TGFβ SUPERFAMILY SIGNALLING IN PLURIPOTENCY AND DIFFERENTIATION OF HUMAN EMBRYONIC STEM CELLS

by

VARVARA CHALMANTZI

A thesis submitted to
The University of Birmingham
For the degree of
DOCTOR OF PHILOSOPHY

School of Biosciences
College of Life and Environmental Sciences
University of Birmingham
January 2019

To my daughter Iouliani for teaching me that human beings can learn, evolve and thrive in every breath they take.

Summary

The discovery of human embryonic stem cells (hESCs) 20 years ago (Thomson et al., 1998), has significantly benefited the field of stem cell biology (Ilic & Ogilvie, 2017). Despite ethical (and legal in some cases) obstacles, therapies utilising hESCs and their derivatives are attempting to tackle several diseases (Ilic & Ogilvie, 2017).

The TGFβ superfamily is known to be implicated in crucial decisions from early development of the embryo to the adult life (Mullen & Wrana, 2017). ACTIVIN A and BMP4 are two members of this superfamily and their activation status has an important impact on hESCs (James et al., 2005). On the one hand, ACTIVIN A is associated with the maintenance of the undifferentiated state (James et al., 2005), (Beattie et al., 2005), while on the other BMP4 is considered a differentiation stimulus (R. Xu et al., 2002), (P. Yu, Pan, Yu, & Thomson, 2011), (Richter et al., 2014).

The aim of this Thesis is to study in depth several key cellular responses of hESCs to BMP4 and investigate the role of SARA, a TGFβ signalling mediator (Tsukazaki et al., 1998) in pluripotency and differentiation. In particular, Chapter 3 establishes the culture system of H1 cells and shows the effect of BMP4 on the expression of several genes associated with differentiation. Moreover, it introduces a new culture system based on a modified version of a commercial medium which is later used for phosphoproteomic studies in Chapter 4.

Chapter 4 describes the experimental procedure followed in order to metabolically label H1 cells and use Mass Spectrometry (MS) to elucidate changes in protein phosphorylation status upon 60min of BMP4 induction. The regulated phosphosites were further categorised depending on three parameters (biological process, molecular function and cellular component). Some of the upregulated phosphosites

were selected and further tested as putative BMP4 targets using phos-tag acrylamide gels.

Chapter 5 focuses on SARA; it investigates the presence of one or more isoforms in human cell lines and the possibility that SARA shuttles to the nucleus. Moreover, it studies the expression levels of SARA, both at the mRNA and at the protein level in differentiating H1 cells and the effect it has in SMAD2 phosphorylation. Additionally, it compares the protein levels between fibroblasts and human induced pluripotent stem cells (hiPSCs) and addresses the notion that SARA is degraded by the lysosome. Last, this chapter attempts to generate GFP-SARA stable H1 cell lines, by transfection, and SARA knockout H1 cell lines, using a genome editing technique (CRISPR).

Acknowledgments

«All men by nature desire knowledge» (Aristotle, Metaphysics) and this powerful quote is the best way I found in order to summarise the reasons behind my decision to start a PhD. This endeavour would have never started if it wasn't for my husband Dr. Anastatios Batilas who encouraged me to pursue it. I would like to thank Prof. Theodore Fotsis and Dr. Carol Murphy for giving me the space and time to adapt in an, unknown for me, laboratory environment a couple of years ago and express my scientific curiosity under their supervision. I also want to thank Prof. John Heath for helping me throughout my studies in the University of Birmingham and for his willingness to assist me and supervise me during my last, and most crucial, year of studies.

«When setting out upon your way to Ithaca, wish always that your course be long, full of adventure, full of lore» (Konstantine P. Kavafy, Ithaca) and indeed this PhD has been a long adventurous journey. I would like to thank all the friends and people with whom I have worked all these years, for teaching me new skills and making this journey less daunting. I also want to express my gratitude to my family for standing next to me and supporting me morally when I was facing several adversities.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ACD	Asymmetric Cell Division
ACN	Acetonitrile
AFP	Alpha fetoprotein
ALP	Alkaline Phosphatase
AMD	Age Related Macular Degeneration
APC	Adenomatous Polyposis Coli
ARHGAP1	Rho-GTP-ase activating protein-1
ARP2/3	Actin Related Protein 2/3
BCL9	B Cell CLL/Lymphoma 9
BMP2K	BMP2 Inducible Kinase
BMP4	Bone Morphogenetic Protein 4
BSA	Bovine Serum Albumin
CDX2	Caudal Type Homeobox 2
CM	Conditioned Medium
CME	Clathrin Mediated Endocytosis
CQ	Chloroquine
CRISPR	Clustered, Regularly Interspaced, short Palindromic Repeats/ Crispr associated
CRM1	Chromosomal Maintenance 1
DDR	Homology Directed Repair
DSB	Double Strand Break
DSH	Dishevelld
ECM	Extracellular Matrix
EE	Early Endosome
EEA1	Early Endosomal Antigen-1
EF1a	Elongation Factor-1a
EGA	Embryonic Genome Activation
EMT	Epithelial to Mesenchymal Transition
EPI	Epiblast
EpiSCs	Epiblast Stem Cells
ERC	Endocytic Recycling Compartment
ES	Embryonic Stem

ESCRT	Endosomal Sorting Complexes Required for Transport
FBS	Fetal Bovine Serum
FGF2	Fibroblast Growth Factor 2
FSC	Fetal Calf Serum
FZ	Frizzled
GABA	γ-aminobuteric acid
GCD	Gentle Cell Dissociation Reagent
GSC	Goosecoid
GSK3β	Glycogen Synthase Kinase 3β
hCG	Human Chorionic Gonadotropin
HDAC	Histone Deacetylase
hESCs	Human Embryonic Stem Cells
HRS	Hepatocyte Growth Factor Regulated Tyrosine Kinase Substrate
HSA	Human Serum Albumin
HSC	Hematopoetic Stem Cell
ICM	Inner Cell Mass
IGF-II	Insulin Growth Factor
IGF-IR	Insulin Like Growth Factor Receptor
iPSCs	induced Pluripotent Stem Cells
IR	Insulin Receptor
IRES	Internal Ribosome Entry Site
ITS	Insulin-Transferrin-Selenium
KD	Knock Down
KEA	Kinome Enrichment Analysis
KSR	Knock out Serum Replacement
LAP	Latency Associated Peptide
LE	Late Endosome
LEF/TCF	Lymphoid Enhancer-Binding Factor 1/T Cell Specific Transcription Factor
LiCl	Lithium Chloride
LMB	Leptomycin B
LMPs	Lysosome Membrane Proteins
LRP	Low-Density Lipoprotein-Receptor Related Protein
LTBP	Latency Associated Peptide Binding Protein

MBD2-NuRD	Methyl Binding Domain Protein-2 -Containing Nucleosome Remodeling And Deacetylation
MDLS	Multi-Dimensional Liquid Chromatography
ME	Mesendoderm
MEFs	Mouse Embryonic Fibroblasts
mESC	Mouse Embryonic Stem Cells
MET	Mesenchymal to Epithelial Transition
MINK1	Misshapen-Like Kinase-1
MMPs	Matrix Metalloproteinases
MS	Mass Spectrometry
MS	Mass Spectrometry
MSN	Misshapen
MTOC	Microtubule Organising Centre
MVB	Multivesicular Body
NE	Neuroectoderm
NEEAs	Non Essential Amino-Acids
NHEJ	Non Homologous End Joining
NLS	Nuclear Localisation Signal
NSP	Novel Serine Protease
OCT4	Octamer Binding Transcription Factor-4
P/S	Penicillin/Streptomycin
PAM	Protospacer adjacent motif
PE	Primitive Endoderm
PGCs	Primordial Germ Cells
PGK	Phosphoglycerate Kinase
PH	Phox Homology
PK	Prickle
PP1C	Protein Phosphatase-1c
Pr-crRNA	precursor CRISPR RNA
PS	Primitive Streak
PSC	Pluripotent Stem Cell
PTM	Post Translational Modification
RABs	Ras-Associated Binding proteins
RNAPII	RNA Polymerase II

RPE	Retinal Pigment Epithelia
SC	Stem Cell
SCNT	Somatic Cell Nuclear Transfer
SFP	Specific Pathogen Free
SILAC	Stable Isotope Labelling by Amino Acids in Cell Culture
SOP	Sensory Organ Precursor
SOX	Sry-related HMG box
SSEA-1	Stage Specific Antigen-1
SSEA-3	Stage Specific Antigen-3
TALEN	Transcription Activator-Like Effector Nucleases
TE	Trophectoderm
TFA	Trifluoroacetic acid
TFSII/TCEA	Transcription elongation Factor II S
TGFβ	Transforming Growth Factor β
TLE	Transducin Like Enhancer Protein
trcrRNA	transactivating crRNA
TSS	Transcription Start Site
UIM	Ubiquitin Interacting Motif
XCI	X Chromosome Inactivation
ZFN	Zinc Finger Nucleases

1. INTRODUCTION

1.1 Introduction in Stem Cells (SC)

1.1.1 Stem Cells

The notion that not all cells are equal and that there is a type of cell from which other types can originate goes back to the 19th century (Daley, 2015). Almost 70 years ago Till and McCulloch started shedding light in the properties of hematopoetic stem cells (HSCs) (MCCULLOCH & TILL, 1960),(TILL & McCULLOCH, 1961), thus preparing the background of cellular hierarchy (Daley, 2015). A stem cell is a cell that can not only propagate but can also give rise to a differentiated progeny (Daley, 2015).

Stem cells have a wide range of differentiation potentials and can be further classified according to the degree of restriction in fate acquisition (Daley, 2015). According to the definition although the zygote ultimately forms all the cell types in an organism and therefore has the widest potential in development, it cannot technically be regarded a stem cell because it lacks the self-renewal capacity (Daley, 2015). Pluripotent stem cells (PSCs) have the ability to differentiate into cell types of all three embryonic germ layers (Dulak et al., 2015). Multipotent stem cells, usually known as adult or somatic stem cells have a more restricted potential as they can give rise to the types of cells that are found in the tissue in which they reside (Dulak et al., 2015). Unipotent stem cells are the most restricted ones as they can only give rise to one specialised cell type, for example epidermal stem cells can only differentiate towards keratinocytes (Dulak et al., 2015).

1.1.2 Early Embryogenesis

The question of how a whole organism comes into existence has puzzled philosophers for a very long time before scientific observations could clarify this process. Aristotle, 2,000 years ago proposed arguments that later on gave rise to two opposing theories (Goy, 2018). On the one hand, he suggested that the embryo is successively developed and therefore he should be considered mostly an advocate of epigenesis (Goy, 2018). However, he also believed that development relies on sperm parts which gradually become larger and this is in alignment with the idea of preformation (Goy, 2018). Today, technological advances have enabled a sharp increase in scientific knowledge and have deeply transformed the scientific view of embryogenesis.

The beginning of embryonic development occurs when an egg and a sperm fuse, thus generating the zygote (Mitalipov & Wolf, 2009). In mammals, the developing embryo travels through the fallopian tube in order to reach the uterus for implantation (Sozen *et al*, 2014). The period between fertilization and implantation is known as pre-implantation **(Fig. 1-1)** and lasts around 7 days in humans (Sozen *et al*, 2014). Moreover, in humans, this period is highly critical since most of the embryos fail to complete their mission (Sozen *et al*, 2014). The major source of knowledge regarding human embryonic development comes from studying other species as well as using hESCs cultures (De Paepe et al., 2014). Therefore, comparisons are useful but “have to be treated as such”(De Paepe et al., 2014).

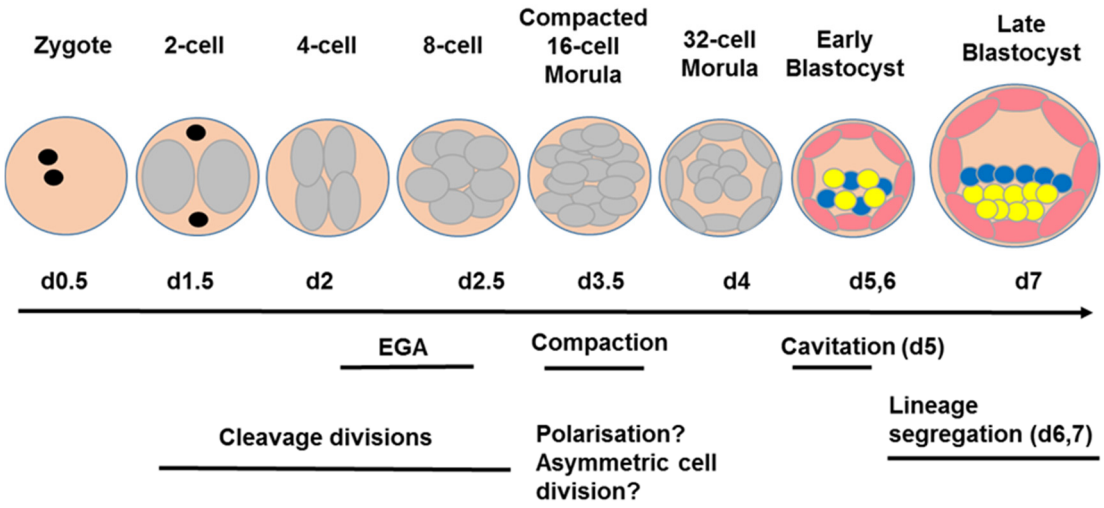


Figure 1-1 Human Preimplantation Development.

The figure follows the most important events in the first week of embryogenesis. The totipotent zygote through a series of divisions reaches the 8 cell stage where the genome becomes transcriptionally active. Compaction transforms the embryo into the morula and later on the blastocyst emerges through cavitation. By the 7th day two events of lineage specification occur thus generating the TE (pink cells) and the ICM (yellow and blue cells) with the latter further segregating into PE (blue cells) and EPI (yellow cells). Adapted from: (Sozen, et al, 2014).

The zygote, as well as the first 4 blastomers that derive from it, are considered to be totipotent, because they can create all embryonic and extraembryonic tissues (Mitalipov & Wolf, 2009). Once formed, the zygote undergoes three cleavage divisions that generate successively first 2, next 4 and last 8 cells (Niakan et al., 2012). These cells are progressively smaller and therefore the total size of the embryo remains the same (Niakan et al., 2012). Initially, the embryo is not transcriptionally active and relies on maternally inherited RNA pools (Dobson et al., 2004). However, by day 3 (4 to 8 cell stage) embryos initiate their own transcription programs (Dobson et al., 2004) and this process is known as Embryonic Genome Activation (EGA) (Niakan et al., 2012).

The next step is compaction and it is usually observed in human embryos at the 8-cell stage and onwards (Iwata et al., 2014). In fact, earlier compaction is associated with abnormalities in development (Iwata et al., 2014). This process begins when blastomers become closely attached thus resulting in flattening of the cell shape (Watson & Barcroft, 2001). The morphological alteration is so obvious that cell boundaries are no longer visible and the arising mass of cells is called Morula (Watson & Barcroft, 2001). In mice, compaction sets the onset for differentiation (Watson and Barcroft, 2001). By the end of it the epithelial cells that are placed on the outside of the embryo become the epithelial trophectoderm (TE) which actively drives the next process of cavitation (Watson & Barcroft, 2001).

The presence of a Na/K-ATPase in the basolateral membrane of these cells facilitates the concentration of Na^+, K^+, Cl^-, Ca^{2+} and Mg^{2+} in the cavity (Watson & Barcroft, 2001). As a result water flows into and remains within the blastocoel (Watson & Barcroft, 2001). Cavitation largely relies on proper formation of tight junctions and E-cadherin junctions (Watson & Barcroft, 2001). Ultimately this process shapes a structure which consists of TE on the outside and Inner Cell Mass (ICM) in the inside and is known as blastocyst (Niakan & Eggan, 2013).

In humans, blastocyst formation is observed between day 4 and 5 (Niakan & Eggan, 2013). Once the blastocyst reaches the uterus it remains there for up to 3 days and eventually is released from the zona pellucida and implants around day 7 (Norwitz, *et al*, 2001). The first lineage segregation in humans occurs on day 5 when cells become either the TE or the ICM (De Paepe et al., 2014). The TE will form extra-embryonic structures (placenta included), while the ICM will form the embryo

(Norwitz, *et al*, 2001). The second segregation occurs on day 7 when the ICM further diverges in epiblast (EPI) and Primitive Endoderm (PE) (De Paepe et al., 2014).

At the time of the implantation the mouse embryo is composed of the TE, the EPI or primitive ectoderm which lies underneath the TE and the PE which faces the blastocoel cavity (Tam et al., 1993). The embryo later becomes a cylindrical structure made up of two layers the EPI inside and the PE outside (Tam *et al*, 1993).

The next crucial step is gastrulation. Gastrulation is a complex process which begins with the formation of the primitive streak (PS) and is completed when the bilaminar embryo becomes trilaminar **(Fig. 1-2)** (Tam *et al*, 1993). The PS is a groove through which epiblast cells start migrating towards the PE (Tam, *et al*, 1993). Cells which replace the pre-existing PE become the definitive endoderm (Tam *et al*, 1993). The space occupied between the EPI and the PE becomes the definitive mesoderm and the cells that remain in the EPI become the definitive ectoderm (Tam *et al*, 1993). In the human embryo, epiblast cells start moving in the primitive streak at day 14 and they are usually called presumptive mesendoderm (ME) in order to define their ability to produce either the mesoderm or the endoderm (P. Yu et al., 2011).

Further along in development, the endoderm will give rise to the epithelia of the liver, pancreas, gastrointestinal and respiratory tract (P. P. L. Tam & Loebel, 2007). Mesoderm will give rise to bone, muscles, cartilage, blood cells and lymph vessels while the ectoderm will give rise to the epidermis and the nervous system **(Fig. 1-3)** (P. P. L. Tam & Loebel, 2007).

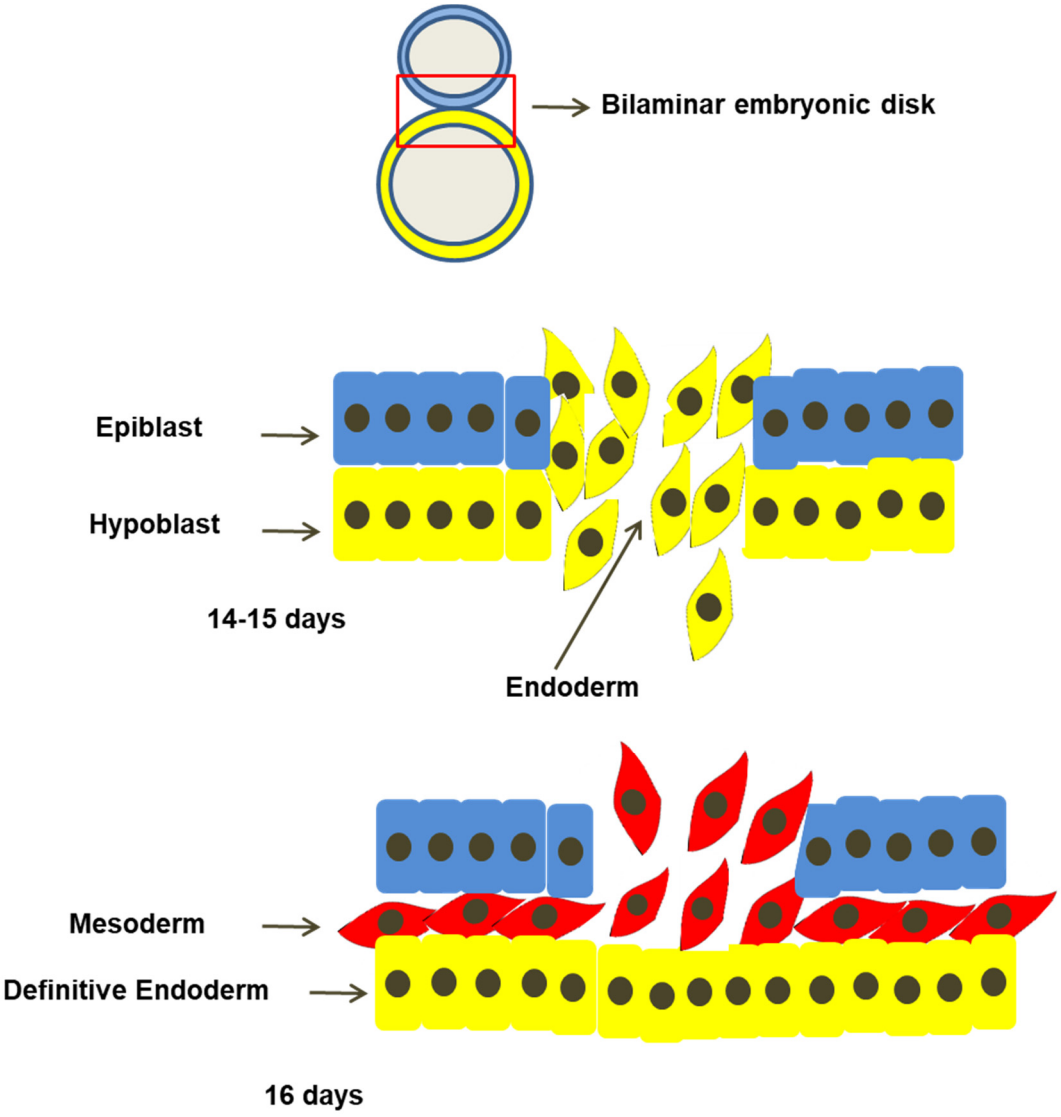


Figure 1-2 Germ layer formation.

The implanted blastocyst consists of two layers; the epiblast (blue cells) and the primitive endoderm or hypoblast (yellow cells). Gastrulation starts by the end of the second week of development. During this process, epiblast cells move through the primitive streak to replace the hypoblast and occupy the space between the epiblast and .hypoblast. The cells remaining on the epiblast become the definitive ectoderm (blue cells) while the cells replacing the hypoblast become the definitive endoderm (yellow cells). The cells occupying the space between the former epiblast and hypoblast become the definitive mesoderm (red cells). Adapted from: (William J. Larsen, Human Embryology, 3rd edition).

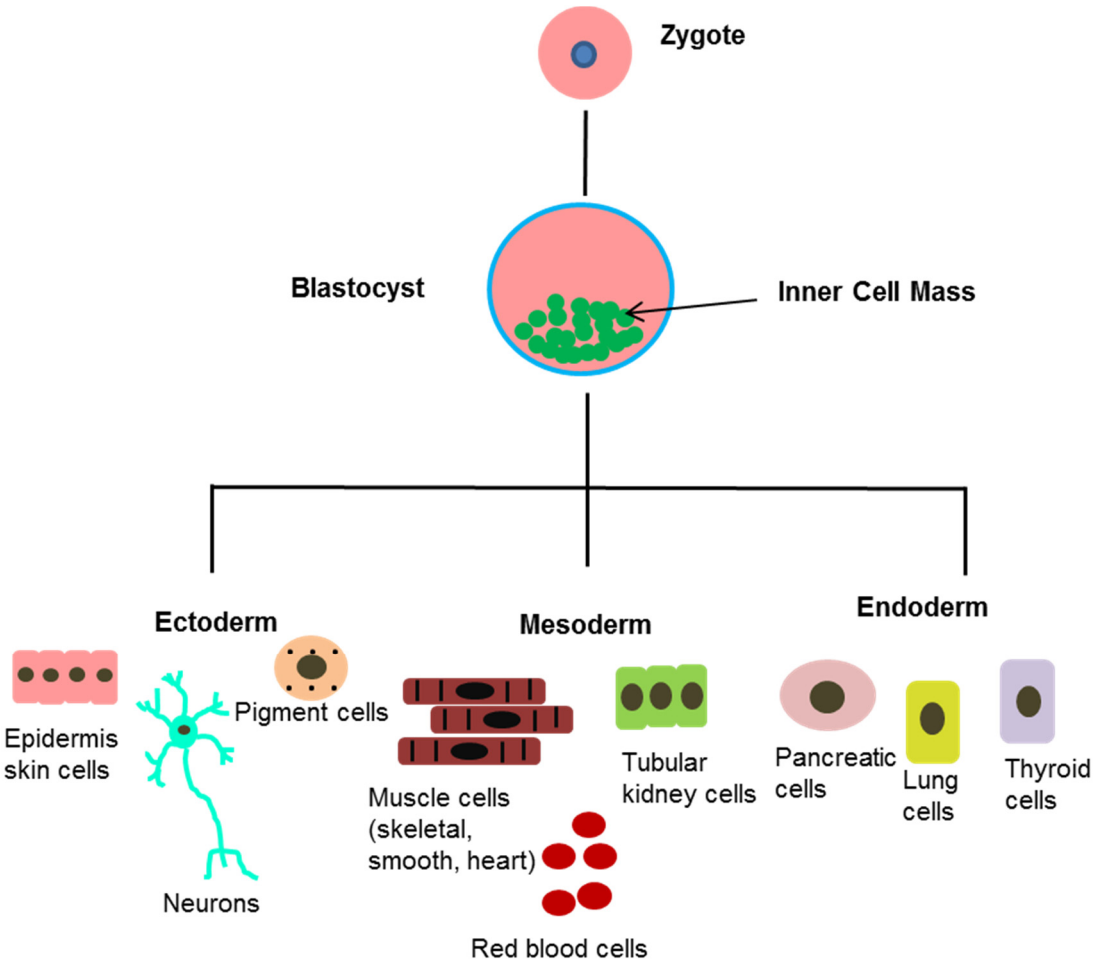


Figure 1-3 Transition from the zygote to the whole organism.

The totipotent zygote, after a series of cleavages, forms the blastocyst. The blastocyst is comprised of an outside layer known as trophectoderm and a population of cells located within the cavity known as the Inner Cell Mass. These cells are pluripotent and give rise to the three germ layers known as ectoderm, mesoderm and endoderm. Further differentiation events will create the specialised types of cells found in the human body. The ectoderm gives rise to cells of the epidermis and the nervous system. The mesoderm specialises to cells of the musculoskeletal system and red blood cells. The endoderm gives rise to the epithelia of the respiratory and gastrointestinal system.

Pluripotency and differentiation associated transcription factors orchestrate lineage specification (Strumpf et al., 2005). In mice, cells that commit to the ICM express NANOG and Octamer binding Transcription factor-4 (OCT4) (Strumpf et al., 2005). At the same time caudal type homeobox 2 (CDX2) is exclusively expressed by TE cells (Strumpf et al., 2005). CDX2 is actually required for the formation of TE since it safeguards the TE fate by repressing NANOG and OCT4 (Strumpf et al., 2005). Similarly, further divergence of the ICM towards the EPI and the PE is governed by mutual exclusion of transcription factors (Sozen *et al*, 2014). Briefly, ICM cells in mice are not identical but they seem to express either NANOG or GATA6 and they later progress into either EPI cells or PE cells respectively (Chazaud et al., 2006). SOX17 seems to follow a similar pattern of expression; first it is randomly expressed in the ICM and later it is confined to the PE (Artus *et al*, 2011).

Immunofluorescence studies in human preimplantation embryos elucidated the spatiotemporal pattern of expression for key pluripotency and lineage specification factors. OCT4 is initially expressed on day 3 (8-cell stage) and is found in both ICM and TE until day 5 (Niakan & Eggan, 2013). However, by day 6 OCT4 is mainly expressed in the ICM (Niakan & Eggan, 2013). CDX2 is first detected on day 5, in the TE where it overlaps with OCT4, while by day 6 TE cells mostly express CDX2 (Niakan & Eggan, 2013). NANOG is also expressed in pre-implantation embryos and is confined to the ICM before OCT4 (Niakan & Eggan, 2013). SOX17 is initially expressed on day 4-5 in the ICM and overlaps partially with OCT4 (Niakan & Eggan, 2013). Gradually, by day 6 SOX17 expression is confined to the area that will form the PE and most of this area is also positive for GATA6 (Niakan & Eggan, 2013).

Another interesting aspect of early embryogenesis is the expression of X chromosomes. Female mammals inactivate at some stage of development either the paternally or the maternally inherited X chromosome (XCI) (LYON, 1961). This process compensates for the existence of only one X chromosome in males (LYON, 1961). XCI is regulated by Xist gene (Penny et al, 1996). Female mice initially inactivate exclusively the paternally derived X chromosome and later in development they reactivate it and randomly inactivate one of the two chromosomes (Heard, et al, 2004). Human female preimplantation embryos display a different pattern of dosage compensation (Petropoulos et al., 2016). Recent data from single cell RNA sequencing indicate that dosage compensation occurs gradually before implantation via a reduction in transcription of both X chromosomes (Petropoulos et al., 2016).

1.1.3 Human embryonic stem cells, origins and properties

The term embryonic stem (ES) cell was first introduced in 1981 when scientists reported the first successful establishment of a stem cell line that displayed pluripotency features and was derived from preimplantation mouse blastocysts (Martin, 1981). This distinguishes them from embryonal carcinoma cells which are pluripotent as well but are derived from teratocarcinomas (Martin, 1981). ES cells should originate either from the preimplantation or the peri-implantation embryo, they should be maintained for several passages in culture without signs of differentiation and they should maintain their ability to give rise to derivatives of the ectoderm, mesoderm and endoderm (Thomson & Marshall, 1998).

In 1998, Thomson and colleagues reported that ES cells were successfully derived from human blastocysts leading to the generation of 5 cell lines (H1, H7, H9, H13,

and H14) (Thomson et al., 1998). Their differentiation potential includes derivatives of all three germ layers (Thomson et al., 1998). Additionally, prolonged differentiation results in the secretion of alpha fetoprotein (AFP) and human chorionic gonadotropin (hCG) (Thomson et al., 1998). These hormones are compatible with trophoblast differentiation (Thomson et al., 1998). Morphologically, hESCs cells grow in colonies and they have a high nuclear/cytoplasmic ratio (Thomson et al., 1998). In terms of marker expression they express the same set of markers found on other primate ES cells such as Stage Specific Antigen-3 (SSEA3), Stage Specific Antigen-4 (SSEA4), Tra-1-60, Tra-1-80 and Alkaline Phosphatase (ALP) (Thomson et al., 1998). A major difference with mESCs is the fact that the latter express SSEA1 but not SEEA3 and SSEA4 (Thomson et al., 1998).

This scientific achievement created the opportunity to gain access to the processes that govern the first stages of embryonic development thus allowing the better understanding of fertility and pregnancy related pathology (Thomson et al., 1998). Furthermore, the ability to generate in culture the cell types of interest fueled the promise that degenerative diseases could find treatment (Thomson et al., 1998).

1.1.4 Human embryonic stem cells, therapeutic potential

Soon after their discovery, hESCs were used in several differentiation protocols and the resulting populations were tested for their safety and efficacy in numerous diseases. In 2005, oligodendrocyte progenitor cells were transplanted in rats with spinal cord injury and they managed to differentiate *in vivo*, facilitate myelin restoration and improve animals' movement (Keirstead et al., 2005). In 2010, Geron wanted to repeat the experiment in humans and started a pioneering clinical trial with

5 patients which was abruptly discontinued due to high cost (Ilic et al, 2015). According to the preliminary results oligodendrocyte precursor transplantation was proven safe but failed to reverse the neurological damage (Ilic et al., 2015).

The first steps towards stem cell therapy in retina diseases started in 2004 with the successful differentiation of hESCs into retinal pigment epithelia (RPE) cells (Klimanskaya et al., 2004). The first presented study that followed patients with Age Related Macular Degeneration (AMD) and Stargardt macular degeneration who underwent stem cell therapy showed that more than half of the patients had improved their visual ability (Schwartz et al., 2015). Another important aspect of the study is that early checks showed no pathology related to the implantation (Schwartz *et al.*, 2012). Similar findings were reported by another study in which 4 patients, 2 with dry AMD and 2 with Stargardt macular degeneration were subjected to replacement therapy with RPE cells (Song et al., 2015). A year after the treatment the patients experienced no compromise in their health associated with the implantation, while in 3 out 4 cases the visual clarity improved (Song et al., 2015).

Type I Diabetes Melitus is one more disease that could benefit from cell therapy. For this reason, hESCs were differentiated towards pancreatic endoderm (Kroon et al., 2008). Upon implantation in mice these cells further differentiate into cells that display features of normal endocrine cells, including response to glucose (Kroon et al., 2008). In 2014, ViaCyte launched a clinical trial with 40 patients that would last for three years in order to separately test the safety and efficacy of a product called VC-01 (Ilic et al., 2015). This product contains pancreatic precursors originating from hESCs and when the product is placed under the skin precursor cells start differentiating into β-pancreatic cells (Ilic et al., 2015).

Recently the first clinical trial involving the use of hESCs for the treatment of heart failure commenced in France (Ilic et al., 2015). The first patient was subjected to a combination of conventional surgical treatment and cardiac progenitor transplantation (Menasché et al., 2015). Three months later, the patient showed improvement in functional assays and symptomatology as well as signs of contractility in the affected myocardium (Menasché et al., 2015).

Patient safety seems to be a vital priority in the design of clinical studies using PSCs and their differentiated progeny. In fact, the International Stem Cell Initiative suggested in 2016 the establishment of a team which will be responsible for assessing the observed alterations of the genome and epigenome of PSCs and the impact these changes may have in their suitability for clinical applications (Andrews et al., 2017).

1.1.5 Human embryonic stem cells, culture systems

Optimizing culture conditions for hESCs has been a parallel to the elucidation of pluripotency regulating mechanisms and an effort to maximize chances for clinical applications. Historically, the first establishment of hESC lines was accomplished in a co-culture system with mouse embryonic fibroblasts (MEFs) (Thomson et al., 1998). For a long time MEFs were critical for the propagation of hESCs and even when they were not used researchers had to condition the media (CM) with MEFs (L. Wang et al., 2007). Conditioning is the exposure of media to MEFs thus enabling the enrichment of media with secreted factors which sustain pluripotency (L. Wang et al., 2007). However, these techniques could involve the risk of contaminating human cultures with animal derived pathogens (Desai *et al*, 2015). Indeed, soon

after the first reports claiming that human cells can also serve as feeders (Cheng et al., 2003), researchers showed that 1/3 of tested MEFs are positive for viruses of the Retroviridae family (Cobo et al., 2008). Moreover, given that some hESC lines are grown in media which contain fetal calf serum (FCS) there is also the possibility of transmitting prions or bovine viruses to the culture (Desai *et al*, 2015).

In 2001, matrigel was successfully introduced as an alternative to feeder cells (C. Xu et al., 2001). Matrigel is a term used to describe extracellular matrix produced by mouse teratocarcinoma cells (Kleinman & Martin, 2005) However, the use of matrigel is accompanied by concerns about efficacy and safety (Desai, *et al*, 2015). From a technical point of view, matrigel composition is not always identical and can therefore cause variation in differentiation efficiency (Desai *et al*, 2015). Furthermore, in 2011, a case study reported the transmission of LDV virus in mice upon injection with contaminated matrigel (Carlson Scholz et al., 2011).

The use of knockout serum replacement (KSR) was a big step towards xeno-free media (Desai *et al*, 2015). However, xeno-free media should be devoid of any non-human components and therefore bovine serum albumin (BSA) was substituted by human serum albumin (HSA) (Desai *et al*, 2015). In addition, the road towards chemically defined media demanded the knowledge of all the components and their concentrations and consequently HSA should be reconstituted rather than plasma derived (Desai *et al*, 2015). Today, these requirements are met by combining commercially available media and substrates (such as TeSR2 combined with laminin fragments or E8 combined with recombinant human vitronectin) and enable a safer propagation of hESC thus improving the chances for therapeutic use (Desai, *et al*, 2015).

1.1.6 Generation of induced pluripotent stem cells (iPSCs)

Traditional embryology viewed embryonic development as a one way process in which cells that specialise inevitably lose their developmental potential. This idea was shaken in 1962 when Gurdon's experiments in Xenopus questioned the irreversibility of differentiation. Briefly, he found that nuclei from intestinal epithelial cells once transplanted in an egg (somatic cell nuclear transfer/SCNT) can develop into a feeding tadpole (whole organism) (GURDON, 1962). Nuclear Reprogramming is a term that was later introduced to describe the process that reverses differentiation and brings a mature cell back to the embryonic state (Hochedlinger & Jaenisch, 2006). For several decades researchers studied SCNT in many organisms (Mitalipov & Wolf, 2009). In 1996, scientists reported the first successful generation of live sheep produced by transplantation of nuclei from a cell line in sheep oocyte (Campbell et al., 1996).

The most ground-breaking progress on nuclear reprogramming was published 10 years later. Scientists reported that mouse embryonic and adult fibroblasts can turn into embryonic stem cell like cells upon introduction of OCT4, SOX2, c-MYC and KLF4 (Takahashi & Yamanaka, 2006). This powerful achievement could open the road to personalised regenerative therapy bypassing ethical issues associated with hESCs and avoiding the problem of donor-recipient incompatibility (Takahashi & Yamanaka, 2006). The next year, two independent groups accomplished the generation of induced pluripotent stem cells (iPSCs) from human somatic cells. Yamanaka's group reprogrammed adult human fibroblasts using the same transcription factors that were successfully tested in mice (Takahashi et al., 2007).

