FACTORS CONTROLLING THE PLURIPOTENT STATE OF HUMAN EMBRYONIC STEM CELLS

Maria del Pilar Vazquez Arango

St. Peter's College

University of Oxford



A thesis submitted for the degree of Doctor of Philosophy at the

University of Oxford, Michaelmas Term 2012

Abstract

FACTORS CONTROLLING THE PLURIPOTENT STATE OF HUMAN EMBRYONIC STEM CELLS

Maria del Pilar Vazquez Arango, St. Peter's College

A thesis submitted for the degree of Doctor of Philosophy at the University of Oxford, Michaelmas Term 2012

In the field of regenerative medicine, human embryonic stem (hES) cells offer the potential to treat degenerative diseases and replace damaged or non-functional tissue. However, it is not fully clear how hES cells retain a pluripotent state or differentiate into specific cell lineages. Although most of previous research has found that various proteins are responsible for the balance in hES cell maintenance and differentiation, there is mounting evidence for non-coding RNAs as important players in this task. Data from our laboratory shows that there are genes that encode for variants of the non-coding RNA U1 small nuclear (sn)RNA, a key component of the spliceosome. These variant (v)U1 snRNA genes are transcriptionally active and differentially expressed. Moreover, interfering with the activity of a specific vU1 snRNA affects the expression of a subset of genes at the level of pre-mRNA 3' end processing (O'Reilly et al., 2012). I have analysed the expression of vU1 snRNA genes throughout hES cell differentiation and shown that vU1 snRNA genes are down regulated upon differentiation into macrophages. Interestingly, analyses of steady state levels of specific vU1 snRNAs throughout hES cell differentiation revealed a characteristic pattern where specific vU1 snRNAs are more stable in the final differentiation step (i.e. macrophages). Thus, strongly supporting the idea that the vU1 snRNA genes are being regulated throughout differentiation both at the level of transcription and snRNA stability. Furthermore, vU1 snRNAs are up regulated upon reprogramming of primary human skin fibroblasts into induced pluripotent stem (iPS) cells, whose pattern of expression is similar to hES cells. Therefore, we hypothesize that vU1 snRNAs play a significant role in establishing pluripotent stem cells and that their differential expression may be key during development and cell reprogramming.

"Wisdom lies neither in fixity nor in change, but in the dialectic between the two."

Octavio Paz

I dedicate this thesis to the memory of my grandparents:

Dedico esta tesis a la memoria de mis abuelos:

Modesta, Miguel, Benjamin & Amor

Acknowledgements

I would first like to deeply thank my supervisor Dr Shona Murphy for welcoming me in her lab in the middle of my DPhil studies. I am equally grateful to Dr Dawn O'Reilly, from Dr Murphy's lab, whose constant support and guidance throughout this project have been invaluable. I feel very fortunate to have had the opportunity to be mentored by them.

I am very grateful to Prof Stephen Kennedy, Head of the Nuffield Department of Obstetrics and Gynaecology and Prof Ian Sargent, Director of Graduate Studies, and Dr Paul Fairchild, Co-Director of the Oxford Stem Cell Institute, for all their support during my DPhil studies.

During these four years I was funded by the Medical Research Council and the Nuffield Department of Obstetrics and Gynaecology, which made the work presented here possible.

For their support in the beginning of my DPhil studies, I would like to thank my former supervisor Prof Helen Mardon and the past members of her lab, especially Dr Frances Brook and Janet Carver, who not only kindly shared precious knowledge but also a sincere friendship that developed over many morning coffees. Thanks to Dr Chandramouli Chillakuri, Dr Zhaohui Li, Dr Javier Periz and Dr Youn-Jung Kang for being there to offer advice and excellent technical expertise (especially with my endless Western blots), and to Dr Rebecca Dragovic for assistance with flow cytometry.

I am also very grateful to Prof William James for allowing me to collaborate with his lab, most of this DPhil project is the result of this fruitful collaboration. Especially, I want to thank Dr Sally Cowley, Cathy Browne and Jane Vowles, at the James Martin Stem Cell Facility, for allowing me to work with them and for always willing to share their knowledge and provide advice. I would also like to thank Dr Kenny Moore for advice and for providing some of the plasmids used in Chapter 5.

Thanks to the current and past members of the Murphy lab, who patiently helped this cell biologist entering the field of molecular biology. Thank you to Alice Taylor

and Dr Martin Dienstbier for sharing their expertise. I feel very fortunate to have encountered in my work environment friends like Justyna, whose strength, dedication and selflessness have become an example to follow; Hadeel, *habibti*, never failing to make me smile and always so exceptionally caring; and Clélia, who gets me as if we knew each other for much longer. Thanks *ángeles*, you have made every day so unique.

I want to sincerely thank special friends who have made my Oxford experience one not to forget: thank you Antoniya for your peculiar sincerity, Surhi for those nice long chats that brushed up our Spanish, Muna for your lively personality, Leila for your unique sense of humour, Sarah for your lovely smile and contagious energy and, Walaa, whose continuous friendship since the beginning of this journey has not failed to teach and surprise me every day, *muchos shukran linda*. I also want to thank those who have been there long before: Paola, Georgia and Elka, for not letting time and space get on the way of our friendship.

A very special thanks to my best friend Elin, who has been by my side for over ten years, thanks for letting me in your life and for sharing so many adventures, *gracias superwoman-A*.

Ευχαριστώ πολύ Lefteri *μου*, I was not expecting you but now you have become my special person.

Last but not least I want to thank my family: *Mi Mamá*, *mi Papá,* Maye and Migue, I do not have enough words to thank you, I am here because of you, thank you for never giving up on me and for teaching me the true meaning of love and friendship. *Los quiero mucho*.

Table of contents

Abstract i

Dedication iii

Acknowledgments iv

List of Figures x

List of Tables xii

Abbreviations xiii

Chapter 1| INTRODUCTION 1

1.1 Stem cells 2

1.2 Pluripotent stem cells 5

1.2.1 Embryonal carcinoma (EC) cells 5

1.2.2 Embryonic stem (ES) cells 5

1.2.3 Embryonic germ (EG) cells 8

1.2.4 Induced pluripotent stem (iPS) cells 9

1.3 Human embryonic stem (hES) cells 10

1.4 Extrinsic regulators of pluripotency in hES cells 11

1.5 Intrinsic regulators of pluripotency in hES cells 18

1.5.1 The core transcription factors of pluripotency 18

1.5.2 Circuitry established by OCT4, NANOG and SOX2 23

1.6 An additional level of regulation of pluripotency in hES cells 25

1.6.1 U1 and mRNA processing 26

1.6.2 The role of alternative splicing in embryonic development and hES cells 28

1.6.3 Splicing-independent roles of U1 30

1.6.4 Variant U1 snRNAs 32

1.7 Project aim and description 37

Chapter 2| MATERIALS AND METHODS 39

2.1 Cell culture 40

2.1.1 Feeder cells (MEFs, CCD and Hs27) 40

2.1.2 Human ES cells (H1, HUES1-4 and OxF1) 41

2.1.3 Cell freezing and thawing 42

2.1.4 Embryoid body preparation 43

2.1.5 Production of monocytes and macrophages from hES cell line HUES2 44

2.1.6 HeLa cells 45

2.1.7 Bacterial cells 45
2.2 Preparation of ECM substrates from feeders 45
2.3 Western blotting 46
2.3.1 Preparation of lysates 46
2.3.2 SDS-polyacrylamide gel electrophoresis (SDS-PAGE) 47
2.3.3 Wet gel transfer (for high molecular weight proteins) 47
2.3.4 Semi-dry gel transfer (for low molecular weight proteins) 48
2.3.5 Antibody (1^{o} and 2^{o}) incubation 48
2.3.6 Protein detection 49
2.3.7 Stripping and re-probing membranes 49
2.4 Immunocytochemistry 50
2.5 Flow cytometry analyses 50
2.6 Transfection of HeLa and HUES2 cells 51
2.6.1 Plasmids 51
2.6.2 2'-O-methyl phosphorothioate oligonucleotides 52
2.7 RNA isolation 52
2.7.1 Preparation of cells 52
2.7.2 RNA extraction 53
2.7.3 RNA re-precipitation 53
2.8 First-Strand complementary (c)DNA synthesis 54
2.9 Real-time quantitative polymerase chain reaction (qPCR) 54
2.9.1 Primers used 56
2.9.2 Statistical analyses 57
2.10 Molecular cloning 57
2.10.1 Polymerase Chain Reaction (PCR) 57
2.10.2 DNA digestion with restriction enzymes 58
2.10.3 DNA phenol/chloroform extraction and EtOH precipitation 59
2.10.4 DNA gel electrophoresis 59
2.10.5 DNA gel extraction 59
2.10.6 DNA ligation 60
2.10.7 Transformation of competent cells with plasmid DNA 60
2.10.8 Mini-prep 60
2.10.9 Maxi-prep 61
2.10.10 DNA sequencing 62

Chapter 3| ESTABLISHING FEEDER-FREE CULTURE FOR HUMAN EMBRYONIC STEM CELLS 63

3.1 Introduction 64

3.2 Different feeders support hES culture differently 70

3.3 The composition of ECM proteins varies among feeders 72

3.4 Conditions for the preparation of ECM proteins from feeders 77

3.5 ECM proteins derived from feeders support hES cell culture 80

3.6 Discussion 86

Chapter 4| CHARACTERIZING vU1 snRNA GENE EXPRESSION THROUGHOUT hES CELL DIFFERENTIATION 91

4.1 Introduction 92

4.2 Matrigel™ is not contaminated with vU1 snRNA 94

4.3 Pluripotent stem cells express similar levels of vU1 snRNAs 97

4.4 Expression of vU1 snRNA genes is reduced upon differentiation 100

4.5 Discussion 111

Chapter 5| ROLE OF vU1 snRNA GENES IN HUMAN PLURIPOTENT STEM CELLS 115

5.1 Introduction 116

5.2 Variant U1 snRNA genes are up regulated in iPS cells 117

5.3 Transfection efficiency in hES cells 121

5.4 Down regulation of vU1 snRNA genes in hES cells 124

5.5 Design and validation of a Tetracycline-inducible system for the expression of vU1 snRNA genes 132

5.6 Discussion 141

Chapter 6| GENERAL DISCUSSION 143

6.1 Feeder-free hES cell culture system supports hES cells 145

6.2 Variant U1 snRNA genes are differentially expressed during hES cell differentiation 146

6.3 Established tools to manipulate vU1 snRNA gene expression 149

6.4 Future directions 151

6.4.1 Proteomic analyses of feeder-derived ECM substrate and feeder CM 151

6.4.2 Assessment of vU1 snRNA gene expression in different cell types 151

6.4.3 Correlation vU1 snRNA expression to specific gene targets in hES cells 152

6.4.4 Identification of the regulatory elements that control vU1 snRNA gene expression in hES cells..... 153
6.4.5 Characterization of the snRNP complexes associated with vU1 snRNAs..... 154
6.4.6 Manipulation of vU1 snRNA levels during iPS cell reprogramming 155
6.5 Summary 155
REFERENCES 157

List of Figures

Figure 1-1. Schematic representation of mammalian embryonic development..... 4
Figure 1-2. Schematic diagram summarizing the signalling cascades activated by extrinsic factors in hES cells..... 17
Figure 1-3. Diagrammatic representation of the regulatory circuitry established by the core transcription factors of hES cell pluripotency. 24
Figure 1-4. RNA-encoding sequence of vU1 snRNA genes. 35
Figure 3-1. Human ES cells grown on two different mitotically inactivated mouse or human feeders. 71
Figure 3-2. Detection of ECM proteins in NH4OH-treated ECM preparations from feeders..... 75
Figure 3-3. Expression of ECM proteins in MEFs and CCD treated with NH_4OH, DOC and CDB. 79
Figure 3-4. Human ES cells OxF1 grown on ECM plus CM from fully supporting and non-supporting feeders. 82
Figure 3-5. Expression of pluripotency markers in hES cells OxF1 grown on MEF ECM plus MEF CM. 83
Figure 3-6. Expression of pluripotency markers on hES cells OxF1 grown on ECM and CM from supporting and non-supporting feeders. 84
Figure 3-7. Human ES cells grown on the presence or absence of ECM/CM from MEFs. 85
Figure 4-1. Testing the levels of vU1 snRNAs in Matrigel^{TM}. 96
Figure 4-2. Pre-vU1 snRNA levels in pluripotent stem cells: Human ES cells HUES1, HUES2 and HUES4. 99
Figure 4-3. Diagrammatic representation of a protocol designed by Karlsson et al (2008) to differentiate hES cell into macrophages. 101
Figure 4-4. Levels of cell specific markers in HUES2, EBs, monocytes and macrophages. 102
Figure 4-5. Pre-vU1 snRNA levels in HUES2, EBs, monocytes and macrophages. 104
Figure 4-6. Mature-vU1 snRNA levels in HUES2, EBs, monocytes and macrophages. 106

Figure 4-7. Comparison of the levels of pre- and m-vU1 and U1 snRNAs in HUES2, EBs, monocytes and macrophages. 110
Figure 5-1. Mature (m)-vU1 snRNA levels in primary human skin fibroblasts and iPS cells. 119
Figure 5-2. Mature (m)-vU1 snRNA levels in iPS cells and HUES2. 120
Figure 5-3. Testing hES cell transfection efficiency with vU1 snRNA expressing constructs. 123
Figure 5-4. Optimizing vU1.8 snRNA knockdown in hES cells..... 127
Figure 5-5. Knockdown of vU1.8 and vU1.20 snRNAs in hES cells. 128
Figure 5-6. Testing antisense (as)_vU1 snRNA expressing construct for the knockdown of vU1.8 snRNA. 131
Figure 5-7. Tetracycline repressor-regulated vU1 snRNA gene expression system. 137
Figure 5-8. Plasmid vector containing the tetracycline repressor-regulated vU1 snRNA gene expression system. 138
Figure 5-9. Validation of TetR/Puro construct. 139
Figure 5-10. Validation of pGEM4/tetO-vU1 construct. 140

List of Tables

Table 3-1. Human ES cell culture systems that use feeder layers. 66

Table 3-2. Human ES cell culture on ECM protein substrates. 69

Table 3-3. Expression of ECM proteins in NH_4OH-treated ECM preparations from feeders..... 76

Abbreviations

AKT	protein kinase B
BMP	Bone Morphogenetic Protein
bp	base pairs
cDNA	complementary DNA
CM	conditioned media
DMEM	Dulbecco's Modified Eagle Medium
DMSO	dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DOC	sodium deoxycholate
DSE	distal sequence element
EBs	embryoid bodies
EC	embryonal carcinoma
ECM	extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EG	embryonic germ
EGF	Epidermal Growth Factor
ERK	extracellular signal-regulated kinase
ES	embryonic stem
FCS	foetal calf serum
FGF	Fibroblast Growth Factor

gDNA	genomic DNA
GFP	green fluorescent protein
GSK	glycogen synthase kinase
h	hour
hES	human ES
HFFs	human foreskin fibroblasts
ICM	inner cell mass
IGF	Insulin-like Growth Factor
IL	interleukin
iPS	induced pluripotent stem
IRES	internal ribosome entry site
KSR	knock-out serum replacement
LIF	Leukemia Inhibitory factor
LIFR	LIF receptor
lncRNA	long ncRNA
MAPK	mitogen-activated protein kinase
MEFs	mouse embryonic fibroblasts
MEKs	MAPK/ERK kinases
mES	mouse ES
miRNA	microRNA
NaCl	sodium chloride
ncRNA	non-coding RNA

OCT	Octamer
PAP	Poly (ADP-ribose) polymerase I
PBS	phosphate buffered saline
PCPA	premature cleavage and polyadenylation
PFA	paraformaldehyde
PI3-K	phosphatidylinositol -3 kinase
Pol II	RNA Polymerase II
PPT	polypyrimidine tract
PS	primitive streak
PSE	proximal sequence element
qPCR	quantitative real time polymerase chain reaction
RNA	Ribonucleic acid
RRM	RNA recognition motif
SCNT	Somatic Cell Nuclear Transfer
SDS	sodium dodecyl sulphate
SF	splicing factor
SMAD	mothers against decapentaplegic homolog
snRNA	small nuclear RNA
Sox	SRY-box
snRNP	small nuclear ribonucleoproteins
SRY	Sex determining region Y
SSEA	Stage Specific Embryonic Antigen

STAT	Signal Transducer and Activator of Transcription
TAF	transcription binding protein-associated factor
Tet	tetracycline
tetO	tet operator
TF	transcription factor
TGF	Transforming Growth Factor
Tra	Tumor rejection antigen
U	urydine-rich
vU1	variant U1
WNT	wingless-Int

Chapter 1| INTRODUCTION

1.1 Stem cells

A stem cell (a term coined by the scientist Alexander Maximow in the early 1900's) is defined as a cell that has both the capacity to give rise to another cell of its kind and to at least one cell that has a specialized function in a given tissue (i.e. differentiated cell). These two characteristics are referred to as self-renewal and differentiation potential or potency, respectively; and together are what makes stem cells hold a great potential for clinical applications.

Stem cells were first characterized in the context of bone marrow transplantation in mice and humans (Becker et al., 1963, Gatti et al., 1968) and of germ cell tumours, where malignant stem cells were identified (Kleinsmith and Pierce, 1964). Thereafter, stem cells known as haematopoietic stem cells were found in human cord blood (Prindull et al., 1978).

It was then attempted to isolate cells from the embryo, given it has the potential to give rise to all the differentiated cell types present in the adult organism. In mammals, they were obtained from the blastocyst stage (at approximately 4.5 days following fertilization and before implantation); this consists of a hollow sphere delineated by trophoblast cells and a compact population of cells known as the inner cell mass (ICM). It was the isolation of the ICM for *in vitro* culture that led to the generation of embryonic stem (ES) cells. During normal mammalian embryonic development (*in vivo*), the ICM will generate a bottom layer that forms the hypoblast or primitive endoderm (part of the extra embryonic tissues) and a top layer or epiblast, which will give rise to part of the extra embryonic tissues (extra embryonic ectoderm) and to all the tissues of the embryo

proper. During gastrulation, the blastocyst undergoes a series of cell movements that ultimately lead to the formation of three primary germ layers known as ectoderm, mesoderm and endoderm; all the somatic tissues of the embryo originate from these three layers. Meanwhile, the germ line arises from a population of cells at the posterior part of the epiblast, which are referred to as primordial germ cells (PGCs) and will later migrate towards the genital ridges. The *in vitro* culture of these cells will eventually give rise to embryonic germ cells (Fig. 1-1).

Given the potential of ES cells, it was soon investigated whether stem cells existed in other differentiated tissues besides the bone marrow. Stem cells have been found in the hair follicle, the intestinal crypts, neural tissues, adipose tissue and dental pulp (Cotsarelis et al., 1990, Reynolds and Weiss, 1992, Gronthos et al., 2000, Zuk et al., 2001, Potten et al., 2002).

However, stem cells found in adult somatic tissues are restricted in their differentiation potential. While ES cells have the potential to form every differentiated cell type found in the adult organism (pluripotent), adult stem cells only have the potential to give rise to a single cell type (unipotent) or to all cells derived from a single primary germ layer (multipotent). Consequently, the clinical applications of adult stem cells are limited.

More recently, another type of pluripotent stem cell has been characterized and is known as induced pluripotent stem (iPS) cell, which is generated *in vitro* from adult somatic cells; therefore, bypassing the use of embryos (Takahashi and Yamanaka, 2006, Takahashi et al., 2007).

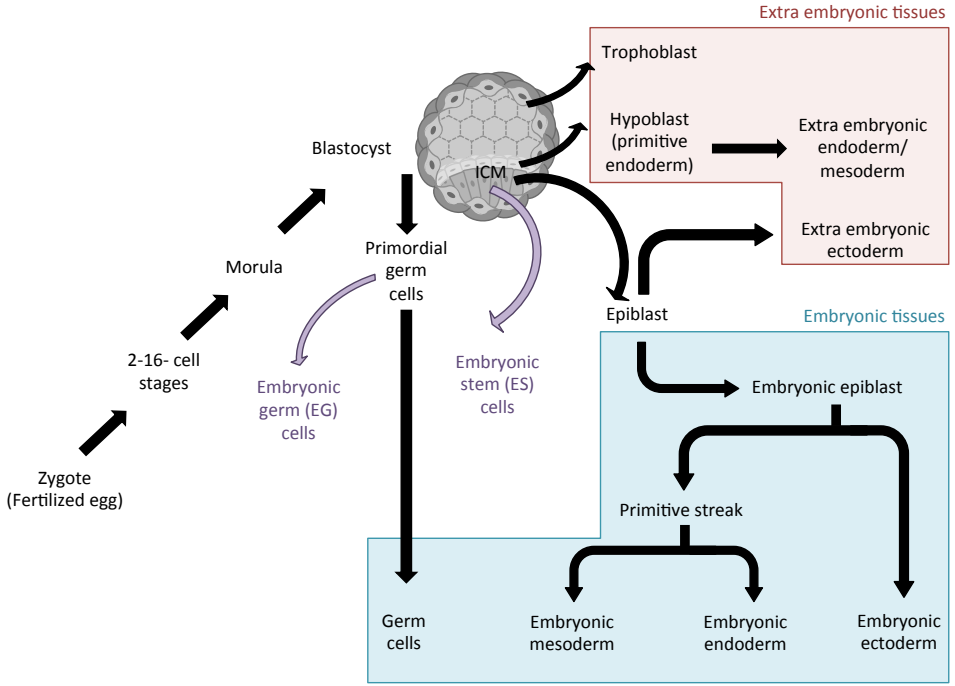


Figure 1-1. Schematic representation of mammalian embryonic development.

A fertilized egg (zygote) begins to undergo cell divisions and when it generates ~16 cells (blastomeres) it becomes a morula, formed by a group of internal cells (the inner cell mass, ICM) and external cells (trophoblast). Subsequently, the blastomeres of the morula pump fluid inside and thus create a blastocoel cavity and give rise to the blastocyst. This consists of a hollow sphere delineated by trophoblast cells and a discrete ICM in the inside. The ICM will then separate into two layers, the hypoblast or primitive endoderm, which contributes to part of the extra embryonic tissues (pink box) and the epiblast, which will give rise to the extra embryonic ectoderm and to all the tissues of the embryo proper (blue box). The germ line originates from a population at the posterior part of the epiblast, which later on migrates towards the genital ridges. The *in vitro* culture of the ICM and primordial germ cells (precursors of the germ line) will eventually give rise to embryonic stem cells and embryonic germ cells, respectively (highlighted in purple).

1.2 Pluripotent stem cells

1.2.1 Embryonal carcinoma (EC) cells

Embryonal carcinoma (EC) cell is the name designating stem cells found within germ cell tumours (Kleinsmith and Pierce, 1964). These tumours, referred to as teratocarcinomas, were originally identified in the testis of the 129 strain of mice and comprised a mix of various somatic tissues (Stevens and Little, 1954). Since then, various mouse and human EC cell lines have been established (Pierce et al., 1957, Kahan and Ephrussi, 1970, Fogh et al., 1977). Like ES cells, EC cells express pluripotency markers such as stage specific embryonic antigen (SSEA)3, SSEA4, tumor rejection antigen (TRA)-1-60, octamer transcription factor (OCT)4, SRY (sex determining region Y)-box transcription factor (SOX)2 and NANOG (Andrews et al., 1984, Thomson et al., 1998, Matin et al., 2004, Greber et al., 2007, Jung et al., 2010). Therefore, EC cells serve as an *in vi*tro model for the study of stem cell carcinogenesis.

1.2.2 Embryonic stem (ES) cells

Embryonic stem cells were first isolated from mouse blastocysts (Evans and Kaufman, 1981, Martin, 1981). Thereafter, ES cells were derived from other mammalian blastocysts (Piedrahita et al., 1990, Strojek et al., 1990, Moreadith and Graves, 1992, Sukoyan et al., 1993, Cherny et al., 1994, Iannaccone et al., 1994), including non-human primates (Thomson et al., 1995) and humans (Thomson et al., 1998). Embryonic stem cells are able to proliferate for prolonged periods of time (self-renew) *in vitro* whilst maintaining their

capacity to differentiate into all differentiated cell types in the adult tissues, including germ cells and cells from the three primary germ layers (Fig. 1-1).

Although ES cells have been described in various species, most of the research has been done on mouse (m) and human (h)ES cells. ES cells are mainly characterized by the following criteria:

- Expression of key ES cell markers such as: the transcription factors OCT4, SOX2 and NANOG and the cell surface proteins SSEA3 and 4 (hES cells) and SSEA1 (mES cells) and TRA-1-60 and -1-81 (hES cells).
- Ability to form teratomas *in vivo* in immunocompromised mice or embryoid bodies (EBs) *in vitro*, which differentiate into tissues representative of the three germ layers.
- In the case of mES cells, capacity to incorporate in mouse pre-implantation blastocyst and give rise to chimeras, including the unique ability to contribute to the germline.
- Tetraploid complementation assay, whereby a tetraploid host morula, formed by the original fusion of blastomeres at the 2-cell stage and that would normally form extra embryonic tissues only, is combined with diploid ES cells from a donor embryo and thus give rise to a normal developing embryo, where the extra embryonic tissues come from the tetraploid host and the embryo proper from the diploid ES cells.

Both mouse and hES cells share expression of pluripotency markers as well as the ability to give rise to tissues representative of all three primary germ layers, but soon it was revealed that mouse and hES cells have significant differences. Mouse ES cells have a

round morphology and high clonogenic capacity whilst hES cells are more flattened and show decreased proliferation and clonal propagation. In addition, mouse and hES cells present different culture requirements, which in turn reflect different signalling pathways involved in their maintenance and differentiation.

Human ES cells are maintained in the presence of fibroblast growth factor (FGF)-2 combined with ACTIVIN A/NODAL signalling (Vallier et al., 2005). However, a member of the interleukin-6 (IL-6) family of cytokines known as leukaemia inhibitory factor (LIF), which is essential for mES cell maintenance in culture (Nichols et al., 1990) is dispensable for hES cells (Thomson et al., 1998). Upon binding of LIF to its receptor (LIFR), via interaction with glycoprotein (gp)130, it triggers downstream signalling cascades that include phosphorylation by Janus kinase (JAK), which ultimately activates signal transducer and activator of transcription (STAT)3, responsible for preventing mES cell differentiation (Niwa et al., 1998, Matsuda et al., 1999). Despite LIFR and gp130 being expressed and able to activate STAT3 in hES cells, LIF fails to maintain hES cells in an undifferentiated state (Humphrey et al., 2004). LIF is also known to maintain mES cell pluripotency via interaction with bone morphogenetic protein (BMP)4 (Ying et al., 2003, Qi et al., 2004), a factor that induces trophoblast differentiation in hES cells (Xu et al., 2002). Moreover, down regulation of OCT4 in mES cells leads to trophoblast differentiation (Niwa et al., 2000), but loss of OCT4 in hES cells does not seem to correlate with differentiation into trophoblast lineage (Hay et al., 2004, Matin et al., 2004, Rodriguez et al., 2007).

These differences in main signalling pathways and role of core transcription factors between mouse and hES cells strongly indicate that there are differences in the

mechanisms that underlie pluripotency for each type of ES cells. Given the potential of hES cells in the development of therapeutic treatments, unravelling the signalling pathways that maintain their pluripotency and prevent differentiation is one of the main issues to be addressed. That is why the focus of this thesis will be mainly on hES cells; nonetheless, reference to mES and iPS cells will be done where it is appropriate.

1.2.3 Embryonic germ (EG) cells

During embryonic development, a population of cells known as primordial germ cells form at the posterior region of the epiblast and migrate, through the allantois, towards the genital ridges in order to give rise to the germ line. It has been suggested that primordial germ cells are the cells that eventually give rise to teratocarcinomas (Andrews, 1998).

Embryonic germ (EG) cells originate from the *in vitro* culture of primordial germ cells (Matsui et al., 1992, Resnick et al., 1992, Shamblott et al., 1998). These EG cells also express characteristic pluripotent markers (SSEA1, SEEA3, SSEA4, TRA-1-60, Tra-1-81) and are able to differentiate into tissues representative of the three germ layers (Matsui et al., 1992, Resnick et al., 1992, Shamblott et al., 1998). More recently, a genome-wide analysis has demonstrated a specific gene signature for EG cells when compared to other pluripotent cells such as EC and ES cells (Pashai et al., 2012), therefore providing further insights into the pluripotent state of these cells.

1.2.4 Induced pluripotent stem (iPS) cells

In 1966 Gurdon and Uehlinger showed that when the nucleus of a cell from the adult intestine of Xenopus laevis was injected into an enucleated egg (somatic cell nuclear transfer, SCNT), this gave rise, at low efficiency, to fertile adult frogs (Gurdon and Uehlinger, 1966). Later on, SCNT from an adult mammary gland cell was achieved and produced a viable sheep named Dolly (Wilmut et al., 1997). SCNT was also achieved in other species and in some cases it was used as the means to produce ES cells in order to correct a genetic defect in mouse (Rideout et al., 2002).

This early work on SCNT suggested that the genetic material within adult cells might be able to undergo reprogramming and re-set to the totipotent state of the fertilized egg. These findings paved the way for the generation of mouse iPS cells that, as their name indicates, originated *in vitro* from mouse somatic cells that are induced to become a pluripotent ES-like cell (Takahashi and Yamanaka, 2006). Human iPS cells were originally generated by transducing adult somatic cells with retroviral vectors containing combinations of key transcription factors: SOX2, OCT3/4, Kruppel-like factor (KLF)4 and cellular myelocytomatosis oncogene (c-Myc) (Takahashi et al., 2007) and, SOX2, OCT3/4, NANOG and LIN28 (Yu et al., 2007). Human iPS cells appear to be indistinguishable from hES cells. However there are concerns that they might be different at the epigenetic level; thus further work remains to be done in order to confirm if these cells are truly and fully like their ES cell counterparts.

1.3 Human embryonic stem (hES) cells

Human ES cells, like mES cells, were initially derived and cultured in the presence of mitotically-inactivated mouse embryonic fibroblasts (MEFs) (Thomson et al., 1998). Under these conditions, hES cells maintain their pluripotent state whilst retaining the capacity to differentiate into cells from all the tissues in the body. The proliferation potential and great plasticity of hES cells is what makes them the ideal candidates for cell-replacement therapies in degenerative diseases (e.g. Alzheimer's disease) and in conditions where tissues have been damaged (e.g. myocardial infarction) or have lost certain function (e.g. diabetes). Moreover, hES cells represent an invaluable tool to model certain diseases, study human embryonic development or test pharmacological compounds. Despite the promising potential of hES cells, before being able to scale them up and apply them in the clinical setting, there are certain issues that must be addressed:

- Complete elimination of animal and xenogeneic components in the culture
- Use of fully defined culture conditions.
- A better understanding of the extrinsic and intrinsic factors that regulate the pluripotent state of hES cells.
- Unravelling the extrinsic and intrinsic factors necessary and sufficient for the differentiation of hES cells into the required cell lineage.
- Generation of vehicles, such as bioengineered scaffolds, for the specific delivery of differentiated hES cells into the required tissue(s).
- Prevention of immune rejection from the host.
- Development of techniques that will allow specific detection of incorporation of the inserted differentiated hES cells into the required tissue(s) and assessing their

functionality, as well as ensuring there are no undifferentiated hES cells left behind.

Following the isolation of hES cells, a considerable amount of research has been carried out over the past decades in order to solve these issues. Herein, the focus will be on the first four points listed above.

1.4 Extrinsic regulators of pluripotency in hES cells

The culture conditions that maintain hES cells in a pluripotent state host the molecules that are ultimately responsible for hES cell maintenance. Given that the origin of these molecules is external to hES cells, I will refer to these factors as extrinsic regulators. They appear to achieve hES cell maintenance by triggering downstream signalling cascade events inside hES cells.

As mentioned above, the presence of exogenous LIF (Smith et al., 1988, Williams et al., 1988) or inhibitors for glycogen synthase kinase (GSK)3 and mitogen-activated protein kinase(MAPK)/extracellular signal-regulated kinase (ERK) kinases (MEKs), referred to as '2i', are sufficient to maintain mES cells in a pluripotent state (Ying et al., 2003, Ying et al., 2008).

In the attempt to investigate and reproduce the key components in current hES cell culture systems that maintain hES cell pluripotency, various studies have identified some of the molecules and signalling pathways that are partially responsible for the maintenance of hES cell pluripotency, which include transforming growth factor (TGF)-β,

wingless-int (WNT) and FGF signalling, also known for their role during embryonic development.

The presence of ACTIVIN A, a member of the TGF-β superfamily of ligands, in hES cell culture was found to be necessary and sufficient for hES cell maintenance (Beattie et al., 2005, Xiao et al., 2006). Inhibiting ACTIVIN A with Follistatin leads to hES cell differentiation and loss of the pluripotency markers OCT4, NANOG and TRA-1-60 (Beattie et al., 2005, Vallier et al., 2005, Xiao et al., 2006); replacing ACTIVIN A with BMP4 also induces a similar effect (Beattie et al., 2005). This is consistent with BMP4 inducing hES cell differentiation into trophoblast lineage (Xu et al., 2002, Vallier et al., 2009). Meanwhile, sustained expression of NODAL, another member of the TGF-β superfamily of ligands, is also able to support hES cells but independently of ACTIVIN A (Vallier et al., 2005).

TGF-β/ACTIVIN/NODAL ligands represent one of the two main branches of the TGF-β signalling pathway, which, upon binding to their Type II receptors and recruiting Type I receptor, phosphorylate and activate mothers against decapentaplegic homolog (SMAD)2/3 that is then translocated into the nucleus and interacts with various transcription factors to regulate gene expression (Derynck and Zhang, 2003). In hES cells, ACTIVIN/NODAL signalling via SMAD2/3 is required for NANOG expression (Vallier et al., 2009). SMAD2/3 was found to bind the promoter of NANOG and regulate its expression (Xu et al., 2008, Vallier et al., 2009), which in turn prevented neuroectoderm differentiation but not differentiation into extra embryonic lineages (Vallier et al., 2009). Moreover, a luciferase reporter assay containing a SMAD-responsive promoter (Jonk et al., 1998) demonstrated that NANOG diminishes the transcriptional activity induced by

ACTIVIN/NODAL, perhaps through a negative feedback loop caused by the interaction between NANOG and SMAD2/3 (Vallier et al., 2009). In addition, NANOG and SMAD2/3 are able to bind the same regions of genes of the TGF-β signalling pathway, including LEFTYA (a NODAL antagonist) (Vallier et al., 2009). This correlates with the limited up regulation of LEFTYA and NODAL (known targets of SMAD2/3) that was observed during mesendoderm differentiation of hES cells that over express NANOG, where NANOG prevents further specification into endoderm (Vallier et al., 2009).

Besides having a role in hES cell maintenance, ACTIVIN/NODAL signalling has also been implicated in hES cell differentiation. The specification of progenitors from the primitive streak (PS) into either endoderm or mesoderm requires a balance in the interaction between BMP, ACTIVIN/NODAL and the canonical WNT signalling pathways. Mesoderm induction by β-CATENIN (intracellular signal transducer in the canonical WNT signalling pathway) in hES cells requires BMP inhibitors and ACTIVIN/NODAL signalling, whereas endoderm induction requires β-CATENIN as well as BMP inhibitors and active ACTIVIN/NODAL signalling (Sumi et al., 2008). Concomitant with a role for WNT/β-CATENIN pathway in hES cell differentiation, OCT4 was recently found to repress this pathway and without the intervention of NANOG (Davidson et al., 2012). Moreover, in hES cells, OCT4 and SOX2 have been shown to block the expression of differentiation markers of primitive streak and definitive endoderm such as EOMES (Teo et al., 2011).

This dual role of ACTIVIN/NODAL signalling in maintenance of hES cell via NANOG and mesoderm/endoderm specification via the canonical WNT/β-CATENIN and BMP signalling pathways has recently been reconciled. The level of phosphatidylinositol -3 kinase (PI3-K)/AKT (protein kinase B) activity correlates with the switch in function of ACTIVIN A,

from maintaining hES cells to promoting their differentiation (Singh et al., 2012). In the presence of high PI3-K/AKT 1 (one of the three mammalian isoforms of AKT) activity, ACTIVIN A, through SMAD2/3, is able to maintain hES cells. Maintaining AKT1 activity is critical for suppression of ERK1/2 and maintenance of GSK3-β activities (Singh et al., 2012). Conversely, the absence of PI3-K activation causes the loss of downstream AKT1 activity and the phosphorylation ERK1/2, which is responsible for the phosphorylation and inactivation of GSK3-β that in turn leads to the activation of WNT activators, such as β-CATENIN, and the up regulation of the mesendoderm markers eomesodermin (EOMES), homeobox proteins goosecoid (GSC) and MIXL1; expression of MIXL1 was prevented by ACTIVIN A inhibition, suggesting that activation of MIXL1 promoter by β-CATENIN perhaps occurs in synergy with SMAD2/3 (Singh et al., 2012).

In addition, the combination of ACTIVIN A/NODAL and FGF-2 allows hES cell maintenance, but FGF-2 was dependent on strong ACTIVIN A/NODAL activity (Vallier et al., 2005). Use of exogenous FGF-2 has been shown to aid hES cell maintenance (Xu et al., 2001). Moreover, FGF-2 appears to be important for hES cell pluripotency, instead of proliferation, by co-operating with BMP antagonists and repressing BMP signalling (Xu et al., 2005) or activating PI3-K/AKT pathway (Dvash et al., 2004, Armstrong et al., 2006, Li et al., 2007, Eiselleova et al., 2009, Ding et al., 2010). Nonetheless, there is increasing evidence that suggests FGF-2 also plays a role in hES cell differentiation: 1) into the neural lineage, which is opposed by NANOG (Vallier et al., 2009, Cohen et al., 2010) or 2) into mesendoderm, via ERK signalling (Yu et al., 2011). This suggests a role for FGF-2 in balancing hES cell pluripotency.

Another potential external regulator of hES cell maintenance is the environment created outside the cell, which is mediated by a complex molecular arrangement known as the extracellular matrix (ECM). Among these molecules are:

- proteins that are the structural base of the ECM (collagens, laminins, fibronectin, vitronectin and elastin).
- specialized proteins, which are present in the ECM but rather than forming its structure influence cell function (growth factors, small matricellular proteins and small integrin-binding glycoproteins).
- proteoglycans are heavily glycosylated proteins that fill the rest of the ECM. The negative charge conferred by the acidic sugar residues allows the recruitment of sodium ions, which lead to hydration and thus contributes to the mechanical properties of the ECM.

Therefore, the ECM not only provides a scaffold for cells/tissues but is also a dynamic structure that gathers bioactive molecules that influence cell/tissue homeostasis. As it will be discussed in Chapter 3, the role of the ECM in hES cell maintenance is poorly understood.

In mES cells and their EBs, deletion of SMAD4, a protein that binds the TGF-β-activated SMAD2/3 and together activates gene transcription, has been shown to cause an increase in the expression of components of ECM proteins such as laminin α-1 and decreased ECM remodelling, as observed by low expression of matrix metalloproteinases; this results in basement membrane deposition and prevents endoderm lineage differentiation (Costello et al., 2009).

Moreover, it is known that certain growth factors such as FGF and TGF-β interact with ECM proteins (Taipale and Keski-Oja, 1997). In the case of TGF-β, once its C-terminal domain is cleaved, a mature TGF-β remains covalently linked to a latency-associated peptide (LAP) and together is referred to as the small latent complex (SLC). This is then bound to one of the four ECM glycoproteins known as latent TGF-β-binding proteins (LTBPs), which bind the SLC inside the cell and mediate their export into the ECM as a large latent complex (LLC). More recently, it has been shown that in osteoblast and fibroblast cultures the assembly of fibronectin and fibrilins, respectively, is required for sequestering LTBP-1 and TGF-β to the ECM that surrounds these cells (Dallas et al., 2005, Massam-Wu et al., 2010). This supports a role for the ECM in maintaining a reservoir of growth factors that can be readily available to the cell and thus the ECM is able to ultimately influence downstream signalling inside the cell. Therefore it is likely that the ECM present in hES cell culture plays a similar role but this remains to be unravelled.

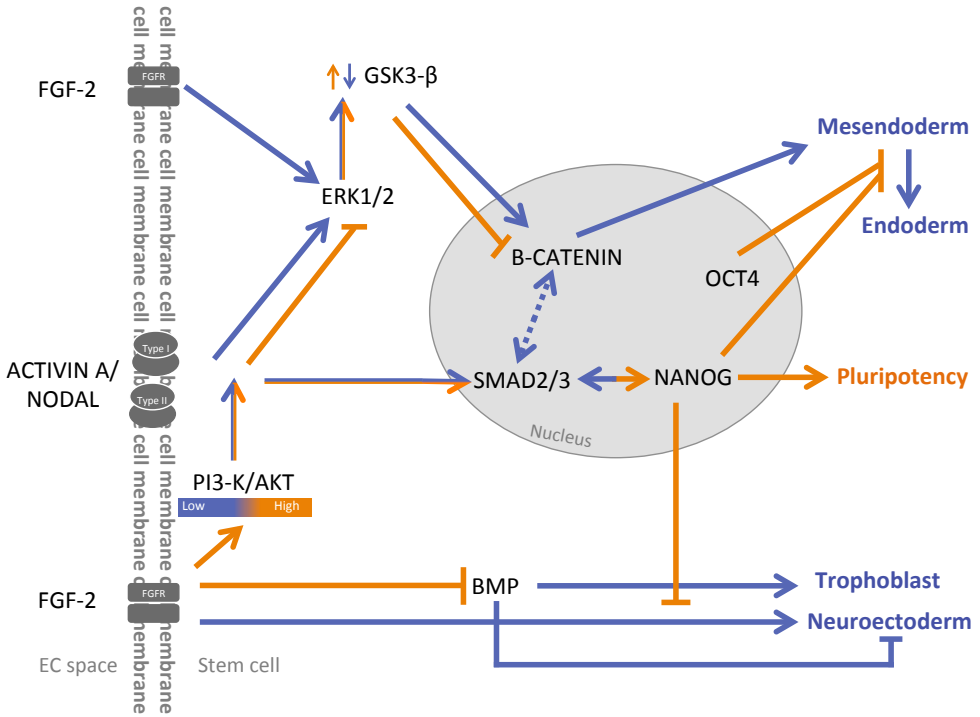


Figure 1-2. Schematic diagram summarizing the signalling cascades activated by extrinsic factors in hES cells.

