TURUN YLIOPISTON JULKAISUJA
ANNALES UNIVERSITATIS TURKUENSIS

SARJA - SER. D OSA - TOM. 1084

MEDICA - ODONTOLOGICA

REGULATION OF SELF-RENEWAL AND DETECTION OF KARYOTYPIC CHANGES OF PLURIPOTENT HUMAN EMBRYONIC STEM CELLS

by

Nelly Rahkonen

TURUN YLIOPISTO
UNIVERSITY OF TURKU
Turku 2013

From the Institute of Biomedicine, 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, M.D., Ph.D.
Turku Centre for Biotechnology
University of Turku and Åbo Akademi University
Turku, Finland

And

Doctor Riikka Lund, Ph.D.
Turku Centre for Biotechnology
University of Turku and Åbo Akademi University
Turku, Finland

Reviewed by

Professor Jari Koistinaho, M.D., Ph.D.
University of Eastern Finland
Kuopio, Finland

And

Professor Olli Silvennoinen, M.D., Ph.D.
University of Tampere,
Tampere, Finland

Opponent

Professor Francis Stewart, Ph.D.
Biotechnology Center of the TU Dresden
Dresden, Germany

The originality of this dissertation has been checked in accordance with the University of Turku quality assurance system using the Turnitin Originality Check service

ISBN 978-951-29-5477-3 (Print)
ISBN 978-951-29-5478-0 (PDF)
ISSN 0355-9483
Painosalama Oy – Turku, Finland 2013

ABSTRACT

Nelly Rahkonen
Regulation of self-renewal and detection of karyotypic changes of pluripotent human embryonic stem cells
Faculty of Biomedicine, 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 2013

Human embryonic stem cells are pluripotent cells capable of renewing themselves and differentiating to specialized cell types. Because of their unique regenerative potential, pluripotent cells offer new opportunities for disease modeling, development of regenerative therapies, and treating diseases. Before pluripotent cells can be used in any therapeutic applications, there are numerous challenges to overcome. For instance, the key regulators of pluripotency need to be clarified. In addition, long term culture of pluripotent cells is associated with the accumulation of karyotypic abnormalities, which is a concern regarding the safe use of the cells for therapeutic purposes.

The goal of the work presented in this thesis was to identify new factors involved in the maintenance of pluripotency, and to further characterize molecular mechanisms of selected candidate genes. Furthermore, we aimed to set up a new method for analyzing genomic integrity of pluripotent cells. The experimental design applied in this study involved a wide range of molecular biology, genome-wide, and computational techniques to study the pluripotency of stem cells and the functions of the target genes. In collaboration with instrument and reagent company Perkin Elmer, Karyolite™ BoBs™ was implemented for detecting karyotypic changes of pluripotent cells.

Novel genes were identified that are highly and specifically expressed in hES cells. Of these genes, L1TD1 and POLR3G were chosen for further investigation. The results revealed that both of these factors are vital for the maintenance of pluripotency and self-renewal of the hESCs. Karyolite™ BoBs™ was validated as a novel method to detect karyotypic abnormalities in pluripotent stem cells.

The results presented in this thesis offer significant new information on the regulatory networks associated with pluripotency. The results will facilitate in understanding developmental and cancer biology, as well as creating stem cell based applications. Karyolite™ BoBs™ provides rapid, high-throughput, and cost-efficient tool for screening of human pluripotent cell cultures.

Keywords: human embryonic stem cell, pluripotency, L1TD1, POLR3G, karyotype

TIIVISTELMÄ

Nelly Rahkonen
Pluripotenttien ihmisen alkion kantasolujen uusiutumiskyvyn säätely ja karyotyypin muutosten määrittäminen
Biolääketieteen laitos, Lääketieteellinen mikrobiologia ja immunologia, Turun yliopisto
Turun Biotekniikan keskus, Turun yliopisto ja Åbo Akademi
Turun biolääketieteen tutkijaohjelma
Annales Universitatis Turkuensis
Turku 2013

– –

Ihmisen alkion kantasolut ovat pluripotentteja soluja, joilla on kyky uudelleenmuodostua ja erilaistua kaikiksi ihmiskehon solutyypeiksi. Ainutlaatuisen uudistumiskykynsä ansiosta pluripotentit solut tarjoavat uusia sovellusmahdollisuuksia tautimallinnukselle, kudoskorvaushoitojen ja muiden terapioiden kehittämiselle. Ennen kuin pluripotentteja soluja voidaan turvallisesti käyttää lääketieteellisissä sovelluksissa, tulee solujen ominaisuudet ja turvallisuus tuntea yksityiskohtaisesti.

Tämän väitöskirjatutkimuksen tarkoituksena oli määrittää uusia tekijöitä, jotka ovat osallisena pluripotenssin säätelyssä, sekä tarkemmin tutkia näiden tekijöiden vaikutus-mekanismeja. Lisäksi tavoitteena oli optimoida ja validoida uusi tehoseulontamenetelmä pluripotenttien solujen genomisen eheyden määrittämiseksi. Työssä käytettiin laaja-alaisesti molekyylibiologian eri menetelmiä sekä genominlaajuisia mikrosiru- ja syvä-sekvensointitekniikoita pluripotenssin mekanismien tutkimiseksi. Karyolite™ BoBs™ menetelmä testattiin solujen karyotyypin määrittämiseksi.

Tutkimuksessa identifioitiin uusia geenejä, jotka ilmentyvät korkeasti pluripotenteissa soluissa ja joiden ilmentyminen laskee solujen erilaistuessa. Geenien joukosta valittiin L1TD1 ja POLR3G jatkotutkimusten kohteiksi. Tutkimuksen tulokset osoittivat molempien geenien olevan välttämättömiä pluripotenssin ja solujen uudelleenmuodostumisen säätelylle. Karyolite™ BoBs™ -menetelmä todettiin ihmisen pluripotenttien solujen genomisen eheyden mittaamiseen soveltuvaksi menetelmäksi.

Tämän tutkimuksen tulokset tarjoavat merkittävää uutta tietoa pluripotenssin molekylaarisista säätelymekanismeista, mikä on tärkeää lääketieteellisten kantasolusovellusten kannalta, sekä kehitysbiologian ja syöpäbiologian osa-alueilla. Karyolite™ BoBs™ -menetelmä auttaa soluviljelmien rutiininomaista analysointia tarjoamalla käyttöön uudenlaisen menetelmän, joka lyhentää analyysiin kuluvaa aikaa ja vähentää niistä johtuvia kustannuksia

Avainsanat: ihmisen alkiokantasolu, pluripotenssi, L1TD1, POLR3G, karyotyyppi

CONTENTS

ABBREVIATIONS

AGO2	argonaute RISC catalytic component 2
BAC	bacterial artificial chromosome
BMP	bone morphogenic protein
cDNA	complementary DNA
CGH	comparative genomic hybridization
ChIP	chromatin immunoprecipitation
CNV	copy number variation
CT	cycle of threshold
CTD	c-terminal domain
DNA	deoxyribonucleic acid
EB	embryonic body
EC	embryonal carcinoma
EGF	epidermal growth factor
EGFP	enhanced green fluorescent protein
EGFR	epidermal growth factor receptor
EpiSC	epiblast derived embryonic stem cell
ESC	embryonic stem cell
FCS	fetal calf serum
FGF	fibroblast growth factor
FISH	fluorescence in situ hybridization
FOXO1	forkhead box O1
G-banding	giemsa staining
GSK3β	glycogen synthase kinase 3 beta
H3K27me3	histone 3 lysine 27 tri methylation
H3K4me3	histone 3 lysine 4 tri methylation
hESC	human embryonic stem cell
HRP	horseradish peroxidase
ICM	inner cell mass
IGF	insulin-like growth factor
IGFBP	insulin-like growth factor binding protein
IGFR	insulin-like growth factor receptor
IP	immunoprecipitation
iPSC	induced pluripotent stem cell
KLF4	Kruppel-like factor 4
L1TD1	LINE 1 type transposase domain containing 1
LIF	leukemia inhibitor factor
LSD1	lysine (K)-specific demethylase 1A
MAPK	mitogen activated protein kinase
mESC	mouse embryonic stem cell
miRNA	micro RNA
mRNA	messenger RNA
ncRNA	noncoding RNA
nt	nucleotide
O_2	oxygen
OCT4	octamer-binding transcription factor 4
PABP	poly A binding protein
P-body	processing body

PCR	polymerase chain reaction
PI3K	Phosphatidylinositide 3-kinases
POL	RNA polymerase
POLR3G	polymerase (RNA) III (DNA directed) polypeptide G (32kD)
PRDM14	PR domain containing 14
pre-miRNA	precursor-microRNA
pri-miRNA	primary-microRNA
PSE	proximal sequence element
RA	retinoic acid
RBP	RNA binding protein
RHA	RNA helicase A
RISC	RNA induced silencing complex
RNA	ribonucleic acid
RNAi	RNA interference
RNP	ribonucleoprotein complex
RRM	RNA recognition motif
rRNA	ribosomal RNA
RTK	receptor tyrosine kinase
RT-qPCR	reverse transcriptase quantitative polymerase chain reaction
shRNA	short hairpin RNA
siRNA	small interfering RNA
SNAPc	small nuclear RNA activating protein complex
SNP	single nucleotide polymorphism
snRNA	small nuclear RNA
SOX2	sex determining region Y-box protein 2
TBP	TATA box binding protein
TF	transcription factor
TGF	transforming growth factor
tRNA	transfer RNA
TSS	transcription start site
XaXa	two active X chromosomes
XaXi	inactivated X chromosome
XCI	X chromosome inactivation

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following original publications, which are referred to in the text by roman numerals (I-III).

I Lund R.*, Rahkonen N.*, Malonzo M., Reddy EM., Kivinen V., Närvä E., Laiho A., Gyenesei A., Lähdesmäki H., Skottman H., Hovatta O., Rasool O., Nykter M., Lahesmaa R. Early Human Embryonic Stem Cell Marker POLR3G Regulates Intracellular Programs Vital for Self-Renewal and Early Development. Manuscript (*Equal contribution)

II Närvä E.*, Rahkonen N.*, Reddy EM.*, Lund R., Pursiheimo J., Näsli J., Autio R., Rasool O., Denessiouk K., Lähdesmäki H., Rao A., Lahesmaa R. (2012) 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 30:452-460 (*Equal contribution)

III Lund R.*, Nikula T.*, Rahkonen N., Närvä E., Baker D., Harrison N., Andrews P.W., Otonkoski T., Lahesmaa R. (2012) High-Throughput Karyotyping of Human Pluripotent Stem Cells. Stem Cell Res. 9:192-195 (*Equal contribution)

1 INTRODUCTION

Stem cell research is one of the most rapidly developing areas of biomedicine. The goal is to yield information about the complex events that take place during the development, and understand the unique properties of the cells giving rise to the new organism. Pluripotent stem cells extracted from the embryo have the property of self-renewal and capability to differentiate to all specialized cell types found in the adult organism (Evans and Kaufman, 1981; Martin, 1981). Pluripotent cells can also be artificially reprogrammed from somatic cells with the use of a set of defined inducing factors (Takahashi and Yamanaka, 2006). Pluripotent stem cells offer the potential for several new avenues of biomedical applications and research; for instance, a way to study and understand diseases, aid in toxicology screening and drug development, and even perhaps be used in regenerative medicine. However, before the pluripotent cells can be used in therapeutic purposes, there is a need to understand the molecular nature of stemness.

The mouse is a widely used organism to model many biological processes, including embryonic development. Thus, there is relatively large amount of data on mouse pluripotent stem cells and early developmental processes. However, considering therapeutic applications or disease modeling, data on human cells would be absolutely essential. There are numerous studies reporting the species specific differences in signaling systems of pluripotent stem cells (Cai et al., 2010; Ng and Surani, 2011), emphasizing the importance of human studies for future stem cell based applications. Moreover, current knowledge indicates that pluripotency is highly complex state regulated by many different factors. Hence, the molecular mechanisms of pluripotent human stem cells are not known and further research is required.

Artificial manipulation and long term culture of pluripotent stem cells is associated with culture-adaptation and accumulation of chromosomal changes (Lund et al., 2012). Chromosomal aberrations are also linked to oncogenesis and tumor progression, which is a concern regarding the safe use of pluripotent cells for therapeutic purposes. Hence, it is essential that the cells are thoroughly characterized prior to clinical applications.

In order to promote the use of pluripotent stem cells in biomedicine, the basic properties of the cells need to be defined. The objective of the work presented in this thesis was to study the mechanisms of pluripotency and self-renewal of human embryonic stem cells (hESC) and to identify novel proteins regulating the pluripotent properties of hESCs. Two novel proteins and their functions in pluripotent hESCs were thus characterized. In addition, our aims included setting up and validating a high-throughput analysis method for detecting chromosomal changes in the *in vitro* maintained stem cell cultures.

2 REVIEW OF THE LITERATURE

2.1 Pluripotent embryonic stem cells

Human body consists of numerous different cell types with a diverse range of functions. All the cells in a human body are derived, stemned, form one single cell, an oocyte fertilized by a sperm cell. This fertilized egg is called a zygote. The zygote is totipotent, in other words, it can create all possible cell types needed in the development of an individual. The zygote develops into morula, which further develops to a blastocyst. The blastocyst has a structure of inner cell mass (ICM) surrounded by trophoblast cells. The cells of ICM are pluripotent, i.e. they can differentiate to all three embryonal lineages, endo-, meso-, and ectoderm, and are the source of all cell types of developing organism. The ICM cells, with the capacity to produce differentiated cell progeny and proliferate indefinitely, are called as stem cells. The unique property of stem cells for indefinite proliferation is termed self-renewal. (Scott, 2006) (Figure 1)

The first pluripotent embryonic stem cell lines were extracted from mouse blastocyst ICM in 1981 (Evans and Kaufman, 1981; Martin, 1981) and were termed as mouse embryonic stem cells (mESCs). Almost two decades later, the first human embryonic stem cell lines (hESCs) were established by Thomson *et al*. in 1998 (Thomson et al., 1998). Cells isolated from the embryo ICM represent the developmentally transient state *in vivo,* which can be maintained *in vitro* in defined culture conditions. These cells can self-renew, have differentiation capacity comparable to ICM *in vivo*, can contribute to all somatic cell types, and form teratomas containing differentiated derivatives of all three germ layers (Thomson et al., 1998).

Pluripotent self-renewing cells have unique properties and are characterized by various factors. Pluripotent ES cells have, for instance, a unique cell cycle when compared to somatic cell types. The ES cell cycle is shorter, with a majority of the cells existing in the S-phase while the G_1-phase is significantly shorter (Becker et al., 2006; Fluckiger et al., 2006). Moreover, high telomerase activity is an important property of pluripotent cells. A telomerase enzyme maintains telomere repeats in the chromosome ends, and give ES cells the ability to replicative life span and cellular immortality (Wright et al., 1996; Yang et al., 2008). In addition, in female cells, inactivation of another X-chromosome (XCI) occurs during early phases of development. This complex process of dosage compensation in mouse cells takes place upon differentiation, and pluripotent mES ICM cells contains two active X chromosomes (XaXa). In pluripotent hES cells, the XCI status can vary, and three different classes of X chromosome activation can be detected: two active chromosomes (XaXa), fully inactivated X (XaXi), and partially inactivated X (Bruck and Benvenisty, 2011; Dvash et al., 2010; Fan and Tran, 2011). The chromatin status and epigenetic landscape of the undifferentiated ES cells is also unique and comprises a pluripotency specific transcriptional machinery (Hawkins et al., 2010). Certain transcription factors, OCT4, SOX2, and NANOG, are highly expressed in pluripotent cells and are considered as the master regulators of pluripotency (Avilion et al., 2003; Boyer et al., 2005; Mitsui et al., 2003; Nichols et al., 1998). Chromatin status and the core transcriptional unit suppress the expression of key developmental regulators and keep cells self-renewing. Pluripotent ES cells also express a unique set of cell surface markers that can be used for determining the status of the cells. Markers of pluripotent hESCs include: SSEA-3, SSEA-4, and SSEA-5, TRA-1-60 and TRA-

1-81, as well as GCTM2, GCTM343, Thy1, CD9, FZD7, CD24, and TRA-2-49. During cell differentiation the cell surface markers are regulated, and induction of SSEA-1 and A2B5 can be detected. (Draper et al., 2002; Draper et al., 2004; International Stem Cell Initiative et al., 2007; Wright and Andrews, 2009)

2.2 Different pluripotent states of embryonic cells

Current knowledge indicates that multiple pluripotent states can exist in *in vitro* cultured embryonic cells. mESCs and hESCs are both derived from blastocyst stage embryos, but differ from each other in several ways, such as, morphology, marker gene expression, transcription factor binding activity, culture requirements, and X chromosome inactivation status (Figure 1) (Xue et al., 2011). These differences have been considered to be species specific. Nevertheless, another type of pluripotent cells have been derived from mouse embryos, referred to as epiblast stem cells (EpiSC), post-implantation epiblast derived stem cells (Brons et al., 2007; Tesar et al., 2007). EpiSCs express core pluripotency genes and can differentiate into multiple lineages *in vitro,* and are considered to be pluripotent. EpiSCs, however, are not competent to contribute to chimeras, and are therefore considered to exhibit limited differentiation potential, and are developmentally and functionally distinct from mESCs (Nichols and Smith, 2009; Tesar et al., 2007). EpiSC resemble more closely the pluripotent state of hESCs than mESCs, as EpiSCs and hESCs share similar morphology, sensitivity to single cell passaging, growth factor dependence (Brons et al., 2007), epigenetic regulation of transcription, and X-chromosome inactivation (Tesar et al., 2007). It has been proposed that mESCs represent the naïve, ground developmental state, whereas hESCs and EpiSCs are 'primed' pluripotent cells, already primed for lineage specification and commitment (Nichols and Smith, 2009). (Figure 1)

Instead of dividing pluripotent cells into two distinct states, an explanation to the observed differences could be that pluripotent cell cultures exist on a continuous spectrum of pluripotent states (Gallagher et al., 2009). The cells are heterogeneous in nature and comprise at least two diverse cell populations with different developing potentials. Several studies support this idea and distinct subpopulations of cells within the cultured cells (mESC, EpiSC, and hESC) have been identified (Enver et al., 2005; Furusawa et al., 2004; Han et al., 2010; Hayashi et al., 2008; Henderson et al., 2002; Stewart et al., 2006; Toyooka et al., 2008). Moreover, the early development of rodent embryos differ from other mammalian embryos, with rodents having egg cylinder structure and diapause during the early embryogenesis (Nichols and Smith, 2009; Nichols and Smith, 2012), which may explain part of the differences observed between the cell types representing early embryo development.

