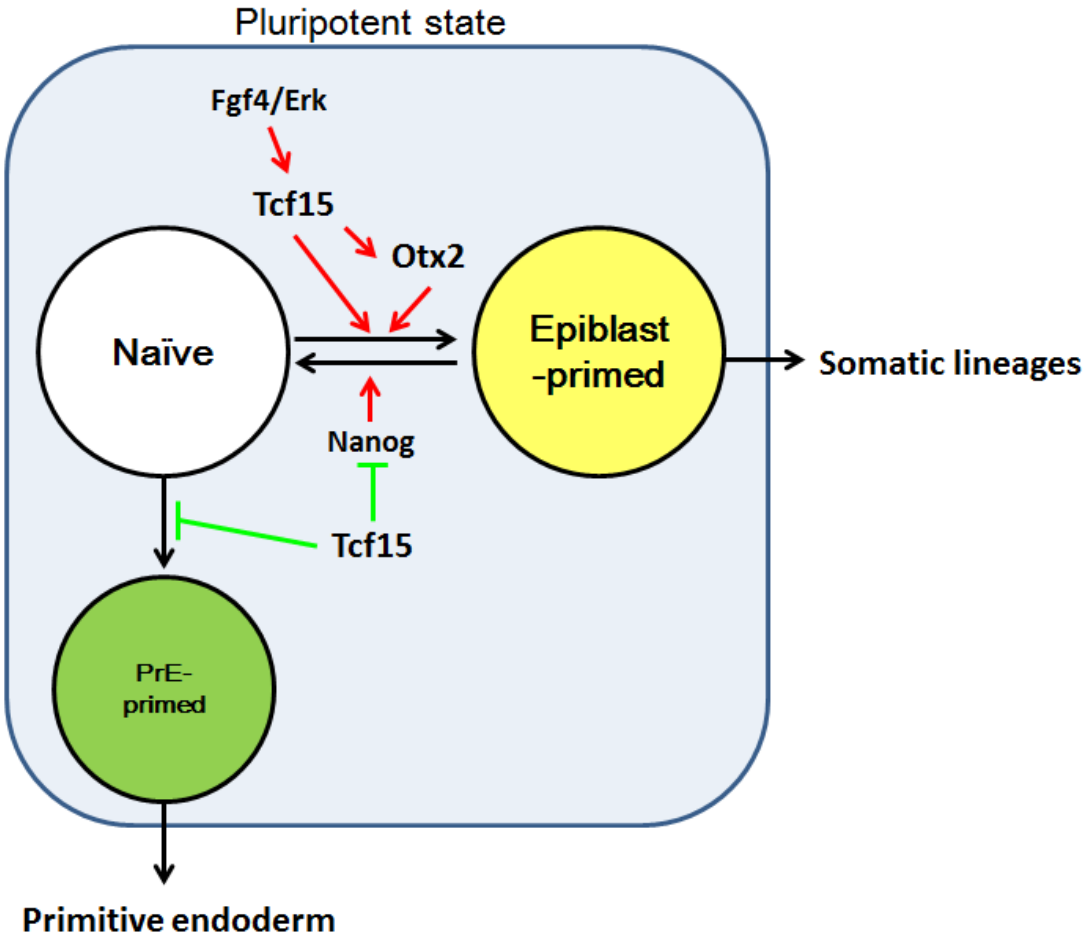


Figure 5.13: Proposed model of Tcf15 function in pluripotent ES cells

FGF4/ERK signaling drives expression of *Tcf15*. Tcf15 thus maturated naïve pluripotent stem cells toward epiblast-primed state, by activating epiblast-determinant gene *Otx2*, or by suppressing *Nanog*.

5.2.3: Adhesive properties of Tcf15-Venus-high and Tcf15-Venus-low subpopulations

As described in section **3.2.2.3.1.6**, the Tcf15-Venus-high subpopulation is primed to undergo neural differentiation more efficiently than the Tcf15-Venus low subpopulation (Figure 3.15). It also been observed that gene expression of adhesion molecules changed in response to Tcf15 activation, and expression of *Itga3*, *Itgb1* and *Zyx* were enriched in Tcf15-Venus high cells (Figure 5.10B). Therefore, it would be interesting to ask whether the Tcf15-high subpopulation showed any obvious difference in colony morphology that may reflect differences in adhesive properties.

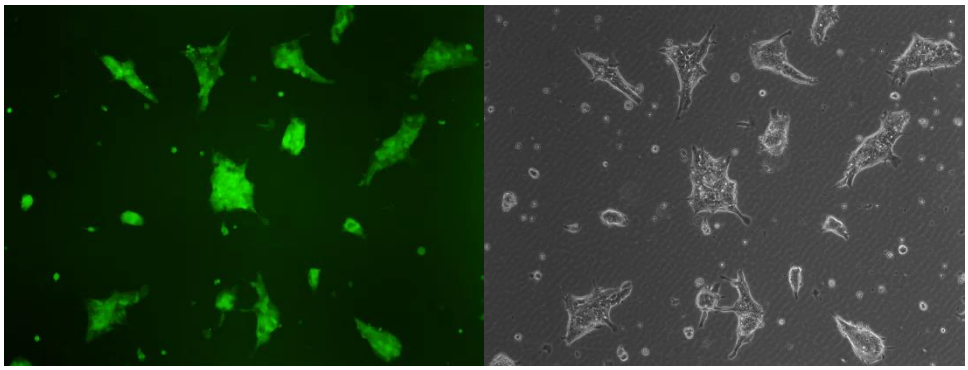
In order to test this, Venus-high and Venus-low populations were FACS sorted and plated for one day in serum-free media (N2B27) in order to allow Id1 to be downregulated and thus allow the endogenous Tcf15 to become activated. Clear morphological differences were seen between the two subpopulations (Figure 5.14). Tcf15-Venus-high cells attached to the plate more quickly than Tcf15-Venus-low cells and tend to form more spreading colonies containing more cells with epiblast-like epithelium morphology (Figure 5.14A). In contrast, Tcf15-Venus-low cells formed more compact colonies containing fewer cells with naïve "dome"-like morphology (Figure 5.14B).

qPCR analysis was performed to further characterise expression of genes encoding adhesion regulators in the two subpopulations. Epithelial-mesenchymal-transition (EMT) markers *Zeb2* (but not *Zeb1*), *Vimentin* and *N-cadherin* were enriched in FACS sorted Tcf15-Venus-high cells (Figure 5.15A),

as well as the polarity regulator *Crb2* (Figure 5.15B).

These data suggest that changes in Tcf15 expression are associated with changes in the expression of particular regulators of adhesion and morphology, and with a change in the morphological appearance of cell colonies. These preliminary observations could be an interesting area for future study, as discussed below.

A) Venus-high cells



B) Venus-low cells

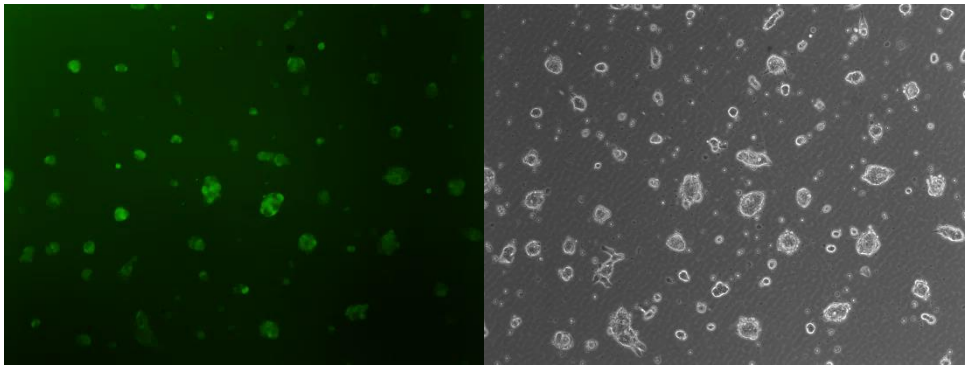


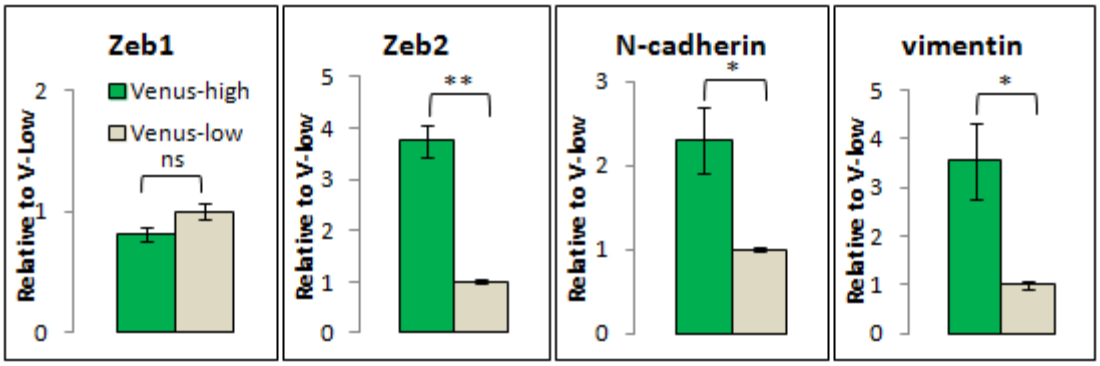
Figure 5.14: Morphological differences between Venus-high and Venus-low cells after 1 day of neural differentiation

Venus expression of FACS sorted Tcf15 reporter cells after 24 hours culture in serum-free N2B27 medium.

(A) Venus-high cells formed larger flatten colonies containing more cells with epithelium morphology.

(B) Venus-low cells formed more compact colonies containing fewer cells with naïve "dome"-like morphology.

A) EMT markers



B) Polarity regulator

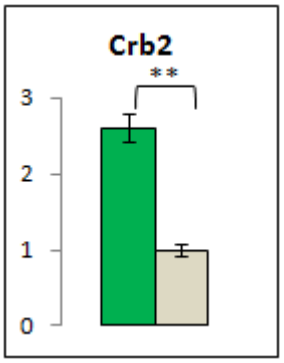


Figure 5.15: Differential expression of EMT-related genes and polarity regulators in FACS sorted Tcf15-Het reporter cells

Tcf15-Het reporter cells were cultured in LIF and serum before sorting. The top and bottom 30 % of cells were sorted from PECAM1-positive population and labelled as Venus-high or Venus-low, respectively.

(A) Differential gene expression of EMT markers in FACS sorted Tcf15-Het cells.

(B) Expression of polarity regulator *Crb2* in FACS sorted Tcf15-Het cells.

ns: non-significant; *: $p<0.05$; **: $p<0.01$

5.3: Discussion

5.3.1: ERK and GSK-3β signalling regulating Tcf15 expression

ES cells can be maintained in a naïve state if the cells are shielded from autocrine FGF/ERK signalling and subjected to GSK-3β inhibition by using two pharmacological inhibitors PD03 and Chiron (Ying et al., 2008). Expression of *Tcf15* was downregulated in 2i culture (Figure 3.4A). Therefore, Tcf15-Het reporter cells were used to investigate the effect of 2i components on the expression of *Tcf15*. To prevent over degeneration and differentiation of cells culturing in serum-free N2B27 medium supplemented only with PD03 or Chiron (Ying et al., 2008), the experimental strategy was designed to add inhibitors into LIF + FCS culture conditions. Addition of PD03 and Chiron into LIF + FCS cultures reduced *Tcf15* expression, as measured by Venus protein expression (Figure 5.1A and 5.3) and by qPCR analysis of both *Tcf15* and *Venus* expression (Figure 5.1B). This down-regulation of *Tcf15* and *Venus* suggested that *Tcf15* is expressed biallelically within ES cells, as both *Tcf15* alleles in the Tcf15-Het cells ($Tcf15^{VenusFRT/+}$) were down-regulated at similar extent.

PD03 and Chiron did not downregulate *Tcf15* expression completely, as some cells still express Venus. One possibility could be the MEK and GSK-3β pathways were not fully inhibited in the presence of serum in the culture, as the concentration used in these experiments were adapted from the serum-free 2i protocol (Ying et al., 2008). Other possibilities are that PD03 or Chiron alone is not sufficient to downregulate *Tcf15*, or, there might be other factors controlling basal transcription of *Tcf15* within LIF + FCS culture.

GSK-3β is inactivated by the canonical Wnt signalling pathway, resulting in the nuclear accumulation of β-catenin (Wu and Pan, 2010). Although inhibition of GSK-3β might have a Wnt-independent function in promoting proliferation and cell survival in 2i ES culture (Ying et al., 2008), it is possible that *Tcf15* expression is downstream of canonical Wnt pathway, as has been proposed in the somites (Linker et al., 2005). It would be interesting to assess if Tcf15-Venus expression could be modulated by addition of Wnt agonist or antagonists (Dkk or WIF-1) into self-renewing cultures.

Autocrine FGF4/ERK signalling is needed for ES cells to exit from self-renewal and initiate differentiation (Kunath et al., 2007). Both PD03 (MEK inhibitor) and PD17 (FGFR inhibitor) were capable of suppressing *Tcf15* expression (Figure 5.3B). However, addition of FGF4 protein directly into cells cultured in self-renewing conditions such as LIF + FCS or N2B27 + LIF + BMP failed to upregulate *Tcf15* expression (data not shown), which is perhaps due to autocrine production of saturating levels of FGF4 by undifferentiated ES cells (Ma et al., 1992, Rathjen et al., 1999).

5.3.2: Downstream responsive genes of Tcf15

Genes which were up-regulated by induction of Tcf15-E47 were manually categorized into 4 groups, as discussed below. Candidate genes which satisfied at least two of the following experiments were considered as possible Tcf15 direct or indirect targets: responds to Tcf15 activation, enriched in Tcf15-high subpopulation or displayed similar expression profile to Tcf15 during neural differentiation.

5.3.2.1: Transcription factors/regulators

Hes6, Id1 and Id3 are basic helix-loop-helix (bHLH) transcriptional factors/regulators. Hes6 alone does not bind to DNA but acts as an inhibitor of Hes1 to suppress Hes1 from repressing transcription and promotes neuronal cell differentiation (Bae et al., 2000). Id1 and Id3 are dominant negative antagonists of the bHLH family of transcription factors, which positively regulate differentiation in many cell lineages. It has been reported that BMP4-induced Id proteins inhibit ES cells from differentiating toward the neural lineage, and sustain self-renewal of ES cells in collaboration with LIF/STAT3 (Ying et al., 2003). A recent study suggested that Id proteins are also able to block the transition of ES cells to epiblast stem cells (EpiSCs) (Zhang et al., 2010).

IRF-1 is associated with regulation of interferon α and β transcription. IRF-1 activates a set of target genes associated with regulation of the cell cycle, apoptosis and the immune response (Chen et al., 2013).

Otx2 is a homeobox transcription factor, which is required at multiple steps in brain development and neuronal differentiation, especially in anterior neuroectoderm specification (Acampora et al., 1995). It has been reported that Otx2 acts as a novel intrinsic determinant controlling ES cells transition into EpiSCs (Acampora et al., 2012).

WT1 is a zinc-finger-containing transcription factor. It is essential for the development of certain mesoderm-derived tissues and also for the proliferation of certain neuronal progenitors (Wagner et al., 2002). The biological function of WT1 appears paradoxical. WT1 induced

mesenchymal–epithelial transition (MET) in the developing kidneys, but in the developing heart, WT1 drove the epithelial-mesenchymal transition (EMT) in the epicardium (Hohenstein and Hastie, 2006). These data suggested that WT1 is involving in the cellular epithelial-mesenchymal balance.

In summary: qPCR analyses validated that Tcf15-E47 induced transcription factors *Hes6*, *Id1*, *Id3*, *Irf1*, *Otx2* and *WT1*. Of these genes, only *Hes6*, *Id1* and *WT1* were enriched in Tcf15-Venus-high cells. *Hes6*, *Otx2* and *WT1* showed a similar expression profile to *Tcf15* during neural differentiation. Although further experiments, including chromatin immuno-precipitation (ChIP) need to be performed to confirm if these genes were direct Tcf15 targets, these data might suggest the hypotheses that Tcf15 specifically favors epiblast rather than primitive endoderm differentiation by up-regulating *Otx2*, a known epiblast determinant (Acampora et al., 2012), and that Tcf15 promotes robust neural differentiation by up-regulating *Hes6*, a known pro-neural factor (Bae et al., 2000). It is also speculated that the ability of Tcf15 to modulate expression of the MET regulator *WT1*, which may relate to the changes in adhesion that accompany the formation of the post-implantation epiblast (Rossant and Tam, 2009) (see also chapter 6).

5.3.2.2: Signalling modulators

Dual specificity phosphatases 4 (DUSP4) and DUSP6 were two signaling molecules up-regulated after inducing Tcf15-E47 expression and both were downstream transcriptional targets of the FGF pathway (Lanner et al., 2010).

In mouse ES cells, DUSP4 and DUSP6 are down-regulated when FGF signalling is inhibited by the FGFR inhibitor, generating a more naïve and homogenous ES cells state (Lanner et al., 2010). This indicated that active FGF signalling not only promotes a primed state but also induces the expression of inhibitors of FGF signalling to generate a negative-feedback loop that would facilitate reversion back to the original naïve state (Figure 5.16) (Lanner and Rossant, 2010).