Upon binding of extracellular ligands to their corresponding receptors, these activate a series of signalling cascade evens that lead to pluripotency and differentiation, depicted by orange and blue lines, respectively. Arrows represent activation, while lines ending with a vertical line represent inhibition. Mixed-colour arrows indicate influence on pluripotency or differentiation via the same cascade. The double–headed arrows, continuous or dashed, represent interaction or possible interaction, respectively, between factors. AKT, protein kinase B; EC, extracellular; ERK1/2, extracellular signal-regulated kinase 1/2; FGF-2, fibroblast growth factor-2; FGFR, FGF receptor; GSK3-β, glycogen synthase kinase 3-β; PI3, phosphatidylinositol -3 kinase; SMAD2/3 mothers against decapentaplegic homologs 2 and 3.

1.5 Intrinsic regulators of pluripotency in hES cells

1.5.1 The core transcription factors of pluripotency

OCT4, NANOG and SOX2 have been identified as the key regulators of ES cell pluripotency. OCT4, also known as OCT3 and encoded by the *POU5F1* gene, is a transcription factor from the POU (Pituitary-specific Pit-1, Octamer transcription factors OCT1 and OCT2 and the neural Unc-86 transcription factor) family that binds the octamer motif (ATGCAAAT) within the promoter and enhancer regions of various genes via its POU domain (Okamoto et al., 1990). NANOG, named after "The Land of the Young" Tir na nOg from Celtic mythology, is a homeobox transcription factor, (Chambers et al., 2003, Mitsui et al., 2003). SOX2 belongs to the *SOX* (SRY-related HMG box, where SRY stands for sex-determining region of Y chromosome and HMG stands for high mobility group) gene family of transcription factors (Gubbay et al., 1990).

Initial clues pointing to a role of these transcription factors in the maintenance of hES cell pluripotency first came from early work performed *in vivo* in the developing human and mouse embryo and *in vitro* in mouse and hEC and EG cells and mES cells. *In vivo*, OCT4, NANOG and SOX2 are detected at various stages of development but only OCT4 is found in mouse and human oocytes (Rosner et al., 1990, Abdel-Rahman et al., 1995). In the mouse, upon fertilization, NANOG and SOX2 are first expressed at the morula stage and become localized to the ICM of the pre-implantation blastocyst (Avilion et al., 2003, Chambers et al., 2003). OCT4 continues to be expressed at the early stages of embryonic development and it is also particularly obvious in the ICM of the pre-implantation blastocyst (Palmieri et al., 1994). Expression of NANOG becomes even further restricted,

to the epiblast fate of the ICM, and is considerably down regulated upon implantation and not found at all in adult somatic or germline tissues (Chambers et al., 2003, Mitsui et al., 2003, Hart et al., 2004, Hyslop et al., 2005). Similarly, OCT4 expression is down regulated during gastrulation but its expression is maintained in the germline (Rosner et al., 1990). In the case of SOX2, following gastrulation, its expression is confined to the neuroectoderm and by later stages is found throughout the developing central nervous system (CNS), sensory placodes, branchial arches and gut endoderm (Wood and Episkopou, 1999, Avilion et al., 2003).

In addition, *in vitro* expression of OCT4 and NANOG is detected in the nucleus of mouse and hEC and EG cells as well as in ES cells (Okamoto et al., 1990, Rosner et al., 1990, Yeom et al., 1996, Chambers et al., 2003, Mitsui et al., 2003, Wang et al., 2003, Hyslop et al., 2005).

The fact that the expression of these transcription factors is mainly restricted to the ICM of the pre-implantation blastocyst, from which ES cells are derived, and to pluripotent cells such as EC and EG cells, provides a good indication of their potential role in specifying early cell fate decisions in the developing embryo. Further studies have confirmed this by means of manipulating the expression of OCT4, NANOG and SOX2 in mouse embryos and hES cells (Nichols et al., 1998, Avilion et al., 2003, Mitsui et al., 2003). Mouse blastocysts that carry a homozygous mutation for OCT4 ($OCT4^{-/-}$) fail to implant and therefore are not viable (Nichols et al., 1998), whereas mouse embryos null for SOX2($^{-/-}$), despite presenting a normal phenotype at the blastocyst stage, die around the time of implantation (Avilion et al., 2003). NANOG($^{-/-}$) null mice have a mixed phenotype, most embryos appear normal in morphology and few contained disorganized

extra embryonic tissues (Mitsui et al., 2003). Moreover, isolation and further culture of these mutant embryos revealed that $OCT4^{-/-}$ and $SOX2^{-/-}$ follow a trophoblastic lineage (Nichols et al., 1998, Avilion et al., 2003) whereas $NANOG^{-/-}$ embryos form parietal endoderm (Mitsui et al., 2003). Together, these experiments clearly demonstrate that expression of OCT4 is imperative to the formation of the ICM and that NANOG and SOX2 are also required for this task. Furthermore, given that the ICM is ultimately the source of ES cells, these transcription factors are likely to participate in the events that maintain pluripotency in hES cells.

Down regulation of OCT4 in hES cells causes a decrease of pluripotency markers SOX2 and NANOG (Chew et al., 2005, Babaie et al., 2007), SSEA3 and TRA-1-60 (Matin et al., 2004) and an increase of differentiation markers of various lineages such as extra embryonic endoderm (Hay et al., 2004) and embryonic endoderm and mesoderm (Rodriguez et al., 2007). More recently, it has been demonstrated that the ability of OCT4 to prevent hES cell differentiation into those various lineages is dependent on BMP4 signalling, which in turn appears to be hES cell line-specific (Wang et al., 2012). The lack of OCT4 leads to neuroectoderm differentiation and to extra embryonic ectoderm and primitive endoderm when BMP4 is present (Wang et al., 2012).

Meanwhile, loss of NANOG in hES and hEC cells results in down regulation of pluripotency markers SSEA4 and OCT4 but not SOX2 (Hyslop et al., 2005, Vallier et al., 2009) and up regulation of neuroectoderm and neural crest markers (Vallier et al., 2009, Wang et al., 2012); the up regulation of markers of embryonic and extra embryonic endoderm and trophoblast (Hyslop et al., 2005) and extra embryonic ectoderm (Wang et al., 2012) lineages may be induced by BMP4 signalling (Vallier et al., 2009, Wang et al., 2012). On

the other hand, down regulation of SOX2 in hES cells leads to decreased levels of OCT4, NANOG, SSEA3, SSEA4, TRA-1-60 and TRA-1-81 and increased expression of trophoblast markers (Chew et al., 2005, Fong et al., 2008, Adachi et al., 2010). Interestingly, when SOX2 levels are down regulated, a member of the same family, SOX3 is up regulated and contributes to hES cell plutipotency but not to self-renewal (Wang et al., 2012).

Over expression of OCT4 in hES cells induces their differentiation into endoderm lineage and may suppress mesoderm differentiation (Rodriguez et al., 2007). However, it has been shown that when high levels of OCT4 are combined with BMP4, this drives mesoderm differentiation (Wang et al., 2012).

Interestingly, up regulation of NANOG in hES cells does not seem to affect their maintenance and differentiation potential. However, expression of some genes is misregulated, including down regulation of ECM-encoding genes and up regulation of genes from the TFG-β signalling pathway, including NODAL (Darr et al., 2006, Fischer et al., 2010). Similar results have been reported in mES cells, where NANOG null cells, although likely to differentiate, retain pluripotency and the capacity to differentiate (Chambers et al., 2007). In the case of SOX2, its over expression in hES cells causes reduction in OCT4 and NANOG and up regulation of trophoblast markers, along with no apparent change in the expression of endodermal, mesodermal and neuronal markers (Adachi et al., 2010).

Thus, down regulation and over expression of the core transcription factors of pluripotency in hES cells often disrupts pluripotency and leads to differentiation; therefore specific levels of these transcription factors are ultimately responsible for hES cell maintenance. This is further illustrated in elegant experiments by Smith and

colleagues, which showed that precise expression levels of OCT4 are essential but not sufficient to maintain the pluripotent state in mES cells. Levels of OCT4 were manipulated by means of introducing a tetracycline-induced *OCT4* transgene system. Loss of OCT4 led to trophoblast differentiation while only two-fold over expression of OCT4 was sufficient to induce mES cell differentiation into primitive endoderm and mesoderm lineages (Niwa et al., 2000). In addition, more recent work has suggested that, in the case of NANOG, fluctuation in the expression levels can determine if ES cells differentiate or not (Chambers et al., 2007). Using a green fluorescent protein (GFP) reporting system, it was observed that mES cells consist of a mix of cells that either express or do not express NANOG, and that the latter differentiated but were also capable of re-gaining an undifferentiated status. This is consistent with earlier studies regarding the ability of NANOG to act on signalling via SMAD2/3, which is able both to maintain hES cell pluripotency and induce their differentiation (Vallier et al., 2009). Therefore, it is suggested that the fluctuating levels of NANOG act as a gatekeeper, with low levels opening a window for differentiation.

In summary, for OCT4, NANOG and SOX2 to be able to maintain the pluripotent state of hES cell their expression levels must be precisely tuned. In addition, the interaction of these transcription factors with members of signalling pathways observed in specific hES cell lines may account for the differences in the tendency of some hES cell lines to differentiate into specific cell lineages (Bock et al., 2011).

1.5.2 Circuitry established by OCT4, NANOG and SOX2

Given the importance of OCT4, NANOG and SOX2 in maintaining the pluripotent state of hES cells, it was soon investigated whether they interacted together and/or targeted common genes and signalling pathways involved in hES cell pluripotency. A summary of this circuitry is illustrated in Fig. 1-3.

Chromatin immunoprecipitation (ChIP) analyses and electrophoretic mobility shift assays have shown that OCT4 and SOX2 bind as a complex to their own enhancer regions in hES cells (Chew et al., 2005). OCT4/SOX2 complex was found to bind the (SRR2) region of the *SOX2* promoter and the octamer motif (oct)-sox element within the conserved region (CR)4 of the distal enhancer (DE) of the *POU5F1* promoter (Chew et al., 2005, Okumura-Nakanishi et al., 2005); down regulation of either OCT4 or SOX2 hampered the activity of the *POU5F1* promoter (Chew et al., 2005). Likewise, NANOG expression in hES cells was also regulated via binding of OCT4/SOX2 complex to the oct-sox element within the *NANOG* promoter (Kuroda et al., 2005, Rodda et al., 2005).

The fact that misregulation of the core transcription factors often leads to down regulation of at least two of them, and that OCT4/SOX2 interact as a complex strongly suggest a feedback loop mechanism of control of their gene expression. Nevertheless, recent work has shown that such feedback loops are not connected; in the absence of SOX2, the expression of OCT4 and NANOG retained hES cell pluripotency, whereas the lack of OCT4 and NANOG did not affect SOX2 expression, which led to neuroectoderm differentiation (Wang et al., 2012).

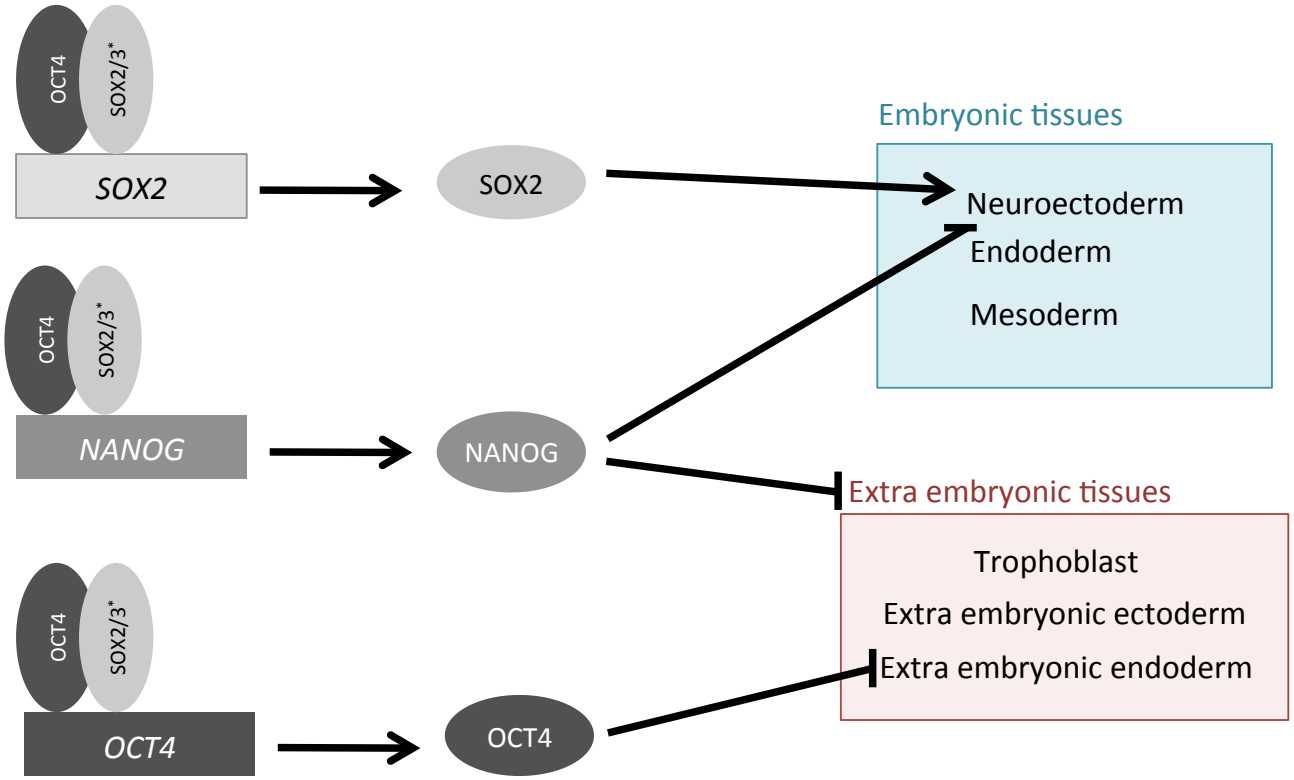


Figure 1-3. Diagrammatic representation of the regulatory circuitry established by the core transcription factors of hES cell pluripotency.

Human ES cell pluripotency is mainly maintained by the repression of cell differentiation into embryonic and extra embryonic lineages by the core transcription factors NANOG and OCT4; conversely, SOX2 modulates differentiation into the neural lineage. Nonetheless, the regulation of expression of these transcription factors (circles) is, in part, achieved by the interaction of OCT4 and SOX2 at the promoter region of their genes (boxes), thus establishing a feedback loop. It is noteworthy that in the absence of SOX2, SOX3 has been found to substitute for the role of SOX2; perhaps also influencing the expression of the core transcription factor, though this has not been directly demonstrated (*).

1.6 An additional level of regulation of pluripotency in hES cells

The extrinsic and intrinsic factors contributing to the maintenance of hES cells that have been described here have in common that they are all products of protein-coding genes. However, these genes only account for ~1% of the total human genome while the rest is represented by genes that instead encode RNAs that are collectively known as non-coding (nc)RNA. For sometime, it was thought that ncRNAs only had a limited and even negligible biological function.

Over the past decade there has been an increasing amount of data that suggests that ncRNAs also play a role in maintaining human pluripotent stem cells. Such is the case of the micro(mi)RNA-145 (whose levels are reduced in hES cells but elevated upon differentiation) that regulates the levels of core pluripotency transcription factors OCT4, SOX2 and KLF4 at the post-transcriptional level in hES cells (Xu et al., 2009a). Moreover, OCT4 was shown to regulate the promoter of the same miRNA (Xu et al., 2009a), thus suggesting a feedback loop mechanism that ultimately contributes to the balance of hES cell maintenance and differentiation.

Another type of ncRNAs, long (l)ncRNAs, have also been implicated in ES cell pluripotency. The lncRNA_ES1 and lncRNA_ES2 have been shown to interact with the core transcription factor of pluripotency SOX2 but not with OCT4 (Ng and Stanton, 2013).

Therefore, these studies suggest that ncRNAs provide an additional layer of regulation of hES cell maintenance. Among ncRNAs, there is a group of RNAs that are found in the nucleus of eukaryotic cells and thus are referred to as small nuclear (sn)RNAs. They are mainly responsible for processing protein-coding genes but their role in hES cell

pluripotency is largely unknown. Small nRNA genes lack introns and instead possess an RNA-encoding region flanked by a 5' promoter region composed of a distal and proximal sequence elements (DSE and PSE) required for transcription initiation and a 3' box element needed for 3' end processing.

Our laboratory has recently identified and characterized a group of snRNA genes that are variants of the known uridine-rich (U)1 snRNA gene, named vU1 snRNA genes (O'Reilly et al., 2012). In order to describe the work that has been done on the vU1 snRNA genes, I shall first introduce U1 snRNA.

1.6.1 U1 and mRNA processing

The U1 snRNA, in association with specialized proteins (U1C, U1A, U1-70K) and Sm (Smith) proteins (B, D1, D2, D3, E, F and G), common to all U small nuclear ribonucleoproteins (U snRNPs), forms the U1 snRNP complex (Lerner and Steitz, 1979). The U1 snRNA gene is transcribed by RNA polymerase II in the nucleus as a 5' capped precursor transcript that is exported into the cytoplasm for further maturation by the survival of motor neuron (SMN) protein complex, which mediates the assembly of Sm proteins into a heptameric ring structure around the Sm site of the U1 snRNA (Fischer et al., 1997, Liu et al., 1997). Prior to re-export into the nucleus, the 5' cap of U1 snRNA becomes hypermethylated and the 3' is trimmed (Mattaj, 1986). Final assembly of the Sm-containing U1 snRNA with the specialized proteins occurs in the nucleus, in part, within the non-membranous nuclear structures known as Cajal bodies by a process that is not yet fully understood (Kiss, 2004, Nesic et al., 2004).

In turn, a fully-mature U1 snRNP is part of a group of U snRNPs: U2, U4, U5 and U6 that form a complex known as the major spliceosome (Grabowski et al., 1985). Equivalently, U11, U12, U4atac, U5 and U6atac form the minor spliceosome, which is responsible for splicing a class of introns (1 in 300 in humans) that possess the conserved AT-AC (as compared to GT-AG) sequence at the 5' and 3' splice sites (Tarn and Steitz, 1996).

Meanwhile, the job of the major spliceosome is to process newly transcribed mRNA transcripts, also referred to as precursor (pre)-mRNA, in an event known as splicing. Pre-mRNAs consist of exons, which are the sequences that encode proteins, and introns or intervening sequences (as their name indicates, interspersed among the former), which do not code for proteins. In order to produce a fully functional protein this pre-mRNA undergoes splicing, whereby the introns of a pre-mRNA are removed and exons are joined together as a continuous mRNA sequence that encodes for a protein. In some cases, exons are excluded or introns/intron parts are included; this variation of splicing is referred to as alternative splicing. Thus a single pre-mRNA can give rise to different mRNAs that are eventually translated into different proteins.

U1 snRNP is responsible for initiating splicing; the 5' end of the U1 snRNA within U1 snRNP recognizes a consensus sequence (GU/GURUGA) at the exon-intron boundary. In turn, this leads to the sequential loading of U2, U4, U5 and U6, which form intermediate complexes. U1 snRNP bound to the 5' splice site associates with splicing factors to form the early (E) complex independently of ATP; within the intron, splicing factor (SF)1 binds the branch point (BP) and the U2AF binds the polypyrimidine tract (PPT) and 3' splice site (Ruskin et al., 1988, Berglund et al., 1998). Thereafter, an ATP-dependent complex A, formed by the addition of U2 snRNP (Gozani et al., 1998), prepares the pre-mRNA for the

first transesterification reaction (i.e. nucleophilic attack to the 5' splice site by the 2' hydroxyl of a conserved adenine residue within the BP). Further integration of U2 and the U4/U6 and U5 as a tri-snRNP into complex A gives rise to complex B. This becomes activated (B^{act} complex) following the exit of U1 and U4, which allows further arrangements that include the interaction of U2 and U6 as a catalytic structure (Valadkhan and Manley, 2001) to form complex B*, where the 5' splice site gets cleaved. This generates complex C1, which contains the free exon and an intermediate product formed by the 2'–5' phosphodiester branch (lariat). Finally, complex C1 undergoes a conformational change into complex C2, which allows the joining of exons (Sawa et al., 1988, Hirose et al., 2006). Ultimately, alternative splicing provides the means to amplify the repertoire of proteins from a fixed number of genes in a given cell. This is particularly useful during embryonic development, and consequently during hES cell differentiation, when cells with the same genetic information must acquire completely different phenotypes to give rise to the various tissues present in the adult organism.

1.6.2 The role of alternative splicing in embryonic development and hES cells

The differences in gene expression in different tissues, during development and ES cell differentiation have been well documented at the transcriptome level (Brandenberger et al., 2004, Calhoun et al., 2004, Dvash et al., 2004, Miura et al., 2004, Cai et al., 2006, Xu et al., 2009b, Fathi et al., 2011). However, over 90% of human genes undergo alternative splicing (Wang et al., 2008); many of which cannot be detected by conventional microarray platforms.

Nevertheless, more recent and sensitive genome-wide technologies, including exon arrays and RNA sequencing, have focused on analysing alternative splicing in different cell types. A computational approach combined with experimental validation showed that alternative splicing is more frequent in tissue-specific genes than in constitutively expressed genes (Pritsker et al., 2005).

Recent work in the developing mouse embryo indicates that a great proportion of genes that are differentially expressed throughout three developmental stages, embryonic day (E)8.5, 9.5 and 11.5, are also differentially spliced; these include factors that regulate splicing itself such as RNA binding proteins and genes that are active in cell attachment to the ECM (Revil et al., 2010).

In addition, in the case of ES cells, a few studies have shown that the patterns of alternative splicing in mouse and hES cells and differentiated cell types differ. Alternative splicing appears to be high in hES cells and gradually decreases along the neural differentiation pathway (Yeo et al., 2007, Wu et al., 2010).

All these data provide increasing evidence that alternative splicing events are common and might constitute an additional layer of regulation of gene expression. Future work in this field will provide more information on specific spliced variants, including their relevance in hES cell maintenance, and the splicing factors that influence alternative splicing.

1.6.3 Splicing-independent roles of U1

Initiation of splicing by recognition of the 5' splice site is one of the best-characterized functions of U1 snRNP. However, it has been shown that splicing activity is not fully supressed in the absence of U1 snRNP; instead, additional proteins such as the snRNP protein U1C or the non-snRNP SR (serine/arginine-rich) proteins are able to compensate for the lack of a functioning U1 snRNP (Crispino et al., 1994, Tarn and Steitz, 1994, Du and Rosbash, 2002). Moreover, in some splicing-related diseases that show aberrant 5' splice sites usage, splicing occurs independently of the U1 snRNP, reinforcing the notion that factors other than the U1 snRNA can regulate splice site selection (Raponi and Baralle, 2008, Raponi et al., 2009).

The fact that splicing can occur independently of U1 snRNPs has prompted the question of whether U1 snRNA could play additional roles besides its role in splicing. It has been suggested that U1 snRNP is important at the level of transcription, including pre-initiation, initiation and elongation, and at the level of polyadenylation.

U1 snRNA was found to interact with the transcription binding protein-associated factor (TAF)-15 (Jobert et al., 2009), which in turn interacts with TFIID (Bertolotti et al., 1996) required for recognizing promoters and assembling the pre-initiation complex. It is noteworthy that the interaction between U1 snRNA and TAF-15 required the Sm motif within U1 but independent of the U1-associated proteins, suggesting this interaction may require association of other proteins and thus form an alternative U1 snRNP complex (Jobert et al., 2009). Similarly, it was shown that U1 snRNA, but not as U1 snRNP, interacts with TFIIH, from the pre-initiation complex; this interaction was disrupted when the 5' splice site was mutated (Kwek et al., 2002). On the other hand, work using an HIV-1

mRNA also supports a role for U1 snRNA in enhancing transcription, but as part of a U1 snRNP; this appears to be dependent on U1 snRNA binding to splice donor sites and having an intact stem loop II (Alexander et al., 2010). Conversely, U1 snRNA, but not U2, U4 and U6, was found in association with transcription sites even when splicing of β-globin intron was deficient (Spiluttini et al., 2010). Furthermore, the U1-associated proteins U1-70k and Sm protein B were found to interact with the human transcription elongation factor TAT-SF1, via the first RNA recognition motif (RRM) of TAT-SF1 (Fong and Zhou, 2001).

Some of the first studies that suggested U1 snRNA could be involved in events other than splicing came from work in viruses, where U1 snRNP inhibits cleavage and polyadenylation. This is caused by binding of U1 snRNP to 5' splice sites located downstream of polyadenylation sites (Ashe et al., 1995, Gunderson et al., 1998). *In vitro* experiments showed that inhibition of polyadenylation occurs via interaction of U1 snRNP with poly(A) polymerase (PAP). Moreover it is dependent, in part, on the C-terminal region of PAP and the U1-70K protein (Gunderson et al., 1998).

More recently, it has been shown that U1 snRNP is important in preventing premature cleavage and polyadenylation (PCPA) (Kaida et al., 2010). Genome-wide analysis in HeLa cells where U1 snRNA had specifically been knocked down showed that for ~9% genes there was an accumulation of intronic sequences, indicative of a splicing defect, while for a higher proportion of genes their transcripts were terminated prematurely. Therefore, it has been proposed that during transcription U1 also binds 5’ splice-like sequences within introns and blocks cleavage/polyadenylation occurring at cryptic poly(A) sites within

these introns by interfering with the polyadenylation complex, possibly via its U1-70K protein (Kaida et al., 2010, Berg et al., 2012).

Taking advantage of the role of U1 snRNA in preventing polyadenylation, it was described that gene expression could be inhibited by U1 snRNA (Fortes et al., 2003, Abad et al., 2008). The transient or stable transfection of a plasmid that expresses a U1 snRNA that is mutated at its 5' end achieves inhibition through its interaction with a target site of 10-11nt in length, where positions 3 to 8 are critical (Abad et al., 2008).

1.6.4 Variant U1 snRNAs

The U1 snRNA gene belongs to a multi-gene family that comprises both functional genes (usually referred to as *bona fide* or true genes) and pseudogenes that have no ascribed function. U1 snRNA genes and pseudogenes have been identified in different species including human (Denison et al., 1981, Monstein et al., 1982, Monstein et al., 1983), mouse (Nojima and Kornberg, 1983), rat (Watanabe-Nagasu et al., 1983), chicken (Kristo et al., 1984) and fruit flies (Lo and Mount, 1990). In humans, U1 snRNA genes are localized to the short arm of chromosome 1 (Lund et al., 1983, Naylor et al., 1984). The described number of copies of U1 snRNA genes per haploid genome ranges from 30-150, whilst the estimated number of pseudogenes is an order of magnitude higher (Denison and Weiner, 1982). These U1 snRNA pseudogenes have been grouped in three classes according to their structure and the mechanism by which they may have originated (Denison and Weiner, 1982):

- Class I pseudogenes, localized to the long arm of chromosome 1 bands *q12-q22* (Lindgren et al., 1985), are the closest in sequence to the U1 snRNA genes, the 5' and 3' flanking regions as well as the RNA-encoding sequence are closely related; therefore it has been suggested that they originated from a DNA duplication mechanism.

- Class II and III pseudogenes show greater divergence in the flanking sequences, which suggests they originated from RNA-mediated mechanisms that possibly included cDNA as an intermediate step. Recently, three variants of U1, namely U1A5, U1A6 and U1A7, that appear to belong to these class of pseudogenes (due to lack of conservation to the 5' and 3' flanking regions of U1), have been identified as actively transcribed in various adult and embryonic tissues (Kyriakopoulou et al., 2006).

Note that despite the closely conserved regions between the prototypic U1 snRNAs and the class I pseudogenes, these had not been investigated further until now. As previously mentioned, work from our group has demonstrated that there are vU1 snRNAs, actively expressed and functional, which correspond to the misclassified class I pseudogenes (O'Reilly et al., 2012).

A search using the UCSC genome browser BLAST-like alignment tool for sequences similar to the U1 snRNA revealed 4 near perfect and 21 putative hits at the *1p36* and *1q12-21* chromosome loci, respectively. The later were classified further into 9 groups according to >95% sequence homology (O'Reilly et al., 2012).

The alignment of U1 and vU1 snRNAs showed that they are fairly conserved (~50% homology) at the 5' promoter region (-300bp), with 80% and 70% of sequence homology at the DSE and PSE, respectively. Conversely, they are highly divergent at the 3' end region (+300bp). On the other hand, within the snRNA-encoding region (164bp) few nucleotide changes that range from 1 to up to 30, and some deletions/insertions, are observed (O'Reilly et al., 2012). Variant U1.18, vU1.7+ 9 and vU1.10 are the least variable whereas vU1.4+5, vU1.20 and vU1.3+12 are the most variable; vU1.8 is the only variant that presents a considerable deletion within the snRNA-encoding region, that is localized in the stem loop II where the U1A protein binds (Fig. 1-4).

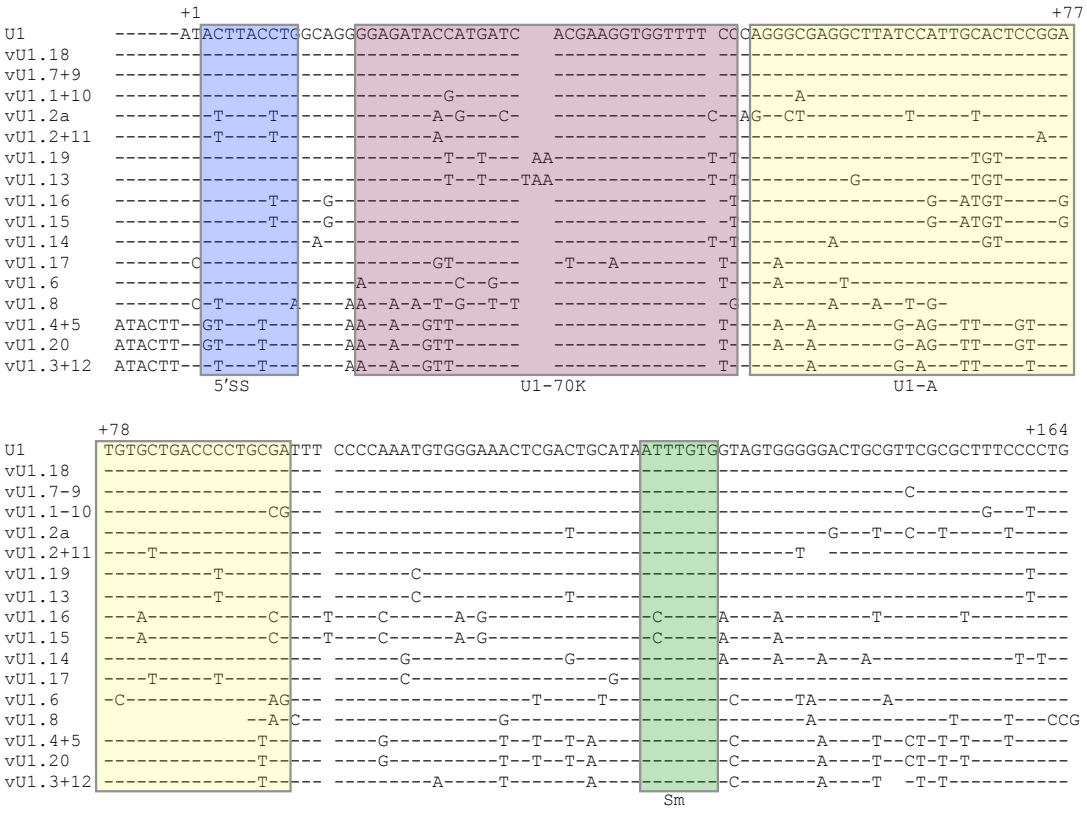


Figure 1-4. RNA-encoding sequence of vU1 snRNA genes.

Alignment of the snRNA-encoding sequence of the U1 snRNA genes (U1.1-4) with the 21 vU1 snRNA genes. Paired vU1 snRNAs have identical sequences. Non-conserved bases are denoted with the letter code representing one of the 4 nucleotide bases. – indicates the presence of a corresponding base, and a gap indicates the absence of a base at that position. The bases are numbered, beginning with 1 for the first base of the U1 snRNA sequence. Important features of the U1 snRNA are boxed and indicated underneath the alignment: 5' splice site recognition motif (5'ss, blue box), U1-70K protein binding region (U1-70K, purple box); U1-A protein binding region (U1-A, yellow box); Sm binding site (Sm, green box) (Taken from O'Reilly et al., 2012).

Assessment of the gene expression of each of the 9 groups of vU1 snRNAs in HeLa cells revealed by ChiP analyses that at least 5 of these groups (vU1.6, vU1.7+9, vU1.8, vU1.13-16+19 and vU1.18) have marks indicative of active snRNA transcription (O'Reilly et al., 2012), including RNA Pol II and the phosphorylation of serine (Ser)7 within its C-terminal domain, a mark associated with snRNA gene transcription (Egloff et al., 2007), PSE-binding transcription factor (PTF) and acetylated histone H3. Moreover, further comparison with hES cells showed that these marks are also present in the majority of vU1 snRNA genes in this cell type (O'Reilly et al., 2012).

In addition, it was demonstrated that vU1 snRNA genes generate nascent transcripts (i.e. transcription going beyond the 3' box) *in vivo* (O'Reilly et al., 2012). HeLa and hES cells express various levels of vU1 nascent transcripts, with vU1.1+10, vU1.2a,2+11 and vU1.3-5,12+20 and vU1.8 mainly detected in hES cells (O'Reilly et al., 2012). Furthermore, when hES cells are differentiated into EBs, RNA Pol II occupancy is considerably reduced but the steady state levels of some of these vU1 snRNAs are only mildly down regulated (O'Reilly et al., 2012).

Finally, interrogation of the Affymetrix Human Exon ST 1.0 array following knock down (89%) of vU1.8 in HeLa cells revealed that the main defect is not in splicing, as initially suspected, but premature 3' end processing (O'Reilly et al., 2012), suggesting that this vU1 snRNA plays a role in protecting a subclass of pre-mRNAs from premature cleavage and polyadenylation, as already shown for U1 snRNA (Kaida et al., 2010, Berg et al., 2012). Thus, vU1 snRNAs regulate gene expression at the level of 3' processing.

1.7 Project aim and description

As mentioned before, identifying the factors that regulate the balance between pluripotency and differentiation in hES cells as well as understanding the molecular mechanisms that trigger these events constitute a crucial step for ultimately developing hES cell-based therapeutic treatments. Following the isolation of hES cells in 1998, there has since been a considerable amount of work on the factors (mainly proteins) that maintain hES cells in a pluripotent state or lead to their differentiation. However, the mechanisms and factors (including ncRNAs) that regulate the expression of those proteins are less well understood. To our knowledge, a direct link between snRNAs and hES cell maintenance has not yet been described.

As described above, recent work from our laboratory has provided evidence that vU1 snRNAs are fully-processed transcripts that are differentially expressed in the cancerous HeLa cell line, in hES cells and their EBs. We hypothesize that vU1 snRNAs play a significant role in regulating gene expression in hES cells and that their differential expression may be key to hES cell differentiation (O'Reilly et al., 2012). Therefore the main aim of my project is to fully characterize the expression of vU1 snRNAs in hES cells and throughout their differentiation to ultimately identify the role vU1 snRNAs play in hES cell maintenance. The work towards this goal is presented in Chapters 3, 4 and 5 and has been divided, respectively, as follows:

1. **Establishment of a feeder-free culture system for hES cells.** Partial deconstruction of the components in hES cell culture that requires the use of MEFs and other fibroblasts to allow the reproduction of such conditions in the

absence of MEFs/fibroblasts. Under these conditions, RNA analyses of hES cells will not be contaminated with xenogeneic RNA.

2. **Profiling of vU1 snRNA expression in hES cells and their differentiated cell types.** Three different hES cell lines were screened for vU1 snRNA expression and further differentiated following a step-wise protocol that involves differentiation of hES cells into EBs, from which monocytes are produced and further induced to become macrophages. RNA was extracted from each hES cell line and each cell type at every stage in differentiation and subjected to quantitative real-time PCR (qPCR) analysis using vU1 snRNA gene-specific primers to establish the expression profiles of the vU1 snRNA in the different cell populations.
3. **Manipulation of vU1 snRNA expression in hES cells.** Variant U1 snRNAs were targeted for knockdown or over expression using a tetracycline-inducible gene expression system. Combined with deep RNA sequencing analysis, this will provide insights into their pre-mRNA targets.

Together, these experiments improve our understanding of the role of these snRNAs in hES cell maintenance in addition to helping to characterizing the changes in expression of genes that underlie the pluripotent state and differentiation potential of hES cells.

These results will also help develop more defined hES cell culture conditions and protocols for differentiation into specific cell lineages, thus bringing hES cell-based therapies a step closer.

Chapter 2| MATERIALS AND METHODS

2.1 Cell culture

2.1.1 Feeder cells (MEFs, CCD and Hs27)

All fibroblast feeder cell lines were grown at 37^{o}C in 5% CO_2 in sterile feeder medium consisting of 180ml of Dulbecco's Modified Eagle Medium (DMEM) containing supplemented with 20ml heat-inactivated foetal calf serum (FCS), 2ml of 200mM L-glutamine (Invitrogen), 2.5ml of 100x non-essential amino acids (NEAA) and 0.2ml of 50mM mercaptoethanol.

2.1.1.1 Feeder cell preparation for hES cell culture

Six-well tissue culture treated plates were coated with 1ml of sterile 0.1% gelatin solution (from porcine skin Type A, from Sigma-Aldrich) for at least 20 minutes at 37^{o}C. Gelatin was subsequently aspirated and 0.17x10^{6} γ–irradiated or 1x10^{6} mitomycin C-treated feeder cells in 2ml of feeder medium were added to each well of a 6-well plate.

2.1.1.2 Preparation of feeder conditioned media (CM)

Feeder medium was aspirated from prepared feeders (MEFs or Hs27), cells were washed with phosphate buffered saline (PBS, from Sigma-Aldrich) once and corresponding volume of hES cell medium (80% DMEM F12 or 78% KnockOut DMEM, 20% KnockOut Serum Replacement, 1.25ml of 200mM L-glutamine or 2.5ml GlutaMAX, 2.5ml NEAA and 0.5ml of 2μg/ml Basic Fibroblast Growth Factor, bFGF) was added to the cells. After 24h,

hES cell medium from these feeders was collected and filtered through a 0.20µm filter. Conditioned media (CM) was used or stored at -20^{o}C.

2.1.2 Human ES cells (H1, HUES1-4 and OxF1)

2.1.2.1 Manual passage

Old ES cell medium was replaced with fresh one early on the day of the passage. Individual colonies were manually cut with a 21G 1½" needle attached to a 1ml syringe, forming a square mesh. Thereafter, each patch was lifted up with the tip of a sealed glass pipette. The floating patches (approximately 20 colonies per well of a 6-well plate) were mixed with fresh hES cell medium and split equally into 2 wells o a 6-well plate containing irradiated feeder cells prepared as previously described. Human ES cells were fed every other day by removing 2ml and adding 2ml of fresh hES cell medium or fed everyday by removing 1ml and adding 1ml fresh hES medium. Human hES cells were ready to be split after 5-6 days in culture.

2.1.2.2 Enzymatic passage

Old hES cell medium was discarded, cells were then washed with PBS once and 600µl or 1ml (per well of a 6-well plate) of previously warmed (to 37^{o}C) TrypLE Express (GIBCO) were added. Cells were incubated at 37^{o}C for 5-7 minutes and shaken halfway through. After incubation, cells were neutralized with 400µl (per well of a 6-well plate) of fresh hES cell medium or 1ml of PBS. Thereafter, cells from all wells were collected and mixed well,

2x 10µl samples were taken for cell counting. The required amount of cells ($0.9x10^5$ or $1x10^6$ cells per well of a 6-well plate) was centrifuged at 1000*g* for 3-5 minutes and re-suspended with hES cell medium, sterile MEF CM or mTeSR1 medium (STEMCELL Technologies) to make a total volume of 2-2.5ml per well of a 6-well plate. Finally, the required volume of cells was placed in a well previously plated with a feeder layer, prepared ECM substrates (see below) or Matrigel™ (BD Biosciences) placed in plates for at least 1h. Finally, 10mM Rho Kinase (ROCK) inhibitor (Sigma-Aldrich) was added to each well of a 6-well plate.

2.1.3 Cell freezing and thawing

2.1.3.1 Freezing

Cells were initially treated the same way as when they were passaged except that, once the cells or patches were counted, they were spun down at 1000*g* for 5 minutes, re-suspended in half volume of corresponding ice-cold culture medium and half volume of 2x concentrated freezing medium (10% dimethyl sulfoxide, DMSO and 30% ES-FCS) in order to make final concentrations of 100 patches/ml or $2x10^6$ cells/ml. Final volumes were transferred into a cryotube and this was placed at -80^oC in a precooled freezing container rack (NALGENE) filled with isopropanol. Note that for longer storage, cryotubes were transferred onto liquid nitrogen tanks.

2.1.3.2 Thawing

Plates were first prepared with corresponding substrate. Cells were thawed in the water bath at 37^{o}C and quickly placed in a tube with 10ml PBS. Cells were centrifuged at 1000*g* for 4 minutes at 4^{o}C and re-suspended in 2ml of corresponding medium (plus 10mM ROCK inhibitor in the case of hES cells). Finally, plates were placed in the incubator at 37^{o}C in 5% CO_2.

2.1.4 Embryoid body preparation

2.1.4.1 Using Aggrewell™ plates

Prior to setting hES cells for embryoid body (EB) preparation, plates were set as follows: each Aggrewell (STEMCELL Technologies) was rinsed with 1ml PBS once, which was then replaced with 1ml mTeSR1 medium with 5mM ROCK inhibitor. The plate was centrifuged at 3000g for 1-2 minutes and placed in the incubator.

Human ES cells growing on Matrigel^{TM} in mTeSR1 medium were prepared in the same way as when they were passaged, except that following centrifugation at 1000*g* for 3-5 minutes cells were re-suspended at a final concentration of 4.8x10^{6} cells/ml (per Aggrewell) in mTeSR1 medium with 5mM ROCK inhibitor. This was then added to the prepared Aggrewell and the plate was centrifuged at 800*g* for 3 minutes and placed in the incubator.