The cells, however, are not locked to these pluripotent states. EpiSCs can be reverted to a mESC-like, or naïve state by culturing them in leukemia inhibitory factor (LIF) containing media and boosting the reversion by expressing pluripotency factors Klf4, Klf2, Nanog, or c-Myc (Bao et al., 2009; Guo et al., 2009). Equally, mESC can be converted to EpiSCs with the activation of FGF2 and Activin signaling. Just the change of culture conditions is sufficient to execute the molecular changes and convert mESC to EpiSC (Greber et al., 2010). Moreover, mESC like hESCs have been created by over expressing KLF4, KLF2, and OCT4 with LIF culture conditions (Hanna et al., 2010) or by changing the oxygen concentration (Lengner et al., 2010). These mES-like hES cells have the X chromosome status of XaXa, which is considered as the hallmark of the naïve pluripotent state. Importantly, the early development of mammalian embryos occurs in low oxygen concentration, and studies suggest that hESC

cultures maintain a more naïve status in physiological oxygen (5% O_2). A lower oxygen level has been shown, for example, to prevent spontaneous hESC differentiation and reduce chromosomal aberrations, thus helping to keep the cells in pluripotent state (Ezashi et al., 2005; Forsyth et al., 2006). Nonetheless, all the pluripotent embryonic cell types, mESC, hESC, and EpiSC, have specific features and can be considered as potentially distinct cell types. Furthermore, pluripotency is a dynamic state that can be modified with manipulation of signaling pathways and culture conditions (Ng and Surani, 2011). (Figure 1)

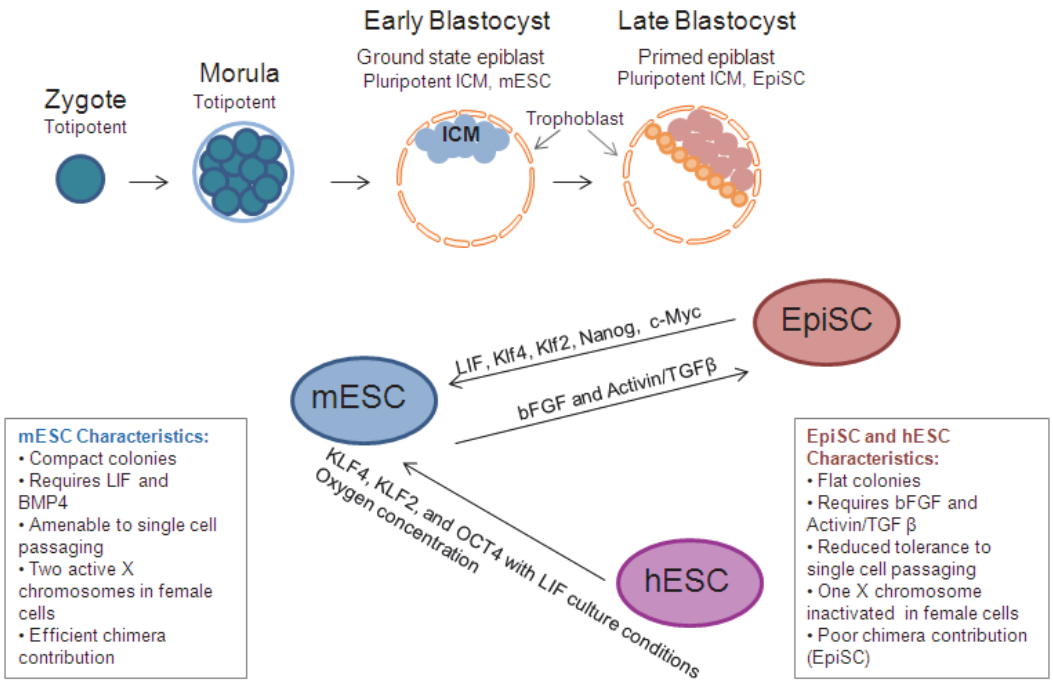

Figure 1. Summary of the pluripotent embryonic stem cells, and conversion between ground and primed pluripotent states. Adapted and modified from Gonzales and Ng 2011 (Gonzales and Ng, 2011).

2.3 Induced pluripotent stem cells

The generation of pluripotent stem cells without the ethical concerns of using embryos has been the interest of many scientists. An alternative means to acquire ES cell-like cells was published in 2006 when Takahashi and Yamanaka reported a breakthrough study of cell reprogramming and generation of induced pluripotent stem cells (iPSC) from mouse somatic fibroblasts by using Oct3/4, Sox2, c-Myc, and Klf4 as inducing factors (Takahashi and Yamanaka, 2006). A year later, iPSCs from human origin were generated with similar method (Takahashi et al., 2007) and also by using OCT4, SOX2, NANOG, and LIN28 as the reprogramming factors (Yu et al., 2007). iPS cells satisfy all the original criteria defined for ES cells (Thomson et al., 1998), thus being equivalent to the stem cells of embryonal origin. iPSCs and ESCs share the same pluripotency markers, are self-renewing, and able to produce differentiated progeny from all three germ layers.

iPS cells provide an enormous potential for medical applications. iPSCs can possibly be used in drug development and screening in the medical industry. Moreover, the technique enables creation of patient and disease specific cell lines, which could further help in studying the mechanisms of certain diseases. Also, the research of early development and possible regenerative and transplantation therapies will benefit from this remarkable method (Figure 2).

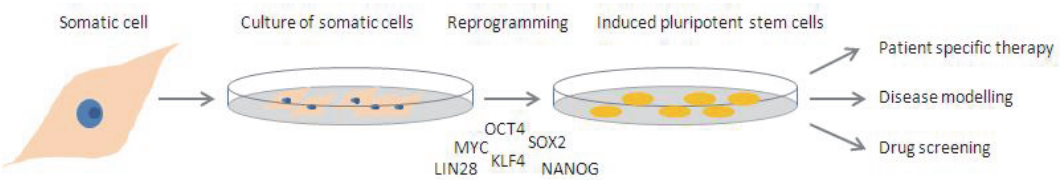

Figure 2. Generation of induced pluripotent stem cell cultures with defined transcription factors, and the promise to medical applications. Adapted and modified from Yamanaka and Blau 2010 (Yamanaka and Blau, 2010).

The first iPS cell lines were reprogrammed using retroviral and lentiviral vectors that integrate into the genomic DNA. In order for the cells to be safe from the risk of tumorigenesis, there is a need to avoid reprogramming methods that involve vector integration to host genome. Therefore, alternative methods for the reprogramming have been tested and developed. Nonintegrating virus and plasmid vectors, microRNAs, isolated proteins, and small molecules have been used to establish iPS cell lines with transient expression of reprogramming factors. However, these non-integrating methods are extremely inefficient compared to integrating viral vectors when looking at the numbers of reprogrammed cells. Sometimes the reprogramming can also be partial, or inefficient, so that the epigenome is not fully reverted to the pluripotent state, but has epigenetic memory of the original cell type. These latter concerns of integrating vectors and inefficient reprogramming are challenges that need to be solved before the therapeutic use of iPS cells. (Miyazaki et al., 2012; Okita and Yamanaka, 2011; Yamanaka, 2012)

The ultimate goal of the reprogramming method is to achieve cells of a particular somatic cell type. The technique to switch cell identity also enables transdifferentiation, a direct conversion of one somatic cell type to other using defined factors (also named as direct reprogramming). Transdifferentiation allows direct differentiation of cells without undergoing a transitional pluripotent state. These transdifferentiated cells could be used for example for cell replacement therapies, thus providing alternative manner to utilize reprogramming and to produce the cells of interest (Gonzales and Ng, 2011).

There are still conflicting conclusions whether iPCs and ESCs are functionally relevant, and which cells would be better for clinical applications. Either way, reprogramming and iPS cells have made a major impact on science and medicine by providing promise for personalized medicine, and for basic research in the broad field of biology. The similarities and differences of pluripotent stem cells are summarized in Table 1.

Table 1. Key characteristic of pluripotent stem cell types. Abbreviations: mESC: mouse embryonic stem cells, EpiSC: epiblast stem cells, miPSC: mouse induced pluripotent stem cells, hESC: human embryonic stem cells, hiPSC: human induced pluripotent stem cells, -: not tested

	mESC	**EpiSC**	**miPSC**	**hESC**	**hiPSC**
Similarities:					
Pluripotency factors	Oct4, Sox2, Nanog	Oct4, Sox2, Nanog	Oct4, Sox2, Nanog	Oct4, Sox2, Nanog	Oct4, Sox2, Nanog
Differentiation potential	All germ layers	All germ layers	All germ layers	All germ layers	All germ layers
Teratoma formation	yes	yes	yes	yes	yes
Differences:					
Cell morphology	compact	flat	compact	flat	flat
X-chromosomes	XaXa	XaXi	XaXa	XaXi / XaXa	XaXi / XaXa
Pluripotency growth factors	LIF/BMP	FGF/ TGFβ	LIF/BMP	FGF/ TGFβ	FGF/ TGFβ
Chimera contribution	yes	no	Yes, but less efficiently than mESCs	-	-

2.4 Signaling of hESC pluripotency

Pluripotent hES cells can be maintained *in vitro* for long periods of time, and different culture systems have been developed for this purpose. Embryonic stem cells represent a state that is present in the embryo for a relatively short time. Maintenance of pluripotent cells *in vitro* therefore requires the integration of signaling inputs to maintain this identity. The first hESC lines were derived using mitotically inactivated fibroblast feeder layers and serum-containing medium (Thomson et al., 1998). Afterwards, diverse culture conditions have been established for the cell maintenance, including fibroblast free culture condition systems (International Stem Cell Initiative Consortium et al., 2010; Ludwig et al., 2006; Ludwig et al., 2006). Extracellular signaling and growth factors have an immense effect on the downstream signaling pathways, and different culture systems can activate a variety of factors and pathways. The determination of a detailed understanding of how signaling pathways maintain pluripotency of hESCs has been hindered by a number of confounding issues. Firstly, the use of inconsistent cell culture conditions has led to context-dependent observations. In addition, variation in experimental design is frequently ignored when interpreting the data, and the focus is often on a specific pathway rather than interpreting the cross-talk between different pathways. Moreover, the effects of signaling pathways can depend on the level of activation, and thus add complexity in the interpretation of results. (Chen et al., 2012; Dalton, 2012; Gonzales and Ng, 2011; Yu et al., 2011)

Pluripotency in mESCs and hESCs is regulated through different pathways. While mESC pluripotency is maintained by leukemia inhibitory factor (LIF) and bone morphogenic protein (BMP) (Ying et al., 2003), the pluripotency of hESCs is maintained by transforming growth factor β (TGF β) and fibroblast growth factor (FGF) signaling (Ng and Surani, 2011). For the culture of mESCs it has been shown that extrinsic signaling pathways are dispensable, and pluripotency can be maintained with the use of chemical inhibitors targeting specific signaling pathways. For this purpose, a medium, called 2i, has been created. The principle of this 2i medium is the combination of GSK3β and MEK inhibitors, influencing WNT and FGF pathways, respectively (Li et al., 2008; Nichols et al., 2009; Ying et al., 2008). The 2i

medium has also been tested for hESCs and EpiSC, but the cells do not survive in these conditions (Guo et al., 2009; Nichols and Smith, 2009; Roode et al., 2012; Van der Jeught et al., 2012). Differences in the extrinsic requirements between mESC and hESC will most probably require alternative inhibitory signal for hESC culture when considering inhibitor based pluripotency maintenance. Tsustui *et al*. developed a culture system for hESC using three chemical inhibitors (ROCK; GSK3β, and MEK). However, bFGF supplementation was still required for the cells to survive (Tsutsui et al., 2011).

TGFβ, FGF, epidermal growth factor (EGF), insulin-like growth factors (IGF), and WNT family signaling pathways have been demonstrated to be involved in the pluripotency maintenance of hESCs. The signaling in pluripotent cells is complex, and the pathways do not function independently, but form a cross-talk network. The details of this network, however, are not understood.

2.4.1 TGFβ signaling

TGFβ signaling has important role in the cell fate decisions in many developmental models. The signal transmission of TGFβ is propagated through receptor activated SMAD proteins through two main branches: BMP activated SMAD1/5/8 and TGFβ/Activin/Nodal activated SMAD2/3 (Massague, 2000). Receptors for both branches are expressed in hESCs (Bendall et al., 2007; Sato et al., 2003; Sperger et al., 2003). Upon activation and phosphorylation, SMADs bind to common SMAD4 and translocate into the nucleus as active transcriptional complexes to regulate gene expression together with other transcription factors (James et al., 2005).

Signaling through the TGFβ pathway has been shown to be crucial for regulating hESC identity. SMAD2/3 is in an activated phosphorylated state in undifferentiated cells, and signaling through this branch has been shown in several studies to be crucial for maintenance of pluripotency and undifferentiated status. Pharmacological inhibition of TGFβ/Activin receptor signaling reduces SMAD2/3 phosphorylation and induces differentiation.(Avery et al., 2010; Beattie et al., 2005; Bendall et al., 2007; Besser, 2004; James et al., 2005; Vallier et al., 2005). Activation of the SMAD1/5/8 branch of TGFβ signaling in turn correlates with cell differentiation (Besser, 2004; James et al., 2005). BMP induces rapid differentiation of hESCs (Pera et al., 2004; Xu et al., 2002) and BMP suppression is used in some culture conditions. Suppression of BMP alone, however, is not sufficient to sustain pluripotency in hESCs (Xu et al., 2005).

TGFβ signaling is directly linked to hESC pluripotency machinery, since *NANOG* proximal promoter is shown to be a direct target of SMAD signaling in hESCs. TGFβ/Activin together with FGF promotes pluripotency by sustaining NANOG expression, whereas BMP binding is decreasing *NANOG* promoter activity (Xu et al., 2008). However, a recent study demonstrated that SMAD2/3 has the ability to regulate the balance between alternative cell states. The threshold of SMAD2/3 signaling was determined to be important in the activation of differentiation genes, linking the SMAD2/3 pathway to the regulation of differentiation (Singh et al., 2012). The study identified that the threshold levels of Activin A/SMAD2/3 in the activation of different sets of target genes is regulated by PI3K/Akt signaling. Moreover, TGFβ/Activin/Nodal and WNT pathways are shown to act in collaboration in the maintenance of pluripotency, and WNT signaling has an effect on SMAD2/3 activation (Besser, 2004; James et al., 2005). In addition, β-catenin, a member of WNT, can act in conjunction with SMAD2/3 to activate mesendoderm genes (Singh et al., 2012).

Recently, Ozair *et al*. (Ozair et al., 2012) reported that SMAD7, an inhibitor of TGFβ signaling, can induce hESC differentiation in the neural direction, hence contributing to the pluripotency regulation network. Moreover, SMAD4, a common component shared by both branches, TGFβ/Activin/Nodal and BMP, has also been shown to have a complex role in the regulation of stem cell fate. SMAD4 is not required for the maintenance of pluripotency, since the knockdown of SMAD4 do not induce rapid cell differentiation or any changes in the morphology or cell growth. However, Avery *et al*. demonstrated that SMAD4 knockdown reduced the stability of hESCs (Avery et al., 2010). In this study they suggested that SMAD2/3 signaling is not promoting pluripotency, but is involved in the stabilization of hESC cultures. Similar conclusions were also made in other studies, suggesting that the balance between SMAD2/3 and SMAD1/5/8 could modulate the selection between undifferentiated state and committed cells (Jiang and Ng, 2008; Xu et al., 2008).

The summary of signaling networks maintaining the pluripotent state of hESCs is presented in Figure 3.

2.4.2 FGF signaling

FGF signaling is mediated through receptor tyrosine kinases (RTK) which through a chain of phosphorylation events activate the mitogen activated protein kinase family (MAPK) (RAS/RAF/MEK1/2/ERK1/2) or PI3K/Akt pathway. Several studies have shown that FGF supports self-renewal and is needed for the maintenance of pluripotency in hESCs (Dvorak and Hampl, 2005; Eiselleova et al., 2009; Greber et al., 2007; Sato et al., 2003; Sudheer et al., 2012). FGF2 is expressed in hESCs, and autocrine FGF signaling is crucial for the maintenance of proliferative hESC in the undifferentiated state (Dvorak et al., 2005). FGF supplementation into the hES culture medium is required to maintain the stem cell properties, and FGF signaling has been directly linked to pluripotency regulation through maintaining the expression of NANOG (Yu et al., 2011).

A recent article, however, reported that FGF2 in hESC culture medium can be replaced by Heregulin and IGF1 (Singh et al., 2012). Moreover, different studies have reported that loss of FGF is linked to cell differentiation, but FGF has only indirect role in hES cell maintenance and is rather involved in inducing the production of supportive factors including TGFβ and IGF II (Bendall et al., 2007; Greber et al., 2007). FGF signaling has been shown to be necessary but not sufficient to maintain pluripotency of hESCs (Vallier et al., 2005), and the effects of FGF signaling in hESCs have been shown to be dependent on the TGFβ pathway. (Greber et al., 2007; Greber et al., 2010; Xu et al., 2008; Yu et al., 2011).

FGF signaling has been reported to activate MEK/ERK and PI3K/Akt downstream signaling pathways, and both MEK/ERK and PI3K/Akt have been shown to regulate hESC pluripotency. Li *et al*. 2007 showed in hESCs that MEK/ERK kinases are targets of the FGF pathway and high MEK/ERK activity is required for undifferentiated hESC status. Moreover, MEK/ERK signaling cooperates with PI3K/Akt in the maintenance of pluripotency (Li et al., 2007). Consistent results were also obtained by Armstrong *et al.* when they showed PI3K/Akt, MAPK/ERK, and NFκβ signaling to be required for pluripotency maintenance, and inhibition of the pathways to induce hESC differentiation (Armstrong et al., 2006). Akt signaling was also shown to maintain pluripotency both in mouse and primate ES cells (Watanabe et al., 2006). Singh *et al*. concluded that PI3K/Akt signaling has important role in

the regulation of hESC pluripotency. Their results indicated that PI3K/Akt regulates levels of ActivinA/SMAD2/3 signaling, which impacts cell fate decisions. (Singh et al., 2012) Contrary results were, however, published by Na *et al*. who used chemically defined culture systems and concluded ERK1/2 to permit differentiation of hESCs. The use of ERK1/2 inhibitor reduced hESC differentiation and did not affect the pluripotency of the cells under these defined culture conditions. However, using other culture conditions activation of ERK1/2 was required to suppress BMP signaling and prevent differentiation (Na et al., 2010). These diverse results again highlight the important role of the culture environment, since the inconsistencies in the outcomes obtained could be explained by different cell culture conditions. Alternatively, the explanation may be that low levels of ERK maintain self-renewal while elevated ERK signaling promotes differentiation (Dalton, 2012).

2.4.3 EGF and IGF signaling pathways

EGF and IGF families of RTKs are important for the development of an organism. The four human EGF receptor (EGFR) homologs are called HER (HER1-HER4) or ERBB (ERBB1-ERBB4). EGFR ligands include, for example, EGF, TGFα, and several isoforms of heregulin (Leahy, 2004). The actions of IGF are mediated through IGF cell surface receptors (IGFR) and with IGF binding proteins (IGFBPs) (Paye and Forsten-Williams, 2006). Several EGF and IGF RTKs have been shown to be highly expressed in hESCs, including IGF1R, ERBB2, and ERBB3 (Sperger et al., 2003), and to be required for hESC self-renewal (Bendall et al., 2007; Chen et al., 2012; Singh et al., 2012; Wang et al., 2007). Insulin/IGF is known to activate the PI3K/Akt pathway that is crucial for the maintenance of pluripotency.

Bendall *et al*. performed a proteomic screen of a fibroblast conditioned medium and identified the most prominent family of growth factors to be the IGF family, especially IGF II. They demonstrated that IGF1R expression correlated with pluripotency markers and underscored the importance the IGF II/IGF1R axis in hESCs (Bendall et al., 2007). IGF1R and ERBB2 were also identified as essential RTKs for hESC self-renewal by Wang *et al*. (Wang et al., 2007). Another growth factor, TGFα, was demonstrated to be secreted from fibroblast feeders and to be important for the undifferentiated hESCs and maintenance of pluripotency. The p44/p42 MAPK downstream pathway was reported to be activated in response to TGFα activation (Chen et al., 2012). IGF has direct role in ES cell maintenance, and IGF signaling via PI3K/Akt activation has been shown to be critical for hESC self-renewal (Bendall et al., 2007). In addition to IGF, EGF family member Heregulin is considered as potent activator of PI3K/Akt signaling (Chen et al., 2012; Singh et al., 2012). Singh *et al*. showed that Heregulin and IGF1 synergistically activate PI3K/Akt signaling and suppress Raf/MEK/ERK activity, and this way promotes hESC self-renewal (Chen et al., 2012; Singh et al., 2012).

2.4.4 WNT signaling

The canonical WNT pathway signaling involves WNT binding to Frizzled receptor, which activates Dishevelled protein. Activated Dishevelled inhibits GSK3β enzyme that is responsible for the degradation of β-catenin. When GSK3β is inhibited, the accumulation of β-catenin enables the translocation into the nucleus and function as a transcription factor. (Cadigan and Nusse, 1997)

Previously reports have concluded that active WNT signaling participates in the maintenance of hESC pluripotency by inhibition of GSK3β (Sato et al., 2004), or by supporting cell proliferation (Cai et al., 2007; Dravid et al., 2005). However, several recent studies show opposite results, and the role of active WNT signaling in the regulation of pluripotency remains controversial. Recent results suggest that WNT signaling promotes cell differentiation and is not involved in the maintenance of the undifferentiated state of hESCs (Davidson et al., 2012; Dravid et al., 2005). WNT signaling has been suggested to act as an antagonist to self-renewal and to promote cell differentiation (Singh et al., 2012).