Thomson's group kept OCT4 and SOX2 but replaced the other two factors with NANOG and LIN28 (J. Yu et al., 2007).

However, soon after iPSCs' discovery researchers raised concerns about their similarity with hESCs and their safety (Koyanagi-Aoi et al., 2013). In 2013, a large study compared 49 hiPSC lines with 10 hESC lines in terms of mRNA expression and DNA methylation and found no significant differences (Koyanagi-Aoi et al., 2013). Nevertheless, this study also revealed that some hiPSC lines show resistance to neuronal differentiation and therefore induce tumour formation upon transplantation in mice (Koyanagi-Aoi et al., 2013). The Human Induced Pluripotent Stem Cell Initiative, generated 711 hiPSC lines from 301 healthy donors and addressed the cause of the observed heterogeneity in hiPSC lines (Kilpinen et al., 2017). According to their findings a significant proportion of differences is due to the fact that they originate from different donors (Kilpinen et al., 2017).

1.2 Dissecting Pluripotency

1.2.1 Key pluripotency transcription factors

Several transcription factors are responsible for maintaining mESC pluripotency (Loh & Lim, 2011). Three of them, NANOG, OCT4 and SOX2 are believed to be the core centre of pluripotency (J. Silva & Smith, 2008).

OCT4 is encoded by Pou5f1 gene and can be found with the alternative names OCT3, OCT3,4 , OTF3, NF-A3 (Zeineddine et al., 2014). OCT4 belongs to a family of proteins that bind the octamer motif ATGCAAAT and is found in unfertilised oocytes, ESCs and Primordial Germ Cells (PGCs) (Schöler et al., 1989). Perturbation of OCT4 levels in mESCs results in differentiation (Niwa *et al*, 2000). In the absence of OCT4

mESCs give rise to TE cells while an increase in OCT4 leads to the expression of mesodermal and endodermal associated markers (Niwa *et al*, 2000).

LIF maintains mESCs in culture in an undifferentiated state (Williams et al., 1988) and the central transcription factor that acts downstream of LIF is STAT3 (Niwa et al., 1998). NANOG is a homeoprotein that maintains pluripotency in mESCs independently of LIF (Mitsui et al., 2003), (Chambers et al., 2003) but at the same time is dispensable for self-renewal (Chambers et al., 2007). Furthermore, NANOG levels vary in mESCs and the purpose for these fluctuations can be explained though the rheostat model (Chambers et al., 2007). This model proposes that mESCs can reversibly transit from a NANOG high state to a NANOG low state (Chambers et al., 2007). The latter one is permissive for differentiation as long as there are stimuli to direct fate acquisition, otherwise, mESCs return to their normal status (Chambers et al., 2007).

SOX2 belongs to the SOX (Sry-related HMG box) family of proteins which can bind to DNA via their HMG domain (Kamachi, Uchikawa, & Kondoh, 2000). Mutant mice lacking SOX2 form blastocysts but die after the implantation (Avilion et al., 2003). Mutant blastocysts grown in culture fail to properly specify their ICM and give rise to TE and PE (Avilion et al., 2003). The presence of maternal SOX2 in the ICM implies that SOX2 could have an essential role earlier in development (Avilion et al., 2003). Further studies targeting both the maternal and the embryonic SOX2 are needed in order to truly elucidate the developmental importance of this transcription factor (Chambers & Smith, 2004). Such studies would conclude about the similarity of SOX2 mutant phenotype and OCT4 mutant phenotype (Chambers & Smith, 2004). SOX2 is essential for pluripotency maintenance in mESCs as SOX2 null mESCs are

driven to TE cells (Masui et al., 2007). Indeed this phenotype is similar with the one observed in Oct4 null mESCs (Niwa et al., 2000). Furthermore, SOX2 seems to affect pluripotency by affecting the expression of OCT4 (Masui et al., 2007). Briefly, when SOX2 is absent, ES cells upregulate the expression of OCT4 negative regulators such as Nr2f2 (Masui et al., 2007). At the same time they downregulate Nr5a2 which is a positive OCT4 regulator as well as OCT4 itself (Masui et al., 2007).

NANOG, OCT4 and SOX2 seem to act as a team in transcriptional regulation of hESCs (Boyer et al., 2005). This is supported by the fact that there is significant overlap on the target genes while all three transcription factors bind together to the genes that encode them (Boyer et al., 2005). Despite their role in pluripotency these key factors are also involved in differentiation **(Fig. 1-4)**. In fact, each one of them favours the acquisition of a different fate (Wang et al., 2012). High OCT4 levels maintain self-renewal but when BMP4 is presented to cells they give rise to ME populations (Wang et al., 2012). On the contrary, low OCT4 levels are not compatible with pluripotency (Wang et al., 2012). OCT4 knock down (KD) hESCs differentiate towards embryonic ectoderm in the absence of stimulation but show extraembryonic differentiation upon BMP4 induction (Wang et al., 2012). Finally, NANOG inhibits the formation of Neuroectoderm (NE) and neural crest and SOX2 inhibits ME differentiation (Wang et al., 2012). Once more these results imply that development utilizes the same machinery to accomplish distinct goals according to the needs of the developing organism.

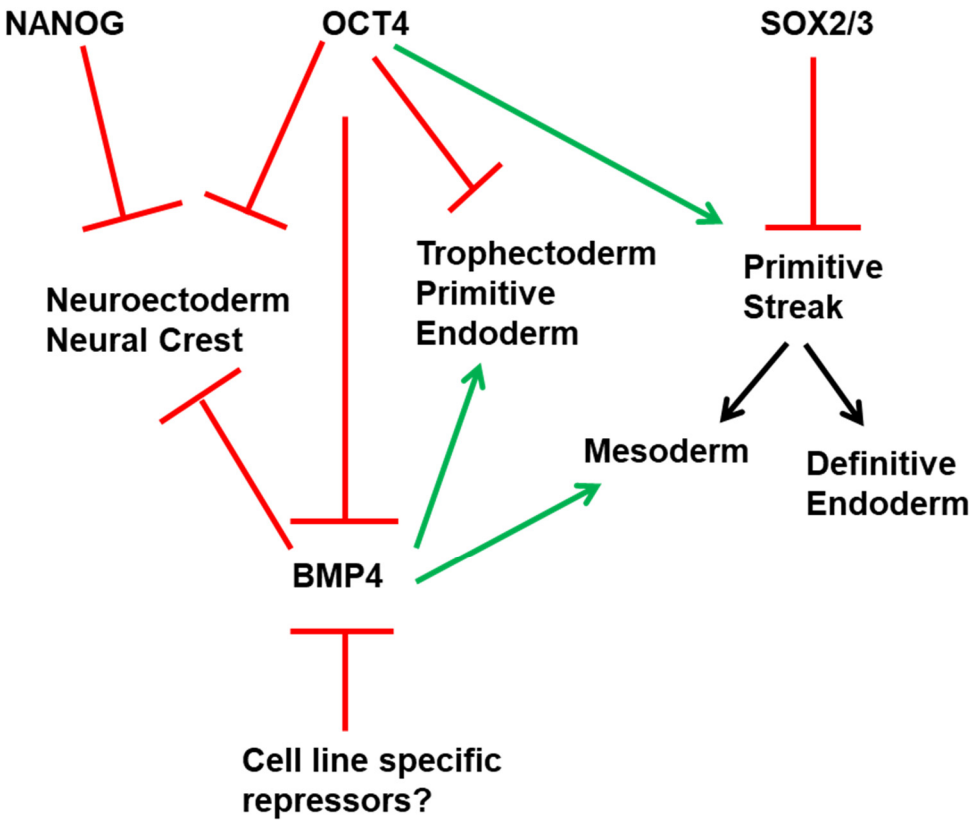


Figure 1-4 Pluripotency transcription factors in lineage specification.
Scientific findings support a pluripotency model in which each one of the key pluripotency factors plays a role in cell fate acquisition. The green arrows indicate a favoured outcome, while red lines indicate inhibition. NANOG blocks neuroectodermal and neural crest differentiation. SOX2 blocks PS and OCT4 blocks NE, neural crest, TE and PE. BMP4 can induce several fates both embryonic and extraembryonic in accordance with OCT4 levels. In the presence of high OCT4 levels BMP4 favours ME differentiation. Meanwhile, when OCT4 levels are reduced, BMP4 is able to induce TE differentiation. Adapted from Wang et al., 2012.

1.2.2 Naïve and Primed Pluripotency

Both hESCs and mESCs are considered pluripotent and have similar origin (Brons et al., 2007). However, pluripotency is supported by different signalling networks in the two species (Brons et al., 2007). The combination of Activin A and FGF2 can

successfully maintain hESCs in culture without the need of feeders, CM or SR (Vallier et al., 2005). On the other hand, mESCs' self-renewal capacity relies on LIF (Williams et al., 1988) and BMP4 with the latter growth factor inhibiting NE differentiation (Ying et al., 2003).

The derivation of pluripotent ES cells from post implantation embryos (Brons et al., 2007), (Tesar et al., 2007) offers an understanding of the observed differences. Epiblast SCs (EpiSCs) can be isolated from the epiblast of post implantation mouse and rat embryos and are maintained in culture in Activin A containing medium (Brons et al., 2007). Interestingly they grow in flat colonies, they do not tolerate single cell passaging and they have a compromised ability for chimera formation (Brons et al., 2007). All these features are reminiscent of hESCs rather than mESCs (Brons et al., 2007). Scientists therefore propose that pluripotency is not a unique state but can exist either as naïve or as primed (Nichols & Smith, 2009). Ground state naïve pluripotency is represented by the epiblast of the pre-implantation blastocyst (Nichols & Smith, 2009). The post-implantation epiblast transforms into an egg cylinder and becomes ready for differentiation, thus representing the Primed Pluripotency (Nichols & Smith, 2009). At this stage of development female embryos randomly inactivate one of their X chromosomes (Heard, 2004).

Until recently the existence of a naïve state for human pluripotent stem cells was not known (De Los Angeles et al., 2012). Finally, in 2013 researchers reported that naïve pluripotency can be accessed directly from blastocysts, from primed hESCs cells or via reprogramming (Gafni et al., 2013). This state is maintained stable via a cocktail of growth factors (LIF, TGFβ, and FGF2) and inhibitors (ERK1/2i, GSK3bi, JNKi, p38i) (Gafni et al., 2013). Among other surprising features, naïve hiPS cells

can effectively contribute in chimeras (Gafni et al., 2013). This capacity raises the expectations for possible utilization of chimeric models to study diseases and assist regenerative therapy (Gafni et al., 2013).

The intense interest for naïve pluripotency necessitates the establishment of criteria that should be fulfilled in order to define cells as naïve pluripotent SCs (Theunissen et al., 2016). Researchers compared molecular features between naïve cells (grown in the presence of LIF, ACTIVIN A and 5 inhibitors) and human embryos (Theunissen et al., 2016). According to their findings the trasposcriptome, DNA methylation and X-chromosome status of naïve cells is reminiscent of that of morulaes and early blastocysts (Theunissen et al., 2016). Additionally, researchers noted the inability of naïve cells to contribute in chimeras (Theunissen et al., 2016). This result is in contrast to previous findings (Gafni et al., 2013) and suggests that this functional assay is not credible enough to distinguish naïve pluripotency (Theunissen et al., 2016).

1.3 TGFβ superfamily signalling

1.3.1 Ligands and antagonists

The TGFβ pathway is implicated in a wide range of processes from the embryo development to maintaining homeostasis in adult tissues while perturbations are associated with diseases such as fibrosis and metastasis (Massagué, 2012). Despite the simplicity of the pathway, there are numerous factors fine tuning the response and even 50,000 published papers are insufficient to finally define the way in which context affects the signalling outcome (Massagué, 2012).

The TGFβ superfamily is comprised of approximately 30 growth factors which are classified in two groups (Weiss & Attisano, 2013). TGFβs, ACTIVINs, NODALs and some of the GDFs belong to the TGFβ group while BMPs, AMH and the majority of GDFs belong to the BMP group (Weiss & Attisano, 2013).

TGFβs are secreted from the cells in the form of a complex which consists of the latency-associated peptide (LAP) and the ligand (Saharinen et al., 1999). Usually, the LAP is carrying the latent TGFβ-binding protein (LTBP), which facilitates the secretion and proper folding of the ligand and mediates interaction of the ligand with the extracellular matrix (ECM) (Saharinen et al., 1999). However, the mature ligand needs to be separated from the LAP in order to exert its functions (Saharinen et al., 1999). ACTIVINs and BMPs are not secreted as latent complexes and therefore they skip this step (Miyazono, Maeda, & Imamura, 2005). One of the key structural features of mature TGFβ ligands is the cystine knot, which is comprised of 6 cysteines bound together with disulphide bonds (Weiss & Attisano, 2013). An additional cysteine which is present in the majority of ligands facilitates ligand dimerization (Weiss & Attisano, 2013).

TGFβ superfamily is negatively regulated by the presence of antagonists (Miyazono, 2000). There are two types of antagonists, those which inhibit the ligand via direct binding and those which compete with the ligand for receptor binding (Miyazono, 2000). NOGGIN, CHORDIN and CERBERUS are some of the numerous identified BMP binding antagonists and FOLLISTATIN is a well-known ACTIVIN binding antagonist (Miyazono, 2000). On the other hand, LEFTY-1 and LEFTY-2 block NODAL by competition (Miyazono, 2000).

1.3.2 Receptors and SMADs

The human genome encodes 12 Serine/Threonine Kinase Receptors which belong to the TGFβ superfamily (Manning et al., 2002). They are categorized in Type I and Type II receptors (Shi & Massagué, 2003). Receptors encode an extracellular domain (N-terminal) which mediates ligand binding, an intracellular domain (C-terminal) which exerts the kinase activity and a transmembrane domain which connects the former two (Shi & Massagué, 2003). An important difference between Type I and Type II receptors is the presence of a GS sequence in the former (Shi & Massagué, 2003). Phosphorylation of the GS region by Type II receptor leads to initiation of the pathway (Massagué, 1998).

BMPs bind first to the Type I receptors and the arising complex binds to Type II receptors (Shi & Massagué, 2003). On the contrary, TGFβs and ACTIVINs bind to Type II receptors and once the binding occurs Type I receptors can recognize the ligand and join the complex (Massagué, 1998). The signalling cascade continues with SMADs' phosphorylation (Shi & Massagué, 2003). Phosphorylated R-SMADs associate with SMAD4 and enter the nucleus in order to modulate gene expression (Shi & Massagué, 2003). The term SMAD was introduced in 1996 in order to describe 5 proteins involved in TGFβ signalling in human, mouse and Xenopus (Derynck et al., 1996). The name is a combination of two genes known as Sma and Mad which are involved in the TGFβ pathway in C.Elegans and Drosophila respectively (Derynck et al., 1996). There are three categories of SMADs: R-SMADs which are regulated by the Type I receptors, Co-SMAD (SMAD4) which facilitates signalling and I-SMADs (SMAD 6,7) which block the function of the other two groups (Shi & Massagué, 2003). R-SMADs diverge further as SMAD2,3 propagate signal

from ACTIVINs and TGFβs while SMAD1,5,8 propagate signal from BMPs. (Shi & Massagué, 2003).

SMADs display structural features that endow them with their functional properties **(Fig. 1-5)**. All SMADs share a similar C-terminal MH2 domain that mediates binding to the receptor and interaction with SMADs (Weiss and Attisano, 2013). The MH2 domain of R-SMADs contains an SXS motif which becomes phosphorylated by the receptor thus leading to R-SMAD activation (Abdollah et al., 1997),(Souchelnytskyi et al., 1997). Moreover, it contains an L3 loop which sets the basis for the specific interaction between R-SMADs and the receptors (Y. G. Chen et al., 1998). The N-terminal (MH1) domain is highly similar in R-SMADs and SMAD4 but not in I-SMADs (Moustakas, Souchelnytskyi, & Heldin, 2001). The existence of a β hairpin in the MH1 domain of SMAD3 allows association with SMAD Binding Elements (SBE) in the genome (Y. Shi et al., 1998) thus suggesting that R-SMADs and SMAD4 can use the same structure to achieve DNA binding (Shi & Massagué, 2003).

In all SMADs the N-terminal domain and the C-terminal domain are connected through a linker region which contains multiple sites amenable to phosphorylation from several signalling pathways (Shi & Massagué, 2003). Thus, the linker region serves as a platform for interplay between pathways (Shi & Massagué, 2003). Additionally the presence of a PY motif that is recognised by SMURF1 and SMURF2 facilitates the proteasomal degradation of SMADs (Shi & Massagué, 2003).

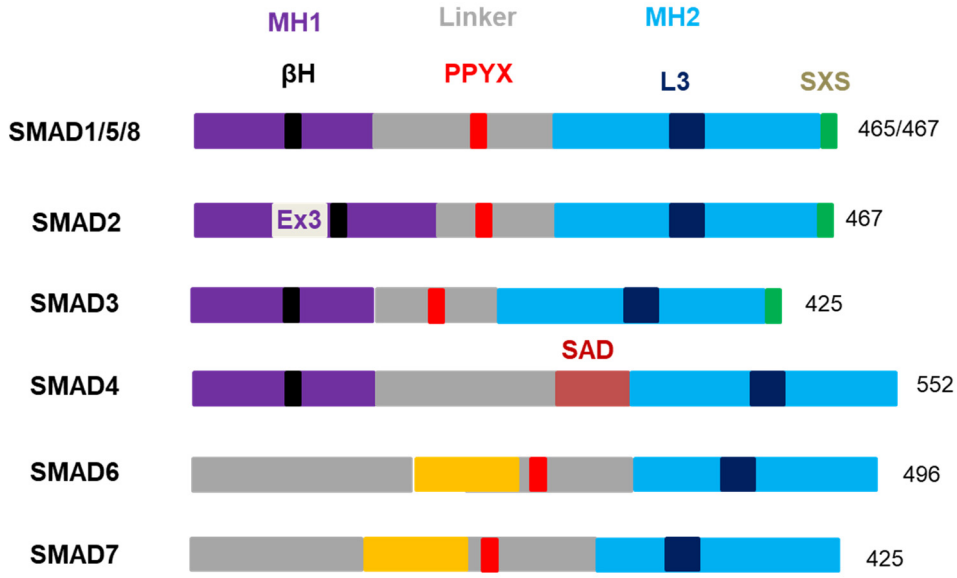


Figure 1-5 Key structural components of SMADs.

All SMADs share a highly conserved MH2 domain (light blue). The L3 loop (dark blue) allows specificity of receptor-SMAD interaction. R-SMADs contain an SXS motif (green) for receptor phosphorylation. R-SMADs and SMAD4 share a similar MH1 domain (purple). A β hairpin structure (black) mediates DNA binding for all R-SMADs and SMAD4. One SMAD2 isoform that contains an additional exon (Ex3) does not bind to DNA. All SMADs contain a linker region (grey) containing a PPYX motif (red) which mediates interaction with E-3 ligases. SMAD4 also contains a domain dedicated to transcriptional regulation (SAD).The name and number of aminoacids for each SMAD is on the left and right side of each structure respectively. Adapted from: (Weiss & Attisano, 2013).

Although, SMADs enter the nucleus to exert their functions the mechanism used for this purpose is elusive. Researchers have proposed two models, one which is mediated by the MH1 domain and is karyopherin dependent and another one that relies on interaction between nucleoporins and the MH2 domain (Hill, 2009). Collective findings seem to suggest that SMADs constantly move inside and outside the nucleus and that upon stimulation, SMADs enter the nucleus in a karyopherin dependent manner (Hill, 2009). The MH1 domain of SMAD1,3,4 contains nuclear

localisation signal (NLS) like motifs necessary for the import (Hill, 2009). Despite the fact that, Co-SMAD and R-SMADs (except for a SMAD2 isoform) can directly interact with DNA (Massagué, Seoane, & Wotton, 2005), this interaction is not strong enough and SBEs are so frequently found across the genome that they seem general (Beyer, Narimatsu et al., 2013). SMADs need tight transcriptional control of specific genes and they bypass this obstacle by associating with guiding transcription factors which are context dependent (Beyer, Narimatsu et al., 2013).

I-SMADs can bind to the Type I receptors and block R-SMADS binding thus inhibiting the initiation of signalling cascade (Xu, Liu, & Derynck, 2012). SMAD6 is mainly used in BMP response while SMAD7 can inhibit both branches (Xu, Liu, & Derynck, 2012). In the case of TGFβ, SMAD7 ensures signal termination in several ways. First, SMAD7 binds the Type I receptor and doesn't allow access to SMAD2,3 for activation (Xu, Liu, & Derynck, 2012). Second, SMAD7 recruits SMURF2 an E-3 ligase to the receptor and therefore targets it for lysosomal and proteasomal degradation (Kavsak et al., 2000). Third, SMAD7 mediates the interaction of the Type I receptor with the protein phosphatase-1C (PP1C) thus leading to receptor deactivation (W. Shi et al., 2004). Finally, in the nucleus, SMAD7 competes with SMAD complexes for DNA binding and therefore attenuates the TGFβ instructed regulation of gene expression (S. Zhang et al., 2007).

1.3.3 Non SMAD signalling

It is generally accepted that SMADs have the most prominent role in orchestrating the cellular response to the TGFβ superfamily signalling. Nevertheless, SMADs are not the unique mediators (Moustakas & Heldin, 2005). A growing body of evidence

suggests that there are several more mediators such as ERK, JNK, p38, PI3K, Akt and Rho GTPases that make up the non-SMAD pathway (Y. E. Zhang, 2017). These proteins can contribute in signalling via three ways: first, non-SMAD mediators can modify SMADs, second, SMADS can modify non-SMAD mediators thus affecting other cascades and third, receptors can activate both SMADs and non-SMAD mediators (Moustakas & Heldin, 2005).

1.3.4 Endocytosis overview

Endocytosis refers to the process of internalising part of the plasma membrane with the accompanying extracellular matter within the cell (Scita & Di Fiore, 2010). In general, large particles (bigger than 500nm) enter the cell via phagocytosis while smaller particles can use several other endocytic pathways (Kumari et al., 2010). These pathways further diverge according to the type of coat that is used to protect the internalized pit (e.g. clathrin, caveolin) and the mechanism that allows scission of the newly formed vesicle from the membrane (dynamin dependent or dynamin independent) (Kumari et al., 2010).

Research focused for a long time on clathrin mediated endocytosis (CME) which was considered the main route for entrance in the cell (Doherty & McMahon, 2009). Clathrin was named after its property to form structures that resemble a cage (Kirchhausen, 2000). CME consists of 4 steps: first the cargo is presented to the cell, then the cell forms a coat made up of clathrin, next, the plasma membrane internalises in the form of a coated pit and last the pit is separated from the membrane by scission and moves in the cell as a coated vesicle (Kumari et al., 2010).

However, advances in microscopy indicated that the presence of clathrin in membrane invaginations is not the rule thus implying that there are other routes used in parallel to CME (Doherty & McMahon, 2009). The plasma membrane is not homogenous and is composed of lipid rafts rich in cholesterol and glycospingolipids displaying resistance to non-ionic detergents such as Triton X-100 (Pike, 2003). Caveolae are a subgroup of lipid rafts which are observed as invaginations and contain caveolin-1 (Pike, 2003).

Regardless of the internalization route used, cargo visits first the sorting compartment of the cell, the early endosome (EE) (Jovic et al., 2010). In this compartment decisions are made that direct cargo to their next destination (Jovic et al., 2010). The available options are: return to the cell surface, progression to the lysosome for degradation and transport to the Trans-Golgi network (Jovic et al., 2010).

EEs display the following features (Naslavsky & Caplan, 2018): they are highly abundant in the lipid PI3P (Kobayashi, Gu, & Gruenberg, 1998), they are not static but instead they are reshaped by fusion with other EEs or internalised vesicles (Diaz, Mayorga, & Stahl, 1988). Additionally, they have a slightly acidic pH (R. F. Murphy, Powers, & Cantor, 1984) and several Ras-associated binding (RAB) proteins are bound on their membrane (Delevoye & Goud, 2015).

PI3P is a phosphorylated derivative of phosphatidylinositol which is generated via the action of PI3K (Stenmark, Aasland, & Driscoll, 2002). There are two known domains that enable PI3P binding: the Phox homology (PH) domain and the FYVE domain (Stenmark, Aasland and Driscoll, 2002). The latter one is named after the initials of the first 4 proteins that were found containing this domain (Fab1p, YOTB, Vac1p and

EEA1) and today it is known that there almost 30 FYVE proteins in the human proteome (Stenmark, Aasland and Driscoll, 2002).

RABs are small GTPases with significant roles in endosomal functions (Zhen & Stenmark, 2015). RABs exist in two forms, the active GTP bound that enables interaction with other proteins known as RAB effectors and the inactive GDP form (Zhen & Stenmark, 2015). RAB5 is the most well studied RAB protein found on the EE (Jovic et al., 2010). RAB5 actively shapes the composition of the EE since its presence results in PI3P production and subsequent attraction of PI3P binding proteins (Naslavsky & Caplan, 2018). One of them, Early Endosomal Antigen-1 (EEA1) is considered to be a typical marker for the compartment (Jovic et al., 2010). EEA1 is specifically found on the EE (Mu et al., 1995) and mediates fusion of endosomes as induced by RAB5 (Simonsen et al., 1998).

Receptor return to the plasma membrane can occur via two main routes: the fast and the slow route (Naslavsky & Caplan, 2018). In the first pathway, the receptors are separated from the EE in the form of a vesicle which moves away towards the cell surface (Naslavsky & Caplan, 2018). In the second pathway, receptors travel first to the Endocytic Recycling Compartment (ERC) which is found close to nucleus, neighbouring with the Microtubule Organizing Centre (MTOC) (Naslavsky & Caplan, 2018). From there receptors travel in recycling endosomes to their final destination. RAB11, RAB8 and RAB22a are usually found on this type of endosome (Naslavsky & Caplan, 2018).

Proteins destined for degradation usually have ubiquitin signals (Piper & Katzmann, 2007) and follow a route which involves the multivesicular bodies (MVB), the late endosome and finally the lysosome (Scott, Vacca, & Gruenberg, 2014). MVBs arise

when the endosomal membrane invaginates inwards thus generating vesicles within the lumen of the endosome which contain cargo targeted for degradation (Piper & Katzmann, 2007). MVB formation is regulated by the Endosomal Sorting Complexes Required for Transport (ESCRT) machinery (Schmidt & Teis, 2012). There are 5 types of ESCRT complexes (ESCRT-0,I,II,III and VPS4 complex) with distinct composition and function (Schmidt & Teis, 2012). ESCRT-0 initiates the process by binding of Hepatocyte Growth Factor Regulated Tyrosine Kinase Substrate (HRS) on the endosomal PI3P (Schmidt & Teis, 2012). HRS associates with cargo tagged for degradation via its Ubiquitin Interacting Motif (UIM) and this binding is further supported by the recruitment of EPS15 and STAM2 (Jovic et al., 2010). In addition, HRS interacts with TSG01 a component of ESCRT-I thus initiating the subsequent enrichment of the initial assembly with the successive addition of ESCRT-II and ESCRT-III (Jovic et al., 2010).

Late endosomes (LE), which also contain vesicles should not be confused with MVBs (Scott et al., 2014). They have a richer morphology than MVBs and like EEs they can undergo homotypic fusion (Scott et al., 2014). Additionally, like EEs, LEs contain an ATP-dependent proton pump that acidifies their lumen but their pH is more acidic than EEs (Scott et al., 2014). Previous findings suggest that RAB5 and RAB7 regulate the progression from the EE to the LE respectively (Poteryaev et al., 2010) In other words, the compartment gradually changes as it becomes positive for RAB7 and negative for RAB5 (Poteryaev et al., 2010). However, it seems LE share more similarities with lysosomes than with EEs (Naslavsky & Caplan, 2018). RAB7 is typically found on LEs while RAB4 and RAB5 are usually not present (Naslavsky & Caplan, 2018).

The traditional idea is that late endosomes and lysosomes merge, thus generating an endo-lysosome which further progresses in a lysosome (Scott et al., 2014). However, there are also findings suggesting that cargo delivery from the one compartment to the other can occur fast through the so called kiss-and-run approach (Scott et al., 2014). On that note, distinguishing the two compartments at the molecular level is challenging, since lysosomes share mostly the same proteins with LEs (Scott et al., 2014). The lysosome is devoted to degradation thus serving as the catabolic centre of the cell (Saftig & Klumperman, 2009). More than 50 hydrolases deliver the important task of degradation (Saftig & Klumperman, 2009). Additionally, several Lysosome Membrane Proteins (LMPs) are responsible for numerous other tasks such as pH regulation of the organelle and bidirectional delivery of proteins between the lysosome and the cytoplasm (Saftig & Klumperman, 2009).

1.3.5 TGFβ superfamily trafficking

TGFβ receptors localize in the basolateral membrane in polarized cells (S. J. Murphy et al., 2004). TGFβ receptors can enter the cell using both clathrin and caveolin mediated endocytosis (Di Guglielmo et al., 2003). These two types of endocytosis seem to dictate distinct fates for the receptors and therefore regulate pathway activity (Di Guglielmo et al., 2003). Clathrin mediated endocytosis is facilitated by AP2 protein and is regarded as the shortest route to the EE (Yakymovych, Yakymovych, & Heldin, 2018). In more detail, TGFβ receptors that use CME are first found in early endosomes ($EEA1^+$) and then mostly in recycling endosomes ($RAB11^+$) while almost no receptors seem to enter the late endosomes (Di Guglielmo et al., 2003). These data imply that CME is a choice that allows receptor recycling to the plasma

membrane and minimizes receptor degradation. On the other hand, receptors using caveolae have a different fate, as they meet SMAD7 and SMURF2 (Di Guglielmo et al., 2003) which mediate their degradation (Kavsak et al., 2000). Degradation can occur in both lysosomes and proteasomes and in the case of Type I receptor, DAPPER2 is reported to mediate delivery from the LE to the lysosome (F. Huang & Chen, 2012). Overall, these findings compose a model which links endocytosis routes with signalling **(Fig. 1-6)**. Clathrin mediated endocytosis enables signal transduction while caveolae-mediated endocytosis aims to turn the signal off (Di Guglielmo et al., 2003).

However, this view becomes blurred under the light of more recent findings that associate lipid rafts with non-SMAD signalling (Yakymovych et al., 2018). HaCaT cells seem to rely on lipid rafts in order to activate ERK and p38, in the presence of the ligand (Zuo & Chen, 2009). Additionally, low levels of caveolin-1 (siRNA) in hepatocytes are associated with an impaired ability of TGFβ to phosphorylate AKT (Meyer et al., 2013). Several factors are shown to favour the one route of endocytosis or the other **(Fig. 1-6)**. For example, IL-6 guides the formation of receptor complexes in non-lipid rafts thus positively affecting the SMAD pathway (X. L. Zhang et al., 2005). Similarly, ADAM12 guides the type II TGFβ receptor to the early endosome and increases SMAD2 activation (Atfi et al., 2007). On the other hand, receptors in HK2 cells are shifted from $EEA1^{+}$ compartments to caveolin-1 lipid rafts in the presence of hyaluronan (Ito et al., 2004). Therefore, the conclusion of a unified model is still tricky as once more an aspect of the TGFβ pathway shows diversity.

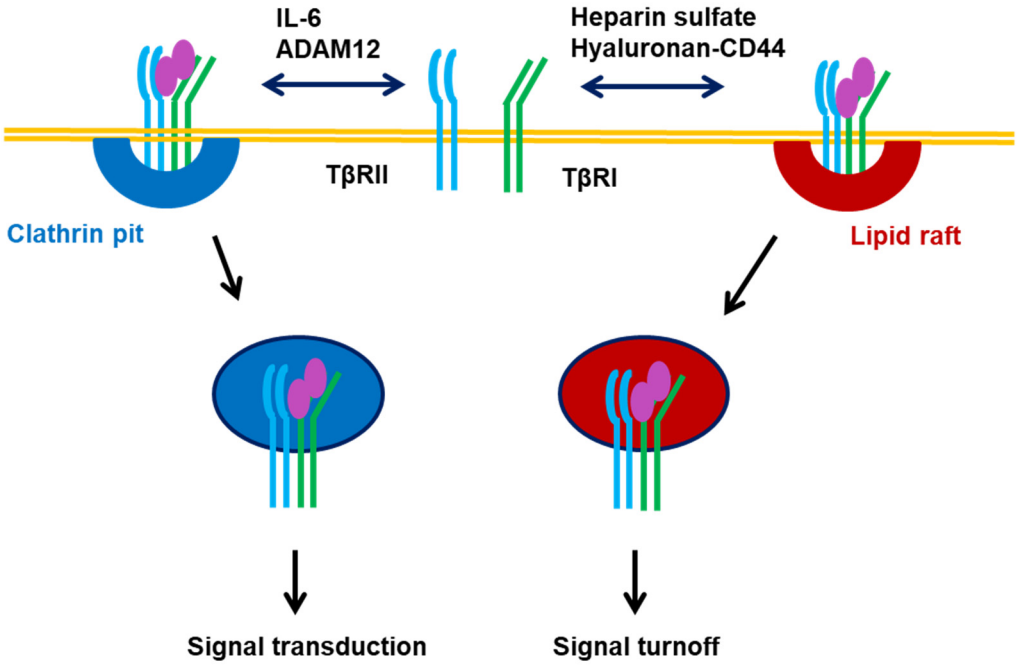


Figure 1-6 Internalization routes of TGFβ receptors.

TGFβ receptors can internalise via two routes, clathrin mediated endocytosis and caveolin mediated endocytosis. TGFβ receptors residing in non-raft parts of the cell membrane use clathrin coated pits (in blue). This route is believed to positively influence signalling and is favoured by the presence of IL-6 and ADAM12. Meanwhile, TGFβ receptors residing in lipid rafts use caveolae (in red). This route is believed to dampen signalling and is favoured by the presence of heparin sulfate and hyaluronan. Furthermore, recent studies enhance the role of lipid rafts suggesting that they are required in non-SMAD signalling. Adapted from: (Y. Chen, 2009).

There is a notion that TGFβ signalling is initiated at the level of the EE. This was introduced because SARA and ENDOFIN are found on the compartment **(Fig. 1-7)** and they positively regulate SMAD signalling (Yakymovych et al., 2018). SARA, is supposed to present SMAD2,3 to the receptors for subsequent phosphorylation and therefore activation (Tsukazaki et al., 1998). Endofin, is involved in the next step of the pathway by possibly bringing together SMAD2 and SMAD4 (Y. Chen et al., 2007). Ongoing research on the numerous and sometimes unexpected aspects of endocytosis in general, has led scientists to suggest the term ‘endocytic matrix’ in order to describe the presence of an endocytic network which is deeply ‘aware’ of the signalling identity of the cell (Scita & Di Fiore, 2010).

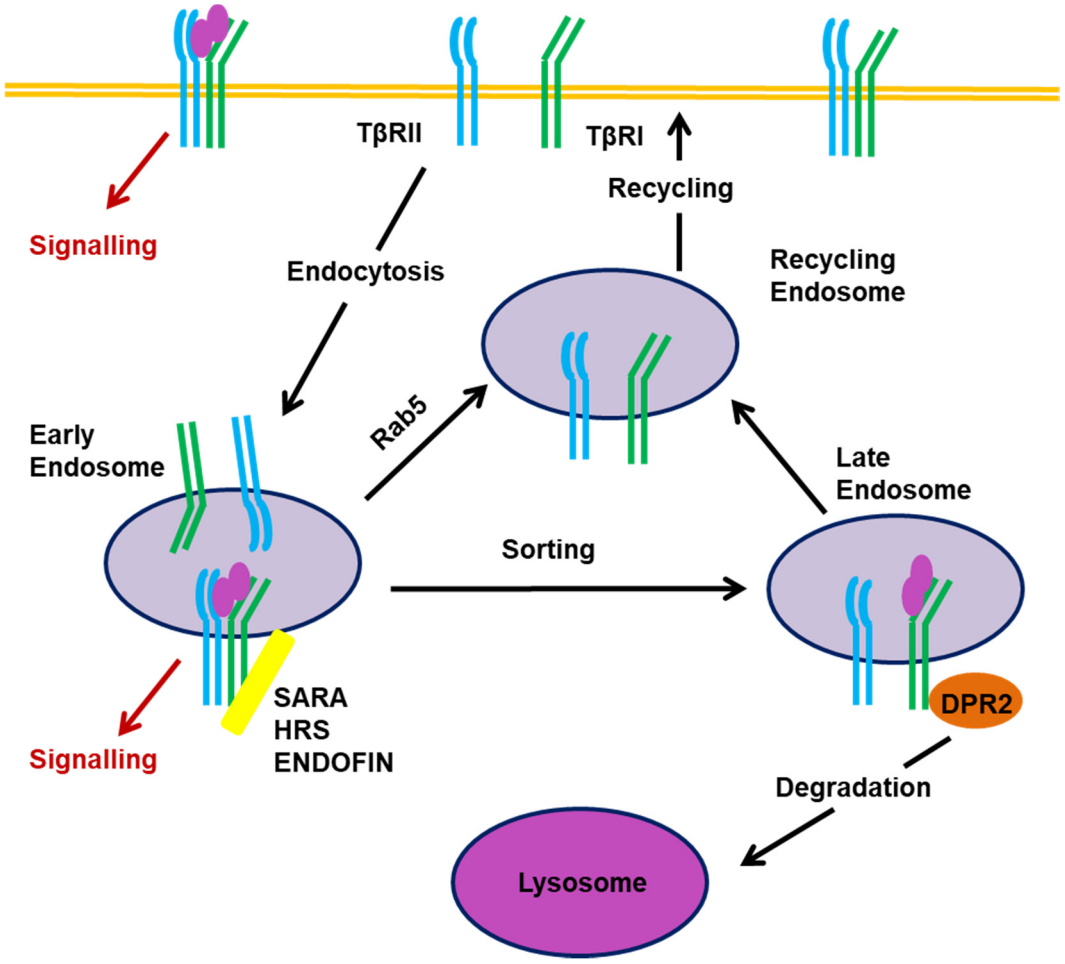


Figure 1-7 TGFβ receptors' trafficking.

