TURUN YLIOPISTON JULKAISUJA
ANNALES UNIVERSITATIS TURKUENSIS

SARJA - SER. D OSA - TOM. 1035

MEDICA - ODONTOLOGICA

PLURIPOTENCY AND GENETIC STABILITY OF HUMAN PLURIPOTENT STEM CELLS

by

Elisa Närvä

TURUN YLIOPISTO
UNIVERSITY OF TURKU
Turku 2012

Department of Medical Microbiology and Immunology, University of Turku;
Turku Centre for Biotechnology;
University of Turku and Åbo Akademi University;
Turku Doctoral Programme of Biomedical Sciences, University of Turku

Supervised by

Professor Riitta Lahesmaa
Turku Centre for Biotechnology
University of Turku and Åbo Akademi University
Turku, Finland

Reviewed by

Professor Tapio Visakorpi
Institute of Biomedical Technology
University of Tampere
Tampere, Finland

and

Docent Katriina Aalto-Setälä
Institute of Biomedical Technology
University of Tampere
Tampere, Finland

Opponent

Professor Christine Mummery
Department of Anatomy & Embryology
Leiden University Medical Center
Leiden, The Netherlands

ISBN 978-951-29-5136-9 (PRINT)
ISBN 978-951-29-5137-6 (PDF)
ISSN 0355-9483
Painosalama Oy – Turku, Finland 2012

The secret of eternal youth is arrested development
Alice Roosevelt Longworth

ABSTRACT

Elisa Närvä
Pluripotency and genetic stability of human pluripotent stem cells

Department of Medical Microbiology and Immunology, University of Turku
Turku Centre for Biotechnology,
University of Turku and Åbo Akademi University
Turku Doctoral Programme of Biomedical Sciences
Annales Universitatis Turkuensis, Turku, 2012

Pluripotent cells have the potential to differentiate into all somatic cell types. As the adult human body is unable to regenerate various tissues, pluripotent cells provide an attractive source for regenerative medicine. Human embryonic stem cells (hESCs) can be isolated from blastocyst stage embryos and cultured in the laboratory environment. However, their use in regenerative medicine is restricted due to problems with immunosuppression by the host and ethical legislation. Recently, a new source of pluripotent cells was established via the direct reprogramming of somatic cells. These human induced pluripotent stem cells (hiPSCs) enable the production of patient specific cell types. However, numerous challenges, such as efficient reprogramming, optimal culture, directed differentiation, genetic stability and tumor risk need to be solved before the launch of therapeutic applications.

The main objective of this thesis was to understand the unique properties of human pluripotent stem cells. The specific aims were to identify novel factors involved in maintaining pluripotency, characterize the effects of low oxygen culture on hESCs, and determine the high resolution changes in hESCs and hiPSCs during culture and reprogramming. As a result, the previously uncharacterized protein L1TD1 was determined to be specific for pluripotent cells and essential for the maintenance of pluripotency. The low oxygen culture supported undifferentiated growth and affected expression of stem cell associated transcripts. High resolution screening of hESCs identified a number of culture induced copy number variations and loss of heterozygosity changes. Further, screening of hiPSCs revealed that reprogramming induces high resolution alterations. The results obtained in this thesis have important implications for stem cell and cancer biology and the therapeutic potential of pluripotent cells.

Key words: human embryonic stem cell, human induced pluripotent stem cell, pluripotency, L1TD1, hypoxia, genetic stability, copy number variation

TIIVISTELMÄ

Elisa Närvä
Ihmisen pluripotenttien kantasolujen geneettinen stabiilisuus ja pluripotenssin säätely

Lääketieteellinen mikrobiologia ja immunologia, Turun Yliopisto
Turun Biotekniikan keskus, Turun Yliopisto ja Åbo Akademi
Turun Biolääketieteellinen tutkijakoulu
Turun Yliopiston Julkaisuja, Turku, 2012

Pluripotenttilla solulla on kyky erilaistua ihmiskehon jokaiseksi solutyypiksi. Erilaistuttuaan useat solutyypit menettävät kykynsä uusiutua, jonka vuoksi useat elimistömme kudoksista ovat korvaamattomia. Pluripotentteja kantasoluja voidaan eristää varhaisista alkioista, viljellä erilaistumattomana ja erilaistaa laboratoriolosuhteissa, mahdollistaen solukorvaushoidon useaan parantumattomaan sairauteen. Alkiosta eristettyjen solujen käyttö on kuitenkin eettisesti ongelmallista ja voi herättää kudoshyljintäreaktion. Nykyisin pluripotentteja kantasoluja voidaan indusoida myös suoraan somaattisista soluista, mahdollistaen potilasspesifisten solutyyppien regeneraation. Ennen terapeuttisia sovelluksia, pluripotenttien kantasolujen indusointi, viljely, erilaistaminen, geneettinen stabiilisuus ja mahdollinen syöpäriski vaativat kuitenkin intensiivistä tutkimustyötä.

Tämä väitöskirjatyö keskittyi pluripotenttien kantasolujen tutkimiseen. Tavoitteena oli tunnistaa uusia kantasoluspesifisiä tekijöitä, jotka ovat tärkeitä pluripotenssin säätelyssä, selvittää miten matalahappinen viljely vaikuttaa kantasolujen ominaisuuksiin ja tutkia miten laboratorioviljely ja pluripotenssin indusointi vaikuttavat kantasolujen geneettisiin ominaisuuksiin. Väitöskirjatyön tuloksena identifioimme uuden kantasoluspesifisen proteiinin, L1TD1:n joka osoittautui pluripotenssin ylläpidolle välttämättömäksi. Lisäksi, L1TD1 havaittiin ilmenevän kivessyövän kantasoluissa, joissa sen eliminointi vaikutti syöpäsolujen kasvuun. Matalahappinen viljely tuki kantasolujen erilaistumatonta kasvua ja aiheutti muutoksia kantasoluspesifisten geenien ilmentymisessä. Lisäksi, laboratorioviljelyn ja somaattisten solujen uudelleen ohjelmoinnin kantasoluiksi osoitettiin indusoivan kopiolukumuutoksia kantasolujen genomiin. Tässä väitöstutkimuksessa esitetyt tulokset lisäävät merkittävässä määrin tietoa pluripotenttien kantasolujen säätelystä, kasvatuksesta ja geneettisestä stabilisuudesta.

Avain sanat: ihmisen alkion kantasolu, pluripotenssi, L1TD1, geneettinen stabiilisuus, kopioluku muutos

CONTENTS

ABBREVIATIONS

ALP	alkaline phosphatase
APC	adeomatosis polyposis coli
ATP	adenosine triphosphate
BFB	breakage-fusion-bridge
bFGF	basic fibroblast growth factor
BMP	bone morphogenetic protein
CC	coiled coil
cDNA	complementary deoxyribonucleic acid
CFS	common fragile site
CGH	comparative genomic hybridization
ChiP	chromatin immunoprecipitation
CM	conditioned media
CN	copy number
CNV	copy number variation
CPP	cell-penetrating peptide
CTD	C-terminal domain
DOX	doxycycline
FC	fold change
FCS	fetal calf serum
FISH	fluorescence in situ hybridization
FoSTeS	fork stalling and template switching
GSK3β	glycogen synthase kinase-3β
hCG	chorionic gonadotropin
hESC	human embryonic stem cell
hiPSC	human induced pluripotent cell
H	histone
HFF	human foreskin fibroblast
HIF	hypoxia-inducible factor
HRE	hypoxia response element
HRP	horseradish peroxidase
IPTG	isopropyl-beta-D-thiogalactopyranoside
EB	embryoid body
EC	embryonal carcinoma
ECC	embryonal carcinoma cell
EGFP	enhanced green fluorescent protein
ESC	embryonic stem cell
FGF	fibroblast growth factor
FZD	Frizzled
ICM	inner cell mass
IVF	*in vitro* fertilization

K	lysine
LINE	long interspersed element-1
L1TD1	LINE 1 type transposase domain containing 1
LC	liquid chromatography
LOH	loss of heterozygosity
me^3	trimetylation
MEF	mouse embryonic fibroblast
mRNA	messenger ribonucleic acid
miRNA	microRNA
MS-MS	tandem mass spectrometry
NAHR	nonallelic homologous recombination
NHEJ	nonhomologous end-joining
ORF	open reading frame
P4	progesterone
PAGE	polyacrylamide gel electrophoresis
P-body	processing body
PCR	polymerase chain reaction
PHD	prolyl hydroxylase domain
pVHL	von Hippel-Lindau protein
RBP	RNA binding protein
RC	retrotransposition competent
RNA	ribonucleic acid
ROS	reactive oxygen species
RRM	RNA recognition motif
RT	real-time
SDS	sodium dodecyl sulfate
SG	stress granule
shRNA	small hairpin ribonucleic acid
siRNA	small interfering ribonucleic acid
SNP	single nucleotide polymorphisms
SSEA	stage-specific embryonic antigen
TGCT	testicular stem cell tumor
TGFβ	transforming growth factor β
TRA	keratin sulphate-related antigen
UTR	untranslated region
VPA	valproic acid

LIST OF ORGINAL PUBLICATIONS

I. Elisa Närvä*, Nelly Rahkonen*, Maheswara Reddy Emani*, Riikka Lund, Juha-Pekka Pursiheimo, Juuso Nästi, Reija Autio, Omid Rasool, Konstantin Denessiouk, Harri Lähdesmäki, Anjana Rao and Riitta Lahesmaa. *RNA Binding Protein L1TD1 Interacts with LIN28 via RNA and Is Required for Human Embryonic Stem Cell Self-renewal and Cancer Cell Proliferation.* STEM CELLS 2012 Mar;30(3):452-60.

II. Elisa Närvä*, Juha-Pekka Pursiheimo*, Asta Laiho, Kirsti Laurila, Roosa Sahla, Riikka Lund, Harri Lähdesmäki, Panu Jaakkola and Riitta Lahesmaa. *Hypoxia induced responses on human embryonic stem cells.* (Manuscript)

III. Elisa Närvä, Reija Autio, Nelly Rahkonen, Lingjia Kong, Neil Harrison, Danny Kitsberg, Lodovica Borghese, Joseph Itskovitz-Eldor, Omid Rasool, Petr Dvorak, Outi Hovatta, Timo Otonkoski, Timo Tuuri, Wei Cui, Oliver Brüstle, Duncan Baker, Edna Maltby, Harry D. Moore, Nissim Benvenisty, Peter W. Andrews, Olli Yli-Harja & Riitta Lahesmaa. *High-resolution DNA analysis of Human Embryonic Stem Cell lines reveals culture-induced copy number changes and loss of heterozygosity.* Nature Biotechnology 2010 Apr;28(4):371-7.

IV. Samer M. Hussein*, Nizar N. Batada*, Sanna Vuoristo, Reagan W. Ching, Reija Autio, Elisa Närvä, Siemon Ng, Michel Sourour, Riikka Hämäläinen, Cia Olsson, Karolina Lundin, Milla Mikkola, Ras Trokovic, Michael Peitz, Oliver Brüstle, David P. Bazett-Jones, Kari Alitalo, Riitta Lahesmaa, Andras Nagy & Timo Otonkoski. *Copy number variation and selection during reprogramming to pluripotency.* Nature 2011 Mar 3;471(7336):58-62.

V. Riikka Lund*, Elisa Närvä*, Riitta Lahesmaa. *Genetic and Epigenetic Stability of Human Pluripotent Stem Cells.* Nature Reviews Genetics 2012 (published online 11 September 2012).
* = equal contribution

1. INTRODUCTION

Stem cell research has advanced tremendously in recent years. The past decade has provided fundamental findings from the isolation of human embryonic stem cells (hESCs) to the induction of pluripotency in human somatic cells. Stem cells have a unique capacity to divide and renew themselves without any limit. The remarkable potential of pluripotent stem cells is their ability to develop into different cell types. The potential future applications of hESCs in regenerative medicine would provide a revolutionary new way to treat disease and injury, with wide-ranging medical benefits.

Pluripotent hESC lines can be derived from blastocyst stage embryos. The pluripotent inner cell mass cells of the blastocyst can be isolated and cultured on specific matrix in the presence of growth factors. Similarly, somatic cells can be reprogrammed into human induced pluripotent stem cells (hiPSCs), enabling construction of patient specific cell lines. *In vitro*-maintenance of human pluripotent cells in culture has allowed scientist to study the first events of human development and unique properties of self-renewing pluripotent cells. Although core factors regulating the pluripotent stage have been identified, there are large numbers of genes, proteins and small RNAs with unknown functions that are specifically expressed in pluripotent cell types, implying that the understanding of the regulation of pluripotency is only in the beginning.

Pluripotent forever young cell lines can be maintained in culture for extended periods as they do not face senescence. However, like any other cell line, pluripotent human lines can adapt to the culture and obtain oncogenic properties. Similarly, induced pluripotent cells can gain genomic changes during reprogramming and culture increasing the tumor risk of cells to be used in the regenerative medicine.

While scientists aim at solving how to guide differentiation in a controlled manner and solve substantial problems, such as tumor risk and production of regenerative therapy suitable patient specific cells, basic unique properties of stem cells has to be understood and studied. The objective of this thesis was to elucidate factors contributing to the regulation of pluripotency and characterize the genetic stability of human pluripotent cell lines. The aim was to find novel protein specific for pluripotent cells and identify its functional role. In addition, the effect of low oxygen culture conditions to the biology of hESCs was determined. Furthermore, high resolution genomic stability of hESCs and hiPSCs, in culture and reprogramming, has been addressed.

2. REVIEW OF THE LITERATURE

2.1 Pluripotency

2.1.1 Definition and discovery of pluripotent stem cells

The potency of a cell can be defined by its potential to differentiate. Human development starts when a sperm fertilizes an egg and creates a single totipotent cell, known as a zygote. The totipotent cell has the potential to differentiate into any cell type and form an entire organism. As a zygote divides, it develops into a 16-cell stage morula, which further differentiates into a structure called a blastocyst. Blastocyst consists of an outer layer of trophoblasts cells hosting the inner cell mass (ICM) cells in the hollow blastocyst cavity called the blastocoele. Trophoblast cells of the blastocyst form extraembryonic tissue, which is vital for supporting the development of the fetus. The cells of the ICM are pluripotent and maintain the ability to differentiate into all three germ layers: endoderm, mesoderm and ectoderm, and are the source cells of a new organism. **(Figure 1)** (Gilbert, 2003).

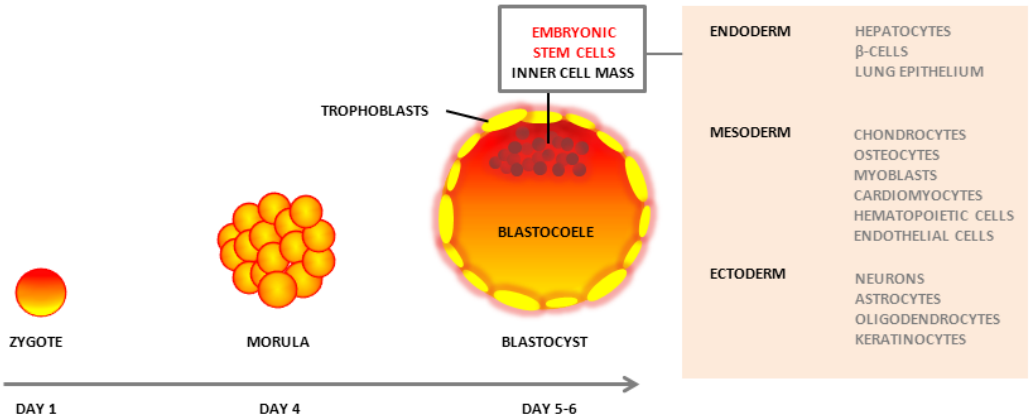

Figure 1. The beginning of the human development. Totipotent zygote is the first cell of the embryo. During the first days of development the zygote divides into the 16-cell stage morula, which further develops into the blastocyst stage embryo. The blastocyst consist of outer layer of trophoblasts cells hosting the inner cell mass (ICM) cells in the hollow blastocyst cavity, blastocoele. The ICM of the blastocyst can be isolated to generate pluripotent stem cell lines, which maintain the ability to differentiate into three germ layers; endoderm, mesoderm and ectoderm, and further to their derivatives.

The term ‘stem cell’ defines a cell that can proliferate to replace itself indefinitely in a process called self-renewal, but can also differentiate into a daughter cell with a restricted capacity to proliferate. The study of embryo-like stem cells started from teratomas when Askanazy proposed a theory that somatic tissues of teratoma develop by embryonic differentiation from a single cell or a group of cells representing embryonic

germinal layers (Askanazy, 1907). Later, Kleinsmith & Pierce (1964) showed that a single embryonal carcinoma cell (ECC) from teratocarcinoma could indeed reform a complex tumour in a new host (Kleinsmith and Pierce, 1964).

Pluripotent embryonic stem cells were first isolated from mouse blastocysts by Evans & Kaufman (Evans and Kaufman, 1981). The term 'embryonic stem cell' (ESC) was launched by Gail Martin when he established a new mouse embryonic stem cell line. It was noted that these cells had the essential features of ECCs derived from mouse teratocarcinomas, such as growth rate, morphology, and the ability to form teratomas (Martin, 1981). The first primate embryonic stem cell line, from a rhesus monkey embryo, was derived by Thomson et al. (Thomson et al., 1995). Three years later, Thomson et al. established first human embryonic stem cell (hESC) lines (H1, H13, H14, H7 and H9) from the ICM cells of the blastocyst stage embryos produced by *in vitro* fertilization (IVF) for clinical purposes. Like rhesus monkey ESCs, these cells had a normal karyotype, a high ratio of nucleus to cytoplasm, prominent nucleoli, telomerase activity, and the ability to form teratomas with all three embryonic germ layers. Interestingly, rhesus monkey ESCs, human ESC and EC cells are able to differentiate also into trophoblast (Thomson et al., 1998).

Captivatingly, somatic cell can be reversed back to pluripotent by activating factors maintaining stemness. Nuclear transfer and fusion experiments prove that factors present in the mammalian oocyte or embryonic stem cells, can reprogram somatic nuclei back to an undifferentiated state (Tada et al., 2001; Wilmut et al., 1997). The search of these factors lead to the break through study revealing that mouse ESCs can be reprogrammed from fibroblast cells simply by the overexpression of only four transcription factors *Oct4*, *Sox2*, *c-myc*, and *Klf4* (Takahashi and Yamanaka, 2006). Nuclear reprogramming of human cells can be also achieved by fusion of somatic cells with hESCs (Cowan et al., 2005). Finally, human induced pluripotent stem cells (hiPSCs) were reprogrammed from somatic cells by lentiviral overexpression of *OCT4*, *SOX2*, *NANOG*, and *LIN28* (Yu et al., 2007) and by retroviral overexpression of *OCT4*, *SOX2*, *KLF4* and *MYC* (Takahashi et al., 2007).

2.1.2 Properties of pluripotent stem cells

Determination of pluripotency

The pluripotency of a cell can be determined by several experimental procedures. First of all, pluripotent cells have a high cloning efficiency, which means that a single cell has a capacity to form a colony of clonal derivatives, which maintain the capacity to differentiate. Moreover, pluripotent cells can participate in the development of an organism. When ESCs are combined with normal pre-implantation embryos, which are placed on the uterus, they contribute to the normal development and to the germ line in chimeras (Bradley et al., 1984). Due to ethical issues this ability has not been tested for hESCs. In addition, pluripotent cells have a high expression of telomerase (Thomson et al., 1998; Wright et al., 1996). Telomerase is a DNA polymerase that adds telomere

repeats to the chromosome ends, and is thus necessary for telomere length maintenance, extending the replicative life-span of stem cells and protecting them from replicative senescence in culture (Harley et al., 1990).

One of the most important properties of pluripotent cells is their potential to differentiate. The differentiation potential is tested *in vivo* by injecting immunodeficient athymic mice with a suspension of cells, which form a teratoma containing derivatives of all three primary germ layers within few weeks. *In vitro,* differentiation potential can be tested by growing cells in suboptimal culture conditions, such as the lack of supporting matrix or growth factors. This leads to the formation of differentiated rounded aggregates called embryoid bodies (EBs), which grow in suspension (Martin, 1981). In principle, human ESCs can be differentiated to all cell types. After commitment to germ cell layers, specific differentiation protocols can be used for the induction of special cell types. For example, endodermal derivatives can be further induced to hepatocytes, insulin producing β cells, and lung epithelium. Mesodermal derivatives can be used to generate chondrocytes, osteocytes, skeletal myoblasts, hematopoietic cells, endothelial cells and cardiomyocytes, whereas ectodermal precursors can be induced to keratinocytes, retinal pigment epithelium, oligodendrocytes, astrocytes and mature neurons **(Figure 1)**. (Vazin and Freed, 2010)

The differentiation status of cells can be monitored by analysing specific markers of the differentiation pathway in question. However, before appearance of cell type specific markers, common to all pathways, is the induction of embryonic germ cell specific markers that can be used to indicate early onset of differentiation and loss of pluripotency. The most commonly used markers are as follows: for the trophectoderm, *GATA3*, *GATA2*, *CDX2*, *hCGα*, *hCGβ*, *PL-1*, *GCM1*, *CD9*, and *HLA-G*; for the endoderm, *AFP*, *GATA4*, *GATA6*, *LAMININ B1*, *ALBUMIN*, *PDX1*, *FOXA2*, and *SOX17*; for the mesoderm *BRACHYURY*, *RUNX1*, *VEGFR2*, *αMHC*, and *MSX1*; and for the ectoderm *NF-200*, *NF68*, *NCAM*, *FGF5*, *PAX6*, *NEUROD1*, *MAP2*. (Cameron et al., 2008; Fong et al., 2008; Hay et al., 2004; Hyslop et al., 2005; Matin et al., 2004; Melchior et al., 2008; Takahashi et al., 2007).

Surface antigens

In order to study pluripotency, one has to define markers of the pluripotent state and the changes that take place when hESCs are committed towards differentiation. Cell surface markers are highly important for the identification of cell state, but relative slow reacting. Human ESCs have high expression of alkaline phosphatase (ALP) enzyme, stage-specific embryonic antigens: SSEA-3, SSEA-4, and SSEA-5 oligosaccharides, and keratin sulphate-related antigens: TRA-1-60, TRA-1-81 glycoproteins (Andrews, 2011; Thomson et al., 1998). In addition, F15-14-1 (Thy1), GCTM2, GCTM343, CD9, FZD7, CD24 and TRA-2-49/TRA-2-54 (liver/bone/kidney isozyme of alkaline phosphatase), HLA-A,-B, and -C can be used to identify pluripotent state of hESCs (Draper et al., 2004; International Stem Cell Initiative et al., 2007).

In hESC differentiation, the expression of SSEA-3 starts to decline in the early stages, whereas SSEA-4, TRA-1-60, TRA-1-81, TRA-2-54 and Thy1 disappear further on during differentiation (Enver et al., 2005). Recently, a new marker, SSEA-5 was shown to decline even faster than SSEA-3 (Tang et al., 2011). The SSEA-1 and A2B5 are markers of differentiated cells and can be used to define the loss of the pluripotent stage. However, their reaction kinetics is dependent on the differentiation direction and both are induced relatively slowly (Draper et al., 2002). Interestingly, although surface markers are tightly controlled in development and excellent markers, SSEA-3 and SSEA-4 for example are shown not to be essential for pluripotency maintenance of hESCs (Brimble et al., 2007). Furthermore, individuals who are unable to synthesize these antigens appear healthy, although women with this genotype suffer from high rates of early spontaneous abortions (Cantin and Lyonnais, 1983).

Unique cell cycle

The median cell cycle time for hESCs is 16 hours, whereas active somatic cells require 24-32 hours to complete cell division (Becker et al., 2006). In an undifferentiated hESC population, the majority of the cells are in the S phase of the cell cycle with a short G_1-phase that shifts to a long G_1-phase in differentiation (Becker et al., 2006; Card et al., 2008; Qi et al., 2009). In addition, hESCs exhibit symmetric cell division during differentiation, as both daughter cells commit for cell lineage specification (Zwaka and Thomson, 2005).

Unique epigenetic signature

Epigenetic regulation, such as DNA methylation and histone modifications, affect gene expression without altering the DNA sequence. After fertilization, the methylation patterns of the sperm and oocyte are erased following *de novo*-methylation. At the blastocyst stage the embryo undergoes major epigenetic modifications, such as DNA methylation, X-chromosome inactivation and chromatin remodeling. (Allegrucci et al., 2005; Morgan et al., 2005).

The methylation of CpG sites of DNA is commonly associated with more permanent silencing. The methylation pattern of CpG sites in hESCs is unique compared to differentiated and somatic cells (Bibikova et al., 2006). The methylation status is highest in undifferentiated hESCs following global reduction in differentiation (Laurent et al., 2010).

The activity of chromatin structure is partly regulated by modifications of the histone proteins. The most prevalent histone modification in hESCs is the histone H3 lysine 4 trimethylation ($H3K4me^3$) which is associated with active promoters (Guenther et al., 2007; Zhao et al., 2007) and inversely correlated with DNA methylation status (Hawkins et al., 2010). On the contrary, the $H3K27me^3$ mark is associated with inactive promoters. Some areas in hESCs are bivalent carrying both $H3K4me^3$ and $H3K27me^3$ histone marks (Laurent et al., 2010; Pan et al., 2007). These promoters are considered poised for

activation or repression at later stages of development. Similarly, central transcription factors of pluripotent cells OCT4, NANOG and SOX2 carry active H3K4me^3 mark in undifferentiated cells but convert bivalent (H3K4me^3 and H3K27me^3) in differentiation (Pan et al., 2007).

X-chromosome inactivation, a whole chromosome silencing, compensates for the extra X-chromosome dosage relative to male. Human ESCs, have two active X-chromosomes of which one undergoes X-inactivation in differentiation (Dhara and Benvenisty, 2004). However, several undifferentiated cultured hESC lines have only one active X-chromosome (Hoffman et al., 2005; International Stem Cell Initiative et al., 2007; Silva et al., 2008).

2.1.3 Regulation of pluripotency

Preservation of the pluripotent state requires maintenance of self-renewal and inhibition of differentiation. The identification of individual factors crucial for self-renewal, is a key to understanding the regulation of pluripotency. In addition, activation of these crucial factors requires signal transmission outside of the cells. Certain proteins, miRNAs and pathways have already been shown to be highly important for the maintenance of pluripotency of hESCs.

Regulators of pluripotency

Studies of mESCs identified that the transcription factors Oct4, Sox2 and Nanog are essential for the maintenance of pluripotency and mammalian development (Avilion et al., 2003; Chambers et al., 2003; Mitsui et al., 2003; Nichols et al., 1998; Niwa et al., 2000; Scholer et al., 1990; Wang et al., 2003; Yuan et al., 1995). Similarly, OCT4 (*POU5F1*), NANOG and SOX2 are highly expressed in hESCs (Fong et al., 2008; Hay et al., 2004; Hyslop et al., 2005; Matin et al., 2004; Zaehres et al., 2005; Zafarana et al., 2009).

Silencing of OCT4 (Hay et al., 2004; Matin et al., 2004; Zaehres et al., 2005; Zafarana et al., 2009), NANOG (Hyslop et al., 2005; Zaehres et al., 2005) or SOX2 (Adachi et al., 2010; Fong et al., 2008) leads to morphological differentiation of hESC colonies. In addition, surface antigen expression of TRA-1-60, SSEA-3, SSEA-4 and TRA-1-81 decreases and expression of SSEA-1 increases (Adachi et al., 2010; Fong et al., 2008; Hyslop et al., 2005; Matin et al., 2004; Zaehres et al., 2005). Silencing of any of the factors has been reported to lead to the induction of trophectoderm and endoderm associated transcripts (Adachi et al., 2010; Fong et al., 2008; Hay et al., 2004; Hyslop et al., 2005; Matin et al., 2004; Zaehres et al., 2005; Zafarana et al., 2009). In contrast to these reports, Wang et al. recently published that SOX2 is dispensable for maintenance of pluripotency in hESCs. Interestingly, although SOX3 is normally moderately expressed it was found to be induced in SOX2 depleted hESCs, and self-renewal was lost only upon simultaneous depletion of both SOX3 and SOX2. Importantly, a large variation

between different cell lines in their responsiveness to the depletion of core factors was observed, especially in the expression of SOX2 (Wang et al., 2012).

The overexpression of the core factors in hESCs does not induce differentiation, although the differentiation capacity is affected (Wang et al., 2012). The NANOG overexpressing hESCs can be maintained for a long-term in feeder-free culture system in the absence of conditioned media (Darr et al., 2006), and have a faster cell cycle (Zhang et al., 2009b), upregulate certain markers of definitive endoderm and are unable to differentiate into neuroectoderm (Wang et al., 2012). Overexpression of OCT4 does not induce any lineage markers, and suppresses differentiation into neuroectoderm (Wang et al., 2012). Overexpression of SOX2 has been reported to affect transcripts of proliferation and development, but not lineage markers (Wang et al., 2012), whereas one report measured induction of trophectoderm (Adachi et al., 2010).