In addition, it is found that several Tcf15 target candidates were down-regulated in a previous microarray analysis, where cells were cultured in another FGFR inhibitor SU5402 which generates a naïve and homogenous ES cell state (Lanner et al., 2010) (Table 5.5). All of these genes, including *Dpp4, Hes6, WT1, DUSP4, DUSP6* and *Nanog*, showed negatively correlated expression between induction of Tcf15-E47 and addition of FGFR inhibitor, suggesting these genes might play certain roles in FGF-regulated naïve to epiblast-primed transition.

Taken together, although DUSPs function at the protein level to negatively regulate the ES cell to EpiSC transition, these data might indicate that Tcf15, as a FGF-inducing gene, might also trigger negative feedback mechanisms in FGF signalling pathway. Alternatively, these changes in DUSP4 and DUSP6 may occur as an indirect consequence of Tcf15-activated cells converting to a more epiblast-like state, which may perhaps be more responsive to autocrine FGF and therefore have higher expression of these FGF target genes. It will therefore be interesting in future to test whether DUSP4 and DUSP6 are direct or indirect targets of Tcf15, and whether other readouts of FGF activity accompany Tcf15 activation.

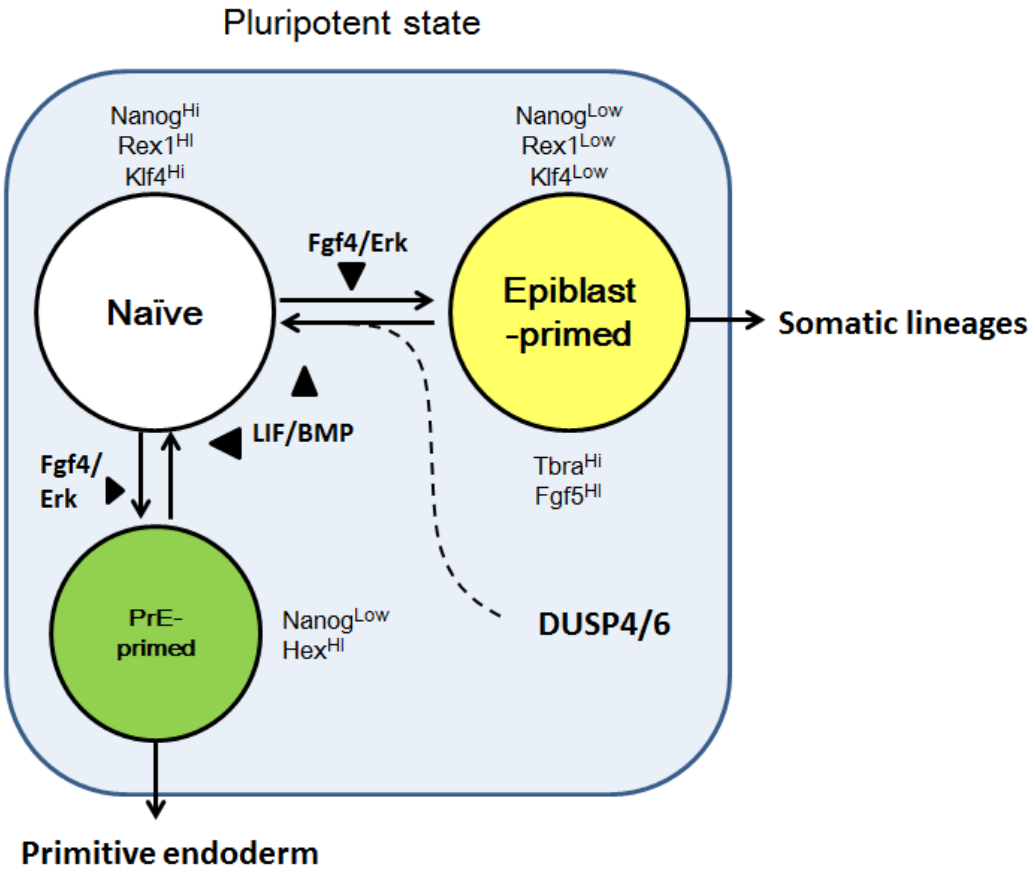


Figure 5.16: Diagram of possible function of DUSPs in negative regulating pluripotent cells transit to epiblast-primed state (Adapted from Lanner et al., 2010)

As described in chapter 3, pluripotent ES cells exhibit great heterogeneity, whereby some cells appear to be transiently primed for differentiation to either extraembryonic endoderm or somatic lineages. FGF4 signalling has been shown to maintain both primed states. DUSP4 and DUSP6 may promote the reversion of primed ES cells to the naïve state, indicated by dashed lines.

Table 5.5: List of Tcf15 and Tcf15-responsive genes which expression was under FGFR signalling pathway

Gene Of Interest	TTE15 + Dox	ES + FGFRi
Tcf15	up	down
Dpp4	up	down
Hes6	up	down
WT1	up	down
DUSP4	up	down
DUSP6	up	down
Nanog	down	up

Data of ES + FGFR inhibitor was retrieved from Lanner et al., 2010.

5.3.2.3: Adhesion molecules

Several adhesion molecules were up-regulated after inducing Tcf15-E47 expression.

Members of the integrin family are expressed in spatially discrete patterns that reflect their function in early mouse embryogenesis (Hynes, 2002). *Integrin α3* mRNA was expressed from late-blastocyst/ peri-implantation stage and mice lacking *Itga3* expression could survive until perinatal stage (Sutherland et al., 1993). Integrin β1 is essential for inner cell mass development and deletion of *Itgb1* result in peri-implantation lethality (Stephens et al., 1995, Fassler and Meyer, 1995).

Zyxin, a focal adhesion-associated LIM protein, is essential for actin reorganization for cell migration in TGF-β1-Twist1–induced EMT in normal murine mammary gland cells (Mori et al., 2009). Although *zyxin*-null mice are viable and fertile (Hoffman et al., 2003), there are no published data regarding zyxin function in ES cells differentiation and early mouse development.

Cdh3 is a known surface marker for trophectoderm (TE) (Niwa et al., 2005), but its function in the pluripotent epiblast has not been reported.

Although Tcf15-E47 up-regulated *Cdh3*, *Itga3*, *Itgb1* and *Zyx* expression, only *Itga3*, *Itgb1* and *Zyx* were enriched in Tcf15-Venus-high cells. In addition, all the genes showed dynamic changes during neural differentiation. Therefore, it is at present not clear whether these genes are Tcf15 targets or not, and further experiments are required to test this.

Whether or not these adhesion molecules are direct targets of Tcf15, they

are interesting as candidate mediators of the adhesive changes that accompany epiblast maturation (Rugg-Gunn et al., 2012). Since adhesion molecules function at the protein level and are often subject to considerable post-transcriptional and post-translational regulation, it would be more informative to analyze, in future studies, whether protein expression of these molecules correlated with Tcf15 expression, by flow cytometry analysis on surface expression of Cdh3, Itga3 or Itgb1, by immunostaining to identify cellular localization of Zyxin, or Western blot analysis on FACS sorted populations.

5.3.2.4: Other genes

Ccno (cyclin O) is a cyclin protein family member which can bind and activate CDK2 to respond to DNA damage-induced intrinsic apoptosis stimuli in in mouse lymphoid cells (Roig et al., 2009). Ccno was reported as a maternal factor gene that mainly transcribed in oocytes and zygotes and played an important role in oocyte meiotic resumption in mouse oocytes (Ma et al., 2013).

EMP-1 (also named as tumour-associated membrane protein, TMP) is a junctional protein in blood-brain barrier and in liver (Bangsow et al., 2008, Lee et al., 2005). EMP-1 is highly expressed in undifferentiated ES cells, but its expression is markedly down-regulated when ES cells are differentiated into embryoid bodies (Ben-Porath and Benvenisty, 1996).

VEGF-C has been characterized as an essential lymphangiogenic growth factor that promotes cancer metastasis by activating phosphorylation of its

receptor VEGF-R3 and triggers downstream activation of PI3K-Akt and PKC-MAPK pathway; thus protecting lymphatic endothelial cells from apoptosis and stimulating proliferation and migration *in vitro* (Makinen et al., 2001). It has been reported that VEGF-C was required for the proliferation of neuroepithelial cells expressing VEGF-R3 in both *Xenopus* and mouse developing brain (Le Bras et al., 2006).

There is no literature reporting function of Prr19 and Raet1b to my knowledge.

The expression of these genes requires further characterisation, as it is currently unclear whether they are expressed in ES cells and how the expression changes during differentiation.

5.3.2.5: Tcf15 down-regulates the naïve pluripotency marker *Nanog*

In chapter 3, a negative relationship between Tcf15-Venus and Nanog in ES cells was identified at both mRNA and protein level. In chapter 4, it was demonstrated that modulating Tcf15 expression level within ES cells changed expression of *Nanog* but not *Oct4*, thus influencing the naïve to primed-epiblast states within cultures of pluripotent cells (Figure 4.10 and 4.18). In this chapter, down-regulation of *Nanog* mRNA expression upon induction of Tcf15-E47 was observed (Figure 5.12). Taken together, this suggested that Tcf15 might regulate *Nanog* expression, either directly or indirectly. Future experiments will be conducted to test how Tcf15 regulates *Nanog* expression. The first step will be to confirm that Tcf15 down-regulates Nanog expression at the protein level. A time course co-immunostaining of Nanog and Flag-

Tcf15-E47 after doxycycline induction would be a good approach for analysing protein expression at the single cell level. Alternatively, to test if Tcf15 directly regulates *Nanog* expression at transcription level, ChIP analysis of Flag-Tcf15-E47 on *Nanog* regulatory elements could also be performed. The relationship between Tcf15, Id1 and Nanog will be discussed in chapter 6.

5.3.3: Adhesive properties of Tcf15-Venus-high and Tcf15-Venus-low subpopulations

Changes in Tcf15 expression are seen to be associated with changes in the morphological appearance of cell colonies (Figure 5.14), which might due to differential expression of adhesion molecules, EMT-related genes or the polarity regulator, *Crb2*, pre-existing in the Tcf15-Venus-high population. Crumbs homolog 2 (Crb2) is one of the three mammalian homologues of the *Drosophila* Crumbs (Crb) protein that involves in the control of cell–cell adhesion and epithelial cell polarity. In *Crb2* deficiency mice, the primary defect appears to be disturbed epiblast cell polarity, which affects the EMT at the primitive streak (Xiao et al., 2011).

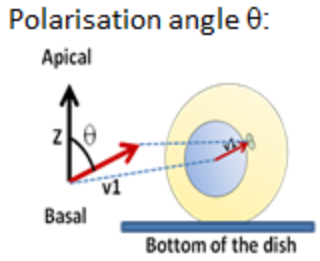
To address whether Tcf15-Venus-high cells exhibit different adhesive properties than Tcf15-Venus-low cells, an image quantification tool is now being developed in the lab. In brief, this macro plug-in in ImageJ software could detect cell spreading by identifying cells in phase contrast images and track them over time to follow the increase in surface area. To address if Tcf15-Venus-high cells exhibit more epithelialized/polarized character, an automated image analysis tool developed by Dr. Guilluame Blin within our

lab could be used, which could measure the orientation of each individual cell within a group of cells. In short, this software calculates the angel θ of epithelialisation/polarisation (Figure 5.17A), which is the angle between nucleus to centrosome vector (v1) and the unit vector in the Z dimension (in 2D culture: from bottom to top). The epithelialization/ polarisation angel θ of each cell within a colony will be reported on a rose diagram (Figure 5.17B) to determine the extent of epithelialization of a cell population.

In summary, it was been observed that Tcf15-Venus-high cells displayed more epiblast-like epithelium morphology. Moreover, Tcf15 induced genes which were associated with cell adhesion/cell polarity, suggesting possible function of Tcf15 to mature ES cells toward primed epiblast-like state might be regulating epithelialization during this transition.

The adhesive changes in Tcf15 cells and how this relates to adhesive changes in the early embryo will be discussed in chapter 6.

A)



B)

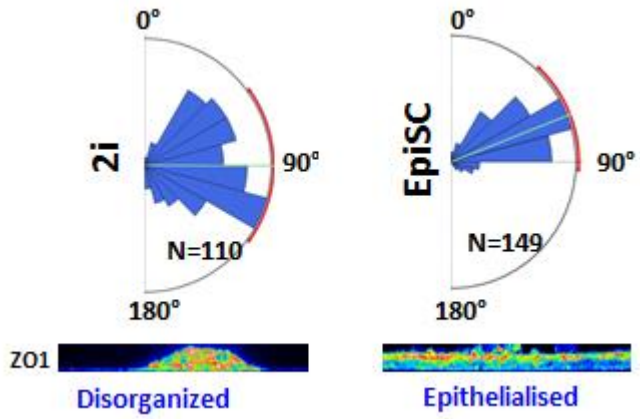


Figure 5.17: Measuring the orientation of cells using homemade image analysing software (developed by Dr. Guillaume Blin)

(A) Single cell orientation can be assigned using the nucleus to centrosome vector v1 (red arrow). The polarisation angle θ is defined by the angle between v1 and the top (apical) of the culture dish (vector Z). The more epithelialized a cells is, the smaller angle θ it has.

(B) Rose diagrams to address orientation of a cell population (top), showing here as disorganized naïve cells culturing in 2i and epithelialized EpiSCs. Each blue pie piece represents a range (here for example, 15° interval of angle θ) and the size is the percentage of occurrences. (Bottom) Confocal images of tight junction protein ZO1 staining to confirm the EpiSCs are more epithelialized than 2i cells.

CHAPTER 6

General discussion

In this study, the function of transcription factor Tcf15 in pluripotent cells was characterized by analysing its expression pattern followed by perturbing its expression in mES cells and identifying its upstream regulators and downstream targets. My data suggested that Tcf15 is acting as pro-differentiation factor priming pluripotent cells for differentiation.

6.1: Function of Tcf15 from naïve to primed pluripotency

6.1.1: Hypothesis: Tcf15 regulates cell adhesion and epithelialisation during naïve to primed transition

Embryonic stem cells are heterogeneous mixture of cells in a metastable state, shifting between naïve ICM-like (ground state) and epiblast-like (primed state) states while remaining pluripotent (Chambers et al., 2007; Hayashi et al., 2008; Toyooka et al., 2008). ES cells navigate a route toward lineage commitment through a complex network of positive and negative cues. The negative cues that restrain cells in a naïve pluripotent state include transcription factors Nanog and Klf4 and signalling pathways activated by BMP and LIF. In contrast, less is known about the positive gene regulatory mechanisms that drive cells toward commitment to differentiation.

The transition from naïve to mature differentiation-primed epiblast spans the peri-implantation period and is therefore technically challenging to study *in vivo*. The *in vitro* conversion of ES cells to EpiSCs in defined culture

recapitulates the progression from naïve pre-implantation epiblast to post-implantation epithelialized egg cylinder (Rossant and Tam, 2009) and provides a useful tool for investigating this process. This system has identified Id1 as a key negative regulator during ES cells to EpiSCs transition (Zhang et al., 2010), which implicating unknown bHLH transcription factors in priming pluripotent cells for differentiation.

Our previous yeast-two-hybrid screening identified Tcf15 as a possible pro-differentiation transcription factor which activity is likely to be negatively regulated by Id1 in ES cells. Tcf15 was heterogeneously expressed within the pluripotent culture and Tcf15-high cells were predominantly associated with the Nanog-low/Klf4-low population but not enriched in Hex and PDGFRα expression, suggesting the Tcf15-high cells were not in a naïve pluripotent state nor primed for extraembryonic endoderm. I proposed that Tcf15 is marking cells priming for somatic differentiation. Two key pieces of my data support this hypothesis: First, a FACS sorted Tcf15-Venus-high subpopulation was found to undergo differentiation more efficiently than the Tcf15-Venus-low subpopulation. Second, using doxycycline-inducible, Id-resistant dominant active form of Tcf15-E47 fusion protein, it was demonstrated that Tcf15 promotes somatic differentiation, possibly by accelerating the ES cell to EpiSCs transition.

How to link these findings to development of the epiblast *in vivo*? Tcf15 is well known as a regulator of somite development (Burgess et al., 1996), and somitogenesis shares some intriguing parallels with epiblast formation. In somites, Tcf15 is not required for cell fate specification; rather, Tcf15 is necessary and sufficient for remodelling of mesenchyme into an N-cadherin$^+$

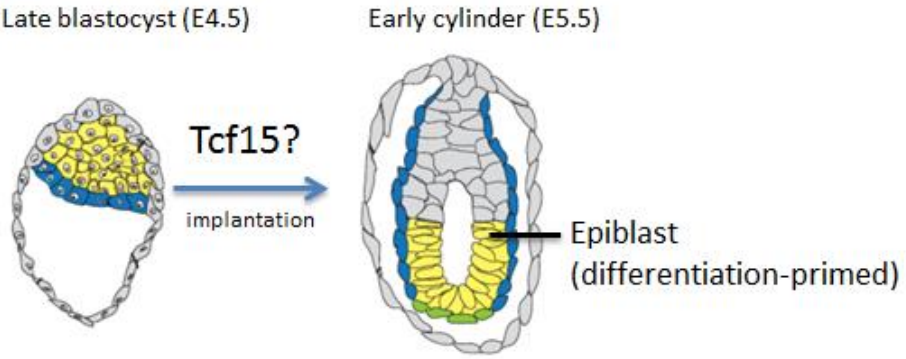
polarized epithelial structure (Burgess et al., 1996; Linker et al., 2005; Takahashi et al., 2005). Epithelialization in the somites is a consequence of N-cadherin-mediated cell-cell adhesion (Duband et al., 1987). Transition of the ICM of the blastocyst undergoes a similar remodelling process as it matures into post-implantation epiblast, including the upregulation of N-cadherin (Bao et al., 2009; Hatta et al., 1987) and formation of a polarized N-cadherin/ E-cadherin double-positive epithelium. Therefore, Tcf15 may play a similar role in both scenarios to help trigger the morphological remodelling events that characterize these changes in cell identity.