Internalised TGFβ receptors reach the EE where they are sorted either for recycling or degradation. For the first decision, receptors return to the cell surface using recycling endosomes. Meanwhile, TGFβ receptors targeted for degradation move towards the late endosome and finally they are delivered to the lysosome possibly through DAPPER2. In this compartment, TGFβ receptors are degraded via the action of specific enzymes. The presence of positive TGFβ regulators (SARA, HRS and ENDOFIN) on the EE raises the question of whether signalling occurs at the plasma membrane and/or the EE and whether this fact affects the signalling outcome. Adapted from: (Y. Chen, 2009).

In general, BMP and TGFβ receptors display similar endocytosis features (Ehrlich, 2016). In detail, BMP receptors undergo constitutive ligand-free internalization, they preferentially enter the cell in clathrin coated vesicles, they signal from the cell surface and they use endosomes to regulate signalling (Ehrlich, 2016). In accordance with the above, BMP initiates SMAD dependent signalling on the non-lipid raft regions of the plasma membrane (Hartung et al., 2006). Although, clathrin mediated endocytosis occurs, it is necessary mostly for downstream regulation of gene expression rather than SMAD activation per se (Hartung et al., 2006). Finally, in an analogy with the TGFβ pathway, lipid rafts are important for the non-SMAD response (Hartung et al., 2006).

In the case of ACTIVIN several interesting suggestions are based on studies in Xenopus. To begin with, internalization of both ligands and receptors is necessary for signalling (Jullien & Gurdon, 2005). Furthermore, the aspect of the time spent in the endocytic compartments becomes a regulator in fate acquisition (Jullien & Gurdon, 2005). This phenomenon, suggests that ligand-receptor complexes that spent a significant amount of time in the compartments before they are degraded, build memory (Jullien & Gurdon, 2005). The advantage of memory is that cells placed in a morphogen gradient, know which decisions to make even when the cells are no more exposed to the growth factor (Jullien & Gurdon, 2005).

1.4 Key Signalling pathways regulating pluripotency and differentiation

1.4.1 TGFβ/ACTIVIN/NODAL signalling

There are several key signalling pathways regulating pluripotency and differentiation and the TGFβ/ACTIVIN/NODAL pathway is one of the most studied. ACTIVIN was

first purified from porcine gonads in 1986 as a protein that induces FSH release, as opposed to Inhibin which was known to inhibit FSH release (Vale et al., 1986), (Ling et al., 1986). There are 4 inhibin subunits (βa, βb βc, βe) which are combined either as homodimers or heterodimers thus forming a range of ACTIVINS (Pauklin & Vallier, 2015). Both ACTIVIN and NODAL, signal through the same cell surface receptors (ALK4 or ALK7 and ACTRIIA or ACTRIIB) and motivate the same intracellular mediators (SMAD2 or 3) (Papanayotou & Collignon, 2014). Given the similarities these two ligands are often presented as a single pathway (Pauklin & Vallier, 2015).

ACTIVINs are present in early mouse embryogenesis. ACTIVINs decline during cleavage divisions and they recover at the morula stage, possibly reflecting degradation of maternal pools and de novo production of embryonic pools respectively (Albano, *et al*, 1993). Additionally, the expression of ACTIVINs shifts from the ICM on day E3.5 to the TE on day E4.5 (Albano, *et al*, 1993). Loss of function experiments for components of the ACTIVIN/NODAL pathway, produce several phenotypes from mild defects in viable mice (e.g. ACTIVIN A) to abnormal gastrulation (e.g. NODAL) (Papanayotou & Collignon, 2014).

A valuable tool to reveal the association between this pathway and pluripotency is offered by the use of hESCs. Extensive studies have established a central role of TGFβ/ACTIVIN/NODAL in hESCs biology. Early studies involved NODAL in the protection of hESCs against spontaneous NE differentiation (Vallier, *et al*, 2004). When scientists investigated the ability of MEFs to support hESCs they found that they secrete ACTIVIN A precursor (Beattie et al., 2005) thus implying that ACTIVIN A could be responsible for the maintenance of pluripotency. Indeed, ACTIVIN A

(Beattie et al., 2005), (Xiao, Yuan, & Sharkis, 2006) or combination of ACTIVIN A with FGF2 (Vallier et al., 2005) enables the undifferentiated propagation of hESCs bypassing the need for MEFs and MEF secreted factors.

Another indirect evidence demonstrating the significance of the pathway is revealed by the comparison of normal hESCs and chromosomally abnormal hESCs with advanced self-renewal capacity (Xiao et al., 2006). The data show that the latter cell line has higher expression of ACTIVIN A and NODAL and lower expression of FOLLISTATIN (an ACTIVIN inhibitor) (Xiao et al., 2006). Accordingly, FOLLISTATIN treatment perturbs two key pluripotency transcription factors (Xiao et al., 2006).

Given the fact that SMAD2,3 are active in undifferentiated hESCs and they decrease upon differentiation (James et al., 2005) it seems reasonable to hypothesize that pluripotency could be SMAD dependent. Interestingly, SMAD2,3 can bind and enhance the activity of Nanog's proximal promoter (R. H. Xu et al., 2008). NANOG is considered a central guard of pluripotency by blocking both the FGF induced NE differentiation as well as the BMP4 induced progression from ME to Endoderm (L. Vallier et al., 2009)

However, the TGFβ/ACTIVIN/NODAL pathway accounts also for ME (P. Zhang et al., 2008) and endodermal differentiation (Brown et al., 2011) and recent findings started dissecting this paradox. One explanation uses the idea of a PI3K switch which regulates cell fate decisions. Briefly, when PI3K/AKT is high then both ERK and WNT are blocked and therefore SMAD2,3 has no other option but to maintain self-renewal (Singh et al., 2012). On the other hand, when PI3K/AKT is compromised, then SMAD2,3 in combination with WNT are free to regulate ME fate acquisition (Singh et al., 2012). Similarly, SMAD2,3 seem to regulate the expression

of different set of genes in pluripotency and endodermal differentiation (Brown et al., 2011) thus implying that the differential effect could be orchestrated by distinct transcription factors utilised in the one condition or the other.

Another interesting insight in ACTIVIN's function comes from epigenetic data in hESCs. This type of cells is known to display a distinct pattern of epigenetic signatures since many important genes have two opposing histone marks, one for activation (H3K4me3) and one for inactivation (H3K27me3) (Bertero et al., 2015). This phenomenon enables hESCs to quickly regulate gene expression (Bertero et al., 2015). COMPASS complex mediates the deposition of H3K4me3 on genes and surprisingly, NANOG and SMAD2,3 seem to direct DPY30, a component of the COMPASS complex to a group of genes associated with pluripotency and a group of genes associated with ME commitment (Bertero et al., 2015). Moreover, a recent study discovered that in hESCs, SMAD2,3 interacts with a complex (METTL3–METTL14–WTAP) which facilitates RNA methylation (m^6A) and directs it towards developmentally important genes (Bertero et al., 2018). Modulated genes display compromised stability and are therefore prone for downregulation with the onset of differentiation (Bertero et al., 2018).

Finally, although SMAD2 and SMAD3 are usually presented together in literature, it should be noted that perhaps they are not of equal importance in hESCs. In fact, only SMAD2 is necessary for pluripotency thanks to its dual ability to maintain NANOG and block BMP (Sakaki-Yumoto, Liu, Ramalho-Santos, Yoshida, & Derynck, 2013). This becomes evident, when SMAD2 decreases and hESCs are vulnerable to autocrine BMP signals which drive them towards ME, TE and germ cell lineages (Sakaki-Yumoto et al., 2013).

1.4.2 BMP signalling

Marshall Urist, first described the generation of bone in muscle through the action of an unknown molecule which he named Bone Morphogenetic Protein (Wagner et al., 2010). Decades of research reveal that the BMP family consists of around 30 cytokines and they exert a wide range of functions from early development to the adult life thus implying that the term Body Morphogenetic Protein would be a more accurate description (Wagner et al., 2010). BMP4 belongs to the osteogenic group (which also contains BMP2, 6, 7 and 9) and has a very crucial developmental role (Wang *et al.*, 2014). BMP4 knock out can be lethal for mice embryos since most of them die early and the observed abnormalities pertain to gastrulation and ME differentiation (Winnier et al., 1995).

LIF can support undifferentiated propagation of mESCs via STAT3 (Niwa et al., 1998). However, LIF alone cannot bypass the need for serum as mESCs differentiate (Ying et al., 2003). Interestingly, the combination of LIF and BMP4 can protect pluripotency in the absence of feeders and serum (Ying et al., 2003). This function of BMP4 is exerted via SMADs which ultimately regulate the transcription of Id genes (Ying et al., 2003). Moreover, BMP4 seems to sustain the undifferentiated state by regulating ERK (Z. Li et al., 2012). Indeed, SMAD dependent signalling induces the expression of DUSP9, a phosphatase that inhibits ERK's activity thus balancing out LIF's effect on ERK (Z. Li et al., 2012). These findings establish BMP as a pluripotency compatible factor in mESCs.

On the contrary, BMP4 is a differentiation stimulus for hESCs and the explanation for this seems to be the fact that these two types of cells reflect distinct *in vivo* stages of development (Z. Li & Chen, 2013). More specifically, mESCs correspond

to the ICM, while hESCs share more similarities with the epiblast (Z. Li & Chen, 2013). The effect of BMP4 on hESCs has raised a lot of discrepancy in the literature since numerous studies have linked BMP4 with different fates both embryonic and extra-embryonic **(Fig. 1-8).**

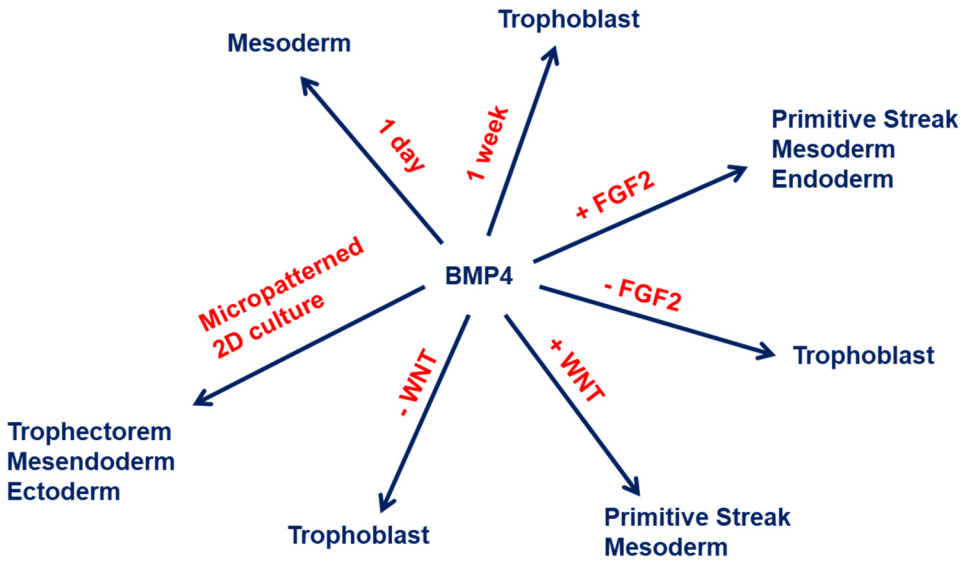


Figure 1-8 BMP4 induced fates in hESCs.
BMP4 is associated with several fates and continuous studies dissect the underlying mechanisms responsible for the observed diversity. The effect of BMP4 depends on several factors including the duration of induction, the presence of FGF2 and WNT activity. In detail, BMP4 induction for 1 day favours mesodermal fates, while 1 week induction favours trophoblastic differentiation. BMP4 leads to mesodermal differentiation when FGF2 is present, but specifies trophoblastic fates when FGF2 is absent. BMP4 induces mesodermal differentiation when WNT activity is high but when it is low trophoblast cells arise. These effects are observed in conventional culture but the addition of BMP4 on micropatterned hESCs gives rise to cells of the three germ layers as well as trophectodermal cells.

Initial experiments revealed a link between the duration of treatment and the derived outcome. Briefly, short exposures (24hrs) of hESCs to BMP4 result in mesodermal

populations which express Brachyury (P. Zhang et al.,2008) while longer exposure (7days) gives rise to trophoblastic cells (R. Xu et al., 2002) or a combination of populations which express trophoblast markers (hCG alpha and beta) as well as markers of extra-embryonic endoderm (AFP, SOX7) (P. Zhang et al.,2008).

BMP4 coordinates fate acquisition in collaboration with other signalling pathways thus creating further complexity. Findings indicate that BMP4 outcome is modulated by MEK/ERK activity and WNT activity. In detail, when FGF2 is present and therefore MEK/ERK is active, BMP enables ME differentiation while in the absence of FGF2 it favours trophoblastic differentiation (P. Yu et al., 2011). Furthermore, endogenous WNT signals in culture are responsible for the diversity observed after BMP4 induction. Briefly, BMP4 induces WNT3 which in turn orchestrates PS and mesodermal differentiation (Kurek et al., 2015). However, upon WNT inhibition, BMP4 directs the expression of trophoblast associated genes (Kurek et al., 2015).

The above findings describe the effect of BMP on conventional 2D cultures. Nevertheless, hESCs can also be cultured on micropatterned surfaces which allow cell growth on a confined circular area (Warmflash et al., 2014) Surprisingly, micropatterned colonies exposed to BMP4 (50ng/ml) generate four distinct populations (Warmflash et al., 2014). Cells from the periphery to the centre of the colony display successively features of the following fates : trophoblast ($CDX2^+$), endoderm ($SOX17^+$), mesoderm ($BRACHYURY^+$) and ectoderm ($SOX2^+$) (Warmflash et al., 2014). This spatial allocation of differentiated cells is reminiscent of developmental events and highlights the great importance of the ligand in fate regulation (Warmflash et al., 2014).

The trophoblast inducing ability of BMP4 has long been considered a paradox, since hESCs are the in vitro counterpart of epiblast and therefore should not be able to differentiate towards extraembryonic tissues (Y. Yang et al., 2015). Recently, scientists reported that short exposure of hESCs to BMP4 combined with ACTIVIN and FGF2 inhibitors can generate an alternative cell state which is considered totipotent (Y. Yang et al., 2015). As a result, cells derived from this treatment are capable of differentiating not only across the three germ layers but also towards trophoblast (Y. Yang et al., 2015). The fact that a well-studied differentiation stimulus expands the developmental potential implies that BMP could have unknown functions in early embryogenesis.

1.4.3 WNT signalling

The canonical WNT pathway contains several intracellular components such as DISHEVELLD (DSH), GLYCOGEN SYNTHASE KINASE 3β (GSK3β), AXIN, ADENOMATOUS POLYPOSIS COLI (APC) and β CATENIN (CTNNB) (Logan & Nusse, 2004). In the basal state, AXIN, APC and GSK3β are associated in a complex which further interacts with CTNNB and targets it for proteasomal degradation (Logan & Nusse, 2004). The receptor complexes on the cell surface are composed of FRIZZLED (FZ) and LOW-DENSITY LIPOPROTEIN-RECEPTOR RELATED PROTEIN (LRP) (Logan & Nusse, 2004). Binding of the ligand to the receptors abolishes CTNNB's degradation (Logan & Nusse, 2004). Under these conditions CTNNB enters the nucleus where, in collaboration with transcription factors such as LYMPHOID ENHANCER-BINDING FACTOR 1/T CELL SPECIFIC

TRANSCRIPTION FACTOR (LEF/TCF) it regulates gene expression (Logan & Nusse, 2004).

The effect of WNT pathway in hESCs is controversial. Pharmacological inhibition of GSK3β and therefore CTNNB rescue from degradation is compatible with the undifferentiated state (Sato et al., 2004). Moreover, a WNT receptor gene (FZD7) is specifically expressed in hESCs rather than differentiated cells (Melchior et al., 2008). In accordance with this, downregulation of the receptor impairs pluripotency associated features (Melchior et al., 2008). However, WNT3a is not sufficient to support pluripotency of hESCs in the absence of feeders and CM (Dravid et al., 2005).

A study utilising WNT reporter hESCs dictates that cells display various degrees of WNT pathway activation (Blauwkamp et al., 2012). When cells are forced to display either high or low levels of activation, preferences in differentiation are observed (Blauwkamp et al., 2012). In detail, high WNT levels favour the formation of endodermal and cardiac cells, while cells with low WNT levels show an inclination towards the neuroectodermal fate (Blauwkamp et al., 2012).

It is generally accepted that the WNT pathway is active when CTNNB regulates transcritpion (Kim et al., 2013). In addition, new findings suggest that WNT pathway can exert dual function in hESCs via the two pools of CTNNB (Kim et al., 2013). Cytoplasmic CTNNB is compatible with self-renewal while nuclear CTNNB initiates differentiation (Kim et al., 2013). This observation may be the one of the reasons causing literature discrepancy (Kim et al., 2013). Another possible explanation for the opposing effects of WNT come from a recent study which suggests that WNT effect is time dependent. Briefly, transient activation enhances the presence of E-

CADHERIN in tight junctions thus promoting PI3K/AKT induced proliferation and sequestering CTNNB away from the nucleus (Huang et al., 2015). On the contrary, when CTNNB levels are too high due to longer activation, only a part of the pool can bind E-CADHERIN while the rest is directed towards the nucleus where it initiates a differentiation programme (Huang et al., 2015).

Under the light of the recently discovered naïve pluripotency, WNT pathway is suggested to be a link between primed and naïve state (Z. Xu et al., 2016). Naïve hESCs have an active WNT pathway compared to primed hESCs which show minimal or no activity (Z. Xu et al., 2016). More importantly when WNT is inhibited (in naïve cells) the differences in the proteome and the metabolome between naïve and primed cells are not that prominent (Z. Xu et al., 2016).

A fascinating aspect of how signalling pathways collaborate in order to shape development comes from a novel publication that identifies for the first time the presence of an organizer in human gastruloids (Martyn et al., 2018). Almost a century ago, Spemann and Mangold discovered a cell population with unique properties, the ‘organizer’ (Spemann & Mangold, 2001). The organizer is characterised by markers such as GOOSECOID (GSC), the ability to dorsalize the ME and orchestrates the formation of the nervous system (Joubin & Stern, 1999). It is located at the one end of the primitive streak and starts functioning at the onset of gastrulation (Joubin & Stern, 1999).

Gastruloids are structures formed when hESCs cultured on micropatterned surfaces are exposed to BMP4 (Warmflash et al., 2014) and they offer a platform for studying development. ACTIVIN, WNT and BMP are three significant pathways in embryogenesis. In mice, ME formation relies on the successive activation of these

pathways; first BMP4 leads to WNT3 expression and WNT3 increases ACTIVIN in the EPI (Ben-Haim et al., 2006). Moreover, studies in several species propose that WNT and NODAL have a critical role in the function of the organizer (Crease, Dyson, & Gurdon, 1998). In accordance with these, human gastruloids display a similar pattern of successive pathway activation (Martyn et al., 2018). Additionally, WNT3 can promote the formation of a primitive streak like region in micropatterned colonies while the combinatorial effect of WNT3 and ACTIVIN generates an area with molecular and functional features of an organizer (Martyn et al., 2018).

1.4.4 FGF signalling

Humans have 22 FGF ligands (Ornitz & Itoh, 2001), FGF2 being widely expressed (Dailey et al., 2005). The receptors have domains with tyrosine kinase activity and are not highly specific for ligand binding (Dailey et al., 2005). The ligand-receptor complex is further stabilised by the presence of heparin or heparin sulfate proteoglycans (Dailey et al., 2005). FGF activates several parallel pathways that seem to regulate each other and ultimately FGF regulates transcription (Dailey et al., 2005).

An indication regarding the importance of FGF2 in hESCs comes from studies which try to optimize media that sustain pluripotency. Since the first derivation of hESCs, Thomson's group has extensively screened growth factors for their ability to maintain hESCs and basic FGF (bFGF or FGF2) was found to be the best candidate (R. Xu et al., 2005). Indeed, the combination of FGF2 with a BMP inhibitor is sufficient to enable hESCs propagation without feeders or CM (R. Xu et al., 2005). Moreover, combined stimulation of hESCs with ACTIVIN A and FGF2 enables long-term

propagation and interestingly enough, the positive effect of FGF2 is exerted through ACTIVIN signalling (Vallier et al., 2005). FGF2 is one of the five essential components which are present in a commercially available pluripotency medium which bypasses the need for feeders, CM and SR (Ludwig *et al.*, 2006).

Further dissecting the role of FGF2 in hESCs reveals that this growth factor can have both direct and indirect effects. FGF2 regulates the expression of several components of the TGFβ pathway in MEFs (Greber, Lehrach, & Adjaye, 2007), thus implying that FGF2 indirectly affects pluripotency in the case of co-culture systems. Undifferentiated hESCs express FGF2 and FGF2 stimulation improves adhesion and moderates cell death during dissociation of colonies in single cells (Eiselleova et al., 2009). Moreover, inhibition of FGF2 results in neuroectodermal differentiation thus suggesting a protective role of FGF2 in maintaining pluripotency (Greber et al., 2011).

The underlying mechanisms enabling FGF2 functions are not well understood. In general, FGF2 regulates several pathways and ERK1/2, p38, PLCγ, PI3K/AKT, and JNK are the most frequently studied (Dailey et al., 2005). In mESCs the MEK/ERK pathway is not compatible with pluripotency (J. Li et al., 2007). However, in hESCs MEK/ERK and PI3/AKT are activated downstream of FGF (J. Li et al., 2007). Both pathways collaborate in supporting pluripotency and PI3K/AKT plays an additional role in enhancing proliferation and survival (J. Li et al., 2007). In accordance with this, when MEK/ERK and PI3K/AKT pathways are perturbed then pluripotency and survival are compromised (Armstrong et al., 2006). Finally, FGF2 seems to maintain NANOG via MEK/ERK (P. Yu et al., 2011), therefore providing an additional link between this growth factor and pluripotency.

1.4.5 IGF, NOTCH and HIPPO signalling

The effects of IGF have not been thoroughly studied in hESCs but it seems that it is associated with self-renewal. When hESCs are stimulated with CM, Insulin Receptor (IR) and Insulin Like Growth Factor Receptor (IGF-1R) become phosphorylated and inhibition of IGF1R results in differentiation (L. Wang et al., 2007). Moreover, IGF-1 and HEREGULIN, which can activate the PI3K/AKT pathway are shown to collaborate with ACTIVIN in maintaining hESCs (Singh et al., 2012).

Likewise, the NOTCH pathway has not drawn much attention in the field of hESCs. The NOTCH pathway seems to be inactive in undifferentiated hESCs but can be activated by exposure to chelating cations such as EDTA (Noggle *et al.*, 2006). Moreover, NOTCH is not necessary for the maintenance of pluripotency (Noggle *et al.*, 2006). However, NOTCH signalling is involved in differentiation. More specifically, when hESCs are treated with NOTCH inhibitors they display a compromised ability to form EBs and this affects the generation of all three embryonic layers (X. Yu et al., 2008).

Similarly, the HIPPO pathway is not extensively studied. Nevertheless, a recent study identified the presence of a complex consisting of SMAD2,3, OCT4 and HIPPO effectors (TAZ/YAP/TEAD) (Beyer, Weiss, et al., 2013). Interestingly, the complex negatively regulates the expression of mesodermal genes while it balances the expression of pluripotency genes (Beyer, Weiss, et al., 2013).

All the major pathways affecting pluripotency and differentiation described above are illustrated in the following figure **(Fig. 1.9)**.

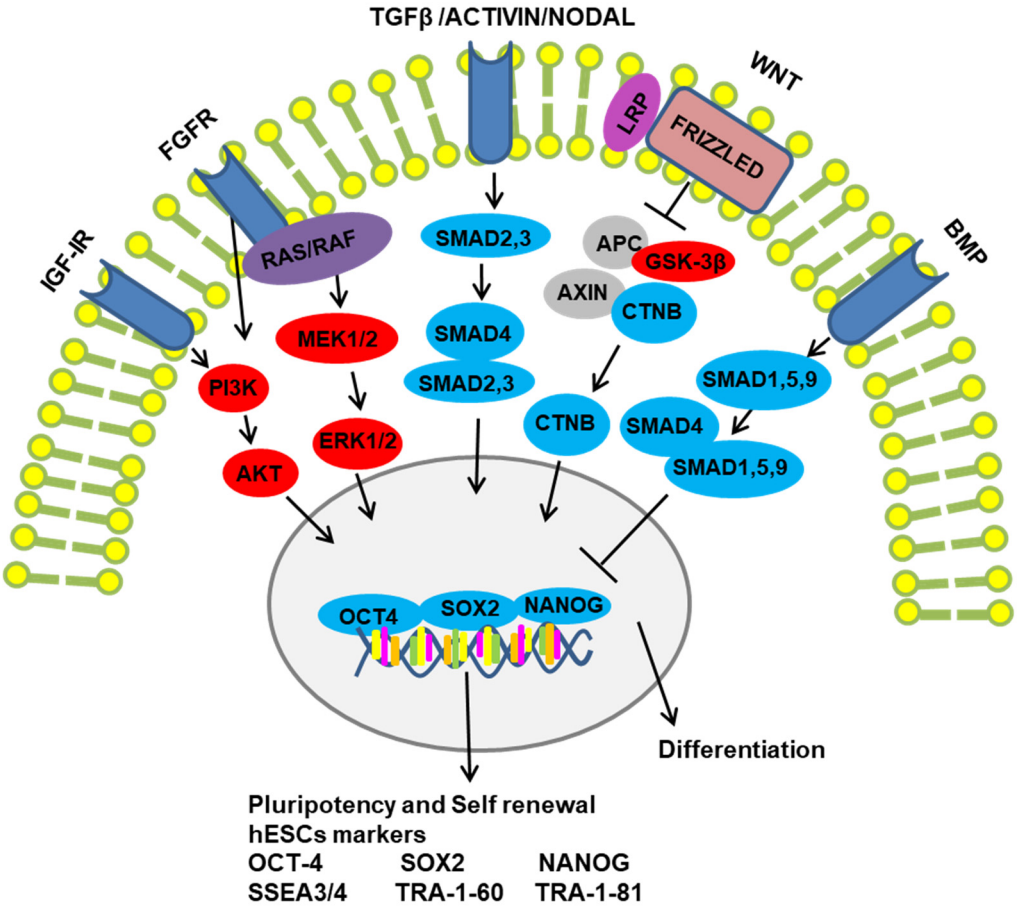


Figure 1-9 Key Signalling Pathways in hESCs.

TGFβ/ACTIVIN/NODAL signalling is established as a crucial pluripotency regulator, since SMAD2,3 are found to maintain NANOG expression. FGF2 is a well-known accompanying factor for ACTIVIN/TGFβ in media that sustain pluripotency. MEK/ERK and PI3K/AKT are possibly involved in FGF's actions. At the same time, IGF can also contribute to pluripotency via PI3K/AKT. BMP is a key differentiation signal that induces SMAD1,5,8 and leads to reduction of NANOG levels. WNT is associated with both pluripotency and differentiation. WNT signalling exerts its functions mainly through CTNNB. When CTNNB is in the cytoplasm self-renewal is maintained, however when CTNNB enter the nucleus it orchestrates alterations in gene expression that initiate differentiation. Adapted from: (1) http://www.cellsignal.com/contents/science-pathway-research-stem-cell-markers/esc-pluripotency-and-differentiation-signaling-pathway/pathways-esc. And from (2) (Zhao & Jin, 2017).

1.5 SARA (ZFYVE9) protein

1.5.1 Introduction in SARA

In 1998, two independent groups published results describing the same new protein and they named it SARA and NSP respectively. Wrana's group identified a SMAD2 interacting protein in Xenopus and named it SARA which stands for SMAD Anchor for Receptor Activation (Tsukazaki et al., 1998). A homolog in humans was also successfully found by the same group (Tsukazaki et al., 1998). Human SARA displays a 62% sequence similarity with Xenopus SARA and consists of several functionally important domains (Tsukazaki et al., 1998). First, SARA can interact with the TGFβ receptor via its C-terminal region (Tsukazaki et al., 1998). Second, a region located in the middle of the protein (665-750AAs) is necessary for binding to the MH2 domain of non-activated SMAD2,3 (Tsukazaki et al., 1998). Last, SARA contains a FYVE domain thus implying that the protein displays membrane anchoring ability (Tsukazaki et al., 1998).

The second group reported the identification of a protein which displayed sequence similarity with serine proteases and therefore named it NSP standing for Novel Serine Protein (Meckelein et al., 1998). Initial experiments revealed that there are several isoforms of the new protein and these could possibly be the result of splicing or cleavage events (Meckelein et al., 1998). Further experiments showed that NSP is expressed by reactive astrocytes, is secreted by smooth muscle cells in the brain and is found together with amyloid in pathologic brains with Alzheimer's disease (Meckelein et al., 1998). All these findings are indicative of a putative role of NSP in the pathogenesis of the disease (Meckelein et al., 1998). However, these findings did

not attract the interest of the following researchers; instead they focused mostly on the signalling aspect of the protein.

1.5.2 SARA in TGFβ/ACTIVIN signalling

The role of SARA in TGFβ signalling is elusive as several studies contradict each other. Wrana's group suggested the regulatory effect of SARA in TGFβ signalling. According to their model **(Fig. 1-10)**, SARA binds non-phosphorylated SMAD2 and presents it to the receptor (Tsukazaki et al., 1998). Subsequent phosphorylation of SMAD2 releases the protein to continue SMAD dependent responses (Tsukazaki et al., 1998). Another FYVE domain containing protein which displays a positive role in signalling is HRS(Miura et al., 2000). HRS, binds SMAD2 and both HRS and SARA facilitate the subsequent binding of the R-SMAD to the receptor (Miura et al., 2000). Interestingly, HRS mutant mice do not survive after day 10.5 and mutant cells show impaired TGFβ/ACTIVIN response which is partly rescued by SARA overexpression (Miura et al., 2000).

The hypothesis that SARA exerts its functions in the endocytic pathway was proposed in 2002. Endogenous SARA resides on the membrane of the EE (Hu et al., 2002). Additionally, overexpressed SARA achieves anchoring on the EE via binding of its FYVE domain with the PI3P lipid (Itoh et al., 2002), (Panopoulou et al., 2002). Proper localization seems to be required for signalling since pharmacological dissociation of SARA from the EE impairs TGFβ response (Itoh et al., 2002). Similarly, ACTIVIN A response is blocked when a SARA mutant that lacks the FYVE domain is expressed in endothelial cells (Panopoulou et al., 2002).

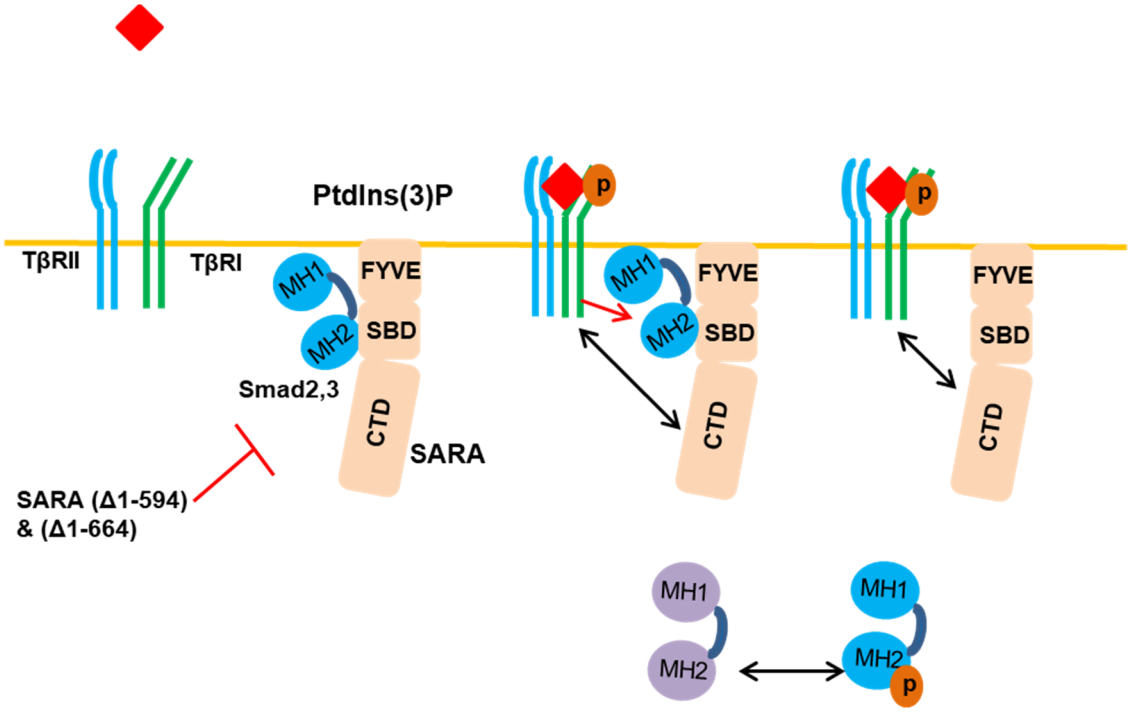


Figure 1-10 Proposed model for the role of SARA in TGFβ signalling.
SARA uses a FYVE domain for membrane anchoring. Unphosphorylated SMAD2 can bind via the MH2 domain to the SBD of SARA. Subsequently, SARA interacts with the TGFβ Type I receptor via the C-terminal region and SMAD2 interacts with the receptor via the MH2 domain. This complex formation facilitates SMAD2 activating phosphorylation. Next, SMAD2 and SMAD4 interact via their MH2 domains and enter the nucleus. SARA mutants (SARA (Δ1-594) and SARA (Δ1-664)) that lack part or the whole FYVE domain fail to properly localize on punctuate structures (early endosomes). As a result, SMAD2 localization is altered and TGFβ signalling is impaired. Adapted from: (Tsukazaki et al., 1998).

Elucidating the role of SARA triggers the long known question regarding the necessity of endocytosis in signalling. In human kidney cells endocytosis seems to define the duration of SMAD2/SARA association (Runyan, *et al*, 2005). In more detail, TGFβ1 stimulation promotes SMAD2/SARA interaction (Runyan, *et al*, 2005).

However, in the absence of endocytosis this interaction remains stable for an extended period thus impairing SMAD2 nuclear transport (Runyan, *et al*, 2005). This paper suggests that signalling initiates at the cell surface where SARA and SMAD2 bind to the receptor and SMAD2 activation occurs (Runyan, *et al*, 2005). Next, phosphorylated SMAD2 and SMAD4 can generate a heterocomplex (Runyan, *et al*, 2005). At this point clathrin mediated endocytosis of the receptor enables SARA's dissociation and allows free movement of SMAD2 to the nucleus (Runyan, *et al*, 2005).

On the other hand, there was a report showing that in Hela cells TGFβ Type I receptor internalises through CME but SMAD2 can be phosphorylated by the receptor directly on the plasma membrane bypassing the need for the EE and SARA (Z. Lu et al., 2002). However, this study can't determine whether there is an additive effect of SARA and EE that could contribute in signalling by affecting other qualitative parameters (Z. Lu et al., 2002). Moreover, in Drosophila, SARA is also responsible for bringing PP1C to the receptor thus negatively affecting signalling (Bennett & Alphey, 2002). A recent study further complicated the picture. In Hela cells, SARA does not interact with SMADs or TGFβ Type I and Type II receptors and SARA siRNA does not affect TGFβ responses (Bakkebø et al., 2012). Researches hypothesize that a number of reasons could account for the observed differences related to key parameters of the experimental setup. First, in the literature there is no unanimity in the ligand used, the concentration or the duration of the induction (Bakkebø et al., 2012). Second, different cells could respond and regulate signalling in different ways (Bakkebø et al., 2012). Third, overexpressing SARA or mutants is

mostly in favour of SARA'S importance while SARA downregulation does not seem to have a marked effect (Bakkebø et al., 2012).