Subsequently, media was replaced everyday for 3 days by doing a 75% mTesR1 medium change. On day 4, EBs were harvested: EBs were manually dislodged from Aggrewells

using a 5ml serological pipette and transferred to a 40µm cell strainer inverted over a 50ml tube, where they were washed with PBS several times. The cell strainer was turned over onto a new 50ml tube; using 4ml mTesR1 medium, EBs were flushed into the tube and counted and were then ready to be processed (see below: RNA isolation).

2.1.4.2 Using manual scraping

For this method, only hES cells growing on MEFs were used. First, half of the medium is replaced with fresh hES medium. Afterwards, using a 21G 1½" needle attached to a 1ml syringe, each whole well of a 6-well plate with about +100 colonies is scored 10 times in one direction and 10 times at 180^{o}. Patches from this mesh were scraped into smaller clumps of cells and transferred to one well of a low-attachment 6-well plate (Costar). Half of the hES media was carefully replaced every day for three days.

2.1.5 Production of monocytes and macrophages from hES cell line HUES2

EBs prepared by manual scraping (20-50 EBs) were collected on day 4 of culture and transferred into a well of an adherent, regular tissue culture 6-well plate adding 3ml of 'FACTORY' media (90% Advanced DMEM, 9% FCS, 1% GlutaMAX and 0.1% of 2-ME). Attached EBs would give rise to 'FACTORIES' that after 10-15 days start producing monocytes in suspension. Before harvesting monocytes, the volume of each well is brought to 6ml with warm 'FACTORY' media; monocytes are then carefully collected ($1.4x10^{6}$ monocytes per well of a 6-well plate) for further analyses (see RNA isolation) or re-suspended in 'MACROPHAGE DIFFERENTIATION' media (90% RPMI, 9% FCS and 1%

GlutaMAX) and transferred into a well of a 6-well plate for further differentiation into macrophages. After 24h, monocytes have attached and over the course of seven days gain a spindle shape; 50% of the media is replaced on day 3. After 7 days, macrophages can be treated for further analyses (see RNA isolation).

2.1.6 HeLa cells

Henrietta Lack's (HeLa) cells were grown at 37^{o}C in 5% CO_2 in sterile medium consisting of 90% DMEM supplemented with 10% heat-inactivated foetal bovine serum (FBS) and 2mM L-glutamine.

2.1.7 Bacterial cells

Bacterial cells NM554 were cultured in 2-YT media shaking or in agar plates overnight at 37^{o}C.

2.2 Preparation of ECM substrates from feeders

One day after irradiated feeder cells were prepared, the feeder medium was replaced with fresh hES cell medium and left in the incubator overnight. The following day, the CM was collected and cells were washed 3x with PBS, followed by incubation with Cell dissociation buffer enzyme-free Hanks'-based (GIBCO) or 0.5% w/v sodium deoxycholate (DOC) (1ml per well of a 6-well plate, 3ml for a T25 or 5-6ml for a T75) at 37^{o}C for 60-90

minutes or (depending on the progression on cell detachment, monitored under the bright field microscope). Finally, the surface was thoroughly washed with PBS until no cell traces were left behind and matrices were kept at 4^{o}C until they were used for hES cell culture or lysed for Western blot analysis (see below).

2.3 Western blotting

2.3.1 Preparation of lysates

2.3.1.1 Cell lysates

Cells were washed with PBS once and the required volume of ice-cold cell lysis buffer consisting of cell lysis solution (2% sodium dodecyl sulphate -SDS-, 1mM EDTA and 1mM EGTA at pH 7.4) supplemented with Protease Inhibitor Cocktail "Complete" (Roche) (one tablet per 10ml of cell lysis solution) was added. The lysed preparation was then scraped from the surface and transferred into a clean Eppendorf tube. Finally, samples were boiled at 100^{o}C for 3 minutes and stored at -20^{o}C.

2.3.1.2 ECM lysates

Prepared ECM substrates were washed with PBS once and lysed with the required volume of cell lysis buffer with aid of a cell scraper. The lysate was then transferred into a clean Eppendorf and re-suspended in 1ml of ice-cold ethanol, followed by freezing down in dry ice for 10 minutes and thawing at 4^{o}C. Thereafter, samples were centrifuged at

1000*g* for 20 minutes at 4°C. Pellets were re-suspended in the required volume of PBS solution containing Protease Inhibitor Cocktail "Complete" (Roche) and stored at -20°C. Finally, the protein concentration of ECM lysates was determined using a Nanodrop spectrophotometer (Thermo Scientific), protein A280 protocol. Each sample was measured three times and the mean value was obtained.

2.3.2 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

The required amount from each sample lysate was mixed with equal volume of 2x loading buffer containing DTT. These samples were heated at 95°C for 5 minutes, briefly centrifuged and loaded into a pre-cast gel 8-16% Precise protein gel (PIERCE) previously placed in a tank (Invitrogen) containing 1x Tris-HEPES-SDS running buffer (100mM HEPES, 3mM SDS and 500ml deionized water at pH 8 ± 0.5). The gel was run at 130V for approximately 45 minutes.

2.3.3 Wet gel transfer (for high molecular weight proteins)

Once the gel was run, it was transferred onto a polyvinylidene difluoride (PVDF) membrane previously activated with methanol and equilibrated with 1x wet blot transfer buffer (25mM Tris base, 190mM glycine and 20% methanol). The gel and membrane were assembled in a cassette in the following order from back to front: anode-oriented grid, sponge, three filter paper sheets, gel, activated PVDF membrane, three filter paper sheets, sponge and cathode-oriented grid. Finally, the fully assembled cassette was

placed into the wet transfer apparatus (Biorad), filled with 1x wet blot transfer buffer and run at 30V and 9mA overnight (19- 20h).

2.3.4 Semi-dry gel transfer (for low molecular weight proteins)

A similar procedure to that of wet gel transfer was followed except that the gel was transferred (onto a PVDF membrane previously activated with methanol) using a semi-dry transfer apparatus (Biorad). In addition, the gel and membrane were placed between 4 filter paper sheets at either side, previously soaked in 1x blotting buffer (10x stock = 30.3g Tris base, 144g glycine and 10g SDS for 1l deionized water). The apparatus was run at 15V for no longer than 1h.

2.3.5 Antibody (1^{o} and 2^{o}) incubation

After gel transfer, membrane was blocked in 30ml of 5% milk (Marvel dried skimmed milk) buffer for 1h at room temperature or at 4^{o}C overnight. Thereafter, the membrane was incubated with correspondent primary antibody (diluted in 1% milk buffer) overnight (no more than 18h) at 4^{o}C. On the other hand, the gel was stained with GelCodeTM blue stain reagent (Thermo Scientific) at room temperature for a couple of hours to verify the quality of the transfer.

The next day, the membrane was washed 5x (5 minutes each time) with 1x Tris-buffered saline (TBS) buffer (10x stock = 88g NaCl and 30g Tris-base in 1l of deionized water at pH7.4) at room temperature to eliminate any excess of primary antibody. After these

washes, the membrane was incubated with the correspondent secondary antibody (diluted in 1x TBS) at room temperature for no more than 1h.

2.3.6 Protein detection

Following incubation with the secondary antibody, the membrane was washed again as previously described. Finally, the membrane was incubated with ECL+ developing solution (GE Healthcare), placed in a cassette and developed on an Amersham Hyperfilm ECL (GE Healthcare).

2.3.7 Stripping and re-probing membranes

Membranes were retrieved from the cassette after development to wash them in 1x TBS 2x 5 minutes. Meanwhile, strip solution (50mM glycine, 0.5M NaCl, 0.1% Triton-X100 and 2% SDS at pH 2.4) was briefly heated up in the microwave. Membranes were then placed in 20 ml Falcon tubes with 5 ml of strip solution. Tubes were left rotating for 10 minutes at room temperature. The procedure was performed 3 times and membranes were briefly washed in 1x TBS, blocked in milk buffer and incubated with primary and secondary antibody and developed as it has been previously described.

2.4 Immunocytochemistry

Cells cultured on coverslips were washed 2x with PBS, fixed with 3.7% paraformaldehyde (PFA) fixation buffer for 10 minutes, washed and incubated with PBS at 4^{o}C. Once cells were fixed, coverslips were rinsed with PBS once and then 3x 5 minutes with PBT (2ml 10% Tween-20 and 98ml $Ca^{2+}Mg^{2+}$-free PBS), followed by blocking for 2h at room temperature in PBT+10% correspondent serum (blocking solution) (50µl per coverslip were added). Thereafter, the primary antibody diluted in blocking solution was added and left overnight at 4^{o}C. The following day, coverslips were rinsed with PBT five times: quick rinse, 5 minutes, quick rinse, 5 minutes and quick rinse. The secondary antibody diluted in blocking solution was then added and left at room temperature for 1h. After that, coverslips were washed five times as previously described, and incubated with DAPI stain solution for 5 minutes at room temperature. Finally, coverslips were mounted on slides, analysed on the fluorescence microscope and stored at 4^{o}C.

2.5 Flow cytometry analyses

Cells were washed with PBS once and centrifuged at 2000*g* for 5 minutes; the supernatant was removed and cells were fixed with 3.75% PFA in order to achieve a final concentration of 1 x10^{7} cells/ml and incubated at room temperature for 20 minutes. Afterwards, cells were washed with 1ml of 1 x Perm/wash buffer (BD Biosciences) and centrifuged at 1300*g* for 5 minutes. Cell pellets were re-suspended in 100µl 1x Perm/wash buffer and incubated at room temperature for 10 minutes. Permeabilized

cells were incubated with 20µl of corresponding antibodies or isotype controls (Human Pluripotent Stem Cell Transcription Factor Analysis kit from BD Biosciences):

- Antibodies: 20µl PE hNanog + 20µl PerCP-Cy5.5 Oct3/4 +20µl AlexaFluor®647 Sox2
- Isotype controls: 20µl PE isotype control + 20µl PerCP-Cy5.5 isotype control + AlexaFluor®647 isotype control

Samples were then analysed using a BD LSRII flow cytometer equipped with a 488nm and 633nm lasers.

2.6 Transfection of HeLa and HUES2 cells

Transfections were carried out using Lipofectamine 2000 according to manufacturer's instructions (Invitrogen).

2.6.1 Plasmids

HeLa cells were plated one day before at a density of $2x10^5$ cells per well of a 6-well plate whereas HUES2 were plated 2-3 days before at a density of $0.5x10^6$ cells per well of a 6-well plate. The day of transfection cells were at 80-90% confluence. Each plasmid (1µg) was incubated with Opti-MEM (GIBCO) reduced serum media (250µl) at room temperature for 5 minutes. Lipofectamine 2000 (Invitrogen) was also incubated with the same volume of Opti-MEM media for the same time. Thereafter, mixtures were

combined and incubated at room temperature for 20-30 minutes. Thereafter, each sample was gradually added, by slow dispensing, to the corresponding well of cells and gently shacked before placing them in the incubator at 37^{o}C. Culture media was changed after 4-6h and cells were harvested 24h after.

2.6.2 2'-O-methyl phosphorothioate oligonucleotides

HeLa cells were plated one day before at a density of 4.6x10^{5} cells per 22cm^{2} dish whereas HUES2 were plated 2-3 days before at a density of 1.2x10^{6} cells per 22cm^{2} dish. The day of transfection cells were at 40-60% confluence. Each oligo (150pmol) was incubated with Opti-MEM (GIBCO) reduced serum media (500µl) at room temperature for 5 minutes. Lipofectamine 2000 (Invitrogen) was also incubated with the same volume of Opti-MEM media for the same time. Thereafter, mixtures processed in the same way as for the transfection of plasmids. Note that in this case, cells were harvested at different time points that ranged from 10-48h.

2.7 RNA isolation

2.7.1 Preparation of cells

Cells growing in a monolayer were left to reach 90% confluence. Cell culture medium was aspirated and cells were lysed using 1ml of TriZOL reagent (GIBCO) per well of a 6-well plate. Cells in suspension (EBs and monocytes) were collected in 50ml Falcon tubes and

left to sink. Supernatant was removed and cells were lysed with 2ml of TriZOL reagent. Finally, samples were processed immediately or stored at -80^{o}C.

2.7.2 RNA extraction

Lysed cells in TriZOL reagent (1ml) were set at room temperature for 5 minutes, vigorously mixed with 200µl of chloroform, incubated at room temperature for 3 minutes and centrifuged at 12,000g for 15 minutes at 4^{o}C. In order to precipitate RNA, the aqueous phase was then transferred onto a new Eppendorf tube and mixed with 500µl of isopropanol, followed by incubation at room temperature for 10 minutes and centrifugation at 12,000g for 15 minutes at 4^{o}C. The obtained pellet was washed with 1ml 75% ethanol (EtOH) once and centrifuged at 7,500g for 5 minutes at 4^{o}C. EtOH was fully removed and pellets were briefly dried at room temperature and re-suspended in 20µl of RNase-free water. Samples were brought to 90µl, 10µl of DNase buffer (Roche) and 1µl of DNase enzyme (Roche) were added, followed by incubation at 37^{o}C for 30 minutes to 2h.

2.7.3 RNA re-precipitation

Equal volume of phenol-chloroform was added to the DNase-treated RNA samples. Samples were properly mixed, incubated at room temperature for 5 minutes and centrifuged at 12,000g for 15 minutes at 4^{o}C. The supernatant was then transferred onto a new Eppendorf tube, gently mixed with 2.5 x starting volume of 100% EtOH and 1/10 of 3M sodium acetate (NaOAc) pH5.5, incubated on ice for 15 minutes and centrifuged at

12,000g for 15 minutes at 4^{o}C. Afterwards, the RNA pellet was washed with 500µl 75% EtOH once, centrifuged at 7,500g for 5 minutes at 4^{o}C, briefly dried and re-suspended in 20µl of RNase-free water.

2.8 First-Strand complementary (c)DNA synthesis

Complementary DNA for each RNA samples was synthesized as follows: in a fresh microtube, 50-250ng (1µl) random hexamers (Roche), 500ng RNA sample (RNA concentration and quality was previously measured using a Nanodrop) and 1µl 10mM deoxyribonucleotide triphosphate (dNTP) mix were combined. This mixture was incubated at 65^{o}C for 5 minutes and on ice for 1 minute. Afterwards, each sample was mixed with 4µl 5x First-strand buffer, 1µl 0.1M DTT, 0.5µl RNaseOUT (Roche) and 1µl SuperScript III Reverse Transcriptase (Invitrogen). Note that negative controls (-RT) were prepared by replacing the reverse transcriptase with RNase-free water. Finally, samples were incubated in the PCR machine using the following programme: 25^{o}C for 5 minutes, 50^{o}C for 1h, 70^{o}C for 15 minutes and 4^{o}C overnight (if required).

2.9 Real-time quantitative polymerase chain reaction (qPCR)

Real time PCR measures the accumulation of PCR product during the exponential phase of the PCR reaction and facilitated the quantification of endogenous U1 snRNA and vU1 snRNA transcript levels. Typical volume of 9µl from master mix (primer stock -1:10 dilution of 100µM forward and reverse primers-, RNase-free water and SYBR green -

QIAGEN-) was combined with 1µl of corresponding cDNA sample. The qPCR was carried out on a Rotorgene 6000 thermal cycler (Corbett). Thermal cycling was performed as follows: 95^{o}C for 15 minutes, followed by 40 cycles at 94^{o}C for 15 seconds, 57^{o}C for 15 seconds and 72^{o}C for 30 seconds. Since SYBR green binds all double stranded (ds) DNA molecules (including artefacts such as primer dimers or non-specific products that may arise from cross hybridization of primer sequences) all amplification were followed by a 10 minutes melt cycle to confirm that a single PCR product was generated per reaction. This involves ramping the temperature from 50^{o}C to 99^{o}C by 1^{o}C every 5 seconds. Each amplicon denatures at a specific temperature and thus a single amplicon should yield a clear peak when the melt curve is plotted at the end of the run. For the quantification of snRNA genes in different cell lines, quantification was performed by the relative standard curve method. A standard curve was constructed for each primer pair using 1/5 step-wise dilutions of human genomic DNA, and all other samples were quantitated relative to this, giving a 'No. copies/ µl' value for each sample. At the end of the run a threshold value is set for all samples and standards over the linear range of the PCR reaction. The Rotorgene software converts the cycle number (Ct) for each standard to copies/ml and generates a standard curve. The ΔΔCt values for each sample are quantitated relative to this standard. For transfections of oligos or plasmids in HeLa or HUES2 fold changes in gene expression were analysed using the Ct method of quantification, whereby samples of transfected cells were compared to mock-transfected cells values after individual samples were normalized to internal 7SK levels.

2.9.1 Primers used

	Forward	Reverse
Name	**U1 snRNA nascent primers**	
pre-U1	5'-gaaactcgactgcataatttgtggtag-3'	5'-cttggcgtacagtctgtttttgaaactc-3'
pre-vU1.1,10	5'-gaaactcgactgcataatttgtggtag-3'	5'-cacatacaatctgcgttttccagcac-3'
pre-vU1.2a, 2,11	5'-gaaactcgactgcataatttgtggtag-3'	5'-ggagcgtacggtctatttttgaaacc-3'
pre-vU1.3-5,12,20	5'-gaatctcgactacataatttgtggcag-3'	5'-cagcaaacaaaatccacttttaaaacgac-3'
pre-vU1.6	5'-gaaacttgactgtataatttgtggcagtg-3'	5'-gcagcaaacaaaacctatttttaaaacgac-3'
pre-vU1.7,9	5'-gaaactcgactgcataatttgtggtag-3'	5'-ccttataggggagtctatttttttcattac-3'
pre-vU1.8	5'-gaaactcgactgcataatttgtggtag-3'	5'-catgcagtctatctttccacactgc-3'
pre-vU1.13-16,19	5'-ggaactcgactgcataacttgtgatagt-3'	5'-cggctcacctactattcgcaccac-3'
pre-vU1.17	5'-gaaactcgactgcataatttgtggtag-3'	5'-ctttgcataccaactgctttttgcactg-3'
pre-vU1.18	5'-gaaactcgactgcataatttgtggtag-3'	5'-cccttgctgtgtagactattctttaaac -3'
	7SK coding primers	
7SK	5'-ctgatctggctggctaggcggg-3'	5'-gaagaccggtcctcctctatcgg-3'
	U1 snRNA and vU1 snRNA-encoding primers	
m-U1	5'-caggggagataccatgatcacgaag-3'	5'-ggtcagcacatccggagtgcaatgg-3'
m-vU1.8	5'-gaagctaattcgtgcaacttccc-3'	5'-cgggggaaaagagcgaacgcag-3'
m-vU1.2a	5'- tttacttggcaggggagataacgt -3'	5'- cagggaaaagcacggacacagc-3'
m-vU1.3	5'-tctggcagaagaaatgttatgatc-3'	5'-gggaaagcacaaacagttcc-3'
m-vU1.20	5'-gtttatctggcagaagaaatgttatg-3'	5'-ggtcagcacatccagagtaaaattgc-3'
	Cell markers primers	
OCT4	5'-cgaaagagaaagcgaaccag-3'	5'-aaccacactcggaccacatc-3'
CD14	5'-atcgaccatggagcgcgcgt-3'	5'-aaggcttcggaccagtcggg-3'
CD68	5'-gtcatggaaatgccacggttc-3'	5'-aggttgggctaggactcgg-3'

	Nuclear snRNA knockdown (2'O-Methyl Phosphorothioate oligonucleotides)
Scrambled control	5'-gaaggaactagcgtacgacgg -3'
vU1.8	5'-guaaacacgatttttcttcugcu-3'
vU1.20	5'-aacauttcttctgccagauaaac-3'

2.9.2 Statistical analyses

Where indicated, statistical analysis was performed using GraphPad Prism (version 5.01) for Windows (GraphPad Software, San Diego California USA, www.graphpad.com). Statistical tests used included one-way analysis of variance (ANOVA) with Bonferroni's posttest and unpaired two-tailed Student's t-test. P values were classified as not significant ($p>0.05$) or significant (*; $p= 0.001-0.05$).

2.10 Molecular cloning

2.10.1 Polymerase Chain Reaction (PCR)

DNA products for cloning experiments were amplified by PCR. Each reaction uses 2x of a Taq-containing mastermix (HotStarTaq Mastermix, QIAGEN), which consisted of a heat-activated Taq enzyme, dNTPs and $MgCl_2$-containing reaction buffer. The corresponding amount of DNA (100ng of gDNA or 10ng of plasmid) was used as template plus 1µl of a 10pmoles/µl primer stock (forward and reverse primers) in a total volume of 20µl. For all cloning experiments, 2U of a proofreading enzyme, usually PFU Turbo, was added per reaction after the first step of the thermal incubation. Thermal cycling was performed as

follows: 95°C for 15 minutes, 5 cycles of 94°C for 15 seconds, 45°C for 15 seconds, 72°C for 40 seconds, 25 cycles of 94°C for 15 seconds, 57°C for 15 seconds, 72°C for 40 seconds, followed by 72°C for 7 minutes. DNA products were then analysed by electrophoresis on 1-2% agarose gels (as described below).

2.10.2 DNA digestion with restriction enzymes

The required amount (500ng - 1µg) of plasmid DNA or PCR products was incubated with 1/10 corresponding buffer and 1 unit (U) of corresponding restriction enzyme (never exceeding 10% of final volume) at 37°C for 1h (plasmid DNA) or 4h (PCR products). Note that in the case of PCR products, halfway through this incubation, an extra 1µl of restriction enzyme was added.

For the generation of blunt ends, digested DNA fragments were treated with Klenow Fragment (NEB) or Mung Bean Nuclease (NEB): Klenow was used in combination with dNTPs for filling in overhang ends according to manufacturer's instructions. Previously digested DNA (1µg) was mixed with 10x reaction buffer, 2mM 4dNTP mix and 1µl Klenow fragment. This mix was incubated at 37°C for 10 minutes and heat-inactivated at 75°C for 10 minutes. Thereafter, full sample was run in a gel and purified as previously described.

MBN was used to remove 5' overhangs according to manufacturer's instructions. Previously digested DNA (1µg) was mixed with 10x MBN reaction buffer and 1µl MBN. This mix was incubated at 30°C for 30 minutes, followed by phenol/chloroform extraction and EtOH precipitation (see below).

2.10.3 DNA phenol/chloroform extraction and EtOH precipitation

Samples were mixed 1/10 3M sodium acetate pH5.2 and equal volume of phenol/chloroform pH 8.0. Thorough mixing and centrifugation at 12,000g for 5 minutes followed this. Supernatant was recovered and mixed vigorously with 2.5x starting volume of EtOH and centrifuged at 12,000*g* for 5 minutes. Pellets were then re-suspended with appropriate volume of water.

2.10.4 DNA gel electrophoresis

Sizing of digested DNA products was carried using agarose gel electrophoresis. DNA samples were combined with 2-5x OJ loading dye and loaded on 1-2% agarose gel containing 0.002% Ethidium bromide (10mg/ml stock) along with the 100bp or 1Kb DNA ladder as a sizing marker. The gel was then run at 100mA for 1-2h in 1xTBE buffer (2l= 21.6g Tris, 11g Orthoboric acid and 2g EDTA) containing 1% Ethidium bromide. DNA bands were visualized under UV light; low UV light and minimal exposure was done when DNA would be used in further cloning experiments.

2.10.5 DNA gel extraction

Corresponding DNA fragments were excised from the agarose gel using a scalpel. Gel slices were processed using the QIAEX II® gel extraction kit (QIAGEN), following the instructions from the manufacturer. The agarose from the gel slices is melted at 50^{o}C in the presence of high salt concentration and pH≤7.5, which allows the dissociation of DNA

fragments and aids the binding of DNA to the QIAEX II silica particles. Thereafter, one wash in high salt and two in an EtOH-based buffer remove traces of agarose and salts, respectively. Finally, DNA is purified from the silica particles by elution with low salt concentration and pH 7-8.5.

2.10.6 DNA ligation

DNA fragments with overlapping or blunt ends were incubated with 1x ligation buffer and 1U of T4 DNA ligase (Roche) at room temperature for at least 5 minutes (overlapping ends) or 2h to overnight (blunt ends). Usually, the ratio of insert to vector was 50:10ng; these were incubated in a total reaction volume of 10µl.

2.10.7 Transformation of competent cells with plasmid DNA

Competent NM554 bacterial cells kept at -80^{o}C were thawed on ice. Afterwards, half volume of the ligation mix (typically 10µl) was added to the competent cells and incubated on ice for 10 minutes. This was followed by the introduction of plasmid DNA into the competent cells by heat-shock at 37^{o}C for 5 minutes. Transformed cells were streaked under sterile conditions on a Luria-Bertani (LB) agar plate containing 100µg/ml ampicillin (AMP) and incubated at 37^{o}C overnight.

2.10.8 Mini-prep

Single colonies from freshly streaked plates were grown in 2.5-5ml of LB or 2x YT media containing 100µg/ml AMP for ≥ 5h at 37^{o}C with vigorous shaking. Bacterial cultures were

processed using the QIAprep® Miniprep kit (QIAGEN), following the instructions from the manufacturer. Bacterial cells are lysed in SDS, which solubilizes the proteins and phospholipids of the bacterial cell membrane and NaOH, which denatures the chromosomal and plasmid DNA and proteins released by lysis. The subsequent addition of high salt concentration neutralizes the lysate at the same time that it denatures chromosomal DNA, cell debris and proteins and re-natures plasmid DNA. This is then purified on a silica membrane, washed and eluted as described before (see DNA gel extraction).

2.10.9 Maxi-prep

In order to prepare higher amounts of the desired plasmid DNA, starter cultures (see Mini-prep) were diluted 1:1000 in 100ml of LB medium containing 100x AMP and grown in a conical flask (1l) for 16h at 37^{o}C with vigorous shaking. Bacterial cultures were processed using the QIAGEN Plasmid Purification Maxi kit (QIAGEN), following the instructions from the manufacturer. Processing was similar to that of mini-prep except that plasmid DNA was purified on an anion-exchange resin and eluted at high salt concentration followed by concentration and desalting with medium salt buffers and precipitated with isopropanol.

2.10.10 DNA sequencing

In order to confirm the incorporation of the DNA fragments of interest into corresponding vector, typically 20µl of purified plasmids from mini-prep or maxi-prep along with corresponding primers (diluted 1:10) were sent for sequencing at Source BioScience UK limited, Department of Biochemistry, University of Oxford, UK. Sequencing files were then analysed using the software Chromas version 1.45 (Conor McCarthy. School of Health Sciences, Griffith University Queensland, Australia).

Chapter 3| ESTABLISHING FEEDER-FREE CULTURE FOR HUMAN EMBRYONIC STEM CELLS

3.1 Introduction

Human ES cells were initially successfully derived from the ICM of pre-implantation blastocysts in the presence of both a monolayer of mitotically inactivated MEFs, which acted as “feeders” for the hES cells, and culture media containing serum (Thomson et al., 1998). These conditions allow the maintenance of pluripotent hES cells in culture for prolonged periods of time.

However, the potential use of hES cells to treat degenerative diseases or repair damaged or non-functional tissues has gradually led to the development of hES culture techniques that avoid the use of xenogeneic components. Alternative culture conditions include the use of fibroblasts from human origin and regular hES culture medium, composed of a base medium of Dulbecco’s Modified Eagle Medium: Nutrient mixture F12 (DMEM/F12) supplemented with a serum-free formulation known as knock-out serum replacement (KSR). It has been demonstrated that under these conditions, human foreskin fibroblasts (HFFs) from adult or foetal tissues (Richards et al., 2003, Koivisto et al., 2004, Ilic et al., 2009) as well as fibroblasts derived from hES cells (autogeneic feeders) (Stojkovic et al., 2005, Saxena et al., 2008, Chen et al., 2009) are able to support hES cell self-renewal and pluripotency, albeit not necessarily to the level of maintenance that MEFs provide (Table 3-1).

For the purpose of this project, I aimed to measure the expression patterns and investigate the potential function(s) of vU1 snRNAs in hES cell maintenance and differentiation. As vU1 snRNAs expression patterns are known to vary in different human cell types (O'Reilly et al., 2012), there was a strong possibility that human feeders could

interfere with the genetic analysis of the hES cells. In addition the mouse feeder system also posed a problem, as mice are known to express a vU1 snRNA, known as U1b, during embryogenesis and it is yet unclear whether they also express homologues of the human vU1 snRNA genes. To eliminate any potential contributions these cells may have on the particular analysis used, it was of paramount importance that a defined and reproducible feeder-free system was pursued.

Fibroblast feeders		Culture media	Supplement	hES cell maintenance			Reference
				No. of passages in culture	Pluripotency & normal karyotype	Potential to differentiate	
MEFs		DMEM/F12	20% FCS	40+	+	+	(Thomson et al., 1998)
Human	**Foetal**		20% KSR bFGF	20-25+	+	+	(Richards et al., 2003, Koivisto et al., 2004, Rajala et al., 2007, Ilic et al., 2009, Chin et al., 2010)
	Adult			3-30+			
Autogeneic				17-30+	+	+	(Stojkovic et al., 2005, Saxena et al., 2008, Chen et al., 2009)

Table 3-1. Human ES cell culture systems that use feeder layers.

Summary of the culture conditions that require the use of feeders for hES cell culture. The ability of these feeders to maintain hES cells in culture is indicated by the number of passages in culture, expression of pluripotency markers, presence of a normal karyotype and the retention of their capacity to differentiate. DMEM/F12, Dulbecco's Modified Eagle Medium: Nutrient mixture F12; FCS, foetal calf serum; KSR, knock-out serum replacement; bFGF, basic fibroblast growth factor.

In order to address this question, it was essential to gain an understanding of the factors that support hES cells. It is known that mammalian somatic stem cells (also referred to as 'adult stem cells') are embedded in a specialized microenvironment *in vivo* known as the stem cell niche, which appears to be responsible for maintaining the stem cell pool. The stem cell niche not only provides a dynamic anatomical location for stem cells but creates a physiochemical environment that hosts components such as supporting cells, ECM and secreted factors that ultimately influence the balance between self-renewal and differentiation (Walker et al., 2009).

In the case of hES cells, there is no equivalent stem cell niche *in vivo* because hES cells only exist as an *in vitro* phenomenon; they are obtained over a period of time from the *ex vivo* culture of the ICM, that is manually isolated from a pre-implantation embryo.

Nevertheless, the feeders that support hES cells in culture *in vitro* ultimately provide what herein is referred to as an *in vitro* hES cell niche. Given that this niche is providing the conditions that allow the maintenance and self-renewal of hES cells, a reasonable approach would be to dissect its components and recreate these conditions in culture.

A study by Klimanskaya and colleagues demonstrated that it was possible to derive and maintain hES cells in culture on the ECM deposited by MEFs (Klimanskaya et al., 2005). However, they did not investigate the composition of this ECM.

In addition, separate studies have shown that individual ECM proteins like collagen IV, fibronectin, laminin and vitronectin, or combinations of them are able, to a certain extent, to maintain hES cell in culture (Table 3-2). Furthermore, the ECM proteins tenascin-C (also known as tenascin) and osteopontin have been shown to positively or

negatively regulate the stem cell pool within the neural and hematopoietic stem cell niche in the mouse, respectively (Garcion et al., 2004, Stier et al., 2005).

In the light of these findings, more recent studies have identified some ECM proteins and integrins (the cell membrane receptors for ECM proteins) in hES cultured on feeders (Braam et al., 2008, Miyazaki et al., 2008, Evseenko et al., 2009). They found that the ECM proteins and integrins were expressed at various degrees in the feeders and hES cells (Braam et al., 2008, Miyazaki et al., 2008, Evseenko et al., 2009). Moreover, ECM proteins have also been detected in conditioned media (CM) from MEF and human fibroblasts (Lim and Bodnar, 2002, Chin et al., 2007, Prowse et al., 2007). However, none of these studies fully characterized the ECM composition that supports hES cells in culture. It is clear that the ECM is a key component of the stem cell niche but little is known about its full composition in the *in vitro* hES cell niche or its contribution in hES cell maintenance.

In this chapter my aim was to establish a hES cell culture system that does not require the use of feeders such that I could analyse the expression of human vU1 snRNAs in these cells and their differentiated progeny without the risk of RNA contamination from feeders. In order to achieve this, I investigated the protein composition of the ECM substrate provided by supporting and non-supporting feeders and ultimately isolated an ECM substrate suitable for hES cell culture.

ECM protein substrate	Culture media	Growth factor supplement	hES cell maintenance	Reference
Collagen IV	MEF-CM	hbFGF (4ng/ml)	+*	(Xu et al., 2001)
	DMEM/F12	bFGF (100ng/ml) and N2/B27	-	(Liu et al., 2006)
Fibronectin	DMEM/F12	hbFGF (4ng/ml)	+*	(Xu et al., 2001)
		bFGF (100ng/ml)	++	(Liu et al., 2006)
Laminin	MEF-CM	-	++	(Xu et al., 2001)
	NC-SFM	hbFGF (80ng/ml)	+	(Li et al., 2005)
	Defined SFM	-	++	(Fletcher et al., 2006)
	DMEM/F12	bFGF (100ng/ml)	-	(Liu et al., 2006)
Vitronectin	mTesR1	-	++	(Braam et al., 2008)
	DMEM/F12	bFGF (100ng/ml)	-	(Liu et al., 2006)
Collagen IV+ Fibronectin	DMEM/F12	bFGF (100ng/ml)	++	(Liu et al., 2006)
Fibronectin+ Vitronectin			++	(Liu et al., 2006)
Collagen IV, Fibronectin, Laminin+ Vitronectin	TeSR1	bGFG and TGFβ	+	(Ludwig et al., 2006b)

Table 3-2. Human ES cell culture on ECM protein substrates.

Summary of culture conditions using ECM protein substrates and their ability to maintain hES cells in culture. MEF-CM, MEF conditioned media; NC-SFM, non-conditioned serum free media; hbFGF, human basic fibroblast growth factor; TGF, transforming growth factor. ++, + and – indicate the degree of hES cell maintenance (hES cell morphology, pluripotency and differentiation potential).* low OCT4 expression.

3.2 Different feeders support hES culture differently

Previous work from Prof H. J. Mardon's laboratory (NDOG, University of Oxford) showed that the human ES cell line H1 was routinely passaged on MEFs whereas the HFF cell line Hs27 did not support these hES cells (Brook F., unpublished observation). In order to make sure those culture conditions were not restricted to H1 hES cell line, I tested whether, in addition to H1, the in house-derived hES cell line OxF1 (Brook et al., 2010) could also grow on MEFs and Hs27 feeders. Both H1 and OxF1 hES cells were able to grow on MEFs and retained their morphology of compact and undifferentiated colonies (Fig. 3-1). Meanwhile H1 and OxF1 maintenance was very poor when plated on Hs27 feeders; hES cells were found to be more spread and did not survive the first passage (Fig. 3-1).

In conclusion, my results show that different hES cell lines are maintained similarly in a given feeder but, in agreement with the literature (Table 3-1), different feeders provide different levels of support to hES cells. I next assessed whether the differences in ECM composition from MEFs and Hs27 accounts for the differences in hES cells support.

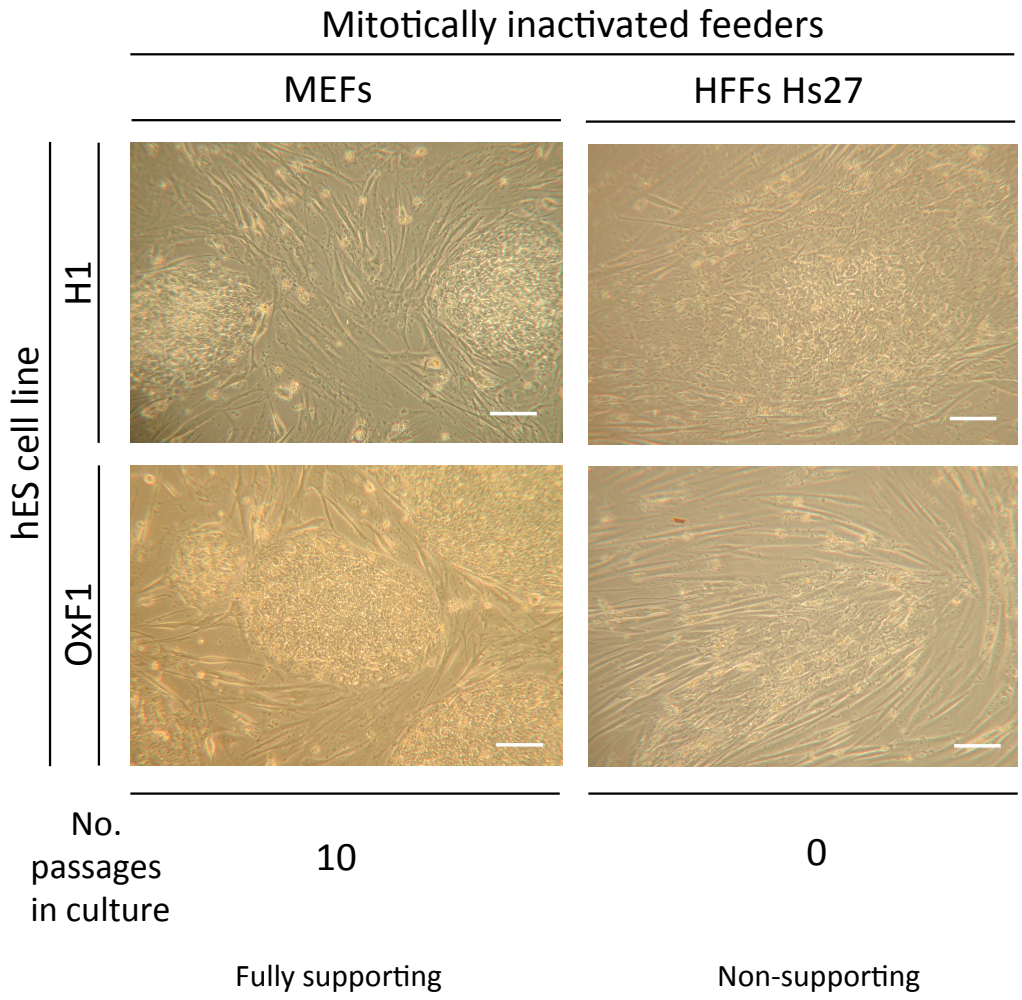


Figure 3-1. Human ES cells grown on two different mitotically inactivated mouse or human feeders.

Phase contrast microscopy analysis of hES cell lines H1 and OxF1 grown on MEFs and human foreskin fibroblasts (HFFs) Hs27. Morphology and number of passages in culture shows them to be fully supporting and non-supporting, respectively. Scale bars: 100µm

3.3 The composition of ECM proteins varies among feeders

Previous studies have identified by immunofluorescence staining expression of some ECM proteins in hES cells and the feeders they are cultured on, MEFs and supporting HFFs (Braam et al., 2008, Evseenko et al., 2009). Nevertheless, the resolution of this approach makes it difficult to discriminate between the ECM proteins that are expressed in the intracellular space and those expressed in the extracellular space. There is only one study that has analysed the expression of the ECM deposited by supporting HFFs, but it only detected four ECM proteins by mass spectrometry: collagen type 6 β3, fibronectin, keratin type 2 and laminin β1 (Meng et al., 2010). All these studies provide qualitative results only and lack a direct comparison with the expression of ECM proteins in non-supporting feeders.

It therefore appears crucial to analyse the protein composition of the ECM deposited by fully supporting MEFs and compare it to that from non-supporting Hs27. In addition, another HFF cell line, CCD-1112SK (here referred to as CCD) was also included in the analyses; HFFs CCD are able to support the growth but not the derivation of hES cells (personal communication, Brook F.), therefore they are considered as 'relatively' supporting feeders.

For the purpose of analysing the ECM from those feeders, a monolayer of mitotically inactivated feeders (0.17 $x10^6$ cells seeded) was treated with a fairly mild detergent (0.02M NH_4OH) to allow partial removal of cells (only a trace of cell membrane is left behind). This treatment ensured that the ECM proteins deposited by feeders are mostly intact for further analyses. The prepared ECMs were then lysed and subjected to Western

blot analyses (Fig. 3-2). MEFs were used as a positive control for mouse samples and endometrial stromal cells (890S) for human samples except for collagen IV and vitronectin blots, where lysates of VB6 cells (cell modified from oral squamous cell carcinomas, over-expressing the integrin receptor αvβ6) or placenta were used. Regular hES cell culture medium was used as a negative control. Expression of β-actin was used to indicate the quality of the ECM preparation.

The analyses showed that the ECM protein collagen IV is expressed at a higher level in MEF ECM and CCD ECM than in Hs27 ECM (Fig. 3-2A). Laminin-1 is expressed at relatively equal levels in the CCD and Hs27 ECM but at lower level than in MEF ECM (Fig. 3-2A). Conversely, two tenascin-C isoforms show the highest expression in CCD ECM and the lowest in MEF ECM (Fig. 2A). Fibronectin was detected at ~100kDa, which may correspond to the central cell-binding fragment of fibronectin generated by proteolytic cleavage; similar levels in MEFs, CCD and Hs27 ECM were observed (Fig. 3-2B). Note that two different fibronectin antibodies were used to recognize human and mouse samples given that an antibody that could recognize both samples was not available.

In the case of osteopontin, its 32kDa matrix metalloproteinase (MMP)-cleaved fragment is detected at higher levels in CCD and Hs27 ECM than in MEF ECM; but a higher band is preferentially detected in the mouse samples, which may correspond to a modified version of that fragment (Fig. 3-2C). For vitronectin, only shorter isoforms of the glycosylated full-length protein dimer (composed of a 65kDa and 10kDa monomers) are detected in the human samples (Fig. 3-2C); CCD ECM and Hs27 ECM express the non-glycosylated isoform of the full-length vitronectin dimer (~55kDa), whereas the positive control (i.e. placenta lysate) expresses a monomer of the glycosylated full-length

vitronectin dimer (65kDa) (Fig. 3-2C). Note that the antibody used for vitronectin could not detect mouse isoforms.

This differential expression of ECM proteins in different feeders already highlights a certain pattern of expression of ECM proteins in fully supporting, relatively supporting and non-supporting feeders. With the exception of fibronectin, which is expressed equally in all feeders, expression of collagen IV, laminin-1, tenascin-C and osteopontin fragments show a reversed pattern between MEFs and human feeders (Table 3-3).

In conclusion, these analyses of the protein composition of ECM deposited by MEF, CCD and Hs27 feeders provided us with some indication of which ECM proteins and what relative levels of them could be important for the support of hES cells *in vitro*.