OCT4, SOX2 and NANOG are termed as the core regulators of pluripotency as they control each other's expression and co-occupy over three hundred (353) gene promoters maintaining the stem cell specific gene expression (Boyer et al., 2005). Interestingly, nearly half of the transcripts occupied by these core factors are transcriptionally inactive, implying that in addition to maintain pluripotency OCT4, SOX2 and NANOG also repress differentiation (Boyer et al., 2005). Based on genome-wide expression profiling after depletion, 111 genes require all three factors for the maintenance of expression in hESCs, whereas a larger set of genes is co-regulated only by NANOG and OCT4 or by SOX2 and OCT4 (Wang et al., 2012). To conclude, OCT4 and NANOG are crucial for the pluripotency of hESCs whereas expression of SOX2 is dispensable in some hESC lines, and presumably replaceable by SOX3.

Genome-wide transcriptome profiling of hESCs has identified additional candidate genes responsible for maintaining pluripotency of hESCs, such as *TERT* (Xu et al., 2002), *LEFTY1, TDGF1*(Sato et al., 2003), *DDPA4* (*FLJ10713*), *DNMT3B* (Sperger et al., 2003), *LIN28* (Richards et al., 2004), *L1TD1 (FLJ10884)*, and *RPC32* (Enver et al., 2005). Based on meta-analysis (28 hESC lines, 38 studies), the only gene specific for hESCs is *OCT4,* while a number of transcripts, including *L1TD1 (FLJ10884), NANOG, OCT4, TDGF1, LECT1, INDO, LEFTY1, ZIC3, DPPA4, DNMT3B, LIN28*, and *GAL,* have a very high expression ratio compared to differentiated counterparts (Assou et al., 2007). In addition, based on polymerase chain reaction (PCR) detection performed on 59 hESC lines, *TDGF1*, *OCT4*, *GABRB3*, *GDF3* and *DNMT3B* have the highest correlation with *NANOG* (International Stem Cell Initiative et al., 2007).

Surprisingly, the characterization of most of these highly expressed candidate factors in hESCs is still relatively limited, although most of them are well studied in other cell types. TDGF1 (CRIPTO-1) is a surface receptor and cofactor of NODAL ligand that activates TGFβ-signaling. Silencing of TDGF1 in EC cells decreases growth rate and levels of SSEA-3 (Baldassarre et al., 1996). DNMT3B is highly used as a marker of pluripotency in hESCs and reported to be a vital DNA methyltransferase responsible for

the *de novo* DNA methylation of the genome during mouse development (Okano et al., 1999) and methylation of Oct4 and Nanog promoters in differentiation of mESCs (Li et al., 2007). Based on a recent publication, silencing of DNMT3B in hESCs following neural differentiation alters neuronal maturation and methylation status of early neural genes (Martins-Taylor et al., 2012).

The RNA binding protein, LIN28 was characterized in hESCs only after it had been used in the reprogramming studies. The LIN28 protein is highly expressed in undifferentiated hESCs (Darr and Benvenisty, 2009). Silencing of LIN28 levels less than 80 % does not influence expression of pluripotency markers (Darr and Benvenisty, 2009; Qiu et al., 2010), whereas more intense silencing affects protein levels of OCT4 and cell growth (Peng et al., 2011; Qiu et al., 2010). Importantly, LIN28 binds to the mRNA of *OCT4* facilitating the expression of OCT4 at the post-transcriptional level (Qiu et al., 2010). Further, LIN28 promotes generation of induced pluripotent cells (Yu et al., 2007) and enhances translation of genes important for growth and survival (Peng et al., 2011).

MicroRNAs

Small noncoding RNA molecules, microRNAs (miRNAs) play important role in development. MicroRNAs regulate one third of protein coding genes post-transcriptionally by binding to the untranslated region (UTR) of the messanger RNA (mRNA) inhibiting the translation of the target mRNAs. MicroRNAs are transcribed as long primary pri-miRNAs that are trimmed into 70 nucleotide stem-loop pre-miRNAs by double-stranded RNA-specific endoribonuclease Drosha in the nucleus. Pre-miRNAs are then exported to cytoplasm and processed by another ribonuclease Dicer into ~22 nucleotide mature functional miRNAs. (Mallanna and Rizzino, 2010; Winter et al., 2009)

Undifferentiated hESCs have high expression of small RNA gene cluster miR-302b*-302b-302c*-302a*-302a-302d-367 on chromosome 4 (Barroso-delJesus et al., 2008; Card et al., 2008; Lakshmipathy et al., 2007; Ren et al., 2009; Rosa et al., 2009; Suh et al., 2004; Xu et al., 2009), which is regulated by OCT4, NANOG and SOX2 (Card et al., 2008). In addition, hESC express miR-17-18a-19a-20a-19b1-92_1 cluster on chromosome 13 (Laurent et al., 2008; Wilson et al., 2009), and miR-371-372-373*-373 (Lakshmipathy et al., 2007; Laurent et al., 2008; Suh et al., 2004; Xu et al., 2009) and miR-520 clusters on chromosome 19 (Ren et al., 2009; Tzur et al., 2008). Interestingly, many of the miRNAs in these clusters contain the same seed sequence AAGTGC and thus are predicted to target the same mRNAs (Laurent et al., 2008). Sequencing based studies have identified a number of additional miRNAs expressed in hESCs (Bar et al., 2008; Morin et al., 2008). Based on these studies the most highly expressed miRNAs in hESCs that are decreased in differentiation, are miR302b and miR302a. Similarly, miR-302 cluster has highest expression in hiPSCs (Wilson et al., 2009).

Based on bioinformatics and experimental analysis, miR-302 targets 150-500 human genes (Lipchina et al., 2011; Rosa et al., 2009). Inhibition of the miR-302 cluster in hESCs

changes the cell cycle profile of hESCs similar to that of differentiated cells and significantly decreases levels of SSEA-3 (Card et al., 2008; Lipchina et al., 2011). In addition, loss of miR-302 in hESCs inhibits mesodermal and endodermal lineage differentiation accelerating formation of neuroectodermal derivatives (Rosa et al., 2009). In fact, miR-302 targets NR2F2, which is one of the earliest markers or neural cells (Rosa et al., 2009). In addition, miR-302 cluster targets TGFβ-pathway inhibitors LEFTY1 and LEFTY2 (Rosa et al., 2009) and regulates directly the cyclin D1 expression (Card et al., 2008). More importantly, miR-302 targets epigenetic regulators *AOF1*, *AOF2*, *MECP1-p66* and *MECP2* and enable reprogramming of somatic cells to pluripotent (Lin et al., 2011).

Perturbation of miRNA pathway by inhibition of Dicer or Drosha in hESCs decreases cell proliferation, increases G_1 phase, and interestingly raises the levels of *NANOG*, *OCT4* and *SOX2* (Qi et al., 2009). In fact, miR-195 directly regulates WEE1 kinase, the negative regulator of the CycB/cyclin dependent kinase (CDK) complex of the G_2/M phase of the cell cycle, whereas miR-372 regulates p21, the CDK inhibitor CDKN1A/p21 that operates mainly in G_1 phase affecting the cell cycle of hESCs (Qi et al., 2009).

Interestingly, miR-145 targets and regulates directly endogenous levels of *OCT4*, *KLF4* and *SOX2*, whereas OCT4 represses the promoter of miR-145 forming a feedback loop. The expression of miR-145 is low in hESCs, but highly induced upon differentiation. Silencing of miR-145 in undifferentiated hESCs leads to upregulation of mRNA levels of *SOX2* and protein levels of OCT4, SOX2 and KLF4, explaining why Dicer inhibition raised expression of the core factors. Thus miR-145 controls the stable levels of pluripotency factors in self-renewal, but also has an important role in repression of pluripotency in differentiation. (Xu et al., 2009)

WNT-signaling

WNT-signaling is activated by hydrophobic cell membrane associated WNT ligands that bind to the Frizzled surface receptors. Frizzled triggering activates Dishevelled protein that inhibits the multiprotein complex, consisting of axin, glycogen synthase kinase-3β (GSK3β) and adeomatosis polyposis coli (APC), which is responsible for marking β-catenin for proteosomal degradation. When this complex is inhibited, β-catenin accumulates in the cytoplasm and travels to the nucleus **(Figure 2)**. (Nusse et al., 2008)

In undifferentiated hESCs, β-catenin is accumulated in the nucleus, on the contrary to cells grown in differentiating conditions (Sato et al., 2004). In addition, Frizzled receptor FZD7 is specifically expressed in undifferentiated hESCs, compared to differentiated cell types, and silencing of FZD7 leads to differentiation and decrease of OCT4 (Melchior et al., 2008). Activation of WNT-signaling, by inhibiting GSK-3 with chemical inhibitor or by adding Wnt3a ligands, has been reported to be sufficient to maintain self-renewal of hESCs in feeder-free unconditioned media in short term cultures (Sato et al., 2004). Furthermore, WNT-signaling has been reported to support survival and proliferation of hESCs (Cai et al., 2007; Dravid et al., 2005). However, based on a recent report, β-catenin

signaling is inactive in undifferentiated hESCs and induction of WNT-signaling leads to differentiation (Davidson et al., 2012). Thus, the role of active WNT-signaling in the maintenance of pluripotency and differentiation remains unclear.

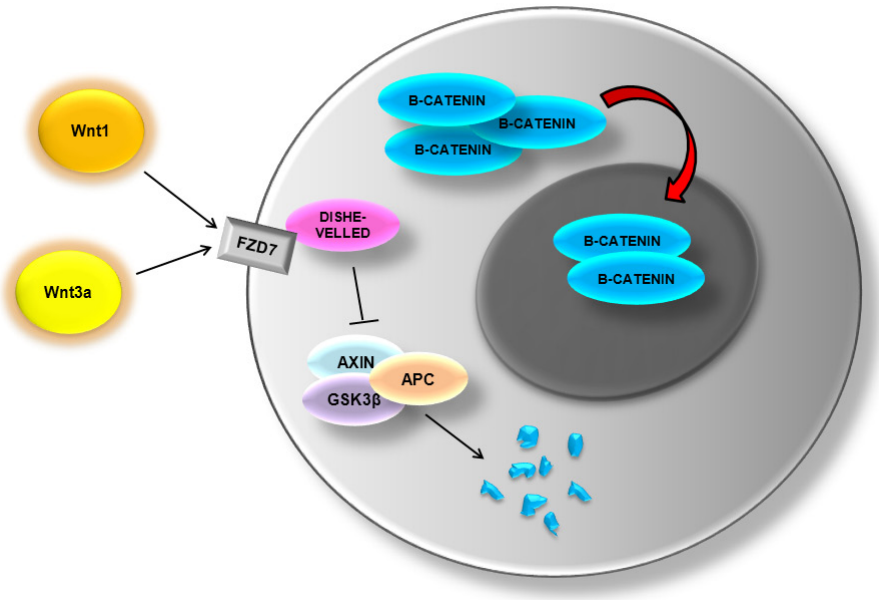

Figure 2. WNT-signaling. Wnt ligands (Wnt1)(Wnt3a) bind to the Frizzled receptor (FZD7) and activate Dishevelled protein, that inhibits the protein complex of AXIN/APC/GSK3β that is responsible for marking β-catenin for proteosomal degradation. Inhibition of this complex leads to accumulation of β-catenin and signal transmission to the nucleus.

FGF-signaling

Addition of fibroblast growth factor FGF (bFGF, FGF2) to the culture medium of hESCs containing serum replacement allows culture of hESCs on fibroblasts (Amit et al., 2000). A high concentration of FGF (100 ng/ml) can support undifferentiated growth also in the absence of supporting fibroblasts (Levenstein et al., 2006). Furthermore, inhibition of FGF-signaling, by chemical inhibitor SU5402, reverses the positive effect of FGF addition (Vallier et al., 2005). In addition, FGF and FGF-receptor are endogenously expressed by hESCs (Dvorak et al., 2005; Kang et al., 2005).

FGF activates MEK1/ERK MAPK cascade, positively affecting the proliferation and survival of hESCs (Kang et al., 2005), thus explaining why FGF addition is so vital for the maintenance of hESC cultures. In addition, FGF sustains expression of NANOG, contributing for the maintenance of pluripotency (Greber et al., 2010; Yu et al., 2011).

TGFβ-signaling

TGFβ-signaling is regulated by soluble ligands surrounding the extracellular space of a cell. The ligands transmit signal by binding to type I and type II receptors that form

heterotetrameric complexes. TGFβ1 binds to the receptor consisting of Alk5 and TβRII dimer, whereas Nodal and Activin bind to and receptor dimer consisting of ActRIIB and Alk4 or Alk7. Binding of Nodal, Activin or TGFβ1 to receptor leads to activation and phosphorylation of SMAD2/3. Another branch of TGFβ-signaling is activated by bone morphogenetic protein (BMP) that binds a receptor dimer consisting of BMPR1A/Alk3 and ActRIIB, and leads to phosphorylatation of SMAD1/5/8. The binding of these ligands to their receptors is regulated by a cofactor, Cripto, and inhibitors, such as Lefty1, Lefty2, Follistatin, Noggin and Cerberus. Activated SMAD2/3 or SMAD1/5/8 binds to SMAD4, and the resulting complex travels to the nucleus (**Figure 3**). (Moustakas and Heldin, 2009)

Undifferentiated hESCs express high levels of components of TGFβ-signaling, such as *NODAL*, *LEFTY1*, *LEFTY2*, *ACTIVIN*, *TGFβ1*, *CRIPTO, ALK-4*, *BMP2* (Beattie et al., 2005; Besser, 2004; Pera et al., 2004; Vallier et al., 2005). In addition, mouse embryonic fibroblasts used in co-culture or production of conditioned media (CM), secrete activin A precursor protein (Beattie et al., 2005) and Noggin (Xu et al., 2005). In undifferentiated hESCs, SMAD2/3 is in an activated phosphorylated state, whereas SMAD1/5 is inactive. Upon differentiation, the phosphorylation status is reversed (Besser, 2004; James et al., 2005; Vallier et al., 2005). Inhibition of type 1 TGFβ-signaling receptors (Alk4/5/7) by the synthetic compound SB431542, precludes SMAD2/3 phosphorylation, activates SMAD1/5/8 and leads to differentiation of hESCs (Beattie et al., 2005; James et al., 2005; Vallier et al., 2005). In addition, if hESCs are cultured in the presence of soluble recombinant TGFβ receptors, ActRIB, hrACtRIIB and hrCripto, which bind free TGFβ-signaling ligands, SMAD2/3 is inactivated (James et al., 2005).

The role of TGFβ-signaling activator and inhibitors in the maintenance of pluripotency has been studied with recombinant ligands. ACTIVIN A is sufficient to maintain SMAD2/3 phosphorylation and undifferentiated status of hESCs in non-conditioned culture media (Beattie et al., 2005; James et al., 2005; Xiao et al., 2006). In addition, hESCs cultured in recombinant NODAL exhibit prolonged expression of pluripotency markers (Vallier et al., 2004) and reduce the differentiation of hESCs in suboptimal culture conditions (James et al., 2005), whereas addition of TGFβ alone cannot (Vallier et al., 2005). Furthermore, inhibition of ACTIVIN by FOLLISTATIN significantly reduces the number of undifferentiated cells (Beattie et al., 2005), but cannot affect hESCs overexpressing NODAL, implying that NODAL and ACTIVIN can act independently to active pluripotency (Vallier et al., 2005). Although, LEFTY2 should inhibit TGFβ-signaling, a high dose of recombinant LEFTY2 protein cannot induce differentiation of hESCs. Similarly, hESCs overexpressing LEFTY2 can be maintained in culture (Vallier et al., 2005). Interestingly, SMAD2/3 activates directly the LEFTY2 promoter and undifferentiated hESCs express LEFTY2 (Besser, 2004). In addition, levels of LEFTY1 and LEFTY2 are inhibited by the miR-302 cluster (Rosa et al., 2009). These results imply that expression levels of LEFTY proteins are controlled in hESCs, but the exact role is still unclear.

Activation of another branch of TGFβ-signaling by addition of BMP4 or BMP2 induces differentiation (Xu et al., 2002) and activates SMAD1/5/8 (Besser, 2004; James et al., 2005; Pera et al., 2004; Vallier et al., 2005; Xu et al., 2005). Addition of soluble BMP receptors or BMP inhibitor protein NOGGIN to the culture media, blocks the differentiation caused by BMP proteins (Xu et al., 2002). BMP-signaling can induce expression of BMP2 and BMP4 ligands in hESCs, forming a positive feedback loop in differentiated cells (Xu et al., 2005). The activation of BMP-signaling leads to induction of trophoblast differentiation (Xu et al., 2002; Xu et al., 2005).

Addition of FGF to the culture medium increases the capacity of ACTIVIN and NODAL to maintain pluripotency (Vallier et al., 2005). In fact, addition of high FGF concentrations (100 ng/ml) to unconditioned culture media can reduce BMP-signaling to that of normal CM media (Xu et al., 2005), implying that TGFβ-signaling together with FGF-signaling is required for pluripotency maintenance. Strikingly, TGFβ-signaling is directly linked to pluripotency maintenance by core transcriptional regulators of pluripotency as both SMAD2/3 and SMAD1/5/8 can bind directly to the proximal promoter of the *NANOG* (Xu et al., 2008). Furthermore, the promoter activity of *NANOG* is enhanced by TGFβ-signaling and FGF-signaling and decreased by BMP-signaling (Xu et al., 2008). Thus the balance of the activated forms of SMAD2/3 and SMAD1/5/8 could determinate the expression of the *NANOG* and hence the pluripotency status of hESCs **(Figure 3)**.

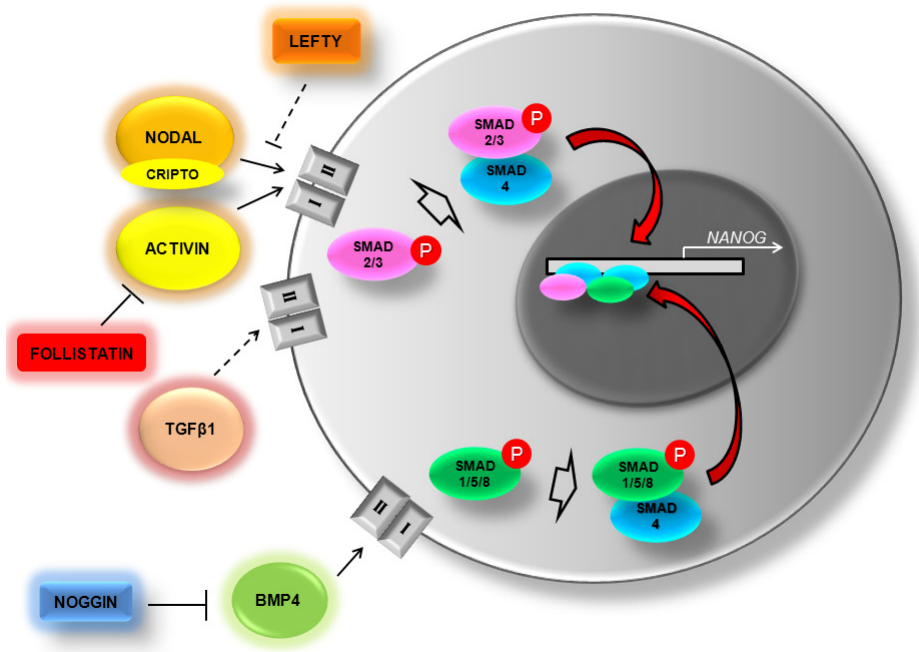

Figure 3. TGFβ-signaling in hESCs. TGFβ-signaling ligands transmit their signal by binding to a receptor consisting of type I and type II dimers. The binding of ligands is regulated by cofactor CRIPTO, and inhibitors such as LEFTY, FOLLISTATIN and NOGGIN. NODAL, ACTIVIN or TGFβ1 binding to receptors leads to activation of SMAD2/3. BMP4 binding activates SMAD1/5/8. Phosphorylated activated SMADs bind to SMAD4 and travel to the nucleus, where they can bind promoter of *NANOG*.

Summary of the pluripotency regulation

The **Table 1.** summarizes the most commonly used markers to detect pluripotency and differentiation. The **Figure 4** outlines the current scheme of the regulation and maintenance of the pluripotency in hESCs.

Table 1. The most common pluripotency and differentiation markers of hESCs.

PLURIPOTENCY	DIFFERENTIATION	
pSMAD2/3	**pSMAD1/5/8**	differentiation
OCT4	**miR-145**	differentiation
SOX2	**SSEA-1**	differentiation
NANOG	**A2B5**	neuroectoderm
DNMT3B	**FGF5**	ectoderm
LIN28	**PAX6**	ectoderm
TDGF1	**NEUROD1**	ectoderm
ALP/TRA-2-54	**MAP2**	ectoderm
SSEA-3	**BRACHYURY**	mesoderm
SSEA-4	**VEGFR2**	mesoderm
TRA-1-60	**αMHC**	mesoderm
TRA-1-81	**MSX1**	mesoderm
THY1	**AFP**	endoderm
CD9	**GATA6**	endoderm
FZD7	**PDX1**	endoderm
SSEA-5	**SOX17**	endoderm
miR-302 cluster	**GATA2**	trophoblast
miR-371 cluster	**CDX2**	trophoblast
miR-17 cluster	**hCG**	trophoblast
miR-520 cluster	**GCM1**	trophoblast

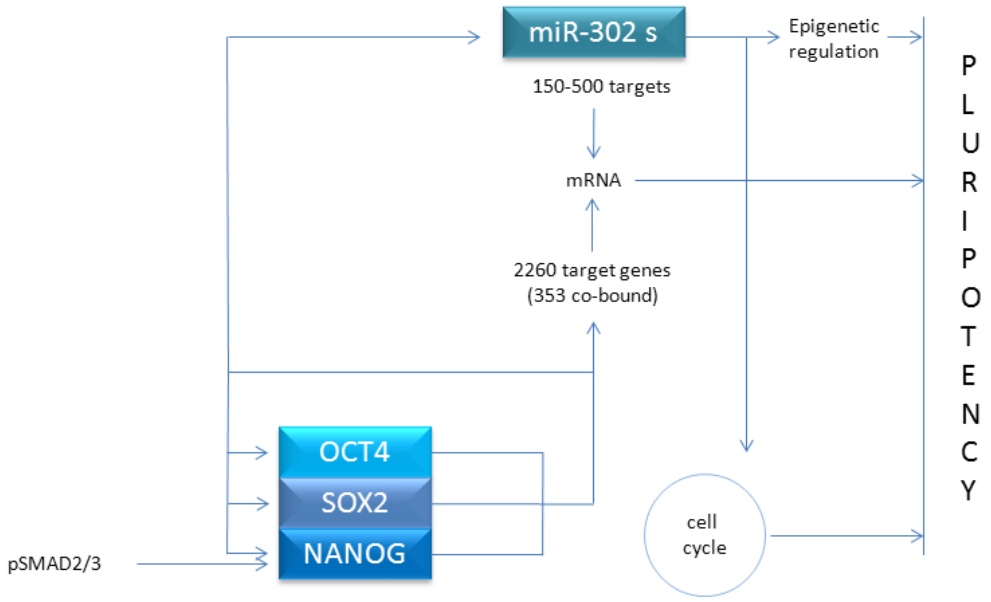

Figure 4. The regulation of pluripotency in hESCs. TGF-β signaling activates pSMAD2/3 that binds directly to the promoter of NANOG. Stem cell specific transcription factors OCT4, SOX2 and NANOG co-regulate a number of target genes maintaining pluripotency and inhibiting differentiation. The translation of these targets is regulated also by stem cell specific micro-RNAs, of which the miR-302 s cluster is highly expressed in hESCs. In addition, hESCs have a unique cell cycle and epigenetic signature.

2.1.4 Induction of pluripotency

After the discovery that human somatic cells can be reprogrammed by introducing lentiviral overexpression of *OCT4*, *SOX2*, *NANOG*, and *LIN28* (Yu et al., 2007) or by retroviral overexpression of *OCT4*, *SOX2*, *KLF4* and *MYC* (Takahashi et al., 2007), a wide range of reprogramming methods has been developed. These include introduction of various types of viruses and plasmids carrying the coding sequence for reprogramming factors or direct introduction of synthetic cell-penetrating peptide (CPP)-fusion proteins, mature double-stranded miRNAs or modified RNAs along with specific inhibitors and activators **(Table 2)** (Kim et al., 2009; Woltjen et al., 2009; Huangfu et al., 2008; Li et al., 2009; Soldner et al., 2009*;* Fusaki et al., 2009; Kaji et al., 2009*;* Yu et al., 2009; Jia et al., 2010; Warren et al., 2010; Zhu et al., 2010; Anokye-Danso et al., 2011; Subramanyam et al., 2011; Miyoshi et al., 2011; Lin et al., 2011). Based on various combinations of pluripotency factors used for reprogramming **(Table 2)**, it is clear that none of the individual factors is essential for the induction of pluripotency. Reprogramming can be achieved if the reprogramming factor is able to trigger on global demethylation and pluripotency maintenance program. The actual process of reprogramming is still largely unknown.

Table 2. Reprogramming methods used for the induction of pluripotency in human cells.

INTRODUCTION OF:	REPROGRAMMING FACTORS:	INHIBITORS/ACTIVATORS:	REFERENCE:
retrovirus	*OCT4, SOX2, KLF4* and *MYC*		(Takahashi et al., 2007)
retrovirus	*OCT4*, *SOX2* and *KLF4*	HDAC inhibitor valproic acid (VPA)	(Huangfu et al., 2008)
retrovirus	*OCT4, SOX2, KLF4, MYC,* Venus, has-miR-302b and has-miR-372		(Subramanyam et al., 2011)
lentivirus	*OCT4*, *SOX2*, *NANOG*, and *LIN28*		(Yu et al., 2007)
lentivirus	*OCT4* and *KLF4*	GSK-inhibitor (CHIR99021) and lysine–specific demethylase 1 (Parnate)	(Li et al., 2009)
lentivirus	*OCT4*	histone deacetylase inhibitor (NaB), activator of PDK1 (PS48), TGFβ receptor inhibitor (A-83-01) and inhibitor of MAPK/ERK (PD0325901)	(Zhu et al., 2010)
lentivirus	miR302/367 cluster		(Anokye-Danso et al., 2011)
lentivirus Cre-recombinase excisable	*OCT4, SOX2,* and *KLF4*		(Soldner et al., 2009)
Sendai virus Non-integrating	*OCT4*, *SOX2*, *KLF4* and *MYC*		(Fusaki et al., 2009)
proteins	OCT4, SOX2, KLF4 and MYC		(Kim et al., 2009)
Episomal vector Non-integrating	*OCT4*, *SOX2*, *NANOG*, *LIN28*, *MYC*, *KLF4*, *SV40LT*		(Yu et al., 2009)
Minicircle vector	*OCT4*, *SOX2*, *NANOG*, and *LIN28*		(Jia et al., 2010)
PiggyPac	*OCT4*, *SOX2*, *KLF4* and *MYC*		(Woltjen et al., 2009)
PiggyPac Single multiprotein cassette	*OCT4*, *SOX2*, *KLF4* and *MYC*		(Kaji et al., 2009)
Synthetic RNA	*OCT4*, *SOX2*, *KLF4*, *MYC* and *LIN28*	interferon inhibitor	(Warren et al., 2010)
miRNA	mir-200c, mir-302 s and mir-369 s		(Miyoshi et al., 2011)
Induciple vector	mir-302 s cluster		(Lin et al., 2011)

The reprogramming of somatic cell to pluripotent typically requires three to four weeks. The use of synthetic RNA delivery can reduce this time to 17 days (Warren et al., 2010), whereas reprogramming with miRNAs can be achieved in 12-14 days (Anokye-Danso et al., 2011). The basic validation of hiPSCs includes assessment of morphology, karyotype, telomerase activity, expression of hESC specific surface antigens, EB and teratoma formation, DNA fingerprinting, gene expression analysis and methylation/histone modification status of *OCT4, SOX2*, and *NANOG* promoters (Takahashi et al., 2007; Yu et al., 2007) **(Figure 5)**.

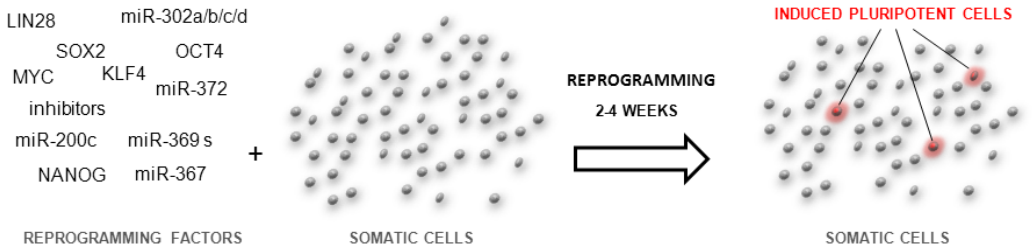

Figure 5. Reprogramming of human somatic cells. Generation of human induced pluripotent cells (hiPSCs) can be achieved by introducing reprogramming factors (such as *OCT4*, *SOX2*, *NANOG*, *LIN28,OCT4*, *SOX2*, *KLF4*, *MYC,* miR302 s, miR-367, miR-200c, miR-369 s) to somatic cells, that trigger on global demethylation and pluripotency maintenance. Only few of the somatic cells (grey) are reprogrammed into induced pluripotent cells (red).