N-cadherin was recently identified as a cell surface maker expressed in the EpiSCs at much higher levels than in ES cells (Rugg-Gunn et al., 2012). Tcf15-Venus-high cells are enriched in *N-cadherin*, as well as EMT-associated genes (*Zeb1, Zeb2* and *Vimentin*) and polarity marker *Crb2*. Microarray analysis also identified several adhesion molecules including *Itga3, Itgb1, Cdh3* and *Zyx* being up-regulated after induction of Tcf15-E47 expression. Tcf15 also up-regulated transcription factor *WT1* expression, which is known regulator controlling the epithelial-mesenchymal balance during kidney and heart development. The function of *WT1* during early ES cells differentiation has not been characterized. Although only mRNA expression analysis was performed, these data had provided clues for supporting follow-up investigation at the protein level. Based on the morphological differences between Tcf15-Venus-high and Tcf15-Venus-low cells upon day 1 of neural differentiation (Figure 5.14), I propose that Tcf15 is likely to regulate cell adhesion and epithelialisation during early time point of differentiation, possibly the ES cells to EpiSCs transition stage *in vitro*, or peri-implantation

stage *in vivo* (Figure 6.1A).

In order to further investigate the hypothetical role of Tcf15 in controlling epithelialization of the epiblast, it would be helpful to have access to an *in vitro* system that better models the morphological changes of the epiblast *in vivo*. It was recently reported that the first major morphogenetic step in the transition of the epiblast cells from pre- to post-implantation stages is its progressive re-organization from a relatively simple ball of cells to a more complex rosette-like structure that is built of polarized cells (Figure 6.1B) (Bedzhov and Zernicka-Goetz, 2014). It was further shown that epithelialization of the epiblast during peri-implantation development can be mimicked by culturing ES cells suspended in a matrix of extra-cellular matrix (ECM) proteins. The ECM/basal membrane-stimulated integrin signalling therefore provide polarization cues for the coordinated orientation of the apical-basal axis of these cells and cells self-organized into spheres recapitulating the epiblast (Bedzhov and Zernicka-Goetz, 2014). This technique could provide a great *in vitro* tool to study whether Tcf15 regulate cell polarisation and epithelialization. I could take advantage of currently available doxycycline-inducible Tcf15-E47 cell line and Tcf15-Het-Venus reporter cell lines to establish self-organizing ES cells spheres. By manipulating Tcf15 expression level, it would be possible to analyse if the formation of polarized rosettes correlates with endogenous or exogenous Tcf15 expression. Therefore, the *in vitro* data would help us to gain more insight into Tcf15 function in peri-implantation embryo development *in vivo*.

A)



B)

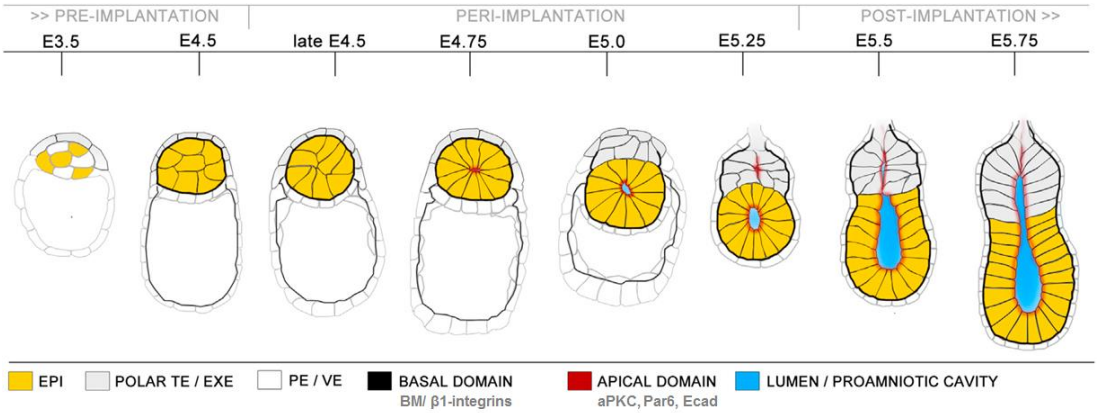


Figure 6.1: Tcf15 might regulate cell polarization and epithelialization *in vivo*

(A) Hypothetical function of Tcf15 in regulating differentiation-primed cellular epithelialisation. The *in vivo* function of Tcf15 might be regulating the epithelialisation of epiblast during peri-implantation stage. Adapted from (Takaoka and Hamada, 2012).

(B) A model for the sequence of the morphogenic events that drive the peri-implantation development (Bedzhov and Zernicka-Goetz, 2014). The pre-implantation blastocyst is comprised of unpolarized epiblast cells (EPI). The cells of the extraembryonic lineages secrete ECM proteins that assemble a basal membrane, which wraps around the EPI. The ECM proteins provide polarization cues that orient the establishment of a basal-apical axis of the EPI cells through β1-integrin receptors on the basal surface of EPI. The EPI cells change their shape as a result of actomyosin constriction, coupled to apical localization of adherens junctions and forming a rosette-like structure (E4.75–E5.0).

6.1.2: Hypothesis: Tcf15 upregulates Otx2 expression during naïve to primed transition

ES cells and EpiSCs are molecularly and epigenetically distinct and therefore represent discrete pluripotent states recently termed naïve and primed pluripotent states, respectively. *In vivo*, development of post-implantation epiblast cells from pre-implantation ICM cells involves significant transcriptional and epigenetic changes, including DNA methylation and X chromosome inactivation (Surani et al., 2007). Recently, it has been reported that the transition between naïve and primed pluripotent states is associated with global re-organization of the chromatin landscape and distinct enhancer element usage (Factor et al., 2014). A key mediator for remodelling the gene regulatory networks as cells exit from ground state pluripotency is Otx2, which drives enhancer activation through affecting chromatin marks and the activity of associated genes in cooperating with Oct4 (Buecker et al., 2014; Yang et al., 2014). What remained unclear is whether Otx2 is acting as a pioneer factor capable of initiating chromatin opening by itself or whether additional factors which are activated by FGF signalling might be required to collaborate with Otx2 to access DNA and lead to enhancer activation (Buecker et al., 2014). In this study, it was found that Tcf15-E47 is able to upregulated *Otx2* expression. Moreover, expression of *Tcf15* is regulated by FGF signalling. Future works including genetic and epigenetic approaches could be performed to analyse if the epiblast-priming function of Tcf15 is Otx2-dependent. Moreover, it would be interesting to investigate whether there is any difference of chromatin marks between

Tcf15-high and Tcf15-low cells, and, if there is difference, does this correlate to known differences between naïve and primed pluripotent cells.

6.2: The interplay between Tcf15, Nanog and Id1

My time-course microarray analysis has identified several genes which expression in response to Tcf15 activity. Among these genes, the naïve pluripotency marker *Nanog* is the transcript displaying the strongest level of downregulation following the induction of Tcf15-E47 expression suggested that Tcf15 might regulate *Nanog* transcription, either directly or indirectly. The negatively correlated expression of Tcf15 and Nanog in pluripotent culture could also imply that Nanog repress *Tcf15* expression. Festuccia et al. performed a microarray on *Nanog*-null ES cells following tamoxifen-induced nuclear translocation of a transgenic Nanog-ERT2 fusion protein. Analysis of this dataset reveals that both *Id1* and *Tcf15* were the transcripts displaying the strongest level of downregulation following the nuclear translocation of Nanog (Festuccia et al., 2012). The observation that Nanog promotes a rapid and strong downregulation of *Id1* and *Tcf15* transcription upon nuclear localisation suggests Nanog can inhibit *Id1* and *Tcf15* expression through the direct binding and repression of the *Id1* and *Tcf15* locus.

Analysis of published ChIP-Seq data (Chen et al., 2008; Marson et al., 2008; Whyte et al., 2013) reveals the potential presence of Nanog binding sites at the *Id1* locus (Figure 6.2). However, ChIP-Seq data analysis of Nanog binding targets in same studies (Chen et al., 2008; Marson et al., 2008; Whyte et al., 2013) does not indicate any binding of Nanog around the *Tcf15* locus.

This suggests Nanog could directly repress *Id1* expression but might not directly repress *Tcf15* expression. Further analysis of published ChIP-Seq data indicates the presence of Oct4 and Esrrb binding sites at the *Tcf15* locus (Chen et al., 2008; Marks et al., 2012) (Figure 6.3). Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells (Festuccia et al., 2012). The putative Oct4 binding site overlapped with RNA polII binding site at transcription start site (TSS) of *Tcf15* locus, whereas potential Esrrb binding sites located slightly upstream of TSS in addition to a small peak at TSS. Therefore, I propose the speculative hypothesis that Nanog itself does not bind to *Tcf15* locus, and, the Nanog-induced downregulation of *Tcf15* is Esrrb-mediated.

The microarray analysis described in chapter 5 showed that Tcf15-E47 down-regulated *Nanog* expression in ES cells cultured in LIF and serum. However, Tcf15-E47 was not able to repress Nanog expression in cells cultured in 2i/LIF (Figure 4.11). The lack of *Nanog* repression by Tcf15-E47 in 2i/LIF culture might due to: a) a strong core pluripotency transcription network presence in the 2i/LIF culture which masks the effect generated by Tcf15-E47, or, b) the possibility that Tcf15-E47 down-regulates *Nanog* expression indirectly or in cooperation with another factor, and this mediator for repressing *Nanog* expression is not expressed or activated sufficiently in 2i/LIF cultures.

Id1 is one of the transcriptional regulators factors up-regulated after inducing Tcf15-E47 expression. Our yeast-two-hybrid screening results of Id1 and E proteins binding partners suggests that in ES cells, Id dimerizes with and sequesters the ubiquitous bHLH transcription factor E12/E47, blocking

its interaction with Tcf15. Therefore, expression of Id within pluripotent cells is likely to render Tcf15 inactive. The Tcf15-induced *Id1* up-regulation is likely to occur through Tcf15-induced Nanog downregulation, as described above. However, the published phenotype of *Id1*-null ES cells included lower expression levels of *Nanog* and *Rex1* and higher expression levels of *T-Brachyury* (Romero-Lanman et al., 2012). Analysis of the microarray data comparing *Id1*$^{-/-}$ ES cells to wild-type ES cells confirms that other post-implantation epiblast markers, such as *Fgf5* and *Otx2*, are upregulated in the knockout cells (Romero-Lanman et al., 2012). Taken together, the gain of EpiSC-like gene expression in *Id1*$^{-/-}$ ES cells might due to the loss of Id1 expression-induced Tcf15-inactive state within pluripotent culture.

The interplay between Tcf15, Nanog and Id1 is summarized in figure 6.4. What remains unclear is that whether Tcf15 could directly control Nanog expression at transcription level. In order to gain further insight into how Tcf15 primes pluripotent cells to differentiation and how Tcf15 drives cells exit pluripotent state, future experiments will be to characterize how the cross-repression between Tcf15 and Nanog is achieved.

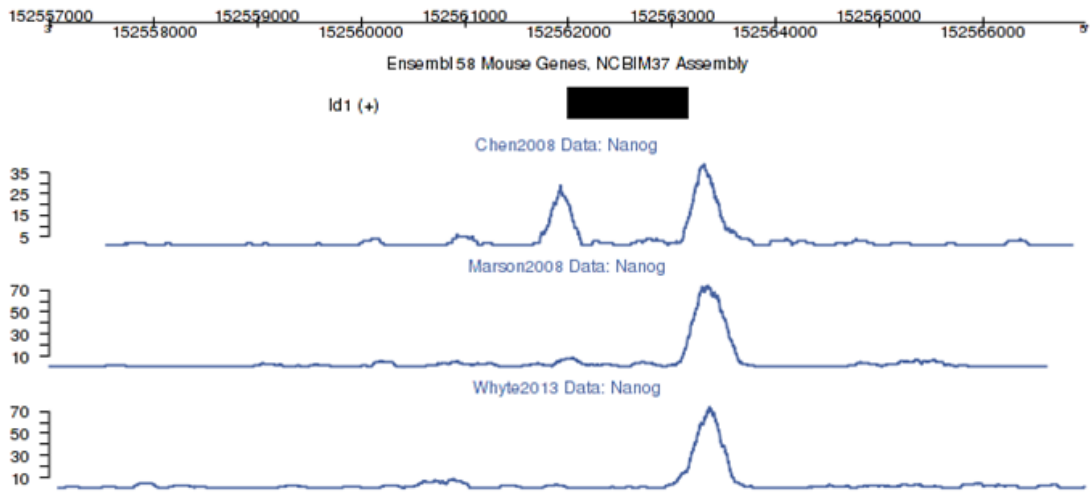


Figure 6.2: ChIP-Seq identified Nanog binding sites at the *Id1* locus

Analysis of the Nanog ChIP-Seq data (Chen et al., 2008; Marson et al., 2008; Whyte et al., 2013) with the GeneProf online resource (Halbritter et al., 2012) reveals peaks of Nanog binding upstream and downstream of the *Id1* exonic DNA (black box).

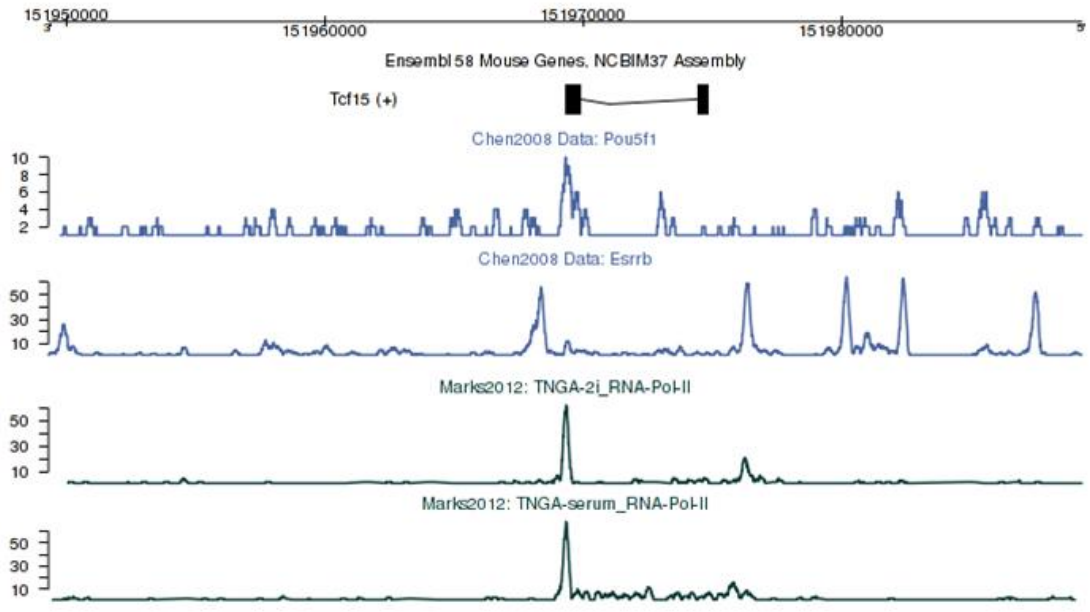


Figure 6.3: ChIP-Seq analysis at the *Tcf15* locus

Analysis of the ChIP-Seq data (Chen et al., 2008; Marks et al., 2012) at *Tcf15* locus using the GeneProf online resource (Halbritter et al., 2012). RNA polII binding is displayed to reveal the transcription start site (TSS) of *Tcf15* locus. ES cells cultured in serum displayed increased RNA polII occupancy over coding body, which fits the observation that *Tcf15* expression is higher in serum culture than in 2i. Putative Oct4 binding site overlapped with TSS, whereas potential Esrrb binding sites located slightly upstream of TSS in addition to a small peak at TSS.

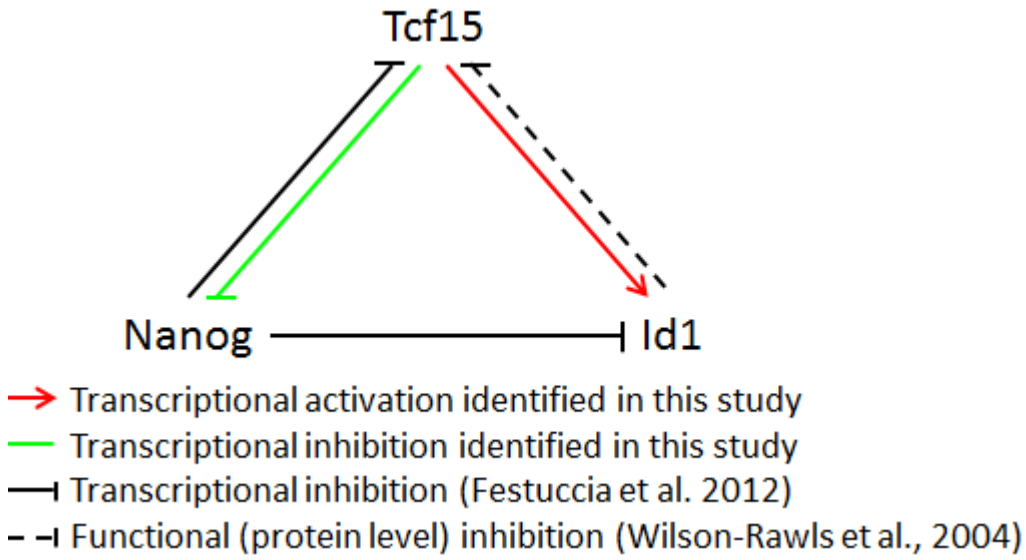


Figure 6.4: Cross repression and activation of Tcf15, Nanog and Id1 expression

Tcf15 activates *Id1* expression and represses *Nanog* expression according to the time-course doxycycline induction of Tcf15-E47 microarray analysis performed in chapter 5. Expression of *Id1* and *Tcf15* were downregulated following the nuclear translocation of Nanog (Festuccia et al., 2012). Tcf15 is active as a heterodimer with E proteins and is therefore functionally inactive in the presence of Id proteins, which bind and sequester E proteins (Wilson-Rawls et al., 2004).