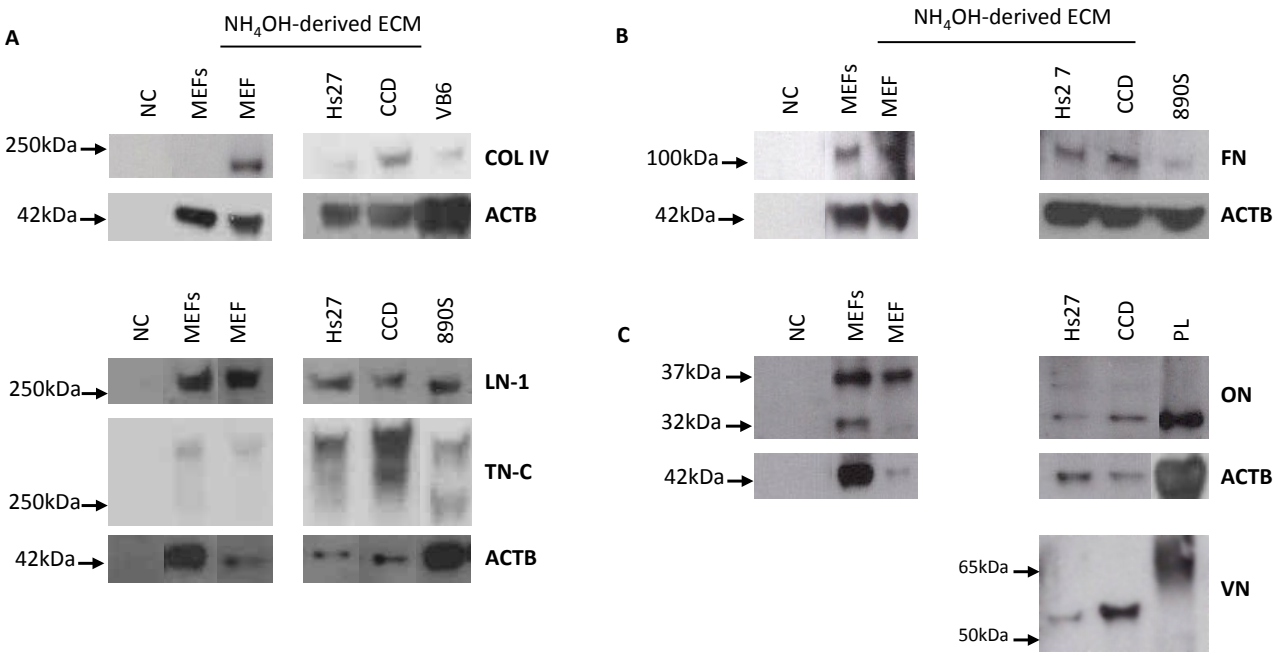


Figure 3-2. Detection of ECM proteins in NH4OH-treated ECM preparations from feeders.

Western blot analyses of MEF, Hs27 and CCD ECM preparations with 0.02M NH_4OH. Membranes were probed with antibodies (indicated on the right) for: (A) collagen IV (COL IV), laminin-111 (LN-1) and tenascin-C (TN-C); (B) fibronectin (FN) and (C) osteopontin (ON) and vitronectin (VN). The protein β-actin (ACTB) was used as a control for the presence of other cells constituents. Human ES cell medium was used as negative control and lysates of MEFs, VB6 modified oral squamous carcinoma cells, 890S endometrial stromal cells and placenta as positive controls. Molecular sizes are indicated on the left.

			Feeders		
			MEFs	**CCD**	**Hs27**
ECM proteins	**COL IV**		+++	++	+
	LN-1		+++	++	++
	TN-C		+	+++	++
	FN		+++	+++	+++
	ON	**MMP-cleaved fragment**	-	+++	+++
		Modified MMP-cleaved fragment	+++	-	-
	VN		n/a	+++	+

Table 3-3. Expression of ECM proteins in NH_4OH-treated ECM preparations from feeders.

Summary of the qualitative protein levels from Western blot analyses (Fig. 2) for the expression of ECM proteins collagen IV (COL IV), laminin-111 (LN-1), tenascin (TN), fibronectin (FN), osteopontin (ON) and vitronectin (VN). MMP, matrix metalloproteinase; +++, ++, +, - indicate high, medium, low or no expression, respectively; n/a, not applicable.

3.4 Conditions for the preparation of ECM proteins from feeders

In order to move towards a feeder-free system that could still provide the same support to hES cell maintenance as MEFs, I tested whether we could recreate the conditions of the hES cell *in vitro* niche without the feeders. This system will allow performing genetic analyses on hES cells without having contamination from feeders.

For this purpose, the ECM derived from feeders had to be completely cleared of any cellular residues. As the presence of β-actin in the NH_4OH–treated feeders indicated contamination with other cell constituents, which could include DNA or RNA from feeders and interfere with further genetic analysis, I decided to test more stringent methods. These were first tested on MEFs and involved treatment with 0.5% w/v sodium deoxycholate (DOC), 4mM ethylenediaminetetraacetic acid (EDTA), 0.5% w/v DOC plus 4mM EDTA and a commercially available (Invitrogen) cell dissociation buffer (enzyme-free solution of salts, chelating agents, and cell-conditioning agents in calcium-free and magnesium-free Hank's Balanced Salt Solution; herein referred to as CDB). Cell detachment was monitored under the phase contrast microscope; cells were fully removed after 15 (DOC method), 50 (EDTA method), 15 (DOC + EDTA method) and 80 (CDB method) minutes. Furthermore, the protein concentration achieved with these four methods was 0.11, 0.07, 0.08 and 0.12 mg/ml, respectively. Given the higher protein yield, DOC and CDB methods were used for further analyses.

In order to test the robustness of these methods, in addition to MEFs, CCD feeders were selected for ECM preparation, given that their ECMs had previously shown a higher expression of ECM proteins compared to Hs27 ECM (Fig. 3-2). Following DOC or CDB

treatment, ECM preparations were subjected to Western blot analysis against selected ECM proteins (Fig. 3-3). These analyses showed that DOC was the most effective method in terms of removing MEFs and CCD cellular residues, as indicated by the lack of β-actin expression. Moreover, I detected expression of collagen IV (Fig. 3-3A) and laminin-1 (Fig. 3- 3B) in CCD ECM and tenascin-C (Fig. 3-3B) in MEF and CCD ECMs prepared with DOC. In addition, other ECM proteins were tested but results were inconclusive (data not shown). Given the more stringent conditions used for ECM preparation, it is likely that such ECM proteins were lost during ECM preparation.

In conclusion, these results show that it is possible to generate an ECM substrate from feeders that is devoid of feeders themselves and that still preserves, to certain extent, its ECM composition.

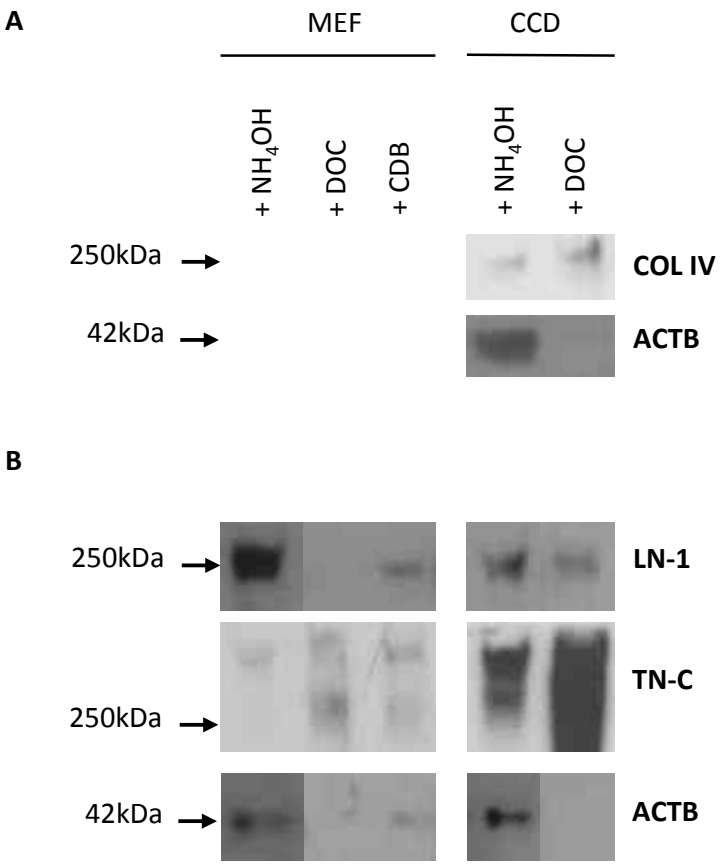


Figure 3-3. Expression of ECM proteins in MEFs and CCD treated with NH_4OH, DOC and CDB.

Western blot analyses of MEF and CCD ECM preparations with 0.02M NH_4OH, 0.5% w/v sodium deoxycholate (DOC) or a commercially available cell dissociation buffer (CDB). Membranes were probed with antibodies (indicated on the right) for: collagen IV (COL IV) (A) and laminin-111 (LN-1) and tenascin-C (TN-C) (B). Note that collagen IV could not be detected in MEF ECM preparations. The protein β-actin (ACTB) was used as a control for the presence of any remaining cells. Molecular sizes are indicated on the left.

3.5 ECM proteins derived from feeders support hES cell culture

Since I was able to detect expression of ECM proteins on the feeder-free ECM preparations, I then cultured hES cells on ECM substrates, prepared with DOC, in the presence of conditioned media (CM) from feeders (regular hES culture medium harvested after being in contact with feeders for ~24 h). The ability of these conditions to support hES cells was assessed by the analysis of hES morphology, characteristic high nuclear to cytoplasmic ratio (Fig. 4), and expression of pluripotency markers (Fig. 3-5 and 3-6).

OxF1 hES cells on MEF- or Hs27-derived ECM and MEF CM were able to grow for 11 passages while retaining typical hES cell morphology (Fig. 3-4A and D) and expression of key pluripotency markers such as the nuclear transcription factors OCT4, NANOG, SOX-2 and cell surface antigen TRA-1-60, as detected by immunocytochemistry (Fig. 3-5) or flow cytometry analyses (Fig. 3-6). On the other hand, OxF1 grown on MEF ECM or Hs27 ECM plus Hs27 CM were maintained under these conditions for only 1 passage and did not present characteristic hES morphology; they showed high content of distinct cytoplasmic granules (Fig. 3-4B and E). Moreover, OxF1 growing under these conditions lost expression of the pluripotency marker and nuclear transcription factor SOX-2, as shown in one flow cytometry analysis (Fig. 3-6). Note that further supplementation of Hs27 CM with 4ng/µl bFGF did not recover the hES cell phenotype (Fig. 3-4C and F). Thus, it appears that growth and maintenance of OxF1 cells is unsuccessful with ECM substrate alone and seems to require an additional component from the CM derived, in particular, from MEFs.

I next tested whether MEF CM alone was sufficient to support hES cells in culture. OxF1 grown on MEF ECM and regular hES culture medium that had not been conditioned by MEFs showed very poor attachment and no proliferation (Fig. 3-7A), whereas OxF1 seeded on uncoated surfaces and MEF CM formed 3-D structures that resembled differentiating EBs (Fig. 3-7B). As expected, OxF1 grown on MEF ECM and MEF CM showed typical hES cell morphology and were able to proliferate in culture (Fig. 3-7C).

This data indicates that culturing hES cells on ECM or CM alone is not sufficient and that and additional component(s) from hES cell culture medium that has been previously conditioned by supporting feeders is critical for maintaining hES cells in culture.

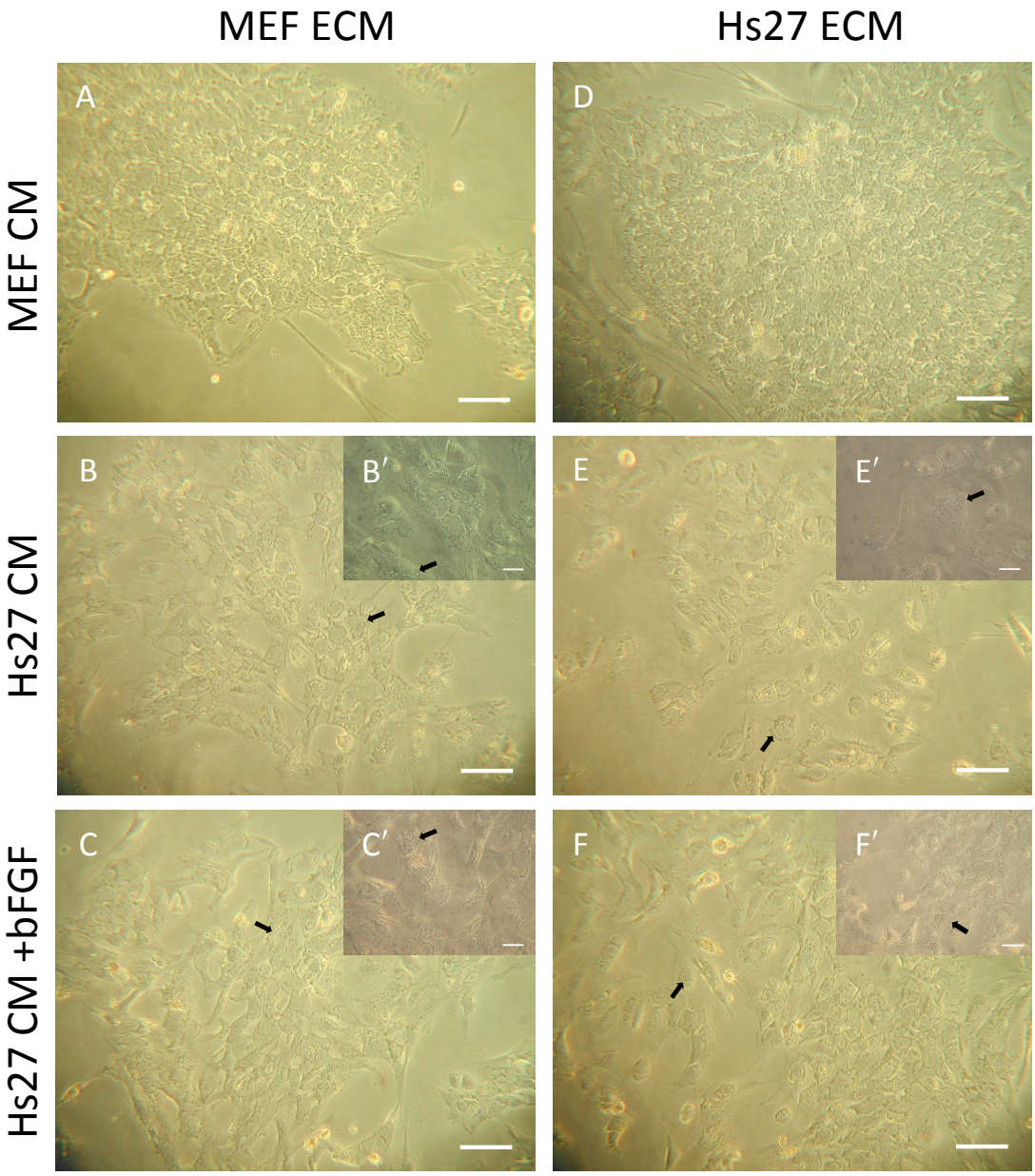


Figure 3-4. Human ES cells OxF1 grown on ECM plus CM from fully supporting and non-supporting feeders.

Phase contrast microscopy analysis of hES cell line OxF1 grown on MEF (A-C) and Hs27 (D-F) ECM plus MEF CM (A, B), Hs27 CM (B, E) or Hs27 CM supplemented with 4 ng/µl bFGF (C, F) conditioned media (CM). The black arrows indicate the hES cells that present granular morphology. Scale bars: 100µm (A-F), 30µm (B'-E') and 50µm (F').

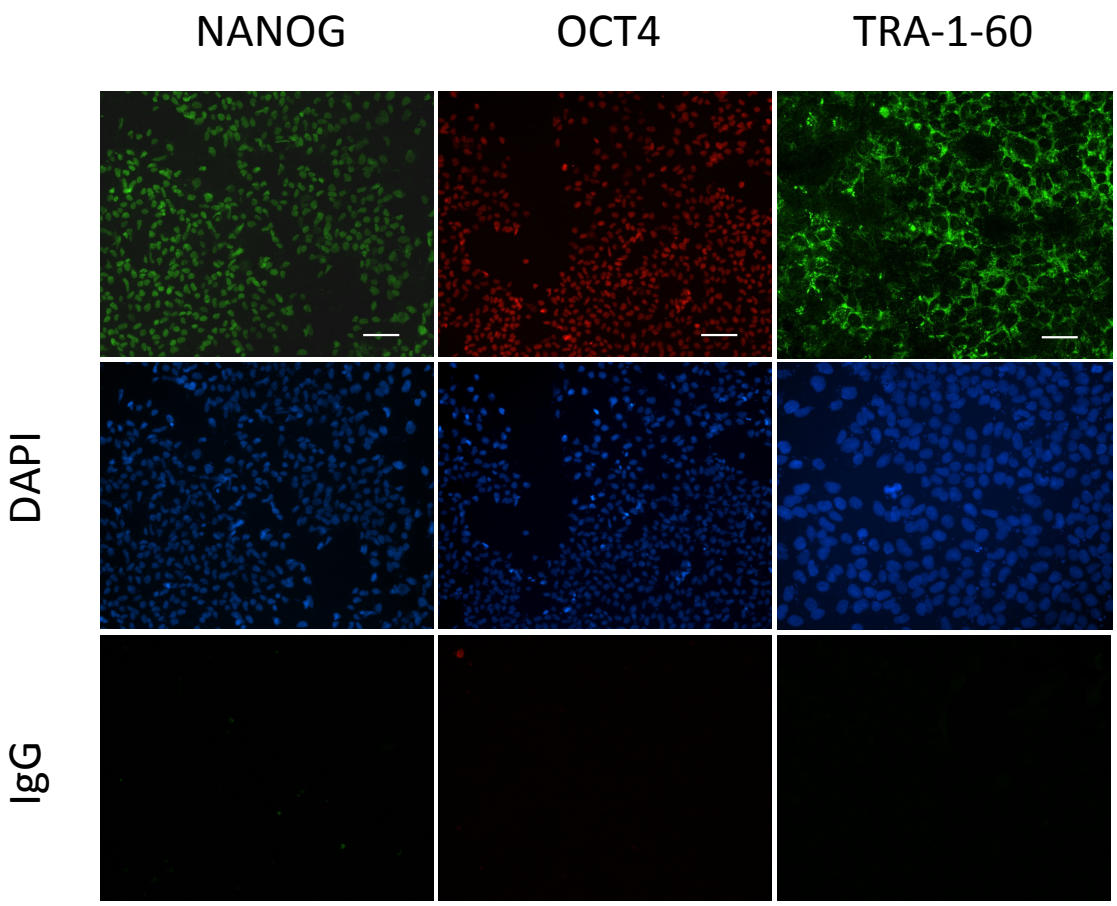


Figure 3-5. Expression of pluripotency markers in hES cells OxF1 grown on MEF ECM plus MEF CM.

Immunofluorescence staining of hES cell line OxF1 grown on MEF ECM plus MEF CM for expression of the nuclear pluripotency transcription factors NANOG (green) and OCT4 (red) and the cell surface receptor TRA-1-60 (green). Nuclear DAPI-staining (blue) was performed to show cell nuclei. Staining using immunoglobulin (Ig)G served as a negative control. Scale bars: 100 µm, except for TRA-1-60, 50 µm.

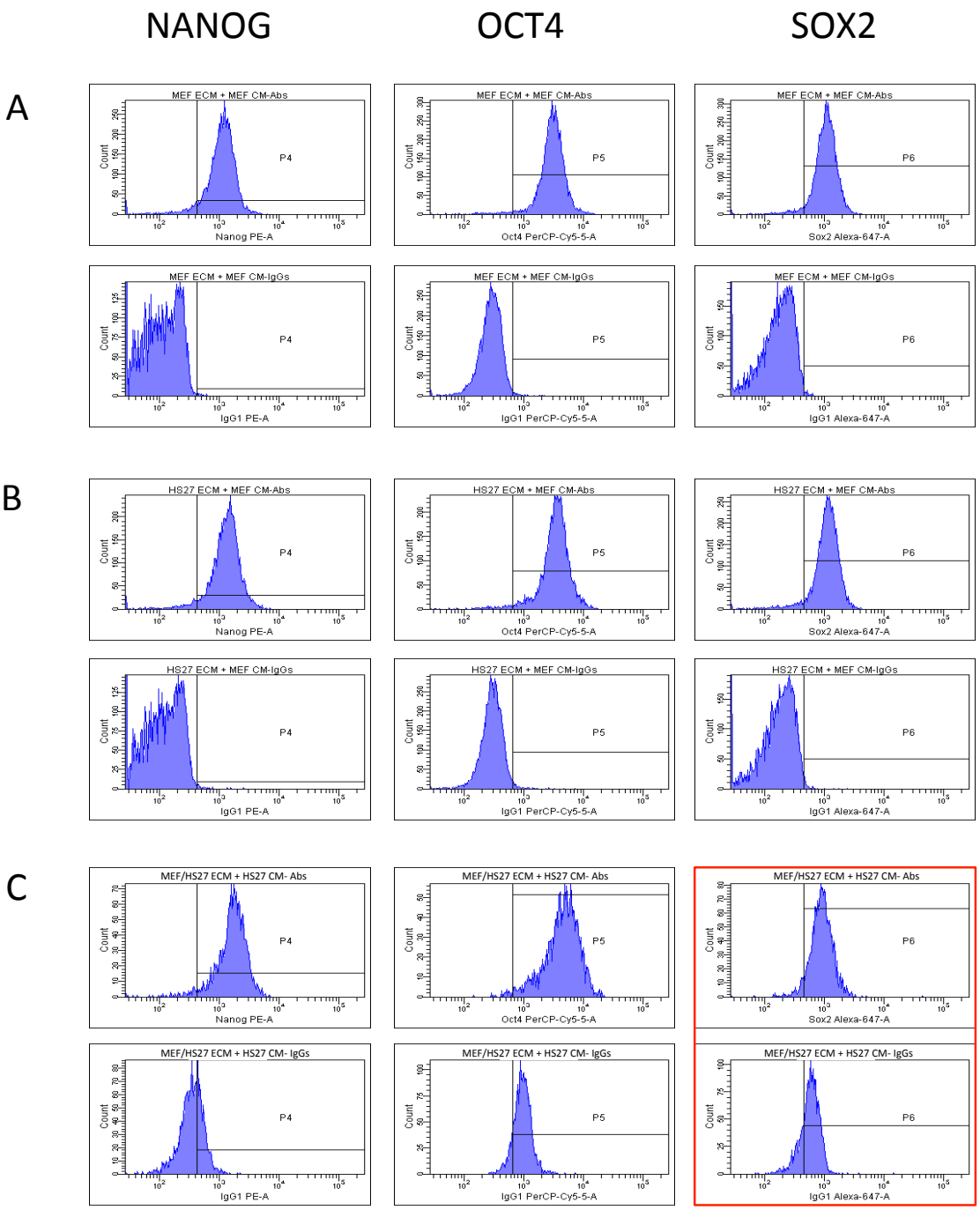


Figure 3-6. Expression of pluripotency markers on hES cells OxF1 grown on ECM and CM from supporting and non-supporting feeders.

Flow cytometry analysis of the expression of the pluripotency markers NANOG, OCT4 and SOX2 on the hES cell line OxF1 grown on MEF ECM plus MEF CM (A), Hs27 ECM plus MEF CM (B) and MEF or Hs27 ECM plus Hs27 CM (C). The panels framed in red highlight the lack of SOX2 expression. Note that due to the low number of cells available, the samples in (C) are a combination of OxF1 cells grown on MEF and Hs27 ECM, which had shown the same morphology (Fig. 4B and E). The Y-axis of each panel indicates the cell count while the X-axis indicates the intensity of fluorescence that is detected. IgG controls are shown at the bottom of each panel.

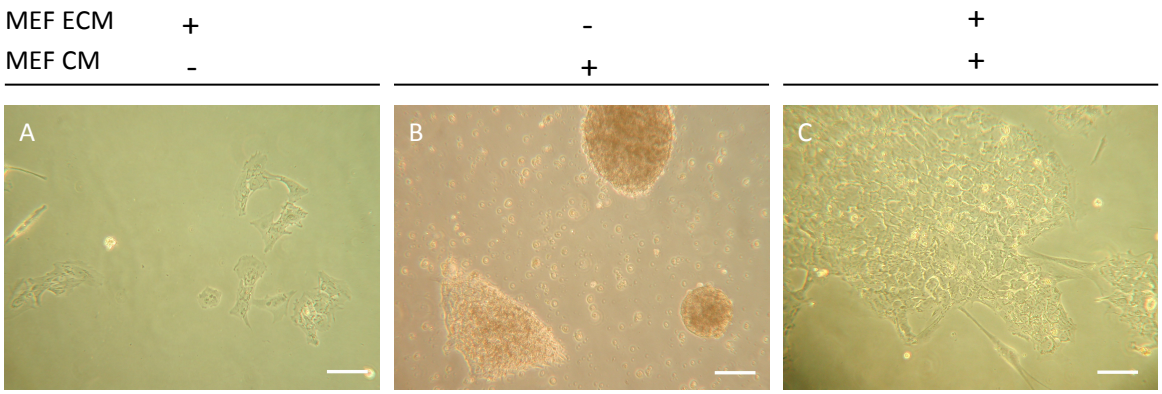


Figure 3-7. Human ES cells grown on the presence or absence of ECM/CM from MEFs.

Phase contrast microscopy analysis of hES cell line OxF1 grown on MEF ECM but no MEF CM (A) and on an uncoated surface and MEF CM (B). OxF1 grown on MEF ECM and CM was used as control (C). Scale bars: 100 µm (A and C) and 50 µm (B).

3.6 Discussion

In order to perform genetic analyses of vU1 snRNAs expression hES cells, a hES cell culture system that was free of MEFs or human feeders was necessary. Towards this goal, I attempted to purify and define the composition of ECM proteins from different feeders and assess whether different ECM protein composition could support maintenance of hES cells in culture.

As shown by previous studies, different feeders provide different levels of hES cell maintenance (Thomson et al., 1998, Richards et al., 2003, Koivisto et al., 2004, Stojkovic et al., 2005, Rajala et al., 2007, Saxena et al., 2008, Chen et al., 2009, Ilic et al., 2009, Chin et al., 2010). I confirmed that two different feeders, MEFs and Hs27, either support or do not support, respectively, two independent hES cell lines (H1 and OxF1). I hypothesized that differences in ECM composition provided by different feeders account for the differences in hES cell maintenance. Thus I isolated the ECM substrate produced by MEFs, Hs27 and CCD and analysed the expression of key ECM proteins by Western blotting.

I showed that different feeders produce an ECM substrate that contains, at least, various relative levels of proteins such as collagen IV, laminin-1, tensacin-C, fibronectin, osteopontin and vitronectin. I found that collagen IV appeared to be expressed in both MEF and CCD ECMs and at much lower level on Hs27 ECM. Studies that tested the ability of collagen IV to support hES cell growth and maintenance showed that it did not support hES cells (Liu et al., 2006) or, in the cases that it did, hES cells showed a decreased expression of OCT4 (Xu et al., 2001). It appeared that only when combined with other ECM proteins, collagen IV was able to maintain hES cells in culture (Liu et al., 2006,

Ludwig et al., 2006b). Furthermore, it has been demonstrated that culture of mES on collagen IV induces their differentiation towards the mesoderm and trophoectoderm lineages (Schenke-Layland et al., 2007, Xiao et al., 2007). These observations support a role for collagen IV in hES cell differentiation rather than hES cell maintenance.

In the case of laminin-1 (also known as laminin-111, as it is composed of the chains α1, β1 and γ1) the highest expression was observed in MEF ECM. Laminin-111, -332 and -511 have been shown to maintain hES cells in culture (Miyazaki et al., 2008) and found at low levels in non-supporting feeders (Hongisto et al., 2012).

For tenascin-C, we observed that two (out of the 20+ described) isoforms were highly expressed in ECM from human feeders, especially in the CCD ECM, compared to MEF ECM. Insights into the role of tenascin-C in stem cell maintenance come from research on the development of the murine central nervous system (mCNS). The work of Garcion and colleagues showed that the neural stem cell pool within the sub-ventricular zone (SVZ) of the developing brain is disturbed in tenascin-C null animals (Garcion et al., 2004). The lack of tenascin-C not only prompted neural stem cells to follow neurogenesis but also delayed the transition from neurons into glial precursors (Garcion et al., 2004). Similarly, further down into CNS development, tenascin-C was shown to be responsible for allowing further glial differentiation into astrocytes in the developing spinal chord (Karus et al., 2011). Interestingly, in both of these studies, it seems clear that tenascin-C regulates the switch between early responsiveness to FGFs and late responsiveness to epidermal growth factor (EGF), which ultimately tunes the balance between neural stem cell proliferation and differentiation (Garcion et al., 2004, Karus et al., 2011).

Full-length fibronectin was highly expressed on CCD ECM compared to Hs27 ECM; in MEF ECM, a fragment of fibronectin was observed instead. Previous studies have shown that, full-length fibronectin alone or combined with other ECM proteins was able to maintain the proliferation of hES cells (Liu et al., 2006, Ludwig et al., 2006b).

Expression of osteopontin was observed in the form of its C-terminal fragments generated by the cleavage of thrombin or MMP. MMP-cleaved osteopontin (32kDa) was detected in MEF, Hs27 and CCD ECM, while thrombin-cleaved osteopontin (~37kDa) was highly expressed in MEF ECM. Clues on the role of osteopontin in stem cell behaviour come from hematopoietic and mesenchymal stem cells in the mouse and rat models, respectively. Mice null for osteopontin showed an increase of hematopoietic stem cells, suggesting a role of osteopontin in restricting the size of the stem cell pool (Stier et al., 2005). On the other hand, in the rat bone marrow, osteopontin induces mesenchymal stem cell migration through activation of focal adhesion kinase (FAK) and ERK signalling pathways via integrin β1 (Zou et al., 2012).

Vitronectrin was only detected in the ECM produced from human feeders. Previous experiments have demonstrated the maintenance of hES cell in culture using vitronectin (Braam et al., 2008).

I have established that the ECM deposited by feeders contains a mixture of these key ECM proteins. Differences in the ECM composition from different feeders (Kueh et al., 2006) or their CM (Prowse et al., 2007, Prowse et al., 2005) have been previously reported but the impact of these differences in hES cell maintenance has not been investigated. To explore this possibility, I cultured hES cells on ECM substrates from MEFs or Hs27 in the presence of MEF or Hs27 CM. My results indicate that MEF and Hs27 ECM

provided a permissive substrate for the maintenance of the hES cell line OxF1. Nonetheless, MEF CM but not Hs27 CM supported OxF1; OxF1 grown on MEF or Hs27 ECM and Hs27 CM presented an obvious granular morphology and down regulation of SOX2 in comparison to OxF1 grown on MEF or Hs27 ECM and MEF CM. Depletion of SOX2 in hES cells leads to down regulation of some pluripotency genes and up regulation of markers of trophectoderm, ectoderm, mesoderm and endoderm differentiation (Adachi et al., 2010, Chew et al., 2005, Fong et al., 2008). Therefore, further characterization of these 'granular hES cells' would need to be done in order to establish if they are going through differentiation and towards which lineage.

My experiments show that both an ECM substrate from supporting or non-supporting feeders as well as CM from supporting feeders are necessary for the growth and maintenance of hES cells. However, its production would be time-consuming and the full composition of MEF CM, more importantly, the component(s) which is/are necessary and sufficient for hES cell maintenance, would remain to be determined. Notwithstanding, by dissecting the *in vitro* hES cell niche, I found that ECM proteins such as collagen IV, laminin-111, tenascin-C, fibronectin and osteopontin are expressed differently in different feeders and therefore may play an important role in allowing the fine balance between hES cell proliferation and differentiation.

Whilst engaged in this study a more defined and less time consuming feeder-free system that allows the growth and maintenance of hES cells over long periods of time was being established in Prof William James laboratory, Oxford. This system requires Matrigel^{TM} as substrate and mTeSR1 as culture medium (Ludwig et al., 2006a). The culture medium mTeSR1 is fully defined and composed of a mixture of amino acids, vitamins, inorganic

salts and additional components (dextrose, hypoxanthine sodium salt, linoleic acid, lipoic acid, phenol red, putrescine hydrochloric acid, sodium pyruvate and thymidine). Matrigel™, according to the manufacturer (BD Bioscience), contains a mixture of ECM proteins that mainly include 60% laminin, 30% collagen IV and 8% entactin; it also contains heparin sulfate proteoglycan and growth factors such as TGF-β, epidermal growth factor (EGF), insulin-like growth factor (IGF), FGF, tissue plasminogen activator and other growth factors.

Consequently, and due to time restraints, I decided to collaborate with Dr. Sally Cowley in the Prof William James laboratory and use their system, enabling me to continue my investigation of the role of vU1 snRNA genes in hES differentiation.

Chapter 4| CHARACTERIZING vU1 snRNA GENE EXPRESSION THROUGHOUT hES CELL DIFFERENTIATION

4.1 Introduction

Human ES cells are defined by their ability to self-renew in culture for prolonged periods of time whilst remaining in a pluripotent state that grants them the capacity to generate all differentiated cell types. These two characteristics make hES cells ideal candidates for cell-replacement therapies in degenerative diseases (e.g. Alzheimer's disease) and in conditions where tissues have been damaged (e.g. myocardial infarction) or are non-functional (e.g. diabetes). In order to be able to generate specific differentiated cell types, it is crucial that we first understand the intrinsic and extrinsic factors and mechanisms that induce or prevent hES cell differentiation.

As described in the previous chapter, a considerable amount of research has been done on establishing the culture conditions (extrinsic factors) that allow hES cell maintenance or direct them to differentiate into specific cell lineages. In parallel, high-throughput approaches (mainly microarray, cDNA library and proteomic analyses) have focused on identifying the molecular signatures (intrinsic factors) of hES cells and their differentiated progeny to identify genes, and therefore possible signalling pathways, involved in hES cell differentiation. These latter studies have revealed that the profile of gene expression, both at the transcriptional and post-transcriptional levels, changes from hES cells to differentiated cell types; specific sets of genes become down regulated whereas others are up regulated (Brandenberger et al., 2004, Calhoun et al., 2004, Dvash et al., 2004, Miura et al., 2004, Cai et al., 2006, Xu et al., 2009b, Fathi et al., 2011). This dynamic profile of gene expression during hES cell differentiation recapitulates embryonic development, as many genes are differentially expressed during oocyte maturation and

throughout the different stages of embryo development in humans (Dobson et al., 2004, Li et al., 2006).

However, the study of differential gene expression in different cell types has been mostly restricted to protein-coding genes. Much less attention has been focused on the expression of genes that do not encode proteins but ncRNAs, which are mainly involved in processing of protein-coding genes. Of special interest here is the urydine-rich (U)1 snRNA. As described in Chapter 1, U1 snRNA and associated proteins form the U1 small nuclear ribonucleoprotein (U1 snRNP), which is responsible for the first step in splicing by recognizing the exon-intron boundary at the 5' splice junction and regulates mRNA length through inhibition of polyadenylation. Recent evidence suggests that the differential protein expression in ES cells compared to their differentiated progeny is, in part, due to differences in the pattern of alternative splicing/polyadenylation (Pritsker et al., 2005, Salomonis et al., 2010, Gabut et al., 2011).

It has been previously shown that levels of U1-U6 snRNAs alter differentially throughout rat embryonic and foetal development and that their expression reaches stability at the adult stage for most of the tissues (Ray et al., 1997). Moreover, their distribution in different tissues/organs throughout development varies considerably (Ray et al., 1997).

Similarly, previous studies in different organisms revealed that different forms of the U1 snRNA gene are differentially expressed in the unfertilized and fertilized eggs, during development and in foetal and adult tissues (Forbes et al., 1984, Lund et al., 1985, Lund et al., 1987, Lobo et al., 1988, Lo and Mount, 1990). These U1 snRNA isoforms are referred to as adult U1 (U1a) and embryonic U1 (U1b) isoforms, which differ by just a few nucleotides in sequence and are mainly expressed in adult or embryonic tissues,

respectively. In addition to U1a and U1b, discrete expression of minor U1 isoforms has also been identified in specific cells such as mature oocytes in Xenopus (Forbes et al., 1984) and human (Lund, 1988).

As described in Chapter 1, variants of the U1 snRNA gene (vU1 snRNA genes) have been recently characterized in our lab (O'Reilly et al., 2012); vU1 snRNA genes were found to be actively transcribed and expressed at different levels in the human cancerous cell line HeLa and the hES cell line, HUES2. Therefore, we speculated that vU1 snRNA genes may also be differentially expressed throughout hES cell differentiation and this could account for differences in the splicing pattern between different cell types. In this chapter, I aim to investigate the profile of expression of vU1 snRNA genes in hES cells and in subsequent differentiation steps. For this purpose, we chose a well-established differentiation system that allows the generation of homogenous monocytes and macrophages from hES cells (Karlsson et al., 2008).

4.2 Matrigel™ is not contaminated with vU1 snRNA

Having chosen a method for hES cell culture that does not require the use of feeders, it remained to be established whether Matrigel™, originally derived from a mouse sarcoma, contained any traces of DNA and/or RNA that could interfere with the analyses of vU1 snRNA gene expression in hES cells.

Although Matrigel™ has a relatively defined composition, it is not DNase- or RNase-treated and thus it is not guaranteed to be free of contaminating DNA and/or RNA. In order to test this possibility, Matrigel™ was prepared in the same way as for hES cell

culture (hES cell line HUES2 was used). Thereafter, RNA was isolated from a 6-well plate containing Matrigel™ alone or from a separate well containing HUES2 growing on Matrigel™. Complementary (c)DNA of each sample was synthesized by reverse transcription using random primers and was then subjected to qPCR analysis using specific gene primers targeting some vU1 snRNAs, vU1.7,9, vU1.8 and vU1.13-16,19 (Fig. 4-1). Results are shown as a percentage related to HUES2; differences in primer specificity between the RNAs were corrected using a human genomic DNA (gDNA) standard. The results show that the level of vU1.7,9, vU1.8 and vU1.13-16,19 snRNAs is very low in Matrigel™ alone compared to HUES2 cells grown on Matrigel™ (Fig. 4-1).

In conclusion, these data demonstrates that the use of Matrigel™ in hES cell culture will not interfere with further analyses of vU1 snRNAs in hES cells.

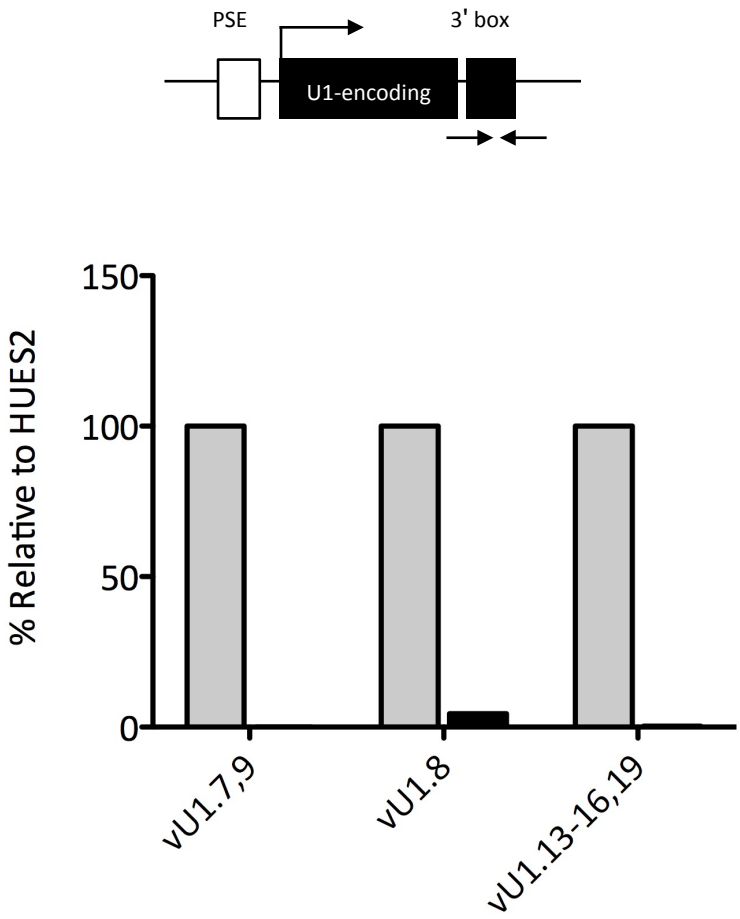


Figure 4-1. Testing the levels of vU1 snRNAs in Matrigel™.

Quantitative real time PCR was performed using specific primers to the 3' box of each vU1 snRNA (black arrows in schematic). Each PCR generated a single product, indicating the specificity of the primers used. The levels of vU1 snRNA detected in Matrigel™ (black bars) are shown relative to HUES2 control (grey bars).

4.3 Pluripotent stem cells express similar levels of vU1 snRNAs

Work in our laboratory has shown that vU1 snRNA genes are detected in the hES cell line HUES2 and the cancer cell line HeLa and that the vU1 snRNAs vU1.1+10, vU1.2a,2+11 and vU1.3-5,12+20 are preferentially expressed in HUES2 (O'Reilly et al., 2012). In order to ensure that this observation was not restricted to a specific hES cell line, in addition to HUES2, the hES cell lines HUES1 and HUES4 were analysed for the expression of vU1 snRNA genes (Fig. 4-2).

Total RNA samples from each hES cell line were isolated and cDNA was synthesized by reverse transcription using random primers. These samples were then subjected to qPCR analysis. Since many of the vU1 snRNAs have high sequence homology in the RNA - encoding regions but more divergent at the 3′ flanking regions, primers within the latter region were chosen to measure nascent levels from all vU1 snRNA genes (pre-vU1 snRNAs; Fig. 4-2). All samples were normalized to 7SK snRNA and relative levels estimated using human gDNA as standard. The specificity of the primers was confirmed using plasmid clones of the vU1 snRNA genes (O'Reilly et al., 2012).