1.5.3 SARA in trafficking

A possible explanation for this could be the fact that overexpressing SARA can alter endosomal function (Bakkebø et al., 2012). Indeed, SARA overexpression results in a significant size increase of EEs, slows down transferrin return to the cell surface thereby decreasing transferrin receptor numbers on the plasma membrane (Hu et al., 2002). This trafficking abnormality is also observed upon overexpression of a constitutively active RAB5 mutant, thus implying that SARA could act downstream of RAB5 (Hu et al., 2002). Several studies investigated in depth the involvement of SARA in trafficking. Interestingly, SARA displays all functional features of HRS. First, SARA identifies cargo tagged for degradation, second it associates with ESCRT-O components (EPS15, STAM2) and third it binds TSG-01 and clathrin (Kostaras et al., 2013). Additionally, SARA binds RNF11 an E-3 ligase which can interact with HRS (Kostaras et al., 2013). Genetic manipulation of SARA and RNF11 levels slows down the degradation of EGFR upon EGF stimulation (Kostaras et al., 2013). It is therefore reasonable to hypothesize that both proteins are part of the endolysosomal pathway and that SARA could possibly be a part of ESCRT-0 complexes (Kostaras et al., 2013).

Membrane trafficking is required for many important processes that take place in neurons (Sann et al., 2009). However, the role of SARA in this type of cells is largely unknown. Endogenous SARA is found in both early and recycling endosomes in rat

hippocampal neurons (Arias, Siri, & Conde, 2015). Perturbing SARA levels alters the localization of recycling endosomes and SARA overexpression impairs the elongation of neurites (Arias et al., 2015). Inversely, minimal to low SARA levels generate elongated axons, increase branching and increase L1 levels on the cell surface (Arias et al., 2015). Since L1 is a cell adhesion molecule, SARA knockdown slows down the movement of neurons in the intermediate zone of mice brain (Mestres et al., 2016). Although, there is not yet a clear mechanism explaining how SARA affects L1, it is possible that in the absence of SARA internalised L1 may follow the fast recycling route to the plasma membrane instead of the $RAB11^{+}$ pathway (Mestres et al., 2016). These results collectively broaden the role of SARA from signalling to trafficking, cellular morphology and developmental processes.

1.5.4 SARA in Epithelial to Mesenchymal Transition

Epithelial to Mesenchymal Transition refers to the conversion of epithelial cells to mesenchymal cells (Kalluri & Neilson, 2003). During this conversion several coordinated events reshape the cells and radically change the interaction of the cells with their neighbouring cells and the surrounding environment (Kalluri & Neilson, 2003). Briefly, tight junctions, adherens junctions and desmosomes are disrupted, polarity is perturbed and cells migrate away from their epithelial niche (Kalluri & Neilson, 2003).

In 2007 and 2008 researchers suggested that there are three types of EMT, which share many similarities but they occur in different contexts and therefore display different features (Kalluri & Weinberg, 2009). Type I EMT forms mesenchymal cells,

Type II results in fibroblasts and Type III generates metastatic tumour cells (Zeisberg & Neilson, 2009). Type I EMT events are often utilized in embryonic development and mesoderm formation is the first observed EMT (Kalluri & Weinberg, 2009). Type II EMT is occurs in adult tissues during fibrosis and is fuelled by ongoing inflammation (Kalluri & Weinberg, 2009). However, validating the emergence of fibroblasts can be tricky since there are no available specific markers (Zeisberg & Neilson, 2009). Type III EMT is considered to be a key stage in epithelial cancers that contributes to a negative prognosis for the patient (Kalluri & Weinberg, 2009). Briefly, epithelial cancer cells which are spatially confined by the boundaries of a basement membrane, enter an EMT which endows them with all the characteristics needed in order to move away from the initial tumour and metastasize (Kalluri & Weinberg, 2009). The phenotypic changes of EMT are reflected at the molecular level thus allowing the establishment of several biomarkers. E-CADHERIN loss is one of the best examples of protein change in the process and more specifically the upregulation of N-CADHERIN and OB-CADHERIN are indicative of Type I/III and Type II EMT respectively (Zeisberg & Neilson, 2009).

Primary epithelial cells grown in culture are long reported to generate fibroblasts (Zeisberg & Neilson, 2009). Some scientists argue that this phenomenon could be induced by the presence of TGFβ in FCS (Zeisberg & Neilson, 2009). Additionally, TGFβ has a defensive role in early stage cancers but in late stage cancers fuels metastasis through EMT (Heldin, Vanlandewijck, & Moustakas, 2012). *In vitro* stimulation of epithelial cells with TGFβ initiates an EMT process which results in a simultaneous decrease of SARA and SMAD2 levels (Runyan et al., 2009). SARA seems to be downregulated at the mRNA level while SMAD2 displays increased

binding with SMURF2 which implies more intense degradation (Runyan et al., 2009). Additionally, in analogy with E-CADHERIN (Zeisberg & Neilson, 2009), SARA loss triggers EMT-like changes thus indicating a protective role of the protein for the epithelial identity (Runyan et al., 2009).

Furthermore, SARA overexpression during high glucose induced EMT is able to reverse the process by positively affecting SMAD2 and changing the kinetics of pSMAD2 and pSMAD3 (Tang et al., 2015). PI3K is also proposed to exert a protective role in fibrosis by regulating SARA levels (Runyan et al., 2012). Briefly, pharmacological inhibition of PI3K, increases the size of endosomes, captivates SARA on them and leads SARA to proteasomal degradation (Runyan, *et al*, 2012). SARA loss is also observed when a constitutively active RAB5 mutant is expressed (Runyan et al., 2012). Although, these findings directly link SARA with PI3K it is unclear whether the effect on SARA levels is specific for PI3K inhibition or whether it is a more general effect linked to alterations in endosomal morphology (Runyan, *et al*, 2012).

1.5.5 SARA in Asymmetric Cell Division

Cell division should ensure that the activity of pathways will be equal in the resulting daughter cells, as small changes can have a dramatic impact during development (Knoblich, 2006). SARA positive endosomes seem to deliver such a task in drosophila developing wing, by interacting with the spindle and segregating equally in the daughter cells (Bökel et al., 2006). These endosomes carry DPP (which is the TGFβ homolog) and its Type I receptor called THICKVEINS (Bökel et al., 2006).

Therefore, the newly formed cells display similar levels of pathway activity (Bökel et al., 2006).

In contrast to the above, SARA endosomes are also utilized across mitosis in order to facilitate unequal signalling activation between daughter cells (Coumailleau et al., 2009). Drosophila Sensory Organ Precursor (SOP) cells divide asymmetrically producing a pIIa and a pIIb cell, which are differentially regulated by NOTCH signalling (Coumailleau et al., 2009). SARA's contribution relies on the fact that SARA endosomes positive for Delta and Notch are segregated to the pIIa cell during mitosis (Coumailleau et al., 2009). Initial experiments implicated the asymmetry of the mitotic spindle for this outcome (Derivery et al., 2015). An unprecedented role of SARA in this process has been recently discovered. Three phosphorylation sites of SARA (S636, S709, S774) determine whether SARA endosomes are driven towards the spindle or detach from it (Loubéry et al., 2017). When SARA is not phosphorylated, spindle targeting is favoured while upon SARA phosphorylation endosomes disengage and segregate unequally in the arising cells (Loubéry et al., 2017).

SARA endosomes positive for NOTCH and DELTA are also asymmetrically distributed during the asymmetric division of Drosophila Intestinal Stem Cells (Montagne & Gonzalez-Gaitan, 2014). In addition, a similar process is observed in zebrafish spinal cord; where again SARA endosomes currying DELTAD preferentially move towards the one daughter cell that will acquire a different fate compared to the other (Kressmann et al., 2015). These findings link endosomes in general and SARA in particular with fate acquisition. It will be interesting to know if SARA can also display similar functions in human cells.

1.6 Thesis Aims and Objectives

The previous parts of this chapter provide the necessary information to understand the scientific framework for this Thesis. This section describes the questions that will be addressed in the current study. The main aim of this Thesis is to study in depth the various effects of TGFβ superfamily members on hESCs. In particular this Thesis addresses the following:

(1) Identification of the phosphoproteome following BMP4 addition to hESCs as the cells differentiate.

(2) The role of SARA, a key regulator of TGFβ/ACTIVIN A signalling and trafficking, in hESCs.

1.6.1 Chapter 3: Introduction to H1 cell culture, differentiation and adaptation to ACTIVIN A medium.

The discovery of hESCs generated a new field of research aiming to study their ability to either self-renew or differentiate. These decisions can be influenced by a plethora of signals and cell fate can be dictated exogenously. Several commercial media enable the propagation of undifferentiated hESCs using combinations of growth factors and most of them contain either ACTIVIN A or TGFβ. Given the prominent importance of ACTIVIN A and BMP4 for hESCs, this chapter attempts to: (1) study the effect of BMP4 on hESCs and explore its ability to induce embryonic and extra-embryonic differentiation (2) Establish an ACTIVIN A-dependent medium for hESCs which allows signalling experiments in pluripotency and differentiation to be performed.

1.6.2 Chapter 4: Quantitative Phosphoproteomics (SILAC)

Signalling pathways operate as a hierarchy of events where one extracellular signal is transmitted intracellularly affecting several effectors (Invergo & Beltrao, 2018). Protein phosphorylation is a key modification frequently used by growth factors in signalling. Revealing phosphorylation changes as a response to a growth factor gives an insight in the signalling employed by the growth factor. Having established in chapter 3 a culture of hESCs cells grown in ACTIVIN A, this chapter attempts to: (1) use SILAC-MS to reveal phosphorylation events in early stage BMP4-induced differentiation, (2) study the phosphorylation status of proteins associated with pluripotency as well as proteins involved in Wnt signalling, (3) validate BMP4-induced targets obtained from the SILAC-MS experiments by utilising phos-tag gels.

1.6.3 Chapter 5: Role of SARA in human embryonic stem cells

Although SARA was discovered two decades ago (Tsukazaki et al., 1998), (Meckelein et al., 1998) several important questions regarding this protein are not yet answered. SARA is known to be implicated in TGFβ signalling (Tsukazaki et al., 1998) and EMT (Runyan et al., 2009) but there is no evidence so far linking SARA with hESCs. This chapter attempts to: (1) explore the presence of SARA isoforms, (2) test the hypothesis that SARA translocates to the nucleus, (3) elucidate the role of SARA in mesodermal differentiation and reveal whether SARA is differentially regulated in EMT and MET, (4) study the hypothesis that SARA undergoes lysosomal degradation.

2. MATERIALS AND METHODS

2.1 Cell Culture and Reagents

2.1.1 PSC culture

Matrigel adapted H1 cells (Thomson et al., 1998) (WiCell, Lot Number WB 0113), HUES1 cells (Cowan et al., 2004)(Harvard HUES cell facility) and hiPSC FORTHi001C (line3 clone5) (Kyrkou et al., 2016) were cultured in mTeSR1 medium (Stem Cell Technologies, 5851) on Matrigel coated dishes (Corning, 354277). For the purposes of SILAC, H1 cells were cultured in mTeSR without select factors (Stem Cell Technologies, 05896) supplemented with 1mM LiCl (Sigma-Aldrich, 62476), 1mM γ-AminoButyric Acid (Sigma-Aldrich, A5835), 0.1mM Pipecolic Acid (Sigma, P2519), 100ng/mL bFGF (ImmunoTools, 11343627) and 0,5ng/mL Activin A (kindly provided by Dr Marko Hyvönen, Cambridge University).

For matrigel coating, one vial (200μl) of aliquoted frozen matrigel (Corning, 354277) was placed on ice and slowly diluted in 18 ml DMEM-F12 (ThermoFisher Scientific, 10770245). The solution was added to 6-well plates (1,5ml/well of a 6well) and left at room temperature for 45min to solidify. The excess coating solution was removed, dishes were washed with DMEM-F12 and fresh media was added.

Differentiated cells were removed with aspiration during culturing and prior to passaging. Medium was changed daily and cells were passaged every 5-7 days in a ratio 1 to 6 to 1 to 8. Briefly, cells were washed once with DMEM-F12, 1ml of dispase (1mg/ml) (Gibco, 17105-041) was added to a well of a 6well plate and incubated for 5min at 37^{o}C. At the end of the incubation colonies should slightly detach from the plate. Dispase was aspirated and cells were gently washed twice with DMEM-F12.

Finally fresh media were added and cells were removed with gentle scraping (H1, HUES1) or manually dissected with an insulin syringe (hiPSCs, H1 in ACTIVIN A).

For the freezing procedure, subconfluent cells were scrapped or manually dissected as described above and then centrifuged at 900rpm for 3 min. The supernatant was discarded and the pellet was resuspended in 1ml mFresR (Stem Cell Technologies, 5855) and transferred into labelled cryo-vials. Vials were kept overnight at -80^{o}C and transferred the next day to liquid nitrogen.

For the thawing procedure, cryo-vials were placed in a waterbath, cells were gently transferred to a 15ml falcon tube, 10ml of fresh medium supplemented with 5μM ROCKi (Fasudil, LC Laboratories, F-4660) was added slowly and then they were centrifuged at 900rpm for 3min. Supernatant was discarded and pellet was resuspended in 3ml of fresh medium supplemeted with 5μM ROCKi and transferred to 6-well plates.

Differentiation to mesoderm was induced with BMP4 (Gibco, PHC9534). BMP4 powder was reconstituted in 4Mm HCL/0.1% HSA (100ng/μl) and stored at -80^{o}C. TGFβ/ACTIVIN A signalling was abolished with SB431542 (Sellechem, S1067). SB431542 powder was reconstituted in DMSO (Sigma-Aldrich, D2650) to a concentration of 20mM and stored at -80^{o}C. Lysosomal degradation was inhibited with chloroquine diphosphate crystalline (Sigma-Aldrich, C-6628). Chloroquine powder was reconstituted in tissue culture water to a concentration of 50mM and stored at room temperature. Leptomycin B (stock: 1mM in ethanol, kindly provided by Dr. Saverio Brogna) was used in order to block putative SARA nuclear export to a final concentration of 20μM.

2.1.2 HEK-293A Culture

HEK-293A (Invitrogen, R705-07) were cultured in 10cm dishes in DMEM (high glucose, pyruvate, Gibco, 41966), supplemented with 10% fetal bovine serum (FBS) (ThermoFisher Scientific – 10270106) and 100 U/mL Penicillin/Streptomycin (P/S) (ThermoFisher Scientific – 15140122). Cells were passaged every 3 days (1/5 ratio). Briefly, medium was aspirated; cells were washed once with phosphate buffered saline (PBS) (Sigma Aldrich, D8537) and then incubated with Trypsin in the incubator. Trypsinised cells were resuspended in fresh medium and transferred to new dishes.

For the freezing procedure cells were trypsinised as described above and then centrifuged at 900rpm for 5 min. The supernatant was discarded and the pellet was resuspended in 1 ml 90% FBS-10% DMSO (Sigma-Aldrich, D2650). Cells were transferred to cryo-vials which were kept overnight at -80^{o}C and next day were stored in liquid nitrogen.

For the thawing procedure cryo-vials were placed in a waterbath, cells were gently transferred to a 15ml falcon tube, 10 ml of fresh medium was added slowly and they were centrifuged at 1100 rpm for 3 min. Supernatant was discarded and pellet was resuspended in 10 of fresh medium and transferred to new dishes.

2.1.3 Miscellaneous Cell Lines

MCF7, Hela and HaCaT cells were cultured in DMEM (high glucose, pyruvate, Gibco, 41966), supplemented with 10% FBS (ThermoFisher Scientific, 10270106) and 100 U/ml P/S (ThermoFisher Scientific, 15140122). Human fibroblasts (ATCC, CRL 2429)

were cultured in IMDM (30–2005) 10%FBS. Thawing, passaging and freezing procedures were performed as described for HEK-293A cells.

2.1.4 Plasmid and siRNA transfection

HEK-293A cells were passaged as previously described. For DNA transfection, cells were cultured until they reached 50-60% confluence. X-tremeGENE™ 9 DNA Transfection Reagent (Roche, XTG9-RO) was kept for 15 min out of the fridge to reach RT. Next, 3μl of lipid were diluted in 100ml Opti MEM™ I Reduced Serum Medium (ThermoFisher Scientific, 31985047) and then 1μg of DNA was added. The mixture was incubated for 15 min at RT and then added dropwise on a well of a 6-well plate. Next day media were changed and transfection efficiency was assessed 36 hrs later.

For siRNA transfection, HEK-293A cells were cultured until they reached 50-60% confluence. RNAimax was kept out of the fridge for 15 min to reach room temperature. Then two separate eppendorfs were prepared, the one containing 4,5μl lipid diluted in 75μl Opti MEM and the other containing 20pmol siRNA diluted in 75μl Opti MEM. The siRNA mixture was added in the lipid mixture, incubated together for 5min and added dropwise on cells. The medium was changed the next day and transfection efficiency was assessed 72hrs later. In this Thesis, two siRNAs were used, one targeting SARA (Ambion, AM16708) and a scrambled (negative control) siRNA (Ambion, AM4642).

Upon reaching confluence, H1 cells were washed once with 1ml Versene-EDTA 0.02% (Lonza, BE17-711E) on a well of a 6well plate. Cells were incubated in Versene-EDTA for 5min until cells looked slightly more spherical and colonies started

detaching. Versene was aspirated and cells were thoroughly pipetted in 3ml fresh media supplemeted with 5μM ROCKi until they become small cell clumps (3-5 cell/clump). Cells were added to new matrigel-coated dishes. Next day medium was aspirated, cells were incubated in PBS for 5min, PBS was aspirated and 500μl of fresh medium was added. In parallel, the transfection mixture was prepared as described above and added dropwise to the cells. For the generation of stable cell lines, the day following transfection, cells were passaged from a 12well plate to a 6 well plate. One day after the second passaging, cells were treated with 250ng/mL puromycin (Sigma-Aldrich, P7255) until the control population died. At the end of the puromycin selection, antibiotic resistant colonies were picked using cloning rings and further cultured for total cell lysate and RNA extraction.

2.2 Bacterial Transformation and CRSPR Cloning

2.2.1 Bacterial transformation

For plasmid amplification, competent DH1OB E.Coli bacteria (generated, tested and stored at -80^{o}C) were thawed on ice for 20min. Next, 10ng of DNA were mixed with 100μl of competent cells on ice for 20min. The mixture was heat shocked for 4min at 42^{o}C. Then 1ml of antibiotic free LB medium was added and cells were allowed to grow in a shaking incubator at 37^{o}C for 45min. 100μl of the mixture were placed on 10cm LB agar plates with the respective antibiotic resistance. Control mixture (without DNA) was also added in parallel to a separate plate. LB plates were incubated at 37^{o}C overnight. Next day, 1 isolated colony was picked and placed in a 15ml falcon tube, supplemented with 3ml LB media containing antibiotic and allowed

to grow overnight with agitation. Next day, 850µl of the overnight culture were mixed with 150µl 87% glycerol and placed in labelled cryovials at -80^{o}C were they were kept as glycerol stocks. Moreover, 100µl of the overnight culture were used to inoculate 200mls of LB supplemented with antibiotic and allowed to grow overnight with agitation at 37^{o}C. Plasmids were purified using either QIAGEN PLASMID MIDI KIT (112143) according to the manufacturer's protocol. DNA concentration was measured using Nanodrop (ND-1000 V3.8.1) and run on an agarose gel to check quality.

2.2.2 Targeting exon 3 and exon 4 of SARA

Specific guides targeting the CDS of SARA in exon3 and exon4 were designed according to the instructions of: SANGER CRISPR (Hodgkins et al., 2015). Each guide consists of a set of primers, which are annealed in the following reaction: Reaction Mixture: 3µg Forward Primer, 3 µg Reverse Primer, 1X Annealing Buffer (Total Volume 50µl). Annealing Buffer: 1M NaCl, 500Mm HEPES, p.H: 7.4. Annealing Reaction: 4 min at 90^{0} C, 10 min at 70^{0} C, 20 min at 37^{0} C, 10 min at 37^{0} C. Next, px459 vector (Ran et al., 2013) (Addgene, 62988) was linearized with the enzyme BbsI and ligated with the annealing product. Ligation product was used to transform E.Coli competent cells (DH10B). 5 colonies were picked and grown for extraction of plasmid DNA which was then sequenced with the U6 primer to verify incorporation of the guide RNA in the plasmid (U6 primer: GATACAAGGCTGTTAGAGAGAT).

2.2.3 Targeting the Transcription Start Site of SARA

Two specific guides targeting the TSS of SARA were designed using the MIT CRISPR FINDER (http://crispr.mit.edu:8079/). Two vectors, one for each guide (2641, 2642) were generated according to the instructions above. Once they were sequenced, vector 2642 was used as a template in order to PCR out the sequence containing the U6-gRNA. (Forward Primer: TCTAGAGGTACCGAGGGCCTATTTC, Reverse Primer: TCTAGAAAAAAAGCACCGACTCG). The PCR product was ligated in zero-blunt vector and transformed in competent E.Coli cells (2644). Next, both 2641 and 2644 were digested with XbaI, the correct DNA fragments were gel extracted and ligated in a final vector which contained both guides.

2.3 SDS PAGE Electrophoresis and Western Blot

2.3.1 Conventional SDS-PAGE

Cells were washed twice with cold PBS and then lysed in cold PBS based buffer containing 1% SDS (Roth, CN30.3) and 1mM PMSF (Thermo Fisher Scientific, 215740100). Lysates were sonicated twice for 10sec, boiled for 10min, centrifuged at 13,200 rpm for 15min and supernatant was collected. Protein concentration was determined using a BCA kit (Thermo Fisher Scientific, 23225). Samples were separated by electrophoresis on 8, 10 or 12%, SDS- polyacrylamide gel gels, depending on the size of the protein of interest, and transferred to nitrocellulose or PVDF membrane. PVDF membranes were activated in methanol for 25sec before the transfer. Membranes were blocked for 1hr in 5% milk powder in TBS-T and then incubated overnight at 4^0C with primary antibodies **(Table 2-2)**. The next day,

membranes were washed 3X10 min in TBS-T then incubated with the secondary antibody **(Table 2-3)** at room temperature for 3 hrs. Membranes were washed 3X10 min in TBS-T. For protein visualisation membranes were incubated for 1 minute with Amersham ECL Western Blotting Detection Reagent (Amersham, RPN2209), exposed to Fuji Super Rx X-ray film and developed using a Curix-60 processor (AGFA HealthCare), or they were visualised using Odyssey® CLx Li-COR Infrared Imaging system.

2.3.2 Phos-Tag SDS PAGE

Cells were washed twice in ice cold buffer containing 10 mM Tris–HCl pH 7.5 (Thermo Fisher Scientific, 10376743) and 0.10 M NaCl (Thermo Fisher Scientific, S/3160/65). Cells were lysed in ice cold buffer containing 50 mM Tris–HCl pH 7.5, 150mM NaCl, 0.25% sodium deoxycholate, 1% Igepal (Sigma-Aldrich, I3021) supplemented with complete mini EDTA free protease inhibitors (Roche, 4906837001) and phosphatase inhibitors (Roche, 4906837001). For the IP experiments cells were lysed in the above mentioned buffer without sodium deoxycholate. Cells were scraped off the plate and transferred to an eppendorf. Lysates were triturated for 15-20 times using an insulin syringe. Lysates were incubated on ice for 15min, centrifuged at 14,000 rpm for 15min and supernatant was collected.

Lysates were run on the gel **(Table 2-1)** at 160V for as long it was appropriate according to the molecular weight of the protein of interest. Prestained molecular weight markers were not used to minimise band distortion. Next, gels were immersed twice in transfer buffer supplemented with EDTA (1mM) for 10min and then

immersed in transfer buffer for 20min. Gels were transferred to PVDF membrane and transfer was performed at 250mA for the amount of time that was appropriate according to the molecular weight of the protein of interest.

Membranes were immersed in methanol for 25sec, air dried, immersed in methanol for another 25sec and transferred to TBS-T. Next, they were incubated with 5% milk powder in TBS-T for 1hr and subsequently with primary antibody **(Table 2-2)** overnight at 4^{o}C. The next day, membranes were washed 3X10 min in TBS-T and then incubated with the secondary antibody **(Table 2-3)** at room temperature for 3 hrs. Membranes were washed 3X10 min in TBS-T and visualised using Odyssey® CLx Li-COR Infrared Imaging system.

The 3X loading buffer was made of 200 mM Tris, pH 6.7, 6% SDS, 15% glycerol and stored at -20^{o}C, while β-mercaptoethanol (Sigma-Aldrich, M3148) was added prior to use (1/10 of final volume). 5X Running buffer was made of: 0.5 M Tris, 0.5 M MOPS (VWR, A1076, 0250), 0.5% SDS, pH 7.8, while for 1X Running buffer fresh sodium metabisulfate (fisher scientific, 10233970) was added to a final concentration of 5mM. 10X Transfer buffer was made of 250 mM Tris and 1.92 M glycine, while 1X Transfer buffer was supplemented with methanol (10% in the final volume). 20X TBS was made of 200 mM Tris-HCl, 3 M NaCl, pH 7.6, while TBS-T was supplemented with Tween-20 (Acros Organics, 233360010) to a final concentration of 0.1%.

Table 2-1 Table lists the reagents used for the preparation of the stacking and separating gel in phos-tag experiments. Information is given regarding company, catalogue number, concentration and quantity of the reagents used.

Reagent	Separating gel 10%	Stacking gel
Acrylamide (30% Solution at 37.5:1 Ratio)	3.25ml	0.75ml
Bis-Tris-HCl buffer, pH 6.8 Stock:1.4M Bis-tris (VWR, 12A1025,0100)	2.5ml	1.25ml
$ZnCl_2$ Stock: 10mM Acros Organics # 318140100	50µl	N/A
Phos-tag (5mM aqueous solution) Alpha labs, AAL-107 5mM	50µl	N/A
TEMED Sigma-Aldrich, T9281	10µl	5µl
APS 10% Acros Organics, 327081000	100µl	200µl
Millipore water	Up to 10ml	Up to 5ml

2.4 Immunofluorescence (IF)

H1 cells were cultured on ibidi dishes (Thistle Scientific, IB81151) and HEK-293A cells cells were cultured on polylysine coated coverslips. Cells were washed twice with PBS, fixed in 3.7% PFA for 15 min and washed again in PBS for 5min. Next, cells were quenched in 50mM ammonium chloride (Sigma-Aldrich, A4514) for 15

min, permeabilized in 0.5% Triton X-100 (Acros Organics, BP151-500) for 4 min, blocked in 10% FCS for 20 min, incubated for 1 hour with primary antibodies **(Table 2-2)** at room temperature, followed by 3x5 min washes. Next, samples were incubated with secondary antibodies **(Table 2-3)** for 45 min at room temperature, followed by a second round of 3x5min washes. Both the primary and secondary antibodies were diluted in 10% FCS. For nuclear staining **(Table 2-2)**, samples were incubated for 15min with YOYO (Invitrogen, Y3601) or Hoechst 33342 (Cell Signalling Technology, 4082s) diluted 1 in 10,000 in 10% FCS. Next, samples were washed in PBS and mounted in Prolong® Diamond Antifade Mountant (ThermoFisher Scientific, P36970). Samples were imaged using either Nikon Eclipse Ti Inverted Microscope and NIS-Elements Advanced Research Imaging Software, or Leica TCS SP8 and LASAF Imaging Software.

Table 2-2 Table lists the name of the primary antibodies and nuclear dyes used in this report, accompanied by information regarding species, company name, catalogue number and dilution factor.

Primary Antibody/Nuclear Staining	Company/Catalogue Number	Host Antibody type Isotype	Dilution
Anti-human Nanog	Cell Signaling Technology/3580S	Rabbit Polyclonal	WB: 1/1000 IF: 1/100
Anti-human Nanog	Cell Signaling Technology/4893S	Mouse Monoclonal	WB: 1/1000 IF: 1/100

		IgG1	
Anti-human Oct-3/4 (C-10)	Santa Cruz Biotechnology/ sc-5279	Mouse Monoclonal IgG2b	IF: 1/200
Anti-Human/Mouse/Rat SOX2 Antibody	R&D Systems/ MAB2018	Mouse Monoclonal IgG_{2a}	IF: 1/200
Anti-Amphibian, Avian, Fish, Human, Lizard, Mouse, Opossum, Planaria, Rat, Turtle, Zebrafish PAX6	DSHB/ N/A	Mouse Monoclonal MIgG1, kappa light chain	IF: 1/200
Anti-Human/Mouse BRACHYURY	R&D Systems / AF2085	Goat Polyclonal IgG	WB: 1/2000 IF: 1/500
Anti-Human/Rat SARA	Abcam/ ab124875	Rabbit Monoclonal IgG	WB: 1/1000 IF: 1/200
Anti-Human/Mouse/Rat Phospho-Smad2 (Ser465/467)	Cell Signaling Technology/3101S	Rabbit Polyclonal	WB: 1/2000
Anti-Human/Mouse/Rat/Mink Phospho-Smad1 (Ser463/465)/	Cell Signaling Technology/9511S	Rabbit Polyclonal	WB: 1/2000

Smad5 (Ser463/465)/ Smad8 (Ser426/428) Antibody			
Anti-Human/Mouse/Rat/Dog/Chicken β-Catenin	BD/610153	Mouse Monoclonal IgG1	WB: 1/2000
Anti-Human/Mouse E-Cadherin	Novus Biologicals/ AF648-SP	Goat Polyclonal	WB: 1/2000
Anti-MYC	Homemade	NA	IF: 1/50
Anti-FLAG	Sigma Aldrich/ F3165	Mouse Monoclonal IgG1	WB: 1/40.000 IF: 1/4000
Anti-Human/Mouse/Rat TCEA1	Abcam/ ab185947	Rabbit Monoclonal IgG	WB: 1/1000
Anti-Human PEPP2	Abcam/ ab122300	Rabbit Polyclonal	WB: 1/1000
Anti-Human/Mouse MAP4K6	Abcam/ ab86385	Rabbit Polyclonal	WB: 1/1000
Anti-Human/Mouse/Rat BMP2K	Abcam/ ab191532	Rabbit Polyclonal	WB: 1/500
Anti-GFP	Sigma/ G1544	Rabbit Polyclonal	WB: 1/4000

Anti-Rabbit IgG	Upstate/12-370	Rabbit Polyclonal IgG	IP: 1µg/300 µg lysate
Anti-Human/Rat/Dog/Chicken EEA1	BD Biosciences/610456	Mouse Monoclonal IgG1	IF: 1/200
Anti-ACTIN clone4 (All species)	EMD Millipore/ MAB1501	Mouse Monoclonal IgG2BK	WB: 1/100.000
HOECHST 33342 Working Stock	Cell Signaling Technology/4082S	NA	IF: 1/10.000
YOYO	Invitrogen/ Y3601	NA	IF: 1/10.000
Propidium Iodide (PI)	NA	NA	IF: 1/5.000

Table 2-3 Table lists the name of the secondary antibodies used in this report, accompanied by information regarding species, company name, catalogue number and dilution factor.

Secondary Antibody	Company	Catalogue Number	Dilution
Peroxidase-AffiniPure Rabbit Anti-Goat IgG (H+L)	Sigma-Aldrich	A8919	WB: 1/3000
Peroxidase-AffiniPure Goat Anti-Mouse IgG	Jackson Immunoresearch	115-035-062	WB: 1/5000

Peroxidase-AffiniPure Goat Anti-Rabbit IgG (H+L)	Jackson Immunoresearch	111-035-144	WB: 1/5000
IRDye® 800CW Donkey anti-Rabbit IgG (H + L)	LI-COR	926-32213	WB: 1/5000
IRDye® 680RD Goat anti-Mouse IgG (H + L)	LI-COR	926-68070	WB: 1/5000
IRDye® 800CW Donkey anti-Goat IgG (H + L)	LI-COR	925-32214	WB: 1/5000
IRDye® 800CW Donkey anti-Mouse IgG (H + L)	LI-COR	925-32212	WB: 1/5000
Alexa Fluor® 488-AffiniPure Donkey Anti-Rabbit IgG (H+L)	Jackson ImmunoResearch	711-545-152	IF: 1/200
FITC)-AffiniPure Donkey Anti-Mouse IgG (H+L)	Jackson ImmunoResearch	715-095-151	IF: 1/200
Alexa Fluor® 594-AffiniPure Donkey Anti-Goat IgG (H+L)	Jackson ImmunoResearch	705-585-147	IF: 1/200
Alexa Fluor® 594-AffiniPure Donkey Anti-Mouse IgG (H+L)	Jackson ImmunoResearch	715-585-151	IF: 1/200

2.5 Quantitative Real-Time PCR

Cells were washed twice with PBS and then lysed in Lysis Buffer (provided in the kit) supplemented with β-mercaptoethanol (55mM, Invitrogen, 21985023). Total RNA was extracted according to the manufacturer's instructions (Macherey Nagel, 119-22402). Sample concentrations were measured using Nanodrop (ND-1000 V3.8.1) and samples were aliquoted and stored at -80^{0}C. Samples for the reaction were

prepared using the appropriate pairs of primers **(Table 2-4)** to a final concentration of 3μM and the QuantiTect SYBR® Green RT-PCR Kit (Qiagen, 204243). The reaction was performed in an AriaMx RealTime PCR System, the data were analysed using AriaMx software and CT values were normalised against GAPDH using the equation: 2^-ΔCt.

Table 2-4 Table lists the name and sequence of primers used in this report.

Primer Description	Primer Sequence
SARA FW	TGAACAAAACGAAGATGAAACAGT
SARA RV	CACACTGGCCAAAGTAGGGT
MSX2 FW	ATGGAGCGGCGTGGATG
MSX2 RV	GCGAGGAGCTGGGATGTG
SNAIL1 FW	ATGCCGCGCTCTTTCCTC
SNAIL1 RV	GGTGGGCCTGGTCGTAG
TWIST W	CAGCGGCGGCGGGAGTC
TWIST V	CGAGGGCAGCGTGGGGATG
GATA6 FW	ACGCCGCCTTCCCCCATCTCT
GATA6 RV	CCCCAGGCGCCGAAGGTC
MIXL1 FW	GAACAGGCGTGCCAAGTCTCG
MIXL1 RV	TTCGGGCAGGCAGTTCACATCTAC
AFP FW	ACACAAAAAGCCCACTCCAGCATC
AFP RV	GTCATAGCGAGCAGCCCAAAGAAG
HAND1 FW	GCCACCAGCTACATCGCCTACCT
HAND1 RV	GCCGCCATCCGCCTTCTTGA

WNT3 FW	CTGGCGAAGGCTGGAAGTGG
WNT3 RV	CGGCCCGCCTCGTTGTTG
CDX2 FW	CACGCAGCCCCGCAGACTACCATC
CDX2 RV	CTTCCGCATCCACTCGCACAGG
GAPDH FW	CGCGCCCCCGGTTTCTAT
GAPDH RV	CCTTCCCCATGGTGTCTGAGC
GAPDH FW	CTCTGCTCCTCCTGTTCGAC
GAPDH RV	ACCAAATCCGTTGACTCCGA

2.6 Karyotype

Subconfluent H1 and HUES1 cells were treated with 0,04ug/ml Karyomax colcemid (10ug/ml, Gibco, 15212) in fresh mTeSR for 2hrs at 37^{o}C. Cells were trypsinized to single cells and the cell pellet was resuspended in hypotonic KCL (0,56% w/v) for 30 min at 37^{o}C. Cells were pelleted again and resuspended in a low volume of KCL (approximately 2ml) and then fixed dropwise with ice cold mixture of Methanol/Acetic Acid (3:1). Fixed cells were kept at -20^{0}C overnight and processed for Giemsa staining.

2.7 Teratoma Assay

Subconfluent H1 and HUES1 cells were manually dissected with an Insulin needle and digested with dispase (1mg/ml). The cell clumps were centrifuged and then resuspended in 100µl matrigel. The mixture was injected through a 25G 7/8 needle (BD, 305124) subcutaneously into the hind legs of immune deficient NOD/SCID mice.

Mice were maintained under specific pathogen free (SPF) conditions, at the animal house of the Biomedical Research Foundation (Academy of Athens, Greece). Animals were sacrificed 10 weeks later and the resulting tumours were dissected and processed for imunohistochemistry.