Towards the clinics

Human iPSCs enable induction of patient specific cell lines. It was firstly believed that autologous hiPSCs would eliminate the concern of immune rejection. However, recent study has demonstrated that syngeneic mouse induced pluripotent cells could be rejected following transplantation (Zhao et al., 2011). Further studies are needed in order to resolve if autologous hiPSCs derived differentiated transplants can also have immunoreactive properties. It is possible that the immunogenicity of induced pluripotent cells can vary depending of the reprogramming and differentiation method used and the origin of somatic cell type (Boyd et al., 2012).

Compared to hESCs, the use of hiPSCs avoids the ethical and accessibility problems concerning the use of embryo derived cell lines. Reprogramming can also facilitate studies of rare genetic disorders as disease specific cell types can be generated. Human iPSCs could also replace animal experiments in drug development and toxicity tests. However, several problems, such as low generation efficiency, time-consuming procedures, induction of genetic alterations, tumor formation, and partial reprogramming, have delayed the potential practical use of hiPSCs in clinical applications.

The reprogramming efficiency, the fraction of iPSCs generated from somatic starting cells, was only 0.01 % for the first generation of viral methods (Takahashi et al., 2007; Yu et al., 2007) and even lower in the case of recombinant proteins 0.001 % (Kim et al., 2009) and plasmid DNA 0.005 % (Jia et al., 2010). Doxycycline (DOX)-inducible lentiviral vectors (Hockemeyer et al., 2008) and use of synthetic modified RNA (Warren et al., 2010) raises the reprogramming efficiency up to 2 %. In general, addition of multiple factors simultaneously (Nakagawa et al., 2008; Takahashi et al., 2007; Yu et al., 2007; Yu et al., 2009), low oxygen culture conditions (Warren et al., 2010; Yoshida et al., 2009) and chemical compounds (Huangfu et al., 2008; Yu et al., 2011; Zhu et al., 2010) improves reprogramming efficiency. Based on the recent publication, reprogramming with miRNAs, instead of transcription factors, increases reprogramming efficiency dramatically (Anokye-Danso et al., 2011).

Reprogramming achieved with lenti- or retroviral overexpression, results hiPSCs containing exogenous overexpression cassettes integrated into the genome (Takahashi et al., 2007; Yu et al., 2007). These exogenous DNA fragments can affect gene expression and block differentiation (Yu et al., 2007), or are silenced in culture (Takahashi et al., 2007). The unpredictable behavior of viral casettes, especially considering the therapeutic approaches, can be solved by use of transgene-free techniques, such as use of recombinant protein (Kim et al., 2009), plasmid that enables transgene removal (Kaji et al., 2009; Woltjen et al., 2009; Yu et al., 2009), minicircle vectors that lack bacterial origin of replication and antibiotic resistance gene (Jia et al., 2010), synthetic RNA delivery (Warren et al., 2010), and transfection of mature double-stranded microRNAs (Miyoshi et al., 2011).

Most of the reprogramming studies have used human fibroblasts as the somatic cell source. From the therapeutic point of view, a less invasive cell source and a cell type with a lower risk for somatic mutations need to be validated for clinical use. Toward this goal, hiPSCs have been reported that are generated from keratinocytes from hair follicles (Aasen and Izpisua Belmonte, 2010), cord blood (Haase et al., 2009), peripheral blood T cells (Loh et al., 2010; Seki et al., 2010; Staerk et al., 2010) and renal tubular cells in the urine (Zhou et al., 2011).

Although, hiPSCs resemble hESCs, they are not exactly analogous. The global gene-expression and epigenetic profiles of hESCs and hiPSCs are similar, but not identical (Chin et al., 2009; Kim et al., 2010; Takahashi et al., 2007). Interestingly, the gene-expression profile of the late-passage hiPSCs is more alike with hESCs than early-passage hiPSCs (Chin et al., 2009). In addition, few miRNAs are consistently expressed differently between ESCs and iPSCs (Chin et al., 2009; Kim et al., 2010; Wilson et al., 2009). Further studies are needed to resolve whether these differences will play a functional role that may potentially risk therapeutic applications.

2.2 Hypoxia

2.2.1 Hypoxia response

Oxygen is essential for the maintenance of cell homeostasis and energy production. Oxidative phosphorylation transfers chemical energy of carbon bonds to the high-energy phosphate bond in adenosine triphosphate (ATP), which is the energy unit of a living cell. The ambient oxygen concentration is 21 %. However, most mammalian cells exist in 2-9 % oxygen. A state were the oxygen level is decreased relative to normal is defined as hypoxia. A hypoxic surrounding can arise because of physiological, pathological or developmental reasons. The changes in oxygen availability triggers hypoxia response, which is mediated by hypoxia-inducible factors (HIFs). HIFs form heterodimers consisting of oxygen labile α-subunits: HIF1α, HIF2α and HIF3α that dimerize with stable β subunit. HIF1α is expressed in all cell types, whereas HIF2α, and HIF3α are

specific for certain tissues. The heterodimer recognize and bind to hypoxia response elements (HREs) in the genome. In normal oxygenated environment, normoxia, HIFα subunits are constantly degraded. This is controlled by prolyl hydroxylase domain-containing enzymes (PHDs), a family consisting of PHD1, PHD2 and PHD3, which hydroxylate proline residues of HIFα. The E3 ubiquitin ligase, the von Hippel-Lindau protein (pVHL) complex recognizes and marks hydroxylated HIFα for degradation. PHD utilizes oxygen as a substrate, leading to inhibited activity in hypoxic conditions. In addition, when oxygen is limited, mitochondria stimulate the production of cellular reactive oxygen species (ROS), which also inhibit PHD activity, stabilizing HIFα and leading to activation of hypoxia responsive genes. However, multiple questions of hypoxia mediated HIF regulation are still open for investigation. In addition, the response varies depending of the level and time of hypoxic exposure. **(Figure 6)**, Reviewed by (Bertout et al., 2008; Jokilehto and Jaakkola, 2010; Majmundar et al., 2010; Semenza, 2012).

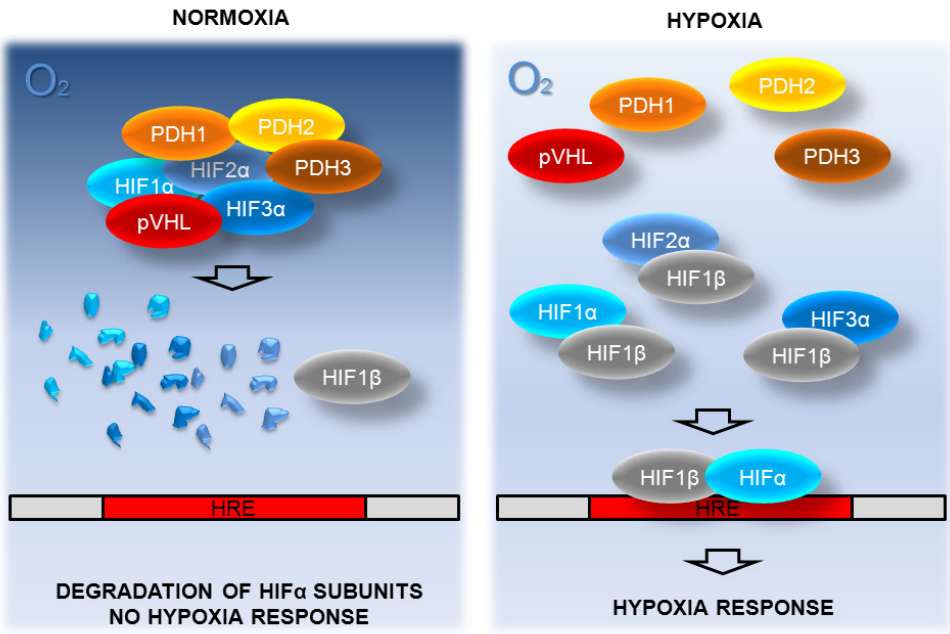

Figure 6. Hypoxia response. In sufficient oxygen environment, normoxia, prolyl hydroxylase domain containing enzymes (PHDs) hydroxylate proline residues of hypoxia-induciple factors (HIFα). The von Hippel-Lindau protein (pVHL) complex recognizes and marks hydroxylated HIFα subunits for degradation. In low oxygen, hypoxia, HIFα subunits form a dimer with HIF1β subunit. Resulting heterodimer binds to hypoxia response elements (HREs) in the genome and trigger hypoxia response.

The purpose of hypoxia response is to activate hypoxia responsive pathways and alter cellular metabolism into anaerobic glycolysis, to ensure energy production and protection against apoptosis. Therefore, HIF1α promotes expression of glucose transporters and glycolytic enzymes. HIF2α has unique target genes that appear to suppress aberrant ROS

accumulation. Hypoxia also stimulates lipid storage and inhibits lipid catabolism. In addition to metabolic changes, hypoxia has an import role in angiogenesis, especially in embryonic vascular development. Importantly, hypoxic response and HIFs play substantial roles in various aspects of cancer development. More importantly, hypoxia has been shown to have a role in cancer stem cells. Reviewed by (Majmundar et al., 2010; Semenza, 2012; Simon and Keith, 2008).

2.2.2 Hypoxic regulation in human embryonic stem cells

Knockout studies in mice have demonstrated that HIF subunits are essential for embryonic development and survival (Iyer et al., 1998; Kotch et al., 1999; Maltepe et al., 1997; Peng et al., 2000; Ryan et al., 1998). In addition, hypoxic microenvironments occur in the developing embryo creating specific stem cell niches, which regulate cellular differentiation. Importantly, before the circulatory system is established, mammalian development occurs in low oxygen concentration. Later, oxygen gradient across growing embryo regulates formation of cardiovascular system. (Simon and Keith, 2008)

In order to improve maintenance of hESCs, Ezashi et al. were first to test hypoxic culture conditions. They reasoned that as naturally conceived embryos are exposed to 1.5-5.3 % oxygen, and because embryos are sensitive for oxidative damage, and blastocyst produced under low oxygen have significantly more ICM cells, hESCs could be cultured under hypoxic environment. They demonstrated that, hESCs cultured under 3-5 % hypoxia, contained statistically less differentiated cells in overgrown cultures analyzed at day 12. In addition, normoxia cultured cells had already produced significantly more chorionic gonadotropin (hCG) in the media after day five and progesterone (P4) after nine days, implying a more differentiated state. Surprisingly, hypoxia cultured hESCs had still enhanced capacity to form EBs (Ezashi et al., 2005). Similarly, Ludwig et al. reported as a side note, that an atmosphere of 5 % oxygen was optimal for undifferentiated hESC proliferation, when physicochemical environment and growth factor supplements were optimized for hESC culture maintenance (Ludwig et al., 2006). Forsyth et al. further supported the hypoxic culture of hESCs, reporting that culture in 2 % oxygen increased clonal recovery. Flow cytometric analysis revealed that hypoxia cultured cells were smaller and less granular. Further, hypoxic culture seemed to decrease the spontaneous chromosomal aberrations detected in early and late passage cultures (Forsyth et al., 2006).

The genome-wide gene expression changes between normoxia and hypoxia has been studied by three groups (Forsyth et al., 2008; Lengner et al., 2010; Westfall et al., 2008). Common to all studies is the observation that hypoxic culture does not have effect on the expression of *OCT4*, *NANOG* or *SOX2*. In addition, the number of differentially expressed genes between normoxia and hypoxia is relative low, varying from 150 to 350 genes that are mostly related to hypoxic metabolic regulation. However, hypoxia cultured cells have higher expression of certain pluripotency associated genes, such as *LEFTY2*, *ENDRB*, *SALL1*, *TRIM2*, *ZIC2*, *FGFR2*, *SERPINE1* and *KLF11*, whereas, some

differentiation associated genes, e.g. *HAND1*, *EOMES*, *H19*, *GATA6* and *MSX2,* have higher expression in normoxia (Westfall et al., 2008). Similarly, an increase of lineage markers is observed when hypoxia cultured cells are exposed to normoxia (Lengner et al., 2010).

Human ESCs can be cultured long-term (over 18 months) in moderate (5 %) hypoxia, but not in 1 % hypoxia (Prasad et al., 2009). Hypoxic culture has not been reported to affect the expression of surface antigens SSEA-1, SSEA-4, TRA-1-60, TRA-1-81 in normal passaging intervals (Forsyth 2006, Chen 2009, Forristal 2010). On the contrary, ALP staining has been observed to be less intensive in hypoxia (Chen et al., 2009). In addition, similarly to the transcriptome analysis no change in the expression of OCT4 has been detected (Chen 2009, Cameron 2008). On the contrary to all earlier reports, one study reported a significant increase of *OCT4*, *NANOG* or *SOX2* transcripts (RT-PCR) after long-term hypoxic (5%) culture (Forristal et al., 2010). Interestingly, Lengner et al. noticed a marked difference in the capacity of different cell lines to maintain expression of OCT4 and SSEA-4 in overgrown cultures between hypoxia and normoxia (Lengner et al., 2010), possibly explaining the differences observed between studies.

The activation of hypoxia response has been validated in two studies, based on which HIF1α is stabilized in 1% and 5 % oxygen in hESCs, but significantly reduced by the fifth and seventh day of exposure and undetectable after long-term culture (Cameron et al., 2008; Forristal et al., 2010). Furthermore, HIF2α and HIF3α are expressed in normoxia, but are induced and localized into the nucleus after the long-term hypoxic culture of hESCs (Forristal et al., 2010). Silencing of HIF2α or HIF3α, but not HIF1α, in hypoxia, leads to the decrease of OCT4, SOX2 and NANOG both at mRNA and protein level. In addition, HIF3α deficient cells can be maintained in culture, whereas HIF2α deficient cells fail to proliferate (Forristal et al., 2010). These results imply that HIF2α and 3α are vital for the long-term hypoxic culture of hESCs.

Interestingly, hypoxia also enhances generation of induced pluripotent cells in both human and mouse (Warren et al., 2010; Yoshida et al., 2009). More importantly, hESC lines derived in hypoxia, maintain the active X-chromosome state, which is lost only upon differentiation (Lengner et al., 2010), implying that hypoxic culture preserves hESCs in a more developmentally immature state. In addition, hypoxic culture has also been reported to enhance endothelium differentiation of hESCs (Prado-Lopez et al., 2010), generation of retinal progenitors (Bae et al., 2011) and hepatocyte-like cells (Si-Tayeb et al., 2010). To conclude, hypoxic culture in 2-5 % oxygen decreases spontaneous differentiation, and chromosomal aberrations, increases clonal recovery and enhances the capability to differentiate. HIF1α is activated in the initial response to hypoxia, but HIF2α is responsible for the survival of hypoxia cultured hESCs in long-term maintenance. However, none of the published studies has identified the molecular mechanism how hypoxia alters the pluripotency regulation of hESCs.

2.3 Genetic stability

2.3.1 Genetic aberrations and detection methods

An abnormal number of chromosomes relative to the reference genome is defined as chromosomal abnormality, aneuploidy. Standard cytogenetic analysis of the karyotype can be used to define the number and structure of the chromosomes. Karyotyping is commonly performed by G-banding, a method that utilizes Giemsa staining (Giemsa, 1904). The resolution obtained with standard karyotyping is 3-20 Mb that results from a Giemsa staining on metaphase spreads available for the microscopic analysis. More specific technique, fluorescence in situ hybridization (FISH) can be used to localize the presence or absence of specific DNA sequences (Langer-Safer et al., 1982). FISH utilizes specific labeled probes that bind only to the areas of DNA to which they show a high degree of sequence complementarity. The benefit of G-banding and FISH is the ability to detect a low level of mosaicism, a situation where a cell population consists of individual cells that have aberrant genome compared to the rest of the cell population (Baker et al., 2007).

The genome of an individual is heterozygous for gene alleles, containing maternal and parental gene copies. Loss of heterozygosity (LOH) is referred to a phenomenon where another allele is lost due to the point mutation or chromosome/gene deletion. Sometimes, loss of one allele leads to re-duplication of the remaining allele, leading to uniparental disomy **(Figure 7)**. LOH changes can cause phenotypic alterations by unmasking recessive alleles or oncogenes. Single base-pair changes in the genome, which are detected in over 1 % of the population, are referred as single nucleotide polymorphisms (SNP). Human genome contains millions of SNPs (http://www.ncbi.nml.nih.gov/SNP/). The SNP profiles, can be used to identify individuals and LOH changes. (Zhang et al., 2009a)

Copy number variations (CNVs) are DNA segments that have a gained copy number state (CN $\geq$ 3) or loss (CN 0 or 1) relative to that of normal reference genome (CN 2) **(Figure 7)**. CNV can affect only part of the single gene or be a large segment containing multiple genes. CNVs are widespread in human genome and are the source of genetic diversity. The Database of Genomic Variants lists 66 741 CNVs (Nov 2, 2010) varying in size from 1 kb to 1 Mb (http://projects.tcag.ca/variation/). Formation of CNVs affect genome by changing gene dosage or disturbing gene sequence creating polymorphic variants that drive genome evolution or rise new complex diseases, such as hemophilia, alzheimer and autism. New CNVs are generated by non-allelic homologous recombination (NAHR), nonhomologous end-joining (NHEJ), retrotransposition and fork stalling and template switching (FoSTeS). Large inverted CNV amplifications common for human cancers can be generated by breakage-fusion-bridge (BFB) cycle. (Hastings et al., 2009; Zhang et al., 2009a)

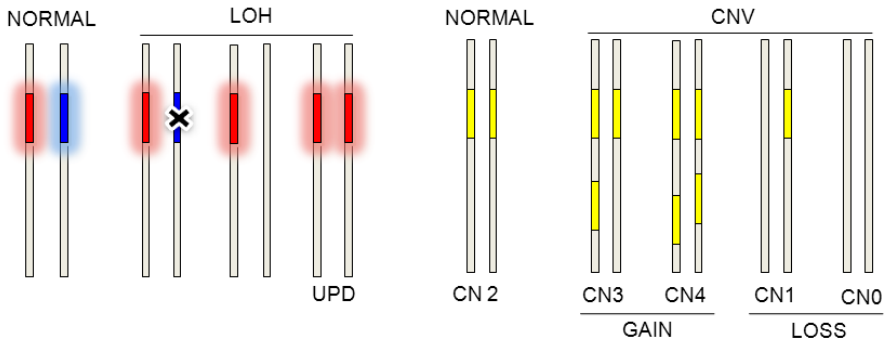

Figure 7. High resolution changes. Loss of heterozygosity (LOH) refers to deletion of maternal or paternal allele of the gene. Reduplication of remaining allele can lead to uniparental disomy (UPD). Copy number variation (CNV) refers to gain (CN ≥ 3) or loss (CN 0 or CN 1) relative to normal (CN 2) copy number of the genes.

High resolution genomic changes can be detected with comparative genomic hybridization (CGH), single nucleotide polymorphism genotyping and sequencing. The resolution achieved with CGH and SNP arrays depends on the number and genomic coverage of the probes. Current array formats can detect CNVs at kb resolution level, but are unable to detect balanced rearrangement and low level of mosaicism. Next-generation sequencing enables kb level resolution and detection of inversions by paired-end techniques. Alterations detected by all of these methods rely on normalization to the reference genome. (Sato-Otsubo et al., 2012)

2.3.2 Genetic integrity of pluripotent cells

The genomic constitution of an individual hESC line originates from the blastocyst used for its derivation. New lines are commonly derived from the pre-embryos considered suboptimal, surplus embryos. However, surplus embryos have a high rate of mosaicism and less than half that reach blastocyst stage are chromosomally normal (Hardarson et al., 2003). Astonishingly, high resolution analysis of normal human good quality blastomeres revealed a high frequency of chromosome instability, mosaicism and uniparental isodisomies, only 9 % being normal (Vanneste et al., 2009). These results indicate that chromosome instability is prevalent in human embryogenesis.

Considering current knowledge of genetic integrity of the human blastocyst, it is unanticipated that the first hESC lines were derived and cultured for extended times as karyotypically normal (Amit et al., 2000; Reubinoff et al., 2000; Thomson et al., 1998). In 2004, Draper et al. reported recurrent gains of chromosomes 17q and 12 in hESC lines (Draper et al., 2004), followed by individual studies describing aneuploidy and mosaicism of various chromosomes in continues cultures (Baharvand et al., 2006; Buzzard et al., 2004; Caisander et al., 2006; Cowan et al., 2004; Heins et al., 2004; Mitalipova et al., 2005; Rosler et al., 2004). Notably, multiple cell lines in these studies were also reported to maintain normal karyotype in extended cultures.

The increasing number of karyotypic changes detected lead to the identification of the most recurrent changes in hESCs, which are gains of chromosomes 12, 17 and, to a lesser extent, X (Baker et al., 2007). The global gene-expression meta-analysis of 38 hESCs and 66 hiPSCs suggested that amplification of chromosome 12 and 12p is also common in hiPSCs (Mayshar et al., 2010). The karyotypic meta-analysis of 552 hiPSC and 1163 hESC cultures revealed that trisomy 12 is a predominant aberration in human pluripotent cells (Taapken et al., 2011). Changes of 12p, i(20)(q10), 8 and X are common in both cell types, whereas trisomy 17 is specific for hESCs. In addition, the incidence for chromosome abnormality (~13 %) is identical for hESCs and hiPSCs. Despite the large number of samples, trisomy 17 has not been detected in hiPSCs (Martins-Taylor et al., 2011; Mayshar et al., 2010; Taapken et al., 2011). Mosaicism is also a common feature for cultures of pluripotent cells (International Stem Cell Initiative et al., 2011).

2.3.3 High resolution changes of pluripotent cells

Karyotyping has relative low resolution (3-10 Mb), leaving high resolution changes undiscovered. In addition, obtaining of high quality metaphases from hESC cultures for karyotyping has proven to be difficult. For these reasons alternative methods has been applied to study genomic integrity of hESCs. The first low resolution CGH analysis was used to verify karyotyping data (Inzunza et al., 2004). As resolution increased to the Mb level, a pilot CGH analysis identified four culture induced CNVs in hESCs (Maitra et al., 2005). Importantly, amplification encompassing *MYC* oncogene and cancer associated gain of 20q11.21 was detected, emphasizing the importance of higher resolution techniques in the quality assessment of hESCs. Later, additional CGH studies reporting numerous individual CNV sites were published (Chin et al., 2009; Elliott et al., 2010; Lefort et al., 2008; Spits et al., 2008; Werbowetski-Ogilvie et al., 2009; Wu et al., 2008). The rapid technical development of CGH arrays led to detection of genomic alterations of hESC and hiPSC genomes at kb resolution level and to the identification of hundreds of CNVs (Hovatta et al., 2010; Hussein et al., 2011; International Stem Cell Initiative et al., 2011; Laurent et al., 2011; Martins-Taylor et al., 2011; Narva et al., 2010).

Based on the CGH studies performed with different array formats on various hESC and iPSC lines, amplification of 20q11.21 is recurrent culture induced genomic hotspot site in pluripotent cells (Elliott et al., 2010; International Stem Cell Initiative et al., 2011; Laurent et al., 2011; Lefort et al., 2008; Maitra et al., 2005; Martins-Taylor et al., 2011; Narva et al., 2010; Spits et al., 2008; Wu et al., 2008). Other sites detected in multiple cell lines are amplification in the area of *NANOG* pseudogene *NANOGP1* (Chin et al., 2009; Laurent et al., 2011; Narva et al., 2010) and reprogramming or early culture induced changes in the areas of 1q31.3, 2p11.2 and 17q21.1 in hiPSCs (Martins-Taylor et al., 2011).

2.3.4 Functional consequences of genomic alterations

Karyotypically abnormal, culture adapted, cell lines express pluripotency markers, retain the ability to differentiate, but are easier to maintain due to the increased growth rate, clonal recovery and reduced apoptosis (Baker et al., 2007; Catalina et al., 2008; Draper et al., 2004; Herszfeld et al., 2006; Moon et al., 2011; Plaia et al., 2006; Sun et al., 2008; Werbowetski-Ogilvie et al., 2009; Yang et al., 2008; Yang et al., 2010). Abnormal hESCs have also reported to maintain an undifferentiated phenotype in the withdrawal of fibroblast growth factor (Herszfeld et al., 2006; Werbowetski-Ogilvie et al., 2009), have higher portion of cells in the S-phase (Yang et al., 2008), express a tumor necrosis factor receptor CD30, which is well established marker of human EC cells (Herszfeld et al., 2006), and undergo abnormal cell division with multiple spindles (Moon et al., 2011). In addition, some studies have indicated that the EBs formed from abnormal cells can have restricted potential to differentiate (Fazeli et al., 2011; Werbowetski-Ogilvie et al., 2009).

Abnormal hESCs are able to form teratomas containing differentiated cell types of all three germ layers, similar to normal diploid hESCs (Sun et al., 2008). However, teratomas of abnormal cells can constitute a greater proportion of undiffentiated cells (Herszfeld et al., 2006), contain immature OCT4 positive areas (Werbowetski-Ogilvie et al., 2009; Yang et al., 2008), have increased abundance for renal development (Gertow et al., 2007), contain additional complex chromosomal changes (Moon et al., 2011), or even lose the ability to form teratomas (Imreh et al., 2006).

It is important to note that genomic changes in hESCs affect gene expression (Baker et al., 2007; Enver et al., 2005; Mayshar et al., 2010; Mitalipova et al., 2005; Narva et al., 2010; Plaia et al., 2006; Werbowetski-Ogilvie et al., 2009; Yang et al., 2008; Yang et al., 2010). For instance, pluripotent cells that have amplification of 12p, significantly overexpress *NANOG* and *GDF3* relative to normal (Mayshar et al., 2010). Also high resolution CNVs alter gene expression. For example, the hotspot CNV gain of 20q11.21 passing *DNMT3B* leads to overexpression of the gene product (Martins-Taylor et al., 2011; Narva et al., 2010; Spits et al., 2008).

2.3.5 Epigenetic stability

Various culture conditions, derivation and reprogramming methods can affect the epigenome of hESCs and alter the phenotype and differentiation capacity of the cells. Indeed, the DNA methylation profiles of individual pluripotent cell lines are not identical (Allegrucci et al., 2007; Bock et al., 2011; International Stem Cell Initiative et al., 2007).

Alterations in DNA methylation pattern occur in prolonged culture (Allegrucci et al., 2007; Bibikova et al., 2006; International Stem Cell Initiative et al., 2011; Maitra et al., 2005). However, meta-analysis performed on a large number of samples has not identified epigenomic hotspot sites shared between cell lines (International Stem Cell Initiative et al., 2011). In addition, the X-chromosome inactivation status may alter in culture

(Silva et al., 2008). Part of the epigenomic changes obtained in culture are transmitted in normal maintenance of the cells and are inherited in differentiation (Allegrucci et al., 2007). Interestingly, karyotypically abnormal cell lines are not more prone for epigenetic abnormalities (International Stem Cell Initiative et al., 2011), implying that genetic instability does not increase epigenomic instability.

3. AIMS OF THE STUDY

The PhD work started from the setup of routine culture maintenance and analysis methods for hESCs. One of the main aims was to select a novel candidate gene that is highly expressed in hESCs, and thus potentially important for pluripotency maintenance, and identify its role in hESCs. In addition, the influence of hypoxic culture to the pluripotent state of hESCs was studied. Another main objective was to study the genetic stability of hESCs and hiPSCs by identifying high resolution changes with a new array based method.

The specific aims of this thesis were to:

- Characterise a novel protein L1TD1 in hESCs.
- Define the functional role of L1TD1 in the pluripotency of hESCs.
- Study the effect of hypoxic culture to the pluripotency of hESCs.
- Identify high-resolution genomic changes in hESCs.
- Study high-resolution genomic changes of the reprogramming process of hiPSCs.

4. MATERIALS AND METHODS

4.1 Ethical consideration

Human ESCs lines were maintained by the permission of Ethics Committee of South-West Finland Hospital District. Human iPSCs were derived and maintained with the permission of the ethics committee of HUS (Helsinki and Uusimaa hospital district). Research was carried out following the good scientific practice and guidelines of the National Advisory Board on Research Ethics.