GSK3β is a serine/threonine protein kinase, that functions in several pathways, including PI3K/Akt, MEK/ERK, and WNT/β-catenin (Singh et al., 2012; Singh et al., 2012; Voskas et al., 2010), the pathways that all are important in the regulation of self-renewal of hESCs. Conflicting reports have suggested that GSK3β has a role in the maintenance of pluripotency, while others report it to be activated in differentiation pathways. It has been observed that different thresholds of GSK3β activity have different biological effects on pluripotent cells, and suggested that different GSK3β-protein complexes, or pools, modify the balance in pluripotent cells (Dalton, 2012). However, the exact role of GSK3β in the regulation of self-renewal remains controversial.

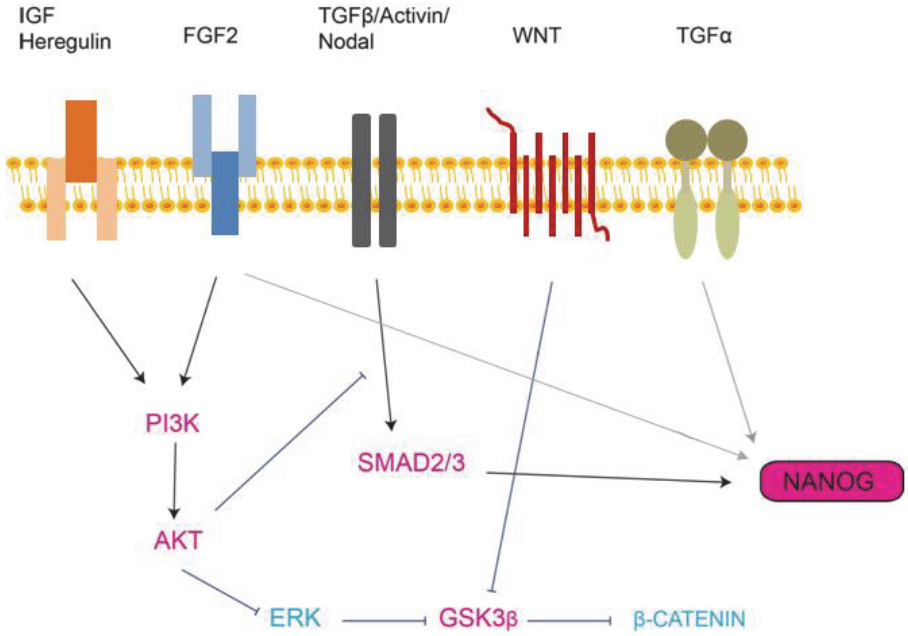

Figure 3. An illustration of the signaling networks maintaining hESC pluripotency. An activating signal is represented with an arrow, whereas the inhibiting signal is represented with a line. Blue color indicates a repressed state in pluripotency, while red indicates the activated state of the signaling molecule. Grey arrows represent unknown mechanism. (Adapted and modified from (Chen et al., 2012; Chen et al., 2012; Dalton, 2012; Singh et al., 2012; Yu et al., 2011).

2.5 Transcriptional control of hESC pluripotency

Every cell derived from a fertilized egg has the same DNA and genetic information. The regulation of gene expression and capacity to differentiate during the development involves

epigenetic regulation, namely DNA methylation and changes in chromatin structure. The chromatin in pluripotent hESCs is in a less compact state, enabling dynamic and flexible reorganization during differentiation (Meshorer et al., 2006). DNA methylation on CpG islands of gene promoters is linked to gene silencing and the methylation in pluripotent hESCs has different patterns than in differentiated cells (Bibikova et al., 2006). The promoters of key housekeeping genes such as NANOG and OCT4 are demethylated, permitting active expression in undifferentiated hESCs (Yeo et al., 2007), whereas the global methylation status in hESCs is high (Laurent et al., 2010). Chromatin structure is also regulated by post-translational modifications, e.g. acetylation or methylation of histone proteins. In pluripotent cells, many development-related genes display bivalent histone marks that correspond to 'poised' status of promoter activation. The bivalent histone mark includes both activating (H3K4me3) and repressing (H3K27me3) histone modifications at the gene promoter, that enables rapid activation of the gene during the induction of differentiation (Bernstein et al., 2006; Pan et al., 2007; Zhao et al., 2007).

2.5.1 Key transcription factors of hESC pluripotency

The undifferentiated hESCs state is controlled by unique transcriptional circuitry. The transcription factors assist in establishing and maintaining the cellular identity by promoting the expression of cell type specific genes and suppressing the expression of irrelevant genes (Gonzales and Ng, 2011). Transcriptional profiling of pluripotent cells has identified genetic networks and several factors to be potential regulators of hESC pluripotency and self-renewal. (Armstrong et al., 2006; Assou et al., 2007; Chia et al., 2010; Enver et al., 2005; Golan-Mashiach et al., 2005; Hirst et al., 2006; International Stem Cell Initiative et al., 2007; Sato et al., 2003; Skottman et al., 2005; Sperger et al., 2003). The core transcription factors maintaining the undifferentiated state in both mESCs and hESCs have been identified to be OCT4, SOX2, and NANOG. All the three factors are highly expressed in undifferentiated ESCs and play a role in the maintenance of pluripotency (Avilion et al., 2003; Fong et al., 2008; Hay et al., 2004; Hyslop et al., 2005; Matin et al., 2004; Mitsui et al., 2003; Nichols et al., 1998; Zaehres et al., 2005; Zafarana et al., 2009). Based on chromatin immunoprecipitation studies, OCT4, SOX2, and NANOG form the core transcriptional network of hESCs by co-occupying over three hundred gene promoters and maintaining the hESC specific gene expression, forming also autoregulatory loops binding to their own promoters. The co-occupied promoters include both active and inactive genes, suggesting that OCT4, SOX2, and NANOG regulate genes in both an activating and repressing manner (Boyer et al., 2005). The basic significance of these factors to hESCs has been demonstrated, yet their mechanistic functions are still not known in detail (Wang et al., 2012). The core regulatory circuitry of hESCs is represented in Figure 4.

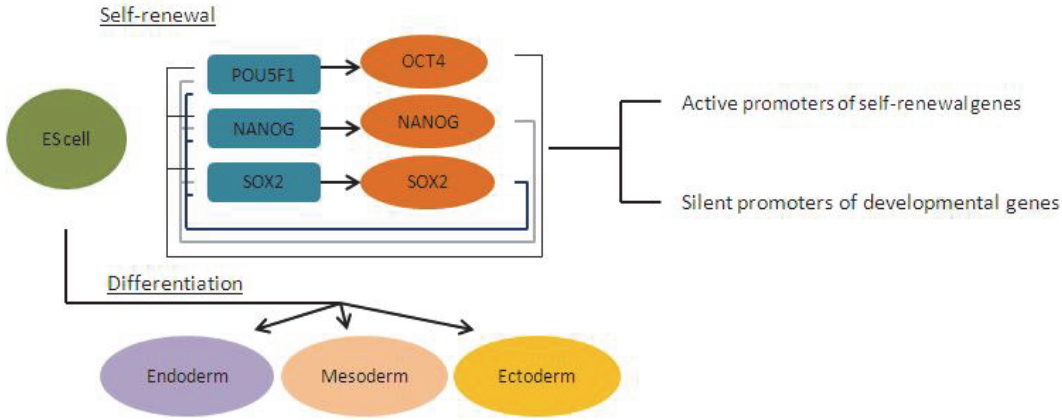

Figure 4. The key transcription factors of hESC pluripotency. OCT4, SOX2, and NANOG form the core transcriptional network and autoregulatory loop regulating stem cell self-renewal. OCT4, SOX2, and NANOG promote the expression of self-renewal genes and suppress the expression of developmental genes. (Adapted and modified from Christophersen et al.2010 (Christophersen and Helin, 2010)).

A recent study from Wang *et al*. highlighted the role of repression of differentiation in the control of hESC status. With the use of loss and gain of function studies on all of the core genes, they reported cell line specific conditions and individual functions for OCT4, SOX2, and NANOG. Thus, instead of being common repressors of differentiation, OCT4, SOX2 and NANOG each control specific cell fates and that the regulatory loop between these genes is uncoupled. Overexpression of any of the factors does not induce differentiation. They identified OCT4 in the regulation of BMP pathway, SOX2 being permissive for self-renewal and being redundant with SOX3, and NANOG being specific repressor of ectoderm differentiation (Wang et al., 2012). Also Yu *et al*. concluded that NANOG has a role in the establishment and maintenance of pluripotency, but in addition can influence the specific lineage commitment as cells exit the pluripotent state (Yu et al., 2011).

In addition to OCT4, SOX2, and NANOG, it seem likely that there must be several other factors that are required for the pluripotency maintenance in hESCs, and recent studies have identified a couple of novel regulators. For example, PRDM14 was shown to be a transcription factor regulating the expression of OCT4 and being a component of the hESC core transcriptional regulatory network (Chia et al., 2010; Tsuneyoshi et al., 2008). In addition, FOXO1 transcription factor was shown to activate OCT4 expression in hESCs (Zhang et al., 2011), and LSD1, a histone demethylase, was shown to regulate bivalent domains of developmentally important genes in hESCs (Adamo et al., 2011).

Although transcriptional profiles of hESCs have been produced in several experiments, only few of the factors identified in these analyses have been functionally analyzed in the context of pluripotency maintenance. Furthermore, hESC and mESC have been reported to show outstanding differences in the OCT4, SOX2, and NANOG target genes (Boyer et al., 2005; Kunarso et al., 2010; Wang et al., 2012), and the transcriptional network in hESCs is not as well studied than that of mESCs, hence there are many unknown factors playing in the hESC regulatory circuitry.

2.5.2 Transcriptional machinery

Chromatin remodelers modify the accessibility of DNA to transcription factors and transcriptional machinery. The pluripotent cell state is controlled by specific transcriptional circuitry in an 'open' chromatin organization that is rapidly modified upon induction of differentiation (Gaspar-Maia et al., 2011). Human genes are transcribed by three DNA directed RNA polymerases, pol I, II, and III. Polymerases are multisubunit complexes that mostly occupy the genome on open and active chromatin. Each of the three human polymerases transcribes specific sets of target transcripts: pol I specialized to transcribe non-coding RNAs (ncRNA) mostly ribosomal RNAs (rRNA), pol II transcribing messenger RNAs (mRNA) and ncRNAs, and pol III synthesizing primarily ncRNAs as well as protein coding genes (Oler et al., 2010; Roeder and Rutter, 1969; Roeder and Rutter, 1970; White, 2008; White, 2011).

Pol II has attracted the most attention of the polymerases, because of its role in transcribing protein encoding mRNAs. However, pol II is also reported to synthesize ncRNAs, including micro RNAs (miRNA) and small nuclear RNAs (snRNA) (Lee et al., 2004; White, 2011). A recent report from Maston *et al*. suggested that the pol II transcriptional machinery of hESCs is unique and essential for pluripotency maintenance. The active pol II gene promoters in pluripotent hESC have a distinct composition of transcriptional components than the differentiated derivatives (Maston et al., 2012).

Pol III transcripts are needed for vital cellular processes, including the regulation of transcription, RNA processing, and translation. Pol III is considered to transcribe a wide range of ncRNAs, including for example tRNA, rRNA, and miRNA (Borchert et al., 2006; Canella et al., 2010; Oler et al., 2010; Teichmann et al., 2010; White, 2011). However, the whole repertoire of pol III genes in the human genome is not known. Interestingly, a tissue and cell type specific expression of pol III genes has been reported (Martignetti and Brosius, 1993; Martignetti and Brosius, 1995; Oler et al., 2010), and decrease in pol III function has been linked to differentiation of mouse embryonal carcinoma cells (White et al., 1989) linking pol III mediated transcription to cell differentiation.

The functions of the polymerases, however, are not completely distinct from each other. Raha *et al*. showed that polymerases II and III can work together to globally coordinate gene expression in human cells (Raha et al., 2010) and that pol II is present at the majority of the loci bound by pol III (Barski et al., 2010; Moqtaderi et al., 2010; Oler et al., 2010). The functions of the polymerases thus can influence one another, and specific components of transcriptional machinery can specify cell type specific gene expression. Taken together, the chromatin structure and gene expression machinery in pluripotent hESC is unique and complex. Various factors are having roles in the regulation of pluripotency associated gene expression.

2.6 Translational control of hESC pluripotency

Extracellular signals and gene transcription have enormous roles in the regulation of hESC specific intracellular programs. However, an additional level of regulation is provided by several factors responsible for post-transcriptional regulation of signaling. mRNA can be regulated in numerous ways in the cytoplasm, for example, by RNA binding proteins (RBP) or miRNAs. Translationally inactive mRNAs can be accumulated in the cytoplasm to dynamic RNA-protein granules, called P-bodies or stress granules, and form

ribonucleoprotein complex (RNP). mRNAs can be transferred between translational polysomes and RNPs, providing additional level of control of gene expression (Balagopal and Parker, 2009). miRNAs are small, ~22 nucleotides (nt) in length, non-coding RNAs that bind to the 3' untranslated region of their target genes and act as translational repressors. miRNAs are transcribed as long primary transcripts (pri-miRNAs) which are processed in the nucleus into precursors (pre-miRNA). Pre-miRNAs are further modified in the cytoplasm to form a mature and functional single stranded miRNAs. (Ambros, 2001)

LIN28 is a conserved RNA binding protein originally identified as a regulator of developmental timing in *Caenorhabditis elegans* (Moss et al., 1997). LIN28 has also been associated with the regulation of Let-7 miRNAs (Heo et al., 2008; Newman et al., 2008; Viswanathan et al., 2008). However, numerous recent studies imply additional, miRNA independent, functions for LIN28 (Cho et al., 2012; Peng et al., 2011; Wilbert et al., 2012). LIN28 is highly expressed in pluripotent hESCs and down regulated in response to differentiation (Darr and Benvenisty, 2009). The major function of LIN28 is proposed to support the growth of hESCs through mRNA regulation in translational machinery in cytoplasmic polysomes (Cho et al., 2012; Peng et al., 2011; Qiu et al., 2010; Wilbert et al., 2012). LIN28 has been shown to enhance the translation of important hESC genes, including OCT4 (Peng et al., 2011; Qiu et al., 2010; Xu et al., 2009). Moreover, complete depletion of LIN28 in hESCs leads to reduced self-renewal, indicating LIN28 to be vital for the maintenance of pluripotency (Peng et al., 2011; Qiu et al., 2010). LIN28 is also one of the factors used in iPS cell reprogramming, further highlighting the role in pluripotency maintenance and regulation (Yu et al., 2007).

A number of studies have identified miRNAs that are highly and uniquely expressed in self-renewing hESCs. Suh *et al*. cloned 36 miRNAs from hESCs and identified most of these newly cloned miRNAs to be expressed in hESC specific manner and to be down regulated in response to cell differentiation. Among the identified miRNAs miR-302-367 cluster was one of the most significant miRNA families (Suh et al., 2004). This miRNA family was identified also from studies by Bar, Morin, and Lakshmipathy *et al*. 2008. Other hESC enriched miRNAs include miR-20b, miR-21, miR-92b, miR-200c, miR-222, and miR-371-373 (Bar et al., 2008; Lakshmipathy et al., 2007; Morin et al., 2008; Suh et al., 2004). Barroso-delJesus *et al.* characterized the promoter and expression profile of the miR-302-367 cluster in hESCs, and suggested core transcription factors, OCT4, SOX2, NANOG, to regulate the expression of the cluster (Bar et al., 2008; Barroso-delJesus et al., 2008). This was later validated by Card *et al*. (Card et al., 2008). Based on computational predictions, miR-302-367 miRNAs have 844 potential target transcripts (www.targetscan.org), many of which are cell cycle regulators (Card et al., 2008; Lipchina et al., 2011). Card *et al.* validated the miR-302-367 cluster to participate in the regulation of the unique stem cell cell cycle, and cyclin D1 to be directly post-transcriptionally regulated by miR-302. Further, Lipchina *et al*. identified 146 high confidence target genes of miR-302-367 cluster and showed it to be positive regulator of self-renewal and promoting G_1/S transition (Lipchina et al., 2011).

Xu *et al*. studied the miRNAs that are involved in the repression of pluripotency genes and are increased during differentiation of hESCs. The study identified miR-145 as a translational repressor directly targeting pluripotency factors OCT4, SOX2, and KLF4. Moreover, induction of miR-145 was able to inhibit self-renewal and promote differentiation. Interestingly, OCT4 was found to bind to miR-145 promoter and act as a transcriptional repressor. Hence, miR-145 and OCT4 form a double negative feedback loop. (Xu et al., 2009)

There seems to be only minimal overlap of the miRNA profiles between pluripotent mESC and hESCs (Bar et al., 2008; Suh et al., 2004). Moreover, naïve mESC and primed EpiSC are reported to express distinct miRNA clusters, and the different pluripotent states can be discriminated by the miRNA profiles (Jouneau et al., 2012). Overall, the study of the roles of miRNAs in embryonic stem cells is still in the very beginning. Very little is known about the miRNAs and their target genes and how they regulate the stem cell properties. The interplay between pluripotency factors and miRNAs reveal a new level of complexity in the regulatory network governing the stem cell properties. The summary of the known regulators of pluripotency is presented in Figure 5.

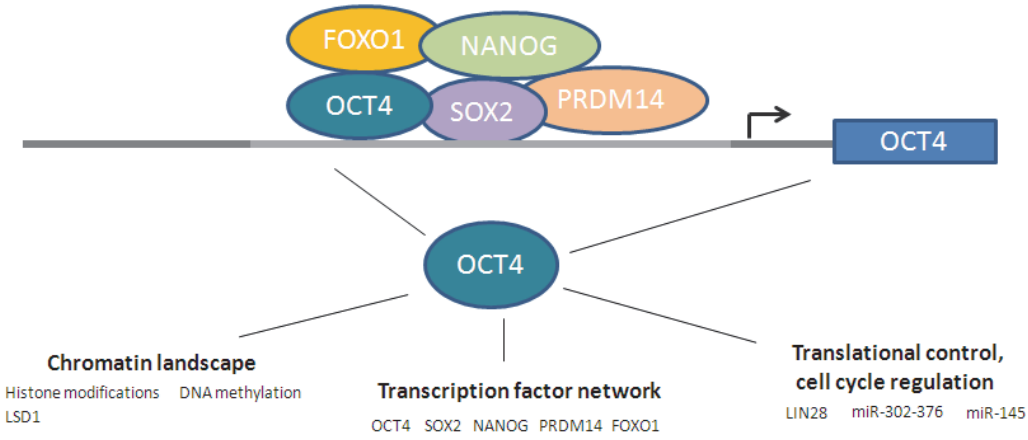

Figure 5. Summary of the factors regulating pluripotency in hESCs. Crosstalk between transcriptional regulatory network, epigenetic network, and translational control of gene expression. OCT4 is one of the best characterized transcription factors in hESCs. Studies have shown a central role for OCT4 in all of the networks and in diverse cellular functions. (Adapted and modified from Ng and Surani 2011 (Ng and Surani, 2011)).

2.7 Pluripotent cells and genomic stability

As previously mentioned, the culture of pluripotent hESC lines involves manipulation of signaling pathways to maintain the dynamic cells in the pluripotent state. Regarding this, it is logical that cultures may contain cells that have adapted to the culture conditions in which they are grown. Studies have revealed that pluripotent cells may gain genomic aberrations in long term *in vitro* maintenance that gives cells distinct characteristics and growth advantages promoting self-renewal and limiting differentiation and apoptosis. Suboptimal culture conditions have been suggested to be involved in creating chromosomal instability in hESCs. (Baker et al., 2007; Draper et al., 2004; Enver et al., 2005; Maitra et al., 2005; Taapken et al., 2011).