6.3: Possible functional redundancy between Tcf15 and Scleraxis during embryo development

As described in section **4.4.2.5**, *Tcf15*-null embryos survive to birth, but die shortly afterward as a consequence of muscular-skeletal defects due to the failure in somite morphogenesis earlier in development (Burgess et al., 1996). The loss of *Tcf15* in highly regulative early embryos, before somitogenesis, might therefore only be delayed epiblast formation during peri-implantation stage and lowering the robustness of early developmental process. Furthermore, any putative *Tcf15* mutant phenotype may be masked by compensatory premature upregulation of the closely related transcription factor Scleraxis (Scx), which shares an almost identical DNA binding domain with Tcf15 (Burgess et al., 1995) (Figure 6.5). Scx has an established role in later stages of epiblast maturation: the epiblast of *Scx* mutant embryo failed to develop further after E6.5 and does not properly epithelialize. At later stages, *Scx* mutant failed to gastrulate properly and mesoderm was unable to form (Brown et al., 1999).

Scx is not consistently detectable by qPCR in ES cells maintained in LIF and serum (data not shown) although it does become upregulated in EpiSCs. As shown in figure 6.6, during the 2i to EpiSCs transition, *Scx* was upregulated at later time point than *Tcf15*. Intriguingly, the *Scx* mutant phenotype begins to manifest itself shortly after *Tcf15* is downregulated after implantation (Brown et al., 1999) (Figure 3.1), supporting the idea that these two transcription factors act sequentially during development and may exhibit some functional redundancy. However, since Scx is not expressed in

ES cells and therefore, unlike Tcf15, does not mark an early epiblast-primed pluripotent state. It will be of interest in future studies to explore the relevance of our ES cells based studies for early embryonic development.

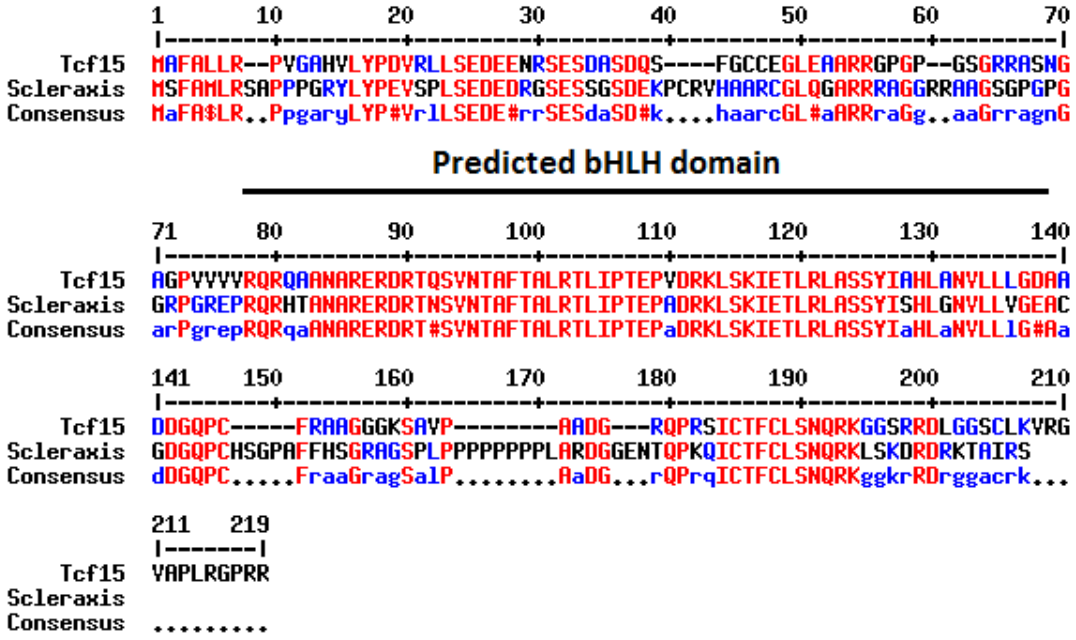


Figure 6.5: Protein sequence alignment between Tcf15 and Scleraxis

The protein sequence of mouse Tcf15 (ENSMUSP00000086511) and Scleraxis (ENSMUSP00000043668) are retrieved from Ensembl genome database. Sequences are aligned using Multalign website (Corpet, 1988). Tcf15 and Scleraxis shared almost identical bHLH DNA binding domain sequences, as computationally predicted by InterProScan software (Jones et al., 2014) and summarized at the Ensembl genome browser.

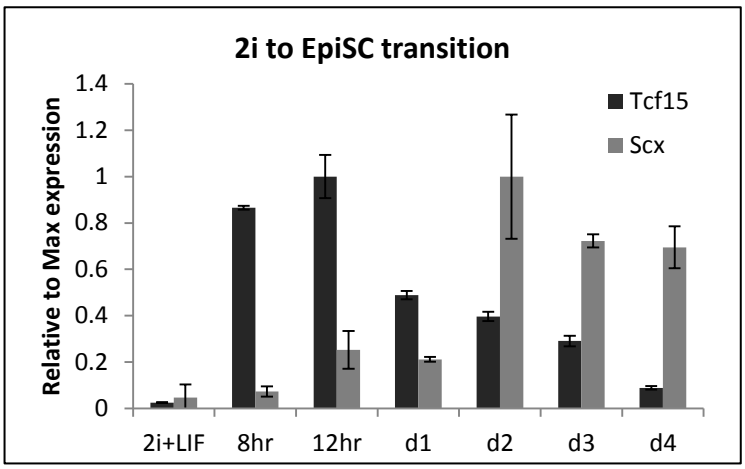


Figure 6.6: Tcf15 is expressed earlier than Scleraxis during the transition to epiblast-like state

qPCR analysis for *Tcf15* (black) and *Scx* (grey) during the transition from naïve pluripotency (2i/LIF) toward epiblast (N2B27 + FGF2 + Activin).

6.4: Concluding remarks

The transcriptional circuits that enable ES cells to maintain naïve pluripotency have been extensively characterized (Jaenisch and Young, 2008; Niwa, 2007). In comparison, less is known about the transcription factors that guide cells out of the pluripotent ground state toward commitment to differentiation. This study identified that Tcf15 could be a marker of a transient cell state that stands midway between naïve pluripotency and somatic lineage commitment.

The work presented in this thesis has shown that *Tcf15* is expressed in the ICM of pre-implantation embryos. In pluripotent ES cells, Tcf15 is heterogeneously expressed and is negatively correlated with naïve pluripotency markers. Tcf15-high cells are not committed to differentiation, but are functionally primed for rapid differentiation. Moreover, a transient expression window of *Tcf15* as cells exit a naïve self-renewing state and before they start to acquire lineage-specific gene expression suggested that Tcf15 may play a role in the early phase of differentiation.

Gain of function analysis revealed that Tcf15 acts as a pro-differentiation factor within pluripotent populations, priming ES cells to differentiate toward somatic lineages. The priming effect of Tcf15 is possibly acting during an early stage of ES cells transition to differentiation-primed epiblast but not extraembryonic endoderm. This rose the question of what would be the mechanism by which Tcf15 expression becomes associated with this epiblast-primed state and how does Tcf15 accelerate ES cells differentiation.

Tcf15 expression is dependent on FGF signalling. Microarray analysis

identified that Tcf15 downregulated expression of the naïve pluripotency marker *Nanog*. Tcf15 activation is also been observed to be associated with changes in the expression of particular regulators of cell adhesion and polarity, and with a change in the morphological appearance of cell colonies. The ability of Tcf15 to prime ES cells toward differentiation could therefore be explained either by its ability to downregulate *Nanog* or through its putative effects on the epithelialization of epiblast (Figure 6.7) or both.

In summary, this study identified Tcf15 as a transcription factor that can be used to monitor exit from the pluripotent state toward somatic lineages and which could serve as a valuable tool for monitoring and interrogating early developmental transitions. Tcf15 acts as a marker of this transition state: *Tcf15* expression is driven by FGF signalling, whereas its activity is suppressed by Id proteins, which are direct targets of BMP signalling (Nakashima et al., 2001; Ying et al., 2003; Wilson-Rawls et al., 2004); this helps explain how these extrinsic signals allow pluripotent cells to become primed for, but restrained from, somatic differentiation.

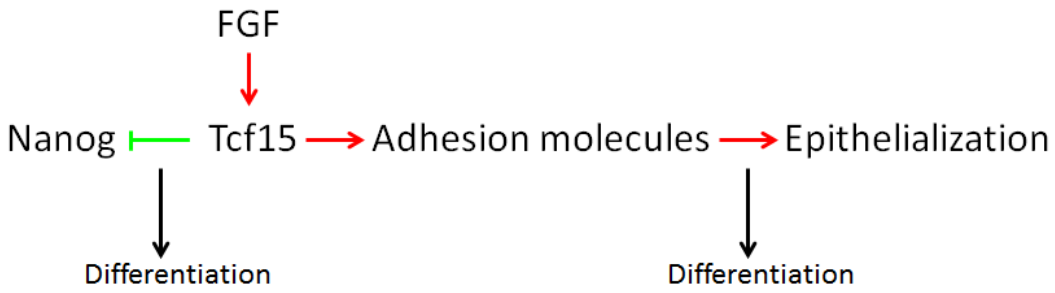


Figure 6.7: Proposed model of Tcf15 function in priming ES cells toward differentiation

Tcf15 expression is driven by FGF4/ERK signalling pathway. The mechanism by which Tcf15 maturates naïve pluripotent stem cells toward epiblast-primed state and further toward differentiation is likely due to: Tcf15 downregulates naïve pluripotency marker Nanog or, Tcf15 activates changes of adhesion molecules expression and thus affects the epithelialization of epiblast thus triggers cell differentiation.

References

ACAMPORA, D., DI GIOVANNANTONIO, L. G. & SIMEONE, A. 2012. Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition. *Development,* 140**,** 43-55.

ACAMPORA, D., MAZAN, S., LALLEMAND, Y., AVANTAGGIATO, V., MAURY, M., SIMEONE, A. & BRULET, P. 1995. Forebrain and midbrain regions are deleted in Otx2-/- mutants due to a defective anterior neuroectoderm specification during gastrulation. *Development,* 121**,** 3279-90.

AIBA, K., NEDOREZOV, T., PIAO, Y., NISHIYAMA, A., MATOBA, R., SHAROVA, L. V., SHAROV, A. A., YAMANAKA, S., NIWA, H. & KO, M. S. 2009. Defining developmental potency and cell lineage trajectories by expression profiling of differentiating mouse embryonic stem cells. *DNA Res,* 16**,** 73-80.

AIBA, K., SHAROV, A. A., CARTER, M. G., FORONI, C., VESCOVI, A. L. & KO, M. S. 2006. Defining a developmental path to neural fate by global expression profiling of mouse embryonic stem cells and adult neural stem/progenitor cells. *Stem Cells,* 24**,** 889-95.

ALEXOPOULOU, A. N., COUCHMAN, J. R. & WHITEFORD, J. R. 2008. The CMV early enhancer/chicken beta actin (CAG) promoter can be used to drive transgene expression during the differentiation of murine embryonic stem cells into vascular progenitors. *BMC Cell Biol,* 9**,** 2.

AMBROSETTI, D. C., BASILICO, C. & DAILEY, L. 1997. Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites. *Mol Cell Biol,* 17**,** 6321-9.

AMBROSETTI, D. C., SCHOLER, H. R., DAILEY, L. & BASILICO, C. 2000. Modulation of the activity of multiple transcriptional activation domains by the DNA binding domains mediates the synergistic action of Sox2 and Oct-3 on the fibroblast growth factor-4 enhancer. *J Biol Chem,* 275**,** 23387-97.

ARONHEIM, A., SHIRAN, R., ROSEN, A. & WALKER, M. D. 1993. The E2A gene product contains two separable and functionally distinct

transcription activation domains. *Proc Natl Acad Sci U S A,* 90**,** 8063-7.
AVILION, A. A., NICOLIS, S. K., PEVNY, L. H., PEREZ, L., VIVIAN, N. & LOVELL-BADGE, R. 2003. Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev,* 17**,** 126-40.
BAE, S., BESSHO, Y., HOJO, M. & KAGEYAMA, R. 2000. The bHLH gene Hes6, an inhibitor of Hes1, promotes neuronal differentiation. *Development,* 127**,** 2933-43.
BANASZYNSKI, L. A., CHEN, L. C., MAYNARD-SMITH, L. A., OOI, A. G. & WANDLESS, T. J. 2006. A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. *Cell,* 126**,** 995-1004.
BANGSOW, T., BAUMANN, E., BANGSOW, C., JAEGER, M. H., PELZER, B., GRUHN, P., WOLF, S., VON MELCHNER, H. & STANIMIROVIC, D. B. 2008. The epithelial membrane protein 1 is a novel tight junction protein of the blood-brain barrier. *J Cereb Blood Flow Metab,* 28**,** 1249-60.
BAO, S., TANG, F., LI, X., HAYASHI, K., GILLICH, A., LAO, K. & SURANI, M. A. 2009. Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. *Nature,* 461**,** 1292-5.
BARNES, R. M. & FIRULLI, A. B. 2009. A twist of insight - the role of Twist-family bHLH factors in development. *Int J Dev Biol,* 53**,** 909-24.
BAYART, E. & COHEN-HAGUENAUER, O. 2013. Technological overview of iPS induction from human adult somatic cells. *Curr Gene Ther,* 13**,** 73-92.
BECK, F., ERLER, T., RUSSELL, A. & JAMES, R. 1995. Expression of Cdx-2 in the mouse embryo and placenta: possible role in patterning of the extra-embryonic membranes. *Dev Dyn,* 204**,** 219-27.
BEDZHOV, I. & ZERNICKA-GOETZ, M. 2014. Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. *Cell,* 156**,** 1032-44.
BEISEL, C. & PARO, R. 2011. Silencing chromatin: comparing modes and mechanisms. *Nat Rev Genet,* 12**,** 123-35.
BEN-PORATH, I. & BENVENISTY, N. 1996. Characterization of a tumor-associated gene, a member of a novel family of genes encoding membrane glycoproteins. *Gene,* 183**,** 69-75.
BEN-SHUSHAN, E., THOMPSON, J. R., GUDAS, L. J. & BERGMAN, Y. 1998. Rex-1, a gene encoding a transcription factor expressed in the early

embryo, is regulated via Oct-3/4 and Oct-6 binding to an octamer site and a novel protein, Rox-1, binding to an adjacent site. *Mol Cell Biol,* 18**,** 1866-78.

BENEZRA, R., DAVIS, R. L., LOCKSHON, D., TURNER, D. L. & WEINTRAUB, H. 1990. The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. *Cell,* 61**,** 49-59.

BERNSTEIN, B. E., MIKKELSEN, T. S., XIE, X., KAMAL, M., HUEBERT, D. J., CUFF, J., FRY, B., MEISSNER, A., WERNIG, M., PLATH, K., JAENISCH, R., WAGSCHAL, A., FEIL, R., SCHREIBER, S. L. & LANDER, E. S. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell,* 125**,** 315-26.

BIRD, A. 2002. DNA methylation patterns and epigenetic memory. *Genes Dev,* 16**,** 6-21.

BOIANI, M. & SCHOLER, H. R. 2005. Regulatory networks in embryo-derived pluripotent stem cells. *Nat Rev Mol Cell Biol,* 6**,** 872-84.

BOULTON, T. G., STAHL, N. & YANCOPOULOS, G. D. 1994. Ciliary neurotrophic factor/leukemia inhibitory factor/interleukin 6/oncostatin M family of cytokines induces tyrosine phosphorylation of a common set of proteins overlapping those induced by other cytokines and growth factors. *J Biol Chem,* 269**,** 11648-55.

BOYER, L. A., LEE, T. I., COLE, M. F., JOHNSTONE, S. E., LEVINE, S. S., ZUCKER, J. P., GUENTHER, M. G., KUMAR, R. M., MURRAY, H. L., JENNER, R. G., GIFFORD, D. K., MELTON, D. A., JAENISCH, R. & YOUNG, R. A. 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell,* 122**,** 947-56.

BRADFORD, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem,* 72**,** 248-54.

BRADLEY, A., EVANS, M., KAUFMAN, M. H. & ROBERTSON, E. 1984. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature,* 309**,** 255-6.

BRENNAN, J., LU, C. C., NORRIS, D. P., RODRIGUEZ, T. A., BEDDINGTON, R. S. & ROBERTSON, E. J. 2001. Nodal signalling in the epiblast patterns the early mouse embryo. *Nature,* 411**,** 965-9.

BRINSTER, R. L. 1974. The effect of cells transferred into the mouse blastocyst on subsequent development. *J Exp Med,* 140**,** 1049-56.