2.8 Quantitative Phosphoproteomics

2.8.1 Cell culture for SILAC

For labelling, H1 cells were cultured in customized mTesR1 without Lysine and Arginine and the medium was supplemented with standard Lys and Arg or heavy labelled forms of Lys and Arg in their original concentration (0.548 mM Arg, 0.391 mM Lys). Labelled amino acids **(Table 2-5)** were kindly provided by Dr. Debbie Cunningham. The medium was supplemented with Proline (500mg/L), in order to avoid a common source of artefact in SILAC experiments caused by conversion of Arginine to Proline.

Table 2-5 Table lists information about labelled aminoacids used for SILAC, including their final concentration in the medium.

Amino acids (MW)	Concentration In mTeSR1	Mass difference	Catalogue Number
L-Lysine 2HCl (223.13) (4,4,5,5-D4)	0.391 mM	4 Da	DLM-2640 (Cambridge)
L-Lysine 2HCl (227.1) (13C6, 15N2)	0.391 mM	8 Da	CNLM-291-H (Cambridge)

L-Lysine	0.391 mM	0 Da	L8662, Sigma
L-Arginine HCl (216.62) (13C6)	0.548 mM	6 Da	CLM-2265-H (Cambridge)
L-Arginine HCl (220.59) (13C6, 15N4)	0.548 mM	10 Da	CNLM-539-H (Cambridge)
L-Arginine	0.548 mM	0 Da	A8094, Sigma
L-Proline	500mg/L	Non applicable	P5607, Sigma

2.8.2 In gel Digestion

40µg of total cell lysate was run on a 12% polyacrylamide gel, the gel was then fixed for 1 hr in 30% Methanol-10% Acetic Acid, stained for 30 min in 0.1% Coomassie Brilliant Blue R-250 (Fischer, BP101-25, in 50% Methanol-10% Acetic Acid) and then destained overnight in 20% Methanol-10% Acetic Acid. Bands were excised with sterile scalpels and further destained in 30% acetonitrile (JT Baker, 9821) for 15 min with agitation and exposed to 50% acetonitrile, 25mM ammonium bicarbonate (Fisher, A/5160/50) for 15 min. This step was repeated until bands were destained. The gel pieces were dehydrated in a vacuum centrifuge for 5 min. Samples were reduced in 10mM DTT (Sigma-Aldrich, D9779), 25 mM ammonium bicarbonate at 56^0C for 45min and then alkylated in 55 mM iodacetamide (Sigma-Aldrich, 91707) and 25 mM ammonium bicarbonate in the dark for 45 min. Then samples were washed with 25 mM ammonium bicarbonate (1 X 10 min with agitation) and 50%

acetonitrile, 25 mM ammonium bicarbonate (2 X 5 min). Next they were dehydrated in vacuum centrifuge for 5 min. Samples were trypsinised overnight in the dark at 37^0C (Promega, V5280). Next day peptides were extracted with 0.5% formic acid (Fisher, A117-50) and remaining peptides were extracted with 50% acetonitrile and 100% acetonitrile. These were pooled together and dried down with vacuum centrifugation. Samples were resuspended in 0.1% Trifluoroacetic acid (TFA), desalted using zip tips (185096), dried down in speed vac, resuspended in 0.1% formic acid and injected in the M/S Orbitrap.

2.8.3 Lysate Preparation for MS

Cells were washed twice with ice cold PBS and lysed with Urea 8M (Sigma-Aldrich, U5128), prepared in HEPES 20mM (Sigma-Aldrich, H4034) supplemented with a cocktail of protease and phosphatase inhibitors. Cells were scraped of the plate, sonicated 3 times for 10sec in ice, with 30 sec intervals between sonication. Next, they were centrifuged for 10min at 4^oC at 14,000 rpm, supernatant was collected and protein concentration was measured.

2.8.4 Trypsin Digestion

Initially, equal amounts of light, medium and heavy lysates were mixed for every biological replicate. Ammonium bicarbonate 0.5M was added to a final concentration of 50mM. Samples were reduced in 8mM DTT, incubated at 56^0C for 45 min, alkylated with 20mM Iodacetamide and transferred to dark for 45 min. In the next step, samples were supplemented with 50mM ammonium bicarbonate to a final

volume of 50ml and then trypsinised (Promega, V5280) overnight at 37^0C with a ratio of 1 µg of enzyme per 100 µg of protein. The next day, TFA was added to a final concentration of 0.5% in order to stop the reaction.

2.8.5 Desalting using SEP-PAK

The peptides that resulted from the trypsin digestion were subjected to a desalting process with the help of SEP-PAK C18 Cartridges. Initially, the cartridge was washed with 100% Acetonitrile (ACN), conditioned with 50% ACN/ 0.5% Acetic Acid and equilibrated with 0.1% TFA. Next, samples were loaded (in 0.4% TFA) and the cartridge was washed first with 0.1% TFA and then with 0.5% Acetic Acid in order to remove TFA. At the end the desalted samples were eluted in 50% ACN/ 0.5% Acetic Acid and dried with the help of vacuum centrifugation.

2.8.6 Strong Cation Exchange Chromatography (SCX).

Dried samples were resuspended in 100µl of mobile phase A solution (10 mM KH3PO4, 20 % acetonitrile, pH 3) and injected with a blunt end Hamilton syringe into the HPLC machine. Eluted peptides (33 fractions) were collected in eppendorfs, dried down and stored at -20^{o}C.

2.8.7 Macrotrap for desalting and concentrating peptides.

The obtained 33 fractions of each biological replicate were combined according to the profile of the SCX in 20 fractions that should contain an approximately similar amount of eluted phosphopeptides. For the next step a C8 Macrotrap cartridge was

initially regenerated with 70% Formic Acid/30% Isopropanol, then washed with 500 µl of 90%ACN/0.1%TFA and equilibrated with 500µl of 2% ACN/0.1 %TFA. Next, each sample was loaded in 100µl of 2% ACN/0.1% TFA, the cartridge was washed again in 2% ACN/0.1% TFA and the desalted peptides were eluted in 300µl of 90% ACN/0.1% TFA. Peptides were dried down and stored at -20^{o}C.

2.8.8 Phosphopeptide Enrichment using Titanium Dioxide.

Peptides were enriched for phosphopeptides using the Titansphere Phos-Tio kit (Hinchrom, 5010-21310). Tips were conditioned in 20µl Buffer A (0.5% TFA, 80% ACN) and then equilibrated in 20µl Buffer B (0.38% TFA, 60% ACN, 25% lactic acid). Each fraction was resuspended in 50µl Buffer B (0.38% TFA, 60% ACN, 25% lactic acid) and loaded twice on the tip. Next the tips were washed with Buffer B once and Buffer A twice. Finally, phosphopeptides were eluted in 5% Ammonia and 5% Pyrrolidine, dried down and stored at -20^{o}C.

2.8.9 Desalting Using Zip Tips.

Samples were resuspended in 10µl 0.1% TFA and then desalted using Zip Tip C18 (Millipore, ZTC 18S 096). Tips were washed twice with 100% ACN and equilibrated with 0.1% TFA. Samples were loaded on the tip, the tips were washed twice with 0.1% TFA and the samples were eluted 10 times in 10µl of 0.1% Formic Acid/ 50% ACN. Finally, samples were dried down, resuspended in 20µl 0.1% Formic Acid and divided equally in eppendorfs that would serve as technical replicates.

3. INTRODUCTION TO H1 CELL CULTURE, DIFFERENTIATION AND ADAPTATION TO ACTIVIN A MEDIUM

3.1 Introduction

Thomson's group first reported the derivation of hESCs from the ICM of human blastocysts in 1998 (Thomson et al., 1998). The group suggested a number of characteristics that should define the identity of hESCs and one of them was the ability to differentiate across the three germ layers (Thomson et al., 1998). Indeed, the first cell lines were capable of producing differentiated progeny both *in vitro* and *in vivo* upon injection in immunocompromised mice (Thomson et al., 1998). As a result, differentiation protocols and teratoma formation assays have been used as a prevalent way to show that the differentiation potential of PSCs is intact.

NANOG, OCT4 and SOX2 are believed to be the main orchestrators of mESC pluripotency (Silva & Smith, 2008). The same transcription factors are also involved in the maintenance of hESCs by blocking differentiation (Z. Wang et al., 2012). Therefore, their presence in culture is considered to be indicative of the undifferentiated state.

Despite the initial observations that hESCs could be cultured for several passages without acquiring karyotypic abnormalities (Thomson et al., 1998), through the years several studies revealed the tendency of this type of cells to present chromosomal alterations after prolonged passaging (Baker et al., 2007), (Maitra et al., 2005), (Spits et al., 2008). The prevalence of genetic changes can offer growth advantage in culture, thus benefiting the abnormal cells (Baker et al., 2007). A large screening of more than 100 hESC lines and 11 hiPSCs revealed that although the majority of cell lines are not affected, there is propensity for changes in 4 chromosomes (International Stem Cell Initiative, 2011). Giemsa staining is a typical method used to screen metaphasic hESCs

for the morphology and number of their chromosomes (Spits et al., 2008). Nevertheless, one of its limitations is the inability to reveal small changes which can be studied using other methods (Spits et al., 2008).

TGFβ/ACTIVIN A and BMP4 play a prominent role in hESCs' biology. TGFβ/ACTIVIN A combined with FGF are supportive of the undifferentiated state (Vallier et al., 2005) and they are also components in numerous commercially available pluripotency media (Desai *et al*, 2015). On the contrary, BMP4 confers upon hESCs the acquisition of mesodermal fate (P. Zhang et al., 2008) as well as extraembryonic fates (R. Xu et al., 2002), (P. Zhang et al., 2008). In mice embryos, T (Brachyury) gene is expressed by the mesoderm and the cells that will form the definitive endoderm at the onset of gastrulation (Wilkinson *et al*, 1990). Brachyury means short tail and was first described in 1927 when researchers studied heterozygous mice displaying this phenotype (P. Yu et al., 2011). This gene is important for differentiation since Brachyury mutants show impaired mesoderm formation ability (Wilkinson *et al*, 1990).

Although, TGFβ/ACTIVIN, FGF, WNT and BMP pathway are implicated in mesoderm formation, BMP4 seems to be the best factor in promoting Brachyury expression in hESCs (P. Zhang et al.,2008). In addition to Brachyury, BMP4 is reported to increase the expression of Wnt3 and Mixl1 genes, which are expressed in the PS, ME and endoderm (P. Zhang et al.,2008). Moreover, longer BMP4 treatments lead to the expression of Cdx2 an early TE marker and Gata6 an early extra-embryonic endoderm marker (P. Zhang et al.,2008). This is followed by the expression of the late extra-embryonic endoderm markers Sox7 and Afp and the expression of hCGa and hCGb which are tropboblast markers (P. Zhang et al.,2008). A similar report demonstrates that in the presence of chemically defined medium, BMP4 upregulates the expression of PE

genes and TE genes (Vallier *et al.*, 2009). Gata4, Gata6 and Sox7 belong to the former group while Cdx2, Eomesodermin and Hand1 belong to the latter (Vallier *et al.*, 2009). Despite the fact that Cdx2 is typically related to trophoblast, recent studies suggest that BMP4 generates populations which express both BRACHYURY and CDX2 (Bernardo et al., 2011), (Warmflash et al., 2014). It seems that in conventional cultures CDX2 acts downstream of BRACHYURY and they are both involved in mesoderm formation (Bernardo et al., 2011). However, in micropatterned cultures upon BMP4 stimulation, CDX2 expression is found in BRACHYURY positive and negative cells (Warmflash et al., 2014). This finding implies that the former cells are mesodermal while the latter are extra-embryonic populations (Warmflash et al., 2014).

Furthermore, hESCs respond to BMP4 by an increase in the expression of Msx2 and Slug while the start moving away from the colony (Richter et al., 2014). These findings suggest that mesodermal commitment reflects EMT (Richter et al., 2014). Interestingly, Msx2 is regulated earlier than Brachyury upon BMP4 stimulation and seems to regulate ME commitment by altering SOX2 and NODAL pathway activity (Q. Wu et al., 2015).

The aims of this chapter are the following: First to introduce the culture system routinely used for hESCs in this Thesis. Second, to dissect the effect of BMP4 on gene expression, protein expression and morphology of hESCs. Third, to establish a culture using a slightly modified version of mTeSR1 in which ACTIVIN A substitutes TGFβ.

3.2 Results

3.2.1 hESCs maintain the expression of pluripotency factors and normal karyotype in culture.

H1 cells were maintained in culture for several passages under serum free and feeder free conditions. During this period there were no visible signs of excessive differentiation and cells were tested for the expression of pluripotency factors as well as for the presence of chromosomal abnormalities. Briefly, cells were seeded on ibidi dishes and when they reached subconfluence they were fixed in 3.7% PFA and stained with antibodies against NANOG, OCT4 and SOX2, while the nuclei were stained with HOECHST. All colonies were positive for the abovementioned markers **(Fig. 3-1)** indicating the maintenance of the undifferentiated state. Additionally, H1 and HUES1 cells were cultured in 60mm dishes and they were harvested for karyotype check before reaching confluence. Giemsa staining and karyotype check was performed by Dr. Marika Syrrou (Associate professor in General Biology-Medical Genetics, Medical School, University of Ioannina). At least 20 metaphases were screened for every sample and both cell lines were found to display normal karyotype **(Fig. 3-2)**.

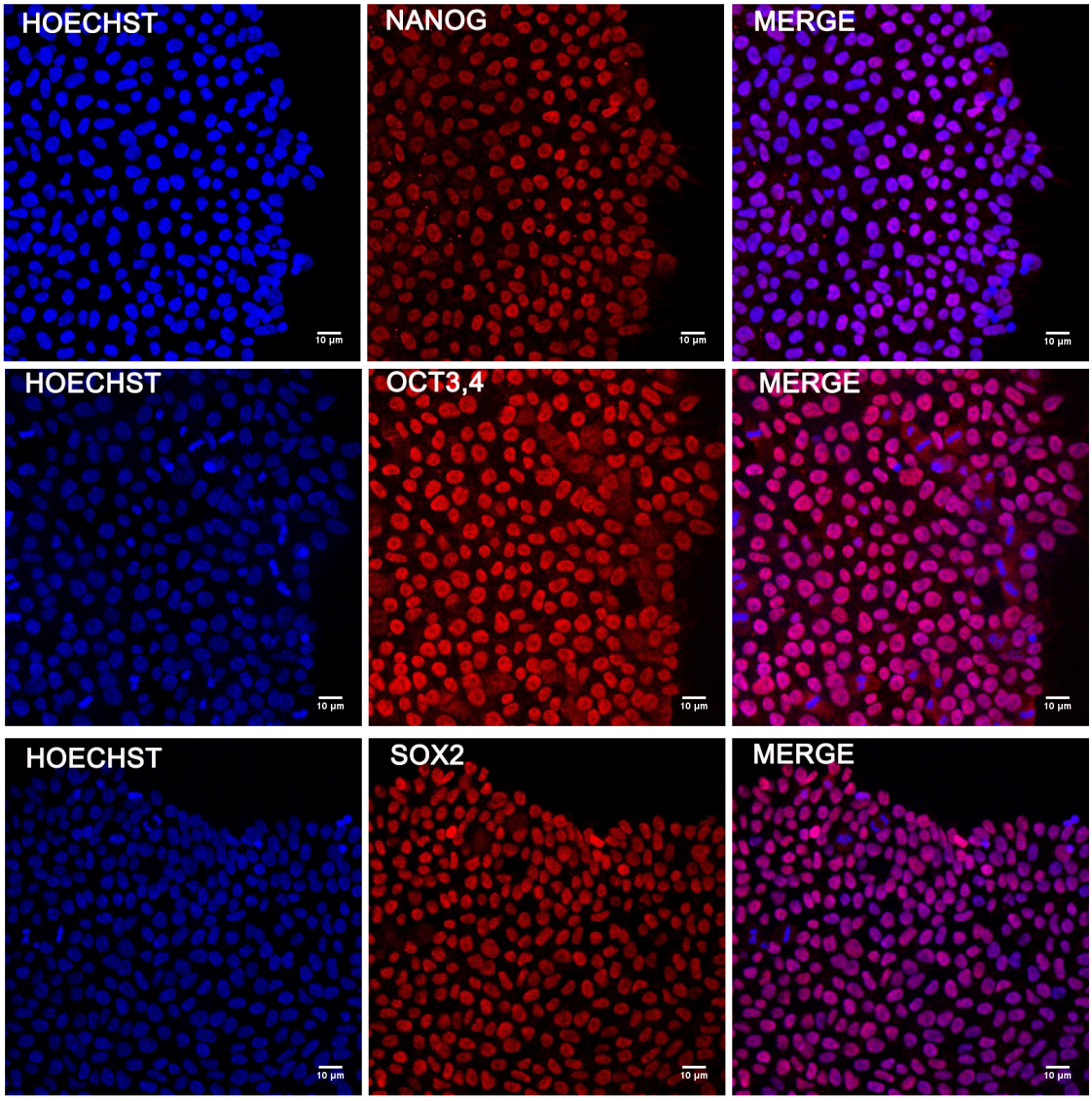


Figure 3-1 H1 cells express pluripotency markers.

H1 cells were seeded on ibidi dishes and before reaching confluence they were fixed in 3.7% PFA and stained for NANOG, OCT4 and SOX2 while nuclei were stained with HOECHST. Images were taken using an SP8 Leica microscope, 40X objective with oil. Images are Z projections and the scale bar is set at 10μm. The left panel is nuclear staining, the middle panel is the transcription factor of interest and the right panel is the merge of the previous two panels.

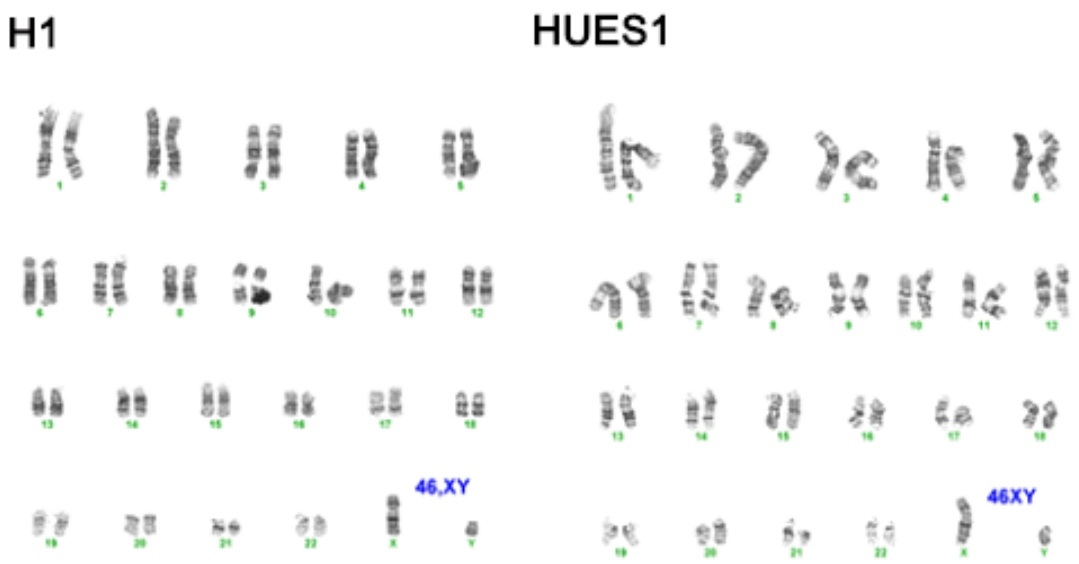


Figure 3-2 H1 and HUES1 cells maintain normal karyotype.

H1 and HUES1 cells were cultured on 60mm dishes and before reaching confluence they were incubated with Karyomax for 2hrs at 37^{o}C. At the end of the incubation period, cells were trypsinised to single cells and then incubated in hypotonic KCL for 30min. Next, cells were pelleted and resuspended in KCL and fixed in methanol/acetic acid. Giemsa staining (at least 20 metaphases were screened) revealed no obvious alterations in chromosomal number or morphology. The left figure demonstrates Giemsa staining in H1 cells, while the right figure demonstrates Giemsa staining in HUES1 cells.

3.2.2 BMP4 pulses induce BRACHYURY expression

Given the significance of BMP4 in differentiation; it seems plausible to explore the effect of this stimulus in hESCs cultures. For this reason, we decided to test the effect of short and long BMP4 pulses in the expression of BRACHYURY in H1 cells. In detail, H1 cells were seeded on 6-well plates and two days after passaging they were stimulated for 1,

3, 6, 24 or 48hrs with BMP4 (50ng/ml), while control cells were not induced. In cells treated for less than 48hrs, the medium was removed at the end of the time point, cells were washed once with DMEM-F12 and then they were cultured for the rest of the experiment in mTeSR1 medium. Cells were lysed in 1% SDS 48 hrs after the initial induction; lysates were run on an SDS-PAGE gel and immunoblotted against BRACHYURY and ACTIN.

According to the results, BRACHYURY was found to be increasingly expressed as the induction period increased. In fact, the difference between time points and the control becomes statistically significant at 6hrs of induction and stays significant for 24hrs and 48hrs **(Fig. 3-3)**. This gradual increase indicates possible association between the duration of the induction and the efficiency of differentiation. Moreover, data suggest that even short treatments are sufficient to induce BRACHYURY expression thus implying that BMP4 triggers the process but does not have to be constantly present in order for differentiation to occur.

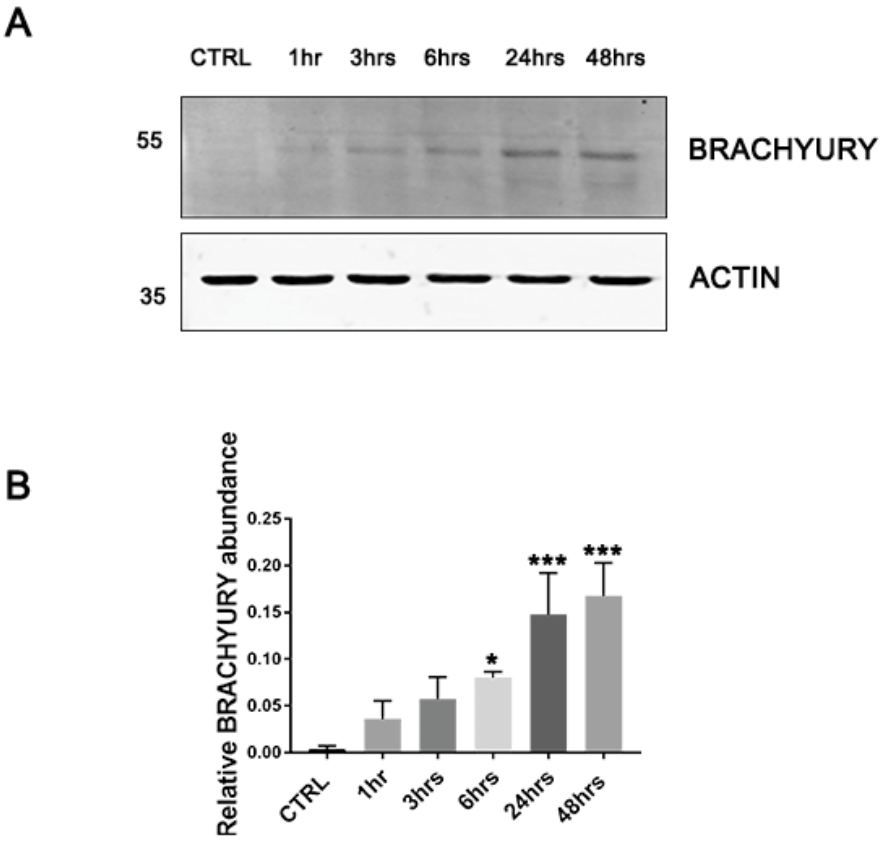


Figure 3-3 BMP4 pulses induce BRACHYURY.

H1 cells were induced with BMP4 for 1, 3, 6, 24 and 48hrs, while control cells were left untreated. For inductions shorter than 48hrs, media were removed at the end of the time point; cells were washed once in DMEM-F12 and then cultured in mTeSR1 for the rest of the study. Cell lysates were extracted in 1% SDS lysis buffer 48hrs after the induction. Lysates were run on an SDS-PAGE gel and immunoblotted against BRACHYURY and ACTIN. **A.** Western blot showing BRACHYURY expression in control H1 cells and cells treated for various time intervals with BMP4. **B.** Plot shows normalised BRACHYURY expression. Statistical significance was tested using Sidak's test in Graph Pad (N=3, *: p-value<0.05, **:p-value<0.01, ***:p-value<0.001).

3.2.3 BMP4 induces morphological changes in H1 colonies and the expression of differentiation associated genes.

The next step was to observe changes in morphology induced by BMP4. For this reason, H1 cells were seeded on ibidi dishes and two days after passaging they were incubated with BMP4 (50ng/ml) for 48hrs. At the end of the induction, cells were fixed in 3.7% PFA and stained with E-CADHERIN while nuclei were stained with HOECHST. The IF showed that the pattern of E-CADHERIN expression changes in differentiated cells as colonies start losing their shape and cells become larger, more loosely attached and migrate away from their initial niche **(Fig. 3-4)**. These signs are not observed in control (untreated) cells and could indicate EMT.

In order to address the changes in gene expression H1 cells were seeded on 6-well plates and stimulated two days after passaging with BMP4 (50ng/ml). 48hrs later RNA was extracted and used to check the levels of expression of several genes associated with differentiation. The tested genes can be categorised in two groups; the first group consists of PS and mesodermal genes (Mixl1, Wnt and Msx2) and there was observed a statistically significant upregulation of Msx2 **(Fig. 3-5A).** The second group contains TE genes (Cdx2, Hand1) and Extraembryonic Endoderm genes (Gata6, Sox7, Afp). According to the data, BMP4 induction leads to a significant rise in Hand1 expression **(Fig. 3-5B)**. These findings are in agreement with previous publications which suggest that BMP4 can induce both mesodermal and extra-embryonic markers in hESCs.

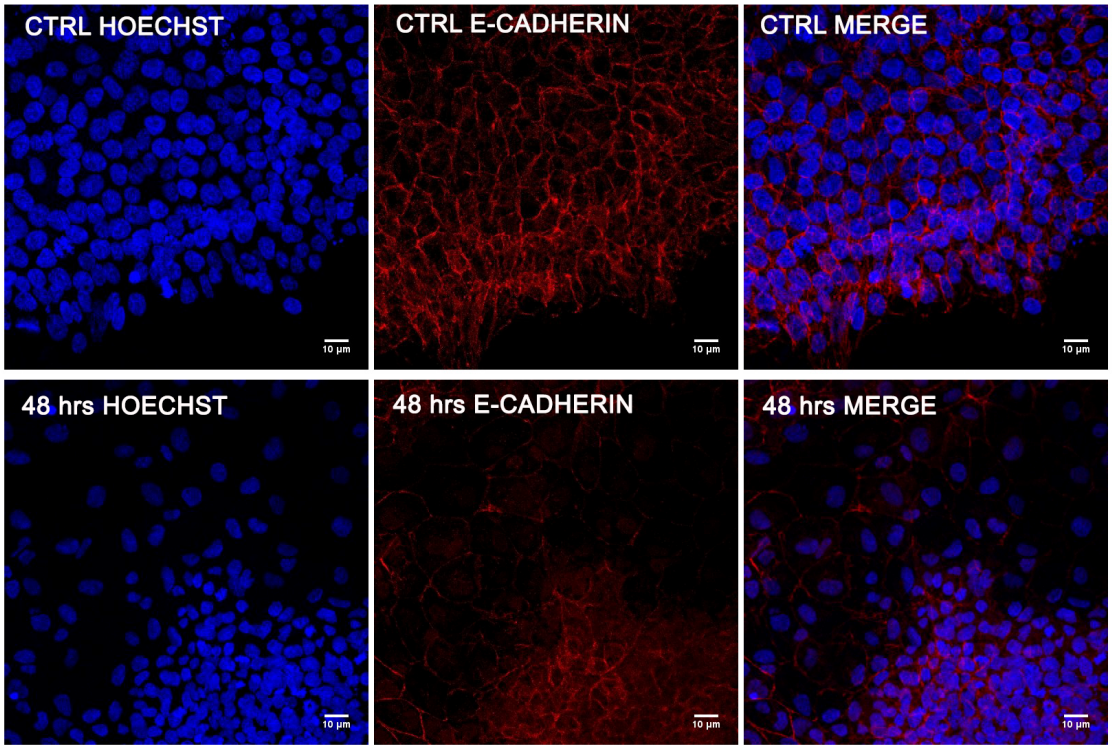


Figure 3-4 BMP4 pulses alter colony morphology.

H1 cells were seeded on Ibidi dishes and two days after passaging they were induced with BMP4 (50ng/ml). After 48hrs cells were fixed in 3.7% PFA and stained against E-CADHERIN while nuclei were stained with HOECHST. Images were taken using an SP8 Leica microscope, 40X objective with oil. Images are Z projections and the scale bar is set at 10µm. The left panel is HOECHST, the middle panel is E-CADHERIN and the right panel is the merge of the previous two.

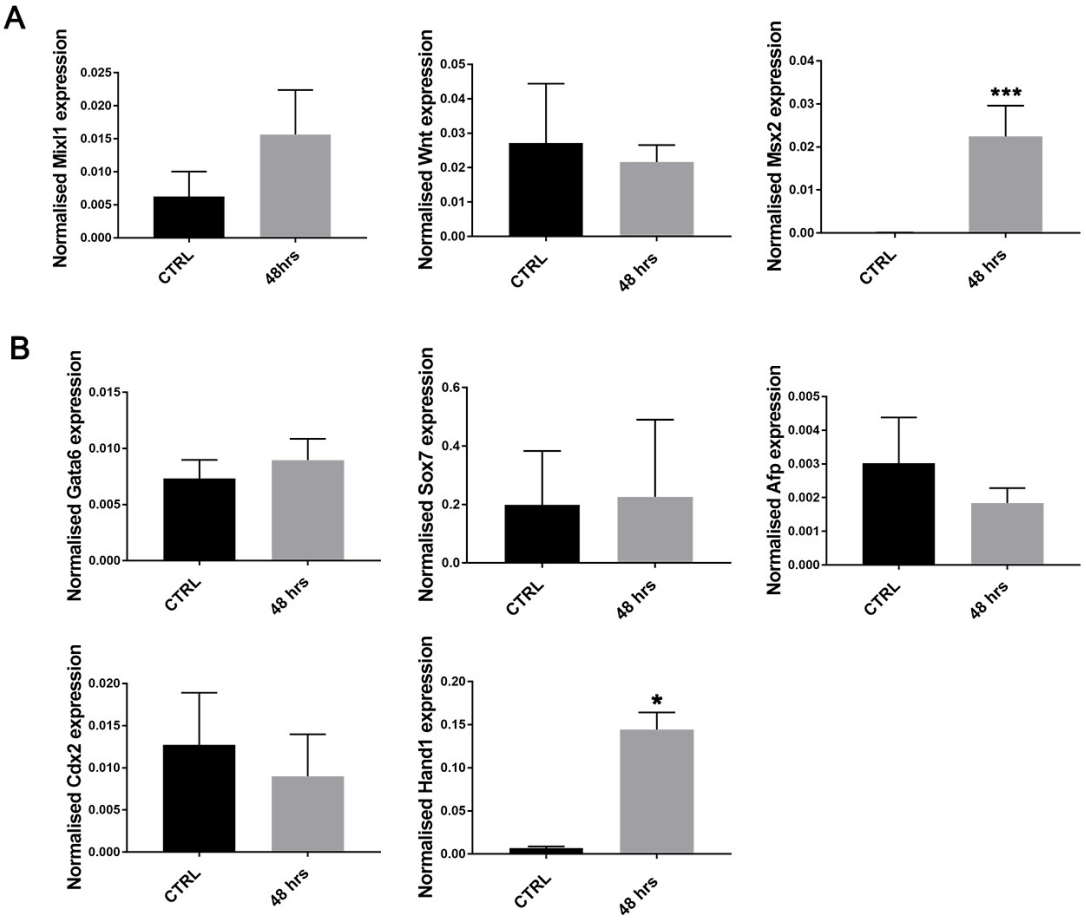


Figure 3-5 Changes in gene expression caused by BMP4.

H1 cells were seeded on 6-well plates and two days after passaging they were induced with BMP4 (50ng/ml). 48 hrs later, RNA was extracted and qRT-PCRs were performed in an AriaMx RealTime PCR System. **A.** Relative expression of genes associated with primitive streak and mesodermal differentiation in control H1 cells and H1 cells induced with BMP4 for 48hrs. **B.** Relative expression of genes associated with differentiation towards trophectoderm and extraembryonic endoderm in control H1 cells and H1 .cells induced with BMP4 for 48hrs Data were analysed using AriaMx software and CT values were normalised against GAPDH. Statistical significance was tested using unpaired t-test in Graph Pad (N=3, *: p-value<0.05 and ***: p-value<0.01).

3.2.4 H1 cells can be maintained in an ACTIVIN A medium.

In this thesis H1 cells are routinely cultured in a commercial medium (mTeSR1) which sustains pluripotency through the action of two growth factors, TGFβ (0.5ng/ml) and FGF2 (100ng/ml) and the accompanying effect of three more select factors, namely lithium chloride (LiCl), γ-aminobuteric acid (GABA) and pipecolic acid. However, in order to elucidate the effect of Activin A in pluripotency, hESC should be studied in a feeder and serum free system where the undifferentiated state is supported by ACTIVIN A.

For this reason, an alternative version of this medium, devoid of the abovementioned factors, (mTeSR, Stem Cell Technologies, # 05896) was purchased. Next, LiCl, GABA and pipecolic acid were exogenously added in their original concentrations as mentioned in the manufacturer's protocol (Ludwig et al., 2006). In this context, various combinations of ACTIVIN A (synthesized and kindly provided by Dr Marko Hyvonen, Cambridge University) and FGF2 were added and then tested for their ability to sustain undifferentiated morphology in H1 cells **(Table 3-1)**. Cells were checked daily for two passages until it was obvious that the combination of Activin A 0.5ng/ml and FGF2 100ng/ml was the most efficient. At the same time control H1 cells were grown in the same medium with TGFβ (0.6 ng/ml) and various concentrations of FGF2 in order to ensure that neither FGF2 nor the three exogenously added select factors affected cell morphology.

Table 3-1 Table shows combinations of different concentrations of ACTIVIN A and FGF2 tested in H1 cells.

Activin A 0.5 ng/ml FGF2 10 ng/ml	Activin A 0.5 ng/ml FGF2 50 ng/ml	Activin A 0.5 ng/ml FGF2 100ng/ml
Activin A 1 ng/ml FGF2 10 ng/ml	Activin A 1 ng/ml FGF2 50 ng/ml	Activin A 1 ng/ml FGF2 100 ng/ml
TGFβ1 0.6 ng/ml FGF2 10 ng/ml	TGFβ1 0.6 ng/ml FGF2 50 ng/ml	TGFβ1 0.6 ng/ml FGF2 100 ng/ml

Although, H1 cells were propagated without signs of excessive differentiation, further criteria should be met in order to prove that they remain pluripotent. In the first step, H1 cells were checked for the expression of pluripotency factors by IF. After three passages in the new medium, H1 cells were seeded on Ibidi dishes and before reaching confluence they were fixed in 3.7% PFA and stained with antibodies against NANOG, OCT4 and SOX2. Microscopy findings **(Fig. 3-6)** revealed that these factors were still expressed thus implying that the new medium is not causing loss of pluripotency.