Karyotype is the appearance of the chromosomes during cell division. The changes in the karyotype that can take place include changes in chromosome numbers or in chromosome structure. The aberrations can be chromosomal deletions, gains/insertions, or translocations. Cells that have gained genomic alterations during *in vitro* maintenance are called as culture-adapted cells and have altered state of self-renewal and differentiation (Enver et al., 2005; Harrison et al., 2007; Maitra et al., 2005). These culture induced genomic alterations are a

concern regarding the potential and safety of the cells for medical applications, since the chromosomal changes and instability of genome are hallmarks of cancer (Albertson et al., 2003) and may change the differentiation potential of the cells.

Some genetic abnormalities may exist already in the embryo before the ICM isolation, however, most of the hESC lines have normal karyotypes and changes in the genome occur later during *in vitro* maintenance (Lund et al., 2012). Studies have identified certain genomic areas that are more sensitive to alterations in hESCs. Gains of chromosomes 12 and 17 have been reported in several studies to be the most frequent changes of hES that accumulate during culture (Baker et al., 2007; Cowan et al., 2004; Draper et al., 2004; Taapken et al., 2011). Other chromosomal amplifications have also been reported, for instance gains of chromosomes 8, 17, and X (Baker et al., 2007; International Stem Cell Initiative et al., 2011; Taapken et al., 2011). Moreover, a number of studies have shown that the genomic changes can occur at any stage of culture. Even though it is twice as likely to detect chromosomal abnormalities at higher passage numbers than in the early passages, the changes can arise at any stage (International Stem Cell Initiative et al., 2011; Laurent et al., 2011; Taapken et al., 2011). In addition, there is no association with any particular culturing system and the detected genomic changes (Taapken et al., 2011). The detected changes in the chromosomes has also been reported to alter the expression levels of the genes located in these genomic areas (Enver et al., 2005; Mayshar et al., 2010; Narva et al., 2010), thus affecting the phenotype of the cells. In addition, the abnormal cell cultures are often mosaic cell populations, in other words, the population is consisting of genotypically distinct cell types (International Stem Cell Initiative et al., 2011).

2.7.1 Methods for detecting karyotypic changes

Methods that are currently used for detecting chromosomal aberrations in pluripotent stem cells include cytogenetic analyses as Giemsa staining (G-banding) and fluorescence in situ hybridization (FISH), in addition to microarray and sequencing based methods. G-banding is a method of staining of metaphase chromosomes and is analyzed under the microscope. G-banding creates pattern of bands specific for each chromosome (Drets and Shaw, 1971; Patil et al., 1971). The method is routinely used to analyze the karyotypes of hES cells. The benefit of the methods is the sensitivity for detecting mosaic cell populations and balanced rearrangements. However, the low resolution, small number of cells used in the analysis, and expertise required for the analysis comprise the disadvantages for the method. FISH is a technique based on the localization of specific DNA sequences using fluorescence labeled DNA probes (Langer-Safer et al., 1982). Like G-banding, FISH is able to detect mosaic cell populations. Microarray based methods provide a higher-resolution technique to analyze the genomic stability of the cells. Different array formats enable comparative genomic hybridization (CGH) and detection of copy number variants (CNV) or single nucleotide polymorphism (SNP). The material that is used in the array based methods is genomic DNA, thus representing whole cell population, and incapable of detecting low level mosaicism or balanced rearrangements. In addition to genome wide array methods, new high-throughput next generation sequencing-based methods offer a novel technique to study the genomic alteration of pluripotent cells. Nevertheless, both of these genome wide methods are relatively costly and need expertise for the data analysis. (Maitra et al., 2005; Narva et al., 2010; Wu et al., 2008)

3 AIMS OF THE STUDY

The goal of this study was to identify novel factors that are involved in the maintenance of pluripotent state in hES cells. Two candidate genes were chosen for further studies to identify their roles in the pluripotency maintenance. Furthermore, a new method for analyzing karyotypic abnormalities in pluripotent cells was implemented.

The detailed aims of this thesis were to:

1. Identify novel factors enriched in pluripotent cells using genome wide approaches throughout different cell lines grown in different laboratories with different genetic background and growth conditions.
2. Identify the importance of POLR3G in the regulation of self-renewal and development.
3. Elucidate the function of L1TD1 in the pluripotency maintenance of hESCs.
4. Test and set up a method for routine high-throughput karyotyping to analyze the genetic stability of pluripotent human stem cells.

4 MATERIALS AND METHODS

4.1 Cell culture

4.1.1 Culturing of hES cells (I, II, III)

Human embryonic stem cell line H9 was obtained from WiCell Research Institute (Madison, WI) and cell lines HS293, HS306, HS346, HS360, HS362, HS363, HS368, HS401 were generously provided by Outi Hovatta (Karolinska Institutet, Sweden) and Heli Skottman (Regea, Tampere, Finland). All hES lines were grown as stock culture and expanded on mitotically inactivated human foreskin fibroblasts (ATCC) plated on 0.1% gelatin coated (Sigma-Aldrich) cell culture plates. Dulbecco's modified Eagle's medium, DMEM-F12 (Stem Cell Technologies), supplemented with 20% serum replacement, 2 mM glutamax, 1% nonessential amino acids, 50 U/ml penicillin-streptomycin, 0.1 mM 2-mercaptoethanol (Gibco), and 4 ng/ml basic fibroblast growth factor (bFGF) (R&D Systems) was used as ES culture medium. For experiments, cells were plated in feeder-free conditions on Matrigel (BD Biosciences) and grown in commercial mTeSR1 medium (Stem Cell Technologies) or in fibroblast conditioned ES culture medium. Cells were passaged by using type IV collagenase (Gibco), Accutase (Invitrogen), or mechanical cutting and culture medium was changed daily.

4.1.2 Culturing of embryonal carcinoma cells and cancer cell lines (II)

Embryonal carcinoma cell (EC) lines, 2102Ep, NT2D1, and NT2D1-TetR3 were provided by Dr. Peter W. Andrews or Dr. Jianliang Li (University of Sheffied, UK) (Andrews, 1984; Zafarana et al., 2009). Cells were grown in EC medium: DMEM (Sigma) supplemented with 10% fetal calf serum (FCS) (PromoCell) and 2 mM L-glutamine (Sigma). 2102Ep was passaged using 0.05% trypsin-EDTA, whereas NT2D1 and NT2D1-TetR3 were passaged by scraping. TCam2 cell line was obtained from Dr. Jukka Westermarck (Turku Centre for Biotechnology, Finland) and was grown in RPMI 1640+GlutaMAX (Gibco) supplemented with 10% FCS (PromoCell) and 1% Penicillin-streptomycin (Sigma). TCam2 cells were passaged with 0.05% trypsin-EDTA.

4.1.3 Differentiation of human ES and EC cells (I, II)

Spontaneous EB differentiation series were made by scraping the cells from the plate and transferring to non-coated cell culture plates. The cells were grown in normal ES culture medium or in ES medium without bFGF. Medium was changed every 2-3 days. Induced cell differentiation was performed using hESCs plated on Matrigel in feeder free conditions. The medium was supplemented with 13.7μM retinoic acid (RA) (Sigma) and was changed daily (II) or every 2-3 days (I).

4.1.4 Analysis of cell numbers (I)

Cedex XS (Innovatis) system was used to determine the numbers of viable and dead cells in response to siPOLR3G treatment based on Trypan Blue staining and cell morphology.

4.2 Cell transfections and gene silencing

Cell transfections were performed according to manufacturer's protocol by using Lipofectamine 2000 or Lipofectamine RNAiMax transfection reagents (Invitrogen). hES cells were plated on Matrigel in mTeSR1 or conditioned ES medium and transfected 24 hours after plating. A second transfection was performed 24 hours from first treatment. Sample collection was started 48 hours after first transfection as day 1 samples. EC and cancer cell transfections were performed 24 hours after plating and having approximately 50% confluency.

4.2.1 Oligonucleotides for RNAi (I, II)

Five sequences were tested for silencing of L1TD1. Sequence number three was selected from the RNAi Codex shRNA database, whereas the others were selected from siRNA Target Finder software of GeneScript. For transient L1TD1 silencing, the sequences were cloned into pSuper-green fluorescent protein GFP-Neo (Oligoengine) vector that produces short hairpin RNA (shRNA), using BglII and XhoI cloning sites. The shRNA vectors were synthesized by DNA Technology A/S (Risskov, Denmark). Sequences with the best knockdown efficiencies, 1, 2, and 5, were ordered also as short interfering RNAs (siRNA) (Sigma). All the sequences used for RNAi are listed in Table 2. Non-targeted siRNA sequences and other siRNAs were chosen from the literature or were commercially available.

Table 2. Sequences used in RNAi

Target	Sequence
L1TD1 1	GCAAGGACGTATCAGCAATTA
L1TD1 2	AGGACAGAGTTTCAGCAAATA
L1TD1 3	GACAGAGTTTCAGCAAATAATC
L1TD1 4	GCCGACCTTTCATCAGCAACA
L1TD1 5	GAGATGAGTCATGATGAGCATA
POLR3G 1	CCAGUACCACUGAAAACAGdTdT
POLR3G 2	UGACGAUGAUGCCGCAGAA
POLR3G 3	Santa Cruzt Biotechnology, sc-43507, pool of 3 sequences
OCT4	AAGGAUGUGGUCCGAGUGUGG
SOX2	AAAACCAAGACGCUCAUGAAG
NANOG	AAGGGUUAAGCUGUAACAUAC
NON-TARGETED 1	GCGCGCUUUGUAGGAUUCG (Mori et al., 2003)
NON-TARGETED 2	AAUUCUCCGAACGUGUCACGU (Hashimoto et al., 2004)
NON-TARGETED 3	CCUACAUCCCGAUCGAUGAUG (Berra et al., 2003)

4.2.2 Inducible stable gene silencing (II)

Clonal cell lines with doxycycline-dependent inducible L1TD1 shRNA were created to study the effect of long term knockdown of L1TD1 (Andrews, 1984; Zafarana et al., 2009). shRNA sequences 1 and 5 were cloned into pSUPERIOR.neo plasmid (Oligoengine) using BglII and XhoI cloning sites. NT2D1 cells stably expressing TetR3 were transfected with L1TD1 shRNA constructs. Cell clones expressing the transgene were selected by using EC medium containing puromycin and G418 (Sigma) at concentrations of 3 µg/ml and 750 µg/ml, respectively. The shRNAs were induced using doxycycline at concentration of 1 µg/ml. Clonal sublines were expanded and screened for the knockdown efficiencies.

4.3 Over expression plasmid constructs (II)

The open reading frame of L1TD1 without stop codon was cloned into pEF6-V5-His-TOPO vector (Invitrogen) by TA cloning to get a V5-tagged L1TD1 over expression vector. This vector was used for optimizing the best short hairpin RNA (shRNA) sequences for L1TD1 and for localization studies. The following cloning primers were used:
5′-GCCATGTCTGATGTATCTACTAG-3′ 5′-AGGTATATTATTCCCCAGTAAT-3′.

$[His]_6$-tagged L1TD1 protein was produced by cloning the open reading frame of L1TD1, including the stop codon, into pET-20b(+) vector (Novagen) using NcoI and XhoI restriction sites. Cloning primers: 5′-CGCGCGCCATGGATTCTGATGTATCTACTAGTGT-3′ 5′-CGCGCGCTCGAGAGGTATATTATTCCCCAGTAA-3′. $[His]_6$-tagged protein was used for L1TD1 specific antibody immunization.

L1TD1-Enhanced Green Fluorescent Protein (EGFP) over expression vector was produced for localization studies, by cloning L1TD1 sequence into pCAGG-EGFP vector (gift from Dr. Peter Andrews, University of Sheffield, UK) with AgeI and XhoI restriction sites with primers: 5′-CGCGCGCTCGAGATGTCTGATGTATCTACTAG-3′, 5′-CGCGCGACCGGTTGAGGTATATTATTCCCCAGT-3′.

4.4 Gene expression analyses

4.4.1 Analysis of the Stem Cell Matrix Data (II)

The Stem Cell Matrix Data was used to analyze the expression of *L1TD1* (http://www.ncbi.nlm.nih.gov/geo/)(GSE11508). Sample preprocessing was done with lumi-package of R using quantile normalization algorithm and probe values were linked to the Ensembl genes (NCBI 36). If several probes were detected within the region of the same gene the probe values were mean centered.

4.4.2 DNA Microarrays (I,II)

In project I the gene expression profiles of undifferentiated hES cells and differentiated cell types were analyzed using Affymetrix U133AB and U133plus2 microarrays. Published datasets used in the analysis are listed in Table 3, in addition, unpublished datasets produced in the group were used in the analysis. GeneSpring data analysis software (Agilent) was used and RMA algorithm was used for normalization. Percentile cut-off 20 was used to filter out the samples with low a expression level throughout the samples. Unpaired t-test and cut-off $p \leq 0.05$ were used to identify statistically differentially expressed genes from the studies with 3 or more available replicates. A fold change cut-off was set at 1.5 between all the replicated condition pairs within the experiment.

Table 3. Published datasets used in the genome wide microarray analysis to identify stem cell enriched genes.

Author	Reference	Array type
Enver et al.	Hum Mol Genet. 2005 1;14(21):3129	U133A
Barberi et al.	PLoS Med. 2005 2(6):e161	U133A
Hirst et al.	Dev Biol. 2006 1;293(1):90	U133AB
Sato et al.	Dev Biol. 2003 15;260(2):404	U133A
Golan-Mashiach et al.	FASEB J. 2005 19(1):147	U133A
Armstrong et al.	Hum Mol Genet. 2006 1;15(11):1894	U133plus2
Skottman et al.	Stem Cells. 2005 23(9):1343	U133AB
Lu et al.	Genome Biol 2007 8(11):R240	U133plus2

In project II the stable L1TD1 knockdown clones were analyzed by using Illumina BeadChips. L1TD1 shRNAs were induced with 1 µg/ml doxycyline (Sigma) for six days. Three different clones expressing two different shRNA sequences and their non-induced controls were used for the analysis and hybridized on Illumina Human HT-12 v.3 Expression Bead Chip. The data was normalized by using quantile normalization, and R-package Limma was used for statistical analysis. Filtering criteria ≥ 1.3 for fold change and ≤ 0.05 for *p*-value were used to identify statistically significant differences in gene expression. The raw data is available in the Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo) with the accession code GSE21275.

4.4.3 Next-Next-Generation RNA Sequencing with Heliscope (I)

The gene expression profiles of undifferentiated hES cells and differentiated EBs were analyzed with Single Molecule RNA Sequencing (Heliscope instrument, Helicos Biosciences). One microgram of total RNA per sample was processed according to the Helicos™ RNA Sequencing Protocol. Filtering and aligning the sequencing reads was done with Helicos Helisphere software version 1.1.498.63. The pipeline was run using default parameters and the reads were aligned against UCSC gene annotations for the hg18 human reference genome assembly. The gene expression values are presented as reads per kilo base of exon model per million mapped reads. RPS13 and RPL27 were used as reference genes for comparison of normalized expression reads of hES and EB samples.

4.4.4 Next-Generation Sequencing with HiSeq2000 (I)

POLR3G siRNA treated knockdown samples and control samples transfected with non-targeted siRNA were analyzed with next-generation mRNA sequencing (HiSeq2000, Illumina). The sample libraries were prepared from 1µg or 0.3 µg of total RNA with TrueSeq RNA Sample Preparation Kit v3 (Illumina) according to manufacturer's protocol. The cluster generation was performed automatically with the c-Bot instrument (Illumina). The samples were multiplexed into a single lane and sequencing was carried out with 2x100bp chemistry and HiSeq2000 platform (Illumina). The reads were aligned against UCSC gene annotations for the hg19 human reference genome assembly using TopHat (Trapnell et al., 2009). RPKM values were computed and genes having less than 0.3 RPKM in more than three of six samples were excluded to filter the data (Ramskold et al., 2009). edgeR (McCarthy et al., 2012) was used to determine differential expression to account for the paired design of the input data. Adjusted *p*-values were computed with Benjamini-Hochberg method and FDR of 0.05 was used as a threshold for significance. Cut-off of 1.5 for fold change was used to extract the differentially expressed genes between the sample groups.

4.4.5 Ingenuity Pathway Analysis (I,II)

Ingenuity Pathway Analysis tool (Ingenuity systems) was used to determine and characterize the functional pathways and signaling networks affected by the silencing of POLR3G and L1TD1.

4.4.6 Quantitative RT-PCT analyses (I,II)

mRNA expression levels were analyzed using Taqman reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) 7900HT Fast Real Time System (Applied Biosystems). RNA was isolated using RNeasy Kit and DNase I digestion in the column was included (Qiagen). The concentration of the samples was measured with a Nanodrop detector (Thermo Scientific). A second round of DNAse treatment was carried out for 500 ng or 1 µg of total RNA with DNase I Amplification Grade (Invitrogen). A negative RT-qPCR control run was performed by measuring levels of the reference gene EF1α to verify that no genomic DNA was present in the sample. Complementary DNA (cDNA) was synthesized using Superscript II kit (GIBCO). The cycle of threshold values (CT) were compared to reference gene (EF1α or actin beta) to obtain normalized expression levels for the transcripts (ΔCt). The primers and probes were designed using Universal ProbeLibrary Assay Design Center (Roche). The sequences are listed in table 4. The probe sequence for EF1α is 5´-(FAM)-AGCGCCGGCTATGCCCCTG-(TAMRA)-3´.

Table 4. Primers and probes used in the RT-qPCR analyses.

Target	Primer Forward	Primer Reverse	Probe
ACTIN B	ccaaccgcgagaagatga	ccagaggcgtacagggatag	64
DNMT3B	ggaaattagaatcaaggaaatacga	aatttgtcttgaggcgcttg	83
EF1a	ctgaaccatccaggccaaat	gccgtgtggcaatccaat	
L1TD1	tcccacaaaaggaagaaataaatc	gctctatgctttgagtctattaggg	69
LIN28	aagcgcagatcaaaaggaga	ctgatgctctggcagaagtg	23
NANOG	cctgaacctcagctacaaacag	gctattcttcggccagttgt	87
OCT4	agcaaaacccggaggagt	ccacatcggcctgtgtatatc	35
OCT4 [2]	agcaaaacccggaggagt	ccacatcggcctgtgtatatc	69
OCT4 [3]	cttcgcaagccctcatttc	gagaaggcgaaatccgaag	60
OCT4 [4]	cctgtctccgtcaccactct	ggcacaaactccaggttttc	52
OCT4 [5]	ctttgaggctctgcagcttag	ctgctttgcatatctcctgaag	69
OCT4 [6]	ccgccgtatgagttctgtg	caggctgagaggtctccaa	57
PAX6	tcaccatggcaaataacctg	cagcatgcaggagtatgagg	20
POLR3G	cgcagaacaggaggaatatga	cactgtctgcgccaaaatc	35
POLR3GL	gcccagtacatttcaagttgg	gggcctgggtattcagagat	62
SOX2	atgggttcggtggtcaagt	ggaggaagaggtaaccacagg	19
TUBB3	gcaactacgtgggcgact	cgaggcacgtacttgtgaga	78

4.5 Protein expression analyses (I,II)

To analyze protein expression levels, samples were lysed in a cell lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5% TX-100, 5% glycerol, 1% sodium dodecyl sulfate (SDS), 1 mM Na_3VO_4, 10 mM NaF, and 1 mM phenylmethanesulfonyl fluoride). Protein concentrations were determined with DC Protein Assay (Bio Rad), and boiled with 6xSDS sample buffer (0.5 M Tris-HCl pH 6.8, 28% glycerol, 9% SDS, 5% 2-mercaptoethanol,

0.01% bromphenol blue). Samples were run on 10% SDS polyacrylamide gel and transferred to a nitrocellulose membrane. Membranes were incubated with primary antibodies over night at +4°C and after 3x5 min washes with Tris-Buffered Saline and Tween (TBST, Biotop) horseradish peroxidase conjugated secondary antibodies were incubated for 1 hour at room temperature. Enhanced chemiluminescence reagent (Amersham Biosciences) or developing solution (Pierce) was used for detection.