4.2 Methods of pluripotency

4.2.1 Cell Culture (I, II, III, IV)

Human ESC lines were obtained from Outi Hovatta (Karolinska Institutet, Sweden) or from WiCell Research Institute (Madison, WI, US). The cells were maintained on 0.1% gelatin-coated (Sigma-Aldrich) plates on mitomycin C inactivated human foreskin fibroblasts (ATCC). The ES culture media consisted of DMEM-F12 (Stem Cell Technologies) supplemented with 20% Serum Replacement (Gibco), 2 mM glutamax (Gibco), 1% non-essential amino acids (Gibco), 50 U/ml penicillin-streptomycin (Gibco), 0.1 mM 2-mercaptoethanol (Gibco) and 4 ng/ml basic fibroblast growth factor (bFGF) (R&D Systems). In fibroblast-free culture conditions, the cells were cultured on Matrigel (BD Biosciences) and maintained in fibroblast conditioned ES culture media or mTeSR1 media (Stem Cell Technologies). Lines were passaged using type IV collagenase (Gibco) or mechanical cutting. In addition, cell lines were routinely karyotyped and analyzed for mycoplasma.

Human fibroblast lines were cultured in 10% FBS (PromoCell) and GlutaMAX in DMEM (Gibco). Human iPS cells were cultured on mitotically inactivated mouse embryonic fibroblasts (MEFs) in KnockOut DMEM supplemented with 20% KnockOut Serum Replacement (Gibco), 0.1 mM 2-mercaptoethanol (Gibco), 1× GlutaMAX (Gibco), 1× non-essential amino acids (Gibco), 1× ITS liquid media supplement (Sigma) and 6 ng ml^{-1} FGF2 (Sigma). Human iPS cells were passaged using 20 U ml^{-1} type IV collagenase (Gibco), approximately every 5 days.

Embryonal carcinoma cell lines 2102Ep and NT2D1 were obtained from Dr. Peter Andrews (University of Sheffield, UK). The cells were maintained on DMEM (Sigma) supplemented with 10% fetal calf serum (FCS) (PromoCell) and 2 mM l-glutamine (Sigma). TCam2 seminoma cell line was obtained from Dr. Jukka Westermarck (Turku Centre for Biotechnology, Finland). Cells were grown in RPMI 1640+GlutaMAX (Gibco) medium supplemented with 10% FCS (PromoCell) and 1% penicillin streptomycin

(Sigma). 2102Ep and TCam2 cells were passaged using 0.05% trypsin-EDTA and NT2D1 lines by scraping.

4.2.2 Human iPSC generation (IV)

Human foreskin fibroblasts (HFFs; CRL-2429, ATCC) and human lung embryonic fibroblasts (IMR90; CCL-186, ATCC) were reprogrammed to hiPSCs as previously described (Takahashi et al., 2007). Briefly, retroviral constructs pMXs-OCT4, pMXs-SOX2, pMXs-KLF4, pMXs-NANOG and pMXs-LIN28 were obtained by cloning the human cDNA encoding each of the factors into the pMXs retroviral vector. pMXs constructs were transfected separately into the 293-GPG packaging cell line (Ory et al., 1996) (10^6 cells per 100-mm-diameter culture dish) to produce retroviral supernatant. Fibroblast lines, seeded overnight, were infected twice with different, but equally mixed, combinations of viral supernatants (0.5 ml each supernatant, 4×10^5 cells per 60-mm-diameter dish), over the course of 2 days. The following day, the medium was changed to fibroblast medium. On day 4, infected cells were collected and reseeded on mitotically inactivated MEFs. Next day, the medium was changed to human ES cell medium containing FGF2. The medium was replenished every 2 days. At 20–30 days post transduction, depending on colony size, colonies with human ES-cell-like morphology were picked and expanded for further analysis.

For the *piggyBac*-transposon-generated hiPSC lines, HFF cells were seeded in 60-mm-diameter plates at a density of 4×10^5 cells per plate. After 24 h culturing, cells were trypsinized, and electroporated using a 100-µl tip and program number 20 in the Neon Transfection System (Invitrogen) with 250 ng each transposon construct, 500 ng PB-rtTA construct and 500 ng pCyL43 PB transposase plasmid, details (Kaji et al., 2009; Woltjen et al., 2009). After 24 h, the medium was supplemented with doxycycline (day 0) and was then changed to human ES cell medium after 48 h of transfection. Cells were fed every 2 days with doxycycline-containing medium (1.5 µg ml^{-1}) for 20–30 days. Doxycycline was removed one passage after picking hiPSC clones. Human iPSC colonies were picked and cultured as described above for retrovirus-derived hiPSCs. For sample collection and genomic DNA extraction, cells were scraped in collagenase or dispase (1 mg ml^{-1}) and centrifuged twice at a low speed to pellet the cells as small colonies and remove the majority of MEFs, which remain as single cells in suspension and are aspirated with the medium.

The produced hiPSCs were validated by standard procedures as described (Hussein et al., 2011).

4.2.3 Hypoxic culture (II)

In hypoxia experiments matrigel (BD) coated plates and defined commercial media mTeSR1 (StemCell Technologies) was used. Plates were transferred to a hypoxia incubator (Invivo2 400, Ruskinn technologies, UK), which allowed media change and

handling to be performed under hypoxic environment. The hypoxia balanced media was changed immediately when cells were introduced to hypoxia incubator. The oxygen balanced media was changed daily in both conditions. Harvesting of the hypoxia treated cells was performed in the hypoxia incubator after which samples were immediately frozen with dry ice.

4.2.4 Microscopy of the morphology (I, II)

The morphology of the colonies was observed and captured with SteREO Lumar V12 (Zeiss).

4.2.5 Karyotyping (I, II, III, IV)

Karyotyping was performed 3-4 days after plating. Firstly, 10µg of ethidium bromide (Q-Biogene) (1mg/ml) per ml of media was added to the culture plates and incubated for 30 min, 37 °C. 200 ng/ml of Colcemid (Gibco) was added and incubated 4 hours at 37 °C. Cells were trypsinized and centrifuged 5 min, 200 g. The resulting pellet was re-suspended and exposed drop-wise (5-10 ml) to warm (37 C) hypotonic solution (0.0375 M KCl) and incubated 25 min at 37 °C. Cells were fixed drop-wise (10-20 drops) by adding fresh fixative (75 % methanol, 25 % glacial acetic acid), mixed, incubated 5 min and centrifuged 5 min at 200g. The resulting pellet was re-suspended and fixed with dropwise addition of fixative (5 ml), incubated 30 min and centrifuged 5 min at 200g. The fixation step was repeated twice without incubation. The sample was stored in the fixative at -85 °C. The samples were sent for analysis at the Sheffield Diagnostic Genetic Services, Sheffield Children's NHS Trust (Sheffield, UK) or to the Department of Genetics, University of Turku (Turku).

4.2.6 hESC Differentiation (I)

Human ESCs were plated on Matrigel and the conditioned media was supplemented with 13.7 µM of retinoic acid (Sigma). The media was changed daily.

For spontaneous differentiation and embryonic body formation, hESCs were grown on uncoated plates without fibroblasts in normal ES culture medium. The medium was changed every three days.

4.2.7 RT-PCR (I, II)

The ribonucleic acid (RNA) was isolated using the RNeasy Kit (Qiagen). To eliminate genomic DNA from RNA samples, DNase I (Qiagen) on column digestion was included during the processing. The concentration of the samples was measured with a ultraviolet absorbance spectrophotometer (Nanodrop, Thermo Scientific). A second round of DNAse treatment was carried out for 500-1000 ng of total RNA with Amplification Grade DNase I (Invitrogen). To verify that no genomic DNA was present, negative

real-time polymerase chain reaction (RT-PCR) control was performed by measuring the levels of housekeeping gene EF1α. Subsequently, complementary DNA (cDNA) was prepared using a Superscript II kit (Gibco).

Gene expression levels were measured using the 7900HT Fast Real-Time PCR System (Applied Biosystems) using 2 µl of the template in a 10 µl reaction volume. The primers and probes used were designed using Universal ProbeLibrary Assay Design Center (Roche). The primers designed for the analysis were first validated to respond by standard curve estimation. All measurements were performed in duplicate in two separate runs, and repeated if necessary to produce four Ct (threshold cycle) values for each gene for which the standard deviation was below 0.5. The results were normalized with the expression values of housekeeping gene EF1α. ΔCt for each gene was calculated ΔCt = Ct(gene) – Ct(housekeeping gene). The primer and probe sequences used are presented in the **Table 3**.

Table 3. The primer and probe sequences used for RT-PCR and PCR. *These primers were used to detect mRNA in the immunoprecipitated RNA-fraction (Report I, Figure 4B). #These primers were used in the ChIP experiment (Report I, Figure 2B).

Target	Primer forward 5'-3'	Primer reverse 5'-3'	Probe
MYC	CACCAGCAGCGACTCTGA	GATCCAGACTCTGACCTTTTGC	34 (Roche)
EGLN3	CGAAGTGCAGCCCTCTTACG	TTTTGGCTTCTGCCCTTTCTT	61 (Roche)
GLUT1	GTGGGCATGTGCTTCCAGTA	AAGAACAGAACCAGGAGCACAGT	67 (Roche)
OCT4	AGCAAAACCCGGAGGAGT	CCACATCGGCCTGTGTATATC	35 (Roche)
NANOG	CCTGAACCTCAGCTACAAACAG	GCTATTCTTCGGCCAGTTGT	87 (Roche)
SOX2	ATGGGTTCGGTGGTCAAGT	GGAGGAAGAGGTAACCACAGG	19 (Roche)
EF1α	CTGAACCATCCAGGCAT	GCCGTGTGGCAATCCAAT	AGCGCCGGCTATGCCCCTG
L1TD1	TCCCACAAAAGGAAGAAATAAATC	GCTCTATGCTTTGAGTCTATTAGGG	69 (Roche)
*OCT4 **	AGCAAAACCCGGAGGAGT	CCACATCGGCCTGTGTATATC	69 (Roche)
*OCT4**	CTTCGCAAGCCCTCATTTC	GAGAAGGCGAAATCCGAAG	60 (Roche)
*OCT4**	CCTGTCTCCGTCACCACTCT	GGCACAAACTCCAGGTTTTC	52 (Roche)
*OCT4**	CTTTGAGGCTCTGCAGCTTAG	CTGCTTTGCATATCTCCTGAAG	69 (Roche)
*OCT4**	CCGCCGTATGAGTTCTGTG	CAGGCTGAGAGGTCTCCAA	57 (Roche)
DNMT3B	GGAAATTAGAATCAAGGAAATACGA	AATTTGTCTTGAGGCGCTTG	83 (Roche)
LIN28	AAGCGCAGATCAAAAGGAGA	CTGATGCTCTGGCAGAAGTG	23 (Roche)
L1TD1#	AGGTGACCTTGGGGTTCAG	TCCCCGGAAATCGCATTC	-

4.2.8 Western blotting (I, II)

Cells were lysed in lysis buffer (50mM Tris-HCl pH7.5, 150mM NaCl, 0.5% TX-100, 5% glycerol, 1% SDS, 1mM Na3VO4, 10mM NaF, 1mM PMSF) and incubated 10 min at +95 °C. Protein concentrations were determined with *DC* Protein Assay (Bio-Rad) after which 6xSDS sample buffer (0.5M Tris-HCl pH6.8, 28% glycerol, 9% SDS, 5% 2-mercaptoethanol, 0.01% bromophenol blue) was added. Lysates were electrophoresed on a 10% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gel and transferred to a nitrocellulose membrane. Membranes were incubated overnight at

+4 °C with primary antibodies and washed 3 x 10 min in TBST buffer. Horseradish peroxidase (HRP)-conjugated secondary antibodies were incubated for 1 hour and second washing step performed. For detection, ECL (Amersham Biosciences) reagent or Pierce developing solution (Pierce) was used. Detailed list of antibodies is presented in the **Table 4**.

Table 4. Antibodies used for western blot measurements

Target protein	Antibody	Company	Dilution
HIF1α	610959	BD Transduction Laboratories	1:3000
β-actin	A5441	Sigma-Aldrich	1:10 000
PHD1	NB100-310	Novus Biologicals	1:1000
PHD2	NB 100-137	Novus Biologicals	1:3000
PHD3	NB 100-139	Novus Biologicals	1:2000
HIF2α	NB-100-122	Novus Biologicals	1:200
MYC	NB600-302 (9E10)	Novus Biologicals	1:200
L1TD1	HPA028501	Sigma-Aldrich	1:1000
OCT4	sc-9081	Santa-Cruz Biotechnology	1:500
NANOG	sc-33759	Santa-Cruz Biotechnology	1:500
NANOG	AF1997	R&D Systems	1:1000
SOX2	MAB2018	R&D Systems	1:1000
GAPDH	5G4	HyTest Ltd	1:20 000
LIN28	ab46020	Abcam	1:1500
DNMT3B	ab13604	Abcam	1:150
pSMAD1/5/8	9511L	Cell Signaling	1:200
RHA	ab54593	Abcam	1:1000
PABP	sc-32318	Santa-Cruz Biotechnology	1:1000
Anti-rabbit-HRP	554021	BD Pharminged	1:10 000
Anti-mouse-HRP	sc-2005	Santa-Cruz Biotechnology	1:10 000

4.2.9 Immunofluorecence (I)

Mitotically inactivated fibroblasts were plated on microscopy coverslips placed in cell culture dishes. The hESCs were plated on the following day. Transfections of the cells with Lipofectamine 2000 were performed 24 h after plating. The immunocytochemistry was carried out 48 h after transfections. For the staining, cells were fixed with 4% paraformaldehyde for 10-20 min at room temperature. The cells were stained for surface markers, after which the cells were permeabilized for intracellular stainings with 0.1-1% Triton-X-100 for 20 min. After intracellular staining, incubation with DAPI (Invitrogen) (1 µg/ml) was performed. Fluorescence images were captured with Zeiss AxioVert 200M or with Leica TCS Sp2 confocal microscope (Spectral Physics). Co-location was analysed with Leica TCS confocal software. Detailed list of antibodies is presented in the **Table 5**.

Table 5. Antibodies used for immunofluorescence.

Target protein	Antibody	Company	Dilution
V5	P/N 46-0705	Invitrogen	1:50-200
AGO2	011-22033	Wako	1:50-200
LIN28	ab75483	Abcam	1:50-200
GW182	ab70522	Abcam	1:50-200
DCP1A	ab57654	Abcam	1:50-200
EEA1	ab70521	Abcam	1:50-200
TIA1	ab2712	Abcam	1:50-200
A555 anti-mouse IgG (H+L)	A21424	Invitrogen	1:50-200
A488 anti-rabbit	A21441	Invitrogen	1:50-200
A594 anti-rabbit	A11037	Invitrogen	1:50-200
A488	A11001	Invitrogen	1:50-200

RNase A treatment was performed on cell permeabilized 5 min with 0.5 % Triton X-100 in the presence of protease inhibitors, after which the cells were incubated in PBS containing 10 µg/ml RNase A (Qiagen) at RT for 20 min.

4.2.10 Flow Cytometry (I, II)

Cells were harvested from feeder free plates with trypsin. Trypsin was inactivated with media supplemented with 50% FCS and cells were filtrated following washings with cold buffer (D-PBS, 2% FCS, 0.01% atzide). Primary antibodies were incubated for 30 min at +4 °C, after which cells were washed. Secondary antibodies were incubated for 30 min at +4 °C. The cells were washed and run with FACS Calibur (BD) and analyzed with Cell Quest FACS diva or Cyflogic Version 1.0.2 (Perttu Terho CyFlo Ltd). Detailed list of antibodies is presented in the **Table 6**.

Table 6. Antibodies used in the flow cytometry measurements. (P.W.A) Peter W. Andrews, Centre for Stem Cell Biology and the Department of Biomedical Science, University of Sheffield, Sheffield, UK.

Target	Antibody	Company	Dilution
P3X	Sheffield	Gift from P.W.A	1:50
SSEA-3	Sheffield	Gift from P.W.A	1:50
SSEA-4	Sheffield	Gift from P.W.A	1:50
TRA-1-81	Sheffield	Gift from P.W.A	1:50
TRA-1-85	Sheffield	Gift from P.W.A	1:50
TRA-1-60	MAB4360	Millipore	1:100
SSEA-1	Sheffield	Gift from P.W.A	1:50
A2B5	Sheffield	Gift from P.W.A	1:50
TRA-2-54	Sheffield	Gift from P.W.A	1:50
A488 IgG	A11001	Invitrogen	1:200
FITC anti-mouse IgG + IgM	M30801	Caltag Laboratories	1:150

4.2.11 Immunoprecipitation (I)

Cells were washed twice with cold PBS and lysed into NP40-buffer (20-50mM Tris, 150 mM NaCl, 0.5 % Sodium Deoxylate, 0.5 % NP-40) containing PhosSTOP (Roche) and Complete EDTA-free (Roche) inhibitors. Lysates were treated in the presence or absence of 20-1000 µg/ml RNase A (Qiagen). Immunoprecipitation was carried out using Bio-Adembeads (PAG 0463) (Ademtech) or proteinG Dynabeads (Invitrogen) according to manufacturer's protocol.

4.2.12 RNA interference and transfection (I)

L1TD1 small hairpin RNA (shRNA) constructs were generated by cloning specific shRNA sequences into pSuper-GFP-Neo (Oligoengine) using *BglII* and *XhoI* cloning sites. The sequence for small interfering L1TD1 (siL1TD1 #3) was selected from the RNAi Codex shRNA database, others from siRNA Target Findersoftware of GeneScript. ShRNA oligos were synthesized by DNA Technology A/S. The L1TD1 shRNA sequences 1, 2, and 5 were selected for synthesizing siRNA oligonucleotides (Sigma) for the knockdown experiments. Transfections were performed according to the manufacturer's protocols using Lipofectamine 2000 reagent (Invitrogen) with a 150-200 nM concentration of siRNAs. hESC double transfections were performed two and three days after plating, using cells cultured in feeder free conditions. 2102Ep cells were transfected one day after plating. The sequences used are listed in the **Table 7**.

Table 7. Sequences used for RNA interference.

Target siRNA/shRNA	Sequence
siL1TD1 1	GCAAGGACGTATCAGCAATTA
siL1TD1 2	AGGACAGAGTTTCAGCAAATA
siL1TD1 3	GACAGAGTTTCAGCAAATAATC
siL1TD1 4	GCCGACCTTTCATCAGCAACA
siL1TD1 5	GAGATGAGTCATGATGAGCATA
siOCT4	AAGGAUGUGGUCCGAGUGUGG
siSOX2	AAAACCAAGACGCUCAUGAAG
siNANOG	AAGGGUUAAGCUGUAACAUAC
si-nontargeting 1	GCGCGCUUUGUAGGAUUCG
si-nontargeting 2	AAUUCUCCGAACGUGUCACGU
si-nontargeting 3	CCUACAUCCCGAUCGAUGAUG

For inducible knockdown of L1TD1, shRNA sequences 1 and 5 were cloned into pSUPERIOR.neo plasmid (Oligoengine) between *BglII* and *XhoI* cloning sites. To establish clonal cell lines with doxycycline-dependent inducible expression of shRNAs, NT2D1 cells stably expressing TetR3 (a generous gift from Dr. Jianliang Li, University of Sheffield) were transfected with L1TD1 shRNA constructs using Lipofectamine 2000 reagent. EC medium containing 3 µg/ml puromycin (Sigma) and 750 µg/ml G418 (Sigma) was used for selecting the clones that express the transgene. From the emerging

colonies, clonal sublines were expanded and screened for the most efficient inducible knockdown of the L1TD1 mRNA. The shRNAs were induced using doxycycline (Sigma) at a concentration of 1 µg/ml.

4.2.13 Vectors (I)

The open reading frame sequence of *L1TD1* was PCR amplified from cDNA prepared from hESC mRNA and cloned into the following plasmids and restriction sites: pET-20b(+), *XhoI* and *NcoI,* (Novagen); pEF6/V5-His-TOPO, ligated by TA-cloning, (Invitrogen); pCAGG-EGFP, *Age*I and *Xho*I, a gift from Dr. Peter Andrews (University of Sheffield, UK). The primers used are listed in the **Table 8**.

Table 8. Primers used for vector cloning

Vector	Primer 5'-3'	Primer 5'-3'
pET-20b(+)	CGCGCGCCATGGATTCTGATGTATCTACTAGTGT	CGCGCGCTCGAGAGGTATATTATTCCCCAGTAA
pEF6/V5-His-TOPO L1TD1	GCCATGTCTGATGTATCTACTAG	AGGTATATTATTCCCCAGTAAT
pCAGGEGFP	CGCGCGCTCGAGATGTCTGATGTATCTACTAG	CGCGCGACCGGTTGAGGTATATTATTCCCCAGT

4.2.14 Antibody generation (I)

To generate antibodies, L1TD1 was overexpressed as a [His]6-tagged protein in pET20b vector in *Escherichia coli* strain BL21(DE3)C43 (Avidis). After induction with 0.4 mM isopropyl-beta-D-thiogalactopyranoside (IPTG, AppliChem), the expressed protein was isolated from inclusion bodies, solubilised and purified with His-tag based Talon metal affinity resin (Clontech). Protein antigen was further purified by size separation on 10% SDS gel and extracted. The identity of protein product was verified by tandem liquid chromatography (LC-MS/MS). Purified L1TD1 protein was used as an antigen to generate rabbit polyclonal antibody. Immunization was performed by the BioGenes company (Berlin, Germany). Another antibody was produced by peptide immunization with a custom designed ISKERQRDIEERSR peptide and affinity purified by BioGenes.

4.2.15 Chromatin immunoprecipitation (ChIP) (I)

ChIP assays were performed as previously described by Li et al. (Li et al., 2003). Briefly, 500 µg of sonicated chromatin, 10 µg of antibodies, and anti-rabbit #112.04 or protein G #112.02 (Dynal Biotech) magnetic beads were used. Reverse cross-linking was performed at 65°C for 12 h. The DNA was treated sequentially with Proteinase K and RNase A, and purified (QIAquick PCR purification kit, #28706, Qiagen). Binding of L1TD1 was tested by PCR with Phusion® High-Fidelity DNA Polymerase (F-530L, Finnzymes) using primers upstream of *L1TD1* (**Table 3**). The antibodies used are listed in the **Table 9**.

Table 9. Antibodies used in the ChIP experiments

Target	Antibody	Company	µg/sonicated chromatin
NANOG	AF1997	R&D Systems	10 / 500
OCT4	sc-8628x	Santa Cruz	10 / 500
SOX2	#5024s	Cell Signaling	10 / 500
Rabbit IgG	#2729s	Cell Signaling	10 / 500
Goat IgG	sc-2028	Santa Cruz	10 / 500

4.2.16 Proliferation and colony forming assays (I)

The cell growth was monitored using CellTiter 96® Non-Radioactive Cell Proliferation Assay (Promega), with a 1 h solubilisation step. For colony formation assays, siRNA transfected cells were plated, cultured for 12 days, and stained with Crystal violet.

4.2.17 Protein-RNA crosslinking (I)

Protein-RNA crosslinking was carried out using the protocol described by Hafner et al. (Hafner et al., 2010). Cells were incubated 16-20 hours with 100-500 µM 4-thiouridine (Sigma) before UV crosslinking and harvesting. For RT-PCR analysis cDNA synthesis from RNA was prepared using 1st strand cDNA synthesis kit (Roche) with Random hexamer primers.

4.2.18 Analysis of the stem cell matrix data (I)

The expression of *L1TD1* was analyzed in the Stem Cell Matrix data (http://www.ncbi.nlm.nih.gov/geo/): accession code GSE11508 (Muller et al., 2008). Samples were preprocessed with the lumi-package of R (Du et al., 2008) utilizing the quantile normalization algorithm (Bolstad et al., 2003). The probe values were linked to the Ensembl genes (NCBI 36) and in cases where several probes were detected within the region of the same gene, the probe values were mean centered.

4.2.19 Transciptome analysis with Illumina BeadChips (I)

For expression analysis, RNA was isolated using RNeasy Kit (Qiagen). To eliminate DNA from RNA samples DNase I (Qiagen) digestion was performed. The Concentration of the samples was measured with Nanodrop (Thermo Scientific). Samples were prepared by the Finnish DNA Microarray Centre, at Turku Centre for Biotechnology, accordingly to manufactures protocol and hybridized on Illumina Human HT-12 v.3 Expression BeadChip. The data was normalized by using quantile normalization. R-package Limma was used for statistics. Statistically significantly differentially expressed genes were selected based on the filtering criteria for fold change FC $\geq$ 1.3 and for P-value $\leq$ 0.05.

4.2.20 Affymetrix Exon 1.0 ST analysis (II)

For expression analysis RNA was isolated using RNeasy Kit (Qiagen). To eliminate DNA from RNA samples, DNase I (Qiagen) digestion was performed. The concentration of the samples was measured with Nanodrop (Thermo Scientific). Samples were prepared by the Finnish DNA Microarray Centre, at Turku Centre for Biotechnology according to manufactures protocol and hybridized on GeneChip Human Exon 1.0 ST Arrays (Affymetrix).

The data was preprocessed with Chipster open source analysis platform (http://chipster.csc.fi/) using RMA method. The probes were re-annotated using the custom CDF package provided by Brainarray (Dai et al., 2005) and the expression signals were summarized at gene level based on Entrez gene annotations. The downstream data analysis was carried out with R/Bioconductor software (Gentleman et al., 2004; R Development Core Team, 2008). The minimum pearson's correlation values between replicates in each group were between 0.93 and 0.95 signifying good reproducibility. The statistical analysis for detecting the global differences in the gene expression between the groups was carried out using Bioconductor's Limma package (Smyth, 2005). For filtering out the differentially expressed genes, the minimum FC limit was set at 1.5 and the significance level at 0.01. Ingenuity Pathway analysis (http://www.ingenuity.com/) was used to study HIF1α pathway.

4.3 High resolution DNA studies

4.3.1 Affymetrix SNP 6.0 arrays (III, IV)

Human ESC plates were washed with cold D-PBS, harvested and frozen immediately. Later genomic DNA was extracted using QIAamp DNA Mini Kit (Qiagen). The concentration and quality of the samples was measured with a Nanodrop spectrophotometer (Thermo Scientific) and gel electrophoresis using reference DNA as a control. Samples were hybridized in the Finnish DNA Microarray Centre, at Turku Centre for Biotechnology using Genome-Wide Human SNP Nsp/Sty 6.0 protocol and SNP 6.0 arrays (Affymetrix).

4.3.2 SNP 6.0 array data analysis (III, IV)

The data was analysed using Affymetrix Genotyping Console 3.0.1 and the Birdseed v2-algorithm. Samples were normalized against 40 International HapMap samples (The International HapMap Consortium, 2003), which were also hybridized in house to decrease technical variation. Sample codes for HapMap samples used are presented in the Report III, Supplementary Table 1. For the CN analysis regional GC correction was used. For the area to be counted as CNV, at least 10 markers were required to be found within the region and the size of the region was determinate to be at least 50 kb (Report III) or10 kb (Report IV). Genotyping Console Browser (Affymetrix) was

used to illustrate changes detected. CNVs, in which the average distribution between markers was > 20 kb (Report III), > 40 kb (Report IV), were considered as false positive, in addition to CNVs affecting Y chromosome in female samples, and excluded from the analysis. False positive estimates were studied by hybridizing 3 different HapMap samples in four replicates. For further validation of CNVs, CNVs from three ES cell lines were also confirmed by running the same samples on an Illumina Human 610-Quad Chip platform. The Illumina Data were analysed for log Bayes factors greater than 10 using QuantiSNP software (http://www.well.ox.ac.uk/QuantiSNP).

The Ensembl (build 49) database was utilized to find the genes within the CNV areas (Hubbard et al., 2007). The genes were further linked to HGNC gene symbols (Eyre et al., 2006). To compare CNV profiles of hESCs to the normal human genome, an additional 90 CEPH samples (Caucasians, Utah residents with Northern and Western European ancestry from the CEPH collection) from the International HapMap Project (www.hapmap.org) were analysed with identical settings to our own. The CEPH samples were chosen because they represented best the same sample origin as the hESC lines used in the Report III.

R (v2.9.2) software and the program Microsoft Excel 2008 (v12.2.3) were used for *in silico* data analysis and CNV data parsing. R and StatPlus for Microsoft Excel (v5.8.3.8) were used for statistical analysis and *P*-value calculations.

4.3.3 Exon array analysis (III)

The probe values of the array were directly linked to Ensembl genes (build 49) (Hubbard et al., 2007) using alternative CDF-files, version 11 (Dai et al., 2005). The aroma. affymetrix package (Bengtsson et al., 2008) was used in analyzing the gene values of the expression measurements. RMA (Bolstad et al., 2003) was used for pre-processing the Exon array values.

4.3.4 Integration of genomic changes and gene expression (III)

To find the genes of which CNV is associated with increased or decreased gene expression level, we performed an integration analysis. First, we labelled the gene values into two groups; ‘gain’ and ‘no gain’. For each gene we computed a weight value (*W)*

$$W_G = (m_{G1}\text{-}m_{G0})/(\sigma_{G1} + \sigma_{G0})$$

where G is the gene in question, *mG1* and *σG1* denote the mean value and standard deviation of the gene expression values of the samples, in which the gene was found to be gained, and *mG0* and *σG0* the mean and standard deviation of the samples, in which the gain was not detected (Hautaniemi et al., 2004). To associate the lost CN values with the low gene expression values, we labelled the genes into groups ‘loss’ and ‘no loss’, respectively, and computed the weight value for the association between a loss in CN and a low gene expression value. Secondly, we obtained a p-value for the weight

value of each gene by performing 10,000 permutations (Hautaniemi et al., 2004). Thus, we could identify genes with significant association between CN and gene expression value. Thirdly, the resulting p-values were adjusted with Benjamini Hochberg's multiple comparison method (Benjamini and Hochberg, 1995). All the associations with over two fold change between the mean values of the expression levels of groups 'gain' and 'no gain', or 'loss' and 'no loss' and the adjusted p-value below 0.05 were considered to be significant (Jarvinen et al., 2006).