BRONS, I. G., SMITHERS, L. E., TROTTER, M. W., RUGG-GUNN, P., SUN,

B., CHUVA DE SOUSA LOPES, S. M., HOWLETT, S. K., CLARKSON, A., AHRLUND-RICHTER, L., PEDERSEN, R. A. & VALLIER, L. 2007. Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature,* 448**,** 191-5.

BROWN, D., WAGNER, D., LI, X., RICHARDSON, J. A. & OLSON, E. N. 1999. Dual role of the basic helix-loop-helix transcription factor scleraxis in mesoderm formation and chondrogenesis during mouse embryogenesis. *Development,* 126**,** 4317-29.

BUECKER, C., SRINIVASAN, R., WU, Z., CALO, E., ACAMPORA, D., FAIAL, T., SIMEONE, A., TAN, M., SWIGUT, T. & WYSOCKA, J. 2014. Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell Stem Cell,* 14**,** 838-53.

BUEHR, M., MEEK, S., BLAIR, K., YANG, J., URE, J., SILVA, J., MCLAY, R., HALL, J., YING, Q. L. & SMITH, A. 2008. Capture of authentic embryonic stem cells from rat blastocysts. *Cell,* 135**,** 1287-98.

BURDON, T., STRACEY, C., CHAMBERS, I., NICHOLS, J. & SMITH, A. 1999. Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. *Dev Biol,* 210**,** 30-43.

BURGESS, R., CSERJESI, P., LIGON, K. L. & OLSON, E. N. 1995. Paraxis: a basic helix-loop-helix protein expressed in paraxial mesoderm and developing somites. *Dev Biol,* 168**,** 296-306.

BURGESS, R., RAWLS, A., BROWN, D., BRADLEY, A. & OLSON, E. N. 1996. Requirement of the paraxis gene for somite formation and musculoskeletal patterning. *Nature,* 384**,** 570-3.

CANHAM, M. A., SHAROV, A. A., KO, M. S. & 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., COLBY, D., ROBERTSON, M., NICHOLS, J., LEE, S., TWEEDIE, S. & SMITH, A. 2003. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell,* 113**,** 643-55.

CHAMBERS, I., SILVA, J., COLBY, D., NICHOLS, J., NIJMEIJER, B., ROBERTSON, M., VRANA, J., JONES, K., GROTEWOLD, L. & SMITH, A. 2007. Nanog safeguards pluripotency and mediates germline development. *Nature,* 450**,** 1230-4.

CHAMBERS, I. & SMITH, A. 2004. Self-renewal of teratocarcinoma and

embryonic stem cells. *Oncogene,* 23**,** 7150-60.
CHAMBERS, I. & TOMLINSON, S. R. 2009. The transcriptional foundation of pluripotency. *Development,* 136**,** 2311-22.
CHAZAUD, C., YAMANAKA, Y., PAWSON, T. & ROSSANT, J. 2006. Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. *Dev Cell,* 10**,** 615-24.
CHEN, F. F., JIANG, G., XU, K. & ZHENG, J. N. 2013. Function and mechanism by which interferon regulatory factor-1 inhibits oncogenesis. *Oncol Lett,* 5**,** 417-423.
CHEN, J. R., CHENG, J. G., SHATZER, T., SEWELL, L., HERNANDEZ, L. & STEWART, C. L. 2000. Leukemia inhibitory factor can substitute for nidatory estrogen and is essential to inducing a receptive uterus for implantation but is not essential for subsequent embryogenesis. *Endocrinology,* 141**,** 4365-72.
CHEN, X., FANG, F., LIOU, Y. C. & NG, H. H. 2008a. Zfp143 regulates Nanog through modulation of Oct4 binding. *Stem Cells,* 26**,** 2759-67.
CHEN, X., XU, H., YUAN, P., FANG, F., HUSS, M., VEGA, V. B., WONG, E., ORLOV, Y. L., ZHANG, W., JIANG, J., LOH, Y. H., YEO, H. C., YEO, Z. X., NARANG, V., GOVINDARAJAN, K. R., LEONG, B., SHAHAB, A., RUAN, Y., BOURQUE, G., SUNG, W. K., CLARKE, N. D., WEI, C. L. & NG, H. H. 2008b. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell,* 133**,** 1106-17.
CHENOWETH, J. G., MCKAY, R. D. & TESAR, P. J. 2010. Epiblast stem cells contribute new insight into pluripotency and gastrulation. *Dev Growth Differ,* 52**,** 293-301.
CHRISTY, B. A., SANDERS, L. K., LAU, L. F., COPELAND, N. G., JENKINS, N. A. & NATHANS, D. 1991. An Id-related helix-loop-helix protein encoded by a growth factor-inducible gene. *Proc Natl Acad Sci U S A,* 88**,** 1815-9.
CONNERNEY, J., ANDREEVA, V., LESHEM, Y., MERCADO, M. A., DOWELL, K., YANG, X., LINDNER, V., FRIESEL, R. E. & SPICER, D. B. 2008. Twist1 homodimers enhance FGF responsiveness of the cranial sutures and promote suture closure. *Dev Biol,* 318**,** 323-34.
CORAUX, C., HILMI, C., ROULEAU, M., SPADAFORA, A., HINNRASKY, J., ORTONNE, J. P., DANI, C. & ABERDAM, D. 2003. Reconstituted skin

from murine embryonic stem cells. *Curr Biol,* 13**,** 849-53.
CORPET, F. 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res,* 16**,** 10881-90.
COSTA, Y., DING, J., THEUNISSEN, T. W., FAIOLA, F., HORE, T. A., SHLIAHA, P. V., FIDALGO, M., SAUNDERS, A., LAWRENCE, M., DIETMANN, S., DAS, S., LEVASSEUR, D. N., LI, Z., XU, M., REIK, W., SILVA, J. C. & WANG, J. 2013. NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature,* 495**,** 370-4.
DAVIES, O. R., LIN, C. Y., RADZISHEUSKAYA, A., ZHOU, X., TAUBE, J., BLIN, G., WATERHOUSE, A., SMITH, A. J. & LOWELL, S. 2013. Tcf15 primes pluripotent cells for differentiation. *Cell Rep,* 3**,** 472-84.
DESBAILLETS, I., ZIEGLER, U., GROSCURTH, P. & GASSMANN, M. 2000. Embryoid bodies: an in vitro model of mouse embryogenesis. *Exp Physiol,* 85**,** 645-51.
DI-GREGORIO, A., SANCHO, M., STUCKEY, D. W., CROMPTON, L. A., GODWIN, J., MISHINA, Y. & RODRIGUEZ, T. A. 2007. BMP signalling inhibits premature neural differentiation in the mouse embryo. *Development,* 134**,** 3359-69.
DODGE, J. E., KANG, Y. K., BEPPU, H., LEI, H. & LI, E. 2004. Histone H3-K9 methyltransferase ESET is essential for early development. *Mol Cell Biol,* 24**,** 2478-86.
DUBAND, J. L., DUFOUR, S., HATTA, K., TAKEICHI, M., EDELMAN, G. M. & THIERY, J. P. 1987. Adhesion molecules during somitogenesis in the avian embryo. *J Cell Biol,* 104**,** 1361-74.
DUNNING, M. J., SMITH, M. L., RITCHIE, M. E. & TAVARE, S. 2007. beadarray: R classes and methods for Illumina bead-based data. *Bioinformatics,* 23**,** 2183-4.
EFRONI, S., DUTTAGUPTA, R., CHENG, J., DEHGHANI, H., HOEPPNER, D. J., DASH, C., BAZETT-JONES, D. P., LE GRICE, S., MCKAY, R. D., BUETOW, K. H., GINGERAS, T. R., MISTELI, T. & MESHORER, E. 2008. Global transcription in pluripotent embryonic stem cells. *Cell Stem Cell,* 2**,** 437-47.
EPSZTEJN-LITMAN, S., FELDMAN, N., ABU-REMAILEH, M., SHUFARO, Y., GERSON, A., UEDA, J., DEPLUS, R., FUKS, F., SHINKAI, Y., CEDAR, H. & BERGMAN, Y. 2008. De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat Struct Mol Biol,* 15**,** 1176-83.

EVANS, M. J. & KAUFMAN, M. H. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature,* 292**,** 154-6.

FACTOR, D. C., CORRADIN, O., ZENTNER, G. E., SAIAKHOVA, A., SONG, L., CHENOWETH, J. G., MCKAY, R. D., CRAWFORD, G. E., SCACHERI, P. C. & TESAR, P. J. 2014. Epigenomic comparison reveals activation of "seed" enhancers during transition from naive to primed pluripotency. *Cell Stem Cell,* 14**,** 854-63.

FASSLER, R. & MEYER, M. 1995. Consequences of lack of beta 1 integrin gene expression in mice. *Genes Dev,* 9**,** 1896-908.

FAZZIO, T. G. & PANNING, B. 2010. Control of embryonic stem cell identity by nucleosome remodeling enzymes. *Curr Opin Genet Dev,* 20**,** 500-4.

FELDMAN, N., GERSON, A., FANG, J., LI, E., ZHANG, Y., SHINKAI, Y., CEDAR, H. & BERGMAN, Y. 2006. G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat Cell Biol,* 8**,** 188-94.

FENG, B., JIANG, J., KRAUS, P., NG, J. H., HENG, J. C., CHAN, Y. S., YAW, L. P., ZHANG, W., LOH, Y. H., HAN, J., VEGA, V. B., CACHEUX-RATABOUL, V., LIM, B., LUFKIN, T. & NG, H. H. 2009. Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat Cell Biol,* 11**,** 197-203.

FESTUCCIA, N., OSORNO, R., HALBRITTER, F., KARWACKI-NEISIUS, V., NAVARRO, P., COLBY, D., WONG, F., YATES, A., TOMLINSON, S. R. & CHAMBERS, I. 2012. Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. *Cell Stem Cell,* 11**,** 477-90.

FIDALGO, M., FAIOLA, F., PEREIRA, C. F., DING, J., SAUNDERS, A., GINGOLD, J., SCHANIEL, C., LEMISCHKA, I. R., SILVA, J. C. & WANG, J. 2012. Zfp281 mediates Nanog autorepression through recruitment of the NuRD complex and inhibits somatic cell reprogramming. *Proc Natl Acad Sci U S A,* 109**,** 16202-7.

FINLAY, D., PATEL, S., DICKSON, L. M., SHPIRO, N., MARQUEZ, R., RHODES, C. J. & SUTHERLAND, C. 2004. Glycogen synthase kinase-3 regulates IGFBP-1 gene transcription through the thymine-rich insulin response element. *BMC Mol Biol,* 5**,** 15.

FINLEY, M. F., DEVATA, S. & HUETTNER, J. E. 1999. BMP-4 inhibits neural differentiation of murine embryonic stem cells. *J Neurobiol,* 40**,** 271-87.

FODOR, B. D., SHUKEIR, N., REUTER, G. & JENUWEIN, T. 2010.

Mammalian Su(var) genes in chromatin control. *Annu Rev Cell Dev Biol,* 26**,** 471-501.

FRAIDENRAICH, D., STILLWELL, E., ROMERO, E., WILKES, D., MANOVA, K., BASSON, C. T. & BENEZRA, R. 2004. Rescue of cardiac defects in id knockout embryos by injection of embryonic stem cells. *Science,* 306**,** 247-52.

FRANKENBERG, S., GERBE, F., BESSONNARD, S., BELVILLE, C., POUCHIN, P., BARDOT, O. & CHAZAUD, C. 2011. Primitive endoderm differentiates via a three-step mechanism involving Nanog and RTK signaling. *Dev Cell,* 21**,** 1005-13.

FUJIKURA, J., YAMATO, E., YONEMURA, S., HOSODA, K., MASUI, S., NAKAO, K., MIYAZAKI JI, J. & NIWA, H. 2002. Differentiation of embryonic stem cells is induced by GATA factors. *Genes Dev,* 16**,** 784-9.

FURUSAWA, T., OHKOSHI, K., HONDA, C., TAKAHASHI, S. & 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. *Biol Reprod,* 70**,** 1452-7.

FUTAKI, S., HAYASHI, Y., EMOTO, T., WEBER, C. N. & SEKIGUCHI, K. 2004. Sox7 plays crucial roles in parietal endoderm differentiation in F9 embryonal carcinoma cells through regulating Gata-4 and Gata-6 expression. *Mol Cell Biol,* 24**,** 10492-503.

GAGLIARDI, A., MULLIN, N. P., YING TAN, Z., COLBY, D., KOUSA, A. I., HALBRITTER, F., WEISS, J. T., FELKER, A., BEZSTAROSTI, K., FAVARO, R., DEMMERS, J., NICOLIS, S. K., TOMLINSON, S. R., POOT, R. A. & CHAMBERS, I. 2013. A direct physical interaction between Nanog and Sox2 regulates embryonic stem cell self-renewal. *EMBO J,* 32**,** 2231-47.

GAO, X., YANG, J., TSANG, J. C., OOI, J., WU, D. & LIU, P. 2013a. Reprogramming to Pluripotency Using Designer TALE Transcription Factors Targeting Enhancers. *Stem Cell Reports,* 1**,** 183-97.

GAO, Y., CHEN, J., LI, K., WU, T., HUANG, B., LIU, W., KOU, X., ZHANG, Y., HUANG, H., JIANG, Y., YAO, C., LIU, X., LU, Z., XU, Z., KANG, L., CHEN, J., WANG, H., CAI, T. & GAO, S. 2013b. Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. *Cell Stem Cell,* 12**,** 453-69.

GEARING, D. P., COMEAU, M. R., FRIEND, D. J., GIMPEL, S. D., THUT, C. J., MCGOURTY, J., BRASHER, K. K., KING, J. A., GILLIS, S., MOSLEY, B. & ET AL. 1992. The IL-6 signal transducer, gp130: an oncostatin M receptor and affinity converter for the LIF receptor. *Science,* 255**,** 1434-7.

GENTLEMAN, R. C., CAREY, V. J., BATES, D. M., BOLSTAD, B., DETTLING, M., DUDOIT, S., ELLIS, B., GAUTIER, L., GE, Y., GENTRY, J., HORNIK, K., HOTHORN, T., HUBER, W., IACUS, S., IRIZARRY, R., LEISCH, F., LI, C., MAECHLER, M., ROSSINI, A. J., SAWITZKI, G., SMITH, C., SMYTH, G., TIERNEY, L., YANG, J. Y. & ZHANG, J. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol,* 5**,** R80.

GONTAN, C., ACHAME, E. M., DEMMERS, J., BARAKAT, T. S., RENTMEESTER, E., VAN, I. W., GROOTEGOED, J. A. & GRIBNAU, J. 2012. RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. *Nature,* 485**,** 386-90.

GRAF, T. & STADTFELD, M. 2008. Heterogeneity of embryonic and adult stem cells. *Cell Stem Cell,* 3**,** 480-3.

GU, P., GOODWIN, B., CHUNG, A. C., XU, X., WHEELER, D. A., PRICE, R. R., GALARDI, C., PENG, L., LATOUR, A. M., KOLLER, B. H., GOSSEN, J., KLIEWER, S. A. & COONEY, A. J. 2005. Orphan nuclear receptor LRH-1 is required to maintain Oct4 expression at the epiblast stage of embryonic development. *Mol Cell Biol,* 25**,** 3492-505.

GUENTHER, M. G., FRAMPTON, G. M., SOLDNER, F., HOCKEMEYER, D., MITALIPOVA, M., JAENISCH, R. & YOUNG, R. A. 2010. Chromatin structure and gene expression programs of human embryonic and induced pluripotent stem cells. *Cell Stem Cell,* 7**,** 249-57.

GUO, G., HUSS, M., TONG, G. Q., WANG, C., LI SUN, L., CLARKE, N. D. & ROBSON, P. 2010. Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. *Dev Cell,* 18**,** 675-85.

GUO, G., YANG, J., NICHOLS, J., HALL, J. S., EYRES, I., MANSFIELD, W. & SMITH, A. 2009. Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development,* 136**,** 1063-9.

HALL, J., GUO, G., WRAY, J., EYRES, I., NICHOLS, J., GROTEWOLD, L., MORFOPOULOU, S., HUMPHREYS, P., MANSFIELD, W., WALKER, R., TOMLINSON, S. & SMITH, A. 2009. Oct4 and LIF/Stat3 additively

induce Kruppel factors to sustain embryonic stem cell self-renewal. *Cell Stem Cell,* 5**,** 597-609.

HANNA, J., MARKOULAKI, S., MITALIPOVA, M., CHENG, A. W., CASSADY, J. P., STAERK, J., CAREY, B. W., LENGNER, C. J., FOREMAN, R., LOVE, J., GAO, Q., KIM, J. & JAENISCH, R. 2009. Metastable pluripotent states in NOD-mouse-derived ESCs. *Cell Stem Cell,* 4**,** 513-24.

HANSSON, M., OLESEN, D. R., PETERSLUND, J. M., ENGBERG, N., KAHN, M., WINZI, M., KLEIN, T., MADDOX-HYTTEL, P. & SERUP, P. 2009. A late requirement for Wnt and FGF signaling during activin-induced formation of foregut endoderm from mouse embryonic stem cells. *Dev Biol,* 330**,** 286-304.

HAO, J., LI, T. G., QI, X., ZHAO, D. F. & ZHAO, G. Q. 2006. WNT/beta-catenin pathway up-regulates Stat3 and converges on LIF to prevent differentiation of mouse embryonic stem cells. *Dev Biol,* 290**,** 81-91.