4.6 Antibodies (I,II)

L1TD1-$[His]_6$ -protein produced from pET20b vector was over expressed in *Escherichia coli* strain BL21(DE3)CE4 (Avidis) using 0.4 mM isopropyl-β-D-thiogalactopyranoside (AppliChem) for induction. L1TD1 protein was isolated from inclusion bodies, solubilized, and purified using His-tag Talon metal affinity resin (Clontech). 10% SDS gel was used to size separate protein antigen. L1TD1 protein was extracted from the gel and its identity was verified by liquid chromatography-tandem mass spectrometry. Rabbit polyclonal antibody against this purified L1TD1 protein antigen was produced by BioGenes (Germany). Custom designed peptide ISKERQRDIEERSR was used to produce another polyclonal rabbit antibody with peptide immunization (BioGenes). Commercial and other antibodies used in the studies are listed in table 5.

Table 5. Antibodies used in the studies. Abbreviations: WB, western blotting; FC, flow cytometry; ChIP, chromatin immunoprecipitation; I, immunohistochemistry; IP, protein immunoprecipitation

Target protein	Cat No	Company	Purpose
A2B5		Gift from P.W.Andrews	FC
A488 anti-mouse	A11001	Invitrogen	I
A488 anti-rabbit	A21441	Invitrogen	I
A555 anti-mouse IgG (H+L)	A21424	Invitrogen	I
A594 anti-rabbit	A11037	Invitrogen	I
ACTIN B	A5441	Sigma	WB
AGO2	011-22033	Wako	WB, I
Anti-Goat-HRP	554021	BD Pharmingen	WB
Anti-Mouse-HRP	Sc-2005	SantaCruz Biotechnology	WB
Anti-Rabbit-HRP	Sc-2020	SantaCruz Biotechnology	WB
CASP3	96625	Cell Signaling	WB
DCP1A	Ab57654	Abcam	I
DNMT3B	Ab13604	Abcam	WB
EEA1	Ab70521	Abcam	I
FITC anti-mouse IgG+IgM	M30801	Caltag Laboratories	FC
GAPDH	5G4	HyTest	WB
Goat IgG	Sc-2028	SantaCruz Biotechnology	ChIP
GW182	Ab70522	Abcam	I
L1TD1	HPA028501	Sigma	WB,IP,I
LIN28	Ab46020	Abcam	WB
LIN28	Ab75483	Abcam	I
NANOG	Sc-33759	SantaCruz Biotechnology	WB
NANOG	AF1997	R&D Systems	WB,ChIP
OCT4	Sc-9081	SantaCruz Biotechnology	WB
OCT4	Sc-8628x	SantaCruz Biotechnology	ChIP
P21	sc-817	SantaCruz Biotechnology	WB

Target protein	Cat No	Company	Purpose
P3X		Gift from P.W.Andrews	FC
PABP	Sc-32318	SantaCruz Biotechnology	WB
PAX6		DSHB	WB
POLR3G	Sc-21754	SantaCruz Biotechnology	WB
POLR3GL	HPA027288	Sigma	WB
pSMAD1/5/8	9511L	Cell Signaling	WB
Rabbit IgG	#2729s	Cell Signaling	ChIP,IP
RHA	Ab54593	Abcam	WB
SOX2	AF2018	R&D Systems	WB
SOX2	#5024s	Cell Signaling	ChIP
SSEA-1		Gift from P.W.Andrews	FC
SSEA-3		Gift from P.W.Andrews	FC
TIA1	Ab2712	Abcam	I
TRA-1-60		Gift from P.W.Andrews	FC
TRA-1-60	MAB4360	Millipore	I
TRA-1-81		Gift from P.W.Andrews	FC
V5	P/N46-0705	Invitrogen	I

4.7 Immunofluorescence (II)

Immunofluorescence analysis of protein localizations was performed with cells plated on microscopy cover slips. Mitotically inactivated fibroblasts were plated 24 hours prior to hESC plating. Cell transfections were performed using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol and immunocytochemistry was done 48 hours after transfections. The cells were fixed with 4% paraformaldehyde for 10-20 minutes at room temperature. Cell surface proteins were stained 30 minutes at room temperature. Permeabilization for intracellular protein staining was done with 0.1%-1% Triton-X 100 for 20 minutes at room temperature. Incubation of antibodies against intracellular proteins was 30 minutes at room temperature. Last, incubation of 4',6-diamindino-2-phenylindole (1µg/ml) (Invitrogen) for one minute was done to stain nuclear DNA. Fluorescence images were captured with Zeiss AxioVert 200M or with Leica TCS Sp2 confocal microscope. Colocation of proteins was analyzed with Leica TCS confocal software. Antibodies used are listed in table 5.

4.8 Immunoprecipitations

4.8.1 Chromatin immunoprecipitations (II)

Chromatin immunoprecipitation studies were used to determine the core pluripotency transcription factor binging to L1TD1 promoter region using protocol from Li *et al*. (Li et al., 2003). 500 µg of sonicated chromatin and 10 µg of antibodies were used per reaction. Antibodies used are listed in table 5. Anti rabbit (#112.04) or protein G (#112.02) magnetic beads (Dynal Biotech, Life Technologies) were used for precipitation. Reverse cross linking was done at 65°C for 12 hours and DNA was treated with Proteinase K and RNase A and purified with QIAquick PCR purification kit (Qiagen). Binding of transcription factors was confirmed with PCR using primers upstream of *L1TD1* transcription start site. Primers: Sense: 5'-AGGTGACCTTGGGGTTCAG-3', antisense: 5'-TCCCCGGAAATCGCATTC-3'.

4.8.2 Protein-protein immunoprecipitations (II)

Cells were washed twice with cold Phosphate-Buffered Saline (PBS) and lysed into NP40 cell lysis buffer (20-50 mM Tris, 150 mM NaCl, 0.5% sodium deoxylate, 0.5% NP40) supplemented with PhosSTOP and complete EDTA-free protease inhibitors (both from Roche). Cell lysates were incubated in the presence or absence of 10-1000 µg RNase A (Qiagen) to identify direct protein-protein interactions and RNA-dependent interactions. For precipitations Bio-Adembeads (Ademtech) or proteinG Dynabeads (Invitrogen) were used according to manufacturers' protocol.

4.8.3 Protein-RNA binding assays (II)

The protocol of Hafner *et al*. (Hafner et al., 2010) was used to analyze L1TD1 bound RNA molecules. The cells were cultured to be around 50% confluency when 100-500 µM 4-thiouridine containing medium was added, followed by 16-20 hours incubation at 37°C. Protein-RNA complexes were UV (365 nm) cross linked and harvested. Cells were lysed in NP-40 buffer (chapter 4.8.2) followed by immunoprecipitation with L1TD1 antibody and cross linking to magnetic proteinG Dynabeads (Invitrogen) according to manufacturer's protocol. L1TD1 bound RNA molecules were run on SDS-PAGE gel and extracted. From purified RNA sample cDNA was synthesized using first strand cDNA synthesis kit (Roche) with random hexamer primers. RT-qPCR analysis of target genes was performed as described in 4.4.6.

4.9 Flow cytometry (II)

For flow cytometry analyses cells were collected from Matrigel plates with trypsin and washed with surface staining buffer (D-PBS, 2% FCS, 0.01% azide). Primary antibodies were incubated for 30 minutes at 4 °C followed by two washes with surface staining buffer. Secondary antibody incubation was 30 minutes at 4 °C. Antibodies are listed in table 5. Cells were run with FACS Calibur (BD) and analyzed with Cell Quest FACS Diva.

4.10 Proliferation and colony forming assays (II)

The cell growth was analyzed using CellTiter 96 Nonradioactive Cell Proliferation Assay (Promega) according to manufacturer's protocol with 1 hour solubilization step. Colony forming assays were performed with L1TD1 and non-targeted siRNA treated cells cultured for 12 days and stained with crystal violet.

4.11 Validation of Karyolite™ BoBS™ for detecting genetic changes (III)

KaryoLite™ BoBs™ BACs-on-Beads™ assay (Perkin Elmer) was tested in collaboration with Perkin Elmer for analysis of genomic stability of hES cells. Sample preparation was done according to the manufacturer's protocol. BACs (bacterial artificial chromosomes) are large cloned sequences of human DNA, which in this case are targeted to all human chromosomes, at least two probes targeted to each chromosome arm. BACs are immobilized onto encoded microspheres (Luminex) distinguishable by the Luminex instrument system. Genomic DNA was extracted with the PureLink PCR Purification Kit (Invitrogen) according

to manufacturer's protocols. Purified sample and reference DNAs were labeled with biotin and purified. Labeled DNA was hybridized onto BACs bead mix. Hybridized beads were washed and incubated with the streptavidin-phycoerrythrin reporter dye. BoBsoft™ analysis software (Perkin Elmer) was used for analysis and data interpretation to analyze the relative amount of fluorescent DNA bound to the beads. The BoBsoft™ presents the karyotypes as a ratio plots against female and male references. The KaryoLite™ BoBs™ method was compared to previously published genomic analyses made with Affymetrix SNP6.0 arrays by reanalyzing DNA samples (Narva et al., 2010) in addition to new samples. To analyze the detection threshold of KaryoLite™ BoBs™ for mosaic cell populations, a dilution series of karyotypically abnormal cells was prepared. H9 cells with normal genome were mixed with a H9 line that had trisomy of chromosome 12. Cells were mixed in 0%, 20%, 30%, 50%, and 100% proportions.

4.12 Ethical consideration

Ethics Committee of South-West Finland Hospital District has given the permission for the culture of human embryonic stem cell lines. The research was carried out following good scientific practice and guidelines of the National Advisory Board on Research Ethics.

5 RESULTS

5.1 Identification of novel factors regulating hESC pluripotency (I, II)

Pluripotent stem cells offer a valuable resource for developmental studies, disease modeling, and new possibilities for medical applications and regenerative medicine. Before pluripotent cells can be used in any therapeutic applications in humans, the detailed properties of the cells need to be understood. Identification of new factors involved in the regulation of stemness and understanding the basic properties of the cells are important considerations in these applications.

We have used genome wide Affymetrix microarrays to analyze the transcriptomes of a large set of pluripotent hES and differentiated cells in order to identify novel factors expressed specifically in pluripotent cells. The sample set consisted of 9 different studies of 18 hESC lines, hESC derivatives, and somatic tissues. This analysis resulted in the identification of a list of stem cell specific genes, that included already known stem cell markers and two novel factors, L1TD1 and POLR3G (I: Figure 1). Based on their very specific expression profiles in pluripotent cells, L1TD1 and POLR3G were selected for further studies on their roles in pluripotent cells. POLR3G is a polypeptide of DNA directed RNA polymerase III multisubunit complex, whereas L1TD1 was an expressed sequence tag with no known protein or function at the time of identification in the transcriptome analysis.

5.1.1 POLR3G and L1TD1 are specifically expressed in pluripotent cells

After observing that POLR3G and L1TD1 were highly and specifically expressed in hESCs in the genome wide transcriptome analysis, the results were validated using various methods. First, the high expression of both genes in pluripotent cells and decline in expression in response to differentiation was confirmed experimentally both at the protein and mRNA levels using undifferentiated, RA induced, and spontaneous EB differentiation of hESCs (I: Figure 2B,C and II: Figure 1A-D). Also, Helicos next-next-generation RNA sequencing was able to verify the expression of both factors in pluripotent cells and down regulation in response to differentiation (I: Figure 2A). The expression levels and down regulation kinetics of POLR3G and L1TD1 were comparable to known pluripotency factors, indicating role in the regulation of stem cell fate.

5.1.1.1 POLR3G and POLR3GL have inverse expression profiles

POLR3G has been reported to have an isoform named POLR3GL, and the isoforms have partially different expression profiles in human cell lines (Haurie et al., 2010). We studied the expression profiles of these two isoforms in pluripotent and differentiated hESCs and were able to show reciprocal protein level expression of the isoforms in the undifferentiated state and in response to induced and spontaneous cell differentiation (I: Figure 6B,C). As already mentioned in chapter 5.1.1, the mRNA levels of POLR3G decrease during differentiation. In contrast, mRNA levels of POLR3GL are not significantly changed during differentiation (I: Figure 6A) indicating POLR3GL to be regulated in post-transcriptional manner.

5.1.2 POLR3G and L1TD1 are regulated by core stem cell factors

Because of the stem cell specific expression profiles of POLR3G and L1TD1, we were interested to study whether core transcription factors are responsible for the expression of *POLR3G* and *L1TD1* in hESCs. ChIP-on-chip studies indicate that SOX2 and NANOG are binding to *L1TD1* promoter and OCT4, SOX2, and NANOG are binding to *POLR3G* promoter (Boyer et al., 2005). Moreover, L1TD1 and POLR3G both appear to have activating H3K4me3 epigenetic histone mark on their promoters in undifferentiated hESCs (Zhao et al., 2007).

The occupancy of core stem cell factors on *L1TD1* promoter was analyzed using chromatin immunoprecipitation with OCT4, NANOG, and SOX2 antibodies. The results showed amplification of *L1TD1* transcriptional start site with all factors, demonstrating all the three transcription factors binding to the *L1TD1* promoter (II: Figure 2B). Furthermore, the effects of core stem cell factors on L1TD1 were analyzed by measuring the levels of *L1TD1* mRNA transcripts in response to siRNA mediated silencing of OCT4, NANOG, and SOX2. The results illustrated clear down regulation of *L1TD1*, indicating reduced transcription (II: Figure 2A).

When analyzing the regulation of POLR3G by the core stem cell factors, siRNA silencing and western blotting was used. POLR3G protein levels were always negatively regulated in response to OCT4 siRNA treatment. The effects of NANOG, SOX2, and L1TD1 silencing on POLR3G decrease were not consistent between the replicates, indicating that OCT4 has the master role in the regulation of POLR3G, and the effects of other factors might be indirect. (I: Figure 3)

5.1.3 POLR3G and L1TD1 are required for the pluripotency maintenance and self-renewal

Because of their specific expression profiles in pluripotent cells, we expected both POLR3G and L1TD1 to have an important role in the regulation of stemness and self-renewal. The importance of these factors was studied by using RNAi with siRNA oligonucleotides targeted to POLR3G and L1TD1, and also inducible stable expression of L1TD1 shRNA in clonal NT2D1 cell lines. The effects of POLR3G and L1TD1 silencing were measured with a wide range of experiments.

5.1.3.1 L1TD1

Efficient silencing of L1TD1 was confirmed by analyzing L1TD1 expression levels by western blotting and RT-qPCR (II: Figure 1E,G). Silencing of L1TD1 showed significant effect on known pluripotency markers, OCT4, NANOG, SOX2, and LIN28, suggesting an important role for L1TD1 in the regulation of pluripotency and self-renewal (II: Figure 1E, 1G, 5I). When analyzing cell surface markers with flow cytometry from siL1TD1 treated cells, a decrease of pluripotency associated markers (SSEA-3, TRA-1-60) was detected, whereas the differentiation associated cell surface markers (SSEA-1, A2B5) were induced (II: Figure 1F).

Inducible and stable knockdown of L1TD1 in NT2D1 cell line resulted in similar conclusions as stated above. The down regulation of core pluripotency regulators OCT4, NANOG, and

SOX2 was detected with western blotting. The samples were also analyzed with Illumina BeadChips to identify the genome wide effects of L1TD1 silencing. The array analysis showed the reduction of *SOX2* and *PRDM14,* indicating loss of self-renewal. In addition, pathway analysis revealed up regulation of trophectodermal marker GATA2, and CEBPB, and CDH1, suggesting cell differentiation in response to L1TD1 reduction (data not shown).

Taken together, the data indicated loss of pluripotency and induction of differentiation in response to reduced levels of L1TD1. Our results suggest a vital role for L1TD1 in the pluripotency maintenance of hESCs.

5.1.3.2 POLR3G

The efficient knockdown of POLR3G in response to siRNA treatment was confirmed at the protein and mRNA levels (I: Figure 4C,D). siRNA mediated knockdown of POLR3G led to a significant change in the cell morphology and decreased cell numbers, and that was observed already 24 hours after siRNA introduction (I: Figure 4A,B). Importantly, the effects of POLR3G silencing on known pluripotency regulators demonstrated POLR3G to be required for self-renewal. Decreased expression of *POU5F1*, *NANOG*, and *L1TD1* mRNA and OCT4 and NANOG protein levels was observed (I: Figure 5). Interestingly, analysis of SOX2 levels gave inconsistent results. The mRNA level of *SOX2* was found to increase, whereas the measured protein level was decreased (I: Figure 5A,B). In addition to *SOX2* induction, we also identified neuroectodermal differentiation markers *PAX6* and *TUBB3* mRNAs to be induced in response to POLR3G silencing (I: Figure 5A). Induction of mesoderm markers *GATA2*, *SOX17*, *T*, and *aMHC* were observed in some of the replicates of siRNA cultures, although not consistently. We did not detect up regulation of endo- or trophoblast markers (data not shown). Taken together, silencing of POLR3G resulted in reduced self-renewal and increased expression of neuroectodermal differentiation markers, indicating POLR3G to be vital for pluripotency maintenance of hESCs.

5.2 Functional characterization of POLR3G in hESC (I)

5.2.1 Silencing of POLR3G leads to decreased cell proliferation

We detected a significant effect on cell growth and proliferation in siRNA mediated silencing of POLR3G. Depletion of POLR3G in hESCs resulted in at least 50 % decrease in cell numbers compared to non-targeted siRNA treatment, and the effect correlated with knockdown efficiency (I: Figure 4B-D). However, the number of dead cells was not observed to be increased in response to knockdown, as was calculated with Cedex XS system and trypan blue staining. We analyzed apoptosis and cell cycle arrest proteins from the siPOLR3G samples to see if the decreased proliferation was due to induction of these factors. Proliferation inhibitor protein 21 (p21, CDKN1A) was not detected to be induced in response to siPOLR3G treatment, suggesting that the proliferation block was independent of p21 (I: Figure 5B). Also, we measured the levels of apoptosis related executioner caspase, CASP3, and found the active form of the protein to be down regulated in response to siPOLR3G, supporting the observation that cells were not apoptotic (I: Figure 5B).

5.2.2 POLR3G target genes are involved in vital intracellular programs and early development

We analyzed the possible target genes of POLR3G by performing RNA-seq deep sequencing analysis from the samples where POLR3G was silenced. The goal was to identify transcripts regulated by POLR3G and to identify the pathways regulated in response to loss of POLR3G. The analysis resulted in identification of 1180 of significantly down regulated genes and 369 of up regulated genes, with a cutoff of ≤0.05 for the *p*-value and ≥1.5 for the logFC. The higher number of significantly down regulated genes suggests that POLR3G mostly positively regulates genes.

The cellular functions that were most strongly affected by POLR3G suppression included intracellular programs vital for cellular maintenance and early development (Ingenuity Pathway analysis) (I: Figure 7). Interestingly, CDKN2A, a negative regulator of cell cycle, was strongly induced (logFC 3.9) in the dataset, indicating that it might be a factor responsible for the detected proliferation block. The pathway analysis revealed that potential upstream regulators of affected target genes could be for instance TP53 and FGF1, both of which have important roles in stem cell functions. Importantly, the developmental programs affected included development of organismal, nervous system, and tissue development, as well as morphology of tissue and organs, supporting our findings and showing a crucial role for POLR3G in the regulation of early development.