4.3.5 RT-PCR validation of CN states and expression (III)

To validate genomic CN (copy number) states, DNA from the original samples was used as a template. RT-PCR was performed as described in 4.2.7. The average results of the samples shearing gain (CN 3) or loss (CN 1) was compared to the samples of normal CN state (CN 2) for each gene studied. CN was counted real if the difference measured was in range of expected difference, 0.5 ΔCt for CN state 3 and 1 ΔCt for CN state 1. The two tailed t-test was counted for each result and required to be under 0.05 (*), 0.01(**) or 0.001(***). The primers and probes used are listed in the **Table 10**.

Table 10. Primers and probes used for validation

Target	Primer forward 5'-3'	Primer reverse 5'-3'	Probe
GAPDH	acacccactcctccaccttt	tgacaaagtggtcgttgagg	45 (Roche)
DNMT3B	tgtaatccagtgatgattgatgc	ggtaggttgccccagaagtat	84 (Roche)
RHOJ	gatgagctacgccaacgac	gcatagtggtcaaacacagtgg	6 (Roche)
CTCFL	gtgagaagcctcacctgtgtc	cgcagcagagtgaccgta	13 (Roche)
EGR2	gggtgtgtgcaccatgtc	ggtggcggagagtacaggt	85 (Roche)
MAGEA4	ccaatgagggttccagca	aacaaggactctgcgtcagg	35 (Roche)
ZNF613	ggcaacctccttattcatcg	agcctttcccacattcattg	47 (Roche)
ID1	ccagaaccgcaaggtgag	ggtccctgatgtagtcgatga	39 (Roche)
REV1	ccgggaacaagtagagcaag	tttttgtcgccatgtgactc	56 (Roche)
JARID2	ttcgctcaggaaaaagaagtg	agtcattgaggacgcctttg	63 (Roche)
TNFSF15	acagccagtgtggaaatgct	ccaggcagcaggtgagag	68 (Roche)
JMJD1C	gcaaactggggaatcctttt	ttctcgacacttttgtaaattaggc	18 (Roche)
GOLGA8B	tggcttatttccgaggaatg	caaatgctctaagctaggaaaggt	76 (Roche)

4.3.6 Three-dimensional FISH (IV)

Human iPSCs were cultured on glass slides seeded with MEF cells. Samples were fixed in 2% paraformaldehyde in PBS for 5 min, washed three times with PBS, permeabilized with 0.5% Triton X-100 in PBS for 20 min, and washed three more times with PBS. The slides were then placed in a solution of 20% glycerol in PBS overnight at 4 °C. Slides were frozen in liquid nitrogen, allowed to partly thaw and then placed back into the 20% glycerol solution. This process was repeated five times. After the freeze–thaw procedure, the slides were washed three times in PBS and then placed in a solution of 0.1 M HCl for 5 min. Slides were then washed with 2× SSC and left overnight at 4 °C in a solution

of 50% formamide in 2× SSC. Before hybridization, the slides were denatured in a solution of 70% formamide in 2× SSC at 75 °C for 3 min and then immediately placed in a separate container containing the same denaturation solution that had been kept on ice. Control (Bac clone RP11-788E9) and test (Bac clone RP11-58E1) probes were obtained from the Centre for Applied Genomics (Toronto). Test and control probe region coordinates were chr1: 146,828,351–147,150,258 and chr1: 104,629,600–104,808,778, respectively, based on the human genome assembly of March 2006 (NCBI36/hg18). The test probe was selected based on a cluster of CNVs consisting of mainly deletions within a frequently affected region in chromosome 1 (coordinates Chr1: 145,797,568–147,958,358). The probes were directly labelled with either spectrum green or orange fluorophore-conjugated nucleotides. A hybridization mixture consisting of labelled probe and human Cot-1 DNA in a 2/1 ratio in hybridization buffer (50% formamide, 10% dextran sulphate, 50 nM sodium phosphate buffer, pH 7.0, in 2× SSC) was prepared and denatured at 80 °C for 5 min and then allowed to partially re-anneal at 37 °C for 20 min. This mixture was then applied to the slides that had been kept on ice during the previous step and left to hybridize overnight at 37 °C. After hybridization, the slides were washed in 50% formamide in 2× SSC three times at 42 °C, then once in a solution of 0.5× SSC at 60 °C, and finally in a solution of 2× SSC at room temperature. Slides were mounted with VECTASHIELD containing DAPI (Vector Laboratories) before fluorescence imaging. Images were collected using an IX81 inverted brightfield microscope (Olympus) equipped with a Cascade 512 camera (Photometrics) using a ×60, 1.32 NA, oil-immersion objective and Immersion Oil Type DF (Cargille Labs) imaging medium. Images were collected using MetaMorph Premier 7.7 (Molecular Devices) and analysed with ImageJ (National Institutes of Health).

5. RESULTS AND DISCUSSION

5.1 L1TD1 is essential for the pluripotency maintenance of hESCs (I)

5.1.1 L1TD1 expression is specific for hESCs and for certain cancer types

In order to find novel factors contributing to the regulation of hESC biology, a search for stem cell specific candidates was performed by analyzing array data of 21 hESC lines and differentiated counterparts (Lund et al. manuscript in preparation). From the resulting list of candidates, an uncharacterized hypothetical protein FLJ10884 (L1TD1) was selected for further studies. Similarly, L1TD1 (ECAT11, embryonic stem cell associated transcript) was originally identified as a stem cell specific transcript in mESCs with the digital differential display (Mitsui et al., 2003). Moreover, transcriptome profiling comparing SSEA-3 positive and negative hESCs identified L1TD1 as one of the genes having the highest expression difference between these subgroups (Enver et al., 2005) supporting our findings. Later, genome-wide microarray screens also identified L1TD1 among consensus hESC genes (Assou et al., 2007).

The RT-PCR validation experiments showed that *L1TD1* was specifically expressed in pluripotent undifferentiated hESCs and embryocarcinoma cell lines and rapidly declined in differentiation (Report I, Figure 1A, S1D). We also re-analyzed the transcriptome data of Stem Cell Matrix data (Muller et al., 2008) and verified that *L1TD1* had highest expression difference relative to EBs and fibroblast compared to that of *OCT4*, *NANOG* and *SOX2* and was highly expressed also in hiPSCs (Report I, S1A). In addition, Wong et al. 2011 recently published that *L1TD1* is a marker for undifferentiated hESCs (Wong et al., 2011).

In order to characterize the protein behavior, an antibody was generated and validated (Report I, Supplementary data (S1C). The protein behavior of L1TD1 was in line with the RNA expression and followed same kinetics in differentiation as known core factors of pluripotency (Report I, Figure 1B, C, D). Based on these results, L1TD1 is an excellent marker to characterize undifferentiated status of hESCs both at the RNA and protein level.

In addition, we identified *L1TD1* to be expressed in ovarian germ cell tumors, colorectal carcinomas, Ewings sacromas, non-seminomas and prominently in seminomas: testicular germ cell tumors, based on the In Silico Transcriptome Database (Kilpinen et al., 2008) (Report I, Figure 3A). The expression of L1TD1 at the protein level was validated in non-seminoma cell line 2102Ep and in seminoma cell line TCam2 (Report I, Figure 3B). These results imply that L1TD1 could be a diagnostic marker for these cancers.

5.1.2 Inhibition of L1TD1 leads to differentiation and loss of self-renewal

To identify if L1TD1 has role in maintaining pluripotency, shRNA sequences were validated to silence expression of L1TD1 (Report I, S1B). As optimization of transfection of hESCs proved siRNAs much more efficient, the siRNAs were ordered with optimized shRNA sequences. Depletion of L1TD1 rapidly decreased expression of OCT4 and NANOG, both at the protein and RNA level (Report I, Figure 1E, G), proving that L1TD1 has important role in the maintenance of pluripotency. The effect of L1TD1 silencing to SOX2 varied in early timepoints, but was constantly decreased after long-term silencing. In addition, levels of TRA-1-60 and SSEA-3 started to decline, whereas A2B5 and SSEA-1 levels increased (Report I, Figure 1F) similarly as in the case of OCT4 silencing (Report I, S2), suggesting the onset of differentiation. Although the siL1TD1 samples were RT-PCR analyzed for various germ layer markers, we could not detect significant induction in any of the markers, implying that the cells had lost pluripotency but not directed to lineage specific differentiation in the early timepoints used for the analysis.

The inconsistent behavior of SOX2 could possibly be explained by the indirect effect following L1TD1 silencing. That is, if L1TD1 would regulate OCT4 and NANOG directly, but not SOX2, the levels of SOX2 would be affected only due to the onset of differentiation. It is also known that the expression profile of SOX2 is not as stem cell specific as for OCT4 and NANOG (Assou et al., 2007)(Assou et al., 2009), and SOX2 is expressed in addition to ICM in the differentiated extraembryonic ectoderm of the mouse emryo (Avilion et al., 2003), which could explain slower responsiveness. Recently, it was published that SOX2 is not constantly decreased upon silencing of OCT4 or NANOG either. In addition, SOX2 was found redundant for pluripotency maintenance (Wang et al., 2012).

As OCT4 and NANOG are highly expressed in seminomas (Ezeh et al., 2005), we studied whether silencing of L1TD1 could also affect levels of these transcription factors in these cancer cell lines. Similarly, as in hESCs, the levels of OCT4 and NANOG were reduced (Report I, Figure 3B). Interestingly, Tcam2 cells do not express SOX2, implying that SOX2 is not needed to mediate the functional effects of L1TD1. Moreover, silencing of L1TD1 in NT2D1 cells with inducible system following array screen, mainly affected genes involved in proliferation (Report I, S1E). Consistently, reduced colony formation and proliferation was observed following L1TD1 depletion in non-seminomas and seminoma cell lines (Report I, Figure 3C, S1F). These novel results strongly suggest that L1TD1 has significant role in self-renewal and malignant transformation.

5.1.3 L1TD1 is an RNA-binding protein regulated by core pluripotency factors

L1TD1 codes for a relative large protein (865 amino acids), composed of N-terminal coiled coil (CC) domain, an RNA recognition motif (RRM), and of C-terminal domain

(CTD) (Report I, Figure 4A). Based on the structure, L1TD1 (LINE 1 type transposase domain containing 1) belongs to the same family of RNA binding proteins (RBPs) and has high homology of the RRM-motif with the LINE-1 element ORF1 protein (ORF1p) (Report I, S3). ORF1p is able to bind single stranded RNA and DNA with high affinity (Hohjoh and Singer, 1997; Kolosha and Martin, 1997) and is required for retrotransposition (Moran et al., 1996), but its exact function remains unsolved. These facts strongly suggested that L1TD1 could also bind RNA, which we were able to validate experimentally (Report I, Figure 4B).

The highly stem cell specific expression of L1TD1 suggested that it is regulated by stem cell specific factors. In line with this conclusion, we noticed from the Supplementary data of the Chip-on-Chip study (Boyer et al., 2005), that NANOG and SOX2 would bind to the promoter region of *L1TD1*. To test this, we performed Chip-PCR analysis for all core factors and found NANOG, OCT4 and SOX2 to bind to the promoter of *L1TD1*. In addition, the silencing of any of these factors lead to the decreased expression of *L1TD1*, further validating that they regulate expression of *L1TD1*. Similarly, Wong et al. reported in their study that NANOG regulates promoter of L1TD1 (Wong et al., 2011). Furthermore, the promoter region of L1TD1 is marked with active histone code (Zhao et al., 2007).

5.1.4 L1TD1 is co-localized in P-bodies with LIN28

To determine the cellular location of L1TD1, fusion proteins were generated. This was done already before antibody generation. The V5- and Enhanced Green Fluorescent Protein (EGFP)-tagged fusion proteins were found to be expressed as cytoplasmic round condensates (Report I, S4C, D). Later, the anti-serum produced in the rabbit verified these findings (Report I, S4A, B, E).

To get closer to the function of L1TD1, we performed a correlation analysis for the Stem Cell Matrix data (Muller et al., 2008). Surprisingly, expression of *L1TD1* correlated highest with stem cell associated transcripts *DNMT3B*, *LIN28*, *TDGF1* and *ESRP1*. Out of these LIN28 caught our attention for being a stem cell specific RBP that was reported to localize into processing bodies (P-bodies) and stress granules (SGs) (Balzer and Moss, 2007) with a similar appearance in the immunostainings as we had detected with L1TD1. The following co-staining of L1TD1 verified co-location with LIN28 and a P-body marker AGO2 (Sen and Blau, 2005) (Report I, Figure 5A). Partial co-location was also detected with P-body markers DCP1A and GW182 and SG-marker TIA1, but not with endosome marker EEA1 (Report I, S4F, G, H, I). As P-bodies are dependent on RNA for their formation (Teixeira et al., 2005), a RNase A treatment was further used to validate location of L1TD1, resulting loss of L1TD1 aggregates (Report I, Figure 5B).

5.1.5 L1TD1-RHA-LIN28 complex regulates OCT4

Due to the similar expression, location and RNA-binding capacities, we performed immunoprecipitation studies to determine if LIN28 interacts with L1TD1. As a result, LIN28 was found to immunoprecipitate with L1TD1, but this interaction was lost if RNA interactions were removed (Report I, Figure 5C, E). No interaction was detected with AGO2 (Report I, Figure 5D). Interestingly, the location of LIN28 and AGO2 was disturbed, if L1TD1 was silenced (Report I, Figure 5F, G). As we had already noticed in the initial silencing experiments that the RNA levels of LIN28 were highly decreased after L1TD1 silencing, this was not surprising. The loss of AGO2 could be possibly explained by instability of dynamic P-body structures, which are dependent on RNA and RNA-binding proteins. The rapid loss of highly expressed P-body proteins L1TD1 and LIN28 from hESCs could disturb the structure of P-bodies and location of AGO2 in these structures.

As LIN28 has been shown to regulate protein levels of OCT4 by binding directly to the mRNA of OCT4 and enhancing the translation through interaction with RHA (Qiu et al., 2010), we probed the L1TD1 immunoprecipitated RNA-fraction with multiple probes for OCT4 to detect if L1TD1 could also bind to the mRNA of OCT4. The result was negative. This result does not, however, rule out that OCT4 mRNA might be present in the RNA-fraction, because the short RNA-fragments in the immunoprecipitate may not be long enough for RT-PCR determination. However, we detected direct interaction with L1TD1 and RHA (Report I, Figure 5H) as reported for LIN28 (Qiu et al., 2010), supporting the idea that L1TD1 is part of the complex regulating translation of OCT4. Moreover, the depletion of L1TD1 affected in addition to OCT4, NANOG and SOX2, the protein levels of LIN28 and RHA (Report I, Figure 5I), supporting the functional cooperation between these proteins. Further, based on our findings (Report I, Figure 1D) LIN28 declines much later in differentiation than L1TD1 and OCT4. This predicts that if L1TD1 maintains levels of OCT4, it would have regulatory role in the L1TD1-RHA-LIN28 complex. The current hypothesis of L1TD1 function in hESCs is illustrated in the **Figure 8**.

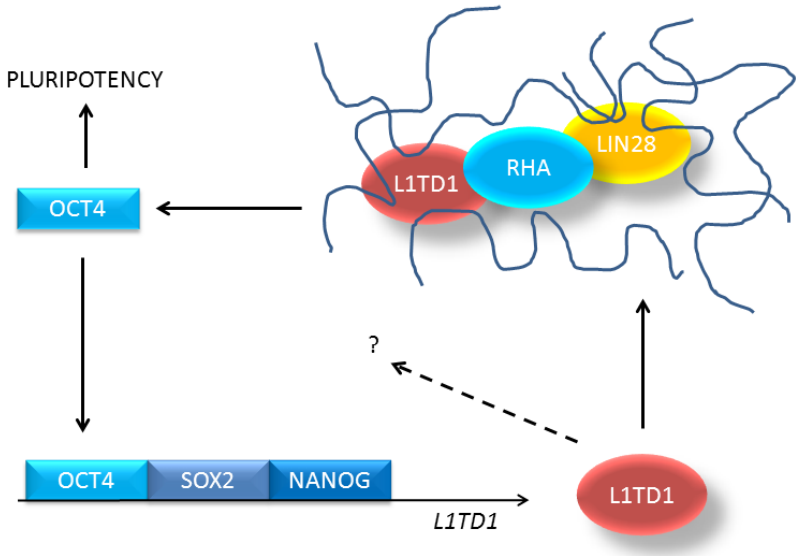

Figure 8. The current hypothesis how L1TD1 regulates pluripotency of hESCs. The core regulators of pluripotency: OCT4, SOX2 and NANOG directly regulate expression of the RNA binding protein L1TD1. The depletion of L1TD1 immediately decreases levels of OCT4, NANOG and LIN28 shifting cells towards differentiation. LIN28 has been shown to regulate translation of OCT4 by association with RHA. As L1TD1 directly binds to RHA and through RNA with LIN28, L1TD1 would be part of this complex and participate in the translation of OCT4, which maintains the pluripotency of hESCs. Furthermore, L1TD1 may influence the regulation of pluripotency by additional ways.

In addition, LIN28 has been shown to inhibit formation of mature let-7 family miRNAs, based on which, one could also expect L1TD1 to participate this interplay. Mouse ESCs express mature let-7 family miRNAs only after differentiation when LIN28 is absent. This has been intensively studied in mESCs and cancer tissues (Heo et al., 2009; Newman et al., 2008; Piskounova et al., 2008; Rybak et al., 2008; Viswanathan et al., 2008; Wulczyn et al., 2007). We measured the mature let-7 miRNA levels in scramble and siL1TD1 treated hESCs and noticed to our surprise that mature-let7 miRNAs were highly expressed already in undifferentiated hESCs (unpublished data). This is also clear from a number of miRNA profiling studies performed on hESCs (Morin et al., 2008; Ren et al., 2009; Suh et al., 2004) and our own sequencing data obtained from undifferentiated hESCs (unpublished data). As LIN28 is highly expressed in undifferentiated hESCs along with mature let-7 miRNAs, LIN28 does not inhibit maturation of let-7 miRNAs in hESCs. Similarly silencing of L1TD1 did not affect to the levels of mature let-7 miRNAs (data not shown). However, L1TD1 can have a role in the regulation of other miRNAs.

Unexpectedly, the L1td1-null mice were recently reported to grow and breed normally. In addition, despite specific expression in mESCs and activation in the induction of pluripotency, L1td1 deficient mESCs proliferated and maintained pluripotency (Iwabuchi et al., 2011). Thus, the L1td1 deficiency results in mice are in total contradiction to what

we have reported in human. This could be explained by differences in the regulation of pluripotency between mouse and human, or due to the capacity of L1TD1 orthologs to bind RNA. The RRM and CTD share 71 % and 57 % of amino acids respectively, while the overall identity between orthologs in only 45 %. Noteworthy, regulation of pluripotency differs highly between human and mouse. For example, while LIF- and BMP-signaling are essential for mESCs, hESCs rely on SMAD2/3 and FGF-signaling and differentiate in the presence of BMP (Ng and Surani, 2011). Moreover, depletion of Lin28 in mESCs does not affect OCT4 protein levels as it does in hESCs (Jin et al., 2011; Qiu et al., 2010). Thus, although L1td1 is not essential for mESCs, it can still be a master regulator in the hESCs.

Interestingly, L1TD1 may also have role in the genomic stability of hESCs. An average human genome contains 80-100 retrotransposition competent (RC) Long interspersed element-1 (LINE-1, L1s) sequences (RC-L1s) that contain two open reading frames, ORF1 and ORF2, whose protein products (ORF1p and ORF2p) are required for retrotransposition of LINE-1 elements into new genomic sites (Babushok and Kazazian, 2007; Feng et al., 1996; Moran et al., 1996). Interestingly, retrotransposition is active in hESCs and activated in reprogramming (Garcia-Perez et al., 2007; Wissing et al., 2012). As L1TD1 has high homology with ORF1p, further studies are needed to resolve whether L1TD1 also has a role in the activation of retrotransposition. Further, as ORF1p also binds to DNA, L1TD1 may also carry this property. In addition, ORF1p trimerizes via its N-terminal CC-domain (Khazina and Weichenrieder, 2009) suggesting that trimerization could be important also for the function of L1TD1.

As RBPs commonly are multifunctional regulators, L1TD1 most likely has additional ways to influence the maintenance of pluripotency. In order to obtain more information, multiple follow up studies are ongoing. As silencing of L1TD1 rapidly shifts cells to differentiation, influencing multiple properties immediately, we need to study L1TD1 in undifferentiated cells. The sequencing and mass-spectrometry identification of RNA and protein binding partners will illuminate the additional roles of L1TD1. In addition, the effect of L1TD1 to the miRNA biology and global transcriptome needs to be studied. As depletion of L1TD1 disturbs location of AGO2, which has important role in small RNA regulation, it is interesting to think how this directly affects pluripotency. Moreover, the clinical significance of L1TD1 in cancer tissues needs to be resolved.

To conclude, we have here characterized for the first time the protein behavior, RNA binding capacity, location and functional significance of L1TD1 to the pluripotency of hESCs. In addition, we have shown that proliferation of non-seminoma and seminomas is dependent of L1TD1 expression. Moreover, the expression of central core factors of pluripotency OCT4 and NANOG are affected immediately in L1TD1 depletion, suggesting that L1TD1 has crucial role in the maintenance of pluripotency in human cells. These results have important impact for the stem cell and cancer biology as L1TD1 can be used as a marker to identify self-renewing human pluripotent cells.

5.2 Effect of hypoxia to the pluripotency of hESCs (II, unpublished)

At the beginning of this study, nothing was known of the hypoxic regulation in hESCs. The culture optimization studies had reported that the maintenance of hESCs in hypoxic conditions (2-5 % oxygen) was favorable due to the lower level of spontaneous differentiation and increased proliferation (Ezashi et al., 2005; Forsyth et al., 2006; Ludwig et al., 2006). This study was designed to study the hypoxia response and resolve which factors were responsible for the improved pluripotency maintenance in the hypoxic environment.

5.2.1 Hypoxia response in hESCs

As none of the earlier studies had validated the onset of hypoxia in hESCs cultures, we first analyzed the protein level stabilization of HIF1α in 4 % hypoxia. As a result we detected HIF1α stabilization at all the timepoints (2h, 24 h and 7d) studied (Report II, Figure 1A). The protein level expression was highest at 24 h, but was clearly decreased after 7 days. This is in agreement with two other more recently published studies (Cameron et al., 2008; Forristal et al., 2010), which reported the total absence of HIF1α in long-term hypoxic culture of hESCs. The RT-PCR validation of HIF1α target gene *GLUT-1* was in agreement with its regulator (Report II, Figure 1D). Moreover, the array profiling and pathway analysis revealed that number of HIF1α responsive genes was highly induced after 2h and 24 h whereas smaller number of targets remained activated in long-term hypoxia (Report II, Figure 3B).

In addition, we detected for the first time that hESCs express stably all PHD isoforms PHD1, PHD2 and PHD3 at the protein level both in normoxia and hypoxia (Report II, Figure 1B). Based on array data, PHD1 also had stable expression at the RNA level. PHD2 was slightly induced, whereas PHD3 was significantly elevated after 24 hours. The induction of PHD3 was validated with RT-PCR (Report II, Figure 1C). As PHD2 and PHD3 has been shown to restrict the excessive HIF activity in long-term hypoxia, reviewed by (Jokilehto and Jaakkola, 2010), PHDs are most likely responsible for the downregulation of HIF1α in long-term hypoxic culture of hESCs. Moreover, PHD3 has been reported to support survival and G_1 to S transition of carcinoma cells in hypoxia (Hogel et al., 2011). As hESCs are characterized by short G_1-phase, PHD3 may have an important role in the regulation of cell cycle and survival of hESCs in hypoxia and possibly also in normoxia.

5.2.2 Hypoxia induced changes in hESCs

In agreement with earlier reports, the morphology of hESC colonies was less differentiated compared to normoxic culture conditions observed at day seven (Report II, Figure 2A). In addition, we did not observe changes in the expression levels of OCT4, NANOG or SOX2 at protein or mRNA level (Report II, Figure 2H,I), similarly to other reports (Cameron et al., 2008; Chen et al., 2009; Forsyth et al., 2008; Westfall et al., 2008), implying that

the improved pluripotency maintenance of hESCs in hypoxia is not directly regulated by these factors. Similarly, we did not detect significant improvement in the levels of SSEA-4, TRA-1-81 or TRA-1-60 (Report II, Figure 2C), which is also in agreement with earlier studies (Chen et al., 2009; Forristal et al., 2010; Forsyth et al., 2008). However, we detected that levels of SSEA-4, TRA-1-81 and TRA-1-60 were consistently higher and differentiation markers SSEA-1 and A2B5 consistently lower in hypoxia (Report II, Figure 2C,D). Furthermore, we were first to detect levels of SSEA-3, which were significantly higher in hypoxia (Report II, Figure 2B,C,E). This change is important as the silencing of core pluripoteny factors induces a similar decrease (15-20 %) on SSEA-3 expression (Fong et al., 2008; Hyslop et al., 2005; Matin et al., 2004; Zaehres et al., 2005) (Report I, Figure 1F, S2). On the contrary, we observed that TRA-2-54, a marker of alkaline phosphatase, was decreased in hypoxia. Lower levels of alkaline phosphatase have also been reported earlier in the immunostaining of hESCs (Chen et al., 2009) and other cell types (Utting et al., 2006) in hypoxic environment. The decrease of TRA-2-54 in the same cells carrying enriched SSEA-3 expression, implies that the surface markers expression validated for undifferentiated cell types in normoxia, may differ in hypoxic conditions.

To solve which factors were responsive for hypoxic conditions, we analyzed the transcriptome of hypoxia and normoxia cultured cells. As a result, we detected induction of hypoxia responsive genes *HK2*, *PFKFB4*, *PDK1*, *DDIT4*, *BHLHE40* and *INSIG2* in all the time points studied. The number of statistically upregulated (193) and downregulated (96) hits was relative low (Report II, Figure 3A), similarly to two earlier studies (Forsyth et al., 2008; Lengner et al., 2010; Westfall et al., 2008). When compared to consensus stem cell and differentiation genes (Assou et al., 2007), no hits were detected at the 2 h timepoint. However, after 24 hours 11 %, and after 7 days 13 %, of the changed genes were among the stem cell associated transcripts (Report II, Figure 3C). Of relative interest, we detected an increase of *LEFTY2* and decrease of *MSX2,* a marker of trophectoderm, which were also detected in the earlier array screen (Westfall et al., 2008). The nodal inhibitor, LEFTY2 has been shown to be highly expressed in undifferentiated hESCs and regulated by SMAD2/3, but inhibited at translational level by miR-302 cluster miRNAs (Barroso-delJesus et al., 2008; Besser, 2004). In addition, *HEY2,* a target of notch-signaling (Katoh and Katoh, 2007), was induced and differentiation markers *PAX6* and *FLRT3* decreased in hypoxia. Moreover, we detected an increase of *MYC* after 24 hours, which was further induced at day seven. The induction of MYC was validated at the RNA and protein level (Report II, Figure 4A,B,C). Interestingly, while HIF1α has been shown to antagonize MYC function, HIF2α promotes MYC activity in other cell types (Gordan et al., 2007). In addition, HIF2α has been shown to be expressed and vital for long-term hypoxic culture of hESCs (Forristal et al., 2010; Westfall et al., 2008). Therefore, we studied if the induction of HIF2α was correlating with induction of MYC. This was clear in embryo carcinoma cells used for the validation (Report II, Figure 4D,E). Further studies are needed to validate HIF2α and MYC expression kinetics in hESCs.

MYC acts as a master regulator having thousands of binding sites in the genome (Seitz et al., 2011). In addition, some degree of MYC overexpression is estimated to occur in

70 % of the human cancers. Moreover, Myc expression is vital for the self-reneval of mESCs (Cartwright et al., 2005). In hESCs, MYC has been shown to be specifically expressed in undifferentiated hESCs (Assou et al., 2007; Sumi et al., 2007), identified as one of the central nodes in the hESC enriched protein interaction network (Zuo et al., 2009), and known to highly induce the generation of hiPSCs (Nakagawa et al., 2008). However, overexpression of chimeric c-MycER protein in hESCs for five days has been reported to lead to apoptosis and differentiation, which was seen as a decrease of *OCT4* and *NANOG* expression and induction of endodermal and trophectodermal genes (Sumi et al., 2007). Based on these studies MYC has role in the regulation of hESCs, however, the exact function of MYC for pluripotency maintenance in hESCs is still unclear.