HATTA, K., TAKAGI, S., FUJISAWA, H. & TAKEICHI, M. 1987. Spatial and temporal expression pattern of N-cadherin cell adhesion molecules correlated with morphogenetic processes of chicken embryos. *Dev Biol,* 120**,** 215-27.

HAYASHI, K., LOPES, S. M., TANG, F. & 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, K., OHTA, H., KURIMOTO, K., ARAMAKI, S. & SAITOU, M. 2011. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell,* 146**,** 519-32.

HEARD, E. 2004. Recent advances in X-chromosome inactivation. *Curr Opin Cell Biol,* 16**,** 247-55.

HENG, J. C., FENG, B., HAN, J., JIANG, J., KRAUS, P., NG, J. H., ORLOV, Y. L., HUSS, M., YANG, L., LUFKIN, T., LIM, B. & NG, H. H. 2010. The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. *Cell Stem Cell,* 6**,** 167-74.

HOFFMAN, L. M., NIX, D. A., BENSON, B., BOOT-HANFORD, R., GUSTAFSSON, E., JAMORA, C., MENZIES, A. S., GOH, K. L., JENSEN, C. C., GERTLER, F. B., FUCHS, E., FASSLER, R. & BECKERLE, M. C. 2003. Targeted disruption of the murine zyxin gene. *Mol Cell Biol,* 23**,** 70-9.

HOHENSTEIN, P. & HASTIE, N. D. 2006. The many facets of the Wilms'

tumour gene, WT1. *Hum Mol Genet,* 15 Spec No 2**,** R196-201.
HOOPER, M., HARDY, K., HANDYSIDE, A., HUNTER, S. & MONK, M. 1987. HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature,* 326**,** 292-5.
HOPFL, G., GASSMANN, M. & DESBAILLETS, I. 2004. Differentiating embryonic stem cells into embryoid bodies. *Methods Mol Biol,* 254**,** 79-98.
HOSLER, B. A., LAROSA, G. J., GRIPPO, J. F. & GUDAS, L. J. 1989. Expression of REX-1, a gene containing zinc finger motifs, is rapidly reduced by retinoic acid in F9 teratocarcinoma cells. *Mol Cell Biol,* 9**,** 5623-9.
HUANG, Y., OSORNO, R., TSAKIRIDIS, A. & WILSON, V. 2012. In Vivo differentiation potential of epiblast stem cells revealed by chimeric embryo formation. *Cell Rep,* 2**,** 1571-8.
HUTCHINS, A. P., CHOO, S. H., MISTRI, T. K., RAHMANI, M., WOON, C. T., NG, C. K., JAUCH, R. & ROBSON, P. 2013. Co-motif discovery identifies an Esrrb-Sox2-DNA ternary complex as a mediator of transcriptional differences between mouse embryonic and epiblast stem cells. *Stem Cells,* 31**,** 269-81.
HYNES, R. O. 2002. Integrins: bidirectional, allosteric signaling machines. *Cell,* 110**,** 673-87.
ITO, S., D'ALESSIO, A. C., TARANOVA, O. V., HONG, K., SOWERS, L. C. & ZHANG, Y. 2010. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature,* 466**,** 1129-33.
IVANOVA, N., DOBRIN, R., LU, R., KOTENKO, I., LEVORSE, J., DECOSTE, C., SCHAFER, X., LUN, Y. & LEMISCHKA, I. R. 2006. Dissecting self-renewal in stem cells with RNA interference. *Nature,* 442**,** 533-8.
JACKSON, M., KRASSOWSKA, A., GILBERT, N., CHEVASSUT, T., FORRESTER, L., ANSELL, J. & RAMSAHOYE, B. 2004. Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. *Mol Cell Biol,* 24**,** 8862-71.
JEN, Y., MANOVA, K. & BENEZRA, R. 1997. Each member of the Id gene family exhibits a unique expression pattern in mouse gastrulation and neurogenesis. *Dev Dyn,* 208**,** 92-106.
JIANG, J., CHAN, Y. S., LOH, Y. H., CAI, J., TONG, G. Q., LIM, C. A., ROBSON, P., ZHONG, S. & NG, H. H. 2008. A core Klf circuitry regulates self-renewal of embryonic stem cells. *Nat Cell Biol,* 10**,** 353-60.

JOHNSON, J., RHEE, J., PARSONS, S. M., BROWN, D., OLSON, E. N. & RAWLS, A. 2001. The anterior/posterior polarity of somites is disrupted in paraxis-deficient mice. *Dev Biol,* 229**,** 176-87.

JONES, P., BINNS, D., CHANG, H. Y., FRASER, M., LI, W., MCANULLA, C., MCWILLIAM, H., MASLEN, J., MITCHELL, A., NUKA, G., PESSEAT, S., QUINN, A. F., SANGRADOR-VEGAS, A., SCHEREMETJEW, M., YONG, S. Y., LOPEZ, R. & HUNTER, S. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics,* 30**,** 1236-40.

KAHAN, B. W. & EPHRUSSI, B. 1970. Developmental potentialities of clonal in vitro cultures of mouse testicular teratoma. *J Natl Cancer Inst,* 44**,** 1015-36.

KAJI, K., NICHOLS, J. & HENDRICH, B. 2007. Mbd3, a component of the NuRD co-repressor complex, is required for development of pluripotent cells. *Development,* 134**,** 1123-32.

KALMAR, T., LIM, C., HAYWARD, P., MUNOZ-DESCALZO, S., NICHOLS, J., GARCIA-OJALVO, J. & MARTINEZ ARIAS, A. 2009. Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells. *PLoS Biol,* 7**,** e1000149.

KAMIYA, D., BANNO, S., SASAI, N., OHGUSHI, M., INOMATA, H., WATANABE, K., KAWADA, M., YAKURA, R., KIYONARI, H., NAKAO, K., JAKT, L. M., NISHIKAWA, S. & SASAI, Y. 2011. Intrinsic transition of embryonic stem-cell differentiation into neural progenitors. *Nature,* 470**,** 503-9.

KARWACKI-NEISIUS, V., GOKE, J., OSORNO, R., HALBRITTER, F., NG, J. H., WEISSE, A. Y., WONG, F. C., GAGLIARDI, A., MULLIN, N. P., FESTUCCIA, N., COLBY, D., TOMLINSON, S. R., NG, H. H. & CHAMBERS, I. 2013. Reduced Oct4 expression directs a robust pluripotent state with distinct signaling activity and increased enhancer occupancy by Oct4 and Nanog. *Cell Stem Cell,* 12**,** 531-45.

KAWASAKI, H., MIZUSEKI, K., NISHIKAWA, S., KANEKO, S., KUWANA, Y., NAKANISHI, S., NISHIKAWA, S. I. & SASAI, Y. 2000. Induction of midbrain dopaminergic neurons from ES cells by stromal cell-derived inducing activity. *Neuron,* 28**,** 31-40.

KEYSE, S. M. 2008. Dual-specificity MAP kinase phosphatases (MKPs) and cancer. *Cancer Metastasis Rev,* 27**,** 253-61.

KIM, J., CHU, J., SHEN, X., WANG, J. & ORKIN, S. H. 2008. An extended transcriptional network for pluripotency of embryonic stem cells. *Cell,*

132**,** 1049-61.
KIM, J. B., SEBASTIANO, V., WU, G., ARAUZO-BRAVO, M. J., SASSE, P., GENTILE, L., KO, K., RUAU, D., EHRICH, M., VAN DEN BOOM, D., MEYER, J., HUBNER, K., BERNEMANN, C., ORTMEIER, C., ZENKE, M., FLEISCHMANN, B. K., ZAEHRES, H. & SCHOLER, H. R. 2009. Oct4-induced pluripotency in adult neural stem cells. *Cell,* 136**,** 411-9.
KIYONARI, H., KANEKO, M., ABE, S. & AIZAWA, S. 2010. Three inhibitors of FGF receptor, ERK, and GSK3 establishes germline-competent embryonic stem cells of C57BL/6N mouse strain with high efficiency and stability. *Genesis,* 48**,** 317-27.
KLEINSMITH, L. J. & PIERCE, G. B., JR. 1964. Multipotentiality of Single Embryonal Carcinoma Cells. *Cancer Res,* 24**,** 1544-51.
KOH, K. P., YABUUCHI, A., RAO, S., HUANG, Y., CUNNIFF, K., NARDONE, J., LAIHO, A., TAHILIANI, M., SOMMER, C. A., MOSTOSLAVSKY, G., LAHESMAA, R., ORKIN, S. H., RODIG, S. J., DALEY, G. Q. & RAO, A. 2011. Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell,* 8**,** 200-13.
KOKUBU, C., HEINZMANN, U., KOKUBU, T., SAKAI, N., KUBOTA, T., KAWAI, M., WAHL, M. B., GALCERAN, J., GROSSCHEDL, R., OZONO, K. & IMAI, K. 2004. Skeletal defects in ringelschwanz mutant mice reveal that Lrp6 is required for proper somitogenesis and osteogenesis. *Development,* 131**,** 5469-80.
KUNATH, T., SABA-EL-LEIL, M. K., ALMOUSAILLEAKH, M., WRAY, J., MELOCHE, S. & SMITH, A. 2007. FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development,* 134**,** 2895-902.
KURODA, T., TADA, M., KUBOTA, H., KIMURA, H., HATANO, S. Y., SUEMORI, H., NAKATSUJI, N. & TADA, T. 2005. Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression. *Mol Cell Biol,* 25**,** 2475-85.
KUROSAWA, H. 2007. Methods for inducing embryoid body formation: in vitro differentiation system of embryonic stem cells. *J Biosci Bioeng,* 103**,** 389-98.
LANNER, F., LEE, K. L., SOHL, M., HOLMBORN, K., YANG, H., WILBERTZ, J., POELLINGER, L., ROSSANT, J. & FARNEBO, F. 2010.

Heparan sulfation-dependent fibroblast growth factor signaling maintains embryonic stem cells primed for differentiation in a heterogeneous state. *Stem Cells,* 28**,** 191-200.

LANNER, F. & ROSSANT, J. 2010. The role of FGF/Erk signaling in pluripotent cells. *Development,* 137**,** 3351-60.

LASSAR, A. B., BUSKIN, J. N., LOCKSHON, D., DAVIS, R. L., APONE, S., HAUSCHKA, S. D. & WEINTRAUB, H. 1989. MyoD is a sequence-specific DNA binding protein requiring a region of myc homology to bind to the muscle creatine kinase enhancer. *Cell,* 58**,** 823-31.

LAWSON, K. A., MENESES, J. J. & PEDERSEN, R. A. 1991. Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development,* 113**,** 891-911.

LE BRAS, B., BARALLOBRE, M. J., HOMMAN-LUDIYE, J., NY, A., WYNS, S., TAMMELA, T., HAIKO, P., KARKKAINEN, M. J., YUAN, L., MURIEL, M. P., CHATZOPOULOU, E., BREANT, C., ZALC, B., CARMELIET, P., ALITALO, K., EICHMANN, A. & THOMAS, J. L. 2006. VEGF-C is a trophic factor for neural progenitors in the vertebrate embryonic brain. *Nat Neurosci,* 9**,** 340-8.

LEE, H. S., SHERLEY, J. L., CHEN, J. J., CHIU, C. C., CHIOU, L. L., LIANG, J. D., YANG, P. C., HUANG, G. T. & SHEU, J. C. 2005. EMP-1 is a junctional protein in a liver stem cell line and in the liver. *Biochem Biophys Res Commun,* 334**,** 996-1003.

LEE, J. E., HOLLENBERG, S. M., SNIDER, L., TURNER, D. L., LIPNICK, N. & WEINTRAUB, H. 1995. Conversion of Xenopus ectoderm into neurons by NeuroD, a basic helix-loop-helix protein. *Science,* 268**,** 836-44.

LEE, M. K., TUTTLE, J. B., REBHUN, L. I., CLEVELAND, D. W. & FRANKFURTER, A. 1990. The expression and posttranslational modification of a neuron-specific beta-tubulin isotype during chick embryogenesis. *Cell Motil Cytoskeleton,* 17**,** 118-32.

LENDAHL, U., ZIMMERMAN, L. B. & MCKAY, R. D. 1990. CNS stem cells express a new class of intermediate filament protein. *Cell,* 60**,** 585-95.

LI, G., LIU, T., TAROKH, A., NIE, J., GUO, L., MARA, A., HOLLEY, S. & WONG, S. T. 2007a. 3D cell nuclei segmentation based on gradient flow tracking. *BMC Cell Biol,* 8**,** 40.

LI, J. Y., PU, M. T., HIRASAWA, R., LI, B. Z., HUANG, Y. N., ZENG, R., JING, N. H., CHEN, T., LI, E., SASAKI, H. & XU, G. L. 2007b. Synergistic

function of DNA methyltransferases Dnmt3a and Dnmt3b in the methylation of Oct4 and Nanog. *Mol Cell Biol,* 27**,** 8748-59.

LI, M., LIU, G. H. & IZPISUA BELMONTE, J. C. 2012. Navigating the epigenetic landscape of pluripotent stem cells. *Nat Rev Mol Cell Biol,* 13**,** 524-35.

LI, P., TONG, C., MEHRIAN-SHAI, R., JIA, L., WU, N., YAN, Y., MAXSON, R. E., SCHULZE, E. N., SONG, H., HSIEH, C. L., PERA, M. F. & YING, Q. L. 2008. Germline competent embryonic stem cells derived from rat blastocysts. *Cell,* 135**,** 1299-310.

LI, Y., ZHANG, Q., YIN, X., YANG, W., DU, Y., HOU, P., GE, J., LIU, C., ZHANG, W., ZHANG, X., WU, Y., LI, H., LIU, K., WU, C., SONG, Z., ZHAO, Y., SHI, Y. & DENG, H. 2011. Generation of iPSCs from mouse fibroblasts with a single gene, Oct4, and small molecules. *Cell Res,* 21**,** 196-204.

LIANG, J., WAN, M., ZHANG, Y., GU, P., XIN, H., JUNG, S. Y., QIN, J., WONG, J., COONEY, A. J., LIU, D. & SONGYANG, Z. 2008. Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. *Nat Cell Biol,* 10**,** 731-9.

LINKER, C., LESBROS, C., GROS, J., BURRUS, L. W., RAWLS, A. & MARCELLE, C. 2005. beta-Catenin-dependent Wnt signalling controls the epithelial organisation of somites through the activation of paraxis. *Development,* 132**,** 3895-905.

LOGAN, C. Y. & NUSSE, R. 2004. The Wnt signaling pathway in development and disease. *Annu Rev Cell Dev Biol,* 20**,** 781-810.

LOH, Y. H., WU, Q., CHEW, J. L., VEGA, V. B., ZHANG, W., CHEN, X., BOURQUE, G., GEORGE, J., LEONG, B., LIU, J., WONG, K. Y., SUNG, K. W., LEE, C. W., ZHAO, X. D., CHIU, K. P., LIPOVICH, L., KUZNETSOV, V. A., ROBSON, P., STANTON, L. W., WEI, C. L., RUAN, Y., LIM, B. & NG, H. H. 2006. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet,* 38**,** 431-40.

LOHMANN, F., LOUREIRO, J., SU, H., FANG, Q., LEI, H., LEWIS, T., YANG, Y., LABOW, M., LI, E., CHEN, T. & KADAM, S. 2010. KMT1E mediated H3K9 methylation is required for the maintenance of embryonic stem cells by repressing trophectoderm differentiation. *Stem Cells,* 28**,** 201-12.

LUNYAK, V. V. & ROSENFELD, M. G. 2008. Epigenetic regulation of stem

cell fate. *Hum Mol Genet,* 17**,** R28-36.
MA, J. Y., OU-YANG, Y. C., LUO, Y. B., WANG, Z. B., HOU, Y., HAN, Z. M., LIU, Z., SCHATTEN, H. & SUN, Q. Y. 2013. Cyclin O regulates germinal vesicle breakdown in mouse oocytes. *Biol Reprod,* 88**,** 110.
MA, Y. G., ROSFJORD, E., HUEBERT, C., WILDER, P., TIESMAN, J., KELLY, D. & RIZZINO, A. 1992. Transcriptional regulation of the murine k-FGF gene in embryonic cell lines. *Dev Biol,* 154**,** 45-54.
MAK, W., NESTEROVA, T. B., DE NAPOLES, M., APPANAH, R., YAMANAKA, S., OTTE, A. P. & BROCKDORFF, N. 2004. Reactivation of the paternal X chromosome in early mouse embryos. *Science,* 303**,** 666-9.
MAKINEN, T., VEIKKOLA, T., MUSTJOKI, S., KARPANEN, T., CATIMEL, B., NICE, E. C., WISE, L., MERCER, A., KOWALSKI, H., KERJASCHKI, D., STACKER, S. A., ACHEN, M. G. & ALITALO, K. 2001. Isolated lymphatic endothelial cells transduce growth, survival and migratory signals via the VEGF-C/D receptor VEGFR-3. *EMBO J,* 20**,** 4762-73.
MALAGUTI, M., NISTOR, P. A., BLIN, G., PEGG, A., ZHOU, X. & LOWELL, S. 2013. Bone morphogenic protein signalling suppresses differentiation of pluripotent cells by maintaining expression of E-Cadherin. *Elife,* 2**,** e01197.
MARKS, H., KALKAN, T., MENAFRA, R., DENISSOV, S., JONES, K., HOFEMEISTER, H., NICHOLS, J., KRANZ, A., STEWART, A. F., SMITH, A. & STUNNENBERG, H. G. 2012. The transcriptional and epigenomic foundations of ground state pluripotency. *Cell,* 149**,** 590-604.
MARSON, A., LEVINE, S. S., COLE, M. F., FRAMPTON, G. M., BRAMBRINK, T., JOHNSTONE, S., GUENTHER, M. G., JOHNSTON, W. K., WERNIG, M., NEWMAN, J., CALABRESE, J. M., DENNIS, L. M., VOLKERT, T. L., GUPTA, S., LOVE, J., HANNETT, N., SHARP, P. A., BARTEL, D. P., JAENISCH, R. & YOUNG, R. A. 2008. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell,* 134**,** 521-33.
MARTELLO, G., BERTONE, P. & SMITH, A. 2013. Identification of the missing pluripotency mediator downstream of leukaemia inhibitory factor. *EMBO J,* 32**,** 2561-74.
MARTIN, G. R. 1981. Isolation of a pluripotent cell line from early mouse

embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A,* 78**,** 7634-8.