As expected, the hES cell line HUES2 showed different comparative levels of expression for all pre-vU1 snRNA genes tested (Fig. 4-2). Moreover, the levels of each pre-vU1 snRNA are fairly comparative among the three hES cell lines; vU1.17 snRNA was consistently not detected and therefore excluded in all subsequent analyses. The levels of pre-vU1 snRNAs in these hES cell lines are comparable to the level of pre-U1 snRNA, thus strongly supporting the idea that high expression of vU1 snRNAs genes may be characteristic of pluripotent stem cells.

In conclusion, my results confirm that vU1 snRNA genes are similarly expressed among different hES cell lines. Consequently, only one hES cell line (HUES2) was chosen for further analyses in the following sections.

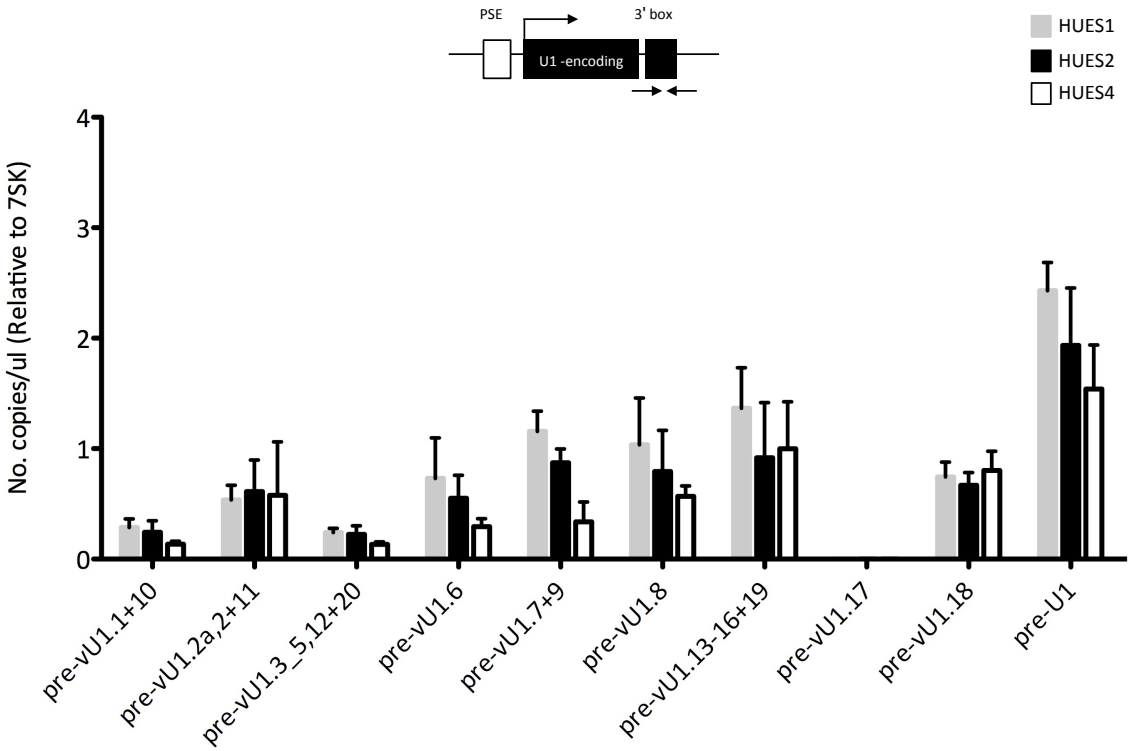


Figure 4-2. Pre-vU1 snRNA levels in pluripotent stem cells: Human ES cells HUES1, HUES2 and HUES4.

Quantitative real time PCR was performed using specific primers to the 3' box of each vU1 and U1 snRNA (black arrows in schematic). Each PCR generated a single product, indicating the specificity of the primers used. Levels were estimated using a gDNA as standard and RNA levels between different cell lines were normalized to 7SK snRNA. The values represent the mean ± SD (n=3).

4.4 Expression of vU1 snRNA genes is reduced upon differentiation

Gene expression patterns in hES cells are dynamic; specific sets of genes are only expressed in undifferentiated and pluripotent hES cells and upon differentiation these genes are down regulated while genes that are involved in differentiation towards specific cell lineages become up regulated (Brandenberger et al., 2004, Calhoun et al., 2004, Dvash et al., 2004, Miura et al., 2004, Cai et al., 2006, Xu et al., 2009b, Fathi et al., 2011). To test whether the vU1 snRNA genes also follow this pattern of expression, I measured their RNA levels in hES cells before and after induction of differentiation.

HUES2 were ultimately differentiated into macrophages, following a gradual differentiation protocol that included EB formation and production of non-adherent monocytes as intermediate steps (Fig. 4-3) (Karlsson et al., 2008). Characterization of each cell type throughout hES cell differentiation was established by the analysis of specific cell markers by qPCR analyses using gene specific primers (Fig. 4-4). As expected, hES cells HUES2 and 4-day-old EBs express the pluripotency marker OCT4 (Fig. 4-4A and B). Moreover, the hES-derived monocytes and hES- derived macrophages express their characteristic monocyte/macrophage differentiation markers CD14 and CD68 antigen (Fig. 4-4D).

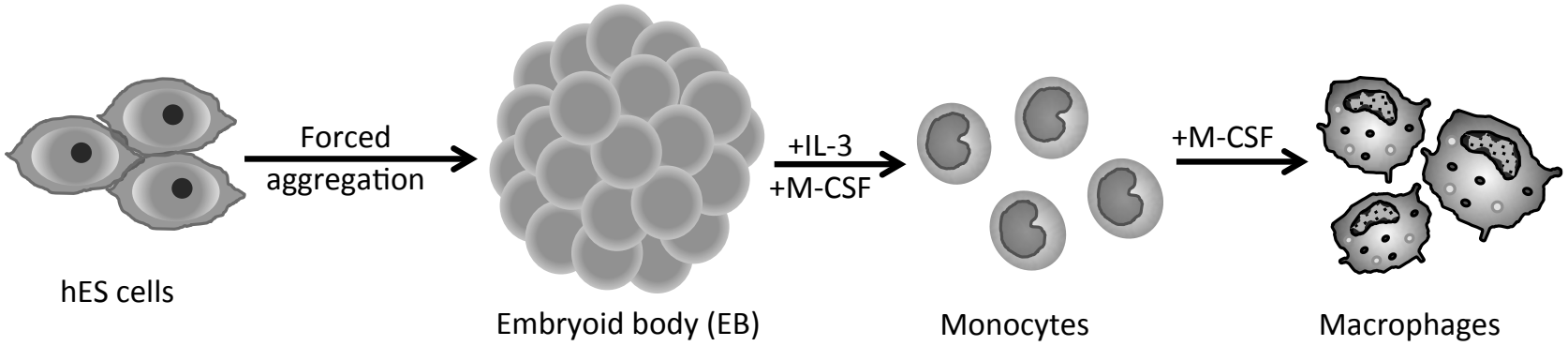


Figure 4-3. Diagrammatic representation of a protocol designed by Karlsson et al (2008) to differentiate hES cell into macrophages.

Human ES cells are first differentiated into embryoid bodies (EBs) by forced aggregation. Following attachment of EBs and supplementation with IL-3 and M-CSF, monocytes are harvested from the supernatant and these are then plated and allowed to differentiate into macrophages. Drawings are not to scale.

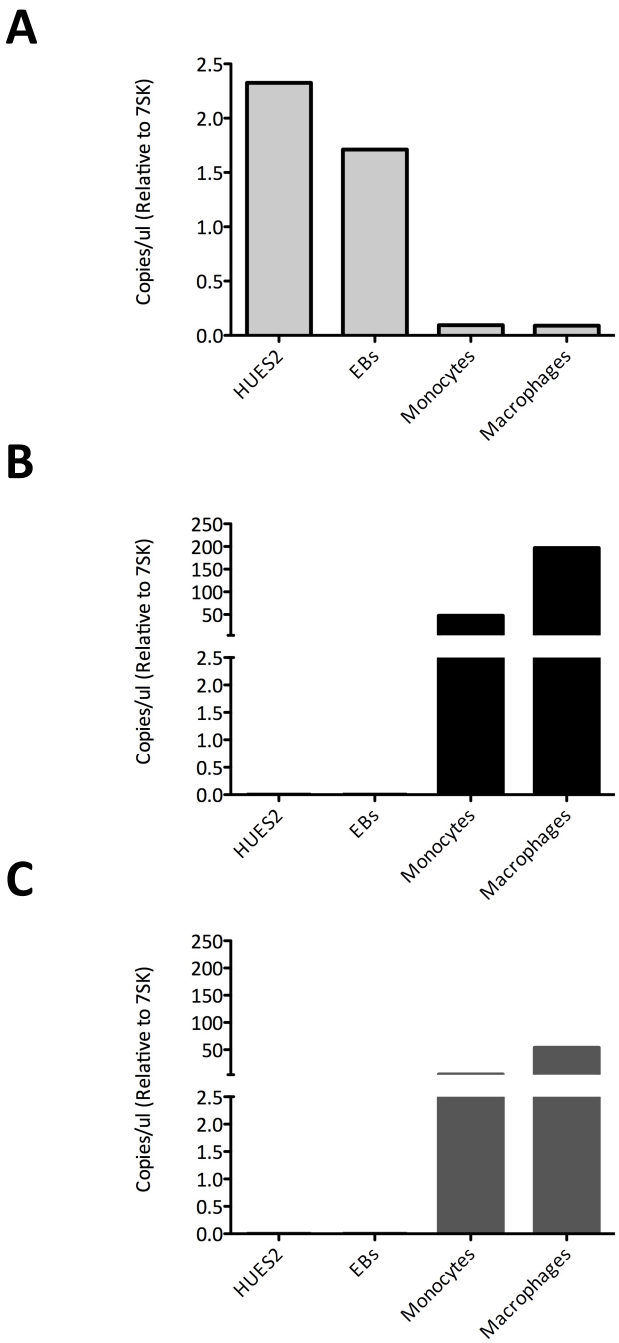


Figure 4-4. Levels of cell specific markers in HUES2, EBs, monocytes and macrophages.

Quantitative real time PCR was performed using specific primers for each gene: OCT4 **(A)**, CD14 **(B)** and CD68 **(C)**. Each PCR generated a single product, indicating the specificity of the primers used. Levels were estimated using a gDNA as standard and RNA levels were normalized to 7SK snRNA.

To investigate changes in the patterns of vU1 snRNA gene expression during differentiation, hES cells HUES2 were first forced to aggregate and form EBs and RNA samples were isolated and processed as described in the previous section. Pre-vU1 snRNA levels in these EBs were compared to levels in HUES2. My results indicate that expression from both the vU1 snRNA genes and the U1 snRNA gene is slightly reduced following differentiation into EBs (Fig. 4-5).

Further culture of plated EBs in the presence of serum-containing DMEM supplemented with IL-3 and macrophage colony stimulating (M-CSF) allows the production of non-adherent monocytes in the supernatant. Harvested monocytes were allowed to adhere in the presence of serum-containing Roswell Park Memorial Institute (RPMI) media supplemented with M-CSF and thus differentiated into macrophages. RNA isolation and subsequent qPCR analysis of nascent levels of vU1 snRNAs, as it has been previously described, was performed for both non-adherent monocytes and their differentiated macrophages. My data show that, with the exception of U1 snRNA and possibly vU1.8 snRNA, expression from all vU1 snRNA genes analysed continues to reduce as hES cells progress through their differentiation into monocytes (Fig. 4-5).

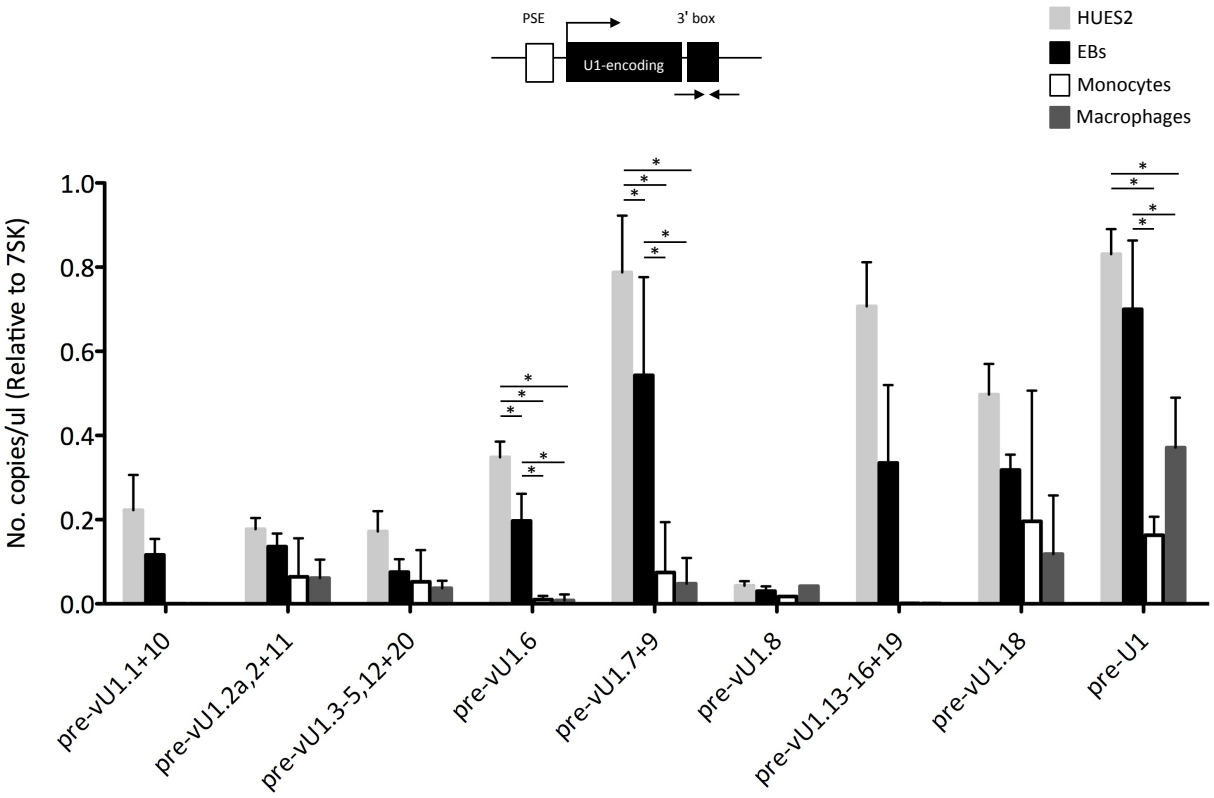


Figure 4-5. Pre-vU1 snRNA levels in HUES2, EBs, monocytes and macrophages.

Quantitative real time PCR was performed using specific primers to the 3' box of each vU1 and U1 snRNA (black arrows in schematic). Each PCR generated a single product, indicating the specificity of the primers used. Levels were estimated using a gDNA as standard and RNA levels between the different cell types were normalized to 7SK snRNA. The values represent the mean ± SD (n=3). *p<0.05

So far, the expression of the vU1 snRNA genes has been analysed only at their nascent levels. This has given us an indication of the pattern of expression for all vU1 snRNA genes but does not provide any information about the generation of a stable snRNA. Typically, a pre-U1 snRNA, of approximately 800bp, is expressed from the U1 snRNA gene (O'Reilly, unpublished observation). A 3' box element directs cleavage of this pre-U1 snRNA a few bases downstream from the mature snRNA and an unknown exonuclease trims back the 3' extension, generating a 164bp mature U1 snRNA (Uguen and Murphy, 2004). This is the functional form of the U1 snRNA, which is packaged into an snRNP complex and participates in pre-mRNA splicing and 3' processing. To determine which of the vU1 snRNAs expressed in hES cells are potentially functional and thus important for hES cell maintenance or differentiation, I first measured the steady state level of U1 snRNA and vU1 snRNAs throughout the differentiation process using primers that target the snRNA-encoding region, herein referred to as mature (m)-U1 snRNA and mature (m)-vU1 snRNA. I then compared this steady state level to nascent levels for each U1 snRNA and vU1 snRNA to estimate the degree of processing that occurs during differentiation.

Given the high degree of sequence homology within the RNA-encoding regions across the U1 snRNA and vU1 snRNA genes, I was only able to specifically measure the steady state levels of a select group of vU1 snRNAs, including vU1.2a, vU1.3, vU1.8 and vU1.20, in addition to the U1 snRNA. Steady state levels of vU1 snRNAs in hES cells, EBs, monocytes and macrophages are illustrated in Fig. 4-6.

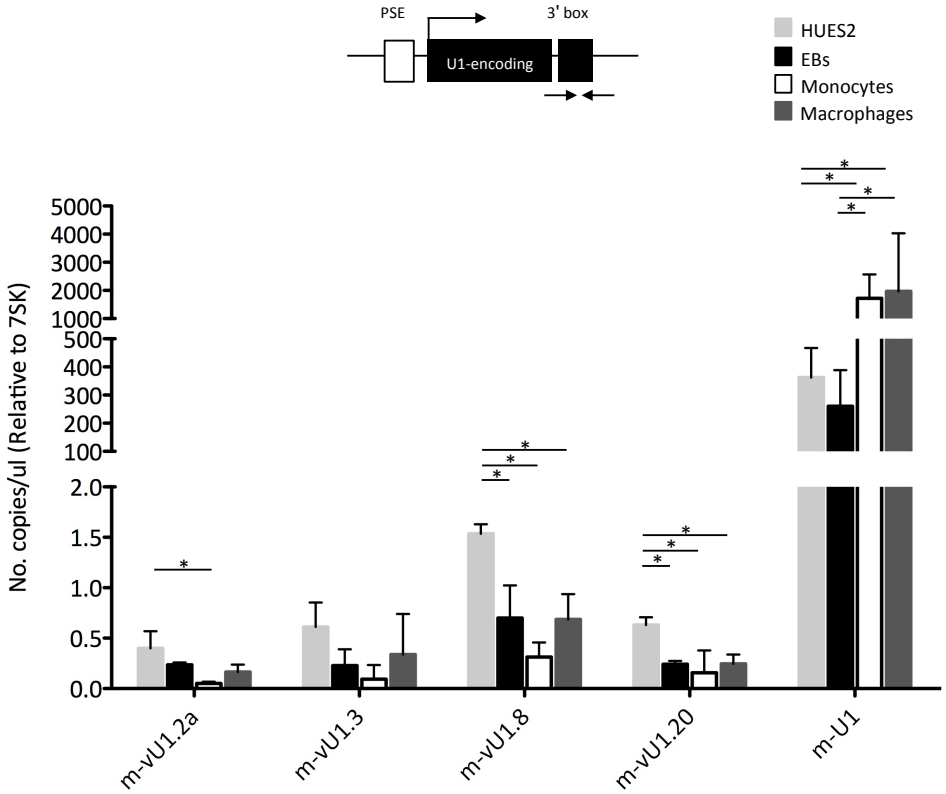


Figure 4-6. Mature-vU1 snRNA levels in HUES2, EBs, monocytes and macrophages.

Quantitative real time PCR was performed using specific primers to the RNA-encoding region of each vU1 and U1 snRNA (black arrows in schematic). Each PCR generated a single product, indicating the specificity of the primers used. Levels were estimated using a gDNA as standard and RNA levels between the different cell types were normalized to 7SK snRNA. The values represent the mean ± SD (n=3, except for m-vU1.2a in monocytes where n=2). *p<0.05

My results show that the steady state levels of all m-vU1 snRNAs tested are at their highest level in HUES2 and continue to reduce as cells differentiate into monocytes. This pattern of expression very much reflects the profile of nascent vU1 snRNAs illustrated in the previous figure (Fig. 4-5), which suggests that vU1 snRNA genes are down regulated during differentiation of hES cells to macrophages. Interestingly, in contrast to vU1 nascent levels, which continue to decrease as hES cells differentiate into macrophages, there appears to be a marginal increase (approximately four-fold) in the m-vU1 snRNA levels as the differentiation process proceeds down the macrophage lineage. It seems that the reduction in nascent levels during differentiation into macrophages is not an indication that the U1 snRNA and vU1 snRNA genes are down regulated but may reflect more efficient processing of the nascent snRNA in macrophages.

To examine this more thoroughly, I compared the levels of pre-U1 snRNA and pre-vU1 snRNAs to their corresponding steady state level (m-U1 and m-vU1, respectively) in all four cell types (Fig. 4-7). In order to do this comparison, the levels of pre-U1 and -vU1 snRNAs had first to be corrected so that they would only represent levels of transcript corresponding to a single U1 or vU1 snRNA gene, respectively. Thus, levels of pre-vU1.2a had to be divided by three, which correspond to the number of vU1 snRNA genes of its group (i.e pre-vU1.2a,2,11). Likewise, levels of pre-vU1.3 and vU1.20 were divided by five (pre-vU1.3-5,12,20) and U1 by four, which correspond to the copies annotated by our group.

Following comparison of corrected pre-U1 and vU1 snRNA levels, as suspected, U1 snRNA is more stable in monocytes (~40,000 fold above nascent levels), followed by macrophages, HUES2 and EBs (~30,000, 1,700 and 1,500 fold above nascent levels,

respectively) (Fig. 4-7E). These data would suggest that the mature U1 snRNA level can vary quite considerably during differentiation and that the mechanisms that control the processing of this snRNA and/or its stability may be key regulatory processes in maintaining appropriate levels in different cell types. As illustrated in Figure 4-7, all vU1 snRNAs undergo some degree of processing in all cell types analysed. Interestingly, with the exception of vU1.8 snRNA, most vU1 snRNAs are hardly processed in monocytes, which is the cell type that expresses the highest level of m-U1 (Fig. 4-7E). In contrast, cells that express reduced levels of m-U1, for example HUES2 and HUES2-dervided macrophages, all show an increase in processing of the vU1 snRNAs. Processing of both the vU1 snRNAs and U1 snRNA is reduced in EBs, which may indicate a general reduction in processing of all gene types at this stage of differentiation.

So far, the results presented here indicate that the vU1 snRNA gene expression profile changes during differentiation of hES cells (HUES2) into macrophages. Analysis of nascent vU1 snRNA levels suggests a general reduction in expression of the vU1 snRNA genes throughout the differentiation process. However, examination of m-vU1/pre-vU1 steady state levels indicates that this reduction in nascent levels is not due to changes in gene expression but either increased efficiency of 3' processing or greater stabilization of the mature vU1. Steady state levels of m-vU1.8 are highest in HUES2, whereas m-vU1.20 snRNA is more efficiently processed in HUES2-derived macrophages, which suggest that each cell type may have its own vU1 snRNA signature.

Although processing of each vU1 snRNA is different in each cell type, by far the biggest difference observed is the relative levels of m-vU1 snRNAs to m-U1 snRNA at each differentiation stage (Fig. 4-6). In monocytes, the m-U1 levels are high (Fig. 4-6) but the

vU1 snRNAs are poorly expressed (Fig. 4-7A-D). In contrast, vU1 snRNA are more stable in HUES2 (Fig. 4-7A-D) but the levels of m-U1 snRNA are ~4 times lower in these cells compared to levels measured in monocytes (Fig. 4-6). There appears to be a reciprocal relationship between m-U1 snRNA and vU1 snRNA processing in each cell type analysed.

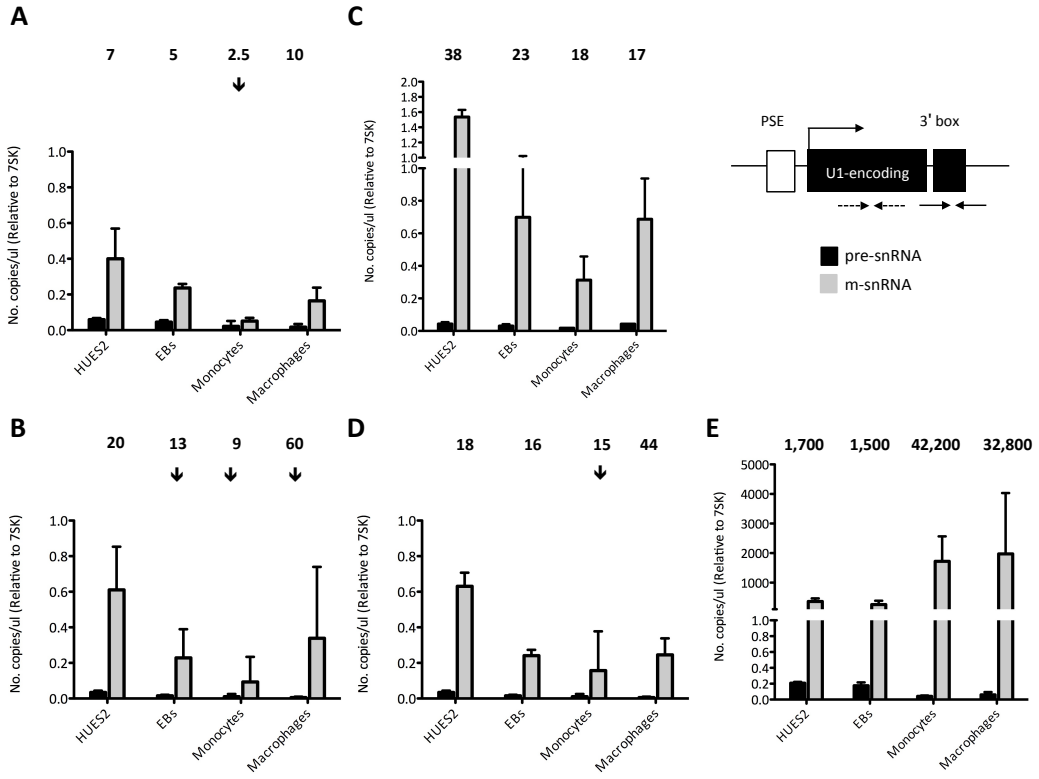


Figure 4-7. Comparison of the levels of pre- and m-vU1 and U1 snRNAs in HUES2, EBs, monocytes and macrophages.

Quantitative real time PCR was performed using specific primers for the RNA-encoding region and 3' box of vU1.2 **(A)**, vU1.3 **(B)**, vU1.8 **(C)**, vU1.20 **(D)** and U1 snRNA **(E)** (dashed and continuous arrows in schematic, respectively). Each PCR generated a single product, indicating the specificity of the primers used. RNA levels were normalized to 7SK snRNA. The values represent the mean ± SD (n=3, except for m-vU1.2a in monocytes where n=2). The numbers in bold above the graphs bars indicate fold-difference between pre- and m-vU1 snRNAs. In most cases there was a statistical significance of $p<0.05$ except where indicated (block arrows).

4.5 Discussion

Human ES cells hold a great potential in the field of regenerative medicine, as they possess the characteristics of being able to self-renew in culture indefinitely whilst retaining the potential to give rise to all cells in the body. However, to fully unleash this potential, it is crucial we identify the factors that either maintain hES cells in a pluripotent state or lead to the differentiation of hES cells into one cell type or another.

The identification of many other factors has come from studies on embryonic development in model organisms. High-throughput techniques have also uncovered more of the genes that are differentially expressed in hES cells and their differentiated cell progeny at the transcriptome and proteome levels. Therefore, it appears that gene expression during differentiation is highly dynamic, and U1 snRNAs are not the exception; different forms of this gene have been described in embryonic (U1b) and adult (U1a) tissues. Studies in plants (Hanley and Schuler, 1991), sea urchin (Santiago and Marzluff, 1989), fruit fly (Lo and Mount, 1990), frog (Forbes et al., 1984, Lund and Dahlberg, 1987, Lund et al., 1984) and mouse (Kato and Harada, 1985, Lerner and Steitz, 1979, Lund et al., 1985) have shown that these U1 snRNAs are also differentially expressed throughout development.

In this chapter it was demonstrated that human vU1 snRNA genes are actively expressed in hES cells and become down regulated upon differentiation. A subset of the newly characterized vU1 snRNA genes appeared to be mainly expressed in the hES cell line HUES2 (O'Reilly et al., 2012); therefore, we aimed to investigate whether their expression is unique to hES cells or changes throughout differentiation.

Before proceeding to understand the pattern of expression of vU1 snRNA genes in hES cell differentiation, I first confirmed that their expression did not vary between different hES cell lines. Indeed, nascent (pre)-vU1 snRNAs levels vary for each vU1 snRNA but these levels were relatively equivalent among the hES cell lines tested (Fig. 4-2).

Previous studies have revealed a dynamic pattern of gene expression throughout development and during hES cell differentiation. Given that vU1 snRNAs are detected in hES cells, I investigated if their pattern of expression also changed over the course of differentiation. For this purpose, we took advantage of a model based on the gradual differentiation of hES cells into macrophages (Karlsson et al., 2008). Variant-U1 snRNA genes were clearly expressed in hES cells (HUES2) at different levels; following four-days differentiation into EBs already showed a slight down regulation of the levels of nascent vU1 snRNAs (Fig. 4-5) and expression of the pluripotency marker OCT4 was also slightly down regulated (Fig. 4-4A). This is consistent with other studies, which showed that EBs at early stages of differentiation showed similar pattern of gene expression to that of hES cells (Bhattacharya et al., 2005). Further differentiation into monocytes and macrophages showed a steeper decrease in pre-vU1 snRNA levels (Fig. 4-5); this was accompanied by up regulation of markers characteristic of monocytes and macrophages (Fig. 4-4B and C).

Similarly to pre-vU1 snRNAs, the steady state levels of m-vU1 snRNAs (only few could be tested due to the high degree of sequence homology) throughout differentiation also decreased upon hES cell differentiation into EBs and monocytes but began to increase at the macrophage stage (Fig. 4-6). More importantly, the comparison of nascent against steady state vU1 snRNA levels provided us with an indication of the stability of the different vU1 snRNAs in the different cell types and/or how efficient the nascent message

was processed in the different cell types (Fig. 4-7). Processing of the vU1 snRNAs that could be tested occurred at different levels throughout differentiation. The maximal processing of vU1 snRNAs occurs in hES cells whereas in monocytes little or no processing was observed for most of the vU1 snRNAs; this poor processing began to be recovered at the macrophage stage (Fig. 4-7A-D). In contrast, the highest level of processing of U1 snRNA occurs in monocytes and is considerably decreased in hES cells and macrophages. In the case of EBs, the levels of processing of both vU1 snRNAs and U1 snRNA are low (Fig. 4-7E). Looking at individual vU1 snRNAs, vU1.8 snRNA shows the highest degree of processing throughout differentiation, followed by vU1.20, vU1.3 and vU1.2a snRNAs (Fig. 4-7).

These data strongly support the idea that the vU1 snRNA genes are being regulated throughout differentiation both at the level of transcription and possibly 3' end processing.

In conclusion, the work presented here demonstrates that:

1) Variant-U1 snRNA genes are mainly expressed in pluripotent hES cells and down regulated in differentiated cell types (EBs, monocytes and macrophages), therefore suggesting a possible role in hES cell maintenance. This would need to be further validated by unravelling target genes involved in pluripotency following genome-wide analyses.

2) Comparison of nascent levels to steady state indicate that expression of vU1/U1 snRNA genes appears to be regulated at the transcription level following differentiation

of hES cells to monocytes and that the 3' end processing may play an additional role in regulating the steady state levels in macrophages.

Given these observations, in the next chapter I aim to manipulate the expression of vU1.8 and vU1.20 snRNAs. The reason for targeting vU1.8 snRNA is that it is the most highly expressed vU1 snRNA analysed and maximally expressed in HUES2 and hES-derived macrophages. Conversely, vU1.20 snRNA is highly expressed in hES cells only.

Chapter 5| ROLE OF vU1 snRNA GENES IN HUMAN PLURIPOTENT STEM CELLS

5.1 Introduction

It has been suggested, but not further investigated, that the differential expression of U snRNAs (and their isoforms) themselves is likely to play a role in regulating the changes in mRNA isoform diversity associated with cells undergoing differentiation (Ray et al., 1997). In the case of U1 snRNA, no studies have unravelled the relevance of the differential expression of U1a, U1b and minor U1 snRNA isoforms during development.

Having established that vU1 snRNA genes are specifically expressed in pluripotent hES cells and that they are differentially expressed throughout hES cell differentiation, the next step is to investigate the biological function of vU1 snRNAs in these different cell types.

In order to perform functional studies, I first had to design constructs that would allow me to express specific sense/antisense vU1 snRNAs at different stages of development in an inducible manner or specifically knockdown their expression in hES cells.

I was also interested in establishing whether vU1 genes are up regulated upon induction of primary human skin fibroblasts to iPS cells as this would give me another indication of a role in cell reprogramming/maintenance of the ES cell state.

Manipulation of vU1 snRNA expression in hES cells, combined with genome-wide analyses, will allow us to gain a better understanding of the role vU1 snRNAs play in maintenance of pluripotency, cell reprogramming or preventing differentiation.

5.2 Variant U1 snRNA genes are up regulated in iPS cells

The forced expression of a set of transcription factors in somatic cells has led to the generation of cells that resemble pluripotent stem cells known as iPS cells (Takahashi et al., 2007). Induced PS cells exhibit some of characteristics that define pluripotent stem cells, including the expression of pluripotency markers and the capacity to differentiate into EBs or form teratomas (Robinton and Daley, 2012).

So far, the results presented here suggest that vU1 snRNA genes may contribute to the maintenance of the pluripotent state of hES cells as the majority of vU1 snRNA genes are mainly expressed in hES cells and become down regulated at each differentiation stage (Fig. 4-5 and 6). I then asked whether the pattern of expression of vU1 snRNA genes changes upon reprogramming of somatic cells into iPS cells as it does during hES cell differentiation.

Total RNA was extracted from iPS cells and from the primary human skin fibroblasts from which they originated. Complementary DNA was synthesized for each sample and analysed by qPCR as it has been previously described. My analyses revealed that levels of all vU1 snRNAs in primary human skin fibroblasts are very low and that these levels increase quite dramatically upon forced expression of pluripotent markers (Fig. 5-1), which ultimately lead to the generation of iPS cells. Conversely, levels of U1 snRNA are higher in primary human skin fibroblasts (approximately threefold) compared to those of iPS cells (Fig 5-1). Furthermore, the majority of vU1 snRNA genes, apart from vU1.20 snRNA, show comparable expression between hES and iPS cells (Fig. 5-2).

These results support the idea that vU1 snRNA genes play a role in cell reprogramming and in maintenance of pluripotent stem cells. In order to test this possibility, I will attempt to manipulate the expression of vU1 snRNA genes in hES cells; further whole genome analyses would then reveal whether genes involved in pluripotency become disrupted.

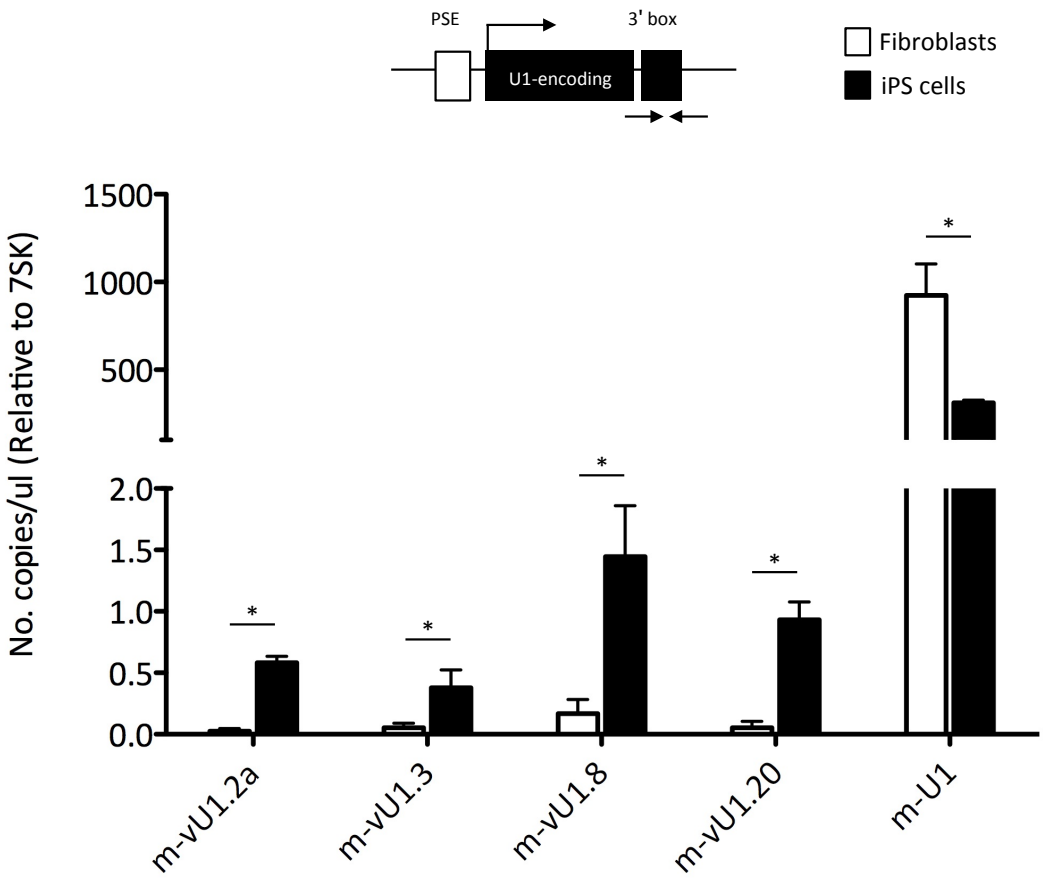


Figure 5-1. Mature (m)-vU1 snRNA levels in primary human skin fibroblasts and iPS cells.

Quantitative real time PCR was performed using specific primers to the RNA-encoding region of each vU1 and U1 snRNA (black arrows in schematic). Each PCR generated a single product, indicating the specificity of the primers used. Levels were estimated using a gDNA as standard and RNA levels were normalized to 7SK snRNA. The values represent the mean ± SD (n=3). *p<0.05

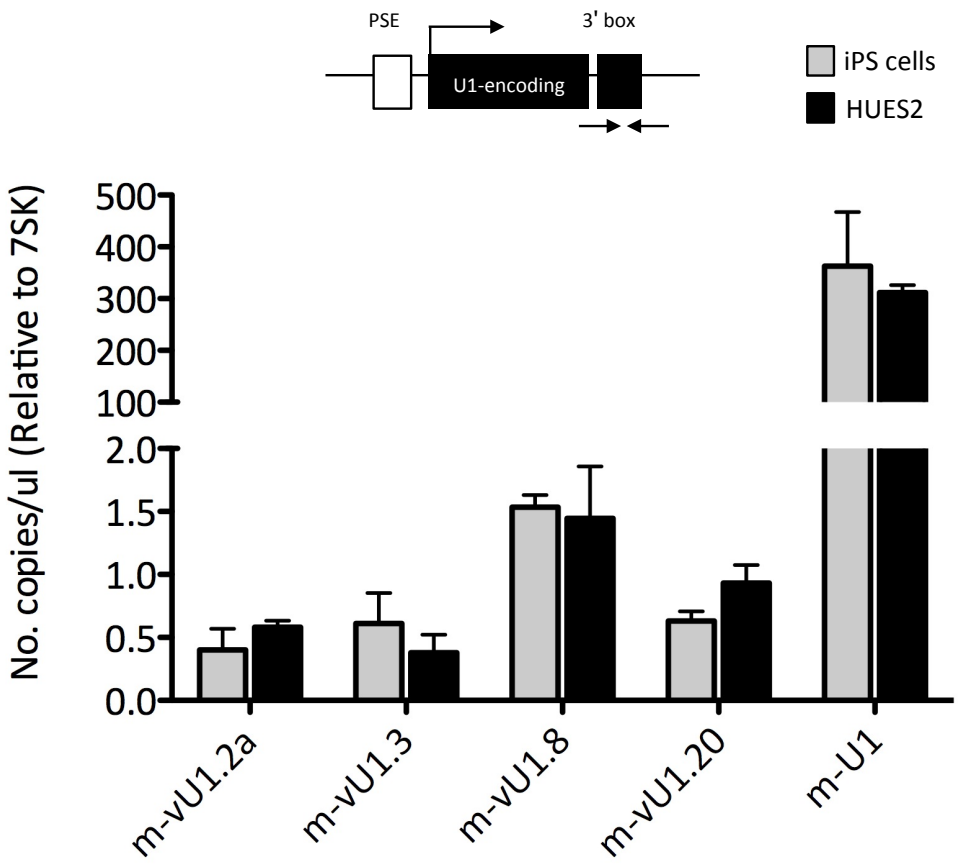


Figure 5-2. Mature (m)-vU1 snRNA levels in iPS cells and HUES2.

Quantitative real time PCR was performed using specific primers to the RNA-encoding region of each vU1 and U1 snRNA (black arrows in schematic). Each PCR generated a single product, indicating the specificity of the primers used. Levels were estimated using a gDNA as standard and RNA levels were normalized to 7SK snRNA. The values represent the mean ± SD (n=3).

5.3 Transfection efficiency in hES cells

Manipulation of vU1 snRNA gene expression in hES cells involves the introduction of genetic material, mainly in the form of either plasmid constructs or oligonucleotides. Therefore, it was imperative to find a suitable method for their transfection into hES cells.

The transfection method routinely used in our laboratory consists of lipofection (i.e. liposomal-based transfection) using the reagent Lipofectamine 2000 (Invitrogen). This has been reported to achieve efficient small interfering (si)RNA-mediated knockdown in hES cells (Hay et al., 2004, Ma et al., 2010).

First, I began testing hES cell transfection efficiency by assessing the incorporation and subsequent expression of green fluorescent protein (GFP). However, this was hardly observed in HUES2 cells, even 48h post-transfection. The lack of GFP fluorescence may not indicate poor transfection but that expression of GFP may be toxic to these cells. Therefore another method to assess transfection efficiency had to be pursued. As constructs that express vU1 snRNAs under the control of the U1 snRNA promoter were available in our laboratory (O'Reilly et al., 2012), I tested whether these could be transiently expressed in hES cells (Fig. 5-3); 1µg of the constructs pGEM4 or pGEM4/vU1.3 or vU1.8 (Fig. 5-3A) was transfected in HUES2, previously seeded at $0.5x10^5$ cells/cm^2 two days before transfection to achieve 60-80% confluence. Following transfection, cells were harvested after 22-24h, RNA was extracted and further processed and analysed as previously described. Real time PCR analyses of these samples using gene specific primers for m-vU1.3 or m-vU1.8 allowed detection of these vU1 snRNAs in HUES2 that had been transfected with the corresponding constructs. Samples were normalized

to 7SK snRNA. Levels of m-vU1.3 or m-vU1.8 were specifically observed at considerably higher levels than the control, approximately 40- and 50-fold increases, respectively, above the endogenous levels detected in controls (Fig. 5-3B). In addition, the level of mRNA for the pluripotency marker OCT4 slightly increases in HUES2 when there is more vU1.3 or vU1.8 snRNA (Fig. 5-3C). In all cases, HUES2 were morphologically similar (data not shown). In conclusion, these results indicate that HUES2 are successfully transfected using Lipofectamine 2000 and that transient transfection does not seem to affect their maintenance.