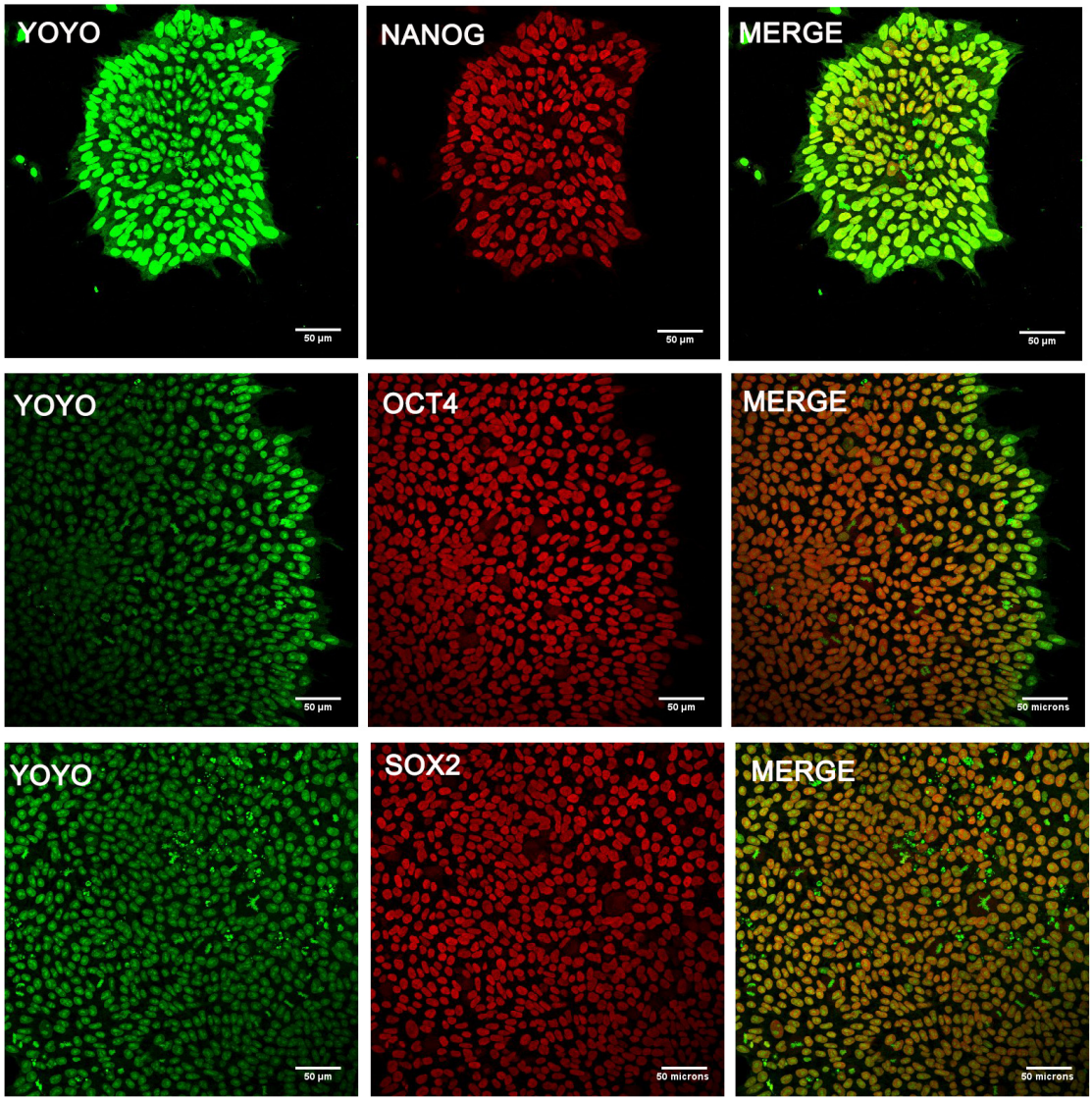


Figure 3-6 H1 in ACTIVIN A maintain NANOG, OCT4 and SOX2 expression.
H1 cells cultured in ACTIVIN A were seeded on Ibidi dishes and before reaching confluence they were fixed in 3.7% PFA and stained against NANOG, OCT4 and SOX2. Nuclei were stained with YOYO and images were taken using an SP5 Leica microscope, 40X objective with oil. Images are Z projections and the scale bar is set at 50μm. The left panel is nuclear staining with YOYO, the middle panel is the transcription factor of interest and the right panel is the merge of the previous two panels.

Moreover, further *in vitro* and *in vivo* experiments were used in order to address their capacity to produce derivatives of the three germ layers. For the *in vitro* experiments H1

cells cultured in TGFβ and ACTIVIN A were subjected to three differentiation protocols. In detail, for the mesodermal differentiation, H1 cells were seeded on 6 well plates and 3 days after passaging the medium was changed to RPMI, 1% glutamax, 1% non essential amino-acids (NEEAs), 1% P/S, 1% insulin-transferrin-selenium (ITS), 1X β-mercaptoethanol supplemented with Activin A 50ng/ml and BMP4 50ng/ml (P. Zhang et al.,2008). Cells were lysed 48 hrs later in 1% SDS, run on and SDS-PAGE gel and immuno-blotted against BRACHYURY and actin **(Fig. 3-7-A)**. According to the Western blot, BRACHYURY is induced in both cultures thus proving the mesodermal producing ability.

For the endodermal differentiation, H1 cells were seeded on 6 well plates and 3 days after passaging the medium was changed to RPMI, 1% glutamax, 1% P/S, 0.5% FBS, Activin A 100ng/ml. One day after the initial induction the medium was renewed and 3 days after the induction it was changed again and serum concentration was increased to 2%. Five days following induction, total RNA was extracted for RT-PCR (Agarwal et al., 2008). The results show a marked increase in the expresison of Sox17 **(Fig. 3-7-B)**, which is an endodermal marker.

Additionally, for the ectodermal differentiation, H1 cells were seeded on non adherent plates in a medium consisting of DMEM F-12, 20% KSR, 1X NEEAs, Glutamine (2mM), 0.1 mM β-mercaptoethanol. Suspension conditions enabled the folding of the colonies and the subsequent formation embryoid bodies. On the second day, medium was changed and on the 4^{th} day the cell aggregates were transferred to a serum free medium containing DMEM F-12, 1X NEEAs, 2μg/μl Heparin. Finally on the 7^{th} day embryoid bodies were transferred to laminin coated ibidi dishes and on the 10^{th} day they

were fixed for IF (Yoo et al., 2011). The staining revealed Pax6 positive cells **(Fig. 3-7-C)**, indicative of ectodermal differentiation.

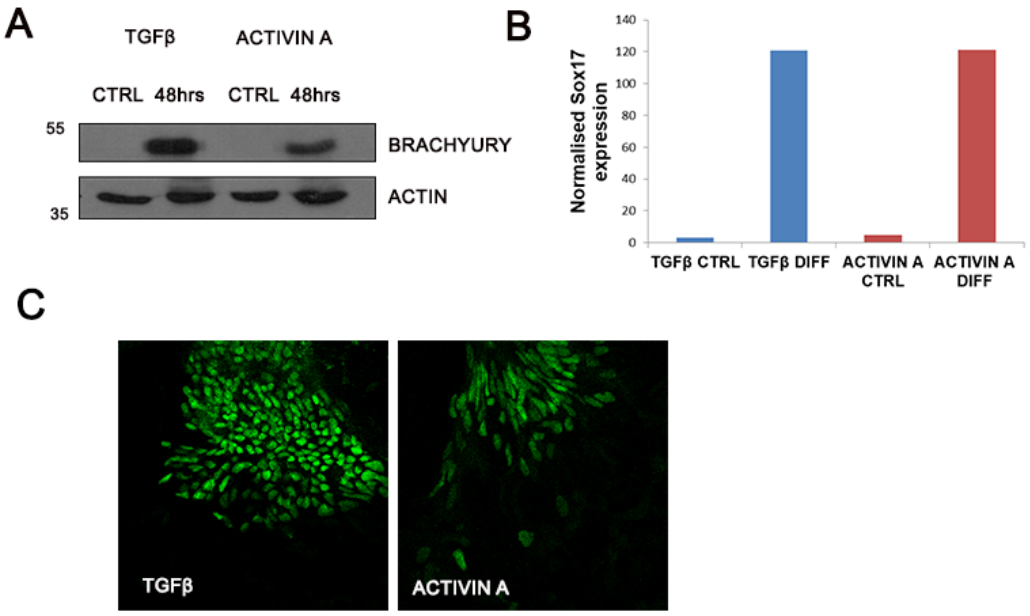


Figure 3-7 In vitro differentiation towards Mesoderm, Endoderm, Ectoderm.
H1 cells were subjected to three distinct differentiation protocols. For Mesodermal differentiation, H1 cells grown in the presence of either TGFβ or ACTIVIN A were induced for 48hrs with combination of ACTIVIN A and BMP4 while control cells were left untreated. For endodermal differentiation, H1 cells grown in the presence of either TGFβ or ACTIVIN A were subjected to a five days differentiation protocol using high doses of ACTIVIN A and serum. Control cells were left untreated. For ectodermal differentiation, H1 cells grown in the presence of either TGFβ or ACTIVIN A were forced to generate embryoid bodies and the derived bodies were further cultured to enable ectodermal differentiation. **A.** Western blot showing BRACHYURY (mesoderm) expression in H1 cells grown in the presence of either TGFβ (T) or ACTIVIN A (A). Control cells (CTRL) were left untreated while differentiated cells (48hrs) were produced by the combined effect of ACTIVIN A and BMP4. .**B.** Plot shows relative Sox17 expression (endoderm) in H1 cells grown in the presence of either TGFβ or ACTIVIN A. Control cells (CTRL) were left untreated, while differentiated cells (DIFF) were induced with high doses of ACTIVIN A and serum. **C.** IF staining reveals the expression of PAX6

(ectoderm) in H1 cells grown in the presence of either TGFβ or ACTIVIN A. Cells were grown in suspension to form embryoid bodies and the derived bodies were further cultured to allow ectodermal differentiation.

For the *in vivo* experiments two cell lines, H1 and HUES1 were cultured in parallel either in TGFβ or ACTIVIN A. Before reaching confluence they were manually dissected with an insulin needle and digested with dispase. The cell clumps were centrifuged and then resuspended in 100μl of hESC-qualified matrigel. This mixture was injected subcutaneously into the hind legs of immune deficient NOD/SCID mice (cells cultured in TGFβ were injected on the back left hind leg and cells cultured in ACTIVIN A were injected on the back right hind leg). The animals developed tumours and 10 weeks following injections they were sacrificed. The obtained tumours were dissected and stained for immunohistochemistry. Animal handling and teratoma analysis was performed by Dr. Dimitris Stellas. Both cell lines in both media were able to generate cells originating from all three germ layers. This further validates the ability of ACTIVIN A to protect a key feature of pluripotency in two distinct cell lines.

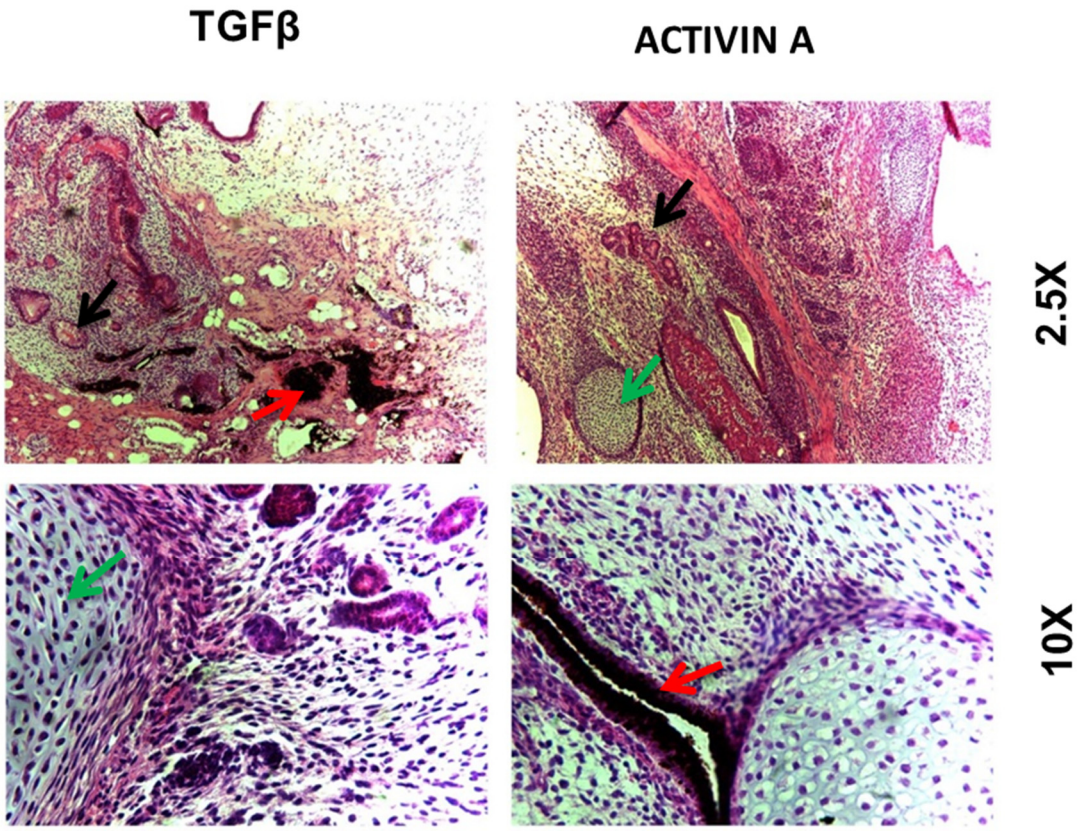


Figure 3-8 H1 cells generate teratomas in NOD/SCID mice.

H1 cells were cultured in either TGFβ or ACTIVIN A and before reaching confluence they were manually dissected, resuspended in matrigel and injected in NOD/SCID mice. Mice were sacrificed on week 10 following the injection and the obtained tumours were dissected, fixed in paraffin and stained with Hematoxylin-Eosin. In both media cells were able to differentiate towards chondrocytes (green arrow) indicative of mesoderm, melanocytes (red arrow) indicative of ectodermal differentiation and glandular formations (black arrow) which are derived from the endoderm.

3.3 Discussion

In this chapter it is seen that H1 cells are safely cultured on matrigel and mTeSR1 for prolonged periods without loss of pluripotency or appearance of chromosomal aberrations. Moreover, they respond to BMP4 by altering colony morphology and by acquiring migratory capacity. All these features are in agreement with previous reports and reflect the quality of the culture. Further dissecting the changes in protein and gene expression, reveals that 48hrs of BMP4 treatment leads to the expression of the mesodermal marker BRACHYURY. Furthermore, H1 cells by 48hrs of BMP4 induction, strongly upregulate the expression of Msx2 and Hand1. Collectively, these data suggest that BMP4 promotes mesodermal and extra-embryonic fates. Whether, BMP4 generates separate populations with distinct expression patterns of the abovementioned genes or whether they are co-expressed by common mesodermal precursors which will further diverge is not known. Additional detailed IF experiments are needed in order to further address this possibility.

An interesting finding of this chapter is the fact that short pulses of BMP4 are able to promote mesodermal commitment. This seems to contradict a previous study which reports that BMP4 drives differentiation through pSMAD1,5,8 which are active only for 6hrs following induction (Richter et al., 2014). The same report suggests that for this reason, differentiation protocols need freshly added BMP4 on a daily basis (Richter et al., 2014). However, another recent study bridges this gap by showing that BMP4 can upregulate Msx2 within 3hrs of treatment and MSX2 is able to lead mesodermal progression (Q. Wu et al., 2015). Therefore, it seems that BMP4 initiates commitment and perhaps downstream factors deliver the task of differentiation without the need for constant presence of BMP4.

Furthermore, the majority of pluripotency media contain TGFβ or combination of TGFβ with ACTIVIN (Desai *et al*, 2015). Another new finding demonstrated in this chapter is the ability of ACTIVIN A to substitute TGFβ in mTeSR1 without impairing pluripotency features in H1 cells. H1 cells cultured for several passages in ACTIVIN A maintain undifferentiated morphology, expression of key pluripotency factors and the ability to produce derivatives of mesoderm, endoderm and ectoderm both *in vitro* and *in vivo.*

4. QUANTITATIVE PHOSPHOPROTEOMICS (SILAC)

4.1 Introduction

Protein phosphorylation is a vital post-translational modification (PTM) and refers to the addition of phosphate molecule(s) to a protein (Macek *et al,* 2009). Briefly, kinases deliver this task by transferring γ-phosphate from a donor, which is usually ATP, to serine, threonine and tyrosine residues of a protein (Braconi Quintaje & Orchard, 2008). It is estimated that approximately one third of proteins are phosphorylated (López et al., 2012) but the degree of phosphorylation varies greatly (Macek *et al,* 2009). Phosphate groups, due to their negative charge, once bound to a protein can alter several aspects of proteins' behaviour, such as structure, interaction and localization (Braconi Quintaje & Orchard, 2008).

The human genome encodes 480 kinases (and 24 atypical kinases) and their action is counterbalanced by phosphatases (Braconi Quintaje & Orchard, 2008). These proteins are responsible for protein dephosphorylation and can be categorised into three groups (Barford *et al,* 1998). Two groups target phosphorylated serine and threonine residues, while the third group contains phosphatases targeting tyrosine residues and dual specificity phosphatases which target all three residues (Barford *et al,* 1998). Given the importance of structure in proper function, phosphorylation is considered as a molecular switch which defines whether a protein will acquire the on state or the off state (Braconi Quintaje & Orchard, 2008).

Studying phosphorylation can be achieved with the use of radioactive phosphorus or antibodies specifically designed against phosphorylated residues in the protein of interest (Macek *et al,* 2009). However, these approaches have several limitations,

while mass spectrometry (MS) is an alternative which enables the detection of phosphopeptides (novel ones included) and can additionally recognize the localization of the phosphate group within the protein (Macek *et al,* 2009).

Cellular responses to extracellular stimuli usually resembles a hierarchy, where molecules are successively regulated via phosphorylation (Invergo & Beltrao, 2018). In fact, kinases responsible for signal transduction are targets for phosphorylation (Invergo & Beltrao, 2018). As a result, revealing the phosphorylation status of proteins within a cell, contributes to the understanding of how signalling networks operate (Macek *et al,* 2009).

Shotgun proteomics refer to a particular workflow which is typically used to identify the proteome of a sample (X. Zhang et al., 2010). The succession of events taking place, begins with the enzymatic digestion of samples into smaller peptides (X. Zhang et al., 2010). These peptides, are first separated using high pressure liquid chromatography (HPLC) and then analysed by a mass spectrometer for further identification and quantitation (X. Zhang et al., 2010). Nevertheless, LC-MS is not an efficient tool for the study of large proteomes (X. Zhang et al., 2010), instead multidimensional liquid chromatography (MDLC)-tandem mass spectrometry (MS/MS) is considered to be the best approach for the analysis of total phosphoproteome or proteome (Tobe et al., 2012).

MDLC stands for the combination of methods (dimensions) used to separate peptide mixtures (Tobe et al., 2012). Briefly, samples are initially digested, usually with trypsin, and the resulting peptides are first separated (dimension 1) with techniques such as strong cation exchange chromatography (SCX) (Tobe et al., 2012). Eluted peptides are both phosphorylated and non-phosphorylated and the former can be

effectively enriched (dimension 2) with methods such as immobilized affinity chromatography (IMAC) and titanium dioxide (TiO_2) based enrichment (Tobe et al., 2012). IMAC distinguishes phosphopeptides by chelating phosphate groups to trivalent metal cations bound on a stationary phase, while TiO_2 based methods, use this material to enable binding of phosphate groups through Lewis acid-base interactions (Tobe et al., 2012). Next, peptides are further separated by reversed phase Liquid Chromatography (RP-LC) (dimension 3), then subjected to electrospray ionization and finally injected in the mass spectrometer (Tobe et al., 2012). These powerful machines measure mass to charge ratio (m/z) and intensity (Tobe et al., 2012). The most abundant peptides are then isolated, fragmented and the obtained product ions are scanned again thus resulting in a tandem-mass spectra (MS/MS) which are used to construct the sequence of the original peptide (Tobe et al., 2012).

Despite the significant advantages accompanied by the use of MS for the study of phosphorylation, MS is not a quantitative method (Macek *et al,* 2009). Therefore, several labelling techniques have been employed in order to address this issue and two of them known as SILAC and iTRAQ are the most preferred ones (Macek *et al,* 2009). SILAC stands for stable isotope labelling by amino acids in culture and was first introduced in 2002 as an efficient way to substitute normal essential amino acids by labelled amino acids in cultured cells (Ong et al., 2002). Ong and co-workers cultured cells either in media containing normal Leucine or media containing deuterated Leucine (Ong et al., 2002). After several doublings, all proteins contained deuterated leucine instead of normal leucine (Ong et al., 2002). The two cell populations were then differentially treated and samples were mixed and quantitatively analysed using MS (Ong et al., 2002). An alternative SILAC scenario,

uses non-radioactive isotopically labelled Lysine and Arginine and can allow the comparison between two or three conditions (Tobe et al., 2012). The labelled amino acids contain heavier isotopes of carbon and nitrogen and therefore they differ in mass number despite sharing the same atomic number (Tobe et al., 2012). Once isotopes are completely incorporated in proteins, samples from different conditions can be mixed and peptides with the same sequence derived from the different conditions can be distinguished according to their mass to charge ratio (m/z) (Tobe et al., 2012).

Studying the phosphoproteome of hESCs is quite a recent endeavour. In 2009, two distinct studies attempted to analyse it using different techniques. In the one case, label free methods were used to compare the phosphorylation status between pluripotent H1 cells and differentiated cells (Brill et al., 2009). Data analysis revealed an association between receptor tyrosine kinases phosphorylation and pluripotency, which was further assessed with *in vitro* experiments (Brill et al., 2009). Additionally, the presence of several putative JNK effectors (phosphorylated) in the undifferentiated state implied that this kinase could regulate pluripotency (Brill et al., 2009). Indeed, inhibition experiments supported a novel role of JNK in the maintenance of the undifferentiated state (Brill et al., 2009).

The other report published in 2009, addressed changes in the phosphoproteome of hESCs using SILAC and focused on a much shorter time window (Van Hoof *et al.,* 2009). Interestingly, the simultaneous removal of FGF2 and addition of BMP4 triggered significant changes in almost half of the identified phosphosites, within one hour of treatment (Van Hoof *et al.,* 2009). Among the most prominent findings was the phosphorylation of SOX2, which possibly enables a second PTM (SUMOylation),

thus decreasing its transcriptional activity (Van Hoof et al., 2009). Furthermore, prediction analysis of the obtained dataset suggested that CDK1/2 is the most prevalent kinase regulating numerous phosphosites in hESCs (Van Hoof et al., 2009).

During the last years a new method has emerged in the investigation of protein phosphorylation. Briefly, in 2004 it was reported that alcoxide-bridged dinuclear zinc (II) complex can recognise a phosphate monoester dianion in aqueous solutions (Kinoshita et al., 2004). These type of complexes were named phos-Tag molecules due to their phosphate binding property (Kinoshita et al., 2006). Phos-Tag molecules can be bound either to Zn^{2+}, or Mn^{2+} and they can bind phosphomonoester dianions which are found on serine, threonine or tyrosine (Kinoshita et al., 2006). Interestingly, Mn^{2+}-Phos-Tag conjugated to acrylamide in an SDS-PAGE enables the distinction between phosphorylated and non-phosphorylated proteins (Kinoshita et al., 2006). This is achieved thanks to the binding property of the phos-Tag molecule to the phosphorylated protein which in turn reduces the migration of the former compared to the non-phosphorylated protein (Kinoshita et al., 2006). Differential migration capacity, alters the final location of proteins on the blot, thus enabling efficient distinction (Kinoshita et al., 2006).

In 2011, it was reported that Zn^{2+}-Phos-Tag SDS-PAGE gels could also be used to study phosphorylation events (Kinoshita & Kinoshita-Kikuta, 2011). Although, the basic principles remain the same, the new system works under neutral pH conditions as opposed to the need for alkakine pH for Mn^{2+}-Phos-Tag gels (Kinoshita & Kinoshita-Kikuta, 2011). Moreover, it is considered to be superior to the old system, given its advanced ability in distinguishing phosphoproteins (Kinoshita & Kinoshita-

Kikuta, 2011). There are several critical steps that need to be strictly followed in order for the phos-Tag protocols to work; one of them being the need to avoid any sample exposure to EDTA (Nagy et al., 2018). Overall, this new tool enables the study of phosphorylation events, bypassing the need for phosphosite specific antibodies and radioactive material while it uses simple rudimentary equipment (Nagy et al., 2018).

The aims of this chapter are the following: First to use the SILAC approach combined with MS in order to reveal differences during differentiation of hESCs caused by BMP4. Second, to use the novel phos-Tag technique in order to study phosphorylation events regarding interesting candidates that will be selected form the obtained dataset.

4.2 Results

4.2.1 Incorporation of labelled amino acids in culture and experimental setup

The first step towards SILAC/MS experiments was the incorporation of differentially labelled aminoacids in cultured H1 cells. Therefore, H1 cells were cultured in customized mTeSR1 without Lysine and Arginine and the medium was supplemented with either (1) standard Lysine and Arginine, (2) medium labelled forms of Lysine and Arginine (R6K4) or (3) heavy labelled forms of Lysine and Arginine (R10K8) as previously described in **Table 2-5**, in their original concentration (0.548 mM Arg, 0.391 mM Lys). After three passages cells were lysed for protein extraction. 40 μg of whole cell lysate were loaded on a 12% polyacrylamide gel. The next day a small piece of the gel was excised with a sterile scalpel and was in gel digested (Materials and Methods). At the end of the procedure samples were dried down in the speed vac, resuspended in 0.1% formic acid and then injected in the M/S

Orbitrap. Data were analysed by Dr. Debbie Cunningham using MaxQuant and showed almost complete incorporation of the labelled amino acids in culture. R6K4 were incorporated in a percentage of 95-100% in 259 out 264 identified peptides **(Figure 4-1A)** and R10K8 were incorporated in a percentage of 95-100% in 891 out of 903 peptides **(Figure 4-1B).**

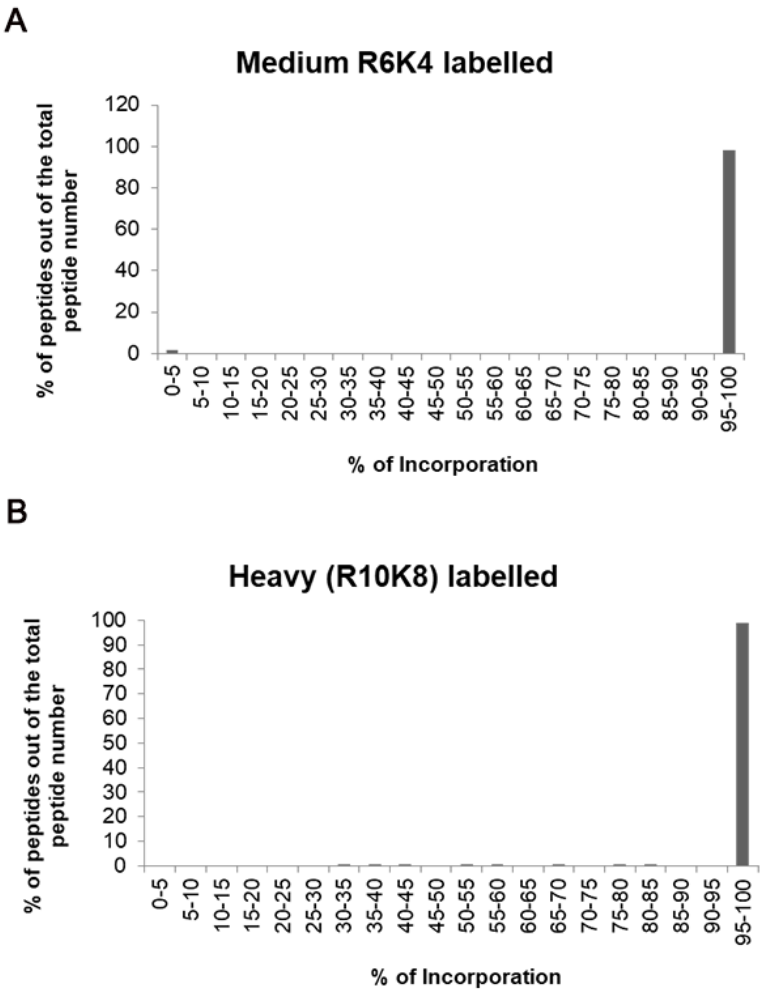


Figure 4-1 Incorporation percentage of labelled Lysine (K) and Arginine (R) in H1 cells.

H1 cells were cultured for three passages in media supplemented with labelled aminoacids. Cells were lysed and 40µg were run on an SDS-PAGE gel. A small piece of the gel was cut, trypsin digested and peptides were extracted from the gel. Next, peptides were zip tipped and run on a MS. The plots show the percentage of incorporation for the labelled peptides. The horizontal axis depicts the percentage of incorporation in identified peptides and the vertical axis depicts the respective percentage of peptides in the total population of the identified peptides. **A.** 95-100% incorporation of R6K4 in 98.1% of the identified peptides in H1 cells. **B.** 95-100% incorporation of R10K8 in 98.6% of the identified peptides in H1 cells.

The purpose of the experiment was to reveal differences in the phosphoproteome of H1 cells upon the addition of BMP4 or the abrogation of ACTIVIN A signalling. For this reason, cells grown in three media containing light, medium and heavy labelled Lysine and Arginine (and therefore called light, medium and heavy medium) were used to explore the three different conditions. Cells grown in the first medium (light labelled) served as the control and were left untreated. Cells grown in the second medium (medium labelled) were used to abrogate ACTIVIN A signaling. In detail, cells were treated with SB431542 10 mM for 3hrs. Last, cells grown in the third medium (heavy labelled) were used for BMP4 induction. In detail, cells were induced with BMP4 50ng/ml for 1 hr. SB431542 is a small molecule inhibitor which blocks ALK4/ALK5/ALK7 thus abrogating TGFβ/ACTIVIN/NODAL signalling (Inman et al., 2002). The concentration and duration of treatments were previously tested **(Fig. A-3,4)**. The experiment was repeated 3 times but the second time (experiment 2) the labels were swapped between the conditions **(Fig. 4-2)**..

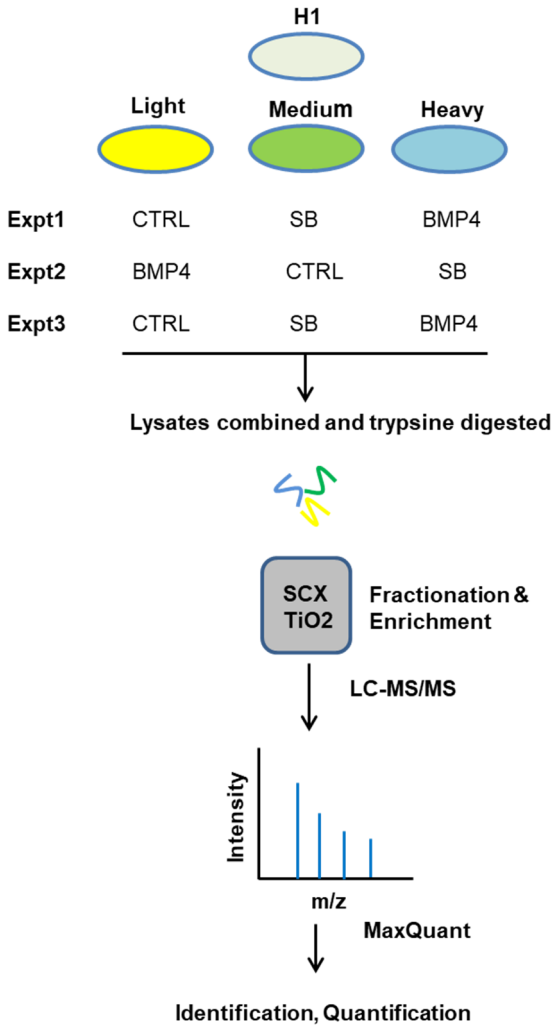


Figure 4-2 Experimental Setup.

First, H1 cells were grown in three media containing differentially labelled arginine and lysine (light, medium and heavy). Once complete incorporation of labelled aminoacids was confirmed, cells were used for differential treatment. Cells grown in light medium served as the control state and were left untreated. Cells grown in medium medium were treated with SB to abrogate ACTIVIN A signalling. Cells grown in heavy medium were induced with BMP4. Experiment 2 is an exception to this description as treatments were swapped. At the end of every treatment cells were lysed. Lysates (light, medium, heavy) were combined in a 1:1:1 ratio and trypsinised to produce small peptides. Next, peptides were fractionated using SCX and samples were enriched for phosphopeptides using TiO2 tips. At the end samples were injected in the MS and the derived spectra and ratios were analysed using MaxQuant.

4.2.2 SILAC data evaluation

The amount of total lysate used in each experiment was 2.7mg, 4.5mg and 5.4mg for experiment 1, 2 and 3 respectively. Experiment 1 and 2 showed a similar SCX profile compared to experiment 3. The maximum UV absorbance for experiment 1, 2 and 3 was 1250 mAU, 750 mAU and 1750 mAU respectively. Phosphopeptide enrichment with the TiO2 kit was reasonably efficient. Experiment 3 was the best in terms of enrichment, according to the percentage of identified phosphopeptides on the total amount of peptides. Note that fractions 16-20 contain a very small amount of peptides in general **(Fig. 4-3)**.

In total 16.308 peptides were identified across all experiments, of these 6006 are phosphopeptides and 7175 unmodified peptides. 1974 phosphopeptides were common across experiments and 1366 displayed a high localization propability (>0.75) and high score difference (>5). 1369 phosphopeptides had M/L ratio and 1322 phosphopeptides had H/L ratio. In order to obtain more information from the already existing samples, the flow through of the zip tip step were enriched for phosphopeptides and further injected in an advanced MS. The additional peptides identified raised the overall numbers. Once reverse hits and potential contaminants were filtered out, the overall number of unmodified peptides was 7,175 and for the phosphopeptides was 9,241. Moreover, 2,092 phosphopeptides displayed M/L ratio, while 2,046 phosphopeptides displayed H/L ratio.

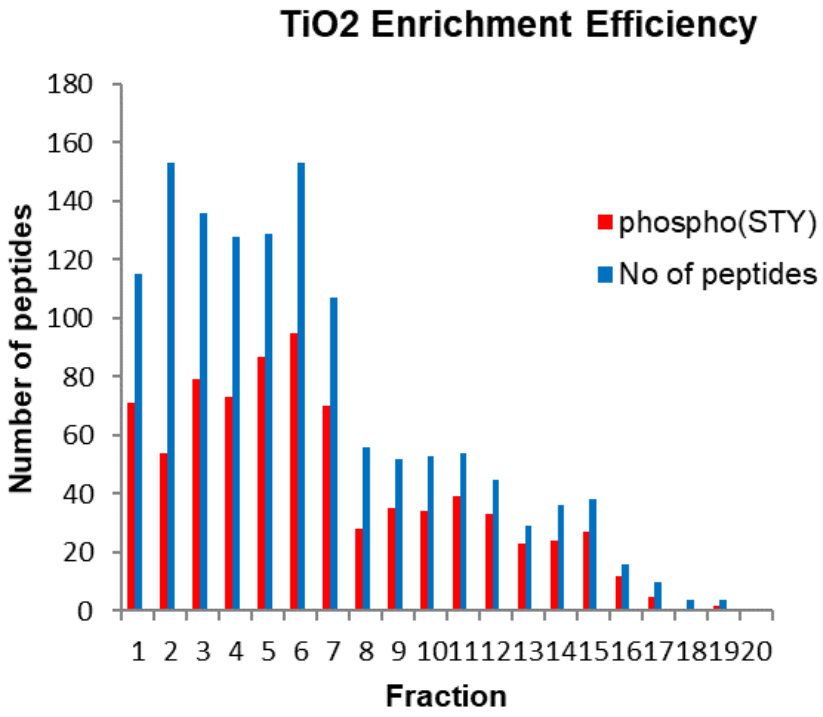


Figure 4-3 TiO2 Enrichment efficiency.

The chart demonstrates the efficiency of TiO2 tips in enriching for phosphopeptides. The vertical axis demonstrates the 20 fractions used, while the horizontal axis presents the number of phosphopeptides (red) and the number of total peptides (blue) identified in each fraction. The data originate from technical replicate 1, biological replicate 3. The overall efficiency is 60% (792 phosphopeptides out of 1318).

The initially identified phosphopeptides were imported into the Perseus software (Tyanova et al., 2016), reverse hits and potential contaminants were removed and the remaining ratios were log2 transformed and then plotted as a multi scatter plot where every experiment was compared to all others. Pearson correlation results revealed medium degree of correlation across experiments 1 to 4 and very low correlation for experiments 5 and 6 (where aminoacids were swapped). It is possible

that the decreased size of the obtained dataset does not offer the possibility to reveal better correlations.

4.2.3 BMP4 induced changes in phosphosites

In order to dissect the effect of BMP4 at the onset of stem cell differentiation, H1 cells were induced with the growth factor (50ng/ml) for 1hr and compared to uninduced cells. The growth factor should elicit a response characterised by increase in BMP4 driven phosphorylations. This comparison is described by the Normalized Ratio H/L. Due to the small size of the obtained dataset and the low reproducibility in the differential phosphorylation of peptides between control and induced cells, the use of t-test was not feasible. Therefore, it was decided, that all phosphosites would be included in the analysis and an arbitrary cutoff was set at 0.5 to mark decrease and at 2.0 to mark increase.

According to this, the addition of BMP4 induced upregulation in 39 phosphosites **(Table 4-1)** (1.9%) (ratio>2.0), while it downregulated 145 phosphosites (7.08%) (<0.5).

Table 4-1 Table lists the identified phosphopeptides upregulated by BMP4.

The first column (left) lists the Uniprot Accession Number of each protein. The second column lists the protein name and gene symbol. The third column lists the identified phosphosite, while the fourth column lists the H/L ratio in descending order.