MARTIN, G. R. & EVANS, M. J. 1975. Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies in vitro. *Proc Natl Acad Sci U S A,* 72**,** 1441-5.

MARUYAMA, M., ICHISAKA, T., NAKAGAWA, M. & YAMANAKA, S. 2005. Differential roles for Sox15 and Sox2 in transcriptional control in mouse embryonic stem cells. *J Biol Chem,* 280**,** 24371-9.

MASSARI, M. E. & MURRE, C. 2000. Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol Cell Biol,* 20**,** 429-40.

MASUI, S., NAKATAKE, Y., TOYOOKA, Y., SHIMOSATO, D., YAGI, R., TAKAHASHI, K., OKOCHI, H., OKUDA, A., MATOBA, R., SHAROV, A. A., KO, M. S. & NIWA, H. 2007. Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol,* 9**,** 625-35.

MASUI, S., OHTSUKA, S., YAGI, R., TAKAHASHI, K., KO, M. S. & NIWA, H. 2008. Rex1/Zfp42 is dispensable for pluripotency in mouse ES cells. *BMC Dev Biol,* 8**,** 45.

MATSUDA, T., NAKAMURA, T., NAKAO, K., ARAI, T., KATSUKI, M., HEIKE, T. & YOKOTA, T. 1999. STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. *EMBO J,* 18**,** 4261-9.

MEISSNER, A. 2010. Epigenetic modifications in pluripotent and differentiated cells. *Nat Biotechnol,* 28**,** 1079-88.

MEISSNER, A., GNIRKE, A., BELL, G. W., RAMSAHOYE, B., LANDER, E. S. & JAENISCH, R. 2005. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res,* 33**,** 5868-77.

MEISSNER, A., MIKKELSEN, T. S., GU, H., WERNIG, M., HANNA, J., SIVACHENKO, A., ZHANG, X., BERNSTEIN, B. E., NUSBAUM, C., JAFFE, D. B., GNIRKE, A., JAENISCH, R. & LANDER, E. S. 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature,* 454**,** 766-70.

MESHORER, E., YELLAJOSHULA, D., GEORGE, E., SCAMBLER, P. J., BROWN, D. T. & MISTELI, T. 2006. Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell,* 10**,** 105-16.

MIKKELSEN, T. S., KU, M., JAFFE, D. B., ISSAC, B., LIEBERMAN, E., GIANNOUKOS, G., ALVAREZ, P., BROCKMAN, W., KIM, T. K., KOCHE, R. P., LEE, W., MENDENHALL, E., O'DONOVAN, A., PRESSER, A., RUSS, C., XIE, X., MEISSNER, A., WERNIG, M., JAENISCH, R., NUSBAUM, C., LANDER, E. S. & BERNSTEIN, B. E. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature,* 448**,** 553-60.

MITSUI, K., TOKUZAWA, Y., ITOH, H., SEGAWA, K., MURAKAMI, M., TAKAHASHI, K., MARUYAMA, M., MAEDA, M. & YAMANAKA, S. 2003. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell,* 113**,** 631-42.

MORI, M., NAKAGAMI, H., KOIBUCHI, N., MIURA, K., TAKAMI, Y., KORIYAMA, H., HAYASHI, H., SABE, H., MOCHIZUKI, N., MORISHITA, R. & KANEDA, Y. 2009. Zyxin mediates actin fiber reorganization in epithelial-mesenchymal transition and contributes to endocardial morphogenesis. *Mol Biol Cell,* 20**,** 3115-24.

MULLIN, N. P., YATES, A., ROWE, A. J., NIJMEIJER, B., COLBY, D., BARLOW, P. N., WALKINSHAW, M. D. & CHAMBERS, I. 2008. The pluripotency rheostat Nanog functions as a dimer. *Biochem J,* 411**,** 227-31.

MURRE, C., BAIN, G., VAN DIJK, M. A., ENGEL, I., FURNARI, B. A., MASSARI, M. E., MATTHEWS, J. R., QUONG, M. W., RIVERA, R. R. & STUIVER, M. H. 1994. Structure and function of helix-loop-helix proteins. *Biochim Biophys Acta,* 1218**,** 129-35.

NAJM, F. J., ZAREMBA, A., CAPRARIELLO, A. V., NAYAK, S., FREUNDT, E. C., SCACHERI, P. C., MILLER, R. H. & TESAR, P. J. 2011. Rapid and robust generation of functional oligodendrocyte progenitor cells from epiblast stem cells. *Nat Methods,* 8**,** 957-62.

NAKAGAWA, M., KOYANAGI, M., TANABE, K., TAKAHASHI, K., ICHISAKA, T., AOI, T., OKITA, K., MOCHIDUKI, Y., TAKIZAWA, N. & YAMANAKA, S. 2008. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat Biotechnol,* 26**,** 101-6.

NAKANO, T., KODAMA, H. & HONJO, T. 1994. Generation of lymphohematopoietic cells from embryonic stem cells in culture. *Science,* 265**,** 1098-101.

NAVARRO, P., FESTUCCIA, N., COLBY, D., GAGLIARDI, A., MULLIN, N.

P., ZHANG, W., KARWACKI-NEISIUS, V., OSORNO, R., KELLY, D., ROBERTSON, M. & CHAMBERS, I. 2012. OCT4/SOX2-independent Nanog autorepression modulates heterogeneous Nanog gene expression in mouse ES cells. *EMBO J,* 31**,** 4547-62.

NAVARRO, P., OLDFIELD, A., LEGOUPI, J., FESTUCCIA, N., DUBOIS, A., ATTIA, M., SCHOORLEMMER, J., ROUGEULLE, C., CHAMBERS, I. & AVNER, P. 2010. Molecular coupling of Tsix regulation and pluripotency. *Nature,* 468**,** 457-60.

NEUHOLD, L. A. & WOLD, B. 1993. HLH forced dimers: tethering MyoD to E47 generates a dominant positive myogenic factor insulated from negative regulation by Id. *Cell,* 74**,** 1033-42.

NICHOLS, J., JONES, K., PHILLIPS, J. M., NEWLAND, S. A., ROODE, M., MANSFIELD, W., SMITH, A. & COOKE, A. 2009a. Validated germline-competent embryonic stem cell lines from nonobese diabetic mice. *Nat Med,* 15**,** 814-8.

NICHOLS, J., SILVA, J., ROODE, M. & SMITH, A. 2009b. Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. *Development,* 136**,** 3215-22.

NICHOLS, J. & SMITH, A. 2009. Naive and primed pluripotent states. *Cell Stem Cell,* 4**,** 487-92.

NICHOLS, J. & SMITH, A. 2012. Pluripotency in the embryo and in culture. *Cold Spring Harb Perspect Biol,* 4**,** a008128.

NICHOLS, J., ZEVNIK, B., ANASTASSIADIS, K., NIWA, H., KLEWE-NEBENIUS, D., CHAMBERS, I., SCHOLER, H. & SMITH, A. 1998. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell,* 95**,** 379-91.

NISHIKAWA, S. I., NISHIKAWA, S., HIRASHIMA, M., MATSUYOSHI, N. & 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**,** 1747-57.

NISHIMOTO, M., FUKUSHIMA, A., OKUDA, A. & MURAMATSU, M. 1999. The gene for the embryonic stem cell coactivator UTF1 carries a regulatory element which selectively interacts with a complex composed of Oct-3/4 and Sox-2. *Mol Cell Biol,* 19**,** 5453-65.

NIWA, H., BURDON, T., CHAMBERS, I. & SMITH, A. 1998. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev,* 12**,** 2048-60.

NIWA, H., MIYAZAKI, J. & SMITH, A. G. 2000. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet,* 24**,** 372-6.

NIWA, H., OGAWA, K., SHIMOSATO, D. & ADACHI, K. 2009. A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature,* 460**,** 118-22.

NIWA, H., TOYOOKA, Y., SHIMOSATO, D., STRUMPF, D., TAKAHASHI, K., YAGI, R. & ROSSANT, J. 2005. Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell,* 123**,** 917-29.

NORTON, J. D. 2000. ID helix-loop-helix proteins in cell growth, differentiation and tumorigenesis. *J Cell Sci,* 113 (Pt 22)**,** 3897-905.

OGAWA, K., NISHINAKAMURA, R., IWAMATSU, Y., SHIMOSATO, D. & NIWA, H. 2006. Synergistic action of Wnt and LIF in maintaining pluripotency of mouse ES cells. *Biochem Biophys Res Commun,* 343**,** 159-66.

OKAMOTO, I., OTTE, A. P., ALLIS, C. D., REINBERG, D. & HEARD, E. 2004. Epigenetic dynamics of imprinted X inactivation during early mouse development. *Science,* 303**,** 644-9.

OKANO, M., BELL, D. W., HABER, D. A. & LI, E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell,* 99**,** 247-57.

OKITA, K., ICHISAKA, T. & YAMANAKA, S. 2007. Generation of germline-competent induced pluripotent stem cells. *Nature,* 448**,** 313-7.

OKITA, K. & YAMANAKA, S. 2011. Induced pluripotent stem cells: opportunities and challenges. *Philos Trans R Soc Lond B Biol Sci,* 366**,** 2198-207.

OKUMURA-NAKANISHI, S., SAITO, M., NIWA, H. & ISHIKAWA, F. 2005. Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. *J Biol Chem,* 280**,** 5307-17.

OSORNO, R., TSAKIRIDIS, A., WONG, F., CAMBRAY, N., ECONOMOU, C., WILKIE, R., BLIN, G., SCOTTING, P. J., CHAMBERS, I. & WILSON, V. 2012. The developmental dismantling of pluripotency is reversed by ectopic Oct4 expression. *Development,* 139**,** 2288-98.

PALMQVIST, L., GLOVER, C. H., HSU, L., LU, M., BOSSEN, B., PIRET, J. M., HUMPHRIES, R. K. & HELGASON, C. D. 2005. Correlation of murine embryonic stem cell gene expression profiles with functional measures of pluripotency. *Stem Cells,* 23**,** 663-80.

PAN, G. & THOMSON, J. A. 2007. Nanog and transcriptional networks in embryonic stem cell pluripotency. *Cell Res,* 17**,** 42-9.

PARISI, S., PASSARO, F., ALOIA, L., MANABE, I., NAGAI, R., PASTORE, L. & RUSSO, T. 2008. Klf5 is involved in self-renewal of mouse embryonic stem cells. *J Cell Sci,* 121**,** 2629-34.

PARK, I. H., ZHAO, R., WEST, J. A., YABUUCHI, A., HUO, H., INCE, T. A., LEROU, P. H., LENSCH, M. W. & DALEY, G. Q. 2008. Reprogramming of human somatic cells to pluripotency with defined factors. *Nature,* 451**,** 141-6.

PELTON, T. A., SHARMA, S., SCHULZ, T. C., RATHJEN, J. & RATHJEN, P. D. 2002. Transient pluripotent cell populations during primitive ectoderm formation: correlation of in vivo and in vitro pluripotent cell development. *J Cell Sci,* 115**,** 329-39.

PESCE, S. & BENEZRA, R. 1993. The loop region of the helix-loop-helix protein Id1 is critical for its dominant negative activity. *Mol Cell Biol,* 13**,** 7874-80.

PLUSA, B., PILISZEK, A., FRANKENBERG, S., ARTUS, J. & HADJANTONAKIS, A. K. 2008. Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. *Development,* 135**,** 3081-91.

POLLARD, S. M., BENCHOUA, A. & LOWELL, S. 2006. Neural stem cells, neurons, and glia. *Methods Enzymol,* 418**,** 151-69.

PORCHER, C., SWAT, W., ROCKWELL, K., FUJIWARA, Y., ALT, F. W. & ORKIN, S. H. 1996. The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages. *Cell,* 86**,** 47-57.

RADZISHEUSKAYA, A., CHIA GLE, B., DOS SANTOS, R. L., THEUNISSEN, T. W., CASTRO, L. F., NICHOLS, J. & SILVA, J. C. 2013. A defined Oct4 level governs cell state transitions of pluripotency entry and differentiation into all embryonic lineages. *Nat Cell Biol,* 15**,** 579-90.

RASTAN, S. 1982. Timing of X-chromosome inactivation in postimplantation mouse embryos. *J Embryol Exp Morphol,* 71**,** 11-24.

RATHJEN, J., LAKE, J. A., BETTESS, M. D., WASHINGTON, J. M., CHAPMAN, G. & RATHJEN, P. D. 1999. Formation of a primitive ectoderm like cell population, EPL cells, from ES cells in response to biologically derived factors. *J Cell Sci,* 112 (Pt 5)**,** 601-12.

RIECHMANN, V. & SABLITZKY, F. 1995. Mutually exclusive expression of two dominant-negative helix-loop-helix (dnHLH) genes, Id4 and Id3,

in the developing brain of the mouse suggests distinct regulatory roles of these dnHLH proteins during cellular proliferation and differentiation of the nervous system. *Cell Growth Differ,* 6**,** 837-43.

RIECHMANN, V., VAN CRUCHTEN, I. & SABLITZKY, F. 1994. The expression pattern of Id4, a novel dominant negative helix-loop-helix protein, is distinct from Id1, Id2 and Id3. *Nucleic Acids Res,* 22**,** 749-55.

RIEHL, R., JOHNSON, K., BRADLEY, R., GRUNWALD, G. B., CORNEL, E., LILIENBAUM, A. & HOLT, C. E. 1996. Cadherin function is required for axon outgrowth in retinal ganglion cells in vivo. *Neuron,* 17**,** 837-48.

ROBERTSON, E., BRADLEY, A., KUEHN, M. & EVANS, M. 1986. Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. *Nature,* 323**,** 445-8.

ROGERS, M. B., HOSLER, B. A. & GUDAS, L. J. 1991. Specific expression of a retinoic acid-regulated, zinc-finger gene, Rex-1, in preimplantation embryos, trophoblast and spermatocytes. *Development,* 113**,** 815-24.

ROIG, M. B., ROSET, R., ORTET, L., BALSIGER, N. A., ANFOSSO, A., CABELLOS, L., GARRIDO, M., ALAMEDA, F., BRADY, H. J. & GIL-GOMEZ, G. 2009. Identification of a novel cyclin required for the intrinsic apoptosis pathway in lymphoid cells. *Cell Death Differ,* 16**,** 230-43.

ROMERO-LANMAN, E. E., PAVLOVIC, S., AMLANI, B., CHIN, Y. & BENEZRA, R. 2012. Id1 maintains embryonic stem cell self-renewal by up-regulation of Nanog and repression of Brachyury expression. *Stem Cells Dev,* 21**,** 384-93.

ROSENTHAL, M. D., WISHNOW, R. M. & SATO, G. H. 1970. In vitro growth and differetiation of clonal populations of multipotential mouse clls derived from a transplantable testicular teratocarcinoma. *J Natl Cancer Inst,* 44**,** 1001-14.

ROSSANT, J. & TAM, P. P. 2009. Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. *Development,* 136**,** 701-13.

RUGG-GUNN, P. J., COX, B. J., LANNER, F., SHARMA, P., IGNATCHENKO, V., MCDONALD, A. C., GARNER, J., GRAMOLINI, A. O., ROSSANT, J. & KISLINGER, T. 2012. Cell-surface proteomics identifies lineage-specific markers of embryo-derived stem cells. *Dev Cell,* 22**,** 887-901.

SASSOON, D., LYONS, G., WRIGHT, W. E., LIN, V., LASSAR, A.,

WEINTRAUB, H. & BUCKINGHAM, M. 1989. Expression of two myogenic regulatory factors myogenin and MyoD1 during mouse embryogenesis. *Nature,* 341**,** 303-7.

SATO, N., MEIJER, L., SKALTSOUNIS, L., GREENGARD, P. & BRIVANLOU, A. H. 2004. Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nat Med,* 10**,** 55-63.

SCHUBELER, D., MACALPINE, D. M., SCALZO, D., WIRBELAUER, C., KOOPERBERG, C., VAN LEEUWEN, F., GOTTSCHLING, D. E., O'NEILL, L. P., TURNER, B. M., DELROW, J., BELL, S. P. & GROUDINE, M. 2004. The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev,* 18**,** 1263-71.

SCHUH, R., AICHER, W., GAUL, U., COTE, S., PREISS, A., MAIER, D., SEIFERT, E., NAUBER, U., SCHRODER, C., KEMLER, R. & ET AL. 1986. A conserved family of nuclear proteins containing structural elements of the finger protein encoded by Kruppel, a Drosophila segmentation gene. *Cell,* 47**,** 1025-32.