5.3 The function of L1TD1 in hESC (II)

5.3.1 L1TD1 has the ability to bind RNA molecules

L1TD1 belongs to the same family of proteins than mammalian LINE-1 element, ORF1 protein. ORF1p is required for retro transposition and able to bind nucleic acids (Martin, 2006). The amino acid sequence of L1TD1 and alignment analysis with ORF1 revealed that L1TD1 is composed of three different domains, including an RNA recognition motif (RRM), coiled-coil domain and c-terminal domain (CTD) (II: Figure 4A). Also the gene expression profile analysis showed that *L1TD1* expression is highly correlated with RNA binding proteins (RBSs), including *LIN28* and *ESRP1* (II: Figure 4C). Based on this information we wanted to test whether L1TD1 is an RBP.

To test if L1TD1 can bind RNA, an immunoprecipitation study was performed. L1TD1 bound molecules were analyzed on SDS-PAGE gel, extracted, and purified. From the purified RNA fraction we were able to amplify *L1TD1* mRNA transcript (II: Figure 4B), which was expected based on the fact that ORF proteins tend to bind their own transcripts (Martin, 1991; Martin, 2006). We also examined if sequences of *POU5F1*, *NANOG, LIN28*, or *EF1α* were present in the RNA fraction, to test if L1TD1 has role in the regulation of their mRNA stability. However the results were negative, at least with the primer pairs used for the analysis (data not shown). *POU5F1* was analyzed with six different primers designed to cover different parts of the amplicon, but the result was still negative. *POU5F1* was carefully tested, since LIN28 has been shown to regulate its stability and translation (Qiu et al., 2010). Nevertheless, we concluded L1TD1 to be an RBP.

5.3.2 *L1TD1 is localized to RNA-protein-complexes and translational machinery with LIN28 and RHA*

After identifying L1TD1 as an RBP, we were interested to identify its interaction partners. Previously, we had identified L1TD1 to localize in the cells in cytoplasm as spherical structures. We were able to identify these structures as processing bodies (P-body) and co-location of L1TD1 with a P-body marker AGO2 (Sen and Blau, 2005) and LIN28. Further, immunoprecipitation analyses showed positive signal demonstrating L1TD1 and LIN28 proteins to interact. The immunoprecipitation with AGO2 was negative (II: Figure 5C,D). Because P-bodies are not membrane enclosed structures, but dynamic structures composed of protein and RNA molecules, we tested if the L1TD1-LIN28 interaction is direct protein-protein interaction. Immunoprecipitation reaction was treated with RNase A to eliminate the RNA molecules. As a result, the LIN28 interaction with L1TD1 was lost, indicating the interaction to be RNA dependent. Furthermore, immunostainings performed in the presence of RNaseA showed spherical L1TD1-AGO2 structures to vanish (II: Figure 5B). Similar results were obtained with siL1TD1 treated cells analyzed by immunofluorescence, where both AGO2 and LIN28 expression were lost, indicating reduction of P-bodies in response to L1TD1 silencing (II: Figure 5F,G). The results suggest a crucial role for L1TD1 in the RNA protein complexes.

LIN28 has been shown to bind RNA helicase A (RHA) and to be involved in the translational machinery (Jin et al., 2011). We were interested to test if also L1TD1 is part of translational machinery. Indeed, we detected direct protein-protein interaction with RHA, and RNA dependent interaction with poly A binding protein (PABP), that is also part of the translational machinery (II: Figure 5H). Moreover, we tested if a decrease in L1TD1 expression has any effect on its interaction partners, and showed these proteins to be down regulated in response to L1TD1 silencing (II: Figure 5I). We concluded that, L1TD1 is an RBP that mediates its functions by operating in the cytocolic RNA protein complexes and translational machinery. The schematic illustration of L1TD1 with its interaction partners is presented in Figure 6.

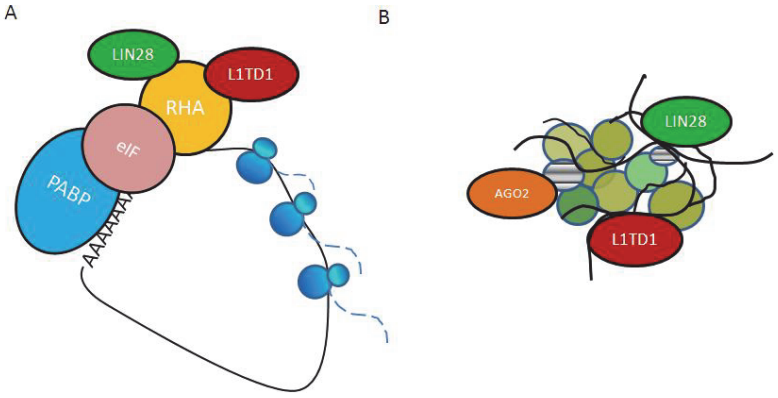

Figure 6. L1TD1 is associated with translational machinery and P-bodies. A) Schematic illustration of the interactions of L1TD1with RHA, PABP, and LIN28 in the translational polyribosome. eIF represents eukaryotic initiation factors. B) Schematic illustration of L1TD1, LIN28, and AGO2 locations to P-body, a dynamic protein-RNA aggregate.

5.3.3 L1TD1 is expressed in certain cancers and affects cell proliferation

In addition to pluripotent stem cells, L1TD1 was identified to be expressed in certain cancers, like many other stem cell specific genes. In Silico Transcriptomics database (Kilpinen et al., 2008) showed L1TD1 to be expressed in colorectal carcinoma, seminoma and non-seminoma of testis, ovarian germ cell tumors and Ewing's sarcoma (II: Figure 3A). We used seminoma cell line TCam2 and nonseminomatous cell line 2102Ep to study the expression of L1TD1 and effect of its knockdown in these cell lines. We detected high L1TD1 expression in both of the cell lines and reduction of self-renewal in response to L1TD1 silencing, as OCT4 and NANOG were down regulated (II: Figure 3B). Moreover, the pathway analysis of NT2D1 EC cells with stable L1TD1 silencing indicated that genes affected by L1TD1 are involved in cancer and cell proliferation. Consistent with this notion we measured reduced colony formation capacity and reduced proliferation in the cells where L1TD1 expression was silenced (II: Figure 3C). Based on our results, L1TD1 is needed also for the self-renewal and proliferation of certain cancer cell lines and could provide a new diagnostic marker for these cancers. However, confirming this hypothesis need further investigation.

5.4 Karyotyping of hESC with KaryoLite™ BoBs™ (III)

Since the genomic stability of pluripotent cell lines in long term culture is shown to associate with karyotypic abnormalities, the cell cultures need to be routinely checked. Previous studies have identified that the most common genomic change occurring during hESC cultures involve gains of partial or whole chromosomes. (Lund et al., 2012). In collaboration with Perkin Elmer, KaryoLite™ BoBs™ was tested for its suitability to be used as method for hESC karyotyping. We compared KaryoLite™ BoBs™, conventional G-banding, and array based methods by reanalyzing samples from previous studies (Narva et al., 2010) in addition to new samples from karyotypically normal and abnormal hESCs. We tested KaryoLite™ BoBs™ for its capacity to detect known abnormalities in sample set that included seven different hESC lines (12 samples), most of which had been previously analyzed with both G-banding and Affymetrix high resolution SNP6.0 arrays. The panel of samples represented normal karyotypes as well as chromosomal aberrations, both losses and amplifications.

KaryoLite™ BoBs™ analysis did not detect genomic abnormalities in samples with normal karyotypes, as predicted. The abnormal samples used in the study and characterized with G-banding and/or SNP6.0 array, were detected to be abnormal also by KaryoLite™ BoBs™ (III: Figure 1). In addition, KaryoLite™ BoBs™ and Affymetrix SNP6.0 detected gain in a particular sample that surprisingly was not detected in conventional G-banding karyotyping analysis. Equally to Affymetrix SNP6.0 array results, balanced translocations that were identified in G-banding were not detected with KaryoLite™ BoBs™. Taken together, KaryoLite™ BoBs™ was able to detect the majority of the karyotypic abnormalities detected using other methods, and concurred with the results obtained from G-banding and Affymetrix SNP6.0 array. KaryoLite™ BoBs™, however, is not appropriate for detecting balanced translocations. We concluded KaryoLite™ BoBs™ to be suitable for routine based high-throughput karyotyping of pluripotent cell lines in a cost-efficient manner.

5.4.1 Detection threshold of KaryoLite™ BoBs™

Cell cultures frequently consist of mosaic populations that contain cells with more than one genotype. We tested the sensitivity of KaryoLite™ BoBs™ to detect karyotypic changes in

heterogenic cell populations. For this purpose, karyotypically normal H9 cells were mixed with cells from the same H9 cell line that had trisomy of the chromosome 12. The proportions analyzed were 0%, 20%, 30%, 50%, and 100%. When the proportion of karyotypically abnormal cells was ≥30%, the amplification was detected with KaryoLite™ BoBs™ (III: Figure 2). However, mosaicism of FES29p37 cell line previously detected with G-banding, was not detected with KaryoLite™ BoBs™ or Affymetrix SNP6.0 (Narva et al., 2010). Taken together, KaryoLite™ BoBs™ enables analysis of whole cell populations and the sensitivity of the assay can detect mosaic cell populations when the percentage of abnormal cells is at least 30%.

6 DISCUSSION

6.1 POLR3G is required for pluripotent hES cells (I)

POLR3G was identified as a hESC specific transcript in our genome-wide transcriptome analysis. The stem cell specific expression profile of POLR3G was also listed in previous studies, emphasizing its universal expression in pluripotent cells and supporting our observation (Armstrong et al., 2006; Assou et al., 2007; Enver et al., 2005). Interestingly, data from Enver *et al*. showed POLR3G to be most highly expressed in the most pluripotent fraction of stem cells, sorted based on SSEA-3 marker, highlighting its role in pluripotent cells.

While our study was in progress, Haurie *et al.* and Wong *et al*. also reported hESC specific expression profile of POLR3G (Haurie et al., 2010; Wong et al., 2011). Wong *et al*. further showed that POLR3G regulates pluripotency of hESCs. Our results are in keeping with those of Wong *et al*. concluding POLR3G to be vital for pluripotency maintenance.

We detected a significant effect on hESC proliferation in response to POLR3G silencing. Enver *at al*. also detected reduced numbers of NTERA2 EC cells as a result of POLR3G depletion, supporting the idea that POLR3G has role in cell growth. Pol III transcription products are reported to be essential for cellular processes, including transcriptional regulation, as well as translation and RNA processing, thus important for cell growth (Teichmann et al., 2010; White, 2011). Pol III complex is composed of 17 subunits, of which POLR3G (RPC32/RPC7) is the only pol III specific subunit with no sequence/structure homology with pol I or pol II transcriptional machineries (Teichmann et al., 2010). The ternary complex of RPC32(POLR3G)/RPC39/RPC62 has been shown to be needed for accurate transcriptional initiation of pol III (Wang and Roeder, 1997). RNA pol III occupies three different types of promoters, two gene internal promoters and one upstream promoter. The ternary complex is involved in all of these three promoter types, while the composition of general transcription factors is different between the promoters (Figure 7). These results further support our findings of the importance of POLR3G to cell proliferation. Loss of POLR3G is likely to influence the whole pol III complex at all promoter types and all pol III transcripts, which further results in the reduced expression of structural molecules vital for cell function and protein synthesis.

Development and differentiation depends on spatially and temporally accurate execution of gene expression. The transcriptional machinery of active pol II gene promoters in pluripotent hESC has a more distinct composition of transcriptional components than the differentiated derivatives (Maston et al., 2012). Apparently, the same applies to pol III transcription. For instance, two paralog subunits (BRF1 and BRF2) of pol III transcription factor IIIB (TFIIIB) occupy different types of gene promoters (Moqtaderi et al., 2010). Cell type specific components of the transcriptional machinery enable complex tissue specific transcriptional programs, and pol III transcription has been associated with stage and tissue specific gene expression (Dieci et al., 2007; Hochheimer and Tjian, 2003). In vertebrates, one pol III transcribed gene (BC1 RNA) has been shown to have a neuron specific expression (Martignetti and Brosius, 1995; Teichmann et al., 2010). Our results on the hESC specific expression profile of POLR3G and the inverse expression with POLR3GL isoform may provide a new pair of factors regulating cell type specific transcription. We propose that

POLR3G is the key component of the pol III complex in pluripotent cells, and during the early differentiation is replaced by an alternative isoform POLR3GL. Comparison of the genome-wide ChIP profiles of the isoforms would provide novel information on the specific transcriptional profiles. Also, whether there are other changes in the pol III complex during hESC differentiation remains to be determined.

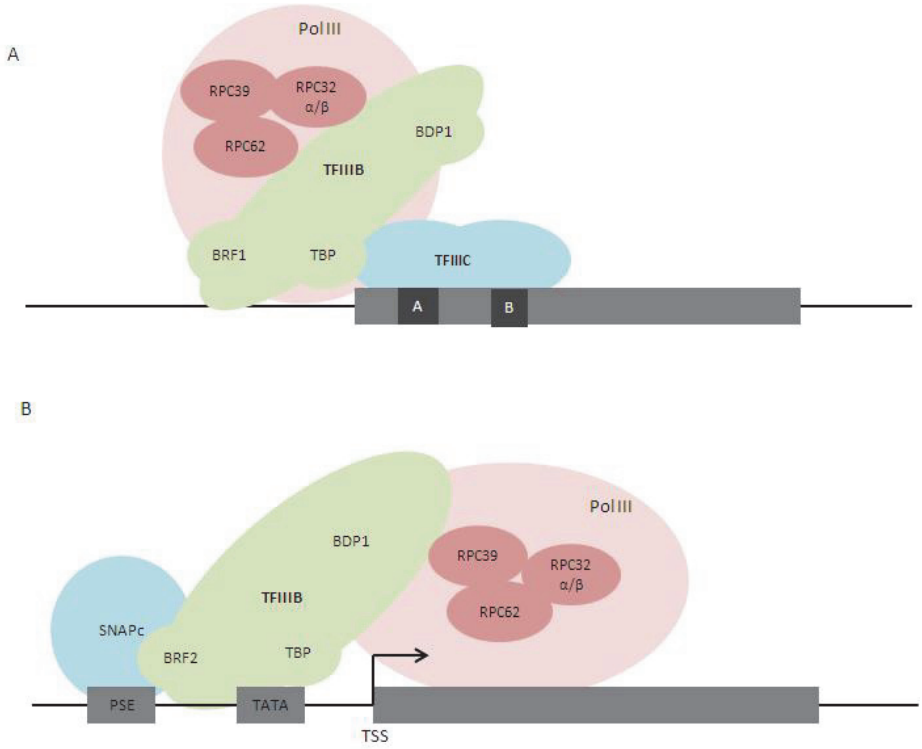

Figure 7. Schematic illustration of POLR3G(RPC32α) and POLR3GL(RPC32β) in the multisubunit complex of pol III together with general transcription factors on the gene promoter areas. A) An example of a gene internal promoter B) An example of an upstream gene promoter. Different promoter types have different compositions of transcription factors. POLR3G containing ternary complex (dark red) of pol III complex (light red) is involved in all types of transcribed promoters. Abbreviations: PSE: proximal sequence element, TBP: TATA box binding protein, SNAPc: small nuclear RNA activating protein complex, TSS: transcription start site, TFIIIC: transcription factor III C. Modified from (Dumay-Odelot et al., 2010; Teichmann et al., 2010; White, 2011).

Silencing of POLR3G led to reduced self-renewal and induced differentiation, indicating that POLR3G containing pol III regulates pluripotency associated transcripts. We used genome-wide RNA sequencing to identify the genes affected by depletion of POLR3G in hESCs. The RNA-seq analysis was made for RNAs extracted to represent long, >200nt, RNA molecules. As a result of Ingenuity Pathway Analysis, we concluded that POLR3G regulates global intracellular programs vital for cellular maintenance and early development. This was consistent with our previous observations from siRNA studies. Based on the pathway analysis, TP53 was among the top candidates of upstream regulators of the genes significantly regulated in response to POLR3G depletion. TP53 and RB have been shown to

regulate pol III transcription by binding to and inactivating TFIIIB (Felton-Edkins et al., 2003). This supports our results indicating that the affected genes represent genes transcribed by pol III. Interestingly, c-Myc has been shown to regulate and activate pol III transcription (Felton-Edkins et al., 2003). Deregulation, and accumulation, of pol III transcripts is frequently associated to cancers, which can be explained by the regulation of these known cancer related proteins. This further links pol III transcription to accelerated cell growth and proliferation (White, 2008).

Pol III is considered to transcribe mostly ncRNAs. However, the study from Oler *et al.* in Hela cells also identified protein coding mRNAs as pol III transcripts (Oler et al., 2010). We identified several mRNAs to be significantly regulated in response to POLR3G silencing. To identify which of the affected genes in our RNA-seq data are direct targets of POLR3G would require ChIP studies (in progress). Comparison of the existing RNA-seq data with ChIP dataset would indicate which RNAs are transcribed by POLR3G containing pol III. In addition, small RNAs, the majority of known targets of pol III transcription, are not included in our analysis. We are in the process of analyzing the small RNA fraction of pluripotent hESCs and in response to POLR3G silencing, with a goal to identify different types of RNA populations and potential pluripotency specific transcripts. Moreover, these sequencing measurements have been performed in great depth, thus also providing information on alternatively spliced pol III genes. The results are expected to offer novel valuable information on pol III function in pluripotent cells.

To conclude, the POLR3G subunit is a crucial component of the pol III complex in the regulation of hESC specific transcriptional programs required for maintenance of hESC self-renewal and pluripotency. The schematic representation of pol III promoters and POLR3G/POLR3GL are presented in Figure 7.

6.2 L1TD1 is required for pluripotent hES cells (II)

L1TD1 was also identified in our genome-wide transcriptome analysis as a hESC specific transcript. L1TD1 has been listed in previous studies to be a pluripotent cell specific gene (Armstrong et al., 2006; Assou et al., 2007; Meshorer et al., 2006; Sperger et al., 2003). L1TD1 was first identified as an ES cell associated transcript (ECAT11) in mouse ES cells based on *in silico* digital differential display by Mitsui *et al*. (Mitsui et al., 2003). During the time of our studies, L1TD1 (FLJ10884) was only recognized as an expressed sequence tag with no known protein or function.

We concluded L1TD1 to be hESC specific gene vital for the pluripotency maintenance. While our study was in progress, Wong *et al*. reported L1TD1 to be a marker for pluripotent hESCs (Wong et al., 2011). Our results correlate well with the results from Wong *et al*. concerning the hESC specific expression profile. Almost at the same time as Wong *et al*., Iwabuchi *et al*. (Iwabuchi et al., 2011) reported a mouse ortholog L1td1 to have stem cell specific expression profile in mouse ES cells, but being dispensable for pluripotency maintenance.

The differences between our results and Iwabuchi *et al*. could possibly be explained by the species specific differences. These orthologous proteins share only 45% identity in their total amino acid sequences, and 71% and 57% identity in RRM and CTD functional domains of the protein, respectively. Based on Iwabuchi *et al*., there are number of L1 sequences

scattered in the mouse genome and these L1TD1 related proteins might compensate L1TD1 function in the knock-out mice model studied by them. There are many differences reported in the regulation of pluripotency in mouse and human ES cells. For instance, the main pathways regulating hESC pluripotency include TGFβ and FGF signaling, whereas in mESCs the BMP and LIF pathways have the main function in the maintenance of self-renewal (Ng and Surani, 2011; Ying et al., 2003). Cross-species clustering analysis identified differences in signaling pathways and regulatory modules (Cai et al., 2010), hence the pluripotency circuitry is completely different between mice and men. In addition to transcriptional networks, the protein networks have been found to be significantly different between mouse and human ES cells; hESCs sharing approximately only 30% similarity in the protein expression profile with mESCs (Van Hoof et al., 2006).

Moreover, individual factors have been determined to behave differentially in mESCs and hESCs. For instance, PRDM14 has been identified as a key transcription factor required for the hESC pluripotency maintenance, whereas the mouse homologue of Prdm14 is not essential for mESC or EpiSC maintenance (Chia et al., 2010). Also, silencing of LIN28 in hESCs causes reduction of OCT4, whereas in mESCs Oct4 levels stay unchanged in response to siLin28 treatment (Qiu et al., 2010; Xu et al., 2009). Thus, many aspects suggest that there are species and cell type specific differences between mESCs and hESCs, and such differences may apply in the role of L1TD1 as well. However, an interesting and important aspect would be to analyze the role and expression of L1td1 in EpiSCs to see whether it resembles more mESCs or hESCs.