To conclude, this study is the first transcriptional level analysis of hESCs grown in the defined hypoxic culture conditions, without exposing cells to oxygen during the media change. We detected over two-fold induction of MYC in long-term hypoxic culture of hESCs, which is presumably regulated by HIF2α. On the contrary to the overexpression study of c-MycER, the endogenous induction of MYC in hypoxia detected in this study did not affect the expression of *OCT4*, *NANOG* or lead towards differentiation, increasing the information of MYC function in hESCs. We did not see a statistically significant difference in the cell numbers or RNA yield at day seven (data not shown), implying that the induction of MYC had not influenced the proliferation of the cells. However, we did detect significantly higher level of SSEA-3 positive cells, which has not been detected or reported in the earlier studies. Furthermore, as MYC overexpression highly induces generation of hiPSCs, our result of endogenous MYC induction in hypoxia possibly explains why generation of hiPSCs is induced in hypoxia. As the role of MYC in hESCs is still unclear, further studies are needed to resolve the meaning of MYC induction in hypoxic conditions. In particular the question remains whether hESCs became more cancerous in hypoxic maintenance or if MYC is also biologically elevated in the inner cell mass of the embryo growing in the hypoxic environment before vascularization and needed to fine tune the pluripotency of hESCs.

5.3 High resolution changes in hESCs (III)

To detect high resolution genomic alterations in hESCs a highest resolution array available at that time, SNP6.0 (Affymetrix) was selected and installed in the Finnish Microarray and Sequencing Centre (Turku, Finland). The project was carried out as an international EU-collaboration with the members of ESTOOLS consortium that enabled analysis of multiple hESC lines from various origins.

5.3.1 CNV statistics in hESCs

When karyotypically normal chromosomes of 29 samples consisting of 17 individual hESC lines were analyzed with the resolution of 50 kb, the genome of hESCs was found to carry on average 29 CNVs, of 221 kb in size (Report III, Supplementary Table 2).

For a comparison, hESCs were compared to somatic cells. Analysis of 90 HapMap samples of similar origin resulted 26 CNVs of 232 kb in size (Report III, Supplementary Table 3) just having similar CNV statistics as hESCs. However, the chromosomal CNV distribution was clearly different between somatic and hES cells (Report III, Figure 1). In addition, whereas CNV distribution of gains and losses was equal in somatic cells, the clear majority (72 %) of the genomic are covered by CNVs in hESCs contributed to gains (Report III, Figure 1).

Our analysis was the first high resolution study performed on hESC lines and first to compare the CNV profile of hESCs to these of somatic cells, as earlier lower resolution studies had only identified larger individual CNV sites making this comparison impossible. This is important, as multiple CNVs detected in hESCs are only part of normal genetic variation and should not be considered as genomic alterations. Later, another high resolution study concentrated only on CNV calls absent from 450 HapMap samples (Laurent et al., 2011). Similarly to our analysis, the average length of duplications was higher in hESCs and hiPSCs than in non-pluripotent cells. In addition, the International Stem Cell Initiative used HapMap comparison to identify 'ES cell associated' CNVs (International Stem Cell Initiative et al., 2011). The HapMap samples represent the normal human genetic variation. However, one should consider that the DNA isolated from HapMap samples can also contain culture-induced changes. That is because, the HapMap samples are fused with cancer cell lines and cultured in order to achieve constitute supply of HapMap DNA. Because of this the HapMap samples may not be ideal for the normalization and somatic comparison or to exclude those CNVs identified in hESCs which are also present in the HapMap data.

5.3.2 Specific alterations

We detected 7 genomic CNVs over 1 Mb in size, of which only one was present also in the HapMap data (Report III, Table 2.). Of particular interest we also detected amplification of 20q22.21, which was also verified with RT-PCR and proven to be the genomic hotspot CNV site in hESCs during both the review process of our study and following CNV reports (Elliott et al., 2010; International Stem Cell Initiative et al., 2011; Laurent et al., 2011; Lefort et al., 2008; Maitra et al., 2005; Martins-Taylor et al., 2011; Narva et al., 2010; Spits et al., 2008; Wu et al., 2008) (Report V, Figure X). This CNV varies in size from 500 to 5000 kb and spans genes of interest, such as *BCL2L1*, *ID1* and sometimes *DNMT3B*, which associate to cell survival and pluripotency. Importantly, the amplification in this area has been shown to be culture induced and lead to increased growth and survival rate and FGF independency (Spits et al., 2008; Werbowetski-Ogilvie et al., 2009). This CNV change proves that small CNV changes can lead to major phenotypic alterations validating the need of high resolution screens before therapeutic applications.

We also detected recurrent loss of 22q11.21 in eight hESC samples. This area was on average 260 kb in size and overlapped the gene coding known tumor suppressor HIC2

(hypermethylated in cancer 2 protein). In addition, 345 kb gain in 1p36.33 overlapping OR4F5 (olfactory receptor) was present in eight hESC samples. Importantly, these sites were present in less than 5 % of the HapMap samples, and the loss of HIC2 was not identified in the early passages samples, identifying this change culture associated.

5.3.3 SNP profiling

The SNP fingerprint analysis identified unique SNP profile for each individual hESC lines derived (Report III, Supplementary Table 1.), verifying the sample identity.

Interestingly, the LOH of 16 q arm was detected in one of the cell lines (Report III, Supplementary Figure 2.). To our knowledge, this is the only chromosomal LOH site identified in hESCs. As earlier passage samples were not available for this cell line, it remained unexplored whether this change was culture induced or inherited from the embryo. Interestingly, LOH of 16q is the most frequent alteration in grade III tumors of breast cancer (Cleton-Jansen et al., 2004; Lindblom et al., 1993), and identified in multiple myelomas and in prostate cancer (Carter et al., 1990; Jenner et al., 2007).

5.3.4 Genomic changes alter gene expression

As correlation between CN state and gene expression is recognized in other cell types (Hughes et al., 2000; Stranger et al., 2007), the transcriptome of nine selected samples was determined in order to analyze how genomic changes affect gene-expression of hESCs. The gained CN state had significantly increased the expression of 30 % of the genes whereas CN loss led to decreased expression in 42 % of the cases. Furthermore, pathway analysis revealed that 45 % of the affected genes were linked to cancer.

Our analysis was the first to examine how high resolution genomic changes affect gene expression in hESCs. Mayshar et al. used this correlation to determine chromosomal integrity of pluripotent cell lines based on the transcriptional profiling (Mayshar et al., 2010).

5.3.5 Culture induced changes

The comparison of sample pairs having early and late passage samples of the same cell lines identified that 66 % of the CNV sites and 24 % of the LOH regions had changed in culture (Report III, Figure 2a.). The high number of alterations can be partly explained by the false-positive estimate, that we detected to be 12.5 % between multiple hybridization trials for the same sample with the analysis settings and array type used (Report III, Supplementary Table 4). Further, when the genes overlapping the altered CNV sites were listed, we identified that 82 % of these corresponded to amplifications. The comparison to the list of oncogenes (http://www.sanger.ac.uk/genetics/CGP/Census/) identified, *PDE4DIP* to be a known translocation gene in myeloproliferative disorder. In addition, we identified that developmental genes *HOXA5, 6, 7, 9, 10, 11* and *13*, affected by 73 kb gain were present only in the late passage of the hESC line H7.

We studied in detail the adaptation process of H7. We determined that the genomic size affected by changes increased significantly with the time in culture (Report III, Figure 2b). The transcriptome analysis of these samples revealed that culture induced CN gains had increased 17-fold the expression of a cancer/testis-specific antigen *MAGEA4,* 10-fold the epigenetic regulator and cancer testis gene *CTCFL* and 2.5-fold of *FGF13*.

In agreement with other studies, we have also observed that in addition to chromosomal gains (Report V, Supplementary Table X, Figure 1), pluripotent cells contain higher levels of amplifications than deletions. This phenomenon is reported to occur typically in cultured cell lines and notified as the process of adaptive amplification (Hastings, 2007). In reality, pluripotent cell lines can be more prone for genetic alterations than other cell lines. It has been reported that hESCs have abnormal DNA repair mechanism, as mitotic spindle assembly checkpoint does not initiate apoptosis (Mantel et al., 2007). In addition, LINE-1 retrotransposition is active creating high resolution variation (Garcia-Perez et al., 2007; Wissing et al., 2012). Further, only the genetic changes that improve survival and desired properties, such as undifferentiated morphology, are selected and enriched in the cell population. As chromosome 12 amplification is the most common for pluripotent cells (Report V, Figure 1), it most probably is linked to expression of NANOG. Moreover, genetic change that increases the capacity of cells to proliferate has obvious parallels to that of malignant transformation. The most common genetic aberrations of hESCs, gains of chromosomes 12, and 17 are also common in testicular germ cell tumors (TGCTs), which develop from pluripotent germ cells (Baker et al., 2007; Harrison et al., 2007), implying that culture induced changes of hESCs would be beneficial for growth also *in vivo*.

To conclude, we detected that the majority of the genomic area covered by CNVs contributes to amplifications, contrasting somatic cells which have an even contribution of deletions and amplifications. In addition, CNV and LOH sites change in culture. Our study remains the only one that has studied the LOH changes in hESCs and identified LOH of 16q. Importantly, we were first to identify that CNVs in hESCs have functional effect as they alter expression of affected genes. The comparison of CNV results with conventional karyotyping proved that CGH array was ideal to determine the exact boundaries of large variations and high resolution changes. However, the array could not detect low level of mosaicism and does not give information were the gained regions have translocated and is unable to detect balanced rearrangements. To conclude, the data obtained via both conventional karyotyping and array screening should be recommended for cells to be used in therapeutic applications.

5.4 Reprogramming induces novel CNVs (IV)

As normal culture induces CNV changes in hESCs, we continued to study how reprogramming affects the genetic integrity of the cells. This project was carried out in collaboration with laboratories specialized in reprogramming and culture of hiPSCs.

5.4.1 CNV statistics of hiPSCs

When the CNV profiles of 22 hiPSC lines and three parental and one unrelated fibroblast lines were compared to those of the normal karyotype hESC lines analyzed in the earlier study (Report III), it was noted that the median number of CNVs was twofold higher in hiPSCs (109) than in hESCs (55) and fibroblasts (53) (Report IV, Supplementary Figure 4a, Supplementary Table 3, 4). The number of CNVs detected was overall higher in this study, because the resolution analysis settings were raised from 50 kb to 10 kb. Similarly to the earlier study (Report III) hESCs and somatic cells (fibroblast) had similar CNV statistics. In addition, the majority (52.4 %) of the CNVs detected in hiPSCs were absent from hESCs and fibroblast (Report IV, Supplementary Figure 4b). Similarly, while this study was in revision, Laurent et al. published that hiPSCs contain higher number of CNVs than non-pluripotent cells (Laurent et al., 2011). These results indicate that *de novo* CNVs are generated during reprogramming.

5.4.2 Early passage hiPSCs are genetically mosaic

Interestingly, early-passage hiPSCs (p ~5) had higher number of CNVs and a larger total CNV size than late-passage hiPSCs (p ~13) (Report IV, Figure 1a, Supplementary Figure 4e). This result was true also in the individual data set consisting of four hiPSC lines (Report VI, Supplementary Table 5.). Moreover, early-passages contained more CNVs also in the individual hESC lines analyzed at different passages, and was not dependent of the reprogramming method used (Report IV, Figure 1b, c, d). This indicates that CNVs are negatively selected during culture after reprogramming.

More importantly, some of the homozygous deletions were found only at the early-passage hiPSCs (Report IV, Figure 2d). Similarly, Laurent et al. 2011 reported that some deletions of hiPSCs receded from the population in long-term culture. Later, Martins-Taylor et al. 2011 also noticed that some CNVs present in the early-passages were lost at higher passages. As the only explanation for the loss of homozygous deletions is that the early-passage hiPSC population contains cells that do not carry the deletion and are enriched later in culture, the possibility for mosaicism was further validated. We found, based on FISH detection, that early-passage hiPSCs were mosaic, 13,5 % carrying the CNV tested, the rest having normal CN-state. Further, the portion of cells carrying tested CNV was decreased to 6.5 % in culture (Report IV, Figure 2a, b), implying that the CN site tested was selected negatively in culture.

In addition, the novel CNVs, not present in the HapMap data, were highly present in hiPSCs (37 %), compared to hESCs (25 %) and fibroblasts (15 %) (Report IV, Supplementary Figure 8a). Importantly, the early passage hiPSCs contained novel CNVs, which were eliminated in culture (Report IV, Supplementary 8b, c). This indicates that the *de novo* generated novel CNVs during reprogramming are not favorable for the maintenance of hiPSCs. However, as SNP-arrays detect large CNVs as multiple smaller CNVs in case of mosaic cell population (Report III, Figure 3b)(Report IV, Figure 3a), the CNV boundaries

and size detected in the early passages are not that exact as the analysis results indicate, reducing the actual number of novel CNVs. Nevertheless, as the overlapping size of CNVs between early- and late-passage does not alter in culture, this supports the idea that CNVs generated during reprogramming are selected against in culture (Report IV, Figure 3b). Interestingly, several deletions in the early passage hiPSCs affected essential genes needed for the maintenance of pluripotency (Supplementary Table 8), explaining the fast elimination of the cells carrying these CNVs in culture. In addition, deletions of hiPSCs occured more frequently in common fragile sites (CFSs) (Report IV, Figure 4a). As CFSs contain late-replicating sequences, reprogramming induced CNVs can be generated due to the replication stress. In addition, the average deletion size within subtelomeric regions was significantly larger in early-passage lines and nearest to the telomeres (Report IV, Supplementary Figure 9a, b).

To conclude, reprogramming is associated with high mutation rates. Early-passage hiPSCs have altered CNV profile due to the presence of aberrant cells, which are rapidly selected against in culture. Interestingly, our results could explain why the gene-expression profile of the late-passage hiPSCs is more alike with hESCs than early-passage hiPSCs (Chin et al., 2009). In addition, the existence of aberrant cells in the early hiPSCs should be considered when hiPSCs are analyzed with various platforms. Importantly, from the therapeutic point of view, cells should not be used immediately after reprogramming. More importantly, although most of the *de novo* generated CNVs are unfavorable for cell growth, reprogramming and long-term culture may also induce cancer promoting mutations, which would enrich in culture leading to population with malignant properties.

5.5 Review: genetic and epigenetic stability of hESCs and hiPSCs (V)

At the time of writing this thesis, a review was requested concerning genetic and epigenetic stability of human pluripotent stem cells.

5.5.1 Major conclusions based on the current literature

Amplification of chromosomes 1, 8, 12, and X are the most common karyotypic abnormalities detected in human pluripotent stem cells (Report V, Figure 1, Supplementary Table S1). Interestingly a gain of chromosome 1q (1q25-1q41) or trisomy of chromosome 17, which are both common in hESCs, are not observed in hiPSCs. The most recurrent changes in mESCs are trisomies of chromosomes 8 and 11 (Report V, BOX2). Interestingly only syntenic areas prone for amplification between human and mouse are chromosome 11 of mouse and chromosome 17 of human. Based on multiple high resolution studies performed on pluripotent cells (Report V, Table 1) CNVs, single nucleotide and LOH changes occur also in prolonged culture. The most recurrent CNV hotspot in pluripotent stem cells is amplification of 20q11.21 (Report V, Figure 1, Supplementary information S2).

The epigenetic state of the pluripotent cells has been intensively studied (Report V, Table 2). Based on these studies pluripotent cells are characterized by a high level of variation in the methylation status of a subset of imprinted genes. In addition, the X chromosome inactivation status varies among different female hESC or hiPSC lines. Comparisons of hESCs and hiPSCs have shown that the variation also in the histone mark patterns. The culture-induced changes are mostly unpredictable and variable between the lines.

The genomic stability of hiPSCs seems not to be dependent on the method used for reprogramming. However, it is clear that reprogramming process itself can induce genomic changes. The *in vitro* maintained pluripotent cells are constantly under selection which eliminates mutated cells carrying deletions but is also prone for adaptive amplification (Report V, Figure 2). Epigenomic instability associated with reprogramming and early passages of hiPSCs correlates with the heterogeneity observed in structural genomic changes. However, the reprogramming process can also be incomplete leaving epigenetic marks from somatic cells used for reprogramming.

Unlike epigenetic modifications, which are sensitive to environmental signals and stress, the genomic instability does not correlate with any particular culturing technique. It is clear that culture-adapted abnormal cells have changed phenotype. The characterization of genomic and epigenomic stability is important as abnormal cells share characteristics with cancer cells and may become malignant after transplantation. Enrichment of genomic and epigenomic changes may occur at any stage during derivation, culturing, reprogramming, differentiation, freezing and thawing (Report V, Figure 3). In the future, it is highly important to study more the instability of cells during various differentiation protocols, as the differentiated cells are the ones to be transferred into patients. In addition, functional demonstrations to confirm which specific alterations have phenotypic consequences and thus implications for the safe clinical use of pluripotent cell-derived therapies needs to further studied.

6. CONCLUSIONS

This thesis concentrated on the study of human pluripotent stem cells. The pluripotency regulation of human pluripotent cells differs from that of mouse in various aspects. However, a number of studies rely on data validated only in the mouse system. From the therapeutic point of view, the pluripotency regulation and genetic stability needs to be characterized in the human system. The aim of this thesis was to study pluripotency regulation in hESCs and to characterize high resolution genomic changes involved in culture and reprogramming.

The work presented in this thesis identifies previously unknown stem cell specific protein L1TD1 to be essential for the maintenance of pluripotency of hESCs. We were able to show that L1TD1 is an RNA-binding protein localized in the cytoplasmic processing bodies. In addition, we determined that the expression of L1TD1 is regulated by stem cell specific core transcription factors OCT4, NANOG and SOX2, increasing the information of the target genes involved in maintaining pluripotency. Moreover, we demonstrated that L1TD1 binds directly to RHA, and through RNA connection with another stem cell specific RNA binding protein, LIN28. As LIN28-RHA complex regulates translation of OCT4, and depletion of L1TD1 immediately decreases levels of OCT4, LIN28 and RHA, we assume that L1TD1 is part of this complex. However, as RNA-binding proteins commonly have multiple targets, L1TD1 most probably can affect maintenance of pluripotency by an additional mechanism. Interestingly, as well as the stem cell specific expression in normal cells, L1TD1 was found to have high expression especially in seminomas. More importantly, depletion of L1TD1 in non-seminoma and seminoma cancer cell lines lead to degreased proliferation and loss of OCT4 and NANOG, implying that L1TD1 has also vital role in the self-renewal of these cancers.

The hypoxic response and effects on the transcriptome of hESCs was determined in this thesis. It was found that hESCs express HIF1α and activate HIF1α responsive genes in the initial exposure to hypoxia, but this response is stabilized in long-term hypoxic culture. In addition, hESCs were found to express all isotypes of PHDs, which can be responsible for the long-term hypoxic tolerance of hESCs. The morphology of the hESCs colonies was found to be less differentiated after long-term exposure to hypoxia. In addition, the expression of surface antigens implied that hypoxia supports undifferentiated growth of hESCs. In particular, the expression of the surface antigen SSEA-3 was higher in hypoxia. Transcriptome analysis revealed that a number of stem cell associated genes had altered expression in hypoxia. Of the relative interest, we detected induction of pluripotency associated genes *LEFTY2*, *HEY2* and *MYC* and decrease of differentiation associated genes *PAX6*, *MSX2* and *FLRT3*. As MYC is a multifunctional master regulator, the endogenous induction of MYC in hypoxia can significantly affect the pluripotency regulation and growth of hESCs.

In this thesis, the high resolution changes were analysed in culture and reprogramming of pluripotent cells. The CNV statistics were similar in hESCs and somatic cells, whereas hiPSCs contained significantly higher number of CNVs. Interestingly, the majority of the genomic area covered by CNVs in hESCs contributed to amplifications, implying that the high resolution changes are also prone for adaptive amplification process. In addition, culture of hESCs induced multiple high resolution changes that altered the gene expression of affected genes. Importantly, reprogramming was found to increase the number of CNVs. We were able to determine that early passage hiPSCs are mosaic, consisting of normal and mutated cells, which are selected out in the subsequent culture, supporting the use of later passage hiPSCs. However, long-term culture presents the risk of generation and enrichment of oncogenic mutations as detected in hESC cultures. Due to this it is difficult to estimate safe passage for the use of hiPSCs. The future challenge is to discover reprogramming methods and culture conditions that would not affect the genomic integrity of reprogrammed cells. However, this can be a major challenge as mutations do occur in living cells and genomic evolution constantly drives the enrichment of favorable properties.

To conclude, in this thesis a number of novel findings on the regulation and genetic stability of pluripotent cells that are important for the field of stem cell research, cancer research and applications of regenerative medicine are presented. The high resolution studies of the genetic stability of pluripotent cells presented in this thesis have already impacted highly the stem cell field and been extensively cited.

ACKNOWLEDGEMENTS

This work was carried out at the Turku Centre for Biotechnology, Department of Medical Microbiology and Immunology, University of Turku during the years 2006-2012. A number of people have impacted the completion of this thesis to who I wish to express my gratitude.

First of all, I would like to thank my supervisor Professor Riitta Lahesmaa. Although many believed that it was impossible to culture stem cells and achieve doctoral thesis in this demanding field of science, you had the courage to set up the team and give me the opportunity to work with human embryonic stem cells. Especially I would like to thank you for the scientific freedom, guidance, travelling, and opportunity to work and interact with number of leading scientist and present my work in international meetings.

I am very grateful for the Turku Doctoral Programme of Biomedical Sciences for my position in the graduate school and funding. The PhD thesis supervisory committee Professors Anjana Rao and Klaus Elenius are acknowledged for guidance and advice during these years. I express my gratitude for Docent Katriina Aalto-Setälä and Professor Tapio Visakorpi for carefully reviewing this thesis and providing excellent comments. In addition, Robert Moulder is acknowledged for the language proofreading.

I would like to deeply acknowledge all the co-authors and collaborators. I would like to especially thank Nelly Rahkonen, a friend and a great scientist for the endless hours that we have shared in the laboratory and office. I express my high gratitude for the excellent work and guidance of the doctors Reija Autio, Riikka Lund, Juha-Pekka Pursiheimo, Maheswara Reddy Emani, and Docents Heli Salminen-Mankonen and Omid Rasool. I deeply acknowledge the collaboration work performed by excellent young scientist Samer Hussein, Sanna Vuoristo, Asta Laiho, Kirsti Laurila, Lingjia Kong, Neil Harrison, Danny Kitsberg, Lodovica Borghese, Milla Mikkola, Ras Trokovic and Duncan Baker. Moreover, I am extremely grateful for the help, discussions and supervision provided by Professors Peter Andrews, Nissim Benvenisty, Outi Hovatta, Oliver Brüstle, Petr Dvorak, Timo Otonkoski, Timo Tuuri, Harri Lähdesmäki, Anjana Rao, Konstantin Denessiouk, Olli Yli-Harja, Panu Jaakkola, Wei Cui, Heli Skottman, and Harry D. Moore. I owe special acknowledge also for the students and past members Juuso Nästi, Anu Neuvonen, Joel Nyström, Ravi Kant, Subhanjan Bhowmik, Roosa Sahla, Anne Mari Tveit and Krista Maurinen for your work in the stem cell team.

The current and past members of ATLAS and stem cell group have provided me significant support in science and life, thank you for that. Especially I want to thank Kirsi Granberg who supervised my pro gradu thesis and prepared me for the future battle of the thesis work. My deepest thanks to Arja Reinikainen, Marjo Hakkarainen, Päivi Junni, and Fezazi Bogata for setting up and maintaining our stock culture laboratory. Without your efforts every day, including Christmas and Midsummer, I would not have reached

this goal. I would also like to thank David Hawkins and his group members Cristina Valensisi and Kalyan Pasumarthy for the close collaboration at CBT. Moreover, I express my gratitude to CBT staff members at administration, technical and computer section. I would like to acknowledge the Finnish DNA Microarray and Sequencing Centre for their excellent technical assistance and preparation of my numerous samples. Special thanks to Cell imaging core, especially Perttu Terho and Jouko Sandholm who have helped me countless times. I owe special acknowledge also to all ESTOOLS members and scientist at Regea Cell and Tissue Center, who have shared their knowledge, research material and friendship.

I wish to express my gratitude to all my friends. Thank you for the memorable moments and great time we have spent together.

My warmest thank to family. Thank you for your support and encouragement to achieve the first doctoral hat in our family. Moreover thank you for your help in other projects in life.

Finally I owe my deepest gratitude to my number one supporters Toni and our daughter Mikaela, the greatest miracle and achievement of my life. Although science is great, you are my sunshine's.

This study was supported by funding for the ESTOOLS consortium under the Sixth Research Framework Programme of the European Union, JDRF, The Academy of Finland, the Finnish Cancer Organizations, Turku Graduate School of Biomedical Sciences, Ida Montin Foundation, EVO: The Hospital District of Southwest Finland, Turun Yliopistosäätiö, SysBio, Turun Mikrobiologien Tiedesäätiö, Turun Yliopisto, and Finnish Cultural Foundation.

REFERENCES

Aasen, T., and Izpisua Belmonte, J.C. (2010). Isolation and cultivation of human keratinocytes from skin or plucked hair for the generation of induced pluripotent stem cells. Nat. Protoc. *5,* 371-382.

Adachi, K., Suemori, H., Yasuda, S.Y., Nakatsuji, N., and Kawase, E. (2010). Role of SOX2 in maintaining pluripotency of human embryonic stem cells. Genes Cells *15,* 455-470.

Allegrucci, C., Thurston, A., Lucas, E., and Young, L. (2005). Epigenetics and the germline. Reproduction *129,* 137-149.

Allegrucci, C., Wu, Y.Z., Thurston, A., Denning, C.N., Priddle, H., Mummery, C.L., Ward-van Oostwaard, D., Andrews, P.W., Stojkovic, M., Smith, N.*, et al.* (2007). Restriction landmark genome scanning identifies culture-induced DNA methylation instability in the human embryonic stem cell epigenome. Hum. Mol. Genet. *16,* 1253-1268.

Amit, M., Carpenter, M.K., Inokuma, M.S., Chiu, C.P., Harris, C.P., Waknitz, M.A., Itskovitz-Eldor, J., and Thomson, J.A. (2000). Clonally derived human embryonic stem cell lines maintain pluripotency and proliferative potential for prolonged periods of culture. Dev. Biol. *227,* 271-278.

Andrews, P.W. (2011). Toward safer regenerative medicine. Nat. Biotechnol. *29,* 803-805.

Anokye-Danso, F., Trivedi, C.M., Juhr, D., Gupta, M., Cui, Z., Tian, Y., Zhang, Y., Yang, W., Gruber, P.J., Epstein, J.A., and Morrisey, E.E. (2011). Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. Cell. Stem Cell. *8,* 376-388.

Askanazy, M. (1907). Die Teratome nach ihrem Bau, ihrem Verlauf, ihrer Genese und im Vergleich zum experimentellen Teratoid.*11,* 39-82.

Assou, S., Cerecedo, D., Tondeur, S., Pantesco, V., Hovatta, O., Klein, B., Hamamah, S., and De Vos, J. (2009). A gene expression signature shared by human mature oocytes and embryonic stem cells. BMC Genomics *10,* 10.

Assou, S., Le Carrour, T., Tondeur, S., Strom, S., Gabelle, A., Marty, S., Nadal, L., Pantesco, V., Reme, T., Hugnot, J.P.*, et al.* (2007). A meta-analysis of human embryonic stem cells transcriptome integrated into a web-based expression atlas. Stem Cells *25,* 961-973.

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

Babushok, D.V., and Kazazian, H.H.,Jr. (2007). Progress in understanding the biology of the human mutagen LINE-1. Hum. Mutat. *28,* 527-539.

Bae, D., Mondragon-Teran, P., Hernandez, D., Ruban, L., Mason, C., Bhattacharya, S.S., and Veraitch, F.S. (2011). Hypoxia Enhances the Generation of Retinal Progenitor Cells from Human Induced Pluripotent and Embryonic Stem Cells. Stem Cells Dev.

Baharvand, H., Ashtiani, S.K., Taee, A., Massumi, M., Valojerdi, M.R., Yazdi, P.E., Moradi, S.Z., and Farrokhi, A. (2006). Generation of new human embryonic stem cell lines with diploid and triploid karyotypes. Dev. Growth Differ. *48,* 117-128.

Baker, D.E., Harrison, N.J., Maltby, E., Smith, K., Moore, H.D., Shaw, P.J., Heath, P.R., Holden, H., and Andrews, P.W. (2007). Adaptation to culture of human embryonic stem cells and oncogenesis in vivo. Nat. Biotechnol. *25,* 207-215.

Baldassarre, G., Bianco, C., Tortora, G., Ruggiero, A., Moasser, M., Dmitrovsky, E., Bianco, A.R., and Ciardiello, F. (1996). Transfection with a CRIPTO anti-sense plasmid suppresses endogenous CRIPTO expression and inhibits transformation in a human embryonal carcinoma cell line. Int. J. Cancer *66,* 538-543.