Uniprot Accession Number	Protein Name/Gene symbol	Identified Phosphosite	H/L Ratio
Q9UKV3	Apoptotic chromatin condensation inducer in the nucleus/ACIN1	Ser384	10.872
Q9HAU0	Pleckstrin homology domain-containing, family A member 5/PLEKHA5	Ser568	9.1447
Q9UKV3	Apoptotic chromatin condensation inducer in the nucleus/ACIN1	Ser388	8.488
P23193	TCEA1/Transcription elongation factor A protein 1	Ser100	6.8439
Q9NSY1	BMP2K/BMP-2-inducible protein kinase	Ser391	5.6686
Q9NSY1	BMP2K/BMP-2-inducible protein kinase	Ser392	5.6686
Q9NSY1	BMP2K/BMP-2-inducible protein kinase	Thr394	5.6686
Q14157	Ubiquitin-associated protein 2-like/UBAP2L	Ser605	3.7076
Q8N4C8	Misshapen-like kinase 1/MINK1	Ser763	3.5149
P40222	TXLNA/Alpha-taxilin	Ser515	3.0453
Q9UKV3	Apoptotic chromatin condensation inducer in the nucleus/ACIN1	Ser386	2.938
Q9NRY4	ARHGAP35/Rho GTPase-activating protein 35	Ser1179	2.8919
Q9UHD8	Septin-9/SEPT9	Thr331	2.5747
Q00839	Heterogeneous nuclear ribonucleoprotein U/HNRNPU	Ser271	2.5333
P35222	Catenin beta-1/CTNNB1	Thr551	2.5117
P27816	Microtubule-associated protein 4/MAP4	Ser358	2.422

P11137	Microtubule-associated protein 2/MAP2	Ser1782	2.3584
Q9BUA3	Spindlin interactor and repressor of chromatin-binding protein/SPINDOC	Ser308	2.3493
Q14247	Src substrate cortactin/CTTN	Thr401	2.318
Q14247	Src substrate cortactin/CTTN	Ser405	2.318
O96013	Serine/threonine-protein kinase PAK 4/PAK4	Ser99	2.2992
Q96GX5	Serine/threonine-protein kinase greatwall/MASTL	Ser453	2.2467
Q9UHB7	AF4/FMR2 family member 4/AFF4	Ser180	2.2417
P07197	Neurofilament medium polypeptide/NEFM	Ser680	2.2322
P07197	Neurofilament medium polypeptide/NEFM	Ser685	2.2322
Q6PKG0	La-related protein 1/LARP1	Ser774	2.2055
P35222	CTNNB1/catenin beta 1	Ser191	2.1466
P53367	ARFIP1/Arfaptin-1	Ser132	2.1459
P08670	Vimentin/VIM	Ser55	2.1387
P54105	Methylosome subunit pICln/CLNS1A	Ser102	2.1363
Q86YP4	GATAD2A/Transcriptional repressor p66-alpha	Ser546	2.1133
P11137	Microtubule-associated protein 2/MAP2	Ser1782	2.1098
Q96C90	Protein phosphatase 1 regulatory subunit 14B/PPP1R14B	Ser32	2.0977
Q7Z417	Nuclear fragile X mental retardation-interacting protein 2/NUFIP2	Ser629	2.0815
P50479	PDZ and LIM domain protein 4/PDLIM4	Ser112	2.0571
Q8IYB3	Serine/arginine repetitive matrix protein 1/SRRM1	Ser260	2.0365
Q9UQ35	SRRM2/ Serine/arginine repetitive matrix protein 2	Thr289	2.01
P08559	Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial/PDHA1	Ser232	2.0076
Q07960	ARHGAP1/Rho GTPase-activating protein 1	Ser51	2.0051

Several of the proteins whose phosphosites were upregulated are related to migration. More specifically, in Hela cells, La-related protein 1 (LARP1) is associated with several cytoskeletal proteins as well as zona occludens-1 (ZO-1) and LARP1 knockdown impairs cell motility (Burrows et al., 2010). Moreover, Rho-GTP-ase activating protein-1 (ARHGAP1) is involved in the onset of EMT in the zebrafish embryo (Clay & Halloran, 2013). ARHGAP1 facilitates the separation of neural crest cells from the epithelium by restricting Rho activity to the apical region of the cells which further leads to detachment (Clay & Halloran, 2013). Invading cancer cells migrate through tissues by generating protrusions known as invadopodia while they also decay the surrounding ECM via the action of matrix metalloproteinases (MMPs)(Yin, Ma, & An, 2017). Cortactin is involved in both processes and PTMs seem to play an important regulatory role (Yin et al., 2017). Our data show that CORTACTIN Thr^{401}, Ser^{405} are upregulated upon BMP4 induction and interestingly, CORTACTIN Ser^{405}, Ser^{418} phosphorylation is found to be ERK-induced and is suggested to promote activation of the actin related protein 2/3 (ARP2/3) complex (Martinez-Quiles et al., 2004). ARP2/3 complex leads to actin polymerization and the resulting bundle pushes the cell membrane in the form of a protrusion (Revach & Geiger, 2014).

Transcription and transcriptional regulation were also enriched terms for Biological Process. CTNNB1 Ser^{191} was upregulated in the dataset and recent findings link this modification with WNT activity (X. Wu et al., 2008). In detail, WNT signaling leads to CTNNB1 phosphorylation at Ser^{191} and Ser^{605} via the action of JNK2 and these phosphorylations are important for CTNNB1 nuclear translocation (X. Wu et al., 2008). GATAD2A Thr^{189} was also upregulated and GATAD2A is known as

component of the methyl binding domain protein-2 (MBD2)-containing nucleosome remodeling and deacetylation (NuRD) complex (Gnanapragasam et al., 2011). NuRD complex is a chromatin modifier and can modulate transcription not only via its action on chromatin but also via its action on gene promoters and enhancers (Basta & Rauchman, 2015). Although, NuRD complex was originally marked as a repressor, it is suggested that it can positively affect transcription (Basta & Rauchman, 2015).

In order to better understand the processes that are regulated by BMP4 in stem cells, the DAVID Functional Annotation tool was used (D. W. Huang, Sherman, & Lempicki, 2009a), (D. W. Huang, Sherman, & Lempicki, 2009b). In detail, the Uniport Accession numbers of the proteins whose phosphosites were found upregulated were uploaded in DAVID website and clustered. The 39 identified phosphosites belong to 31 proteins which can be grouped in 3 clusters. The terms that were enriched for Biological Process were cell-cell adhesion, transcription and transcriptional regulation **(Fig. 4-4A)**. The most enriched molecular function was cadherin binding **(Fig. 4-4B)** and the enriched cellular compartments were adherens junctions and the perinuclear cytoplasmic region **(Fig. 4-4C).**

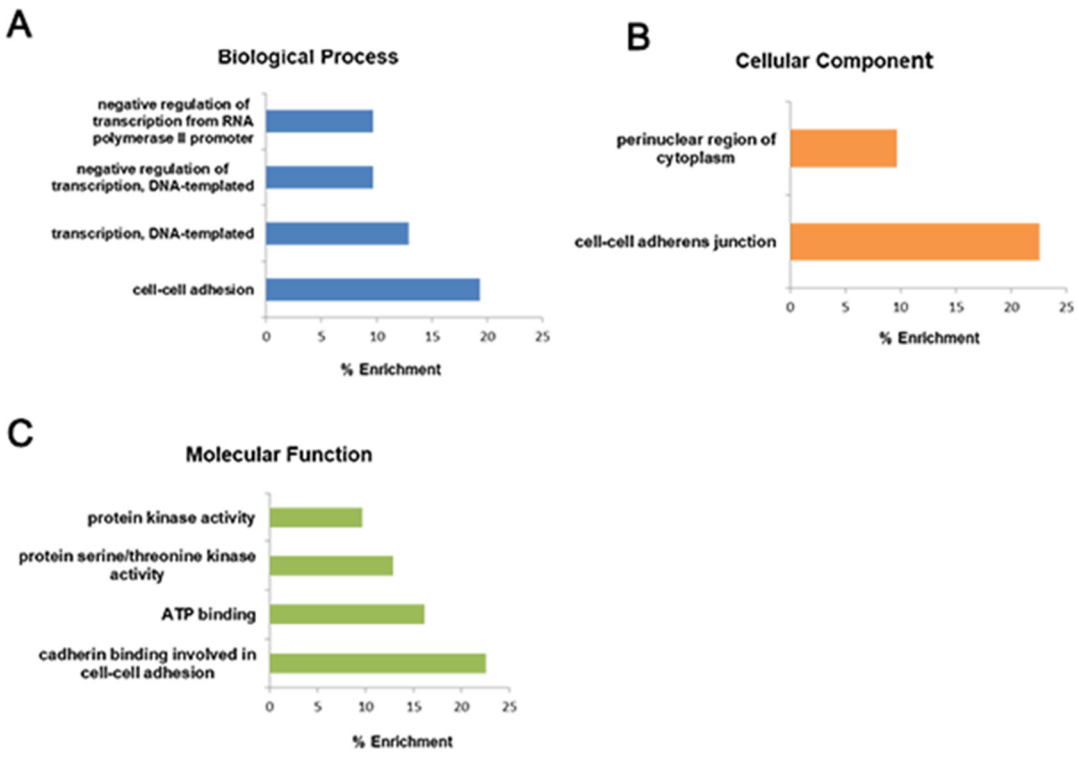


Figure 4-4 Enriched GO terms according to DAVID clustering.

The proteins whose phosphosites were upregulated upon BMP4 addition were uploaded in DAVID Functional Annotation tool. 3 clusters were obtained and the top ten enriched terms were plotted. In the charts, the horizontal axis illustrates the percentage of enrichment while the vertical axis illustrates the enriched terms. **A.** Enriched terms for Biological Process. **B.** Enriched terms for Cellular Component. **C.** Enriched terms for Molecular Function.

Moreover, in order to gain an insight into the credibility of the dataset, the upregulated phosphosites were compared to previously published data where researchers revealed the phosphoproteome differences in hESCs upon BMP4 addition and self-renewal factors' removal (Van Hoof et al., 2009). Out of the 39 phosphosites, 35 have a strong localization propability (>0.75) and a high score difference (>5.0). According to the comparison **(Fig. 4-5)**, 5 phosphosites (14.28%)

were upregulated in both datasets, 5 (14.28%) were found in both cases but were not upregulated in the published dataset. Moreover, 12 phosphosites (34.28%) were not upregulated in the second dataset but other residues of the same protein were found upregulated. Last, 13 (37. 14%) phosphosites belong to proteins which were not found in the other dataset. Interestingly, in the last category are some phosphosites who seem reasonable candidates for regulation by BMP4, such as for example three residues of BMP2 inducible kinase (BMP2K).

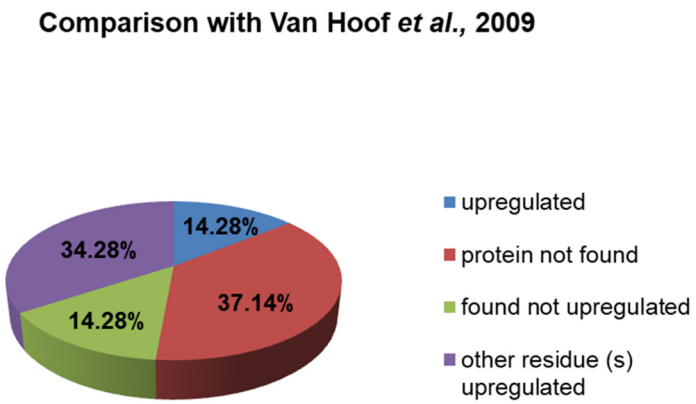


Figure 4-5 Comparison between the obtained upregulated phosphosites and those previously published (Van Hoof et al., 2009).

The phosphosites found upregulated were searched against an online available list of phosphosites regulated by BMP4 addition and removal of self-renewal factors in hESCs (Van Hoof et al., 2009). The phosposites were classified in four categories which are demonstrated in the form of a pie chart with the respective percentage in the total population of upregulated phosphosites. 14.28% of the upregulated phosphosites were also upregulated in the published dataset (blue), while another 14.28% of them were found but were not upregulated in the published dataset (green). Additionally, 34.28% of the upregulated phosphopeptides were found to be phosphorylated at different residue (s) in the published dataset (purple). 37.14% of the upregulated phosphosites belong to proteins which were not identified in the published dataset (red).

Taking into consideration the notion that downregulated phosphosites also carry information regarding the regulatory role of BMP4 in stem cell differentiation, the proteins whose phosphosites were found downregulated were clustered using the DAVID Functional Annotation Tool (D. W. Huang et al., 2009a), (D. W. Huang et al., 2009b). The 145 identified phosphosites belong to 129 proteins which can be grouped in 7 clusters. The most enriched term for biological process was negative regulation of transcription from RNA polymerase II promoter **(Fig. 4-6A)**. Meanwhile, the most enriched term for cellular component was Golgi apparatus **(Fig. 4-6B)** and the most enriched term for molecular function was transcription factor binding **(Fig. 4-6C)**.

Among the downregulated phosphosites were HDAC1 Ser^{393} and HDAC2 Ser^{394}. Histone deacetylases (HDACs) remove acetyl groups from histones, thus inducing a less accessible chromatin and this action is associated with transcriptional repression (Delcuve et al., 2012). Interestingly, phosphorylation of HDAC1 at residues Ser^{393}, Ser^{421} and Ser^{423} as well as phosphorylation of HDAC2 at residues Ser^{394}, Ser^{422} and Ser^{424} is suggested to promote the activity of the enzymes (Delcuve et al., 2012). Surprisingly, in mESCs, NANOG and OCT4 associate with several repressors such as HDAC1/2 and form a complex known as NODE (Liang et al., 2008). This complex can perform deacetylation and mESCs display propensity to differentiation upon knockdown of some of its components (Liang et al., 2008). Although, HDAC2 knockdown does not markedly affect pluripotency, it is suggested that HDAC1 or other repressors could compensate for HDAC2 loss (Liang et al., 2008).

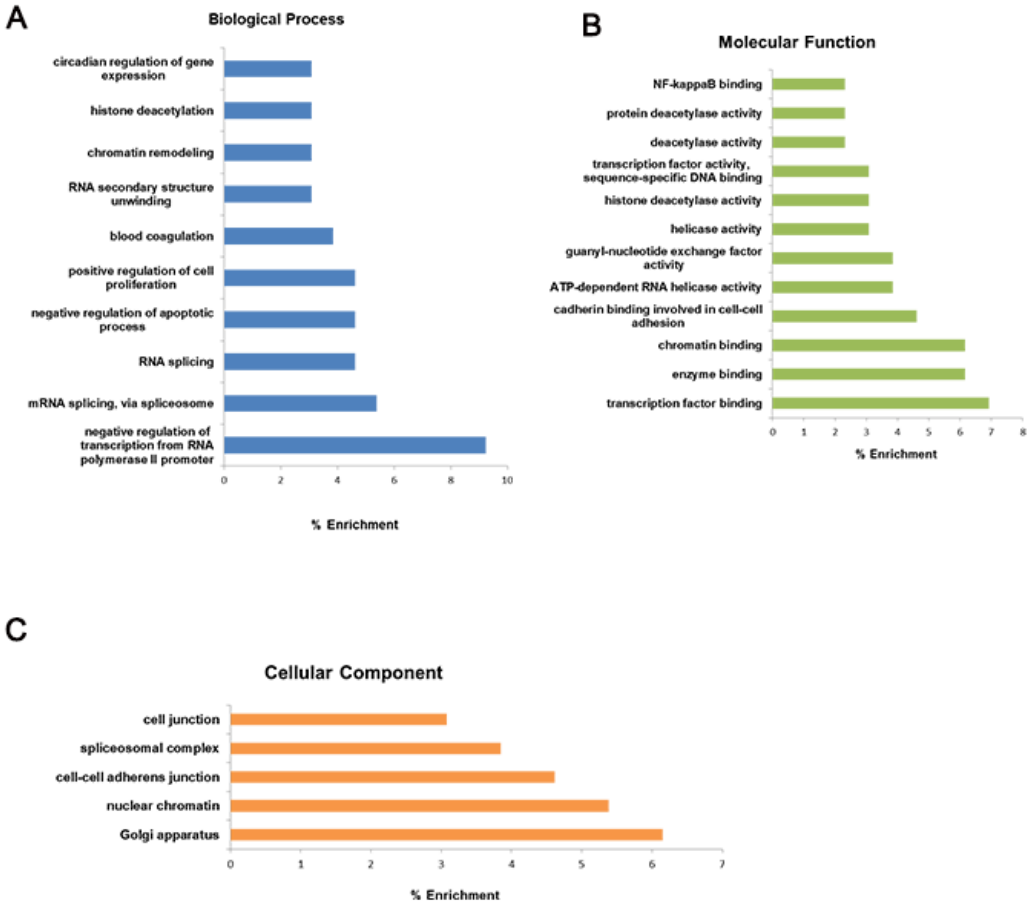


Figure 4-6 Enriched GO terms according to DAVID clustering.

The proteins whose phosphosites were downregulated upon BMP4 addition were uploaded in DAVID Functional Annotation tool. 3 clusters were obtained and the top ten enriched terms were plotted. In the charts, the horizontal axis illustrates the percentage of enrichment while the vertical axis illustrates the enriched terms. **A.** Enriched terms for Biological Process. **B.** Enriched terms for Cellular Component. **C.** Enriched terms for Molecular Function.

4.2.4 Phosphorylation of pluripotency associated proteins

In 2007, the International Stem Cell Initiative published a detailed characterization of 59 hESC lines (International Stem Cell Initiative, 2007). Among few interesting findings, researchers compared the expression levels of several pluripotency and

differentiation associated genes between pluripotent cells and embryoid bodies (International Stem Cell Initiative, 2007). Next, they correlated the data with the expression of Nanog and revealed 5 genes strongly correlated with Nanog as well as 14 genes with a decent degree of correlation (International Stem Cell Initiative, 2007). Our dataset contains phosphosites belonging to 5 proteins out of the described genes **(Table 4-2)**. In particular, there are two phosphorylation sites for DNMT3B which shows a strong Nanog correlation. Moreover, there were identified phosphosites for proteins encoded by genes which display medium to strong correlation, such as Lin28A Ser^{200}, Lin28B Thr^{202}, Ser^{203}, GRBP7 Ser^{361} and SOX2 Ser^{251}. These findings seem to indicate a strong role of phosphorylation events in pluripotency.

Table 4-2 Table lists the identified phosphosites in pluripotency associated proteins. The left column contains the Uniprot Accession number of the protein, the column in the middle lists the official gene name and protein name and the right column contains the identified phosphosite(s). Genes are listed in descending order of Nanog pairwise correlation. Phosphosites in black were identified across all three biological experiments, while phosphosites in blue were only identified in one.

Uniprot Accession Number	Gene Symbol (Protein Name)	Identified Phosphosite(s)
Q9UBC3	DNA Methyltransferase-3B (DNMT3B)	Ser136, Thr383
Q9H9Z2	Lin-28 Homolog A (Lin-28A)	Ser200
Q6ZN17	Lin-28 Homolog B (Lin-28B)	Thr202, Ser203
Q14451	Growth Factor Receptor Bound Protein 7 (GRBP7)	Ser361
P48431	SRY-box2 (SOX2)	Ser251

4.2.5 Phosphorylation of Wnt signalling components

Given the known importance of Wnt pathway in hESCs, the obtained dataset was searched for phosphosites belonging to components of the Wnt pathway. CTNNB is known to exist in two pools in the cell, the first and more stable being the CADHERIN bound and the second being the cytosolic which is amenable to degradation(Daugherty & Gottardi, 2007). CTNNB proteasomal degradation is a phosphorylation driven event, since phosphorylation at Ser^{45} by casein kinase-1a (CK1a) promotes the subsequent phosphorylations at Ser^{33}, Ser^{37} and Thr^{41} by GSK3b (Liu et al., 2002) and these aminoterminal modifications facilitate degradation (Liu et al., 1999). Although, in our dataset we identified 4 phosphosites for CTNNB1 **(Table 4-3)**, they are not in the aminoterminal region and therefore, concluding about the status of CTNNB1 is not straightforward. Meanwhile, GSK3b Tyr^{216} was found and this modification is regarded a product of autophosphorylation (Cole, Frame, & Cohen, 2004) and marks the active status of the protein (Hughes et al.,1993).

Some of the identified CTNNB phosphosites are suggested to play a regulatory role in the function of the protein. More specifically, CTNNB Ser^{191} and Ser^{246} phosphorylation is reported as a result of the action of the CDK5/P35 complex (Muñoz et al., 2007). These modifications promote interaction with Peptidyl Prolyl Isomerase1 (PIN1), which could have a protective role for CTNNB against degradation (Muñoz et al., 2007). Moreover, EGF induced AKT activation leads to phosphorylation of CTNNB1 at Ser^{552} thus promoting translocation from cell junctions towards the cytoplasm and the nucleus (Fang et al., 2007). Additionally, PKA phosphorylates CTNNB Ser^{552} and Ser^{675} thus enhancing CTNNB dependent transcriptional activity (Taurin et al., 2006).

Furthermore, the dataset also contains phosphosites belonging to proteins which modulate WNT activity. In particular, BCL9 Thr^{155}, Ser^{687} were identified and BCL9 is regarded as a Wnt co-activator (Kramps et al., 2002), while three more residues namely TLE1 Ser^{286}, TLE2 Ser^{203} and TLE2 Ser^{286} belong to members of a family which represses Wnt activity (Levanon et al., 1998). All the above seem to indicate that WNT pathway is under regulation in hESCs.

Table 4-3 Table lists the identified phosphosites in important components of the Wnt pathway. The left column contains the Uniprot Accession number of the protein, the column in the middle lists the official gene name and protein name and the right column contains the identified phosphosite(s). Phosphosites in black were identified in two or three biological experiments, while phosphosites in blue were only identified in one.

Uniprot Accession Number	Gene Symbol (Protein Name)	Identified Phosphosite(s)
P35222	Catennin beta-1 (CTNNB1)	Ser191, Ser552, Thr556, Ser675
P49841	Glucogen synthase kinase-3 beta (GSK3b)	Tyr216
O00512	B-cell CLL/lymphoma 9 protein (BCL9)	Thr155, Ser687
Q04724	Transducin-like enhancer protein 1 (TLE1)	Ser286
Q04726	Transducin-like enhancer protein 3 (TLE3)	Ser203, Ser286

4.2.6 Assessing SILAC derived candidates using Phos-Tag gels

The next step after obtaining the dataset was to try to validate some of the phosphosites which were found upregulated by BMP4. Out of the available candidates, TCEA1 Ser^{100} and PLEKHA6 Ser^{568} were selected.

Transcription elongation Factor II S (TFSII/TCEA) is one of several identified elongation factors for RNA Polymerase II (RNAPII) (Reines et al.,1996). RNAPII is responsible for mRNA synthesis (Paule, 2000) and TFIIS assists transcription by facilitating RNAPII passage through arrest sites (Reines et al., 1996). TCEA1 is one of the three TCEA isoforms and is considered to be ubiquitously expressed (Cha et al., 2013), while recent findings link it with proliferation and differentiation in myeloid cells (T. Yang et al., 2018). In particular, TCEA1 knockdown enhances proliferation while it compromises GMSF-induced differentiation of myeloid cells and affects several transcription factors (T. Yang et al., 2018). Although, there is no known underlying mechanism, data seem to indicate that TCEA1 is to some extent involved in myeloid fate specification (T. Yang et al., 2018). In addition, TCEA3 was reported to interact with both Type I and Type II TGBβRs in ovarian cancer cells, and binding to the former seems to induce apoptosis via JNK (Cha et al., 2013).

Pleckstrin homology domain-containing family A member 5 (PLEKHA5) is one of the seven PLEKHA proteins which contain a pleckstrin homology (PH) domain (Eisele et al., 2015). PH domain is one of the most frequently found domains in human proteins and is known to mediate membrane anchoring by binding to phosphoinositides(Lemmon, 2004). PLEKHA5 in particular, was found widely expressed across several tested human tissues (Dowler et al., 2000), while the presence of two transcript variants as well as the in vitro ability to bind PI3P were

later confirmed (Yamada et al., 2007). Although the actual function of the protein is not well understood, PLEKHA5 localization could indicate possible roles (Yamada et al., 2007). Interestingly, PLEKHA5 is mostly found on the plasma membrane and on microtubules, while the accumulation of PLEKHA5 in actin nucleation sites of migrating cells implies an association with cell motility (Yamada et al.,2012).

The kinases responsible for TCEA1 Ser^{100} and PLEKHA5 Ser^{568} phosphorylation were searched using NetworKIN 3.0 (Horn et al., 2014) and Kinase Enrichment Analysis 2 (KEA2) (Lachmann & Ma'ayan, 2009). NetworKIN is a freely available platform which combines information regarding kinase motif recognition and information regarding putative kinase-substrate interaction, in order to predict kinase(s) responsible for phosphorylation events (Horn et al., 2014). KEA2 is a freely available tool which encompasses information regarding kinase-substrate interaction from several sources and can be used to associate a phosphosite with the responsible kinase (Lachmann & Ma'ayan, 2009).

TCEA1 phosphorylation at Ser^{100} was predicted by both tools to be a result of MAPK14 activity. MAPK14 is one of the main active kinases in differentiating H1 cells (Van Hoof et al., 2009). PLEKHA5 Ser^{568} phosphorylation is listed in PhosphositePlus (Hornbeck et al., 2015), however, the online tools could not connect this modification to a specific kinase.

The lack of phosphospecific antibodies for our candidates led us to employ a recently developed method which uses phos-tag acrylamide in order to distinguish phosphorylated from non-phosphorylated protein (Kinoshita et al., 2006). To test whether phos-tag acrylamide works efficiently in our hands we assessed the phosphorylation status of SMAD2 in hESCs. The pluripotent state is supposed to be

associated with elevated pSMAD2 while incubation of the cells with a TGFβ/ACTIVIN/NODAL inhibitor (SB431542) is known to abolish SMAD2 phosphorylation (James et al., 2005).

For this reason, H1 cells were cultured until they reached 50-60% confluence and then were then treated with DMSO or SB431542 (10μM) for 3 hrs. Cells were washed in Tris-HCL buffer and lysed in Tris-HCL based buffer. Lysates were run on phos-Tag containing gels as well as conventional acrylamide gels, transferred to membranes and both membranes were probed with an antibody recognising Total SMAD2,3 and ACTIN. The sample run on a conventional acrylamide gel showed one strong band detected by the Total SMAD2,3 antibody, this strong band is TOTAL SMAD2, and a much weaker lower band corresponds to Total SMAD3 (as expected). The phos-tag gel showed the presence of two strong bands in the phos-tag gels, thus implying that the upper band is phosphorylated SMAD2. For quantification, the intensity of the upper band (pSMAD2) was divided by the sum of the intensity of the upper and lower band (Total SMAD2,3) while the obtained ratio was divided by the respective actin intensity (relative abundance). Although the decrease in pSMAD2 is not markedly strong **(Fig. 4-7C)**, the result is supportive of the notion that phos-tag is efficient in separating phosphoproteins from their non-phosphorylated counterparts.

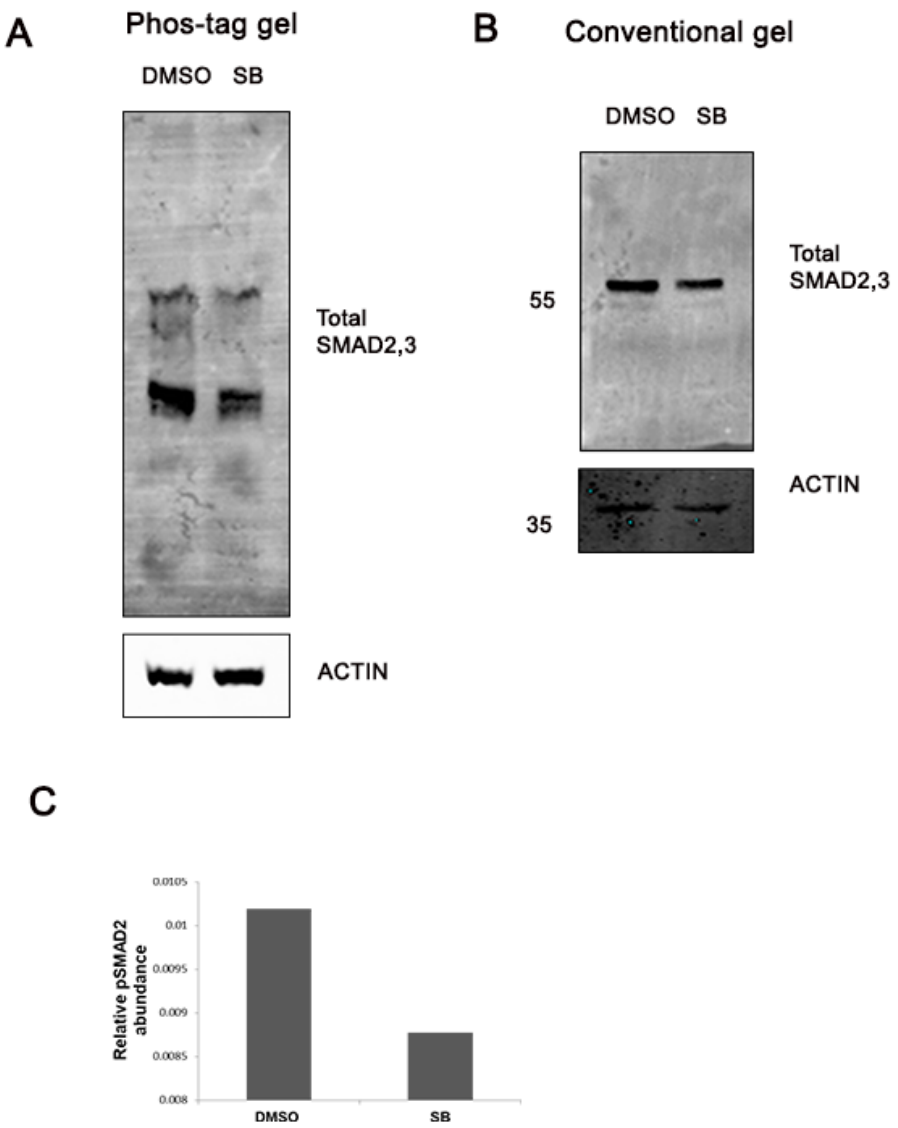


Figure 4-7 Evaluating SMAD2 phosphorylation using Phos-Tag gels.

H1 cells (50-60% confluent) were treated with DMSO or SB431542 (10μM) for 3 hrs. Cells were lysed in Tris-HCL buffer and separated on phos-tag containing gels and conventional gels. Membranes were immunoblotted against total SMAD2,3 and ACTIN. **A.** Western blot showing total SMAD2,3 and ACTIN in vehicle treated and inhibitor treated H1 cells in phos-tag gels. **B.** Western blot showing total SMAD2,3 and ACTIN in vehicle treated and inhibitor treated H1 cells in conventional gel. **C.** Plot showing relative pSMAD2 abundance in phos-tag gel (N=1).

Next, H1 cells were cultured in customised mTeSR media using FGF2/ACTIVIN A to sustain pluripotency as described in SILAC setup. H1 cells were either left untreated or induced with BMP4 for 1 hour. Cells were washed in Tris-HCL buffer and lysed in a Tris-HCL based buffer. Lysates were loaded on phos-tag acrylamide gels and conventional gels and immunoblotted against TCEA1 and ACTIN in the one case and PLEKHA5 and ACTIN in the other case.

Several additional bands were identified higher than the original TCEA1 band **(Fig.4-8B)** in phos tag gel **(Fig. 4-8A)** thus indicating the presence of more than one phosphorylation event. Phos tag gels did not clearly show a band differentially expressed between control and induced cells. Therefore, we decided to take into consideration all additional bands. For quantification, the intensity of the extra bands (pTCEA1) was divided by the intensity of the sum of the extra bands and the original band (total TCEA1) and the obtained ratio was divided by the respective ACTIN intensity (relative abundance). Comparing the two conditions did not result in a statistically significant difference **(Fig.4-8C)** thus implying that pTCEA1 is not found upregulated by BMP4.

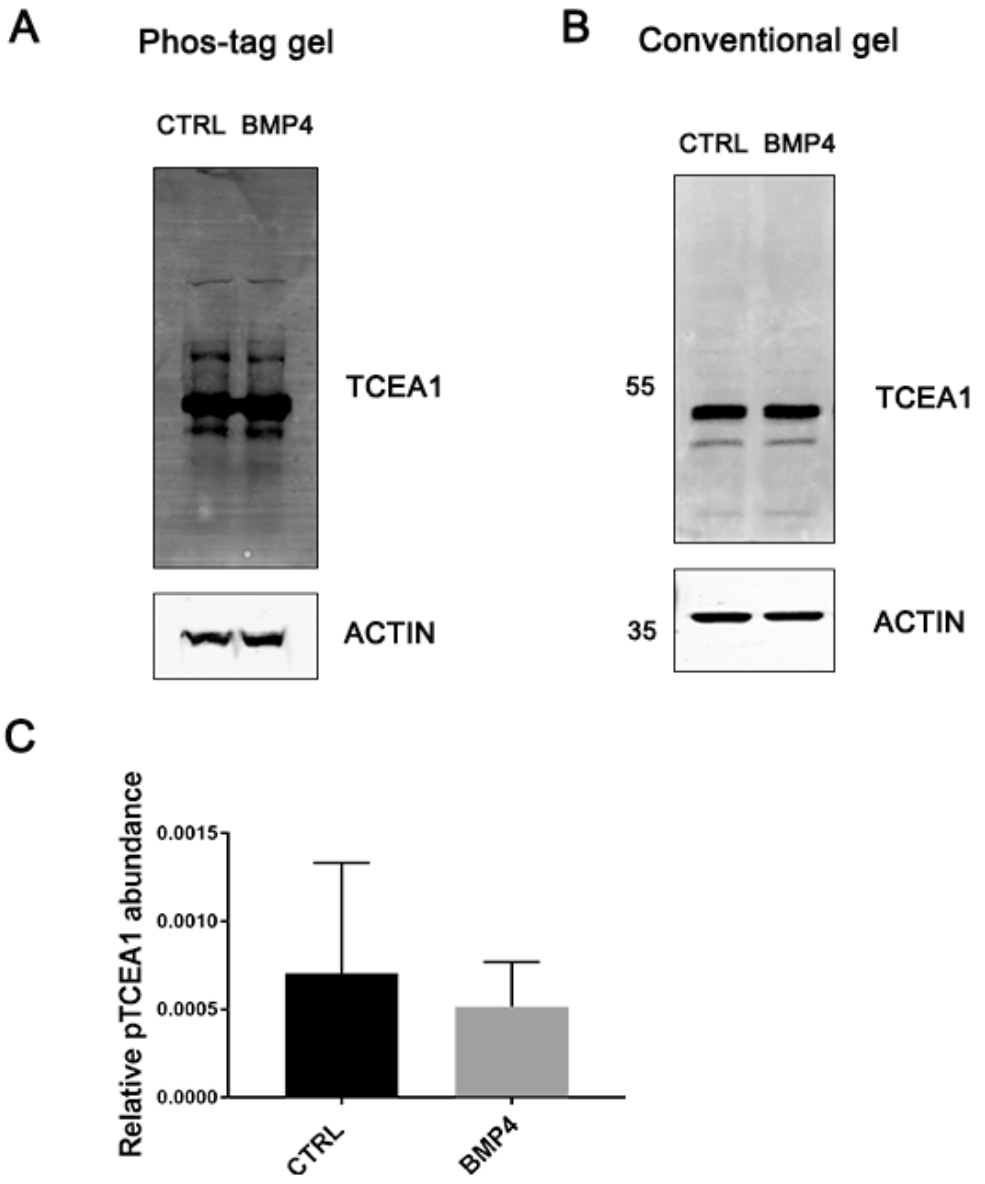


Figure 4-8 Evaluating TCEA1 phosphorylation using Phos-Tag gels.

H1 cells were left untreated or induced with BMP4 (50ng/ml) for 1hr. Cells were washed in Tris-HCL buffer and lysed in Tris-HCL-based buffer. Lysates were loaded on phostag gels and conventional gels and immunoblotted against TCEA1 and ACTIN. **A.** Western blot showing TCEA1 and ACTIN in control and induced H1 cells in phos-tag gels. **B.** Western blot showing TCEA1 and ACTIN in control and induced H1 cells in conventional gels. **C.** Plot showing relative pTCEA1 abundance in phos-tag gel. Statistical significance was tested using unpaired t-test in Graph Pad (N=3, *: p-value<0.05, ***: p-value<0.01).