The location and function of L1TD1 seems to correlate well with LIN28. Both are localized to cytoplasmic P-bodies (Balzer and Moss, 2007) and translational machinery (Peng et al., 2011; Qiu et al., 2010). Interestingly, LIN28 has also been shown to bind its own mRNA as a mode of auto regulation (Wilbert et al., 2012), which further suggest the proteins to function in a similar manner. Since LIN28 has been shown to regulate the translation of OCT4 (Qiu et al., 2010), one intriguing aspect is that if L1TD1 participates to this regulation with LIN28. That would link L1TD1 directly to the core regulatory network of pluripotency, nevertheless, we were not able to demonstrate L1TD1 binding to *OCT4* in our experimental setting.

Interestingly, LIN28 has been shown to be specific marker for testicular and ovarian germ cell tumors (Cao et al., 2011; Xue et al., 2011) and associated with several cancers (Viswanathan et al., 2009). Moreover, LIN28 have been shown to promote cancer cell proliferation (Feng et al., 2012). We observed that L1TD1 interacts with LIN28 and affects its expression. Furthermore, L1TD1 and LIN28 are expressed in same cancers. This suggests that L1TD1 can also be a potential marker for these cancers. High expression of L1TD1 (FLJ10884) in germ cell tumor cell lines was also identified in the data of Sperger *et al*. (Sperger et al., 2003), where L1TD1 was listed as one of the most significantly expressed genes. The role of L1TD1 in cancer, however, needs further investigation.

The results show a crucial role for L1TD1 in the regulation of pluripotency and self-renewal in hESC. L1TD1 was identified as an RBP likely to be involved in post-transcriptional gene regulation, based on cellular location and interaction partners. Our ongoing studies aim at identification of L1TD1 interactome in hESC to reveal the network important for the maintenance of stem cell status. We are in a process of analyzing L1TD1 interacting proteins with mass spectrometry and immunoprecipitations. Further, a deep sequencing analysis of L1TD1 bound RNA molecules is ongoing. The results are expected to reveal the functional role of L1TD1 and to provide novel information on the pluripotency network of hESCs.

6.3 KaryoLite™ BoBs™ can be used to detect karyotypes of hESCs (III)

Conventional karyotyping, G-banding, has been the most commonly used method for screening of cell lines for chromosomal changes. The limitation of this method is the small cell numbers analyzed (around 50 mitosis), and relatively high cost and laboriousness. Array based methods provide higher resolution technique for the analysis of chromosomal stability, but are limited by the cost and complexity of the data analysis. KaryoLite™ BoBs™ (Perkin Elmer) was tested as a fast, simple, and low-cost alternative for routine monitoring of cell cultures.

KaryoLite™ BoBs™ is intended for the detection of gains and losses of DNA in particular chromosomal regions on the level of chromosome arms. The assay covers all 24 chromosomes, 1-22, X, and Y. KaryoLite™ BoBs™ is not suitable for detecting small scale changes or balanced translocations, because the technology of the BACs cover only proximal and terminal regions of the chromosome arms. Given that the majority of the reported chromosomal aberrations during prolonged hESC culture represent gains of partial or whole chromosomes (Lund et al., 2012), the chromosome arm detection level is enough for routine cell culture monitoring. The BACs are designed from human chromosomes, and thus the method is applicable for analyzing human cells only.

The BACs product family created by Perkin Elmer also includes the Prenatal BoBs™ assay that is targeted to detect 9 common microdeletion syndrome regions and most common aneuploidies. The assay has been used in prenatal diagnosis where it has been shown to be a reliable method to identify chromosomal changes (Grati et al., 2012; Kiiski et al., 2012; Paxton et al., 2012; Vialard et al., 2011). In addition to prenatal diagnosis, BACs technology has been used for evaluation of products of conception. Two studies used KaryoLite™ BoBs™ for this purpose, and concluded the method to be beneficial and suitable to be used for genetic analyses of miscarriage samples (Grati et al., 2012; Paxton et al., 2012). In conclusion, the BACs-on-beads™ technology can be used in several applications to detect chromosomal changes.

KaryoLite™ BoBs™ enables analysis of whole cell populations, for which we established the sensitivity for mosaic detection threshold to be 30%. This is in concordance with Vialard *et al*., who used Prenatal BoBs™ assay for detection of trisomy 21, and concluded the detection threshold to be 30%. Moreover, Paxton *et al*. came up with the same 30% threshold with KaryoLite BoBs™. Grati *et al*., however, detected mosaicism starting from at least 50% of mixed samples.

Taken together, the study validated KaryoLite™ BoBs™ as a new method for the analysis of genetic status of cell cultures, providing an alternative tool for analyzing pluripotent human cell lines.

7 CONCLUSIONS

Pluripotent stem cells provide enormous potential for various medical applications. This study focused on pluripotent human embryonic stem cells with a goal to identify and characterize factors involved in the maintenance of pluripotency and self-renewal of hESCs. Several transcriptome studies have been performed and published for this purpose, however, only few of the identified target transcripts have been further characterized. In this study, two novel regulators of hESC pluripotency were identified and functionally studied. In addition, a new improved method for detecting karyotypic abnormalities of hESC cultures was validated.

Using genome-wide analyses, POLR3G and L1TD1 were identified as interesting stem cell specific target genes. Functional studies in hESCs showed POLR3G to be vital for the maintenance of self-renewal and pluripotency. POLR3G is required for the proper function of polymerase III, and is possibly replaced by an isoform protein during cell differentiation. Moreover, genome-wide analysis of target genes showed that POLR3G regulates vital intracellular programs involved in cellular maintenance and early development.

L1TD1 was demonstrated to be an RNA binding protein with a significant role in the regulation of hESC pluripotency. L1TD1 was shown to associate to RNA-protein complexes and translational machinery in the cell cytoplasm. Moreover, an RNA dependent interaction with LIN28, and direct protein-protein interaction with RHA were discovered. L1TD1 is also associated with certain cancers and has an effect on the proliferation of cancer cell lines.

KaryoLite^{TM} BoBs^{TM} assay is based on BACs-on-Beads methodology, targeted to detect DNA sequences in particular chromosomal regions of each chromosome. The assay was tested and validated as a new high-throughput analysis method for detecting chromosomal changes in the long-term hESC cultures. KaryoLite^{TM} BoBs^{TM} is comparable to the other karyotyping methods used for this purpose.

In conclusion, this study provides valuable new information on the molecular mechanisms of pluripotency regulation. Importantly, two novel regulators of hESC stem cell status were identified. Functional analyses revealed the cellular processes affected by these proteins, thus bringing valuable information to the highly complex network of pluripotency maintenance. A new method was validated as a rapid, high-throughput, and cost-efficient tool for screening of human pluripotent cell cultures.

ACKNOWLEDGEMENTS

This study was carried out at the Turku Centre for Biotechnology, University of Turku and Åbo Akademi University and in the Department of Medical Microbiology and Immunology, University of Turku. I would like to express my deepest gratitude to a number of people who have made this achievement possible.

I would like to start by expressing my gratitude to my supervisor Professor Riitta Lahesmaa for all the support and guidance, and for trusting in me and giving me opportunity to do research independently. I am also very grateful to my other supervisor, Dr. Riikka Lund, for all the guidance, kindness, and encouragement. You both have helped me to grow as a person and as a scientist.

I am very grateful to Turku Doctoral Programme of Biomedical Sciences for the financial support and scientific education. My supervisory board members, Professor Sanjeev Galande and Professor Jorma Palvimo, are warmly acknowledged for their valuable comments and advice during these years. I also want to thank Professor Jari Koistinaho and Professor Olli Simell for carefully reviewing the thesis and providing me with valuable comments. Robert Moulder is acknowledged for the language revision of the thesis.

I owe my sincerest gratitude to all my co-authors and collaborators who have worked in the projects of this thesis. I would like to acknowledge Reija Autio, Duncan Baker, Konstantin Denessiouk, Maheswara Reddy Emani, Attila Gyenesei, Neil Harrison, Virpi Kivinen, Asta Laiho, Riikka Lund, Maia Malonzo, Tuomas Nikula, Juuso Nästi, Juha-Pekka Pursiheimo, Omid Rasool, Heli Skottman, and Professors Peter Andrews, Outi Hovatta, Harri Lähdesmäki, Matti Nykter, Timo Otonkoski, and Anjana Rao for the invaluable work, special expertise, and critical comments. Last but not least, my sincerest thanks go to Elisa Närvä, a wonderful person and an exceptional scientist. I have learnt so much from you and it has been a true pleasure to work with you.

Importantly, I would like to thank the skilful personnel and technical staff of Turku Centre for Biotechnology and its core facilities, especially all the people in the office, it-support, technical and instrument maintenance, and Finnish Microarray and Sequencing Centre. I highly appreciate the work all of you have done.

I would like to extend my gratitude to Sarita Heinonen for the technical assistance with the wet lab work. Marjo Hakkarainen and Päivi Junni are most warmly acknowledged for taking good care of the stem cells; without your daily efforts this work would not have been possible.

I want to thank all the current and former members of the ATLAS group. This group with such a blend of personalities is something special; because of you it has never been boring to come to work. Thank you all for that. It has been a privilege to get to know you all.

I also want to thank all my friends for sharing the life's ups and downs with me. It has been a privilege to get to know so many wonderful people through Ari, too numerous to be mentioned here. Anna, Helena, Miia, and Sanna T., thank you and your families for your friendship, and proving that distance and time cannot separate true friends. Eve, Hanna,

Ninni, Sanna L., and especially Laura, thank you for all your support and interest in both my personal life and thesis related matters. You all mean a lot to me.

I am extremely grateful to my dear family. Kiitos äidille, isälle, Nooralle, Janille ja Mericalle tuesta ja avusta tämän tavoitteen saavuttamisessa. Most importantly, my deepest gratitude goes to my beloved Ari. No words can express what you mean to me.

This work was financially supported by Turku Doctoral Programme of Biomedical Sciences, Finnish Cultural Foundation, Ida Montin Foundation, Emil Aaltonen Foundation, Waldemar von Frenckells Foundation, Finnish-Norwegian Medical Foundation, Paulo Foundation, Finnish Cancer Organizations, Hospital District of Southwest Finland, Systems biology research program, Turku Postgraduate School of Health Sciences, JDRF, The Academy of Finland, and ESTOOLS consortium under the Sixth Research Framework Programme of the European Union.

Turku, May 2013

Nelly Rahkonen

REFERENCES

Adamo, A., Sese, B., Boue, S., Castano, J., Paramonov, I., Barrero, M.J., and Izpisua Belmonte, J.C. (2011). LSD1 regulates the balance between self-renewal and differentiation in human embryonic stem cells. Nat. Cell Biol. *13,* 652-659.

Albertson, D.G., Collins, C., McCormick, F., and Gray, J.W. (2003). Chromosome aberrations in solid tumors. Nat. Genet. *34,* 369-376.

Ambros, V. (2001). microRNAs: tiny regulators with great potential. Cell *107,* 823-826.

Andrews, P.W. (1984). Retinoic acid induces neuronal differentiation of a cloned human embryonal carcinoma cell line in vitro. Dev. Biol. *103,* 285-293.

Armstrong, L., Hughes, O., Yung, S., Hyslop, L., Stewart, R., Wappler, I., Peters, H., Walter, T., Stojkovic, P., Evans, J., Stojkovic, M., and Lako, M. (2006). The role of PI3K/AKT, MAPK/ERK and NFkappabeta signalling in the maintenance of human embryonic stem cell pluripotency and viability highlighted by transcriptional profiling and functional analysis. Hum. Mol. Genet. *15,* 1894-1913.

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.

Avery, S., Zafarana, G., Gokhale, P.J., and Andrews, P.W. (2010). The role of SMAD4 in human embryonic stem cell self-renewal and stem cell fate. Stem Cells *28,* 863-873.

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.

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.

Balagopal, V., and Parker, R. (2009). Polysomes, P bodies and stress granules: states and fates of eukaryotic mRNAs. Curr. Opin. Cell Biol. *21,* 403-408.

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.

Bao, S., Tang, F., Li, X., Hayashi, K., Gillich, A., Lao, K., and Surani, M.A. (2009). Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. Nature *461,* 1292-1295.

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.

Barski, A., Chepelev, I., Liko, D., Cuddapah, S., Fleming, A.B., Birch, J., Cui, K., White, R.J., and Zhao, K. (2010). Pol II and its associated epigenetic marks are present at Pol III-transcribed noncoding RNA genes. Nat. Struct. Mol. Biol. *17,* 629-634.

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.

Bendall, S.C., Stewart, M.H., Menendez, P., George, D., Vijayaragavan, K., Werbowetski-Ogilvie, T., Ramos-Mejia, V., Rouleau, A., Yang, J., Bosse, M., Lajoie, G., and Bhatia, M. (2007). IGF and FGF cooperatively establish the regulatory stem cell niche of pluripotent human cells in vitro. Nature *448,* 1015-1021.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., *et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125,* 315-326.

Berra, E., Benizri, E., Ginouves, A., Volmat, V., Roux, D., and Pouyssegur, J. (2003). HIF prolyl-hydroxylase 2 is the key oxygen sensor setting low steady-state levels of HIF-1alpha in normoxia. EMBO J. *22,* 4082-4090.

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.

Borchert, G.M., Lanier, W., and Davidson, B.L. (2006). RNA polymerase III transcribes human microRNAs. Nat. Struct. Mol. Biol. *13,* 1097-1101.

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.

Brons, I.G., Smithers, L.E., Trotter, M.W., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A., and Vallier, L. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature *448,* 191-195.

Bruck, T., and Benvenisty, N. (2011). Meta-analysis of the heterogeneity of X chromosome inactivation in human pluripotent stem cells. Stem Cell. Res. *6,* 187-193.

Cadigan, K.M., and Nusse, R. (1997). Wnt signaling: a common theme in animal development. Genes Dev. *11,* 3286-3305.

Cai, J., Xie, D., Fan, Z., Chipperfield, H., Marden, J., Wong, W.H., and Zhong, S. (2010). Modeling co-expression across species for complex traits: insights to the difference of human and mouse embryonic stem cells. PLoS Comput. Biol. *6,* e1000707.

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.

Canella, D., Praz, V., Reina, J.H., Cousin, P., and Hernandez, N. (2010). Defining the RNA polymerase III transcriptome: Genome-wide localization of the RNA polymerase III transcription machinery in human cells. Genome Res. *20,* 710-721.

Cao, D., Allan, R.W., Cheng, L., Peng, Y., Guo, C.C., Dahiya, N., Akhi, S., and Li, J. (2011). RNA-binding protein LIN28 is a marker for testicular germ cell tumors. Hum. Pathol. *42,* 710-718.

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.

Chen, A.C., Lee, Y.L., Hou, D.Y., Fong, S.W., Peng, Q., Pang, R.T., Chiu, P.C., Ho, P.C., Lee, K.F., and Yeung, W.S. (2012). Study of transforming growth factor alpha for the maintenance of human embryonic stem cells. Cell Tissue Res. *350,* 289-303.

Chen, Y.G., Li, Z., and Wang, X.F. (2012). Where PI3K/Akt meets Smads: the crosstalk determines human embryonic stem cell fate. Cell. Stem Cell. *10,* 231-232.

Chia, N.Y., Chan, Y.S., Feng, B., Lu, X., Orlov, Y.L., Moreau, D., Kumar, P., Yang, L., Jiang, J., Lau, M.S*., et al.* (2010). A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. Nature *468,* 316-320.

Cho, J., Chang, H., Kwon, S.C., Kim, B., Kim, Y., Choe, J., Ha, M., Kim, Y.K., and Kim, V.N. (2012). LIN28A Is a Suppressor of ER-Associated Translation in Embryonic Stem Cells. Cell *151,* 765-777.

Christophersen, N.S., and Helin, K. (2010). Epigenetic control of embryonic stem cell fate. J. Exp. Med. *207,* 2287-2295.

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.

Dalton, S. (2012). Signaling networks in human pluripotent stem cells. Curr. Opin. Cell Biol.

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.

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.

Dieci, G., Fiorino, G., Castelnuovo, M., Teichmann, M., and Pagano, A. (2007). The expanding RNA polymerase III transcriptome. Trends Genet. *23,* 614-622.

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.

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.

Drets, M.E., and Shaw, M.W. (1971). Specific banding patterns of human chromosomes. Proc. Natl. Acad. Sci. U. S. A. *68,* 2073-2077.

Dumay-Odelot, H., Durrieu-Gaillard, S., Da Silva, D., Roeder, R.G., and Teichmann, M. (2010). Cell growth- and differentiation-dependent regulation of RNA polymerase III transcription. Cell. Cycle *9,* 3687-3699.

Dvash, T., Lavon, N., and Fan, G. (2010). Variations of X chromosome inactivation occur in early passages of female human embryonic stem cells. PLoS One *5,* e11330.

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.

Dvorak, P., and Hampl, A. (2005). Basic fibroblast growth factor and its receptors in human embryonic stem cells. Folia Histochem. Cytobiol. *43,* 203-208.

Eiselleova, L., Matulka, K., Kriz, V., Kunova, M., Schmidtova, Z., Neradil, J., Tichy, B., Dvorakova, D., Pospisilova, S., Hampl, A., and Dvorak, P. (2009). A complex role for FGF-2 in self-renewal, survival, and adhesion of human embryonic stem cells. Stem Cells *27,* 1847-1857.

Enver, T., Soneji, S., Joshi, C., Brown, J., Iborra, F., Orntoft, T., Thykjaer, T., Maltby, E., Smith, K., Abu Dawud, R*., 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.

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.

Fan, G., and Tran, J. (2011). X chromosome inactivation in human and mouse pluripotent stem cells. Hum. Genet. *130,* 217-222.

Felton-Edkins, Z.A., Kenneth, N.S., Brown, T.R., Daly, N.L., Gomez-Roman, N., Grandori, C., Eisenman, R.N., and White, R.J. (2003). Direct regulation of RNA polymerase III transcription by RB, p53 and c-Myc. Cell. Cycle *2,* 181-184.

Feng, C., Neumeister, V., Ma, W., Xu, J., Lu, L., Bordeaux, J., Maihle, N.J., Rimm, D.L., and Huang, Y. (2012). Lin28 regulates HER2 and promotes malignancy through multiple mechanisms. Cell. Cycle *11,* 2486-2494.

Fluckiger, A.C., Marcy, G., Marchand, M., Negre, D., Cosset, F.L., Mitalipov, S., Wolf, D., Savatier, P., and Dehay, C. (2006). Cell cycle features of primate embryonic stem cells. Stem Cells *24,* 547-556.

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.

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.

Furusawa, T., Ohkoshi, K., Honda, C., Takahashi, S., and Tokunaga, T. (2004). Embryonic stem cells expressing both platelet endothelial cell adhesion molecule-1 and stage-specific embryonic antigen-1 differentiate predominantly into epiblast cells in a chimeric embryo. Biol. Reprod. *70,* 1452-1457.

Gallagher, M.F., Flavin, R.J., Elbaruni, S.A., McInerney, J.K., Smyth, P.C., Salley, Y.M., Vencken, S.F., O'Toole, S.A., Laios, A., Lee, M.Y*., et al.* (2009). Regulation of microRNA biosynthesis and expression in 2102Ep embryonal carcinoma stem cells is mirrored in ovarian serous adenocarcinoma patients. J. Ovarian Res. *2,* 19.

Gaspar-Maia, A., Alajem, A., Meshorer, E., and Ramalho-Santos, M. (2011). Open chromatin in pluripotency and reprogramming. Nat. Rev. Mol. Cell Biol. *12,* 36-47.