SHCHERBO, D., MURPHY, C. S., ERMAKOVA, G. V., SOLOVIEVA, E. A., CHEPURNYKH, T. V., SHCHEGLOV, A. S., VERKHUSHA, V. V., PLETNEV, V. Z., HAZELWOOD, K. L., ROCHE, P. M., LUKYANOV, S., ZARAISKY, A. G., DAVIDSON, M. W. & CHUDAKOV, D. M. 2009. Far-red fluorescent tags for protein imaging in living tissues. *Biochem J,* 418**,** 567-74.

SHI, W., WANG, H., PAN, G., GENG, Y., GUO, Y. & PEI, D. 2006. Regulation of the pluripotency marker Rex-1 by Nanog and Sox2. *J Biol Chem,* 281**,** 23319-25.

SHIMOSATO, D., SHIKI, M. & NIWA, H. 2007. Extra-embryonic endoderm cells derived from ES cells induced by GATA factors acquire the character of XEN cells. *BMC Dev Biol,* 7**,** 80.

SILVA, J., NICHOLS, J., THEUNISSEN, T. W., GUO, G., VAN OOSTEN, A. L., BARRANDON, O., WRAY, J., YAMANAKA, S., CHAMBERS, I. & SMITH, A. 2009. Nanog is the gateway to the pluripotent ground state. *Cell,* 138**,** 722-37.

SILVA, J. & SMITH, A. 2008. Capturing pluripotency. *Cell,* 132**,** 532-6.

SINGLA, D. K., SCHNEIDER, D. J., LEWINTER, M. M. & SOBEL, B. E. 2006. wnt3a but not wnt11 supports self-renewal of embryonic stem cells.

Biochem Biophys Res Commun, 345**,** 789-95.
SKINNER, M. K., RAWLS, A., WILSON-RAWLS, J. & ROALSON, E. H. 2010. Basic helix-loop-helix transcription factor gene family phylogenetics and nomenclature. *Differentiation,* 80**,** 1-8.
SMITH, A. G., HEATH, J. K., DONALDSON, D. D., WONG, G. G., MOREAU, J., STAHL, M. & ROGERS, D. 1988. Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature,* 336**,** 688-90.
SMITH, A. G. & HOOPER, M. L. 1987. Buffalo rat liver cells produce a diffusible activity which inhibits the differentiation of murine embryonal carcinoma and embryonic stem cells. *Dev Biol,* 121**,** 1-9.
SMYTH, G. K., MICHAUD, J. & SCOTT, H. S. 2005. Use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics,* 21**,** 2067-75.
SOLTER, D., SKREB, N. & DAMJANOV, I. 1970. Extrauterine growth of mouse egg-cylinders results in malignant teratoma. *Nature,* 227**,** 503-4.
SOUTHERN, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol,* 98**,** 503-17.
STAHL, N., BOULTON, T. G., FARRUGGELLA, T., IP, N. Y., DAVIS, S., WITTHUHN, B. A., QUELLE, F. W., SILVENNOINEN, O., BARBIERI, G., PELLEGRINI, S. & ET AL. 1994. Association and activation of Jak-Tyk kinases by CNTF-LIF-OSM-IL-6 beta receptor components. *Science,* 263**,** 92-5.
STEPHENS, L. E., SUTHERLAND, A. E., KLIMANSKAYA, I. V., ANDRIEUX, A., MENESES, J., PEDERSEN, R. A. & DAMSKY, C. H. 1995. Deletion of beta 1 integrins in mice results in inner cell mass failure and peri-implantation lethality. *Genes Dev,* 9**,** 1883-95.
STEVENS, L. C. 1970. The development of transplantable teratocarcinomas from intratesticular grafts of pre- and postimplantation mouse embryos. *Dev Biol,* 21**,** 364-82.
STEVENS, L. C. & LITTLE, C. C. 1954. Spontaneous Testicular Teratomas in an Inbred Strain of Mice. *Proc Natl Acad Sci U S A,* 40**,** 1080-7.
STEWART, C. L., KASPAR, P., BRUNET, L. J., BHATT, H., GADI, I., KONTGEN, F. & ABBONDANZO, S. J. 1992. Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. *Nature,* 359**,** 76-9.
SUN, X. H., COPELAND, N. G., JENKINS, N. A. & BALTIMORE, D. 1991. Id

proteins Id1 and Id2 selectively inhibit DNA binding by one class of helix-loop-helix proteins. *Mol Cell Biol,* 11**,** 5603-11.

SURANI, M. A., HAYASHI, K. & HAJKOVA, P. 2007. Genetic and epigenetic regulators of pluripotency. *Cell,* 128**,** 747-62.

SUTHERLAND, A. E., CALARCO, P. G. & DAMSKY, C. H. 1993. Developmental regulation of integrin expression at the time of implantation in the mouse embryo. *Development,* 119**,** 1175-86.

TAHILIANI, M., KOH, K. P., SHEN, Y., PASTOR, W. A., BANDUKWALA, H., BRUDNO, Y., AGARWAL, S., IYER, L. M., LIU, D. R., ARAVIND, L. & RAO, A. 2009. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science,* 324**,** 930-5.

TAKAHASHI, K., TANABE, K., OHNUKI, M., NARITA, M., ICHISAKA, T., TOMODA, K. & YAMANAKA, S. 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell,* 131**,** 861-72.

TAKAHASHI, K. & YAMANAKA, S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell,* 126**,** 663-76.

TAKAHASHI, Y., SATO, Y., SUETSUGU, R. & NAKAYA, Y. 2005. Mesenchymal-to-epithelial transition during somitic segmentation: a novel approach to studying the roles of Rho family GTPases in morphogenesis. *Cells Tissues Organs,* 179**,** 36-42.

TANAKA, T. S., DAVEY, R. E., LAN, Q., ZANDSTRA, P. W. & STANFORD, W. L. 2008. Development of a gene-trap vector with a highly sensitive fluorescent protein reporter system for expression profiling. *Genesis,* 46**,** 347-56.

TANG, F., BARBACIORU, C., BAO, S., LEE, C., NORDMAN, E., WANG, X., LAO, K. & SURANI, M. A. 2010. Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis. *Cell Stem Cell,* 6**,** 468-78.

TEMPLETON, N. S., ROBERTS, D. D. & SAFER, B. 1997. Efficient gene targeting in mouse embryonic stem cells. *Gene Ther,* 4**,** 700-9.

TESAR, P. J., CHENOWETH, J. G., BROOK, F. A., DAVIES, T. J., EVANS, E. P., MACK, D. L., GARDNER, R. L. & MCKAY, R. D. 2007. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature,* 448**,** 196-9.

THOMAS, K. R. & CAPECCHI, M. R. 1986. Introduction of homologous

DNA sequences into mammalian cells induces mutations in the cognate gene. *Nature,* 324**,** 34-8.

TOMIOKA, M., NISHIMOTO, M., MIYAGI, S., KATAYANAGI, T., FUKUI, N., NIWA, H., MURAMATSU, M. & OKUDA, A. 2002. Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex. *Nucleic Acids Res,* 30**,** 3202-13.

TORRES-PADILLA, M. E. & CHAMBERS, I. 2014. Transcription factor heterogeneity in pluripotent stem cells: a stochastic advantage. *Development,* 141**,** 2173-81.

TOYOOKA, Y., SHIMOSATO, D., MURAKAMI, K., TAKAHASHI, K. & NIWA, H. 2008. Identification and characterization of subpopulations in undifferentiated ES cell culture. *Development,* 135**,** 909-18.

TROPEPE, V., HITOSHI, S., SIRARD, C., MAK, T. W., ROSSANT, J. & VAN DER KOOY, D. 2001. Direct neural fate specification from embryonic stem cells: a primitive mammalian neural stem cell stage acquired through a default mechanism. *Neuron,* 30**,** 65-78.

TSAI, S. Y., BOUWMAN, B. A., ANG, Y. S., KIM, S. J., LEE, D. F., LEMISCHKA, I. R. & RENDL, M. 2011. Single transcription factor reprogramming of hair follicle dermal papilla cells to induced pluripotent stem cells. *Stem Cells,* 29**,** 964-71.

TSUMURA, A., HAYAKAWA, T., KUMAKI, Y., TAKEBAYASHI, S., SAKAUE, M., MATSUOKA, C., SHIMOTOHNO, K., ISHIKAWA, F., LI, E., UEDA, H. R., NAKAYAMA, J. & OKANO, M. 2006. Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes Cells,* 11**,** 805-14.

VAN DEN BERG, D. L., ZHANG, W., YATES, A., ENGELEN, E., TAKACS, K., BEZSTAROSTI, K., DEMMERS, J., CHAMBERS, I. & POOT, R. A. 2008. Estrogen-related receptor beta interacts with Oct4 to positively regulate Nanog gene expression. *Mol Cell Biol,* 28**,** 5986-95.

VON BUBNOFF, A. & CHO, K. W. 2001. Intracellular BMP signaling regulation in vertebrates: pathway or network? *Dev Biol,* 239**,** 1-14.

WAGNER, K. D., WAGNER, N., VIDAL, V. P., SCHLEY, G., WILHELM, D., SCHEDL, A., ENGLERT, C. & SCHOLZ, H. 2002. The Wilms' tumor gene Wt1 is required for normal development of the retina. *EMBO J,* 21**,** 1398-405.

WANG, J., LEVASSEUR, D. N. & ORKIN, S. H. 2008. Requirement of Nanog

dimerization for stem cell self-renewal and pluripotency. *Proc Natl Acad Sci U S A,* 105**,** 6326-31.

WANG, J., RAO, S., CHU, J., SHEN, X., LEVASSEUR, D. N., THEUNISSEN, T. W. & ORKIN, S. H. 2006. A protein interaction network for pluripotency of embryonic stem cells. *Nature,* 444**,** 364-8.

WANG, S. H., TSAI, M. S., CHIANG, M. F. & LI, H. 2003. A novel NK-type homeobox gene, ENK (early embryo specific NK), preferentially expressed in embryonic stem cells. *Gene Expr Patterns,* 3**,** 99-103.

WANG, X. & YANG, P. 2008. In vitro differentiation of mouse embryonic stem (mES) cells using the hanging drop method. *J Vis Exp*.

WERNIG, M., MEISSNER, A., FOREMAN, R., BRAMBRINK, T., KU, M., HOCHEDLINGER, K., BERNSTEIN, B. E. & JAENISCH, R. 2007. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature,* 448**,** 318-24.

WHELAN, J., CORDLE, S. R., HENDERSON, E., WEIL, P. A. & STEIN, R. 1990. Identification of a pancreatic beta-cell insulin gene transcription factor that binds to and appears to activate cell-type-specific expression: its possible relationship to other cellular factors that bind to a common insulin gene sequence. *Mol Cell Biol,* 10**,** 1564-72.

WHYTE, W. A., ORLANDO, D. A., HNISZ, D., ABRAHAM, B. J., LIN, C. Y., KAGEY, M. H., RAHL, P. B., LEE, T. I. & YOUNG, R. A. 2013. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell,* 153**,** 307-19.

WILDER, P. J., KELLY, D., BRIGMAN, K., PETERSON, C. L., NOWLING, T., GAO, Q. S., MCCOMB, R. D., CAPECCHI, M. R. & RIZZINO, A. 1997. Inactivation of the FGF-4 gene in embryonic stem cells alters the growth and/or the survival of their early differentiated progeny. *Dev Biol,* 192**,** 614-29.

WILLIAMS, R. L., HILTON, D. J., PEASE, S., WILLSON, T. A., STEWART, C. L., GEARING, D. P., WAGNER, E. F., METCALF, D., NICOLA, N. A. & GOUGH, N. M. 1988. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature,* 336**,** 684-7.

WILSON-RAWLS, J., RHEE, J. M. & RAWLS, A. 2004. Paraxis is a basic helix-loop-helix protein that positively regulates transcription through binding to specific E-box elements. *J Biol Chem,* 279**,** 37685-92.

WOOD, H. B. & EPISKOPOU, V. 1999. Comparative expression of the mouse

Sox1, Sox2 and Sox3 genes from pre-gastrulation to early somite stages. *Mech Dev,* 86**,** 197-201.

WRAY, J., KALKAN, T., GOMEZ-LOPEZ, S., ECKARDT, D., COOK, A., KEMLER, R. & SMITH, A. 2011. Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nat Cell Biol,* 13**,** 838-45.

WRAY, J., KALKAN, T. & SMITH, A. G. 2010. The ground state of pluripotency. *Biochem Soc Trans,* 38**,** 1027-32.

WU, D. & PAN, W. 2010. GSK3: a multifaceted kinase in Wnt signaling. *Trends Biochem Sci,* 35**,** 161-8.

WU, T., WANG, H., HE, J., KANG, L., JIANG, Y., LIU, J., ZHANG, Y., KOU, Z., LIU, L., ZHANG, X. & GAO, S. 2011. Reprogramming of trophoblast stem cells into pluripotent stem cells by Oct4. *Stem Cells,* 29**,** 755-63.

XIAO, Z., PATRAKKA, J., NUKUI, M., CHI, L., NIU, D., BETSHOLTZ, C., PIKKARAINEN, T., VAINIO, S. & TRYGGVASON, K. 2011. Deficiency in Crumbs homolog 2 (Crb2) affects gastrulation and results in embryonic lethality in mice. *Dev Dyn,* 240**,** 2646-56.

YAMANAKA, Y., LANNER, F. & ROSSANT, J. 2010. FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. *Development,* 137**,** 715-24.

YAN, W., YOUNG, A. Z., SOARES, V. C., KELLEY, R., BENEZRA, R. & ZHUANG, Y. 1997. High incidence of T-cell tumors in E2A-null mice and E2A/Id1 double-knockout mice. *Mol Cell Biol,* 17**,** 7317-27.

YANG, S. H., KALKAN, T., MORISSROE, C., MARKS, H., STUNNENBERG, H., SMITH, A. & SHARROCKS, A. D. 2014. Otx2 and Oct4 Drive Early Enhancer Activation during Embryonic Stem Cell Transition from Naive Pluripotency. *Cell Rep,* 7**,** 1968-81.

YEAP, L. S., HAYASHI, K. & SURANI, M. A. 2009. ERG-associated protein with SET domain (ESET)-Oct4 interaction regulates pluripotency and represses the trophectoderm lineage. *Epigenetics Chromatin,* 2**,** 12.

YEO, J. C., JIANG, J., TAN, Z. Y., YIM, G. R., NG, J. H., GOKE, J., KRAUS, P., LIANG, H., GONZALES, K. A., CHONG, H. C., TAN, C. P., LIM, Y. S., TAN, N. S., LUFKIN, T. & NG, H. H. 2014. Klf2 is an essential factor that sustains ground state pluripotency. *Cell Stem Cell,* 14**,** 864-72.

YEOM, Y. I., FUHRMANN, G., OVITT, C. E., BREHM, A., OHBO, K., GROSS,

M., HUBNER, K. & SCHOLER, H. R. 1996. Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development,* 122**,** 881-94.

YILDIRIM, O., LI, R., HUNG, J. H., CHEN, P. B., DONG, X., EE, L. S., WENG, Z., RANDO, O. J. & FAZZIO, T. G. 2011. Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell,* 147**,** 1498-510.

YING, Q. L., NICHOLS, J., CHAMBERS, I. & SMITH, A. 2003. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell,* 115**,** 281-92.

YING, Q. L., NICHOLS, J., EVANS, E. P. & SMITH, A. G. 2002. Changing potency by spontaneous fusion. *Nature,* 416**,** 545-8.

YING, Q. L. & SMITH, A. G. 2003. Defined conditions for neural commitment and differentiation. *Methods Enzymol,* 365**,** 327-41.

YING, Q. L., WRAY, J., NICHOLS, J., BATLLE-MORERA, L., DOBLE, B., WOODGETT, J., COHEN, P. & SMITH, A. 2008. The ground state of embryonic stem cell self-renewal. *Nature,* 453**,** 519-23.

YUAN, H., CORBI, N., BASILICO, C. & DAILEY, L. 1995. Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Genes Dev,* 9**,** 2635-45.

YUAN, X., WAN, H., ZHAO, X., ZHU, S., ZHOU, Q. & DING, S. 2011. Brief report: combined chemical treatment enables Oct4-induced reprogramming from mouse embryonic fibroblasts. *Stem Cells,* 29**,** 549-53.

ZHANG, K., LI, L., HUANG, C., SHEN, C., TAN, F., XIA, C., LIU, P., ROSSANT, J. & JING, N. 2010. Distinct functions of BMP4 during different stages of mouse ES cell neural commitment. *Development,* 137**,** 2095-105.

ZHANG, X., ZHANG, J., WANG, T., ESTEBAN, M. A. & PEI, D. 2008. Esrrb activates Oct4 transcription and sustains self-renewal and pluripotency in embryonic stem cells. *J Biol Chem,* 283**,** 35825-33.

ZHAO, X. D., HAN, X., CHEW, J. L., LIU, J., CHIU, K. P., CHOO, A., ORLOV, Y. L., SUNG, W. K., SHAHAB, A., KUZNETSOV, V. A., BOURQUE, G., OH, S., RUAN, Y., NG, H. H. & WEI, C. L. 2007. Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. *Cell Stem Cell,* 1**,** 286-98.

ZHOU, X., SMITH, A. J., WATERHOUSE, A., BLIN, G., MALAGUTI, M., LIN, C. Y., OSORNO, R., CHAMBERS, I. & LOWELL, S. 2013. Hes1 desynchronizes differentiation of pluripotent cells by modulating STAT3 activity. *Stem Cells,* 31**,** 1511-22.