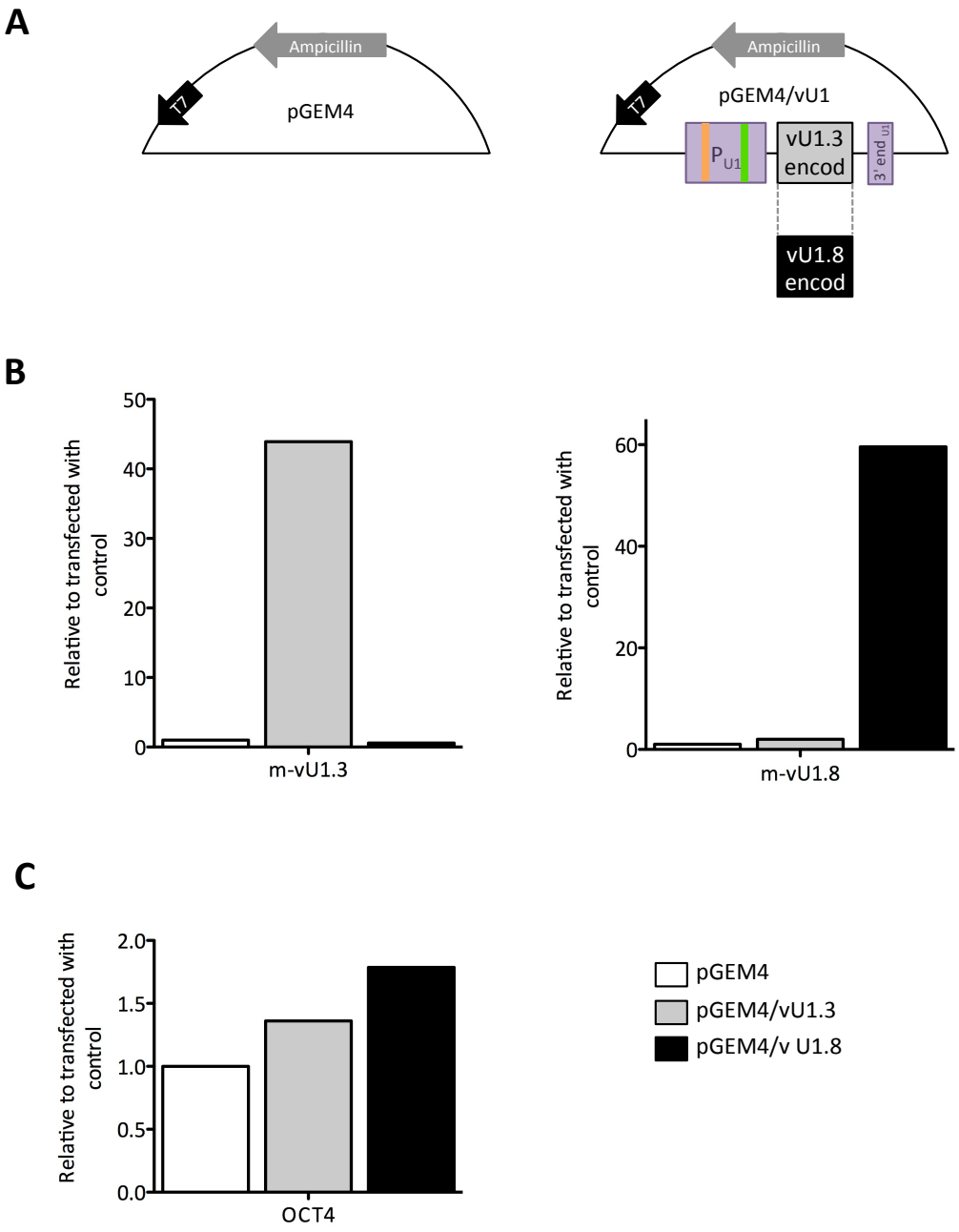


Figure 5-3. Testing hES cell transfection efficiency with vU1 snRNA expressing constructs.

(A) Schematic diagrams of the constructs pGEM4 (left) and pGEM4 containing the vU1 snRNA-encoding region of vU1.3 and vU1.8 flanked by the promoter (P_{U1}) and 3' end of U1 snRNA (right). The orange and green vertical lines within the P_{U1} represent the DSE and PSE, respectively. The black arrow indicates the T7 promoter and the grey arrow represents the ampicillin resistance gene within the vector backbone. **(B)** Following transfection of pGEM4 (white bars), pGEM4/vU1.3 (grey bars) or pGEM4/vU1.8 (black bars) in HUES2, qPCR analyses, using gene specific primers, indicate expression of each vU1 snRNA, vU1.3 (left) and vU1.8 (right), in a specific manner. **(C)** Following transfection of the indicated constructs in HUES2, qPCR analyses, using gene specific primers, indicate expression of OCT4. In all these analyses, levels are normalized to 7SK snRNA and relative to the transfection control.

5.4 Down regulation of vU1 snRNA genes in hES cells

Having established that lipofection using Lipofectamine 2000 allows the transient transfection of foreign episomal genetic material in hES cells, I then asked whether knockdown of vU1 snRNAs in hES cells could be achieved by transfection of antisense oligonucleotides against the 5' end of vU1 snRNAs.

Some of the most common methods used to achieve knockdown of genes of interest involve the RNA interference (RNAi) pathway, whereby small RNA molecules as part of a multiprotein complex down regulate protein expression by means of disrupting translation (miRNAs) or by degrading mRNA (small interfering (si)RNAs).

Given the nuclear localization of vU1 snRNAs, such methods would not be suitable for our purpose. The 2'-O-Methyl phosphorothioate RNA/DNA antisense oligonucleotides provided a better option. They are approximately 20 base pairs in length, formed by central DNA (~18 bases) flanked by RNA (~5 bases). The former allows it to be directed into the nucleus while the latter prevents it from degradation (Kiss, 2004), the modification of the 2'- position of the ribose to 2'-O-Methyl confers the oligonucleotide with higher binding affinity (Lubini et al., 1994) and the conversion of the phosphate backbone to phosphorothioate provides resistance against nucleases and allows the recruitment of RNase H, which ultimate degrades the RNA target (Lapham et al., 1997).

The 2'-O-Methyl phosphorothioate RNA/DNA (herein referred to as OMP) antisense oligonucleotides have achieved successful knockdown of snRNAs in HeLa cells (Ideue et al., 2009). Knockdown of vU1.8 snRNA, using a vU1.8_OMP that targets the 5' end of this vU1 snRNA immediately downstream of the 5' splice site, was effectively achieved in

HeLa cells as early as 10h post-transfection (O'Reilly et al., 2012). Therefore, I tested whether the knockdown of vU1.8 snRNA with vU1.8_OMP was also able to down regulate this vU1 snRNA in HUES2. Two days after being seeded, HUES2 were transfected with 150pmol of vU1.8_OMP or a scrambled OMP oligonucleotide that acted as control. In addition, HeLa cells were transfected in parallel in order to confirm the specificity of the knockdown. Cells were harvested after 10h and their RNA was extracted and processed as previously described. The levels of m-vU1.8 were analysed by qPCR in both cell populations. It was observed that levels of m-vU1.8 were reduced by 85% in HeLa cells compared to 30% in HUES2 (Fig. 5-4A).

This indicates that the antisense oligonucleotides work but that in the case of HUES2 further optimization was required. Therefore, I decided to test the effect of different doses of vU1.8_OMP. HUES2 were transfected with 50, 150 and 300pmol of vU1.8_OMP. Cells were harvested after 10h and processed as it has been previously described. Levels of m-vU1.8 snRNA were measured by qPCR analyses. This showed that down regulation of vU1.8 snRNA to 57 and 54% was achieved at 150 and 300pmol, respectively (Fig. 5-4B).

Having established the conditions for transfection and the correct dose of antisense oligonucleotide. I attempted to knockdown vU1.8 and vU1.20 snRNAs for whole genome analyses. Ten hours post-transfection, levels of m-vU1.8 were down regulated to 28 and 9% using 150 and 300pmol, respectively. However, note that transfection with 300pmol of vU1.20_OMP reduced by half the levels of m-vU1.8 in HUES2 (Fig. 5-5A). Conversely transfection with 300pmol of vU1.8_OMP reduced the levels of m-vU1.20 to 35% (Fig. 5-5B). This suggests that the OMP antisense oligonucleotides have a tendency to cross-react at high concentrations. This is likely due to the similarity in the sequence targeted

by both antisense oligonucleotides as there is only eight-nucleotide difference between them. Moreover, we are restricted to target the 5' end of the vU1 snRNAs given that other sites, although presenting more sequence divergence, are inaccessible due to the secondary structure of vU1 snRNAs and binding to snRNP proteins.

Efficient and specific knockdown of vU1.8 and vU1.20 snRNAs was only observed at 10h post-transfection with an oligonucleotide concentration of 150pmol (Fig. 5-5). Therefore, total RNA of samples under those conditions was sent to the Wellcome Trust Centre for Human Genetics in Oxford for RNA sequencing analysis. At the time of writing this thesis I had not received the data and therefore the analysis is still ongoing.

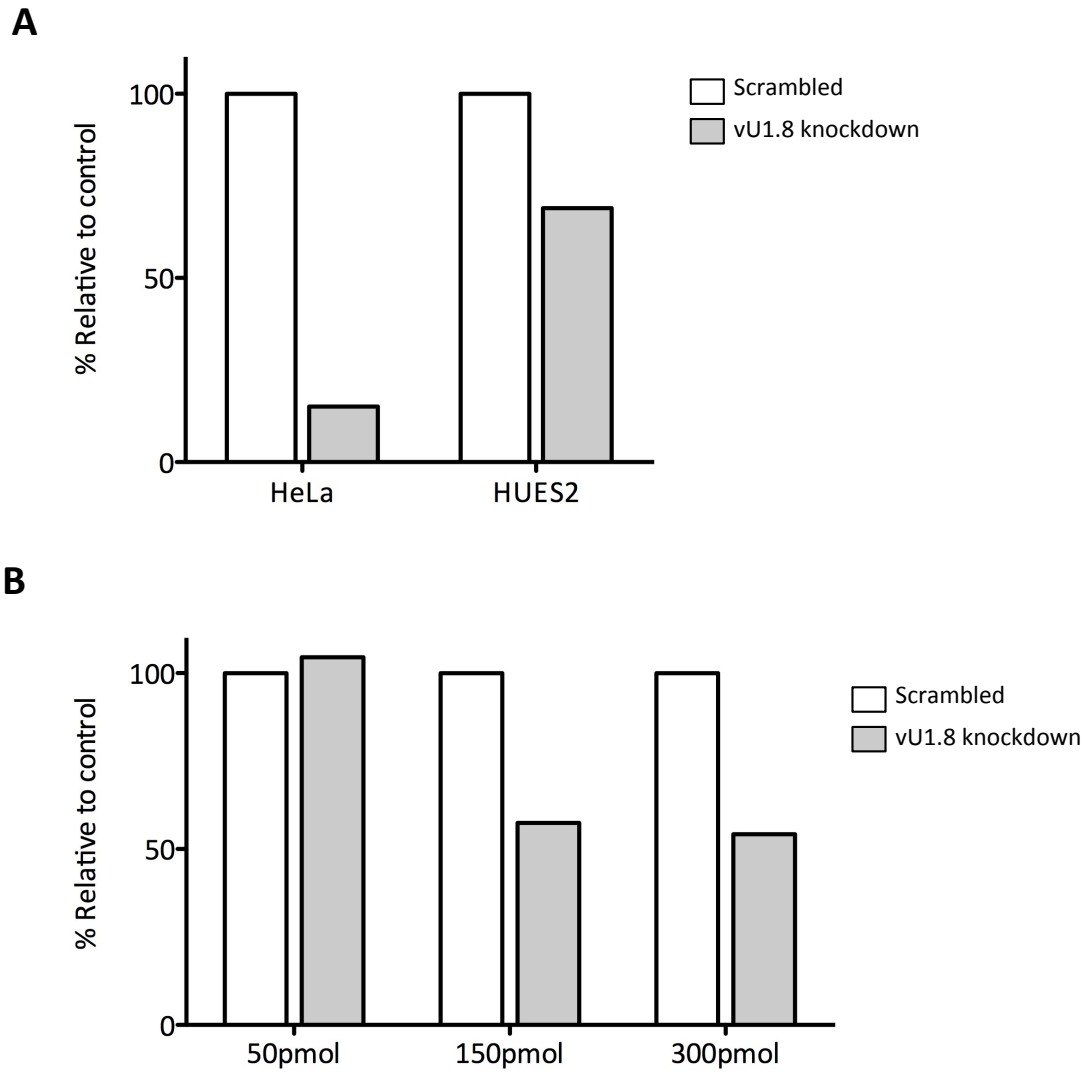


Figure 5-4. Optimizing vU1.8 snRNA knockdown in hES cells.

(A) Quantitative real time PCR analysis of m-vU1.8 snRNA levels in HeLa and HUES2 after vU1.8 snRNA knockdown. **(B)** Quantitative real time PCR analyses of m-vU1.8 snRNA levels in HUES2 after vU1.8 snRNA knockdown with three different doses (50, 150 and 300pmol) of 2'-O-Methyl phosphorothioate RNA/DNA (OMP) antisense oligonucleotide against the 5' end of vU1.8 snRNA. In all these analyses, levels are normalized to 7SK snRNA and relative to the transfection control (scrambled OMP).

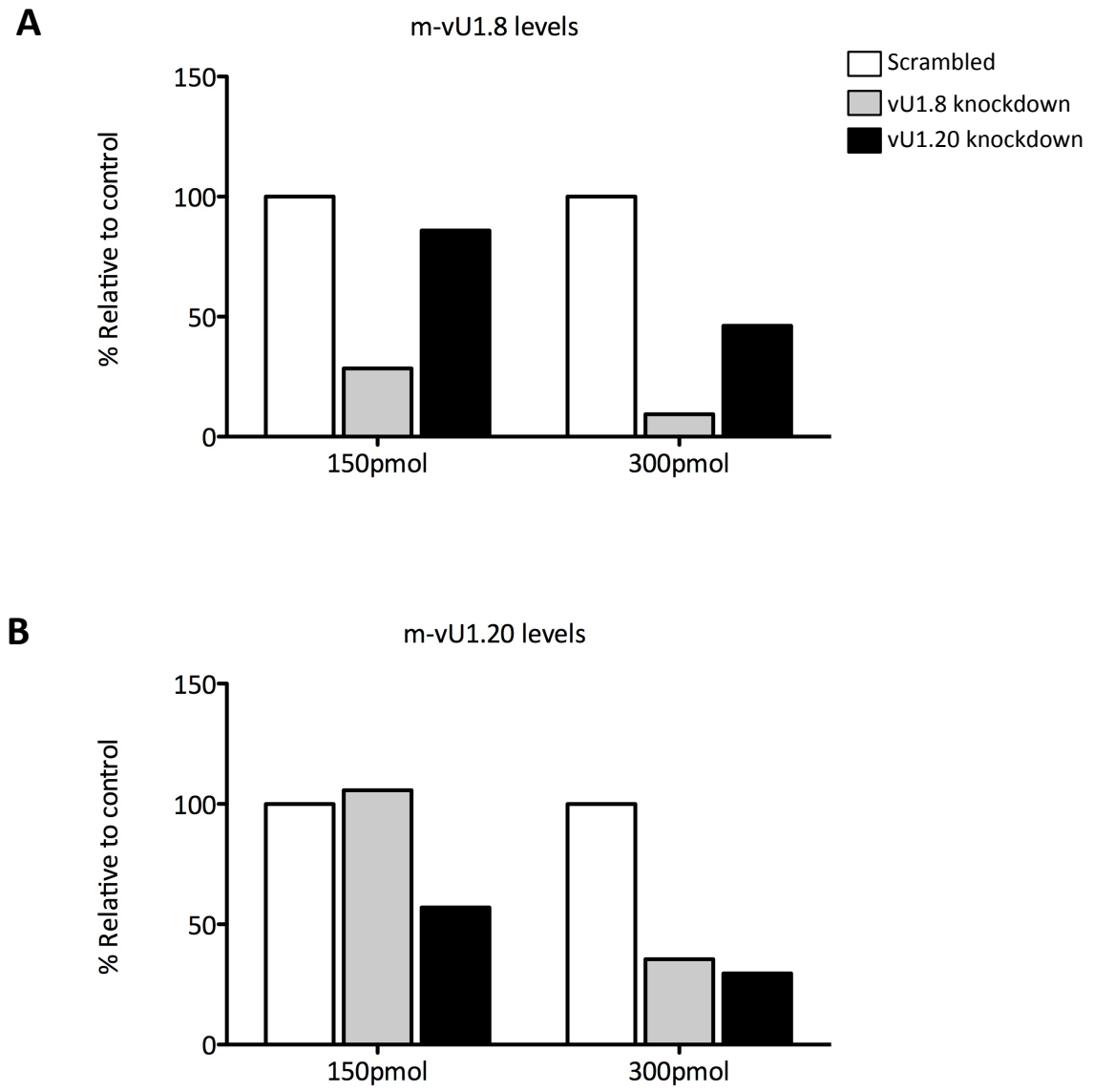


Figure 5-5. Knockdown of vU1.8 and vU1.20 snRNAs in hES cells.

Quantitative real time PCR analysis of m-vU1.8 **(A)** and m-vU1.20 **(B)** snRNA levels in HUES2 cells transfected with two different doses (150 and 300pmol) of 2'-O-Methyl phosphorothioate RNA/DNA (OMP) antisense oligonucleotides against the 5' end of vU1.8 snRNA, vU1.20 snRNA or a scrambled control at 10h post-transfection. Levels are normalized to 7SK snRNA and relative to the transfection control.

While optimizing the use of OMP antisense oligonucleotides for the knockdown of specific vU1 snRNAs, I decided to test another approach for knocking down these genes. Given that transfection of the constructs that express vU1 snRNAs under the control of the U1 snRNA promoter and 3' end had been successful (Fig. 5-3), I focused on building a vector that would express an antisense pre-vU1 snRNA, which could bind the sense vU1 snRNA transcript and thus target its degradation. For this purpose, I PCR-amplified the vU1 snRNA-encoding region from the vU1.8 snRNA gene-expressing construct using specific gene primers and cloned it in reverse orientation in the pGEM4 expressing construct (Fig. 5-6A).

The constructs pGEM4 or pGEM4/antisense (as)_vU1.8 (1 μg of each) were transfected in HeLa cells at 60-80% confluence. Following transfection, cells were harvested after 22-24h, their RNA was extracted and cDNA was synthesized with random primers or with specific primers for the newly transcribed as_vU1.8 and sense(s)_vU1.8 snRNAs. Real-time PCR analyses of these samples using gene specific primers for m-vU1.8 allowed the detection of as_vU1.8 and s_vU1.8 snRNAs in HeLa cells that had been transfected with the corresponding construct. Samples were normalized to 7SK snRNA.

The results show that the levels of as_vU1.8 snRNA are three-fold over the background levels detected in the control (Fig.5-6B). More interestingly, transfection with the as_vU1.8 construct caused reduction in the levels of s_vU1.8 snRNA (which correspond to the endogenous levels) by approximately half compared to the levels observed in the control (Fig. 5-6). Note that the apparently low levels of expression of the as_vU1.8 snRNA compared to the control are possibly due to its degradation along with its target s_vU1.8 snRNA.

The data presented here indicates that expression of an antisense vU1 snRNA within a gene-expressing construct is able to reduce the levels the endogenous sense vU1.8 snRNA, thus constituting a potential approach for the knockdown of vU1 snRNAs.

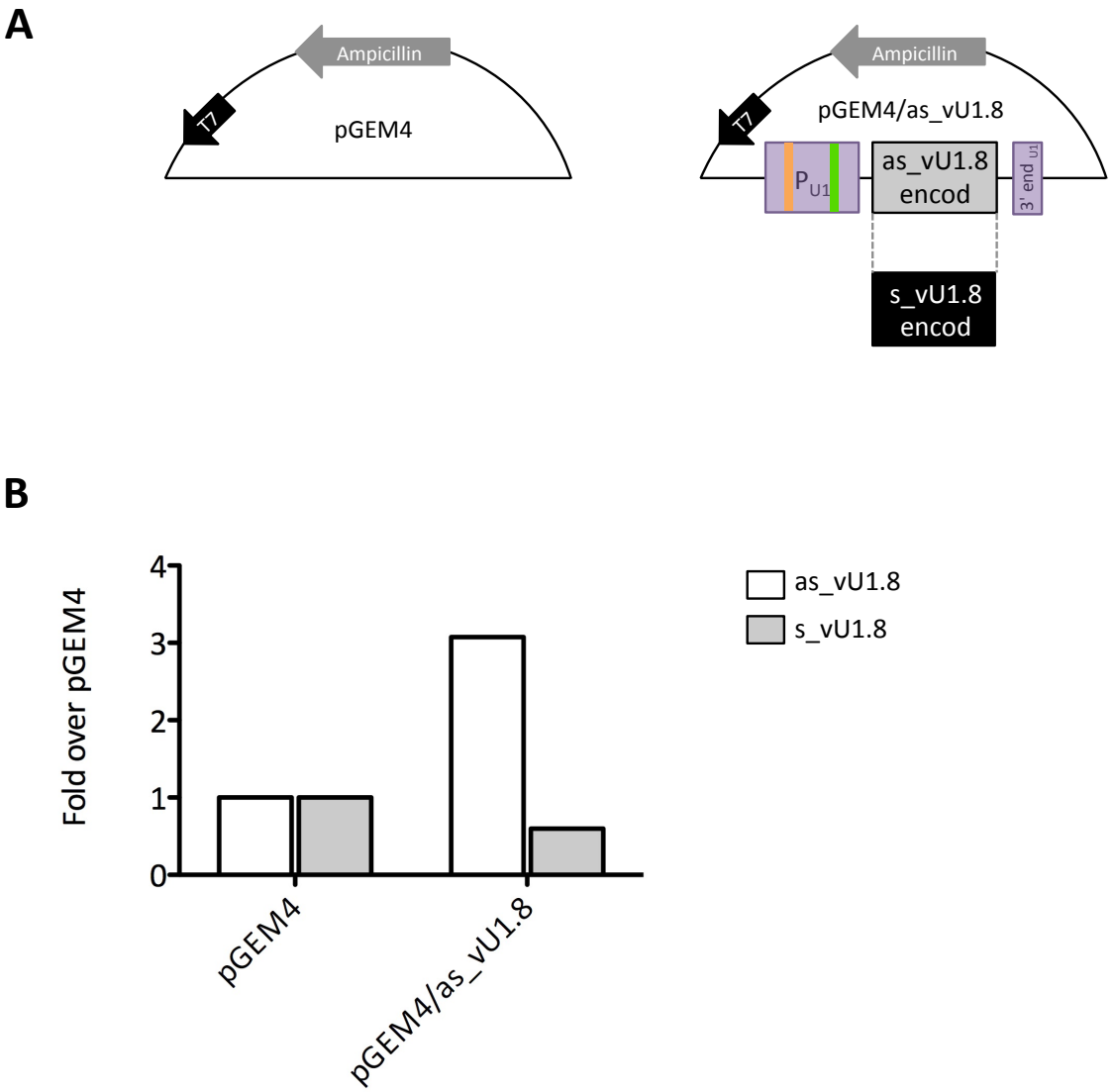


Figure 5-6. Testing antisense (as)_vU1 snRNA expressing construct for the knockdown of vU1.8 snRNA.

(A) Schematic diagrams of the constructs pGEM4 (left) and pGEM4 containing the vU1 snRNA-encoding region of as_vU1.8 flanked by the promoter (P_{U1}) and 3' end of U1 snRNA (right). The orange and green vertical lines within the P_{U1} represent the DSE and PSE, respectively. The black arrow indicates the T7 promoter and the grey arrow represents the ampicillin resistance gene within the vector backbone. **(B)** Following transfection of pGEM4 or pGEM4/as_vU1.8 in HeLa cells, qPCR analyses, using gene specific primers, indicate expression of as_vU1.8 snRNA (white bars) and s_vU1.8 (grey bars), in a specific manner. Levels are normalized to 7SK snRNA and relative to the transfection control.

5.5 Design and validation of a Tetracycline-inducible system for the expression of vU1 snRNA genes

Knocking down vU1 snRNAs in hES cells would provide information on their function in hES cell maintenance. The characteristic differential expression of vU1 snRNAs throughout differentiation suggests a finely tuned mechanism behind their expression. In order to test this possibility, I focused on creating a stable hES cell line that expresses the vU1 snRNA genes in an inducible manner (i.e. dependent on an exogenous inducer/repressor).

One of the widely used systems for gene regulation in eukaryotic cells is the tetracycline-regulated gene expression system, more specifically the tetracycline-controlled *trans* activator (tTA) and reverse *trans* activator (rtTA). Their expression is controlled by doxycycline, a homolog of tetracycline; in its absence or presence, tTA or rtTA, respectively, are able to switch on gene expression upon binding to a specific sequence (tet operator, tetO) inserted in the promoter of the gene of interest. Both tTA and rtTA appear to enhance gene expression as they result from the fusion of the tetracycline repressor (TetR) with the activating domain of the virion protein (VP)16 of herpex simplex virus (Gossen and Bujard, 1992, Das et al., 2004), which is known to promote transcription of protein-coding genes by interacting with transcription factors of the pre-initiation complex, such as TFIIB (Hayashi et al., 1998), but fails to activate the U2 snRNA promoter (Das et al., 1995) and thus will not be suitable for controlling the expression of vU1 snRNA genes (Zaborowska et al., 2012). Therefore, a method that employs the TetR alone was pursued instead. In this case, in the absence of tetracycline/doxycycline the

TetR binds as a dimer to the tetO within the promoter of the gene of interest (in our case, the U1 snRNA promoter) (Fig.5-7). Once bound to the promoter, the TetR prevents the expression of the gene of interest (in this case, vU1 snRNA genes) and upon addition of tetracycline/doxycycline, this is able to bind and sequester the TetR thus allowing expression of the gene of interest (Fig. 5-7).

My plan was to create a single vector that expresses the TetR and an antibiotic resistance gene for selection (in this case puromycin) under the same promoter, thereby ensuring that the selected clones resistant to puromycin also express the TetR. In addition, the same vector will also contain the vU1 snRNA of interest flanked by the U1 snRNA promoter and 3' end, which will allow the efficient expression and processing of the vU1 snRNA; the U1 snRNA promoter will contain tetO sequence to be recognized by the TetR (Fig. 5-8). Previous experiments have investigated how the positioning and number of tetO sequences within the promoter of the U6 snRNA ultimately influence the expression of the genes driven by this promoter. The tightest control of gene expression was observed when: 1) a tetO (O1) sequence was positioned between the PSE and the TATA box (Ohkawa and Taira, 2000), 2) 2x tetO (O2) sequences flanked the TATA box (Lin et al., 2004) and 3) 2-6x tetO sequences were placed at the immediate upstream 5' of the RNA-encoding sequence of U6 (Zhou et al., 2008). In the case of the U1 snRNA promoter, the region between the PSE and the RNA-encoding region consists of 42bp and since it has been previously reported that this distance is critical for expression (Lescure et al., 1992), it restricted the use of only two tetO sequences in my construct.

In order to build this vector, first two vectors were built separately, one containing the TetR and puromycin and a second one, which expresses the vU1 snRNA of interest. In this way, I would also be able to test them individually (see the following section).

For the TetR/puromycin vector, the TetR plus the β-globin intervening sequence (IVS) and SV40 early pA were PCR-amplified from the vector pcDNA6/TR©(Invitrogen) at the restriction sites BamHI, which was converted to a KpnI site, and NotI. This amplified and purified PCR fragment was digested at these sites and inserted into the pcDNA3.1(+) vector (Invitrogen) and validated by sequence analysis. Separately, an internal ribosome entry site (IRES) joined to the puromycin gene was PCR-amplified from the plasmid vector rat insulin promoter (pRIP) (kindly provided by Dr. Kenny Moore, Sir William Dunn School of Pathology, University of Oxford) at the sites BamHI and EcoRI, digested and inserted into the pcDNA3.1 (+) vector; sequence analyses of this construct allowed validation of insertion of the IRES and puromycin sequences. From this vector, the IRES-Puromycin fragment was purified and inserted into the pcDNA3.1 (+) containing β-globin IVS/TetR/ SV40 early pA. In this manner, the IRES-Puromycin fragment was inserted between the TetR and the SV40 early pA, which allows the synthesis of a single stable RNA that would be translated into both TetR and puromycin.

The inducible vU1 snRNA construct was built by the sequential ligation of the U1 snRNA promoter, two tetO sequences, vU1.8 snRNA-encoding region and U1 snRNA 3' end (787bp) and inserted into the pGEM4 vector backbone. The U1 snRNA promoter and 3' end were previously PCR amplified from human gDNA whereas the vU1.8 snRNA-encoding region was amplified from the vU1.8 snRNA expressing construct that was previously used (see section 5.1). Finally, the tetO sequences were synthesized from the

phosphorylation and ligation of separate single-stranded oligonucleotides with incorporated restriction sites. There are two types of tetO sequences, O1 and O2; the latter was chosen given that TetR has higher binding affinity to it (Kleinschmidt et al., 1991).

This last vector would be incorporated into the TetR /puromycin/ pcDNA 3.1(+) construct at the restriction sites NotI and XhoI; in this way their expression would be driven by promoters acting on opposite directions (Fig. 5-8).

Before the two vectors that were built were joined together, each one of them was tested separately. First, the expression of the TetR from the construct consisting of TetR/puromycin/ pcDNA 3.1(+) was tested (Fig. 5-9). HeLa cells, seeded at a density of $0.2x10^5$ cells/cm^2 one day before, were transfected with 1 µg of the TetR/puromycin/pcDNA 3.1(+) construct or pcDNA3.1(+) alone, as control (Fig. 5-9A). Following transfection, cells were harvested after 22-24h and their RNA was extracted and further processed and analysed as previously described. Real time PCR analyses of these samples using gene specific primers for the TetR allowed its detection; cDNA samples that had been prepared in the absence of reverse transcriptase (-RT) served as control for any background expression from the plasmid vector. I observed that the TetR was expressed at considerably high levels in cells that were transfected with the TetR/puromycin/pcDNA 3.1(+) construct in comparison to those cells transfected with pcDNA3.1(+) alone (Fig. 5-9B).

Secondly, I had to ensure that the insertion of the tetO sequence construct did not compromise the activity of the U1 snRNA promoter in the pGEM4/tetO-vU1.8 construct, which drives the expression of the vU1 snRNAs, in this case vU1.8. In order to test this,

HeLa cells were transfected with a pGEM4 control vector or pGEM4/vU1.8 construct with and without the tetO sequence (Fig. 5-10A). Note that within the pGEM4/tetO-vU1.8 construct the insertion of 2x tetO at the BamHI site increased the distance between the PSE and the RNA-encoding region by four nucleotides. Given this could disrupt the expression of the vU1.8 snRNA, I also tested a pGEM4/tetO-vU1.8 construct that had those extra nucleotides removed, pGEM4/tetO (no BamHI)-vU1.8 (Fig. 5-10A). Cells were harvested after 22-24h and processed as previously described. Expression of m-vU1.8 levels by those four constructs was analysed by qPCR (Fig. 5-10B). This showed that efficient expression of vU1.8 snRNA was only observed in those cells transfected with the pGEM4/vU1.8 construct while vU1.8 snRNA expression was severely compromised in the other two constructs (Fig. 5-10B). Maintaining the exact distance between the PSE and the RNA-encoding region only improved expression mildly (Fig. 5-10B), indicating that the sequence within this region is also critical for the function of the U1 snRNA promoter.

In summary, these results show that a Tet-inducible system for the expression of vU1 snRNA genes was hampered by the fact that the insertion of two tetO sequences into the U1 snRNA promoter disrupted its activity and prevented the expression of vU1.8 snRNA gene. Nevertheless, a way to improve this outcome would be to 1) change the U1 snRNA promoter for another snRNA promoter (i.e. U2 snRNA promoter) or 2) place the tetO sequences upstream or between the DSE and PSE as an alternative approach.

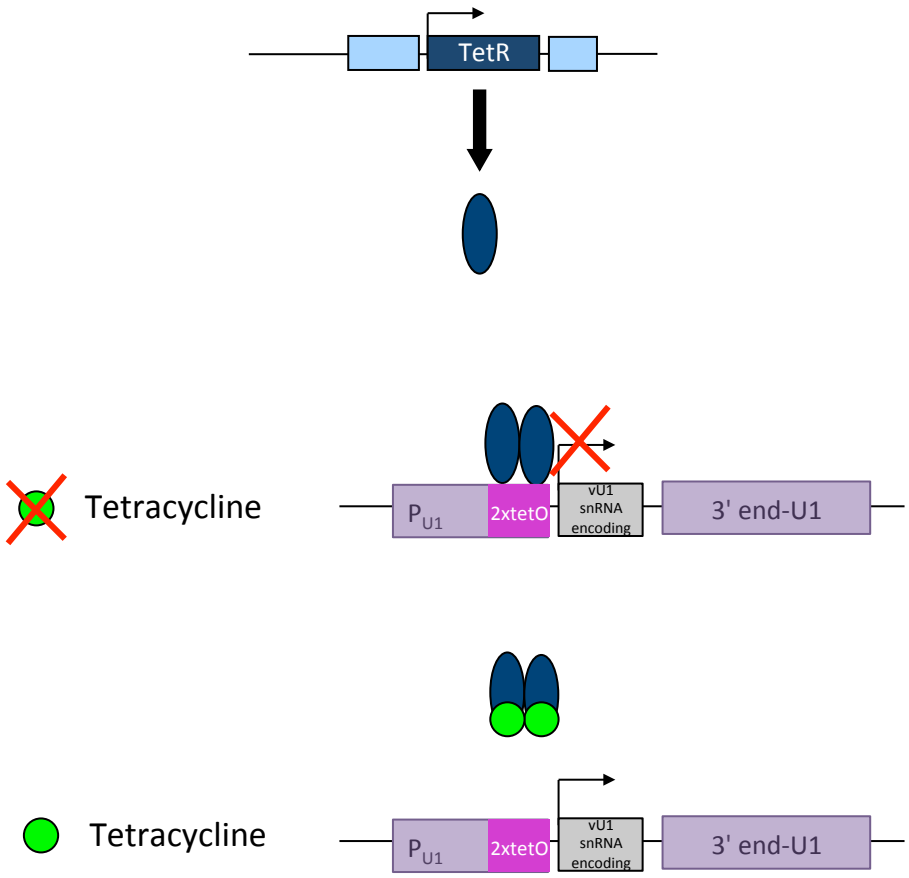


Figure 5-7. Tetracycline repressor-regulated vU1 snRNA gene expression system.

Schematic diagram of the tetracycline-inducible system for the conditional expression of vU1 snRNAs, whereby the tetracycline repressor (TetR) is expressed and is able to bind as a dimer the tet operator (tetO) sequences placed within the promoter used (in this case U1 snRNA promoter), thus preventing expression of the gene of interest (in this case vU1 snRNA). Upon addition of tetracycline, this repression is reversed as the tetracycline binds the TetR, therefore allowing vU1 snRNA gene expression.

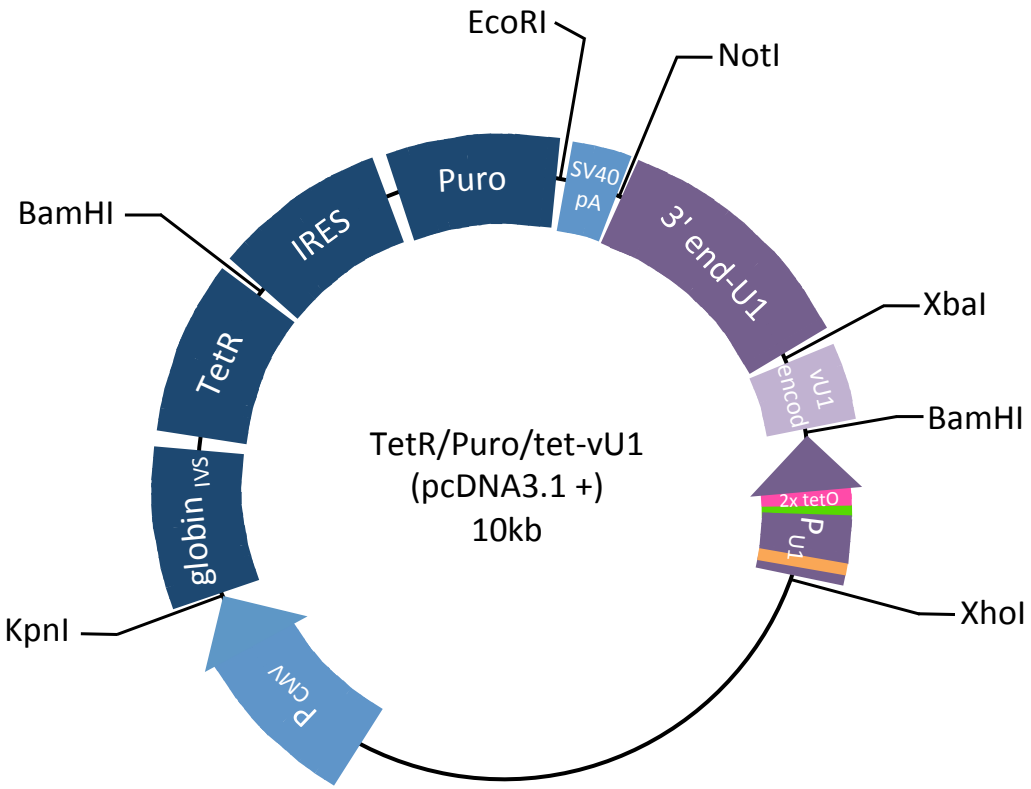


Figure 5-8. Plasmid vector containing the tetracycline repressor-regulated vU1 snRNA gene expression system.

Schematic diagram of the designed plasmid vector, which uses the pcDNA3.1(+) plasmid vector as a backbone and expresses the tetracycline repressor (TetR) on one side and a vU1 snRNA on the other, herein referred to as TetR/Puro/tet-vU1. The expression of TetR is driven by the promoter of the cytomegalovirus (P_{CMV}) and requires the intervening sequence of the β-globin ($globin_{IVS}$); the puromycin (Puro) resistance gene is also expressed under the same promoter and requires for its translation the internal ribosome site (IRES). The polyadenylation signal (pA) of the simian virus (SV)40 ensures termination of the full transcript. Expressed on the opposite direction is the vU1 snRNA, driven by the U1 snRNA promoter and processed by the 3' end of U1 snRNA. The U1 snRNA promoter contains 2x tet operator (tetO) sequences between the PSE (green bar) and the beginning of the vU1 snRNA-encoding region. The restriction sites used for cloning are indicated.

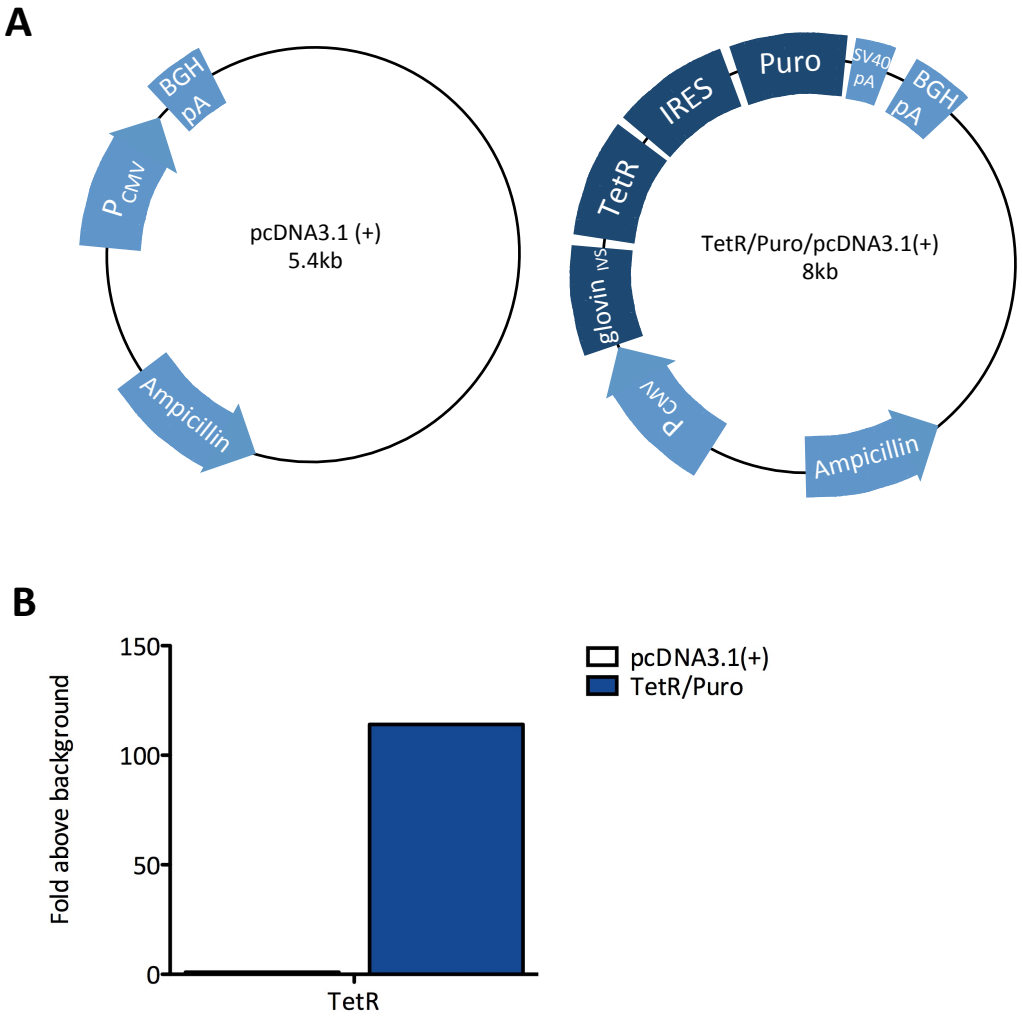


Figure 5-9. Validation of TetR/Puro construct.