Balzer, E., and Moss, E.G. (2007). Localization of the developmental timing regulator Lin28 to mRNP complexes, P-bodies and stress granules. RNA Biol. *4,* 16-25.

Bar, M., Wyman, S.K., Fritz, B.R., Qi, J., Garg, K.S., Parkin, R.K., Kroh, E.M., Bendoraite, A., Mitchell, P.S., Nelson, A.M.*, et al.* (2008). MicroRNA discovery and profiling in human embryonic stem cells by deep sequencing of small RNA libraries. Stem Cells *26,* 2496-2505.

Barroso-delJesus, A., Romero-Lopez, C., Lucena-Aguilar, G., Melen, G.J., Sanchez, L., Ligero, G., Berzal-Herranz, A., and Menendez, P. (2008). Embryonic stem cell-specific miR302-367 cluster: human gene structure and functional characterization of its core promoter. Mol. Cell. Biol. *28,* 6609-6619.

Beattie, G.M., Lopez, A.D., Bucay, N., Hinton, A., Firpo, M.T., King, C.C., and Hayek, A. (2005). Activin A maintains pluripotency of human embryonic stem cells in the absence of feeder layers. Stem Cells *23,* 489-495.

Becker, K.A., Ghule, P.N., Therrien, J.A., Lian, J.B., Stein, J.L., van Wijnen, A.J., and Stein, G.S. (2006). Self-renewal of human embryonic stem cells is

supported by a shortened G1 cell cycle phase. J. Cell. Physiol. *209,* 883-893.

Bengtsson, H., Simpson, K., Bullard, J., and Hansen, K. (2008). aroma.affymetrix: A generic framework in R for analyzing small to very large Affymetrix data sets in bounded memory.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing.*57,* 289-300.

Bertout, J.A., Patel, S.A., and Simon, M.C. (2008). The impact of O2 availability on human cancer. Nat. Rev. Cancer. *8,* 967-975.

Besser, D. (2004). Expression of nodal, lefty-a, and lefty-B in undifferentiated human embryonic stem cells requires activation of Smad2/3. J. Biol. Chem. *279,* 45076-45084.

Bibikova, M., Chudin, E., Wu, B., Zhou, L., Garcia, E.W., Liu, Y., Shin, S., Plaia, T.W., Auerbach, J.M., Arking, D.E.*, et al.* (2006). Human embryonic stem cells have a unique epigenetic signature. Genome Res. *16,* 1075-1083.

Bock, C., Kiskinis, E., Verstappen, G., Gu, H., Boulting, G., Smith, Z.D., Ziller, M., Croft, G.F., Amoroso, M.W., Oakley, D.H.*, et al.* (2011). Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. Cell *144,* 439-452.

Bolstad, B.M., Irizarry, R.A., Astrand, M., and Speed, T.P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics *19,* 185-193.

Boyd, A.S., Rodrigues, N.P., Lui, K.O., Fu, X., and Xu, Y. (2012). Concise review: Immune recognition of induced pluripotent stem cells. Stem Cells *30,* 797-803.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G.*, et al.* (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122,* 947-956.

Bradley, A., Evans, M., Kaufman, M.H., and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature *309,* 255-256.

Brimble, S.N., Sherrer, E.S., Uhl, E.W., Wang, E., Kelly, S., Merrill, A.H.,Jr, Robins, A.J., and Schulz, T.C. (2007). The cell surface glycosphingolipids SSEA-3 and SSEA-4 are not essential for human ESC pluripotency. Stem Cells *25,* 54-62.

Buzzard, J.J., Gough, N.M., Crook, J.M., and Colman, A. (2004). Karyotype of human ES cells during extended culture. Nat. Biotechnol. *22,* 381-2; author reply 382.

Cai, L., Ye, Z., Zhou, B.Y., Mali, P., Zhou, C., and Cheng, L. (2007). Promoting human embryonic stem cell renewal or differentiation by modulating Wnt signal and culture conditions. Cell Res. *17,* 62-72.

Caisander, G., Park, H., Frej, K., Lindqvist, J., Bergh, C., Lundin, K., and Hanson, C. (2006). Chromosomal integrity maintained in five human embryonic stem cell lines after prolonged in vitro culture. Chromosome Res. *14,* 131-137.

Cameron, C.M., Harding, F., Hu, W.S., and Kaufman, D.S. (2008). Activation of hypoxic response in human embryonic stem cell-derived embryoid bodies. Exp. Biol. Med. (Maywood) *233,* 1044-1057.

Cantin, G., and Lyonnais, J. (1983). Anti-PP1Pk and early abortion. Transfusion *23,* 350-351.

Card, D.A., Hebbar, P.B., Li, L., Trotter, K.W., Komatsu, Y., Mishina, Y., and Archer, T.K. (2008). Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells. Mol. Cell. Biol. *28,* 6426-6438.

Carter, B.S., Ewing, C.M., Ward, W.S., Treiger, B.F., Aalders, T.W., Schalken, J.A., Epstein, J.I., and Isaacs, W.B. (1990). Allelic loss of chromosomes 16q and 10q in human prostate cancer. Proc. Natl. Acad. Sci. U. S. A. *87,* 8751-8755.

Cartwright, P., McLean, C., Sheppard, A., Rivett, D., Jones, K., and Dalton, S. (2005). LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. Development *132,* 885-896.

Catalina, P., Montes, R., Ligero, G., Sanchez, L., de la Cueva, T., Bueno, C., Leone, P.E., and Menendez, P. (2008). Human ESCs predisposition to karyotypic instability: Is a matter of culture adaptation or differential vulnerability among hESC lines due to inherent properties? Mol. Cancer. *7,* 76.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell *113,* 643-655.

Chen, H.F., Kuo, H.C., Chen, W., Wu, F.C., Yang, Y.S., and Ho, H.N. (2009). A reduced oxygen tension (5%) is not beneficial for maintaining human embryonic stem cells in the undifferentiated state with short splitting intervals. Hum. Reprod. *24,* 71-80.

Chin, M.H., Mason, M.J., Xie, W., Volinia, S., Singer, M., Peterson, C., Ambartsumyan, G., Aimiuwu, O., Richter, L., Zhang, J.*, et al.* (2009). Induced

pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. Cell. Stem Cell. *5,* 111-123.

Cleton-Jansen, A.M., Buerger, H., Haar, N., Philippo, K., van de Vijver, M.J., Boecker, W., Smit, V.T., and Cornelisse, C.J. (2004). Different mechanisms of chromosome 16 loss of heterozygosity in well- versus poorly differentiated ductal breast cancer. Genes Chromosomes Cancer *41,* 109-116.

Cowan, C.A., Atienza, J., Melton, D.A., and Eggan, K. (2005). Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. Science *309,* 1369-1373.

Cowan, C.A., Klimanskaya, I., McMahon, J., Atienza, J., Witmyer, J., Zucker, J.P., Wang, S., Morton, C.C., McMahon, A.P., Powers, D., and Melton, D.A. (2004). Derivation of embryonic stem-cell lines from human blastocysts. N. Engl. J. Med. *350,* 1353-1356.

Dai, M., Wang, P., Boyd, A.D., Kostov, G., Athey, B., Jones, E.G., Bunney, W.E., Myers, R.M., Speed, T.P., Akil, H., Watson, S.J., and Meng, F. (2005). Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. Nucleic Acids Res. *33,* e175.

Darr, H., and Benvenisty, N. (2009). Genetic analysis of the role of the reprogramming gene LIN-28 in human embryonic stem cells. Stem Cells *27,* 352-362.

Darr, H., Mayshar, Y., and Benvenisty, N. (2006). Overexpression of NANOG in human ES cells enables feeder-free growth while inducing primitive ectoderm features. Development *133,* 1193-1201.

Davidson, K.C., Adams, A.M., Goodson, J.M., McDonald, C.E., Potter, J.C., Berndt, J.D., Biechele, T.L., Taylor, R.J., and Moon, R.T. (2012). Wnt/beta-catenin signaling promotes differentiation, not self-renewal, of human embryonic stem cells and is repressed by Oct4. Proc. Natl. Acad. Sci. U. S. A. *109,* 4485-4490.

Dhara, S.K., and Benvenisty, N. (2004). Gene trap as a tool for genome annotation and analysis of X chromosome inactivation in human embryonic stem cells. Nucleic Acids Res. *32,* 3995-4002.

Draper, J.S., Moore, H.D., Ruban, L.N., Gokhale, P.J., and Andrews, P.W. (2004). Culture and characterization of human embryonic stem cells. Stem Cells Dev. *13,* 325-336.

Draper, J.S., Pigott, C., Thomson, J.A., and Andrews, P.W. (2002). Surface antigens of human embryonic stem cells: changes upon differentiation in culture. J. Anat. *200,* 249-258.

Draper, J.S., Smith, K., Gokhale, P., Moore, H.D., Maltby, E., Johnson, J., Meisner, L., Zwaka, T.P., Thomson, J.A., and Andrews, P.W. (2004). Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. Nat. Biotechnol. *22,* 53-54.

Dravid, G., Ye, Z., Hammond, H., Chen, G., Pyle, A., Donovan, P., Yu, X., and Cheng, L. (2005). Defining the role of Wnt/beta-catenin signaling in the survival, proliferation, and self-renewal of human embryonic stem cells. Stem Cells *23,* 1489-1501.

Du, P., Kibbe, W.A., and Lin, S.M. (2008). lumi: a pipeline for processing Illumina microarray. Bioinformatics *24,* 1547-1548.

Dvorak, P., Dvorakova, D., Koskova, S., Vodinska, M., Najvirtova, M., Krekac, D., and Hampl, A. (2005). Expression and potential role of fibroblast growth factor 2 and its receptors in human embryonic stem cells. Stem Cells *23,* 1200-1211.

Elliott, A.M., Elliott, K.A., and Kammesheidt, A. (2010). High resolution array-CGH characterization of human stem cells using a stem cell focused microarray. Mol. Biotechnol. *46,* 234-242.

Enver, T., Soneji, S., Joshi, C., Brown, J., Iborra, F., Orntoft, T., Thykjaer, T., Maltby, E., Smith, K., Dawud, R.A., *et al.* (2005). Cellular differentiation hierarchies in normal and culture-adapted human embryonic stem cells. Hum. Mol. Genet. *14,* 3129-3140.

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

Eyre, T.A., Ducluzeau, F., Sneddon, T.P., Povey, S., Bruford, E.A., and Lush, M.J. (2006). The HUGO Gene Nomenclature Database, 2006 updates. Nucleic Acids Res. *34,* D319-21.

Ezashi, T., Das, P., and Roberts, R.M. (2005). Low O2 tensions and the prevention of differentiation of hES cells. Proc. Natl. Acad. Sci. U. S. A. *102,* 4783-4788.

Ezeh, U.I., Turek, P.J., Reijo, R.A., and Clark, A.T. (2005). Human embryonic stem cell genes OCT4, NANOG, STELLAR, and GDF3 are expressed in both seminoma and breast carcinoma. Cancer *104,* 2255-2265.

Fazeli, A., Liew, C.G., Matin, M.M., Elliott, S., Jeanmeure, L.F., Wright, P.C., Moore, H., and Andrews, P.W. (2011). Altered patterns of differentiation in karyotypically abnormal human embryonic stem cells. Int. J. Dev. Biol. *55,* 175-180.

Feng, Q., Moran, J.V., Kazazian, H.H.,Jr, and Boeke, J.D. (1996). Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell *87,* 905-916.

Fong, H., Hohenstein, K.A., and Donovan, P.J. (2008). Regulation of self-renewal and pluripotency by Sox2 in human embryonic stem cells. Stem Cells *26,* 1931-1938.

Forristal, C.E., Wright, K.L., Hanley, N.A., Oreffo, R.O., and Houghton, F.D. (2010). Hypoxia inducible factors regulate pluripotency and proliferation in human embryonic stem cells cultured at reduced oxygen tensions. Reproduction *139,* 85-97.

Forsyth, N.R., Kay, A., Hampson, K., Downing, A., Talbot, R., and McWhir, J. (2008). Transcriptome alterations due to physiological normoxic (2% O2) culture of human embryonic stem cells. Regen. Med. *3,* 817-833.

Forsyth, N.R., Musio, A., Vezzoni, P., Simpson, A.H., Noble, B.S., and McWhir, J. (2006). Physiologic oxygen enhances human embryonic stem cell clonal recovery and reduces chromosomal abnormalities. Cloning Stem Cells *8,* 16-23.

Fusaki, N., Ban, H., Nishiyama, A., Saeki, K., and Hasegawa, M. (2009). Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. Proc. Jpn. Acad. Ser. B. Phys. Biol. Sci. *85,* 348-362.

Garcia-Perez, J.L., Marchetto, M.C., Muotri, A.R., Coufal, N.G., Gage, F.H., O'Shea, K.S., and Moran, J.V. (2007). LINE-1 retrotransposition in human embryonic stem cells. Hum. Mol. Genet. *16,* 1569-1577.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J.*, et al.* (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. *5,* R80.

Gertow, K., Cedervall, J., Unger, C., Szoke, K., Blennow, E., Imreh, M.P., and Ahrlund-Richter, L. (2007). Trisomy 12 in HESC leads to no selective in vivo growth advantage in teratomas, but induces an increased abundance of renal development. J. Cell. Biochem. *100,* 1518-1525.

Giemsa, G. (1904). Eine Vereinfachung und Vervollkommnung meiner Methylenazur-Methylenblau-Eosin-Färbemethode zur Erzielund der Romanowsky-Nochtchen Chromatinfärbung. Cetralblatt Für Bakteriologie *37,* 308-311.

Gilbert, F.S. (2003). Developmental Biology (USA: Sinauer Associates Inc).

Gordan, J.D., Thompson, C.B., and Simon, M.C. (2007). HIF and c-Myc: sibling rivals for control of cancer cell metabolism and proliferation. Cancer. Cell. *12,* 108-113.

Greber, B., Wu, G., Bernemann, C., Joo, J.Y., Han, D.W., Ko, K., Tapia, N., Sabour, D., Sterneckert, J., Tesar, P., and Scholer, H.R. (2010). Conserved and divergent roles of FGF signaling in mouse epiblast stem cells and human embryonic stem cells. Cell. Stem Cell. *6,* 215-226.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130,* 77-88.

Haase, A., Olmer, R., Schwanke, K., Wunderlich, S., Merkert, S., Hess, C., Zweigerdt, R., Gruh, I., Meyer, J., Wagner, S.*, et al.* (2009). Generation of induced pluripotent stem cells from human cord blood. Cell. Stem Cell. *5,* 434-441.

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M.,Jr, Jungkamp, A.C., Munschauer, M.*, et al.* (2010). Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell *141,* 129-141.

Hardarson, T., Caisander, G., Sjogren, A., Hanson, C., Hamberger, L., and Lundin, K. (2003). A morphological and chromosomal study of blastocysts developing from morphologically suboptimal human pre-embryos compared with control blastocysts. Hum. Reprod. *18,* 399-407.

Harley, C.B., Futcher, A.B., and Greider, C.W. (1990). Telomeres shorten during ageing of human fibroblasts. Nature *345,* 458-460.

Harrison, N.J., Baker, D., and Andrews, P.W. (2007). Culture adaptation of embryonic stem cells echoes germ cell malignancy. Int. J. Androl. *30,* 275-81; discussion 281.

Hastings, P.J. (2007). Adaptive amplification. Crit. Rev. Biochem. Mol. Biol. *42,* 271-283.

Hastings, P.J., Lupski, J.R., Rosenberg, S.M., and Ira, G. (2009). Mechanisms of change in gene copy number. Nat. Rev. Genet. *10,* 551-564.

Hautaniemi, S., Ringnér, M., Kauraniemi, P., Autio, R., Edgren, H., Yli-Harja, O., Astola, J., Kallioniemi, A., and Kallioniemi, O. (2004). A strategy for identifying putative causes of gene expression variation in human cancers.*341,* 77-88.

Hawkins, R.D., Hon, G.C., Lee, L.K., Ngo, Q., Lister, R., Pelizzola, M., Edsall, L.E., Kuan, S., Luu, Y., Klugman, S.*, et al.* (2010). Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. Cell. Stem Cell. *6,* 479-491.

Hay, D.C., Sutherland, L., Clark, J., and Burdon, T. (2004). Oct-4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers

in human and mouse embryonic stem cells. Stem Cells *22,* 225-235.

Heins, N., Englund, M.C., Sjoblom, C., Dahl, U., Tonning, A., Bergh, C., Lindahl, A., Hanson, C., and Semb, H. (2004). Derivation, characterization, and differentiation of human embryonic stem cells. Stem Cells *22,* 367-376.

Heo, I., Joo, C., Kim, Y.K., Ha, M., Yoon, M.J., Cho, J., Yeom, K.H., Han, J., and Kim, V.N. (2009). TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell *138,* 696-708.

Herszfeld, D., Wolvetang, E., Langton-Bunker, E., Chung, T.L., Filipczyk, A.A., Houssami, S., Jamshidi, P., Koh, K., Laslett, A.L., Michalska, A.*, et al.* (2006). CD30 is a survival factor and a biomarker for transformed human pluripotent stem cells. Nat. Biotechnol. *24,* 351-357.

Hockemeyer, D., Soldner, F., Cook, E.G., Gao, Q., Mitalipova, M., and Jaenisch, R. (2008). A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. Cell. Stem Cell. *3,* 346-353.

Hoffman, L.M., Hall, L., Batten, J.L., Young, H., Pardasani, D., Baetge, E.E., Lawrence, J., and Carpenter, M.K. (2005). X-inactivation status varies in human embryonic stem cell lines. Stem Cells *23,* 1468-1478.

Hogel, H., Rantanen, K., Jokilehto, T., Grenman, R., and Jaakkola, P.M. (2011). Prolyl hydroxylase PHD3 enhances the hypoxic survival and G1 to S transition of carcinoma cells. PLoS One *6,* e27112.

Hohjoh, H., and Singer, M.F. (1997). Sequence-specific single-strand RNA binding protein encoded by the human LINE-1 retrotransposon. EMBO J. *16,* 6034-6043.

Hovatta, O., Jaconi, M., Tohonen, V., Bena, F., Gimelli, S., Bosman, A., Holm, F., Wyder, S., Zdobnov, E.M., Irion, O.*, et al.* (2010). A teratocarcinoma-like human embryonic stem cell (hESC) line and four hESC lines reveal potentially oncogenic genomic changes. PLoS One *5,* e10263.

Huangfu, D., Osafune, K., Maehr, R., Guo, W., Eijkelenboom, A., Chen, S., Muhlestein, W., and Melton, D.A. (2008). Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. Nat. Biotechnol. *26,* 1269-1275.

Hubbard, T.J., Aken, B.L., Beal, K., Ballester, B., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cunningham, F., Cutts, T.*, et al.* (2007). Ensembl 2007. Nucleic Acids Res. *35,* D610-7.

Hughes, T.R., Roberts, C.J., Dai, H., Jones, A.R., Meyer, M.R., Slade, D., Burchard, J., Dow, S., Ward, T.R., Kidd, M.J., Friend, S.H., and Marton, M.J. (2000). Widespread aneuploidy revealed by DNA microarray expression profiling. Nat. Genet. *25,* 333-337.

Hussein, S.M., Batada, N.N., Vuoristo, S., Ching, R.W., Autio, R., Narva, E., Ng, S., Sourour, M., Hamalainen, R., Olsson, C.*, et al.* (2011). Copy number variation and selection during reprogramming to pluripotency. Nature *471,* 58-62.

Hyslop, L., Stojkovic, M., Armstrong, L., Walter, T., Stojkovic, P., Przyborski, S., Herbert, M., Murdoch, A., Strachan, T., and Lako, M. (2005). Downregulation of NANOG induces differentiation of human embryonic stem cells to extraembryonic lineages. Stem Cells *23,* 1035-1043.

Imreh, M.P., Gertow, K., Cedervall, J., Unger, C., Holmberg, K., Szoke, K., Csoregh, L., Fried, G., Dilber, S., Blennow, E., and Ahrlund-Richter, L. (2006). In vitro culture conditions favoring selection of chromosomal abnormalities in human ES cells. J. Cell. Biochem. *99,* 508-516.

International Stem Cell Initiative, Adewumi, O., Aflatoonian, B., Ahrlund-Richter, L., Amit, M., Andrews, P.W., Beighton, G., Bello, P.A., Benvenisty, N., Berry, L.S.*, et al.* (2007). Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. Nat. Biotechnol. *25,* 803-816.

International Stem Cell Initiative, Amps, K., Andrews, P.W., Anyfantis, G., Armstrong, L., Avery, S., Baharvand, H., Baker, J., Baker, D., Munoz, M.B.*, et al.* (2011). Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage. Nat. Biotechnol. *29,* 1132-1144.

Inzunza, J., Sahlen, S., Holmberg, K., Stromberg, A.M., Teerijoki, H., Blennow, E., Hovatta, O., and Malmgren, H. (2004). Comparative genomic hybridization and karyotyping of human embryonic stem cells reveals the occurrence of an isodicentric X chromosome after long-term cultivation. Mol. Hum. Reprod. *10,* 461-466.

Iwabuchi, K.A., Yamakawa, T., Sato, Y., Ichisaka, T., Takahashi, K., Okita, K., and Yamanaka, S. (2011). ECAT11/L1td1 Is Enriched in ESCs and Rapidly Activated During iPSCGeneration, but It Is Dispensable for the Maintenance and Induction of Pluripotency. PLoS One *6,* e20461.

Iyer, N.V., Kotch, L.E., Agani, F., Leung, S.W., Laughner, E., Wenger, R.H., Gassmann, M., Gearhart, J.D., Lawler, A.M., Yu, A.Y., and Semenza, G.L. (1998). Cellular and developmental

control of O2 homeostasis by hypoxia-inducible factor 1 alpha. Genes Dev. *12,* 149-162.

James, D., Levine, A.J., Besser, D., and Hemmati-Brivanlou, A. (2005). TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. Development *132,* 1273-1282.

Jarvinen, A.K., Autio, R., Haapa-Paananen, S., Wolf, M., Saarela, M., Grenman, R., Leivo, I., Kallioniemi, O., Makitie, A.A., and Monni, O. (2006). Identification of target genes in laryngeal squamous cell carcinoma by high-resolution copy number and gene expression microarray analyses. Oncogene *25,* 6997-7008.

Jenner, M.W., Leone, P.E., Walker, B.A., Ross, F.M., Johnson, D.C., Gonzalez, D., Chiecchio, L., Dachs Cabanas, E., Dagrada, G.P., Nightingale, M.*, et al.* (2007). Gene mapping and expression analysis of 16q loss of heterozygosity identifies WWOX and CYLD as being important in determining clinical outcome in multiple myeloma. Blood *110,* 3291-3300.

Jia, F., Wilson, K.D., Sun, N., Gupta, D.M., Huang, M., Li, Z., Panetta, N.J., Chen, Z.Y., Robbins, R.C., Kay, M.A., Longaker, M.T., and Wu, J.C. (2010). A nonviral minicircle vector for deriving human iPS cells. Nat. Methods *7,* 197-199.

Jin, J., Jing, W., Lei, X.X., Feng, C., Peng, S., Boris-Lawrie, K., and Huang, Y. (2011). Evidence that Lin28 stimulates translation by recruiting RNA helicase A to polysomes. Nucleic Acids Res. *39,* 3724-3734.

Jokilehto, T., and Jaakkola, P.M. (2010). The role of HIF prolyl hydroxylases in tumour growth. J. Cell. Mol. Med. *14,* 758-770.

Kaji, K., Norrby, K., Paca, A., Mileikovsky, M., Mohseni, P., and Woltjen, K. (2009). Virus-free induction of pluripotency and subsequent excision of reprogramming factors. Nature *458,* 771-775.

Kang, H.B., Kim, J.S., Kwon, H.J., Nam, K.H., Youn, H.S., Sok, D.E., and Lee, Y. (2005). Basic fibroblast growth factor activates ERK and induces c-fos in human embryonic stem cell line MizhES1. Stem Cells Dev. *14,* 395-401.

Katoh, M., and Katoh, M. (2007). Integrative genomic analyses on HES/HEY family: Notch-independent HES1, HES3 transcription in undifferentiated ES cells, and Notch-dependent HES1, HES5, HEY1, HEY2, HEYL transcription in fetal tissues, adult tissues, or cancer. Int. J. Oncol. *31,* 461-466.

Khazina, E., and Weichenrieder, O. (2009). Non-LTR retrotransposons encode noncanonical RRM domains in their first open reading frame. Proc. Natl. Acad. Sci. U. S. A. *106,* 731-736.

Kilpinen, S., Autio, R., Ojala, K., Iljin, K., Bucher, E., Sara, H., Pisto, T., Saarela, M., Skotheim, R.I., Bjorkman, M.*, et al.* (2008). Systematic bioinformatic analysis of expression levels of 17,330 human genes across 9,783 samples from 175 types of healthy and pathological tissues. Genome Biol. *9,* R139.

Kim, D., Kim, C.H., Moon, J.I., Chung, Y.G., Chang, M.Y., Han, B.S., Ko, S., Yang, E., Cha, K.Y., Lanza, R., and Kim, K.S. (2009). Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. Cell. Stem Cell. *4,* 472-476.

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M.J., Ji, H., Ehrlich, L.I.*, et al.* (2010). Epigenetic memory in induced pluripotent stem cells. Nature *467,* 285-290.

Kleinsmith, L.J., and Pierce, G.B.,Jr. (1964). Multipotentiality of Single Embryonal Carcinoma Cells. Cancer Res. *24,* 1544-1551.

Kolosha, V.O., and Martin, S.L. (1997). In vitro properties of the first ORF protein from mouse LINE-1 support its role in ribonucleoprotein particle formation during retrotransposition. Proc. Natl. Acad. Sci. U. S. A. *94,* 10155-10160.

Kotch, L.E., Iyer, N.V., Laughner, E., and Semenza, G.L. (1999). Defective vascularization of HIF-1alpha-null embryos is not associated with VEGF deficiency but with mesenchymal cell death. Dev. Biol. *209,* 254-267.

Lakshmipathy, U., Love, B., Goff, L.A., Jornsten, R., Graichen, R., Hart, R.P., and Chesnut, J.D. (2007). MicroRNA expression pattern of undifferentiated and differentiated human embryonic stem cells. Stem Cells Dev. *16,* 1003-1016.

Langer-Safer, P.R., Levine, M., and Ward, D.C. (1982). Immunological method for mapping genes on Drosophila polytene chromosomes. Proc. Natl. Acad. Sci. U. S. A. *79,* 4381-4385.

Laurent, L., Wong, E., Li, G., Huynh, T., Tsirigos, A., Ong, C.T., Low, H.M., Kin Sung, K.W., Rigoutsos, I., Loring, J., and Wei, C.L. (2010). Dynamic changes in the human methylome during differentiation. Genome Res. *20,* 320-331.

Laurent, L.C., Chen, J., Ulitsky, I., Mueller, F.J., Lu, C., Shamir, R., Fan, J.B., and Loring, J.F. (2008). Comprehensive microRNA profiling reveals a unique human embryonic stem cell signature dominated by a single seed sequence. Stem Cells *26,* 1506-1516.

Laurent, L.C., Ulitsky, I., Slavin, I., Tran, H., Schork, A., Morey, R., Lynch, C., Harness, J.V., Lee, S., Barrero, M.J., *et al.* (2011). Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture. Cell. Stem Cell. *8,* 106-118.

Lefort, N., Feyeux, M., Bas, C., Feraud, O., Bennaceur-Griscelli, A., Tachdjian, G., Peschanski, M., and Perrier, A.L. (2008). Human embryonic stem cells reveal recurrent genomic instability at 20q11.21. Nat. Biotechnol.

Lengner, C.J., Gimelbrant, A.A., Erwin, J.A., Cheng, A.W., Guenther, M.G., Welstead, G.G., Alagappan, R., Frampton, G.M., Xu, P., Muffat, J., *et al.* (2010). Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. Cell *141,* 872-883.

Levenstein, M.E., Ludwig, T.E., Xu, R.H., Llanas, R.A., VanDenHeuvel-Kramer, K., Manning, D., and Thomson, J.A. (2006). Basic fibroblast growth factor support of human embryonic stem cell self-renewal. Stem Cells *24,* 568-574.

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., and Xu, G.L. (2007). Synergistic function of DNA methyltransferases Dnmt3a and Dnmt3b in the methylation of Oct4 and Nanog. Mol. Cell. Biol. *27,* 8748-8759.

Li, W., Zhou, H., Abujarour, R., Zhu, S., Young Joo, J., Lin, T., Hao, E., Scholer, H.R., Hayek, A., and Ding, S. (2009). Generation of human-induced pluripotent stem cells in the absence of exogenous Sox2. Stem Cells *27,* 2992-3000.