Golan-Mashiach, M., Dazard, J.E., Gerecht-Nir, S., Amariglio, N., Fisher, T., Jacob-Hirsch, J., Bielorai, B., Osenberg, S., Barad, O., Getz, G*., et al.* (2005). Design principle of gene expression used by human stem cells: implication for pluripotency. FASEB J. *19,* 147-149.

Gonzales, K.A., and Ng, H.H. (2011). Choreographing pluripotency and cell fate with transcription factors. Biochim. Biophys. Acta *1809,* 337-349.

Grati, F.R., Gomes, D.M., Ganesamoorthy, D., Marcato, L., De Toffol, S., Blondeel, E., Malvestiti, F., Loeuillet, L., Ruggeri, A.M., Wainer, R*., et al.* (2012). Application of a new molecular technique for the genetic evaluation of products of conception. Prenat. Diagn.

Greber, B., Lehrach, H., and Adjaye, J. (2007). Fibroblast growth factor 2 modulates transforming growth factor beta signaling in mouse embryonic fibroblasts and human ESCs (hESCs) to support hESC self-renewal. Stem Cells *25,* 455-464.

Greber, B., Wu, G., Bernemann, C., Joo, J.Y., Han, D.W., Ko, K., Tapia, N., Sabour, D., Sterneckert, J., Tesar, P., 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.

Guo, G., Yang, J., Nichols, J., Hall, J.S., Eyres, I., Mansfield, W., and Smith, A. (2009). Klf4 reverts developmentally programmed restriction of ground state pluripotency. Development *136,* 1063-1069.

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.

Han, D.W., Tapia, N., Joo, J.Y., Greber, B., Arauzo-Bravo, M.J., Bernemann, C., Ko, K., Wu, G., Stehling, M., Do, J.T., and Scholer, H.R. (2010). Epiblast stem cell subpopulations represent mouse embryos of distinct pregastrulation stages. Cell *143,* 617-627.

Hanna, J., Cheng, A.W., Saha, K., Kim, J., Lengner, C.J., Soldner, F., Cassady, J.P., Muffat, J., Carey, B.W., and Jaenisch, R. (2010). Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. Proc. Natl. Acad. Sci. U. S. A. *107,* 9222-9227.

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.

Hashimoto, S., Onodera, Y., Hashimoto, A., Tanaka, M., Hamaguchi, M., Yamada, A., and Sabe, H. (2004). Requirement for Arf6 in breast cancer invasive activities. Proc. Natl. Acad. Sci. U. S. A. *101,* 6647-6652.

Haurie, V., Durrieu-Gaillard, S., Dumay-Odelot, H., Da Silva, D., Rey, C., Prochazkova, M., Roeder, R.G., Besser, D., and Teichmann, M. (2010). Two isoforms of human RNA polymerase III with specific functions in cell growth and transformation. Proc. Natl. Acad. Sci. U. S. A. *107,* 4176-4181.

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.

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

Henderson, J.K., Draper, J.S., Baillie, H.S., Fishel, S., Thomson, J.A., Moore, H., and Andrews, P.W. (2002). Preimplantation human embryos and embryonic stem cells show comparable expression of stage-specific embryonic antigens. Stem Cells *20,* 329-337.

Heo, I., Joo, C., Cho, J., Ha, M., Han, J., and Kim, V.N. (2008). Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. Mol. Cell *32,* 276-284.

Hirst, C.E., Ng, E.S., Azzola, L., Voss, A.K., Thomas, T., Stanley, E.G., and Elefanty, A.G. (2006). Transcriptional profiling of mouse and human ES cells identifies SLAIN1, a novel stem cell gene. Dev. Biol. *293,* 90-103.

Hochheimer, A., and Tjian, R. (2003). Diversified transcription initiation complexes expand promoter selectivity and tissue-specific gene expression. Genes Dev. *17,* 1309-1320.

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.

International Stem Cell Initiative Consortium, Akopian, V., Andrews, P.W., Beil, S., Benvenisty, N., Brehm, J., Christie, M., Ford, A., Fox, V., Gokhale, P.J.*, et al.* (2010). Comparison of defined culture systems for feeder cell free propagation of human embryonic stem cells. In Vitro Cell. Dev. Biol. Anim. *46,* 247-258.

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.

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 iPSC generation, but it is dispensable for the maintenance and induction of pluripotency. PLoS One *6,* e20461.

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.

Jiang, J., and Ng, H.H. (2008). TGFbeta and SMADs talk to NANOG in human embryonic stem cells. Cell. Stem Cell. *3,* 127-128.

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.

Jouneau, A., Ciaudo, C., Sismeiro, O., Brochard, V., Jouneau, L., Vandormael-Pournin, S., Coppee, J.Y., Zhou, Q., Heard, E., Antoniewski, C., and Cohen-Tannoudji, M. (2012). Naive and primed murine pluripotent stem cells have distinct miRNA expression profiles. RNA *18,* 253-264.

Kiiski, K., Roovere, T., Zordania, R., von Koskull, H., and Horelli-Kuitunen, N. (2012). Prenatal diagnosis of 17p13.1p13.3 duplication. Case Report Med. *2012,* 840538.

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.

Kunarso, G., Chia, N.Y., Jeyakani, J., Hwang, C., Lu, X., Chan, Y.S., Ng, H.H., and Bourque, G. (2010). Transposable elements have rewired the core regulatory network of human embryonic stem cells. Nat. Genet. *42,* 631-634.

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., 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.

Leahy, D.J. (2004). Structure and function of the epidermal growth factor (EGF/ErbB) family of receptors. Adv. Protein Chem. *68,* 1-27.

Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H., and Kim, V.N. (2004). MicroRNA genes are transcribed by RNA polymerase II. EMBO J. *23,* 4051-4060.

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.

Li, J., Wang, G., Wang, C., Zhao, Y., Zhang, H., Tan, Z., Song, Z., Ding, M., and Deng, H. (2007). MEK/ERK signaling contributes to the maintenance of human embryonic stem cell self-renewal. Differentiation *75,* 299-307.

Li, P., Tong, C., Mehrian-Shai, R., Jia, L., Wu, N., Yan, Y., Maxson, R.E., Schulze, E.N., Song, H., Hsieh, C.L., Pera, M.F., and Ying, Q.L. (2008). Germline competent embryonic stem cells derived from rat blastocysts. Cell *135,* 1299-1310.

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.

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.

Ludwig, T.E., Bergendahl, V., Levenstein, M.E., Yu, J., Probasco, M.D., and Thomson, J.A. (2006). Feeder-independent culture of human embryonic stem cells. Nat. Methods *3,* 637-646.

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.

Lund, R.J., Narva, E., and Lahesmaa, R. (2012). Genetic and epigenetic stability of human pluripotent stem cells. Nat. Rev. Genet. *13,* 732-744.

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.

Martignetti, J.A., and Brosius, J. (1995). BC1 RNA: transcriptional analysis of a neural cell-specific RNA polymerase III transcript. Mol. Cell. Biol. *15,* 1642-1650.

Martignetti, J.A., and Brosius, J. (1993). BC200 RNA: a neural RNA polymerase III product encoded by a monomeric Alu element. Proc. Natl. Acad. Sci. U. S. A. *90,* 11563-11567.

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.

Martin, S.L. (2006). The ORF1 protein encoded by LINE-1: structure and function during L1 retrotransposition. J. Biomed. Biotechnol. *2006,* 45621.

Martin, S.L. (1991). Ribonucleoprotein particles with LINE-1 RNA in mouse embryonal carcinoma cells. Mol. Cell. Biol. *11,* 4804-4807.

Massague, J. (2000). How cells read TGF-beta signals. Nat. Rev. Mol. Cell Biol. *1,* 169-178.

Maston, G.A., Landt, S.G., Snyder, M., and Green, M.R. (2012). Characterization of enhancer function from genome-wide analyses. Annu. Rev. Genomics Hum. Genet. *13,* 29-57.

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.

McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. *40,* 4288-4297.

Meshorer, E., Yellajoshula, D., George, E., Scambler, P.J., Brown, D.T., and Misteli, T. (2006). Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. Dev. Cell. *10,* 105-116.

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.

Miyazaki, S., Yamamoto, H., Miyoshi, N., Takahashi, H., Suzuki, Y., Haraguchi, N., Ishii, H., Doki, Y., and Mori, M. (2012). Emerging methods for preparing iPS cells. Jpn. J. Clin. Oncol. *42,* 773-779.

Moqtaderi, Z., Wang, J., Raha, D., White, R.J., Snyder, M., Weng, Z., and Struhl, K. (2010). Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. Nat. Struct. Mol. Biol. *17,* 635-640.

Mori, S., Tanaka, M., Nanba, D., Nishiwaki, E., Ishiguro, H., Higashiyama, S., and Matsuura, N. (2003). PACSIN3 binds ADAM12/meltrin alpha and up-regulates ectodomain shedding of heparin-binding epidermal growth factor-like growth factor. J. Biol. Chem. *278,* 46029-46034.

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.

Moss, E.G., Lee, R.C., and Ambros, V. (1997). The cold shock domain protein LIN-28 controls developmental timing in C. elegans and is regulated by the lin-4 RNA. Cell *88,* 637-646.

Na, J., Furue, M.K., and Andrews, P.W. (2010). Inhibition of ERK1/2 prevents neural and mesendodermal differentiation and promotes human embryonic stem cell self-renewal. Stem Cell. Res. *5,* 157-169.

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. *28,* 371-377.

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

Nichols, J., and Smith, A. (2012). Pluripotency in the embryo and in culture. Cold Spring Harb Perspect. Biol. *4,* a008128.

Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. Cell. Stem Cell. *4,* 487-492.

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.

Okita, K., and Yamanaka, S. (2011). Induced pluripotent stem cells: opportunities and challenges. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *366,* 2198-2207.

Oler, A.J., Alla, R.K., Roberts, D.N., Wong, A., Hollenhorst, P.C., Chandler, K.J., Cassiday, P.A., Nelson, C.A., Hagedorn, C.H., Graves, B.J., and Cairns, B.R. (2010). Human RNA polymerase III transcriptomes and relationships to Pol II promoter chromatin and enhancer-binding factors. Nat. Struct. Mol. Biol. *17,* 620-628.

Ozair, M.Z., Noggle, S., Warmflash, A., Krzyspiak, J.E., and Brivanlou, A.H. (2012). SMAD7 Directly Converts Human Embryonic Stem Cells to Telencephalic Fate by a Default Mechanism. Stem Cells

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.

Patil, S.R., Merrick, S., and Lubs, H.A. (1971). Identification of each human chromosome with a modified Giemsa stain. Science *173,* 821-822.

Paxton, C.N., Brothman, A.R., and Geiersbach, K.B. (2012). Rapid aneusomy detection in products of conception using the KaryoLite BACs-on-Beads assay. Prenat. Diagn. 1-7.

Paye, J.M., and Forsten-Williams, K. (2006). Regulation of insulin-like growth factor-I (IGF-I) delivery by IGF binding proteins and receptors. Ann. Biomed. Eng. *34,* 618-632.

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.

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.

Raha, D., Wang, Z., Moqtaderi, Z., Wu, L., Zhong, G., Gerstein, M., Struhl, K., and Snyder, M. (2010). Close association of RNA polymerase II and many transcription factors with Pol III genes. Proc. Natl. Acad. Sci. U. S. A. *107,* 3639-3644.

Ramskold, D., Wang, E.T., Burge, C.B., and Sandberg, R. (2009). An abundance of ubiquitously expressed genes revealed by tissue transcriptome sequence data. PLoS Comput. Biol. *5,* e1000598.

Roeder, R.G., and Rutter, W.J. (1970). Specific nucleolar and nucleoplasmic RNA polymerases. Proc. Natl. Acad. Sci. U. S. A. *65,* 675-682.

Roeder, R.G., and Rutter, W.J. (1969). Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. Nature *224,* 234-237.

Roode, M., Blair, K., Snell, P., Elder, K., Marchant, S., Smith, A., and Nichols, J. (2012). Human hypoblast formation is not dependent on FGF signalling. Dev. Biol. *361,* 358-363.

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.

Scott, G. (2006). Developmental Biology Sinauer Associates Inc.).

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.

Singh, A.M., Bechard, M., Smith, K., and Dalton, S. (2012). Reconciling the different roles of Gsk3beta in "naive" and "primed" pluripotent stem cells. Cell. Cycle *11,* 2991-2996.

Singh, A.M., Reynolds, D., Cliff, T., Ohtsuka, S., Mattheyses, A.L., Sun, Y., Menendez, L., Kulik, M., and Dalton, S. (2012). Signaling network crosstalk in human pluripotent cells: a Smad2/3-regulated switch

that controls the balance between self-renewal and differentiation. Cell. Stem Cell. *10,* 312-326.

Skottman, H., Mikkola, M., Lundin, K., Olsson, C., Stromberg, A.M., Tuuri, T., Otonkoski, T., Hovatta, O., and Lahesmaa, R. (2005). Gene expression signatures of seven individual human embryonic stem cell lines. Stem Cells *23,* 1343-1356.

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.

Stewart, M.H., Bosse, M., Chadwick, K., Menendez, P., Bendall, S.C., and Bhatia, M. (2006). Clonal isolation of hESCs reveals heterogeneity within the pluripotent stem cell compartment. Nat. Methods *3,* 807-815.

Sudheer, S., Bhushan, R., Fauler, B., Lehrach, H., and Adjaye, J. (2012). FGF Inhibition Directs BMP4-Mediated Differentiation of Human Embryonic Stem Cells to Syncytiotrophoblast. Stem Cells Dev. *21,* 2987-3000.

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.

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.

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.

Teichmann, M., Dieci, G., Pascali, C., and Boldina, G. (2010). General transcription factors and subunits of RNA polymerase III: Paralogs for promoter- and cell type-specific transcription in multicellular eukaryotes. Transcription *1,* 130-135.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature *448,* 196-199.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282,* 1145-1147.

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

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25,* 1105-1111.

Tsuneyoshi, N., Sumi, T., Onda, H., Nojima, H., Nakatsuji, N., and Suemori, H. (2008). PRDM14 suppresses expression of differentiation marker genes in human embryonic stem cells. Biochem. Biophys. Res. Commun. *367,* 899-905.

Tsutsui, H., Valamehr, B., Hindoyan, A., Qiao, R., Ding, X., Guo, S., Witte, O.N., Liu, X., Ho, C.M., and Wu, H. (2011). An optimized small molecule inhibitor cocktail supports long-term maintenance of human embryonic stem cells. Nat. Commun. *2,* 167.

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.

Van der Jeught, M., O'Leary, T., Ghimire, S., Lierman, S., Duggal, G., Versieren, K., Deforce, D., Chuva de Sousa Lopes, S., Heindryckx, B., and De Sutter, P. (2012). The Combination of Inhibitors of FGF/MEK/Erk and GSK3beta Signaling Increases the Number of OCT3/4- and NANOG-Positive Cells in the Human Inner Cell Mass, But Does Not Improve Stem Cell Derivation. Stem Cells Dev.

Van Hoof, D., Passier, R., Ward-Van Oostwaard, D., Pinkse, M.W., Heck, A.J., Mummery, C.L., and Krijgsveld, J. (2006). A quest for human and mouse embryonic stem cell-specific proteins. Mol. Cell. Proteomics *5,* 1261-1273.

Vialard, F., Simoni, G., Aboura, A., De Toffol, S., Molina Gomes, D., Marcato, L., Serero, S., Clement, P., Bouhanna, P., Rouleau, E*., et al.* (2011). Prenatal BACs-on-Beads : a new technology for rapid detection of aneuploidies and microdeletions in prenatal diagnosis. Prenat. Diagn. *31,* 500-508.

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

Viswanathan, S.R., Powers, J.T., Einhorn, W., Hoshida, Y., Ng, T.L., Toffanin, S., O'Sullivan, M., Lu, J., Phillips, L.A., Lockhart, V.L*., et al.* (2009). Lin28 promotes transformation and is associated with advanced human malignancies. Nat. Genet. *41,* 843-848.

Voskas, D., Ling, L.S., and Woodgett, J.R. (2010). Does GSK-3 provide a shortcut for PI3K activation of Wnt signalling? F1000 Biol. Rep. *2,* 82.

Wang, L., Schulz, T.C., Sherrer, E.S., Dauphin, D.S., Shin, S., Nelson, A.M., Ware, C.B., Zhan, M., Song, C.Z., Chen, X.*, et al.* (2007). Self-renewal of human embryonic stem cells requires insulin-like growth factor-1 receptor and ERBB2 receptor signaling. Blood *110,* 4111-4119.

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.

Wang, Z., and Roeder, R.G. (1997). Three human RNA polymerase III-specific subunits form a subcomplex with a selective function in specific transcription initiation. Genes Dev. *11,* 1315-1326.

Watanabe, S., Umehara, H., Murayama, K., Okabe, M., Kimura, T., and Nakano, T. (2006). Activation of Akt signaling is sufficient to maintain pluripotency in mouse and primate embryonic stem cells. Oncogene *25,* 2697-2707.

White, R.J. (2011). Transcription by RNA polymerase III: more complex than we thought. Nat. Rev. Genet. *12,* 459-463.

White, R.J. (2008). RNA polymerases I and III, non-coding RNAs and cancer. Trends Genet. *24,* 622-629.

White, R.J., Stott, D., and Rigby, P.W. (1989). Regulation of RNA polymerase III transcription in response to F9 embryonal carcinoma stem cell differentiation. Cell *59,* 1081-1092.

Wilbert, M.L., Huelga, S.C., Kapeli, K., Stark, T.J., Liang, T.Y., Chen, S.X., Yan, B.Y., Nathanson, J.L., Hutt, K.R., Lovci, M.T.*, et al.* (2012). LIN28 Binds Messenger RNAs at GGAGA Motifs and Regulates Splicing Factor Abundance. Mol. Cell *48,* 195-206.

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.

Wong, R.C., Pollan, S., Fong, H., Ibrahim, A., Smith, E.L., Ho, M., Laslett, A.L., and Donovan, P.J. (2011). A novel role for an RNA polymerase III subunit POLR3G in regulating pluripotency in human embryonic stem cells. Stem Cells *29,* 1517-1527.

Wright, A.J., and Andrews, P.W. (2009). Surface marker antigens in the characterization of human embryonic stem cells. Stem Cell. Res. *3,* 3-11.

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.

Xu, B., Zhang, K., and Huang, Y. (2009). Lin28 modulates cell growth and associates with a subset of cell cycle regulator mRNAs in mouse embryonic stem cells. RNA *15,* 357-361.

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.

Xue, D., Peng, Y., Wang, F., Allan, R.W., and Cao, D. (2011). RNA-binding protein LIN28 is a sensitive marker of ovarian primitive germ cell tumours. Histopathology *59,* 452-459.

Xue, K., Ng, J.H., and Ng, H.H. (2011). Mapping the networks for pluripotency. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *366,* 2238-2246.

Yamanaka, S. (2012). Induced pluripotent stem cells: past, present, and future. Cell. Stem Cell. *10,* 678-684.

Yamanaka, S., and Blau, H.M. (2010). Nuclear reprogramming to a pluripotent state by three approaches. Nature *465,* 704-712.

Yang, C., Przyborski, S., Cooke, M.J., Zhang, X., Stewart, R., Anyfantis, G., Atkinson, S.P., Saretzki, G., Armstrong, L., and Lako, M. (2008). A key role for telomerase reverse transcriptase unit in modulating human embryonic stem cell proliferation, cell cycle dynamics, and in vitro differentiation. Stem Cells *26,* 850-863.

Yeo, S., Jeong, S., Kim, J., Han, J.S., Han, Y.M., and Kang, Y.K. (2007). Characterization of DNA methylation change in stem cell marker genes during differentiation of human embryonic stem cells. Biochem. Biophys. Res. Commun. *359,* 536-542.

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

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

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.

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, X., Yalcin, S., Lee, D.F., Yeh, T.Y., Lee, S.M., Su, J., Mungamuri, S.K., Rimmele, P., Kennedy, M., Sellers, R.*, et al.* (2011). FOXO1 is an essential regulator of pluripotency in human embryonic stem cells. Nat. Cell Biol. *13,* 1092-1099.

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.