(A) Schematic diagrams of the constructs pcDNA3.1(+) (left) and TetR/Puro (pcDNA3.1+) (right). **(B)** Following transfection of pcDNA 3.1(+) or TetR/Puro in HUES2, qPCR analyses, using gene specific primers, indicate expression of TetR. Levels are normalized to 7SK snRNA and relative to the transfection control.

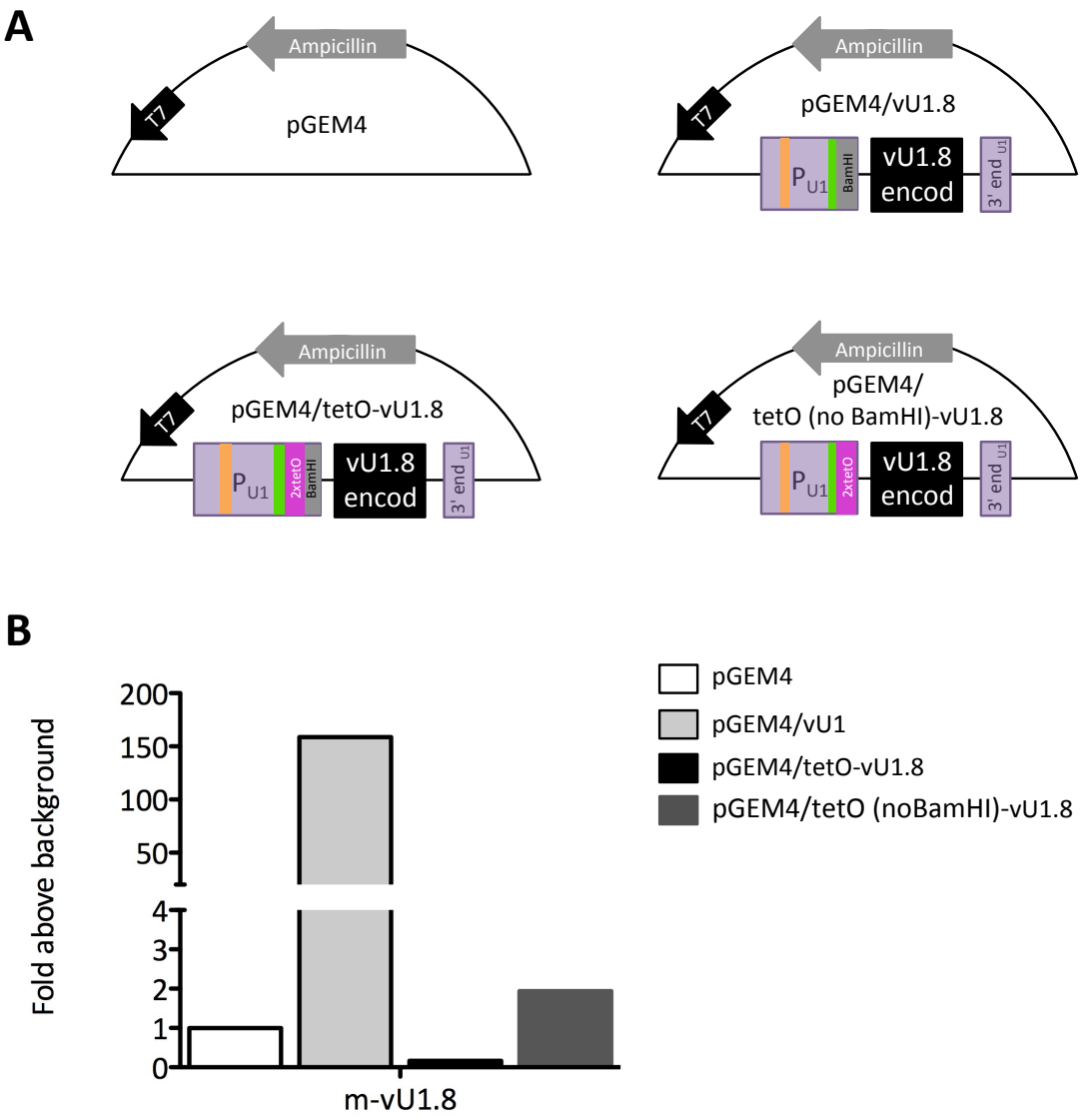


Figure 5-10. Validation of pGEM4/tetO-vU1 construct.

(A) Schematic diagrams of the constructs pGEM4 (top left), pGEM4/vU1.8 (top right), pGEM4/tetO-vU1.8 (bottom left) and pGEM4/tetO (noBamHI)-vU1.8 (bottom right). The black arrow indicates the T7 promoter and the grey arrow represents the ampicillin resistance gene. The orange and green vertical lines within the P_{U1} represent the DSE and PSE, respectively. A BamHI site is located within the P_{U1} (in the case of pGEM4/vU1.8) or outside the P_{U1}, immediately upstream of the vU1 snRNA-encoding region. **(B)** Following transfection of pGEM4 (white bars), pGEM4/vU1.8 (grey bars), pGEM4/tetO-vU1.8 (black bars) or pGEM4/tetO (noBamHI)-vU1.8 (dark grey bars) in HUES2, qPCR analyses, using gene specific primers, indicate expression of vU1.8 snRNA. Levels are normalized to 7SK snRNA and relative to the transfection control.

5.6 Discussion

The manipulation of the expression of genes of interest, including knockdown and overexpression, provides the possibility to investigate the function of such genes in a given cell. In the case of hES cells, this is particularly important as it allows unravelling the potential role of these genes, and thus the signalling pathways, involved in maintaining hES cells in a pluripotent state or in inducing their differentiation. However, genetic manipulation of hES cells is limited by the fact they are a cell type that is hard to transfect, in part due to their low clonogenic capacity and their accessibility, as they tend to grow in tight colonies rather than in a more dispersed manner. Nonetheless, different methods of transfection have achieved various degrees of transfection efficiency and demonstrated successful knockdowns or stable expression of genes of interest. A direct comparison of some of these methods in hES cells demonstrated that lentiviral transduction and nucleofection (a liposomal-based electroporation) might be a better way forward to improve transfection efficiency and hES cell survival rate compared to lipofection and electroporation (Cao et al., 2010).

The transfection method used here was lipofection, which allowed transient transfection of vU1 snRNA expressing constructs (Fig. 5-1) and, to a certain extent, of 2'-O-Methyl phosphorothioate RNA/DNA antisense oligonucleotides, as observed by knockdown of vU1.8 and vU1.20 snRNA (Fig. 5-2, 3). Furthermore, another potential approach to down regulate vU1 snRNA gene expression in hES cells involves the use of a plasmid vector that expresses an antisense vU1 snRNA, which would be able to bind the endogenous sense vU1 snRNA and possibly lead to its degradation.

In the case of the TetR-regulated system for the expression of vU1 snRNAs that has been described here, expression of the TetR was detected but the tetO sequences inserted within the U1 snRNA promoter, driving the expression of vU1.8 snRNA, disrupted its activity. Therefore, it is worth testing whether the vU1 snRNAs can be expressed under the control of the U2 snRNA promoter. Experiments within our laboratory have shown that sequences upstream of the PSE of U2 snRNA can be replaced with the GAL4 binding sites without any effect on promoter activity (Shona Murphy, personal communication). In addition, changing the location of the tetO sequences, either between the DSE and PSE or upstream of the DSE may be a better approach as this was shown to be successful for the U6 promoter (Zhou et al., 2008). While optimizing these tools for the transfection of hES cells I found that transient over expression of vU1 snRNAs leads to an increase in the pluripotency marker OCT4 (Fig. 5-3C). Furthermore when I compared the pattern of vU1 snRNA gene expression in human primary skin fibroblasts and the iPS cells generated from their reprogramming, I observed vU1 snRNA genes were considerably up regulated in iPS cells (Fig. 5-1). These findings confirm that vU1 snRNAs are preferentially expressed in pluripotent stem cells. This is consistent with the fact that both hES and iPS cells showed similar levels of expression for most vU1 snRNAs (Fig. 5-2). The exceptions to this were the levels of vU1.20 snRNA (Fig. 5-2). This perhaps highlights the existence of subtle differences between hES and iPS cells that remain to be addressed.

In conclusion, my results strongly suggest that vU1 snRNAs play a role in the maintenance of pluripotency and in cell reprogramming. Moreover, I have established tools that would allow further manipulation of vU1 snRNA gene expression in hES and iPS cells and thus provide insights into their function.

Chapter 6| GENERAL DISCUSSION

The potential of hES cells to form every cell type in the body makes them hold great promise in the treatment of degenerative diseases or in conditions where tissues are not functional or have been damaged. Moreover, hES cells provide a tool to further study early embryonic development; this is particularly useful for gaining a better understanding of defects that occur during development. Over the past years since the discovery of hES cells in 1998 (Thomson et al., 1998), a considerable amount of research has taken place in order to understand the biology behind these cells. Recently, some of the first clinical trials using hES cell-derived retinal pigment epithelium cells have been initiated (Medina et al., 2011, Schwartz et al., 2012).

Nevertheless, various issues concerning the clinical use of hES cell-derived differentiated cell types remain to be fully addressed. One major limitation is the derivation and culture of hES cells in the presence of xenogeneic and undefined components. Moreover, there is a lack of standardized hES cell culture conditions among the field, thus data from different studies has to be interpreted according to the hES cell culture method used.

Another important issue is the incomplete understanding of the molecular signatures and signalling pathways involved in the maintenance of hES cells and/or their differentiation towards specific cell lineages. Components within the culture trigger downstream signalling pathways that ultimately influence hES cell behaviour. However, most of the key molecular players that have been identified to be involved in hES cell maintenance are genes that encode proteins. There is now increasing evidence that ncRNAs also play a role in influencing hES cell behaviour (Huo and Zambidis, 2012).

In this thesis, I aimed to analyse whether a set of ncRNA genes, the vU1 snRNA genes, newly characterized in our lab (O'Reilly et al., 2012), play a role in maintaining hES cell

pluripotency. For this purpose, I first had to establish a hES cell culture system devoid of feeders, as these would have interfered with further genetic analyses on hES cells. I was then able to show that vU1 snRNA genes are differentially expressed in hES cells and throughout differentiation. These vU1 snRNA genes have been implicated in the 3' end processing of a subset of pre-mRNAs in HeLa cells (O'Reilly et al., 2012). At the time of writing this thesis, the function of vU1 snRNA genes in hES cells was still under investigation.

6.1 Feeder-free hES cell culture system supports hES cells

In order to perform RNA analyses on hES cells, a feeder (mainly MEFs)-free hES cell culture had to be pursued. The approach I took to create such a system consisted of deconstructing the *in vitro* hES cell niche created by such feeders. By dissecting its main components, I managed to establish that the ECM derived from feeders (MEFs and HFFs) provides a sufficient substrate for hES cell culture. Moreover, deconstructing this *in vitro* hES cell niche allowed me to gain some insights into the components that are likely to act as extrinsic factors that would potentially influence hES cell behaviour.

Analyses of the feeder-derived ECMs revealed their varied ECM protein composition. ECM from MEFs, which support hES cell maintenance, showed high expression of the majority of the ECM proteins tested. Furthermore, MEF ECM showed a reversed pattern of ECM protein expression compared to the ECM from feeders that relatively support or do not support hES cell maintenance (Table 3-3). Therefore, this suggests that the expression levels of certain ECM proteins may influence the ability of feeders to support

hES cell maintenance to various degrees or not support it. This is consistent with evidence that supports that the ECM is a dynamic structure that has the potential to influence the behaviour of surrounding cells, perhaps acting as a reservoir of growth factors that ultimately take part in cellular signalling pathways and that in turn determine the biogenesis of the ECM itself (Dallas et al., 2005, Massam-Wu et al., 2010).

In the case of the culture media, I found that only the MEF CM was able to support hES cells in culture. Human ES cells cultured in the presence of CM from non-supporting feeders lost the typical hES cell morphology, as cells appeared to be more granular. In addition, these cells seemed to have lost core pluripotency markers such as SOX2. This strongly suggests the presence of key component(s) in the CM from supporting feeders, which is (are) responsible for hES cell maintenance.

Further work on this area would provide insights into the relevance of particular ECM proteins and/or specific growth factors in hES cell maintenance and/or differentiation. Consequently, unravelling their dynamics in this system may be useful for understanding other stem cell niches or even the microenvironment within tumours.

6.2 Variant U1 snRNA genes are differentially expressed during hES cell differentiation

Work from our laboratory has annotated a set of 21 genes, variants of the U1 snRNA, referred to as vU1 snRNAs on the long arm of chromosome 1 at the locus 1q12-21 (O'Reilly et al., 2012). Variant U1 snRNA genes are actively transcribed and expressed in

HeLa cells and a subset are specifically expressed in the hES cell line HUES2 (O'Reilly et al., 2012). We hypothesized that vU1 snRNA genes play a role in supporting hES cells by means of maintaining their pluripotent state or preventing their differentiation into specific cell types. I have investigated this possibility further.

I have confirmed that the majority of vU1 snRNA genes are expressed at different levels in HUES2; moreover, these levels are comparable to those from other HUES cell lines, HUES1 and HUES4. These hES cells have been derived and grown under identical conditions, so it would be important to extend the analyses to hES cell lines that have been derived under different conditions. Nevertheless, previous studies have revealed that hES cell lines from different backgrounds (derivation and culture) show very similar gene expression patterns when cultured under identical conditions (Bhattacharya et al., 2005, Skottman et al., 2005).

I have also shown that expression of vU1 snRNA genes changes throughout a well-established gradual differentiation protocol of HUES2 into macrophages (Karlsson et al., 2008). I observed that the majority of vU1 snRNA genes are slightly down regulated at the first stage of differentiation into EBs (Fig. 4-5). Thereafter, vU1 snRNA gene expression continues to be down regulated at the monocyte and macrophage stages, with the exception of vU1.8 snRNA, whose expression is possibly up regulated at the macrophage stage (Fig. 4-5). Moreover, analyses of the steady-state levels of vU1 snRNAs (herein referred to as m-vU1 snRNAs) followed a similar expression profile (Fig. 4-6). Steady state levels continue to reduce, for all vU1 snRNAs tested, up to the monocyte stage in line with a corresponding drop in nascent levels. This would suggest that expression of the vU1 snRNA genes is controlled at the level of transcription during these stages of the hES

cell differentiation process. However, following directed differentiation of monocytes to macrophages the nascent levels continue to drop for most vU1 snRNAs but, interestingly, the steady state levels showed an increase. This would suggest that mechanisms which control 3' processing and/or factors that improve the stability of the vU1 snRNAs are key players in regulating vU1 snRNA steady state levels in these cells.

Similar analysis, mapping the expression profile of the U1 snRNA, indicated that cells expressing the highest levels of vU1 snRNAs (i.e. HUES2) express the lowest level of U1 snRNA and conversely, U1 snRNA steady state levels are maximal in monocytes, which express poor levels of vU1snRNAs. This suggests that mechanisms are in place to control the relative levels of U1 snRNA to vU1 snRNAs. Considering the known functional properties of these ncRNAs, variations in their levels could have profound effects on mRNA isoforms expressed in different cells and during different stages of development.

These dynamics in the pattern of expression of vU1 snRNAs are consistent with the changes in gene expression that are observed upon hES cell differentiation and also during early human embryonic development, where specific sets of genes are down regulated and others up regulated (Brandenberger et al., 2004, Calhoun et al., 2004, Dobson et al., 2004, Dvash et al., 2004, Miura et al., 2004, Cai et al., 2006, Xu et al., 2009b, Fathi et al., 2011).

Future work would be focused on understanding the mechanisms that regulate such patterns of gene expression throughout hES cell differentiation. This would ultimately allow us to direct differentiation of hES cells into specific cell lineages.

6.3 Established tools to manipulate vU1 snRNA gene expression

After establishing a pattern of expression of vU1 snRNA genes in hES cells and their differentiated progeny, the next step was to assess whether manipulating these vU1 snRNA genes would have implications for hES cell maintenance/differentiation. The expression of vU1 snRNA genes mainly in hES cells provides an indication of the potential role of vU1 snRNA genes in maintaining the pluripotent state of these cells.

Despite hES cells possess low capacity for transfection I had managed to introduce, via lipofection, vU1 snRNA gene-expressing constructs in hES cells, as well as 2'-O-Methyl phosphorothioate RNA/DNA antisense oligonucleotides. Transfection of the latter led to down regulation of vU1.8 and vU1.20 snRNAs specifically. RNA from these samples was prepared for further RNA sequencing analysis in order to analyse the global effect of vU1 snRNAs down regulation. At the moment of writing this thesis, this analysis is still ongoing.

Following transient transfection of vU1 snRNA gene-expressing constructs (O'Reilly et al., 2012) in hES cells it was noticed that levels of the pluripotency marker OCT4 appeared to be up regulated. Moreover, analyses of the vU1 snRNA gene expression in iPS cells and the primary human skin fibroblasts they originated from revealed that vU1 snRNA genes are preferentially expressed in iPS cells. These data strongly supports the idea that vU1 snRNAs play a role in maintaining pluripotent stem cells (hES and iPS cells) and perhaps contribute to iPS cell reprogramming.

In addition, given the successful transfection of gene-expressing constructs in hES cells, I focused on building various constructs that would allow also manipulating the expression

of vU1 snRNA genes in these cells. The first construct consists of a vU1 snRNA, whose expression would be regulated in a tetracycline-inducible manner. For this purpose, two separate constructs were built, one expressing a vU1snRNA under the control of the U1 snRNA promoter and 3' box and one expressing the TetR. The U1 snRNA promoter was designed with two tet operator sequences, which would be recognized by the TetR and thus prevent vU1 snRNA expression. However, the insertion of these sequences disrupted the expression of vU1 snRNA genes (Fig. 5-9). Therefore future work should focus on utilizing another promoter such as the U6 snRNA promoter, which has been shown to be active following the insertion of tet operator sequences (Lin et al., 2004, Ohkawa and Taira, 2000, Zhou et al., 2008). Such a tet-inducible vU1snRNA gene expression system would potentially allow the sustained expression of vU1 snRNA genes in hES cells or iPS cells, and thus assess whether this prevents their gradual differentiation into macrophages or aid iPS cell reprogramming, respectively.

Additionally, I created a second construct with the potential to knock down vU1 snRNA gene expression. This construct is similar to the vU1 snRNA-gene expressing constructs but differs in that the vU1 snRNA-encoding region has been replaced with an antisense vU1 snRNA-encoding region that expresses the antisense strand of the vU1 snRNA. I hypothesised that this antisense vU1 snRNA would then be able to bind the corresponding sense endogenous vU1 snRNA and target their degradation. Preliminary data suggests vU1.8 snRNA can be down regulated using this approach (Fig. 5-5).

Thus, I have here established some methods for the manipulation of vU1 snRNA gene expression in hES cells and iPS cells, which will aid the elucidation of the role of vU1 snRNA genes in maintaining the pluripotent state of these cells.

6.4 Future directions

6.4.1 Proteomic analyses of feeder-derived ECM substrate and feeder CM

The protein analyses of the feeder-derived ECM substrates that was presented here was restricted to key ECM proteins only. Therefore, it would be important to expand the analyses of the ECM protein composition of feeder-derived ECM substrates by performing proteomic analyses. Further validation of the contribution of individual ECM proteins to hES cell maintenance would include: 1) growing hES cells on purified ECM proteins and 2) blocking individual ECM proteins in feeder-derived ECM using specific antibodies, followed by assessment of hES cell growth on such substrates.

In addition, I have shown that it is the CM from supporting and non-supporting feeders that, irrespective of the feeder-derived ECM substrate, supports or does not support hES cells in culture, respectively. Proteomic analyses of the full protein composition of the CM from supporting and non-supporting feeders should also provide insights on those critical components. These then would then be validated by selective deletion/addition of these components in the corresponding feeder CM and assessing their ability to support hES cells in culture.

6.4.2 Assessment of vU1 snRNA gene expression in different cell types

The analyses of the expression of vU1 snRNA genes that was performed here was restricted to some of the HUES hES cell lines. Various studies have previously shown that different hES cell lines, including HUES cell lines, have different propensity to

differentiate towards a specific cell type or another (Pal et al., 2009, Tavakoli et al., 2009, Mehta et al., 2010). Therefore, it would be useful to extend the analysis of vU1 snRNA gene expression to other hES cell lines and other differentiated cell lineages

Moreover, analyses of vU1 snRNA expression could also be investigated in normal and diseased tissues. Using iPS cell technology would allow modelling certain diseases; somatic tissue (fibroblasts) from patients could be reprogrammed into iPS cells and in turn be directed to form the cell type affected in the disease. Such is the case of motor neurons in patients with spinal muscular atrophy (SMA), where one of the genes that encodes for SMN protein, involved in the assembly of Sm proteins into U snRNPs, (*SMN1*) presents a single nucleotide change that, like SMN2, leads to the formation of a truncated protein that gets degraded, thus ultimately causing the gradual death of motor neurons. In addition, it would be worth analysing whether vU1 snRNAs could account for the function of U1 snRNA in SMN deficient neurons, given some vU1 snRNAs possess mutations at their Sm binding site and thus may still function in the absence of SMN in patients with SMA.

This type of experiment would provide further insights into the role of specific vU1 snRNAs.

6.4.3 Correlation vU1 snRNA expression to specific gene targets in hES cells

The work from our laboratory, following interrogation of the Affymetrix Human Exon ST 1.0 array after knock down of vU1.8 in HeLa cells, revealed premature 3' end processing of various gene targets. Analyses of these genes could be undertaken for complementary

sequences to vU1 snRNAs. In order to assess whether they recognize atypical 5' splice sites, minigene- reporter constructs could be prepared. These constructs would contain an exon and flanking introns of the genes of interest with a consensus 5' splice site recognized by U1 snRNP; using suppressor U1 and vU1 snRNAs we can then determine whether premature 3' end processing is a direct result of vU1 base pairing with the 5' splice site.

6.4.4 Identification of the regulatory elements that control vU1 snRNA gene expression in hES cells

I have here established that the pattern of vU1 snRNA gene expression is dynamic throughout hES cell differentiation. Further work should be focused on understanding the mechanisms behind this differential vU1 snRNA gene expression.

It has been suggested that the differential expression of the mouse adult and embryonic U1 snRNA isoforms (U1a and U1b, respectively) is due to differences in the activity of their promoters (Caceres et al., 1992). They reached this conclusion after introducing chimeric genes that contain 1) U1a promoter and U1b snRNA-encoding region and 2) U1b promoter combined with human U1 snRNA-encoding region, in cells that do not normally express U1b snRNA (Caceres et al., 1992).

This approach could also be applied to analysing the regulation of vU1 snRNA gene expression, by means of exchanging the promoter of U1 snRNA gene with that of vU1 snRNA genes and analysing their expression in cell types that hardly express vU1 snRNA genes (monocytes or primary human skin fibroblasts). In addition, it would be important

to assess whether the stable transfection of hES cells with vU1 snRNAs under the control of U1 snRNA promoter revert the differential vU1 snRNA gene expression observed during hES cell differentiation into macrophages.

More detailed analyses of the sequences within the vU1 snRNA promoters required for their activity would involve the creation of mutations within vU1 snRNA promoters and assessing their activity using a GFP reporter system.

Finally, it would be important to investigate the factors that bind specific elements within the vU1 snRNA promoters. For instance, given that the OCT-1 transcription factor is known to bind the U1 snRNA gene promoter in HeLa cells, ChIP analyses would reveal whether OCT-4, a core transcription factor of pluripotency, is recruited to the vU1 snRNA promoters in hES cells.

6.4.5 Characterization of the snRNP complexes associated with vU1 snRNAs

The expression analyses of vU1 snRNA genes in hES cells and their differentiated progeny showed that some of these genes are being processed into stable forms at each stage of differentiation. However, the nature of vU1 snRNP complexes requires further investigation.

The work of our laboratory has reported that selected vU1 snRNAs are found in snRNP complexes following immunoprecipitation with anti-Sm antibodies in HeLa and hES cells (O'Reilly et al., 2012). Moreover, in HeLa cells these vU1 snRNAs associate with the protein U1-C but not with U1-70K and U1-A, suggesting vU1 snRNAs are likely to interact

with different snRNP complexes (O'Reilly et al., 2012). It will still remain to establish if this is also the case for hES cells.

It would also be informative to perform RNA fluorescence in situ hybridization (FISH) analyses combined with the detection of cellular proteins (RNA immuno-FISH) (Namekawa and Lee, 2011), in this case snRNPs associated with vU1 snRNAs, and thus gain an indication of the spatial and temporal expression of vU1 snRNA genes.

6.4.6 Manipulation of vU1 snRNA levels during iPS cell reprogramming

The data presented here showed that expression of vU1 snRNA genes is up regulated upon reprogramming of primary human skin fibroblasts into iPS cells, suggesting a potential role in establishing the pluripotent state of iPS cells. To explore this possibility, manipulation of vU1 snRNA gene expressing during somatic cell reprogramming into iPS cells may be performed. By transfecting vU1 snRNA gene-expressing constructs, containing sense or antisense vU1 snRNAs, in the presence/absence of constructs expressing the pluripotency markers required for reprogramming into primary human somatic cells (skin fibroblasts), the potential effect of vU1 snRNAs in iPS cell reprogramming could be elucidated.

6.5 Summary

The work presented in this thesis provides evidence to support that extrinsic (ECM proteins) and intrinsic factors (vU1 snRNA genes) are potentially involved in controlling

the pluripotent state of hES cells and thus prevent their differentiation. Further manipulation of these factors in human pluripotent stem cells (hES and iPS cells) would reveal specific gene targets, which would in turn provide a better understanding of the signalling pathways of pluripotency, differentiation or reprogramming. Ultimately, this information would provide insights into early developmental processes and the onset of disease.

REFERENCES

ABAD, X., VERA, M., JUNG, S. P., OSWALD, E., ROMERO, I., AMIN, V., FORTES, P. & GUNDERSON, S. I. 2008. Requirements for gene silencing mediated by U1 snRNA binding to a target sequence. *Nucleic Acids Res,* 36**,** 2338-52.

ABDEL-RAHMAN, B., FIDDLER, M., RAPPOLEE, D. & PERGAMENT, E. 1995. Expression of transcription regulating genes in human preimplantation embryos. *Hum Reprod,* 10**,** 2787-92.

ADACHI, K., SUEMORI, H., YASUDA, S. Y., NAKATSUJI, N. & KAWASE, E. 2010. Role of SOX2 in maintaining pluripotency of human embryonic stem cells. *Genes Cells,* 15**,** 455-70.

ALEXANDER, M. R., WHEATLEY, A. K., CENTER, R. J. & PURCELL, D. F. 2010. Efficient transcription through an intron requires the binding of an Sm-type U1 snRNP with intact stem loop II to the splice donor. *Nucleic Acids Res,* 38**,** 3041-53.

ANDREWS, P. W. 1998. Teratocarcinomas and human embryology: pluripotent human EC cell lines. Review article. *APMIS,* 106**,** 158-67; discussion 167-8.

ANDREWS, P. W., DAMJANOV, I., SIMON, D., BANTING, G. S., CARLIN, C., DRACOPOLI, N. C. & FOGH, J. 1984. Pluripotent embryonal carcinoma clones derived from the human teratocarcinoma cell line Tera-2. Differentiation in vivo and in vitro. *Lab Invest,* 50**,** 147-62.

ARMSTRONG, L., HUGHES, O., YUNG, S., HYSLOP, L., STEWART, R., WAPPLER, I., PETERS, H., WALTER, T., STOJKOVIC, P., EVANS, J., STOJKOVIC, M. & 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**,** 1894-913.

ASHE, M. P., GRIFFIN, P., JAMES, W. & PROUDFOOT, N. J. 1995. Poly(A) site selection in the HIV-1 provirus: inhibition of promoter-proximal polyadenylation by the downstream major splice donor site. *Genes Dev,* 9**,** 3008-25.

AVILION, A. A., NICOLIS, S. K., PEVNY, L. H., PEREZ, L., VIVIAN, N. & LOVELL-BADGE, R. 2003. Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev,* 17**,** 126-40.

BABAIE, Y., HERWIG, R., GREBER, B., BRINK, T. C., WRUCK, W., GROTH, D., LEHRACH, H., BURDON, T. & ADJAYE, J. 2007. Analysis of Oct4-dependent transcriptional networks regulating self-renewal and pluripotency in human embryonic stem cells. *Stem Cells,* 25**,** 500-10.

BEATTIE, G. M., LOPEZ, A. D., BUCAY, N., HINTON, A., FIRPO, M. T., KING, C. C. & HAYEK, A. 2005. Activin A maintains pluripotency of human embryonic stem cells in the absence of feeder layers. *Stem Cells,* 23**,** 489-95.

BECKER, A. J., MC, C. E. & TILL, J. E. 1963. Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. *Nature,* 197**,** 452-4.

BERG, M. G., SINGH, L. N., YOUNIS, I., LIU, Q., PINTO, A. M., KAIDA, D., ZHANG, Z., CHO, S., SHERRILL-MIX, S., WAN, L. & DREYFUSS, G. 2012. U1 snRNP determines mRNA length and regulates isoform expression. *Cell,* 150**,** 53-64.

BERGLUND, J. A., ABOVICH, N. & ROSBASH, M. 1998. A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes Dev,* 12**,** 858-67.

BERTOLOTTI, A., LUTZ, Y., HEARD, D. J., CHAMBON, P. & TORA, L. 1996. hTAF(II)68, a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins

TLS/FUS and EWS is associated with both TFIID and RNA polymerase II. *EMBO J,* 15**,** 5022-31.

BHATTACHARYA, B., CAI, J., LUO, Y., MIURA, T., MEJIDO, J., BRIMBLE, S. N., ZENG, X., SCHULZ, T. C., RAO, M. S. & PURI, R. K. 2005. Comparison of the gene expression profile of undifferentiated human embryonic stem cell lines and differentiating embryoid bodies. *BMC Dev Biol,* 5**,** 22.

BOCK, C., KISKINIS, E., VERSTAPPEN, G., GU, H., BOULTING, G., SMITH, Z. D., ZILLER, M., CROFT, G. F., AMOROSO, M. W., OAKLEY, D. H., GNIRKE, A., EGGAN, K. & MEISSNER, A. 2011. Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell,* 144**,** 439-52.

BRAAM, S. R., ZEINSTRA, L., LITJENS, S., WARD-VAN OOSTWAARD, D., VAN DEN BRINK, S., VAN LAAKE, L., LEBRIN, F., KATS, P., HOCHSTENBACH, R., PASSIER, R., SONNENBERG, A. & MUMMERY, C. L. 2008. Recombinant vitronectin is a functionally defined substrate that supports human embryonic stem cell self-renewal via alphavbeta5 integrin. *Stem Cells,* 26**,** 2257-65.

BRANDENBERGER, R., WEI, H., ZHANG, S., LEI, S., MURAGE, J., FISK, G. J., LI, Y., XU, C., FANG, R., GUEGLER, K., RAO, M. S., MANDALAM, R., LEBKOWSKI, J. & STANTON, L. W. 2004. Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation. *Nat Biotechnol,* 22**,** 707-16.

BROOK, F. A., COWLEY, S. A., EVANS, E. P., TURNER, K., JAMES, W. S. & MARDON, H. J. 2010. Derivation and characterisation of the human embryonic stem cell line, OxF1. *In Vitro Cell Dev Biol Anim,* 46**,** 173-7.

CACERES, J. F., MCKENZIE, D., THIMMAPAYA, R., LUND, E. & DAHLBERG, J. E. 1992. Control of mouse U1a and U1b snRNA gene expression by differential transcription. *Nucleic Acids Res,* 20**,** 4247-54.

CAI, Y., WU, P., OZEN, M., YU, Y., WANG, J., ITTMANN, M. & LIU, M. 2006. Gene expression profiling and analysis of signaling pathways involved in priming and differentiation of human neural stem cells. *Neuroscience,* 138**,** 133-48.

CALHOUN, J. D., RAO, R. R., WARRENFELTZ, S., REKAYA, R., DALTON, S., MCDONALD, J. & STICE, S. L. 2004. Transcriptional profiling of initial differentiation events in human embryonic stem cells. *Biochem Biophys Res Commun,* 323**,** 453-64.

CAO, F., XIE, X., GOLLAN, T., ZHAO, L., NARSINH, K., LEE, R. J. & WU, J. C. 2010. Comparison of gene-transfer efficiency in human embryonic stem cells. *Mol Imaging Biol,* 12**,** 15-24.

CHAMBERS, I., COLBY, D., ROBERTSON, M., NICHOLS, J., LEE, S., TWEEDIE, S. & SMITH, A. 2003. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell,* 113**,** 643-55.

CHAMBERS, I., SILVA, J., COLBY, D., NICHOLS, J., NIJMEIJER, B., ROBERTSON, M., VRANA, J., JONES, K., GROTEWOLD, L. & SMITH, A. 2007. Nanog safeguards pluripotency and mediates germline development. *Nature,* 450**,** 1230-4.

CHEN, H. F., CHUANG, C. Y., SHIEH, Y. K., CHANG, H. W., HO, H. N. & KUO, H. C. 2009. Novel autogenic feeders derived from human embryonic stem cells (hESCs) support an undifferentiated status of hESCs in xeno-free culture conditions. *Hum Reprod,* 24**,** 1114-25.

CHERNY, R. A., STOKES, T. M., MEREI, J., LOM, L., BRANDON, M. R. & WILLIAMS, R. L. 1994. Strategies for the isolation and characterization of bovine embryonic stem cells. *Reprod Fertil Dev,* 6**,** 569-75.

CHEW, J. L., LOH, Y. H., ZHANG, W., CHEN, X., TAM, W. L., YEAP, L. S., LI, P., ANG, Y. S., LIM, B., ROBSON, P. & NG, H. H. 2005. Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol,* 25**,** 6031-46.

CHIN, A. C., FONG, W. J., GOH, L. T., PHILP, R., OH, S. K. & CHOO, A. B. 2007. Identification of proteins from feeder conditioned medium that support human embryonic stem cells. *J Biotechnol,* 130**,** 320-8.

CHIN, A. C., PADMANABHAN, J., OH, S. K. & CHOO, A. B. 2010. Defined and serum-free media support undifferentiated human embryonic stem cell growth. *Stem Cells Dev,* 19**,** 753-61.

COHEN, M. A., ITSYKSON, P. & REUBINOFF, B. E. 2010. The role of FGF-signaling in early neural specification of human embryonic stem cells. *Dev Biol,* 340**,** 450-8.

COSTELLO, I., BIONDI, C. A., TAYLOR, J. M., BIKOFF, E. K. & ROBERTSON, E. J. 2009. Smad4-dependent pathways control basement membrane deposition and endodermal cell migration at early stages of mouse development. *BMC Dev Biol,* 9**,** 54.

COTSARELIS, G., SUN, T. T. & LAVKER, R. M. 1990. Label-retaining cells reside in the bulge area of pilosebaceous unit: implications for follicular stem cells, hair cycle, and skin carcinogenesis. *Cell,* 61**,** 1329-37.

CRISPINO, J. D., BLENCOWE, B. J. & SHARP, P. A. 1994. Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. *Science,* 265**,** 1866-9.

DALLAS, S. L., SIVAKUMAR, P., JONES, C. J., CHEN, Q., PETERS, D. M., MOSHER, D. F., HUMPHRIES, M. J. & KIELTY, C. M. 2005. Fibronectin regulates latent transforming growth factor-beta (TGF beta) by controlling matrix assembly of latent TGF beta-binding protein-1. *J Biol Chem,* 280**,** 18871-80.

DARR, H., MAYSHAR, Y. & BENVENISTY, N. 2006. Overexpression of NANOG in human ES cells enables feeder-free growth while inducing primitive ectoderm features. *Development,* 133**,** 1193-201.

DAS, A. T., ZHOU, X., VINK, M., KLAVER, B., VERHOEF, K., MARZIO, G. & BERKHOUT, B. 2004. Viral evolution as a tool to improve the tetracycline-regulated gene expression system. *J Biol Chem,* 279**,** 18776-82.

DAS, G., HINKLEY, C. S. & HERR, W. 1995. Basal promoter elements as a selective determinant of transcriptional activator function. *Nature,* 374**,** 657-60.

DAVIDSON, K. C., ADAMS, A. M., GOODSON, J. M., MCDONALD, C. E., POTTER, J. C., BERNDT, J. D., BIECHELE, T. L., TAYLOR, R. J. & 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**,** 4485-90.

DENISON, R. A., VAN ARSDELL, S. W., BERNSTEIN, L. B. & WEINER, A. M. 1981. Abundant pseudogenes for small nuclear RNAs are dispersed in the human genome. *Proceedings of the National Academy of Sciences of the United States of America,* 78**,** 810-4.

DENISON, R. A. & WEINER, A. M. 1982. Human U1 RNA pseudogenes may be generated by both DNA- and RNA-mediated mechanisms. *Mol Cell Biol,* 2**,** 815-28.

DERYNCK, R. & ZHANG, Y. E. 2003. Smad-dependent and Smad-independent pathways in TGF-beta family signalling. *Nature,* 425**,** 577-84.

DING, V. M., LING, L., NATARAJAN, S., YAP, M. G., COOL, S. M. & CHOO, A. B. 2010. FGF-2 modulates Wnt signaling in undifferentiated hESC and iPS cells through activated PI3-K/GSK3beta signaling. *J Cell Physiol,* 225**,** 417-28.

DOBSON, A. T., RAJA, R., ABEYTA, M. J., TAYLOR, T., SHEN, S., HAQQ, C. & PERA, R. A. 2004. The unique transcriptome through day 3 of human preimplantation development. *Hum Mol Genet,* 13**,** 1461-70.

DU, H. & ROSBASH, M. 2002. The U1 snRNP protein U1C recognizes the 5' splice site in the absence of base pairing. *Nature,* 419**,** 86-90.

DVASH, T., MAYSHAR, Y., DARR, H., MCELHANEY, M., BARKER, D., YANUKA, O., KOTKOW, K. J., RUBIN, L. L., BENVENISTY, N. & EIGES, R. 2004. Temporal gene expression during differentiation of human embryonic stem cells and embryoid bodies. *Hum Reprod,* 19**,** 2875-83.

EGLOFF, S., O'REILLY, D., CHAPMAN, R. D., TAYLOR, A., TANZHAUS, K., PITTS, L., EICK, D. & MURPHY, S. 2007. Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science,* 318**,** 1777-9.

EISELLEOVA, L., MATULKA, K., KRIZ, V., KUNOVA, M., SCHMIDTOVA, Z., NERADIL, J., TICHY, B., DVORAKOVA, D., POSPISILOVA, S., HAMPL, A. & DVORAK, P. 2009. A complex role for FGF-2 in self-renewal, survival, and adhesion of human embryonic stem cells. *Stem Cells,* 27**,** 1847-57.

EVANS, M. J. & KAUFMAN, M. H. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature,* 292**,** 154-6.

EVSEENKO, D., SCHENKE-LAYLAND, K., DRAVID, G., ZHU, Y., HAO, Q. L., SCHOLES, J., WANG, X. C., MACLELLAN, W. R. & CROOKS, G. M. 2009. Identification of the critical extracellular matrix proteins that promote human embryonic stem cell assembly. *Stem Cells Dev,* 18**,** 919-28.

FATHI, A., HATAMI, M., HAJIHOSSEINI, V., FATTAHI, F., KIANI, S., BAHARVAND, H. & SALEKDEH, G. H. 2011. Comprehensive gene expression analysis of human embryonic stem cells during differentiation into neural cells. *PLoS One,* 6**,** e22856.

FISCHER, U., LIU, Q. & DREYFUSS, G. 1997. The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. *Cell,* 90**,** 1023-9.

FISCHER, Y., GANIC, E., AMERI, J., XIAN, X., JOHANNESSON, M. & SEMB, H. 2010. NANOG reporter cell lines generated by gene targeting in human embryonic stem cells. *PLoS One,* 5.

FLETCHER, J. M., FERRIER, P. M., GARDNER, J. O., HARKNESS, L., DHANJAL, S., SERHAL, P., HARPER, J., DELHANTY, J., BROWNSTEIN, D. G., PRASAD, Y. R., LEBKOWSKI, J., MANDALAM, R., WILMUT, I. & DE SOUSA, P. A. 2006. Variations in humanized and defined culture conditions supporting derivation of new human embryonic stem cell lines. *Cloning Stem Cells,* 8**,** 319-34.

FOGH, J., FOGH, J. M. & ORFEO, T. 1977. One hundred and twenty-seven cultured human tumor cell lines producing tumors in nude mice. *J Natl Cancer Inst,* 59**,** 221-6.

FONG, H., HOHENSTEIN, K. A. & DONOVAN, P. J. 2008. Regulation of self-renewal and pluripotency by Sox2 in human embryonic stem cells. *Stem Cells,* 26**,** 1931-8.

FONG, Y. W. & ZHOU, Q. 2001. Stimulatory effect of splicing factors on transcriptional elongation. *Nature,* 414**,** 929-33.

FORBES, D. J., KIRSCHNER, M. W., CAPUT, D., DAHLBERG, J. E. & LUND, E. 1984. Differential expression of multiple U1 small nuclear RNAs in oocytes and embryos of Xenopus laevis. *Cell,* 38**,** 681-9.

FORTES, P., CUEVAS, Y., GUAN, F., LIU, P., PENTLICKY, S., JUNG, S. P., MARTINEZ-CHANTAR, M. L., PRIETO, J., ROWE, D. & GUNDERSON, S. I. 2003. Inhibiting expression of specific genes in mammalian cells with 5' end-mutated U1 small

nuclear RNAs targeted to terminal exons of pre-mRNA. *Proc Natl Acad Sci U S A,* 100**,** 8264-9.

GABUT, M., SAMAVARCHI-TEHRANI, P., WANG, X., SLOBODENIUC, V., O'HANLON, D., SUNG, H. K., ALVAREZ, M., TALUKDER, S., PAN, Q., MAZZONI, E. O., NEDELEC, S., WICHTERLE, H., WOLTJEN, K., HUGHES, T. R., ZANDSTRA, P. W., NAGY, A., WRANA, J. L. & BLENCOWE, B. J. 2011. An alternative splicing switch regulates embryonic stem cell pluripotency and reprogramming. *Cell,* 147**,** 132-46.