Li, Z., Van Calcar, S., Qu, C., Cavenee, W.K., Zhang, M.Q., and Ren, B. (2003). A global transcriptional regulatory role for c-Myc in Burkitt's lymphoma cells. Proc. Natl. Acad. Sci. U. S. A. *100,* 8164-8169.

Lin, S.L., Chang, D.C., Lin, C.H., Ying, S.Y., Leu, D., and Wu, D.T. (2011). Regulation of somatic cell reprogramming through inducible mir-302 expression. Nucleic Acids Res. *39,* 1054-1065.

Lindblom, A., Rotstein, S., Skoog, L., Nordenskjold, M., and Larsson, C. (1993). Deletions on chromosome 16 in primary familial breast carcinomas are associated with development of distant metastases. Cancer Res. *53,* 3707-3711.

Lipchina, I., Elkabetz, Y., Hafner, M., Sheridan, R., Mihailovic, A., Tuschl, T., Sander, C., Studer, L., and Betel, D. (2011). Genome-wide identification of microRNA targets in human ES cells reveals a role for miR-302 in modulating BMP response. Genes Dev. *25,* 2173-2186.

Loh, Y.H., Hartung, O., Li, H., Guo, C., Sahalie, J.M., Manos, P.D., Urbach, A., Heffner, G.C., Grskovic, M., Vigneault, F., *et al.* (2010). Reprogramming of T cells from human peripheral blood. Cell. Stem Cell. *7,* 15-19.

Ludwig, T.E., Levenstein, M.E., Jones, J.M., Berggren, W.T., Mitchen, E.R., Frane, J.L., Crandall, L.J., Daigh, C.A., Conard, K.R., Piekarczyk, M.S., Llanas, R.A., and Thomson, J.A. (2006). Derivation of human embryonic stem cells in defined conditions. Nat. Biotechnol. *24,* 185-187.

Maitra, A., Arking, D.E., Shivapurkar, N., Ikeda, M., Stastny, V., Kassauei, K., Sui, G., Cutler, D.J., Liu, Y., Brimble, S.N., *et al.* (2005). Genomic alterations in cultured human embryonic stem cells. Nat. Genet. *37,* 1099-1103.

Majmundar, A.J., Wong, W.J., and Simon, M.C. (2010). Hypoxia-inducible factors and the response to hypoxic stress. Mol. Cell *40,* 294-309.

Mallanna, S.K., and Rizzino, A. (2010). Emerging roles of microRNAs in the control of embryonic stem cells and the generation of induced pluripotent stem cells. Dev. Biol. *344,* 16-25.

Maltepe, E., Schmidt, J.V., Baunoch, D., Bradfield, C.A., and Simon, M.C. (1997). Abnormal angiogenesis and responses to glucose and oxygen deprivation in mice lacking the protein ARNT. Nature *386,* 403-407.

Mantel, C., Guo, Y., Lee, M.R., Kim, M.K., Han, M.K., Shibayama, H., Fukuda, S., Yoder, M.C., Pelus, L.M., Kim, K.S., and Broxmeyer, H.E. (2007). Checkpoint-apoptosis uncoupling in human and mouse embryonic stem cells: a source of karyotpic instability. Blood *109,* 4518-4527.

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,* 7634-7638.

Martins-Taylor, K., Nisler, B.S., Taapken, S.M., Compton, T., Crandall, L., Montgomery, K.D., Lalande, M., and Xu, R.H. (2011). Recurrent copy number variations in human induced pluripotent stem cells. Nat. Biotechnol. *29,* 488-491.

Martins-Taylor, K., Schroeder, D.I., Lasalle, J.M., Lalande, M., and Xu, R.H. (2012). Role of DNMT3B in the regulation of early neural and neural crest specifiers. Epigenetics *7,*

Matin, M.M., Walsh, J.R., Gokhale, P.J., Draper, J.S., Bahrami, A.R., Morton, I., Moore, H.D., and Andrews, P.W. (2004). Specific knockdown of Oct4 and beta2-microglobulin expression by RNA interference in human embryonic stem cells and embryonic carcinoma cells. Stem Cells *22,* 659-668.

Mayshar, Y., Ben-David, U., Lavon, N., Biancotti, J.C., Yakir, B., Clark, A.T., Plath, K., Lowry, W.E., and Benvenisty, N. (2010). Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. Cell. Stem Cell. *7,* 521-531.

Melchior, K., Weiss, J., Zaehres, H., Kim, Y.M., Lutzko, C., Roosta, N., Hescheler, J., and Muschen, M. (2008). The WNT receptor FZD7 contributes to self-renewal signaling of human embryonic stem cells. Biol. Chem. *389,* 897-903.

Mitalipova, M.M., Rao, R.R., Hoyer, D.M., Johnson, J.A., Meisner, L.F., Jones, K.L., Dalton, S., and Stice, S.L. (2005). Preserving the genetic integrity of human embryonic stem cells. Nat. Biotechnol. *23,* 19-20.

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

Miyoshi, N., Ishii, H., Nagano, H., Haraguchi, N., Dewi, D.L., Kano, Y., Nishikawa, S., Tanemura, M., Mimori, K., Tanaka, F*., et al.* (2011). Reprogramming of mouse and human cells to pluripotency using mature microRNAs. Cell. Stem Cell. *8,* 633-638.

Moon, S.H., Kim, J.S., Park, S.J., Lim, J.J., Lee, H.J., Lee, S.M., and Chung, H.M. (2011). Effect of chromosome instability on the maintenance and differentiation of human embryonic stem cells in vitro and in vivo. Stem Cell. Res. *6,* 50-59.

Moran, J.V., Holmes, S.E., Naas, T.P., DeBerardinis, R.J., Boeke, J.D., and Kazazian, H.H.,Jr. (1996). High frequency retrotransposition in cultured mammalian cells. Cell *87,* 917-927.

Morgan, H.D., Santos, F., Green, K., Dean, W., and Reik, W. (2005). Epigenetic reprogramming in mammals. Hum. Mol. Genet. *14 Spec No 1,* R47-58.

Morin, R.D., O'Connor, M.D., Griffith, M., Kuchenbauer, F., Delaney, A., Prabhu, A.L., Zhao, Y., McDonald, H., Zeng, T., Hirst, M., Eaves, C.J., and Marra, M.A. (2008). Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. Genome Res. *18,* 610-621.

Moustakas, A., and Heldin, C.H. (2009). The regulation of TGFbeta signal transduction. Development *136,* 3699-3714.

Muller, F.J., Laurent, L.C., Kostka, D., Ulitsky, I., Williams, R., Lu, C., Park, I.H., Rao, M.S., Shamir, R., Schwartz, P.H., Schmidt, N.O., and Loring, J.F. (2008). Regulatory networks define phenotypic classes of human stem cell lines. Nature *455,* 401-405.

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat. Biotechnol. *26,* 101-106.

Narva, E., Autio, R., Rahkonen, N., Kong, L., Harrison, N., Kitsberg, D., Borghese, L., Itskovitz-Eldor, J., Rasool, O., Dvorak, P*., et al.* (2010). High-resolution DNA analysis of human embryonic stem cell lines reveals culture-induced copy number changes and loss of heterozygosity. Nat. Biotechnol.

Newman, M.A., Thomson, J.M., and Hammond, S.M. (2008). Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. RNA *14,* 1539-1549.

Ng, H.H., and Surani, M.A. (2011). The transcriptional and signalling networks of pluripotency. Nat. Cell Biol. *13,* 490-496.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell *95,* 379-391.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat. Genet. *24,* 372-376.

Nusse, R., Fuerer, C., Ching, W., Harnish, K., Logan, C., Zeng, A., ten Berge, D., and Kalani, Y. (2008). Wnt signaling and stem cell control. Cold Spring Harb. Symp. Quant. Biol. *73,* 59-66.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99,* 247-257.

Ory, D.S., Neugeboren, B.A., and Mulligan, R.C. (1996). A stable human-derived packaging cell line for production of high titer retrovirus/vesicular stomatitis virus G pseudotypes. Proc. Natl. Acad. Sci. U. S. A. *93,* 11400-11406.

Pan, G., Tian, S., Nie, J., Yang, C., Ruotti, V., Wei, H., Jonsdottir, G.A., Stewart, R., and Thomson, J.A. (2007). Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. Cell. Stem Cell. *1,* 299-312.

Peng, J., Zhang, L., Drysdale, L., and Fong, G.H. (2000). The transcription factor EPAS-1/hypoxia-inducible factor 2alpha plays an important role in vascular remodeling. Proc. Natl. Acad. Sci. U. S. A. *97,* 8386-8391.

Peng, S., Chen, L.L., Lei, X.X., Yang, L., Lin, H., Carmichael, G.G., and Huang, Y. (2011). Genome-wide studies reveal that Lin28 enhances the translation of genes important for growth and survival of human embryonic stem cells. Stem Cells *29,* 496-504.

Pera, M.F., Andrade, J., Houssami, S., Reubinoff, B., Trounson, A., Stanley, E.G., Ward-van Oostwaard, D., and Mummery, C. (2004). Regulation of human embryonic stem cell differentiation by BMP-2 and its antagonist noggin. J. Cell. Sci. *117,* 1269-1280.

Piskounova, E., Viswanathan, S.R., Janas, M., LaPierre, R.J., Daley, G.Q., Sliz, P., and Gregory, R.I. (2008). Determinants of microRNA processing inhibition by the developmentally regulated RNA-binding protein Lin28. J. Biol. Chem. *283,* 21310-21314.

Plaia, T.W., Josephson, R., Liu, Y., Zeng, X., Ording, C., Toumadje, A., Brimble, S.N., Sherrer, E.S., Uhl, E.W., Freed, W.J.*, et al.* (2006). Characterization of a new NIH-registered variant human embryonic stem cell line, BG01V: a tool for human embryonic stem cell research. Stem Cells *24,* 531-546.

Prado-Lopez, S., Conesa, A., Arminan, A., Martinez-Losa, M., Escobedo-Lucea, C., Gandia, C., Tarazona, S., Melguizo, D., Blesa, D., Montaner, D.*, et al.* (2010). Hypoxia promotes efficient differentiation of human embryonic stem cells to functional endothelium. Stem Cells *28,* 407-418.

Prasad, S.M., Czepiel, M., Cetinkaya, C., Smigielska, K., Weli, S.C., Lysdahl, H., Gabrielsen, A., Petersen, K., Ehlers, N., Fink, T., Minger, S.L., and Zachar, V. (2009). Continuous hypoxic culturing maintains activation of Notch and allows long-term propagation of human embryonic stem cells without spontaneous differentiation. Cell Prolif. *42,* 63-74.

Qi, J., Yu, J.Y., Shcherbata, H.R., Mathieu, J., Wang, A.J., Seal, S., Zhou, W., Stadler, B.M., Bourgin, D., Wang, L.*, et al.* (2009). microRNAs regulate human embryonic stem cell division. Cell. Cycle *8,* 3729-3741.

Qiu, C., Ma, Y., Wang, J., Peng, S., and Huang, Y. (2010). Lin28-mediated post-transcriptional regulation of Oct4 expression in human embryonic stem cells. Nucleic Acids Res. *38,* 1240-1248.

R Development Core Team. (2008). R: A Language and Environment for Statistical Computing (manual).

Ren, J., Jin, P., Wang, E., Marincola, F.M., and Stroncek, D.F. (2009). MicroRNA and gene expression patterns in the differentiation of human embryonic stem cells. J. Transl. Med. *7,* 20.

Reubinoff, B.E., Pera, M.F., Fong, C.Y., Trounson, A., and Bongso, A. (2000). Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. Nat. Biotechnol. *18,* 399-404.

Richards, M., Tan, S.P., Tan, J.H., Chan, W.K., and Bongso, A. (2004). The transcriptome profile of human embryonic stem cells as defined by SAGE. Stem Cells *22,* 51-64.

Rosa, A., Spagnoli, F.M., and Brivanlou, A.H. (2009). The miR-430/427/302 family controls mesendodermal fate specification via species-specific target selection. Dev. Cell. *16,* 517-527.

Rosler, E.S., Fisk, G.J., Ares, X., Irving, J., Miura, T., Rao, M.S., and Carpenter, M.K. (2004). Long-term culture of human embryonic stem cells in feeder-free conditions. Dev. Dyn. *229,* 259-274.

Ryan, H.E., Lo, J., and Johnson, R.S. (1998). HIF-1 alpha is required for solid tumor formation and embryonic vascularization. EMBO J. *17,* 3005-3015.

Rybak, A., Fuchs, H., Smirnova, L., Brandt, C., Pohl, E.E., Nitsch, R., and Wulczyn, F.G. (2008). A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. Nat. Cell Biol. *10,* 987-993.

Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and Brivanlou, A.H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. Nat. Med. *10,* 55-63.

Sato, N., Sanjuan, I.M., Heke, M., Uchida, M., Naef, F., and Brivanlou, A.H. (2003). Molecular signature of human embryonic stem cells and its comparison with the mouse. Dev. Biol. *260,* 404-413.

Sato-Otsubo, A., Sanada, M., and Ogawa, S. (2012). Single-nucleotide polymorphism array karyotyping in clinical practice: where, when, and how? Semin. Oncol. *39,* 13-25.

Scholer, H.R., Ruppert, S., Suzuki, N., Chowdhury, K., and Gruss, P. (1990). New type of POU domain in germ line-specific protein Oct-4. Nature *344,* 435-439.

Seitz, V., Butzhammer, P., Hirsch, B., Hecht, J., Gutgemann, I., Ehlers, A., Lenze, D., Oker, E., Sommerfeld, A., von der Wall, E.*, et al.* (2011). Deep sequencing of MYC DNA-binding sites in Burkitt lymphoma. PLoS One *6,* e26837.

Seki, T., Yuasa, S., Oda, M., Egashira, T., Yae, K., Kusumoto, D., Nakata, H., Tohyama, S., Hashimoto, H., Kodaira, M.*, et al.* (2010). Generation of induced pluripotent stem cells from human terminally differentiated circulating T cells. Cell. Stem Cell. *7,* 11-14.

Semenza, G.L. (2012). Hypoxia-inducible factors in physiology and medicine. Cell *148,* 399-408.

Sen, G.L., and Blau, H.M. (2005). Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies. Nat. Cell Biol. *7,* 633-636.

Silva, S.S., Rowntree, R.K., Mekhoubad, S., and Lee, J.T. (2008). X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. *105,* 4820-4825.

Simon, M.C., and Keith, B. (2008). The role of oxygen availability in embryonic development and stem cell function. Nat. Rev. Mol. Cell Biol. *9,* 285-296.

Si-Tayeb, K., Noto, F.K., Nagaoka, M., Li, J., Battle, M.A., Duris, C., North, P.E., Dalton, S., and Duncan, S.A. (2010). Highly efficient generation of human hepatocyte-like cells from induced pluripotent stem cells. Hepatology *51,* 297-305.

Smyth, G., K. (2005). Limma: linear models for microarray data (incollection).397-420.

Soldner, F., Hockemeyer, D., Beard, C., Gao, Q., Bell, G.W., Cook, E.G., Hargus, G., Blak, A., Cooper, O., Mitalipova, M., Isacson, O., and Jaenisch, R. (2009). Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. Cell *136,* 964-977.

Sperger, J.M., Chen, X., Draper, J.S., Antosiewicz, J.E., Chon, C.H., Jones, S.B., Brooks, J.D., Andrews, P.W., Brown, P.O., and Thomson, J.A. (2003). Gene expression patterns in human embryonic stem cells and human pluripotent germ cell tumors. Proc. Natl. Acad. Sci. U. S. A. *100,* 13350-13355.

Spits, C., Mateizel, I., Geens, M., Mertzanidou, A., Staessen, C., Vandeskelde, Y., Van der Elst, J., Liebaers, I., and Sermon, K. (2008). Recurrent chromosomal abnormalities in human embryonic stem cells. Nat. Biotechnol.

Staerk, J., Dawlaty, M.M., Gao, Q., Maetzel, D., Hanna, J., Sommer, C.A., Mostoslavsky, G., and Jaenisch, R. (2010). Reprogramming of human peripheral blood cells to induced pluripotent stem cells. Cell. Stem Cell. *7,* 20-24.

Stranger, B.E., Forrest, M.S., Dunning, M., Ingle, C.E., Beazley, C., Thorne, N., Redon, R., Bird, C.P., de Grassi, A., Lee, C.*, et al.* (2007). Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science *315,* 848-853.

Subramanyam, D., Lamouille, S., Judson, R.L., Liu, J.Y., Bucay, N., Derynck, R., and Blelloch, R. (2011). Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells. Nat. Biotechnol. *29,* 443-448.

Suh, M.R., Lee, Y., Kim, J.Y., Kim, S.K., Moon, S.H., Lee, J.Y., Cha, K.Y., Chung, H.M., Yoon, H.S., Moon, S.Y., Kim, V.N., and Kim, K.S. (2004). Human embryonic stem cells express a unique set of microRNAs. Dev. Biol. *270,* 488-498.

Sumi, T., Tsuneyoshi, N., Nakatsuji, N., and Suemori, H. (2007). Apoptosis and differentiation of human embryonic stem cells induced by sustained activation of c-Myc. Oncogene *26,* 5564-5576.

Sun, X., Long, X., Yin, Y., Jiang, Y., Chen, X., Liu, W., Zhang, W., Du, H., Li, S., Zheng, Y.*, et al.* (2008). Similar biological characteristics of human embryonic stem cell lines with normal and abnormal karyotypes. Hum. Reprod. *23,* 2185-2193.

Taapken, S.M., Nisler, B.S., Newton, M.A., Sampsell-Barron, T.L., Leonhard, K.A., McIntire, E.M., and Montgomery, K.D. (2011). Karotypic abnormalities in human induced pluripotent stem cells and embryonic stem cells. Nat. Biotechnol. *29,* 313-314.

Tada, M., Takahama, Y., Abe, K., Nakatsuji, N., and Tada, T. (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. Curr. Biol. *11,* 1553-1558.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131,* 861-872.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126,* 663-676.

Tang, C., Lee, A.S., Volkmer, J.P., Sahoo, D., Nag, D., Mosley, A.R., Inlay, M.A., Ardehali, R., Chavez, S.L., Pera, R.R.*, et al.* (2011). An antibody against SSEA-5 glycan on human pluripotent stem cells enables removal of teratoma-forming cells. Nat. Biotechnol. *29,* 829-834.

Teixeira, D., Sheth, U., Valencia-Sanchez, M.A., Brengues, M., and Parker, R. (2005). Processing bodies require RNA for assembly and contain nontranslating mRNAs. RNA *11,* 371-382.

The International HapMap Consortium. (2003). The International HapMap Project. Nature *426,* 789-796.

Thomson, A., J., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, A.M., Swiergiel, J.J., Marshall, S.V., and Jones, M.J. (1998). Embryonic Stem Cell Lines Derived from Human Blastocysts. November *282,* 1145-1147.

Thomson, J.A., Kalishman, J., Golos, T.G., Durning, M., Harris, C.P., Becker, R.A., and Hearn, J.P. (1995). Isolation of a primate embryonic stem cell line. Proc. Natl. Acad. Sci. U. S. A. *92,* 7844-7848.

Tzur, G., Levy, A., Meiri, E., Barad, O., Spector, Y., Bentwich, Z., Mizrahi, L., Katzenellenbogen, M., Ben-Shushan, E., Reubinoff, B.E., and Galun, E. (2008). MicroRNA expression patterns and function in endodermal differentiation of human embryonic stem cells. PLoS One *3,* e3726.

Utting, J.C., Robins, S.P., Brandao-Burch, A., Orriss, I.R., Behar, J., and Arnett, T.R. (2006). Hypoxia inhibits the growth, differentiation and bone-forming capacity of rat osteoblasts. Exp. Cell Res. *312,* 1693-1702.

Vallier, L., Alexander, M., and Pedersen, R.A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. J. Cell. Sci. *118,* 4495-4509.

Vallier, L., Reynolds, D., and Pedersen, R.A. (2004). Nodal inhibits differentiation of human embryonic stem cells along the neuroectodermal default pathway. Dev. Biol. *275,* 403-421.

Vanneste, E., Voet, T., Le Caignec, C., Ampe, M., Konings, P., Melotte, C., Debrock, S., Amyere, M., Vikkula, M., Schuit, F.*, et al.* (2009). Chromosome instability is common in human cleavage-stage embryos. Nat. Med. *15,* 577-583.

Vazin, T., and Freed, W.J. (2010). Human embryonic stem cells: derivation, culture, and differentiation: a review. Restor. Neurol. Neurosci. *28,* 589-603.

Viswanathan, S.R., Daley, G.Q., and Gregory, R.I. (2008). Selective blockade of microRNA processing by Lin28. Science *320,* 97-100.

Wang, S.H., Tsai, M.S., Chiang, M.F., and 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, Z., Oron, E., Nelson, B., Razis, S., and Ivanova, N. (2012). Distinct Lineage Specification Roles for NANOG, OCT4, and SOX2 in Human Embryonic Stem Cells. Cell. Stem Cell. *10,* 440-454.

Warren, L., Manos, P.D., Ahfeldt, T., Loh, Y.H., Li, H., Lau, F., Ebina, W., Mandal, P.K., Smith, Z.D., Meissner, A.*, et al.* (2010). Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. Cell. Stem Cell. *7,* 618-630.

Werbowetski-Ogilvie, T.E., Bosse, M., Stewart, M., Schnerch, A., Ramos-Mejia, V., Rouleau, A., Wynder, T., Smith, M.J., Dingwall, S., Carter, T.*, et al.* (2009). Characterization of human embryonic stem cells with features of neoplastic progression. Nat. Biotechnol. *27,* 91-97.

Westfall, S.D., Sachdev, S., Das, P., Hearne, L.B., Hannink, M., Roberts, R.M., and Ezashi, T. (2008). Identification of oxygen-sensitive transcriptional programs in human embryonic stem cells. Stem Cells Dev. *17,* 869-881.

Wilmut, I., Schnieke, A.E., McWhir, J., Kind, A.J., and Campbell, K.H. (1997). Viable offspring derived from fetal and adult mammalian cells. Nature *385,* 810-813.

Wilson, K.D., Venkatasubrahmanyam, S., Jia, F., Sun, N., Butte, A.J., and Wu, J.C. (2009). MicroRNA profiling of human-induced pluripotent stem cells. Stem Cells Dev. *18,* 749-758.

Winter, J., Jung, S., Keller, S., Gregory, R.I., and Diederichs, S. (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation. Nat. Cell Biol. *11,* 228-234.

Wissing, S., Munoz-Lopez, M., Macia, A., Yang, Z., Montano, M., Collins, W., Garcia-Perez, J.L., Moran, J.V., and Greene, W.C. (2012). Reprogramming somatic cells into iPS cells activates LINE-1 retroelement mobility. Hum. Mol. Genet. *21,* 208-218.

Woltjen, K., Michael, I.P., Mohseni, P., Desai, R., Mileikovsky, M., Hamalainen, R., Cowling, R., Wang, W., Liu, P., Gertsenstein, M.*, et al.* (2009). piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature *458,* 766-770.

Wong, R.C., Ibrahim, A., Fong, H., Thompson, N., Lock, L.F., and Donovan, P.J. (2011). L1TD1 Is a Marker for Undifferentiated Human Embryonic Stem Cells. PLoS One *6,* e19355.

Wright, W.E., Piatyszek, M.A., Rainey, W.E., Byrd, W., and Shay, J.W. (1996). Telomerase activity in human germline and embryonic tissues and cells. Dev. Genet. *18,* 173-179.

Wu, H., Kim, K.J., Mehta, K., Paxia, S., Sundstrom, A., Anantharaman, T., Kuraishy, A.I., Doan, T., Ghosh, J., Pyle, A.D.*, et al.* (2008). Copy number variant analysis of human embryonic stem cells. Stem Cells *26,* 1484-1489.

Wulczyn, F.G., Smirnova, L., Rybak, A., Brandt, C., Kwidzinski, E., Ninnemann, O., Strehle, M., Seiler, A., Schumacher, S., and Nitsch, R. (2007). Post-transcriptional regulation of the let-7 microRNA during neural cell specification. FASEB J. *21,* 415-426.

Xiao, L., Yuan, X., and Sharkis, S.J. (2006). Activin A maintains self-renewal and regulates fibroblast growth factor, Wnt, and bone morphogenic protein pathways in human embryonic stem cells. Stem Cells *24,* 1476-1486.

Xu, N., Papagiannakopoulos, T., Pan, G., Thomson, J.A., and Kosik, K.S. (2009). MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. Cell *137,* 647-658.

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

Xu, R.H., Peck, R.M., Li, D.S., Feng, X., Ludwig, T., and Thomson, J.A. (2005). Basic FGF and suppression of BMP signaling sustain undifferentiated proliferation of human ES cells. Nat. Methods *2,* 185-190.

Xu, R.H., Sampsell-Barron, T.L., Gu, F., Root, S., Peck, R.M., Pan, G., Yu, J., Antosiewicz-Bourget, J., Tian, S., Stewart, R., and Thomson, J.A. (2008). NANOG is a direct target of TGFbeta/activin-mediated SMAD signaling in human ESCs. Cell. Stem Cell. *3,* 196-206.

Yang, S., Lin, G., Tan, Y.Q., Deng, L.Y., Yuan, D., and Lu, G.X. (2010). Differences between karyotypically normal and abnormal human embryonic stem cells. Cell Prolif. *43,* 195-206.

Yang, S., Lin, G., Tan, Y.Q., Zhou, D., Deng, L.Y., Cheng, D.H., Luo, S.W., Liu, T.C., Zhou, X.Y., Sun, Z.*, et al.* (2008). Tumor progression of culture-adapted human embryonic stem cells during long-term culture. Genes Chromosomes Cancer *47,* 665-679.

Yoshida, Y., Takahashi, K., Okita, K., Ichisaka, T., and Yamanaka, S. (2009). Hypoxia enhances the generation of induced pluripotent stem cells. Cell. Stem Cell. *5,* 237-241.

Yu, J., Chau, K.F., Vodyanik, M.A., Jiang, J., and Jiang, Y. (2011). Efficient feeder-free episomal reprogramming with small molecules. PLoS One *6,* e17557.

Yu, J., Hu, K., Smuga-Otto, K., Tian, S., Stewart, R., Slukvin, I.I., and Thomson, J.A. (2009). Human induced pluripotent stem cells free of vector and transgene sequences. Science *324,* 797-801.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., Slukvin, I.I., and Thomson, J.A. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318,* 1917-1920.

Yu, P., Pan, G., Yu, J., and Thomson, J.A. (2011). FGF2 sustains NANOG and switches the outcome of BMP4-induced human embryonic stem cell differentiation. Cell. Stem Cell. *8,* 326-334.

Yuan, H., Corbi, N., Basilico, C., and Dailey, L. (1995). Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. Genes Dev. *9,* 2635-2645.

Zaehres, H., Lensch, M.W., Daheron, L., Stewart, S.A., Itskovitz-Eldor, J., and Daley, G.Q. (2005). High-efficiency RNA interference in human embryonic stem cells. Stem Cells *23,* 299-305.

Zafarana, G., Avery, S.R., Avery, K., Moore, H.D., and Andrews, P.W. (2009). Specific knockdown of OCT4 in human embryonic stem cells by inducible short hairpin RNA interference. Stem Cells *27,* 776-782.

Zhang, F., Gu, W., Hurles, M.E., and Lupski, J.R. (2009a). Copy number variation in human health, disease, and evolution. Annu. Rev. Genomics Hum. Genet. *10,* 451-481.

Zhang, X., Neganova, I., Przyborski, S., Yang, C., Cooke, M., Atkinson, S.P., Anyfantis, G., Fenyk, S., Keith, W.N., Hoare, S.F.*, et al.* (2009b). A role for NANOG in G1 to S transition in human embryonic stem cells through direct binding of CDK6 and CDC25A. J. Cell Biol. *184,* 67-82.

Zhao, T., Zhang, Z.N., Rong, Z., and Xu, Y. (2011). Immunogenicity of induced pluripotent stem cells. Nature *474,* 212-215.

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.*, et al.* (2007). Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. Cell. Stem Cell. *1,* 286-298.

Zhou, T., Benda, C., Duzinger, S., Huang, Y., Li, X., Li, Y., Guo, X., Cao, G., Chen, S., Hao, L.*, et al.* (2011). Generation of induced pluripotent stem cells from urine. J. Am. Soc. Nephrol. *22,* 1221-1228.

Zhu, S., Li, W., Zhou, H., Wei, W., Ambasudhan, R., Lin, T., Kim, J., Zhang, K., and Ding, S. (2010). Reprogramming of human primary somatic cells by OCT4 and chemical compounds. Cell. Stem Cell. *7,* 651-655.

Zuo, C., Liang, S., Wang, Z., Li, H., Zheng, W., and Ma, W. (2009). Enriching protein-protein and functional interaction networks in human embryonic stem cells. Int. J. Mol. Med. *23,* 811-819.

Zwaka, T.P., and Thomson, J.A. (2005). Differentiation of human embryonic stem cells occurs through symmetric cell division. Stem Cells *23,* 146-149.