GARCION, E., HALILAGIC, A., FAISSNER, A. & FFRENCH-CONSTANT, C. 2004. Generation of an environmental niche for neural stem cell development by the extracellular matrix molecule tenascin C. *Development,* 131**,** 3423-32.

GATTI, R. A., MEUWISSEN, H. J., ALLEN, H. D., HONG, R. & GOOD, R. A. 1968. Immunological reconstitution of sex-linked lymphopenic immunological deficiency. *Lancet,* 2**,** 1366-9.

GOSSEN, M. & BUJARD, H. 1992. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci U S A,* 89**,** 5547-51.

GOZANI, O., POTASHKIN, J. & REED, R. 1998. A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. *Mol Cell Biol,* 18**,** 4752-60.

GRABOWSKI, P. J., SEILER, S. R. & SHARP, P. A. 1985. A multicomponent complex is involved in the splicing of messenger RNA precursors. *Cell,* 42**,** 345-53.

GREBER, B., LEHRACH, H. & ADJAYE, J. 2007. Silencing of core transcription factors in human EC cells highlights the importance of autocrine FGF signaling for self-renewal. *BMC Dev Biol,* 7**,** 46.

GRONTHOS, S., MANKANI, M., BRAHIM, J., ROBEY, P. G. & SHI, S. 2000. Postnatal human dental pulp stem cells (DPSCs) in vitro and in vivo. *Proc Natl Acad Sci U S A,* 97**,** 13625-30.

GUBBAY, J., COLLIGNON, J., KOOPMAN, P., CAPEL, B., ECONOMOU, A., MUNSTERBERG, A., VIVIAN, N., GOODFELLOW, P. & LOVELL-BADGE, R. 1990. A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature,* 346**,** 245-50.

GUNDERSON, S. I., POLYCARPOU-SCHWARZ, M. & MATTAJ, I. W. 1998. U1 snRNP inhibits pre-mRNA polyadenylation through a direct interaction between U1 70K and poly(A) polymerase. *Mol Cell,* 1**,** 255-64.

GURDON, J. B. & UEHLINGER, V. 1966. "Fertile" intestine nuclei. *Nature,* 210**,** 1240-1.

HANLEY, B. A. & SCHULER, M. A. 1991. Developmental expression of plant snRNAs. *Nucleic Acids Res,* 19**,** 6319-25.

HART, A. H., HARTLEY, L., IBRAHIM, M. & ROBB, L. 2004. Identification, cloning and expression analysis of the pluripotency promoting Nanog genes in mouse and human. *Dev Dyn,* 230**,** 187-98.

HAY, D. C., SUTHERLAND, L., CLARK, J. & 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**,** 225-35.

HAYASHI, F., ISHIMA, R., LIU, D., TONG, K. I., KIM, S., REINBERG, D., BAGBY, S. & IKURA, M. 1998. Human general transcription factor TFIIB: conformational variability and interaction with VP16 activation domain. *Biochemistry,* 37**,** 7941-51.

HIROSE, T., IDEUE, T., NAGAI, M., HAGIWARA, M., SHU, M. D. & STEITZ, J. A. 2006. A spliceosomal intron binding protein, IBP160, links position-dependent assembly of intron-encoded box C/D snoRNP to pre-mRNA splicing. *Mol Cell,* 23**,** 673-84.

HONGISTO, H., VUORISTO, S., MIKHAILOVA, A., SUURONEN, R., VIRTANEN, I., OTONKOSKI, T. & SKOTTMAN, H. 2012. Laminin-511 expression is associated with the functionality of feeder cells in human embryonic stem cell culture. *Stem Cell Res,* 8**,** 97-108.

HUMPHREY, R. K., BEATTIE, G. M., LOPEZ, A. D., BUCAY, N., KING, C. C., FIRPO, M. T., ROSE-JOHN, S. & HAYEK, A. 2004. Maintenance of pluripotency in human embryonic stem cells is STAT3 independent. *Stem Cells,* 22**,** 522-30.

HUO, J. S. & ZAMBIDIS, E. T. 2012. Pivots of pluripotency: The roles of non-coding RNA in regulating embryonic and induced pluripotent stem cells. *Biochim Biophys Acta*.

HYSLOP, L., STOJKOVIC, M., ARMSTRONG, L., WALTER, T., STOJKOVIC, P., PRZYBORSKI, S., HERBERT, M., MURDOCH, A., STRACHAN, T. & LAKO, M. 2005. Downregulation of NANOG induces differentiation of human embryonic stem cells to extraembryonic lineages. *Stem Cells,* 23**,** 1035-43.

IANNACCONE, P. M., TABORN, G. U., GARTON, R. L., CAPLICE, M. D. & BRENIN, D. R. 1994. Pluripotent embryonic stem cells from the rat are capable of producing chimeras. *Dev Biol,* 163**,** 288-92.

IDEUE, T., HINO, K., KITAO, S., YOKOI, T. & HIROSE, T. 2009. Efficient oligonucleotide-mediated degradation of nuclear noncoding RNAs in mammalian cultured cells. *RNA,* 15**,** 1578-87.

ILIC, D., GIRITHARAN, G., ZDRAVKOVIC, T., CACERES, E., GENBACEV, O., FISHER, S. J. & KRTOLICA, A. 2009. Derivation of Human Embryonic Stem Cell Lines From Biopsied Blastomeres on Human Feeders With Minimal Exposure to Xenomaterials. *Stem Cells and Development,* 18**,** 1343-1350.

JOBERT, L., ARGENTINI, M. & TORA, L. 2009. PRMT1 mediated methylation of TAF15 is required for its positive gene regulatory function. *Exp Cell Res,* 315**,** 1273-86.

JONK, L. J., ITOH, S., HELDIN, C. H., TEN DIJKE, P. & KRUIJER, W. 1998. Identification and functional characterization of a Smad binding element (SBE) in the JunB promoter that acts as a transforming growth factor-beta, activin, and bone morphogenetic protein-inducible enhancer. *J Biol Chem,* 273**,** 21145-52.

JUNG, M., PETERSON, H., CHAVEZ, L., KAHLEM, P., LEHRACH, H., VILO, J. & ADJAYE, J. 2010. A data integration approach to mapping OCT4 gene regulatory networks operative in embryonic stem cells and embryonal carcinoma cells. *PLoS One,* 5**,** e10709.

KAHAN, B. W. & EPHRUSSI, B. 1970. Developmental potentialities of clonal in vitro cultures of mouse testicular teratoma. *J Natl Cancer Inst,* 44**,** 1015-36.

KAIDA, D., BERG, M. G., YOUNIS, I., KASIM, M., SINGH, L. N., WAN, L. & DREYFUSS, G. 2010. U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature,* 468**,** 664-8.

KARLSSON, K. R., COWLEY, S., MARTINEZ, F. O., SHAW, M., MINGER, S. L. & JAMES, W. 2008. Homogeneous monocytes and macrophages from human embryonic stem cells following coculture-free differentiation in M-CSF and IL-3. *Exp Hematol,* 36**,** 1167-75.

KARUS, M., DENECKE, B., FFRENCH-CONSTANT, C., WIESE, S. & FAISSNER, A. 2011. The extracellular matrix molecule tenascin C modulates expression levels and territories of key patterning genes during spinal cord astrocyte specification. *Development,* 138**,** 5321-31.

KATO, N. & HARADA, F. 1985. New U1 RNA species found in Friend SFFV (spleen focus forming virus)-transformed mouse cells. *J Biol Chem,* 260**,** 7775-82.

KISS, T. 2004. Biogenesis of small nuclear RNPs. *J Cell Sci,* 117**,** 5949-51.

KLEINSCHMIDT, C., TOVAR, K. & HILLEN, W. 1991. Computer simulations and experimental studies of gel mobility patterns for weak and strong non-cooperative protein binding to two targets on the same DNA: application to binding of tet repressor variants to multiple and single tet operator sites. *Nucleic Acids Res,* 19**,** 1021-8.

KLEINSMITH, L. J. & PIERCE, G. B., JR. 1964. Multipotentiality of Single Embryonal Carcinoma Cells. *Cancer Res,* 24**,** 1544-51.

KLIMANSKAYA, I., CHUNG, Y., MEISNER, L., JOHNSON, J., WEST, M. D. & LANZA, R. 2005. Human embryonic stem cells derived without feeder cells. *Lancet,* 365**,** 1636-41.

KOIVISTO, H., HYVARINEN, M., STROMBERG, A. M., INZUNZA, J., MATILAINEN, E., MIKKOLA, M., HOVATTA, O. & TEERIJOKI, H. 2004. Cultures of human embryonic stem cells: serum replacement medium or serum-containing media and the effect of basic fibroblast growth factor. *Reprod Biomed Online,* 9**,** 330-7.

KRISTO, P., TSAI, M.-J. & O'MALLEY, B. W. 1984. Characterization of Three Chicken Pseudogenes for U1 RNA. *DNA,* 3**,** 281-286.

KUEH, J., RICHARDS, M., NG, S. W., CHAN, W. K. & BONGSO, A. 2006. The search for factors in human feeders that support the derivation and propagation of human embryonic stem cells: preliminary studies using transcriptome profiling by serial analysis of gene expression. *Fertil Steril,* 85**,** 1843-6.

KURODA, T., TADA, M., KUBOTA, H., KIMURA, H., HATANO, S. Y., SUEMORI, H., NAKATSUJI, N. & TADA, T. 2005. Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression. *Mol Cell Biol,* 25**,** 2475-85.

KWEK, K. Y., MURPHY, S., FURGER, A., THOMAS, B., O'GORMAN, W., KIMURA, H., PROUDFOOT, N. J. & AKOULITCHEV, A. 2002. U1 snRNA associates with TFIIH and regulates transcriptional initiation. *Nat Struct Biol,* 9**,** 800-5.

KYRIAKOPOULOU, C., LARSSON, P., LIU, L., SCHUSTER, J., SODERBOM, F., KIRSEBOM, L. A. & VIRTANEN, A. 2006. U1-like snRNAs lacking complementarity to canonical 5' splice sites. *RNA,* 12**,** 1603-11.

LAPHAM, J., YU, Y. T., SHU, M. D., STEITZ, J. A. & CROTHERS, D. M. 1997. The position of site-directed cleavage of RNA using RNase H and 2'-O-methyl oligonucleotides is dependent on the enzyme source. *RNA,* 3**,** 950-1.

LERNER, M. R. & STEITZ, J. A. 1979. Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. *Proc Natl Acad Sci U S A,* 76**,** 5495-9.

LESCURE, A., MURGO, S., CARBON, P. & KROL, A. 1992. The proximal promoter and the start site cooperate to specify correct U1 snRNA transcription initiation by RNA polymerase II. *Nucleic Acids Res,* 20**,** 1573-8.

LI, J., WANG, G., WANG, C., ZHAO, Y., ZHANG, H., TAN, Z., SONG, Z., DING, M. & DENG, H. 2007. MEK/ERK signaling contributes to the maintenance of human embryonic stem cell self-renewal. *Differentiation,* 75**,** 299-307.

LI, S. S., LIU, Y. H., TSENG, C. N. & SINGH, S. 2006. Analysis of gene expression in single human oocytes and preimplantation embryos. *Biochem Biophys Res Commun,* 340**,** 48-53.

LI, Y., POWELL, S., BRUNETTE, E., LEBKOWSKI, J. & MANDALAM, R. 2005. Expansion of human embryonic stem cells in defined serum-free medium devoid of animal-derived products. *Biotechnol Bioeng,* 91**,** 688-98.

LIM, J. W. & BODNAR, A. 2002. Proteome analysis of conditioned medium from mouse embryonic fibroblast feeder layers which support the growth of human embryonic stem cells. *Proteomics,* 2**,** 1187-203.

LIN, X., YANG, J., CHEN, J., GUNASEKERA, A., FESIK, S. W. & SHEN, Y. 2004. Development of a tightly regulated U6 promoter for shRNA expression. *FEBS Lett,* 577**,** 376-80.

LINDGREN, V., BERNSTEIN, L. B., WEINER, A. M. & FRANCKE, U. 1985. Human U1 small nuclear RNA pseudogenes do not map to the site of the U1 genes in 1p36 but are clustered in 1q12-q22. *Molecular and cellular biology,* 5**,** 2172-80.

LIU, Q., FISCHER, U., WANG, F. & DREYFUSS, G. 1997. The spinal muscular atrophy disease gene product, SMN, and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins. *Cell,* 90**,** 1013-21.

LIU, Y., SONG, Z., ZHAO, Y., QIN, H., CAI, J., ZHANG, H., YU, T., JIANG, S., WANG, G., DING, M. & DENG, H. 2006. A novel chemical-defined medium with bFGF and N2B27 supplements supports undifferentiated growth in human embryonic stem cells. *Biochemical and Biophysical Research Communications,* 346**,** 131-139.

LO, P. C. & MOUNT, S. M. 1990. Drosophila melanogaster genes for U1 snRNA variants and their expression during development. *Nucleic Acids Res,* 18**,** 6971-9.

LOBO, S. M., MARZLUFF, W. F., SEUFERT, A. C., DEAN, W. L., SCHULTZ, G. A., SIMERLY, C. & SCHATTEN, G. 1988. Localization and expression of U1 RNA in early mouse embryo development. *Dev Biol,* 127**,** 349-61.

LUBINI, P., ZURCHER, W. & EGLI, M. 1994. Stabilizing effects of the RNA 2'-substituent: crystal structure of an oligodeoxynucleotide duplex containing 2'-O-methylated adenosines. *Chem Biol,* 1**,** 39-45.

LUDWIG, T. E., BERGENDAHL, V., LEVENSTEIN, M. E., YU, J., PROBASCO, M. D. & THOMSON, J. A. 2006a. Feeder-independent culture of human embryonic stem cells. *Nat Methods,* 3**,** 637-46.

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. & THOMSON, J. A. 2006b. Derivation of human embryonic stem cells in defined conditions. *Nat Biotechnol,* 24**,** 185-7.

LUND, E. 1988. Heterogeneity of human U1 snRNAs. *Nucleic Acids Res,* 16**,** 5813-26.

LUND, E., BOSTOCK, C., ROBERTSON, M., CHRISTIE, S., MITCHEN, J. L. & DAHLBERG, J. E. 1983. U1 small nuclear RNA genes are located on human chromosome 1 and are expressed in mouse-human hybrid cells. *Mol Cell Biol,* 3**,** 2211-20.

LUND, E., BOSTOCK, C. J. & DAHLBERG, J. E. 1987. The transcription of Xenopus laevis embryonic U1 snRNA genes changes when oocytes mature into eggs. *Genes Dev,* 1**,** 47-56.

LUND, E. & DAHLBERG, J. E. 1987. Differential accumulation of U1 and U4 small nuclear RNAs during Xenopus development. *Genes Dev,* 1**,** 39-46.

LUND, E., DAHLBERG, J. E. & FORBES, D. J. 1984. The two embryonic U1 small nuclear RNAs of Xenopus laevis are encoded by a major family of tandemly repeated genes. *Mol Cell Biol,* 4**,** 2580-6.

LUND, E., KAHAN, B. & DAHLBERG, J. E. 1985. Differential control of U1 small nuclear RNA expression during mouse development. *Science,* 229**,** 1271-4.

MA, Y., JIN, J., DONG, C., CHENG, E. C., LIN, H., HUANG, Y. & QIU, C. 2010. High-efficiency siRNA-based gene knockdown in human embryonic stem cells. *RNA,* 16**,** 2564-9.

MARTIN, G. R. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A,* 78**,** 7634-8.

MASSAM-WU, T., CHIU, M., CHOUDHURY, R., CHAUDHRY, S. S., BALDWIN, A. K., MCGOVERN, A., BALDOCK, C., SHUTTLEWORTH, C. A. & KIELTY, C. M. 2010. Assembly of fibrillin microfibrils governs extracellular deposition of latent TGF beta. *J Cell Sci,* 123**,** 3006-18.

MATIN, M. M., WALSH, J. R., GOKHALE, P. J., DRAPER, J. S., BAHRAMI, A. R., MORTON, I., MOORE, H. D. & 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**,** 659-68.

MATSUDA, T., NAKAMURA, T., NAKAO, K., ARAI, T., KATSUKI, M., HEIKE, T. & YOKOTA, T. 1999. STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. *EMBO J,* 18**,** 4261-9.

MATSUI, Y., ZSEBO, K. & HOGAN, B. L. 1992. Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. *Cell,* 70**,** 841-7.

MATTAJ, I. W. 1986. Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding. *Cell,* 46**,** 905-11.

MEDINA, R. J., ARCHER, D. B. & STITT, A. W. 2011. Eyes open to stem cells: safety trial may pave the way for cell therapy to treat retinal disease in patients. *Stem Cell Res Ther,* 2**,** 47.

MEHTA, A., MATHEW, S., VISWANATHAN, C. & SEN MAJUMDAR, A. 2010. Intrinsic properties and external factors determine the differentiation bias of human embryonic stem cell lines. *Cell Biol Int,* 34**,** 1021-31.

MENG, Y., ESHGHI, S., LI, Y. J., SCHMIDT, R., SCHAFFER, D. V. & HEALY, K. E. 2010. Characterization of integrin engagement during defined human embryonic stem cell culture. *FASEB J,* 24**,** 1056-65.

MITSUI, K., TOKUZAWA, Y., ITOH, H., SEGAWA, K., MURAKAMI, M., TAKAHASHI, K., MARUYAMA, M., MAEDA, M. & YAMANAKA, S. 2003. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell,* 113**,** 631-42.

MIURA, T., LUO, Y., KHREBTUKOVA, I., BRANDENBERGER, R., ZHOU, D., THIES, R. S., VASICEK, T., YOUNG, H., LEBKOWSKI, J., CARPENTER, M. K. & RAO, M. S. 2004. Monitoring early differentiation events in human embryonic stem cells by massively parallel signature sequencing and expressed sequence tag scan. *Stem Cells Dev,* 13**,** 694-715.

MIYAZAKI, T., FUTAKI, S., HASEGAWA, K., KAWASAKI, M., SANZEN, N., HAYASHI, M., KAWASE, E., SEKIGUCHI, K., NAKATSUJI, N. & SUEMORI, H. 2008. Recombinant human laminin isoforms can support the undifferentiated growth of human embryonic stem cells. *Biochem Biophys Res Commun,* 375**,** 27-32.

MONSTEIN, H. J., HAMMARSTROM, K., WESTIN, G., ZABIELSKI, J., PHILIPSON, L. & PETTERSSON, U. 1983. Loci for human U1 RNA: structural and evolutionary implications. *Journal of molecular biology,* 167**,** 245-57.

MONSTEIN, H. J., WESTIN, G., PHILIPSON, L. & PETTERSSON, U. 1982. A candidate gene for human U1 RNA. *EMBO J,* 1**,** 133-7.

MOREADITH, R. W. & GRAVES, K. H. 1992. Derivation of pluripotential embryonic stem cells from the rabbit. *Trans Assoc Am Physicians,* 105**,** 197-203.

NAMEKAWA, S. H. & LEE, J. T. 2011. Detection of nascent RNA, single-copy DNA and protein localization by immunoFISH in mouse germ cells and preimplantation embryos. *Nat Protoc,* 6**,** 270-84.

NAYLOR, S. L., ZABEL, B. U., MANSER, T., GESTELAND, R. & SAKAGUCHI, A. Y. 1984. Localization of human U1 small nuclear RNA genes to band p36.3 of chromosome 1 by in situ hybridization. *Somat Cell Mol Genet,* 10**,** 307-13.

NESIC, D., TANACKOVIC, G. & KRAMER, A. 2004. A role for Cajal bodies in the final steps of U2 snRNP biogenesis. *J Cell Sci,* 117**,** 4423-33.

NG, S. Y. & STANTON, L. W. 2013. Long non-coding RNAs in stem cell pluripotency. *Wiley Interdiscip Rev RNA,* 4**,** 121-8.

NICHOLS, J., EVANS, E. P. & SMITH, A. G. 1990. Establishment of germ-line-competent embryonic stem (ES) cells using differentiation inhibiting activity. *Development,* 110**,** 1341-8.

NICHOLS, J., ZEVNIK, B., ANASTASSIADIS, K., NIWA, H., KLEWE-NEBENIUS, D., CHAMBERS, I., SCHOLER, H. & SMITH, A. 1998. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell,* 95**,** 379-91.

NIWA, H., BURDON, T., CHAMBERS, I. & SMITH, A. 1998. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev,* 12**,** 2048-60.

NIWA, H., MIYAZAKI, J. & SMITH, A. G. 2000. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet,* 24**,** 372-6.

NOJIMA, H. & KORNBERG, R. D. 1983. Genes and pseudogenes for mouse U1 and U2 small nuclear RNAs. *Journal of Biological Chemistry,* 258**,** 8151-8155.

O'REILLY, D., DIENSTBIER, M., COWLEY, S. A., VAZQUEZ, P., DROZDZ, M., TAYLOR, S., JAMES, W. S. & MURPHY, S. 2012. Differentially expressed, variant U1 snRNAs regulate gene expression in human cells. *Genome Res*.

OHKAWA, J. & TAIRA, K. 2000. Control of the functional activity of an antisense RNA by a tetracycline-responsive derivative of the human U6 snRNA promoter. *Hum Gene Ther,* 11**,** 577-85.

OKAMOTO, K., OKAZAWA, H., OKUDA, A., SAKAI, M., MURAMATSU, M. & HAMADA, H. 1990. A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. *Cell,* 60**,** 461-72.

OKUMURA-NAKANISHI, S., SAITO, M., NIWA, H. & ISHIKAWA, F. 2005. Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. *J Biol Chem,* 280**,** 5307-17.

PAL, R., TOTEY, S., MAMIDI, M. K., BHAT, V. S. & TOTEY, S. 2009. Propensity of human embryonic stem cell lines during early stage of lineage specification controls their terminal differentiation into mature cell types. *Exp Biol Med (Maywood),* 234**,** 1230-43.

PALMIERI, S. L., PETER, W., HESS, H. & SCHOLER, H. R. 1994. Oct-4 transcription factor is differentially expressed in the mouse embryo during establishment of the first two extraembryonic cell lineages involved in implantation. *Dev Biol,* 166**,** 259-67.

PASHAI, N., HAO, H., ALL, A., GUPTA, S., CHAERKADY, R., DE LOS ANGELES, A., GEARHART, J. D. & KERR, C. L. 2012. Genome-wide profiling of pluripotent cells reveals a unique molecular signature of human embryonic germ cells. *PLoS One,* 7**,** e39088.

PIEDRAHITA, J. A., ANDERSON, G. B. & BONDURANT, R. H. 1990. On the isolation of embryonic stem cells: Comparative behavior of murine, porcine and ovine embryos. *Theriogenology,* 34**,** 879-901.

PIERCE, B., VERNEY, E. L. & DIXON, F. J. 1957. The biology of testicular cancer. I. Behavior after transplantation. *Cancer Res,* 17**,** 134-8.

POTTEN, C. S., OWEN, G. & BOOTH, D. 2002. Intestinal stem cells protect their genome by selective segregation of template DNA strands. *J Cell Sci,* 115**,** 2381-8.

PRINDULL, G., PRINDULL, B. & MEULEN, N. 1978. Haematopoietic stem cells (CFUc) in human cord blood. *Acta Paediatr Scand,* 67**,** 413-6.

PRITSKER, M., DONIGER, T. T., KRAMER, L. C., WESTCOT, S. E. & LEMISCHKA, I. R. 2005. Diversification of stem cell molecular repertoire by alternative splicing. *Proc Natl Acad Sci U S A,* 102**,** 14290-5.

PROWSE, A. B., MCQUADE, L. R., BRYANT, K. J., MARCAL, H. & GRAY, P. P. 2007. Identification of potential pluripotency determinants for human embryonic stem cells following proteomic analysis of human and mouse fibroblast conditioned media. *J Proteome Res,* 6**,** 3796-807.

PROWSE, A. B., MCQUADE, L. R., BRYANT, K. J., VAN DYK, D. D., TUCH, B. E. & GRAY, P. P. 2005. A proteome analysis of conditioned media from human neonatal fibroblasts used in the maintenance of human embryonic stem cells. *Proteomics,* 5**,** 978-89.

QI, X., LI, T. G., HAO, J., HU, J., WANG, J., SIMMONS, H., MIURA, S., MISHINA, Y. & ZHAO, G. Q. 2004. BMP4 supports self-renewal of embryonic stem cells by inhibiting mitogen-activated protein kinase pathways. *Proc Natl Acad Sci U S A,* 101**,** 6027-32.

RAJALA, K., HAKALA, H., PANULA, S., AIVIO, S., PIHLAJAMAKI, H., SUURONEN, R., HOVATTA, O. & SKOTTMAN, H. 2007. Testing of nine different xeno-free culture media for human embryonic stem cell cultures. *Hum Reprod,* 22**,** 1231-8.

RAPONI, M. & BARALLE, D. 2008. Can donor splice site recognition occur without the involvement of U1 snRNP? *Biochem Soc Trans,* 36**,** 548-50.

RAPONI, M., BURATTI, E., DASSIE, E., UPADHYAYA, M. & BARALLE, D. 2009. Low U1 snRNP dependence at the NF1 exon 29 donor splice site. *FEBS J,* 276**,** 2060-73.

RAY, R., RAY, K. & PANDA, C. K. 1997. Differential alterations in metabolic pattern of the six major UsnRNAs during development. *Molecular and Cellular Biochemistry,* 177**,** 79-88.

RESNICK, J. L., BIXLER, L. S., CHENG, L. & DONOVAN, P. J. 1992. Long-term proliferation of mouse primordial germ cells in culture. *Nature,* 359**,** 550-1.

REVIL, T., GAFFNEY, D., DIAS, C., MAJEWSKI, J. & JEROME-MAJEWSKA, L. A. 2010. Alternative splicing is frequent during early embryonic development in mouse. *BMC Genomics,* 11**,** 399.

REYNOLDS, B. A. & WEISS, S. 1992. Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. *Science,* 255**,** 1707-10.

RICHARDS, M., TAN, S., FONG, C. Y., BISWAS, A., CHAN, W. K. & BONGSO, A. 2003. Comparative evaluation of various human feeders for prolonged undifferentiated growth of human embryonic stem cells. *Stem Cells,* 21**,** 546-56.

RIDEOUT, W. M., 3RD, HOCHEDLINGER, K., KYBA, M., DALEY, G. Q. & JAENISCH, R. 2002. Correction of a genetic defect by nuclear transplantation and combined cell and gene therapy. *Cell,* 109**,** 17-27.

ROBINTON, D. A. & DALEY, G. Q. 2012. The promise of induced pluripotent stem cells in research and therapy. *Nature,* 481**,** 295-305.

RODDA, D. J., CHEW, J. L., LIM, L. H., LOH, Y. H., WANG, B., NG, H. H. & ROBSON, P. 2005. Transcriptional regulation of nanog by OCT4 and SOX2. *J Biol Chem,* 280**,** 24731-7.

RODRIGUEZ, R. T., VELKEY, J. M., LUTZKO, C., SEERKE, R., KOHN, D. B., O'SHEA, K. S. & FIRPO, M. T. 2007. Manipulation of OCT4 levels in human embryonic stem cells results in induction of differential cell types. *Exp Biol Med (Maywood),* 232**,** 1368-80.

ROSNER, M. H., VIGANO, M. A., OZATO, K., TIMMONS, P. M., POIRIER, F., RIGBY, P. W. & STAUDT, L. M. 1990. A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. *Nature,* 345**,** 686-92.

RUSKIN, B., ZAMORE, P. D. & GREEN, M. R. 1988. A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell,* 52**,** 207-19.

SALOMONIS, N., SCHLIEVE, C. R., PEREIRA, L., WAHLQUIST, C., COLAS, A., ZAMBON, A. C., VRANIZAN, K., SPINDLER, M. J., PICO, A. R., CLINE, M. S., CLARK, T. A., WILLIAMS, A., BLUME, J. E., SAMAL, E., MERCOLA, M., MERRILL, B. J. & CONKLIN, B. R. 2010. Alternative splicing regulates mouse embryonic stem cell pluripotency and differentiation. *Proc Natl Acad Sci U S A,* 107**,** 10514-9.

SANTIAGO, C. & MARZLUFF, W. F. 1989. Expression of the U1 RNA gene repeat during early sea urchin development: evidence for a switch in U1 RNA genes during development. *Proc Natl Acad Sci U S A,* 86**,** 2572-6.

SAWA, H., OHNO, M., SAKAMOTO, H. & SHIMURA, Y. 1988. Requirement of ATP in the second step of the pre-mRNA splicing reaction. *Nucleic Acids Res,* 16**,** 3157-64.

SAXENA, S., HANWATE, M., DEB, K., SHARMA, V. & TOTEY, S. 2008. FGF2 secreting human fibroblast feeder cells: a novel culture system for human embryonic stem cells. *Mol Reprod Dev,* 75**,** 1523-32.

SCHENKE-LAYLAND, K., ANGELIS, E., RHODES, K. E., HEYDARKHAN-HAGVALL, S., MIKKOLA, H. K. & MACLELLAN, W. R. 2007. Collagen IV induces trophoectoderm differentiation of mouse embryonic stem cells. *Stem Cells,* 25**,** 1529-38.

SCHWARTZ, S. D., HUBSCHMAN, J. P., HEILWELL, G., FRANCO-CARDENAS, V., PAN, C. K., OSTRICK, R. M., MICKUNAS, E., GAY, R., KLIMANSKAYA, I. & LANZA, R. 2012. Embryonic stem cell trials for macular degeneration: a preliminary report. *Lancet,* 379**,** 713-20.

SHAMBLOTT, M. J., AXELMAN, J., WANG, S., BUGG, E. M., LITTLEFIELD, J. W., DONOVAN, P. J., BLUMENTHAL, P. D., HUGGINS, G. R. & GEARHART, J. D. 1998. Derivation of pluripotent stem cells from cultured human primordial germ cells. *Proc Natl Acad Sci U S A,* 95**,** 13726-31.

SINGH, A. M., REYNOLDS, D., CLIFF, T., OHTSUKA, S., MATTHEYSES, A. L., SUN, Y., MENENDEZ, L., KULIK, M. & 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**,** 312-26.

SKOTTMAN, H., MIKKOLA, M., LUNDIN, K., OLSSON, C., STROMBERG, A. M., TUURI, T., OTONKOSKI, T., HOVATTA, O. & LAHESMAA, R. 2005. Gene expression signatures of seven individual human embryonic stem cell lines. *Stem Cells,* 23**,** 1343-56.

SMITH, A. G., HEATH, J. K., DONALDSON, D. D., WONG, G. G., MOREAU, J., STAHL, M. & ROGERS, D. 1988. Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature,* 336**,** 688-90.

SPILUTTINI, B., GU, B., BELAGAL, P., SMIRNOVA, A. S., NGUYEN, V. T., HEBERT, C., SCHMIDT, U., BERTRAND, E., DARZACQ, X. & BENSAUDE, O. 2010. Splicing-independent recruitment of U1 snRNP to a transcription unit in living cells. *J Cell Sci,* 123**,** 2085-93.

STEVENS, L. C. & LITTLE, C. C. 1954. Spontaneous Testicular Teratomas in an Inbred Strain of Mice. *Proc Natl Acad Sci U S A,* 40**,** 1080-7.

STIER, S., KO, Y., FORKERT, R., LUTZ, C., NEUHAUS, T., GRUNEWALD, E., CHENG, T., DOMBKOWSKI, D., CALVI, L. M., RITTLING, S. R. & SCADDEN, D. T. 2005. Osteopontin is a hematopoietic stem cell niche component that negatively regulates stem cell pool size. *J Exp Med,* 201**,** 1781-91.

STOJKOVIC, P., LAKO, M., STEWART, R., PRZYBORSKI, S., ARMSTRONG, L., EVANS, J., MURDOCH, A., STRACHAN, T. & STOJKOVIC, M. 2005. An autogeneic feeder cell system that efficiently supports growth of undifferentiated human embryonic stem cells. *Stem Cells,* 23**,** 306-14.

STROJEK, R. M., REED, M. A., HOOVER, J. L. & WAGNER, T. E. 1990. A method for cultivating morphologically undifferentiated embryonic stem cells from porcine blastocysts. *Theriogenology,* 33**,** 901-13.

SUKOYAN, M. A., VATOLIN, S. Y., GOLUBITSA, A. N., ZHELEZOVA, A. I., SEMENOVA, L. A. & SEROV, O. L. 1993. Embryonic stem cells derived from morulae, inner cell mass, and blastocysts of mink: comparisons of their pluripotencies. *Mol Reprod Dev,* 36**,** 148-58.

SUMI, T., TSUNEYOSHI, N., NAKATSUJI, N. & SUEMORI, H. 2008. Defining early lineage specification of human embryonic stem cells by the orchestrated balance of canonical Wnt/beta-catenin, Activin/Nodal and BMP signaling. *Development,* 135**,** 2969-79.

TAIPALE, J. & KESKI-OJA, J. 1997. Growth factors in the extracellular matrix. *FASEB J,* 11**,** 51-9.

TAKAHASHI, K., TANABE, K., OHNUKI, M., NARITA, M., ICHISAKA, T., TOMODA, K. & YAMANAKA, S. 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell,* 131**,** 861-72.

TAKAHASHI, K. & YAMANAKA, S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell,* 126**,** 663-76.

TARN, W. Y. & STEITZ, J. A. 1994. SR proteins can compensate for the loss of U1 snRNP functions in vitro. *Genes Dev,* 8**,** 2704-17.

TARN, W. Y. & STEITZ, J. A. 1996. A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron in vitro. *Cell,* 84**,** 801-11.

TAVAKOLI, T., XU, X., DERBY, E., SEREBRYAKOVA, Y., REID, Y., RAO, M. S., MATTSON, M. P. & MA, W. 2009. Self-renewal and differentiation capabilities are variable between human embryonic stem cell lines I3, I6 and BG01V. *BMC Cell Biol,* 10**,** 44.

TEO, A. K., ARNOLD, S. J., TROTTER, M. W., BROWN, S., ANG, L. T., CHNG, Z., ROBERTSON, E. J., DUNN, N. R. & VALLIER, L. 2011. Pluripotency factors regulate definitive endoderm specification through eomesodermin. *Genes Dev,* 25**,** 238-50.

THOMSON, J. A., ITSKOVITZ-ELDOR, J., SHAPIRO, S. S., WAKNITZ, M. A., SWIERGIEL, J. J., MARSHALL, V. S. & JONES, J. M. 1998. Embryonic stem cell lines derived from human blastocysts. *Science,* 282**,** 1145-7.

THOMSON, J. A., KALISHMAN, J., GOLOS, T. G., DURNING, M., HARRIS, C. P., BECKER, R. A. & HEARN, J. P. 1995. Isolation of a primate embryonic stem cell line. *Proc Natl Acad Sci U S A,* 92**,** 7844-8.

UGUEN, P. & MURPHY, S. 2004. 3'-box-dependent processing of human pre-U1 snRNA requires a combination of RNA and protein co-factors. *Nucleic Acids Res,* 32**,** 2987-94.

VALADKHAN, S. & MANLEY, J. L. 2001. Splicing-related catalysis by protein-free snRNAs. *Nature,* 413**,** 701-7.

VALLIER, L., ALEXANDER, M. & PEDERSEN, R. A. 2005. Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J Cell Sci,* 118**,** 4495-509.

VALLIER, L., MENDJAN, S., BROWN, S., CHNG, Z., TEO, A., SMITHERS, L. E., TROTTER, M. W., CHO, C. H., MARTINEZ, A., RUGG-GUNN, P., BRONS, G. & PEDERSEN, R. A. 2009. Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. *Development,* 136**,** 1339-49.

WALKER, M. R., PATEL, K. K. & STAPPENBECK, T. S. 2009. The stem cell niche. *J Pathol,* 217**,** 169-80.

WANG, E. T., SANDBERG, R., LUO, S., KHREBTUKOVA, I., ZHANG, L., MAYR, C., KINGSMORE, S. F., SCHROTH, G. P. & BURGE, C. B. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature,* 456**,** 470-6.

WANG, P., BRANCH, D. R., BALI, M., SCHULTZ, G. A., GOSS, P. E. & JIN, T. 2003. The POU homeodomain protein OCT3 as a potential transcriptional activator for fibroblast growth factor-4 (FGF-4) in human breast cancer cells. *Biochem J,* 375**,** 199-205.

WANG, Z., ORON, E., NELSON, B., RAZIS, S. & IVANOVA, N. 2012. Distinct lineage specification roles for NANOG, OCT4, and SOX2 in human embryonic stem cells. *Cell Stem Cell,* 10**,** 440-54.

WATANABE-NAGASU, N., ITOH, Y., TANI, T., OKANO, K., KOGA, N., OKADA, N. & OHSHIMA, Y. 1983. Structural analysis of gene loci for rat U1 small nuclear RNA. *Nucleic Acids Res,* 11**,** 1791-801.

WILLIAMS, R. L., HILTON, D. J., PEASE, S., WILLSON, T. A., STEWART, C. L., GEARING, D. P., WAGNER, E. F., METCALF, D., NICOLA, N. A. & GOUGH, N. M. 1988. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature,* 336**,** 684-7.

WILMUT, I., SCHNIEKE, A. E., MCWHIR, J., KIND, A. J. & CAMPBELL, K. H. 1997. Viable offspring derived from fetal and adult mammalian cells. *Nature,* 385**,** 810-3.

WOOD, H. B. & EPISKOPOU, V. 1999. Comparative expression of the mouse Sox1, Sox2 and Sox3 genes from pre-gastrulation to early somite stages. *Mech Dev,* 86**,** 197-201.

WU, J. Q., HABEGGER, L., NOISA, P., SZEKELY, A., QIU, C., HUTCHISON, S., RAHA, D., EGHOLM, M., LIN, H., WEISSMAN, S., CUI, W., GERSTEIN, M. & SNYDER, M. 2010. Dynamic transcriptomes during neural differentiation of human embryonic stem cells revealed by short, long, and paired-end sequencing. *Proc Natl Acad Sci U S A,* 107**,** 5254-9.

XIAO, L., YUAN, X. & SHARKIS, S. J. 2006. Activin A maintains self-renewal and regulates fibroblast growth factor, Wnt, and bone morphogenic protein pathways in human embryonic stem cells. *Stem Cells,* 24**,** 1476-86.

XIAO, Q., ZENG, L., ZHANG, Z., HU, Y. & XU, Q. 2007. Stem cell-derived Sca-1+ progenitors differentiate into smooth muscle cells, which is mediated by collagen IV-integrin alpha1/beta1/alphav and PDGF receptor pathways. *Am J Physiol Cell Physiol,* 292**,** C342-52.

XU, C., INOKUMA, M. S., DENHAM, J., GOLDS, K., KUNDU, P., GOLD, J. D. & CARPENTER, M. K. 2001. Feeder-free growth of undifferentiated human embryonic stem cells. *Nat Biotechnol,* 19**,** 971-4.

XU, N., PAPAGIANNAKOPOULOS, T., PAN, G., THOMSON, J. A. & KOSIK, K. S. 2009a. MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. *Cell,* 137**,** 647-58.

XU, R. H., CHEN, X., LI, D. S., LI, R., ADDICKS, G. C., GLENNON, C., ZWAKA, T. P. & THOMSON, J. A. 2002. BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nat Biotechnol,* 20**,** 1261-4.

XU, R. H., PECK, R. M., LI, D. S., FENG, X., LUDWIG, T. & THOMSON, J. A. 2005. Basic FGF and suppression of BMP signaling sustain undifferentiated proliferation of human ES cells. *Nat Methods,* 2**,** 185-90.

XU, R. H., SAMPSELL-BARRON, T. L., GU, F., ROOT, S., PECK, R. M., PAN, G., YU, J., ANTOSIEWICZ-BOURGET, J., TIAN, S., STEWART, R. & THOMSON, J. A. 2008. NANOG is a direct target of TGFbeta/activin-mediated SMAD signaling in human ESCs. *Cell Stem Cell,* 3**,** 196-206.

XU, X. Q., SOO, S. Y., SUN, W. & ZWEIGERDT, R. 2009b. Global expression profile of highly enriched cardiomyocytes derived from human embryonic stem cells. *Stem Cells,* 27**,** 2163-74.

YEO, G. W., XU, X., LIANG, T. Y., MUOTRI, A. R., CARSON, C. T., COUFAL, N. G. & GAGE, F. H. 2007. Alternative splicing events identified in human embryonic stem cells and neural progenitors. *PLoS Comput Biol,* 3**,** 1951-67.

YEOM, Y. I., FUHRMANN, G., OVITT, C. E., BREHM, A., OHBO, K., GROSS, M., HUBNER, K. & SCHOLER, H. R. 1996. Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development,* 122**,** 881-94.

YING, Q. L., NICHOLS, J., CHAMBERS, I. & SMITH, A. 2003. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell,* 115**,** 281-92.

YING, Q. L., WRAY, J., NICHOLS, J., BATLLE-MORERA, L., DOBLE, B., WOODGETT, J., COHEN, P. & SMITH, A. 2008. The ground state of embryonic stem cell self-renewal. *Nature,* 453**,** 519-23.

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, II & THOMSON, J. A. 2007. Induced pluripotent stem cell lines derived from human somatic cells. *Science,* 318**,** 1917-20.

YU, P., PAN, G., YU, J. & THOMSON, J. A. 2011. FGF2 sustains NANOG and switches the outcome of BMP4-induced human embryonic stem cell differentiation. *Cell Stem Cell,* 8**,** 326-34.

ZABOROWSKA, J., TAYLOR, A. & MURPHY, S. 2012. A novel TBP-TAF complex on RNA polymerase II-transcribed snRNA genes. *Transcription,* 3**,** 92-104.

ZHOU, H., HUANG, C. & XIA, X. G. 2008. A tightly regulated Pol III promoter for synthesis of miRNA genes in tandem. *Biochim Biophys Acta,* 1779**,** 773-9.

ZOU, C., LUO, Q., QIN, J., SHI, Y., YANG, L., JU, B. & SONG, G. 2012. Osteopontin Promotes Mesenchymal Stem Cell Migration and Lessens Cell Stiffness via Integrin beta1, FAK, and ERK Pathways. *Cell Biochem Biophys*.

ZUK, P. A., ZHU, M., MIZUNO, H., HUANG, J., FUTRELL, J. W., KATZ, A. J., BENHAIM, P., LORENZ, H. P. & HEDRICK, M. H. 2001. Multilineage cells from human adipose tissue: implications for cell-based therapies. *Tissue Eng,* 7**,** 211-28.