Similar results were found in the case of PLEKHA5. Phos-tag gels revealed the presence of 3 bands **(Fig. 4-9A)** migrating slower than PLEKHA5 compared to the conventional gel **(Fig. 4-9B)** which most likely represents phosphorylated PLEKHA5. Since no band seemed to be differentially regulated in one of the two conditions, they were all used for quantification. More specifically, the intensity of the extra bands (pPLEKHA5) was divided by the intensity of the sum of the extra bands and the original band (total PLEKHA5) and the obtained ratio was divided by the respective ACTIN intensity (relative abundance). Comparing the two conditions did not result in a statistically significant difference **(Fig.4-9C)** thus implying that pTCEA1 is not upregulated by BMP4.

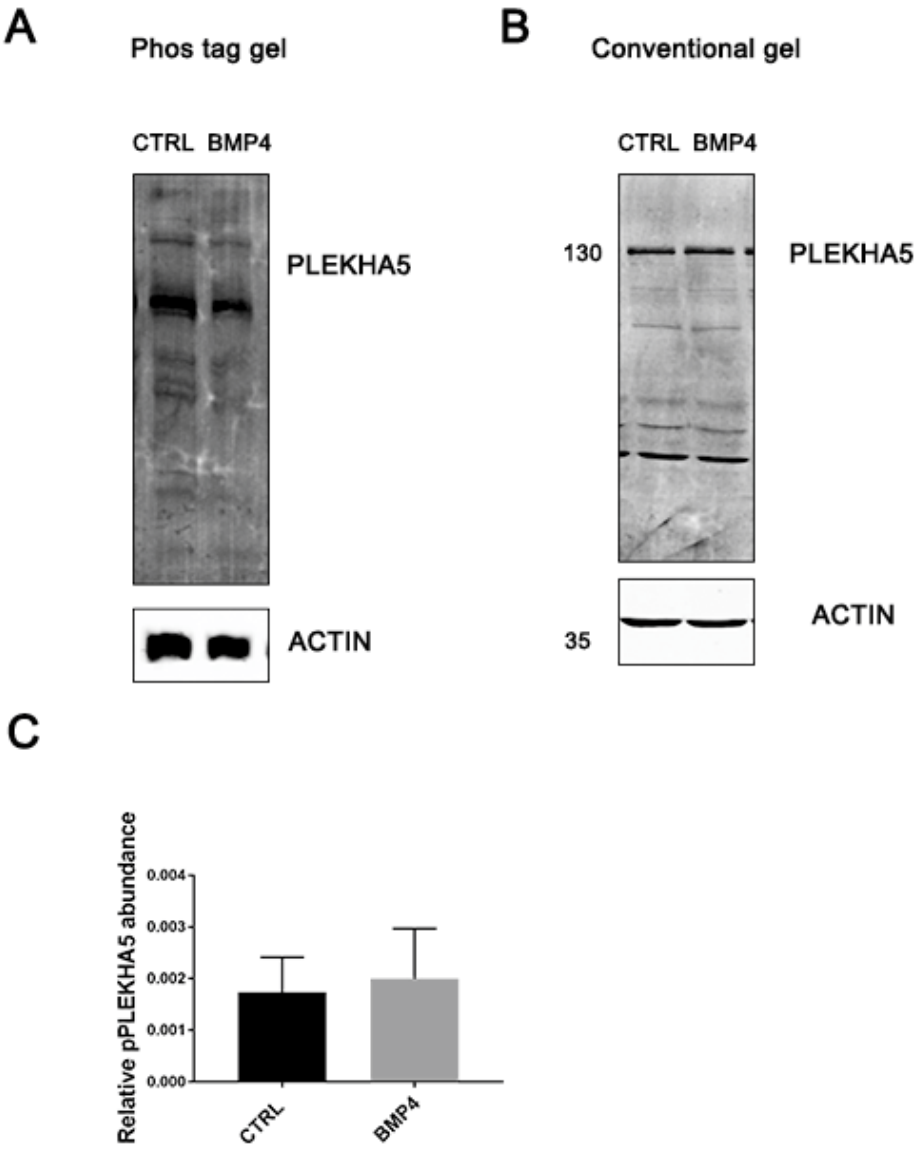


Figure 4-9 Evaluating PLEKHA5 phosphorylation using Phos-Tag gels.

H1 cells were left untreated or induced with BMP4 (50ng/ml) for 1hr. Cells were washed in Tris-HCL buffer and lysed in Tris-HCL-based buffer. Lysates were loaded on phostag gels and conventional gels and immunoblotted against PLEKHA5 and ACTIN. **A.** Western blot showing PLEKHA5 and ACTIN in control and induced H1 cells in phos-tag gels. **B.** Western blot showing PLEKHA5 and ACTIN in control and induced H1 cells in conventional gels. **C.** Plot showing relative pPLEKHA5 abundance in phos-tag gel. Statistical significance was tested using unpaired t-test in Graph Pad (N=3, *: p-value<0.05, ***: p-value<0.01).

4.3 Discussion

In this chapter it is shown that SILAC combined with MS is a useful technique that can be used in order to investigate phosphorylation events in hESCs. This thesis investigates several aspects of BMP4 effects on hESCs and this chapter in particular focuses in revealing phosphoproteome changes. It is noted that the number of regulated phosphosites is relatively small and the reproducibility is low. However, comparing upregulated phosphosites in our dataset with a significantly larger published dataset revealed a considerable percentage of overlap. This enhances data credibility and implies that performing SILAC experiments in a large scale is more likely to provide statistically significant results.

We identified upregulated phosphosites in proteins associated with migration. This result is in accordance with recent reports which show that the exposure of hESCs to BMP4 initiates EMT events, alters the expression pattern of E-CADHERIN and confers upon them migrative capacity (Richter et al., 2014). Besides, BMP4 is associated with EMT events in embryonic development such as gastrulation (Kishigami & Mishina, 2005). Meanwhile, in premigratory NC cells BMP4 triggers WNT1 expression and both growth factors seem to regulate an EMT machinery which leads to disruption of cellular contacts and migration (Theveneau & Mayor, 2012). Similarly, BMP4 seems to promote motility in cancer cells. In particular, BMP4 triggers EMT in ovarian (Guo et al., 2012) and pancreatic cancer cell lines (Hamada et al., 2007). In addition, the ability of BMP4 to promote both migration and invasion of breast cancer cells implies that this growth factor can direct metastasis in breast cancer (Guo et al., 2012).

Data analysis also associates BMP4 with transcriptional regulation and this sounds reasonable enough since it is known that hESCs rapidly change gene expression upon differentiation (Pan et al., 2007). Among the downregulated candidates found in this dataset were HDAC1 Ser^{393} and HDAC2 Ser^{394} and surprisingly, HDAC1 and HDAC3 seem to be involved in mESC mesodermal differentiation (Lv et al., 2014). More specifically, both HDACs can interact with BRACHYURY and knockdown of either HDAC1 or HDAC3 increases mesodermal differentiation efficiency (Lv et al., 2014).

Moreover, it is seen that several pluripotency associated proteins as well as WNT signalling components are phosphorylated in hESCs, thus demonstrating the importance of phosphorylation in pluripotency and pathway regulation. Besides, it is already suggested that networks governing pluripotency and differentiation are amenable to phosphorylation-driven regulation (Tobe et al., 2012).

A novel aspect of this chapter is the employment of phos-tag gels in the evaluation of phosphorylation status of proteins in hESCs. Findings suggest that phos-tag can be used to effectively separate various low and high molecular weight proteins from their phosphorylated counterparts. The presence of more than one phosphorylated bands for TCEA1 and PLEKHA5 indicates that at least in H1 cells both proteins exist in more than one phosphorylated forms. This is in agreement with PhosphositePlus (Hornbeck et al., 2015) where both proteins are found to contain several phosphorylation sites. Despite the effectiveness of phos tag gels, neither of the proteins was found to be differentially regulated upon BMP4 treatment. A possible explanation for this result is the notion that MS is a more sensitive technique enabling

the identification of subtle changes in the abundance of phosphopeptides which are not easily seen via western blot approaches.

Except for TCEA1 and PLEKHA5, two more candidates were selected for evaluation using phos-tag gels. These were BMP2K and Misshapen-like kinase-1 (MINK1). MINK1 is a Misshapen (MSN) mammalian orthologue which can phosphorylate the a-helix 1 region of SMADs (Kaneko et al., 2011). This modification is linked to suppression of TGFβ/BMP signalling (Kaneko et al., 2011), thus implying that in differentiating H1 cells BMP4 could orchestrate a negative feedback in order to cease further BMP4 induced response. Moreover, MINK1 was reported to phosphorylate a cell polarity component named PRICKLE (PK) (Daulat et al., 2012). This modification is required for proper (punctate) localization of PK to the plasma membrane in gastrulating Xenopus embryos (Daulat et al., 2012). Nevertheless, the antibodies recognising these proteins did not perform well neither in normal western blot, nor in Immunoprecipitation protocols **(Fig. A-7,8)**, thus making it difficult to decipher whether these candidates are regulated.

5. ROLE OF SARA IN HUMAN EMBRYONIC STEM CELLS

5.1 Introduction

According to the NCBI (Geer et al., 2010), the official gene name for SARA is zfyve9. It localizes on chromosome 1 and pseudogenes are found on chromosome 2, 15 and X (Geer et al., 2010). The genome browser (Kent et al., 2002) reveals one Transcription Start Site (TSS) for SARA **(Fig. 5-1B)** and Ensembl gene predictions (UCSC Genome Browser on human Feb.2009 GRCh37/hg19 assembly) suggest that there are several transcript variants **(Fig. 5-1A)**. Similar indications are found in NCBI (Geer et al., 2010) which describes two protein coding and two predicted transcripts. Briefly, transcript variant 3 (NM_004799.3) contains 19 exons and encodes for the longest isoform (1425 amino acids) which has a molecular weight of 156kDa. Transcript variant 1 (NM_007324.3) lacks exon 6, and encodes for isoform1 which is identical to isoform3 except for the fact that it lacks the SBD and therefore consists of 1366 amino acids which correspond to a molecular weight of 150kDa. Additionally, there are two predicted variants. XM_017002845.1, has exons 1-13, and in principle should encode for X2 (128kDa). XM_011542437.2 lacks a part of the 1st exon, which is non-coding and therefore results in the same protein as isoform3. (https://www.ncbi.nlm.nih.gov/gene/9372).

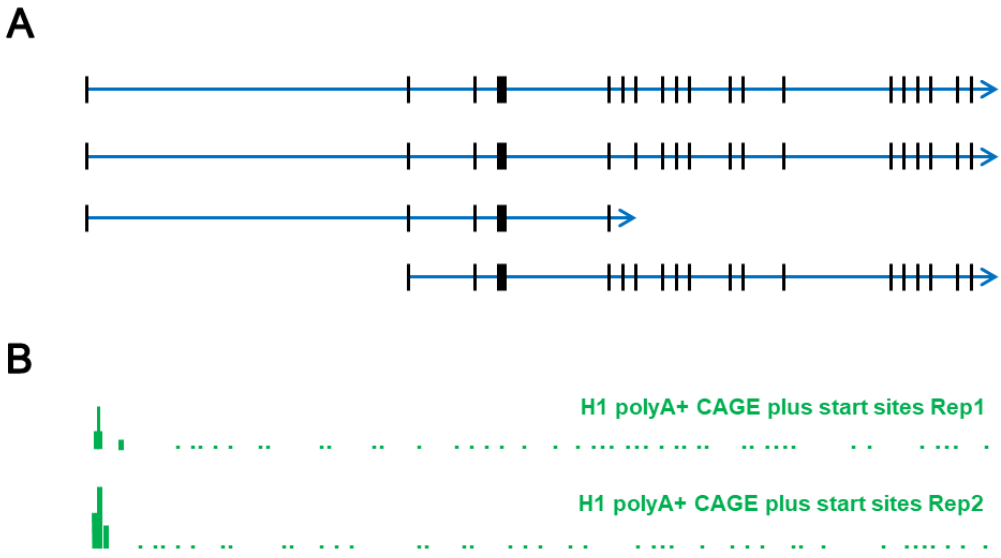


Figure 5-1 SARA transcript variants and Transcription Start Site.

A. The gene predictions according to the ENSEMBL database reveal 4 transcript variants for the zfyve9 gene. The first two are also found in the NCBI database as protein coding variants. The first contains 19 exons and encodes for the full length protein while the second lacks exon 6 which contains the SBD and encodes the second isoform. **B.** CAGE results from H1 cells, available in the Genome Browser webpage show the presence of a single peak at the beginning of the 1st exon. This peak is found in two replicates and stands for the starting site of transcription.

Initially, researchers observed that human SARA existed as three or more isoforms (Meckelein et al., 1998). Mice are predicted to have three SARA isoforms, the first being the full length, the second lacking the SBD and the third lacking the FYVE domain (Chang et al., 2014). The first two isoforms seem to be expressed in all tested mice tissues while the third was not found (Chang et al., 2014). It is not clear whether the same applies for human SARA as well. Wrana's group first suggested that human SARA consists of 1323 amino acids (Tsukazaki et al., 1998). As a result,

several studies in the literature have used tagged versions of SARA and SARA mutants which are slightly shorter than the authentic full length protein.

Transiently transfected SARA seems to be confined to the cytoplasm in a punctate pattern (Tsukazaki et al., 1998). However, mutated SARA versions which lack the FYVE domain are found both in the cytoplasm and the nucleus (Tsukazaki et al., 1998). Recent findings suggest that several proteins with a primary role in endocytosis and sorting can also exert a regulatory role in the nucleus (Pyrzynska *et al.,* 2009). Although, full length SARA or possible isoforms have not been detected in the nucleus it would be interesting to know whether SARA can have a broader role which is not limited on the endosomal membranes.

Early observations showed that SARA is not equally expressed among cells of different origin (Runyan et al., 2009). More specifically, low SARA levels are found in fibroblasts as opposed to epithelial cells (Runyan et al., 2009). This prompted the notion that SARA could be reduced in EMT and indeed SARA downregulation is a vital EMT step for renal tubular epithelial cells (Runyan et al., 2009).

EMT is a critical step during gastrulation (Nakaya & Sheng, 2008). Several reports suggested that hESCs differentiate displaying EMT features. More specifically, hESCs cultured without feeders gradually show inhomogeneity within the colony since peripheral cells start resembling mesenchymal cells in terms of morphology and marker expression (Ullmann et al., 2007). Moreover, hESCs seem to undergo EMT spontaneously when feeders are removed from the culture (Eastham et al., 2007). In fact, researchers propose the idea that hESCs cultures could be a valuable tool for the study of EMT events (Eastham et al., 2007). Furthermore, ME differentiation of hESCs can be initiated by the combined action of growth factors

(FGF2, VEGF, BMP4 and short ACTIVIN A induction) and this commitment alters gene expression reflecting EMT (Evseenko et al., 2010). In addition to the above, the exposure of hESC colonies to BMP4 triggers both mesodermal differentiation and EMT like events (Richter et al., 2014).

On the other hand, mesenchymal to epithelial transition (MET) seems to occur during reprogramming (Samavarchi-Tehrani et al., 2010), (R. Li et al., 2010). The generation of mouse iPSCs from mouse fibroblasts encompasses three stages and the initial one reflects MET in terms of gene expression (Samavarchi-Tehrani et al., 2010). Interestingly, this step is associated with BMP-induced microRNAs which seem to regulate the process (Samavarchi-Tehrani et al., 2010). Reprogramming relies on MET while at the same time EMT is blocked (R. Li et al., 2010). The transcription factors used to erase the differentiated state seem to regulate distinct signals which either promote or impair reprogramming (R. Li et al., 2010). Among their targets seem to be TGF-β1, TGFβ Type II receptor, Snail and E-cadherin (R. Li et al., 2010).

There are several tools which enable altering the levels of a protein in order to study its function. Briefly, short interfering RNAs (siRNAs) are double stranded RNA molecules which can silence homologous genes (McManus & Sharp, 2002). However, their use is accompanied by several drawbacks which impair their suitability for studying phenotypes (Gaj et al., 2013). SiRNAs decrease gene expression at variable levels but cannot completely abolish it (Gaj et al., 2013). Moreover, they only transiently regulate gene expression and have unknown effects on non-target genes (Gaj et al., 2013).

‘Genome editing’ is another valuable strategy which uses targeted endonucleases in order to generate double strand breaks (DSBs) in a genetic location (Urnov et al.,

2010). Zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) are frequently used in genome editing and they can be instructed to identify any desirable target across the genome (Gaj et al., 2013). Both ZFNs and TALENs use proteins to bind DNA, however, another recently discovered technique CRISPR-Cas, is targeted to the DNA via RNA molecules (Ran et al., 2013). This technique is given an acronym meaning Clustered, Regularly Interspaced, short Palindromic Repeats/Crispr associated (Waddington et al., 2016). The original CRISPR/Cas machinery was identified as a protection mechanism utilised by bacteria and archae against intruders (Waddington et al., 2016).

The CRISPR locus consists of short sequences separated by sequences known as 'spacers' and has the ability to integrate foreign DNA first and then target it for destruction (Waddington et al., 2016). Cas genes are located next to the CRISPR locus and play an auxiliary role in the establishment of CRISPR mediated immunity (Waddington et al., 2016). More specifically, when a pathogen invades the cell, cas genes recognize and cleave the foreign DNA (protospacer), next this DNA is integrated in the CRISPR locus where it serves as a spacer (Waddington et al., 2016). The recognition of specific DNA sequences relies on the presence of motifs next to them which are known as Protospacer adjacent motifs (PAMs) (Waddington et al., 2016). Once transcribed, the CRISPR locus generates the precursor CRISPR RNA (pre-crRNA) which is further processed to crRNA (Waddington et al., 2016). Finally, this molecule cooperates with cas proteins and mediates recognition and destruction of invading DNA (Waddington et al., 2016).

There are three known CRISPR systems and type II is the one best studied (Ran et al., 2013). It has three components, namely the nuclease Cas9, the crRNA and the

transactivating crRNA (trcrRNA) which mediates crRNA maturation (Ran et al., 2013). Mature crRNAs contain 20nucleotides and can guide Cas9 to the genome taking advantage of the Watson-Crick pairing (Ran et al., 2013). Wild type Cas9 causes DSBs while a Cas9 mutant gives the opportunity to generate single strand breaks (Ran et al., 2013). Once the DSB occurs, the cell initiates one of the two known repair processes depending on whether there is a repair template or not (Ran et al., 2013). In the first case, as long as the cell divides it can use the homology directed repair (HDR) which is generally less frequent and more accurate. In the second case, the cell uses the non-homologous end joining mechanism (NHEJ) which inevitably generates insertions and deletions (Ran et al., 2013).

The aims of this chapter are the following: First to address whether SARA exists in one or more isoforms in human cell lines and the possibility that SARA localizes in the nucleus. Second, to reveal whether SARA is regulated in mesodermal differentiation of hESCs and whether it is differentially expressed before and after nuclear reprogramming. Third, to attempt to knock out SARA in H1 cells using the novel CRISPR approach.

5.2 Results

5.2.1 SARA isoforms in human cell lines

In order to address the question of how many SARA isoforms can be detected in human cell lines and at what molecular weight, several cell lines were cultured for at least two passages after thawing and before reaching confluence they were lysed in 1% SDS lysis buffer. Cell lysates were run on an SDS-PAGE gel and immunoblotted with anti-SARA and anti-ACTIN antibodies. All the cell lines tested showed the

presence of two bands between 250kDa and 130kDa, while in some cases the presence of additional bands with lower molecular weight were detected **(Fig. 5-2)**.

Full length SARA has a predicted molecular weight of 156kDa and we considered that band2 would most likely correspond to the full length. SARA was proposed to be N-glycosylated when first described and this could increase the molecular weight (Meckelein et al., 1998). Therefore, we considered that this might explain why band1 detected using the anti-SARA antibody and running higher than the predicted mass might indeed be a post-translationally modified version of the full length protein. The presence of several other bands in HEK-293 raised the possibility that some if not all could be either isoforms or cleavage products. Lastly, we needed to check the specificity of the SARA antibody to ensure that all bands were indeed specific.

Therefore, HEK-293A cells were subjected to transient siRNA experiments. Control cells were transfected with scrambled siRNA while cells on separate dishes were transfected with siRNA targeting a part of the 4th exon of SARA which is present in all predicted isoforms. 72hrs following transfection, cells were lysed and the lysates were run on an SDS-PAGE gel and immunoblotted with anti-SARA and anti-ACTIN antibodies. Five bands (red arrows) are recognised in control cells **(Fig. 5-3A)** and two of them are significantly decreased in the siSARA population **(Fig. 5-3B)**. The first band (1st red arrow) is located at approximately 180kDa and most likely is the full length protein while the second band (4th red arrow) is found at approximately 100kDa.

Therefore, from the siRNA experiment we conclude that we have 2 SARA proteins specifically recognised by our antibody at approx. 180 KDa (Band 1) and at 100 KDa (Band 4). Our original thought that Band 2 was full length SARA and Band 1 a post-

translational modification of Band 1 was incorrect. Therefore, it seemed that Band 1 is full length SARA. Moreover, Band4 could be the predicted X2 isoform, an unknown isoform or a cleaved product of the full length protein.

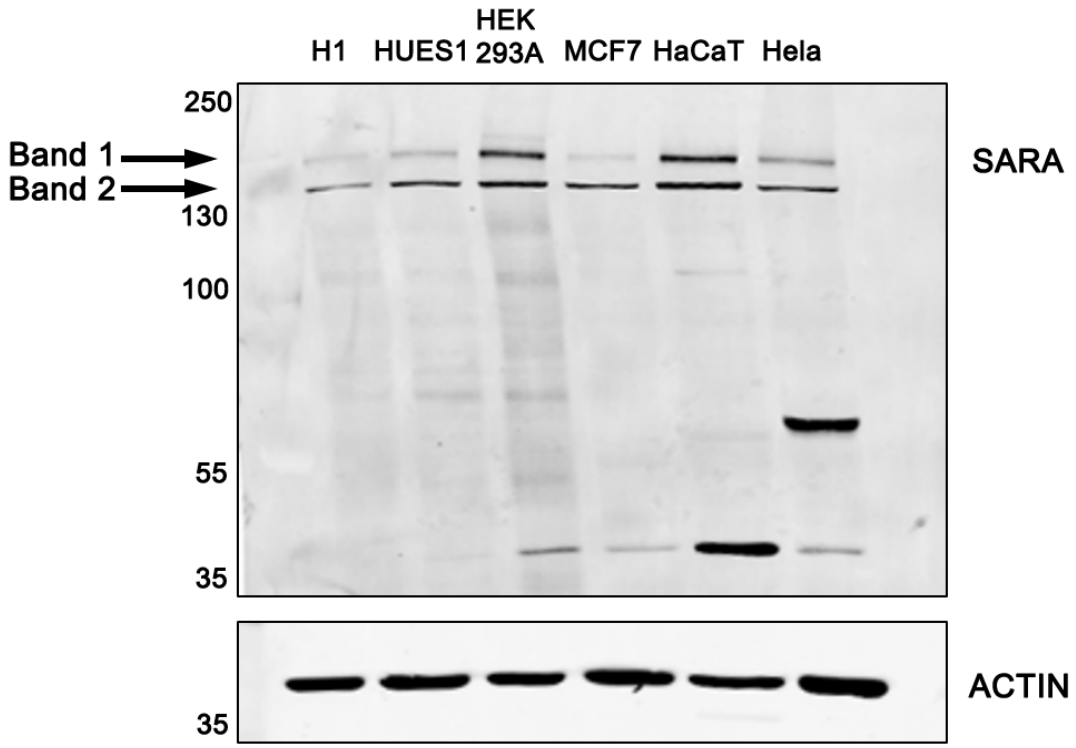


Figure 5-2 SARA expression in different cell lines.

H1, HUES1, HEK-293A, MCF7, HaCaT and Hela cells were seeded on 6well dishes and before reaching confluence they were lysed in 1%SDS. Lysates were run on SDS-PAGE gel and Immunoblotted with anti-SARA and anti-ACTIN antibodies. All cell lines seem to express two or more bands detected by the antibody. (N=1).

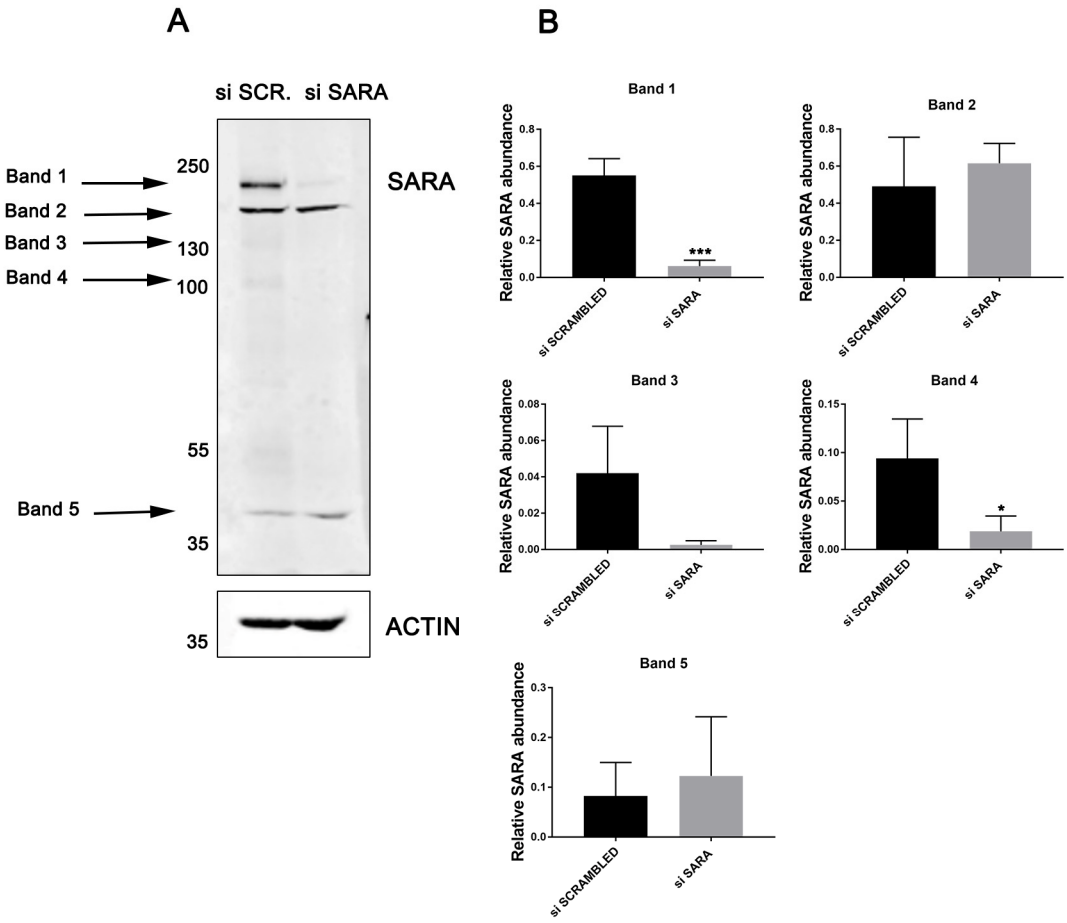


Figure 5-3 SARA knockdown in HEK-293A cells.

HEK-293A cells were seeded on two separate dishes. When they reached 70% confluence they were transfected with either scrambled siRNA or an siRNA targeting all SARA isoforms. Next day media were changed and cells were incubated for 72hrs in total. Cells were then lysed in 1% SDS and lysates were run on an SDS-PAGE gel. **A.** Western blot showing SARA in cells treated with si scrambled and cells treated with si SARA **B.** Plots show the normalised ratios of SARA intensity between the two conditions. Each plot accounts for a distinct band and the number of the band is given arbitrarily according to their molecular weight, Band 1 being the highest and Band 5 being the lowest. Statistical significance was tested using unpaired t-test in Graph Pad (N=3, *: p-value<0.05, ***: p-value<0.01).

An additional approach was also used in order to estimate the correct molecular weight of the full length protein as well as to check whether SARA isoform2 which lacks the SBD could be one of the observed bands. For this reason a GFP-SARA construct was generated. Briefly, the coding sequence of SARA was purchased (abm, ORF011808) and cloned into a GFP expressing vector kindly provided by Dr. Ivana Barbaric. This vector drives the expression of the protein of interest under the control of the chicken β-actin promoter (ACTB). ACTB was chosen because it is considered to be one of the three best constitutive promoters for hESCs together with the elongation factor-1a (EF1a) and phosphoglycerate kinase (PGK) (Norrman et al., 2010). Moreover the vector allows for puromycin selection since the same promoter drives the expression of the puromycin resistance gene. Puromycin resistance and SARA are separately expressed thanks to the presence of an Internal Ribosome Entry Site (IRES2). All these features would be helpful for the generation of stable cell lines in H1 cells.

Additionally, two more plasmids were kindly provided by Wrana's lab. The first plasmid encodes the full length SARA tagged with FLAG at the N-terminus and the second is a SARA mutant plasmid, which has the same tag as the full length but lacks the SBD. It should be noted that both plasmids lack the first 102 nucleotides and are therefore expected to migrate slightly faster than the actual isoforms. Next, HEK-293A cells were seeded on 6well dishes and separately transfected with the following plasmids: GFP-SARA, GFP control vector, FLAG-SARA and FLAG-SARA Δ(SBD). Control cells were left untreated to check endogenous SARA. Cells were lysed 36hrs after transfection and lysates were run on an SDS-PAGE gel. Lysates were loaded three times and membranes were separately immunoblotted with anti-

GFP, anti-SARA and anti-FLAG antibodies. All membranes were immunoblotted anti-ACTIN antibody.

The results show that GFP-SARA runs at approximately 250kDa **(Fig. 5-4A)** and given the fact that GFP has a molecular weight of 27kDa, SARA full length should be located at around 200kDa. Moreover, FLAG-SARA and FLAG-SARA Δ(SBD), run at almost the same size **(Fig. 5-4C)** as expected, since mutant SARA in only 6kDa smaller than the full length protein.

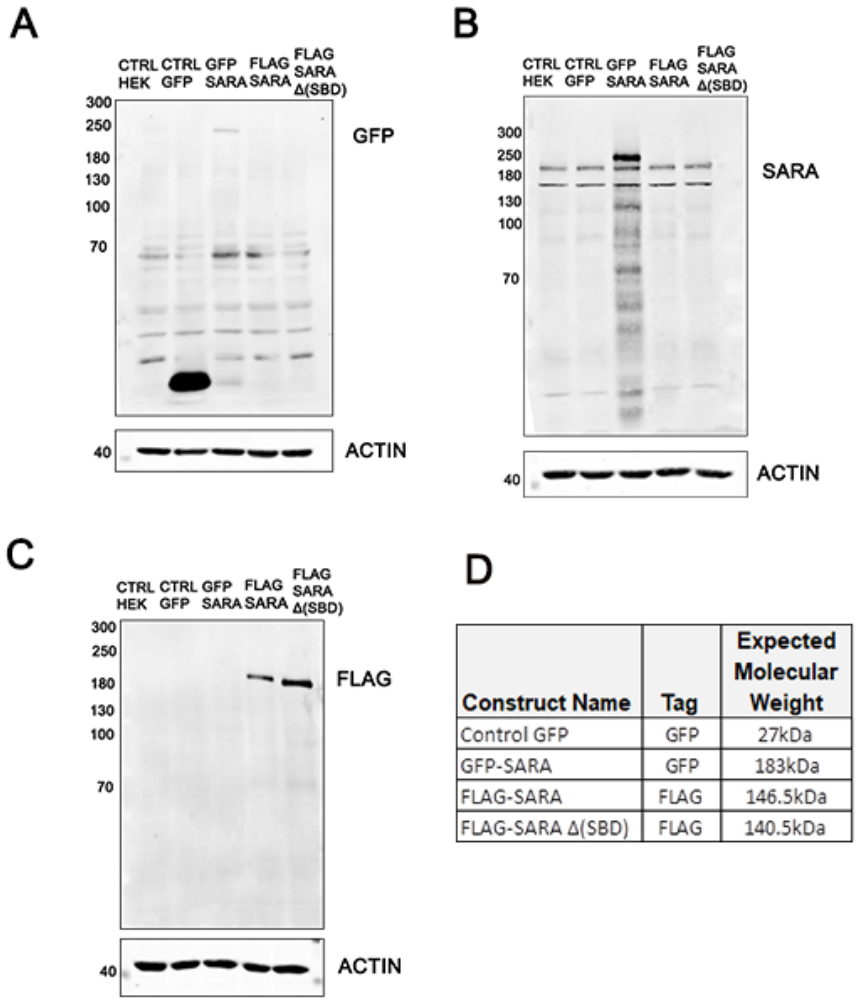
Therefore, from the siRNA experiment and the over-expression experiment we conclude:

(1) Band 1 is full length SARA

(2) Bands 2 and 3 are non-specific

(3) Band4 is specific SARA product but we cannot be sure if it represents an isoform or degradation product

(4) Band 5 is non-specific



Construct Name	Tag	Expected Molecular Weight
Control GFP	GFP	27kDa
GFP-SARA	GFP	183kDa
FLAG-SARA	FLAG	146.5kDa
FLAG-SARA Δ(SBD)	FLAG	140.5kDa

Figure 5-4 Transient expression of GFP control, GFP-SARA, FLAG-SARA and FLAG-SARA Δ(SBD).

HEK-293A cells were seeded on 6-well dishes. When they reached 50-70% confluence they were separately transfected with GFP-control, FLAG-SARA and FLAG-SARA Δ(SBD). 36hrs later, cells were lysed in 1% SDS and lysates were run on SDS-PAGE gels (N=2). **A.** The first membrane was incubated with an anti-GFP antibody. Western blot showing the size of GFP-SARA and GFP-control. **B.** The second membrane was incubated with an anti-SARA antibody. Western blot showing the size of endogenous SARA. **C.** The third membrane was incubated with an anti-FLAG antibody. Western blot showing the size of FLAG-SARA and FLAG-SARA Δ(SBD). **D.** Table shows the name of the constructs used, their tag and their expected molecular weight.

5.2.2 SARA putative nuclear localisation

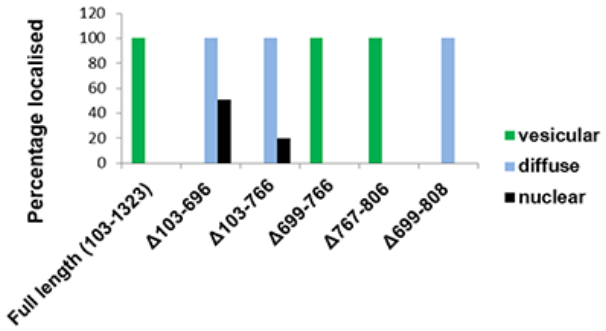
SARA sequence was imported in an online tool (cNLS Mapper) and searched for the presence of NLS (Kosugi, Hasebe, Matsumura, et al., 2009),(Kosugi et al., 2008),(Kosugi, Hasebe, Tomita, & Yanagawa, 2009). Indeed, a putative NLS was found located within the FYVE domain. This NLS begins at position 711 and has a score of 6.1. According to the details of the prediction site, scores 3 to 5 refer to proteins that localise both to the cytoplasm and the nucleus **(Fig.5-5B)**. This finding raised the notion that SARA could potentially translocate to the nucleus.

Several SARA mutants were previously shown to display nuclear localisation (Tsukazaki et al., 1998). The presence of a putative NLS within the FYVE domain, which enables endosomal binding for the protein, prompted us to investigate the localisation of several SARA mutants and whether the localisation they display correlated with the presence of the NLS. Six SARA constructs previously generated and described by Wrana's lab (Tsukazaki et al., 1998) were used to transiently transfect HEK-293A cells. HEK-293A cells were seeded on polylysine coated coverslips and when they reached approximately 50% confluence they were transfected with six SARA constructs. 36hrs later cells were fixed in 3.7% PFA and stained with an anti-FLAG antibody, recognising the FLAG tag on the SARA constructs, while HOECHST was used for nuclear staining.

Microscopy revealed that only two mutants partially localised to the nucleus. One of them SARA Δ(103-696) contains the NLS while the second one SARA Δ(103-766) does not **(Fig. 5-5A)**. The rest of the constructs are either exclusively cytoplasmic exhibiting either a vesicular or a diffuse localisation pattern **(Fig. 5-6)**. This result indicates that the presence of a putative NLS is not sufficient to cause or justify

nuclear localisation. However, an interesting finding is that the two nuclear localised constructs share a similar feature compared to the rest of the tested constructs. More specifically, they both have large deletions of the N-terminal part of the protein.

A



B

Position	Predicted NLS sequence	cNLS Mapper score
711	RFTFTKRRHHCRACGKVFCASCCS LKCKLLYMD	6.1

C

Contruct Number	Deleted area	NLS
801	Full length	yes
802	103-696	yes
803	103-766	no
805	699-766	no
806	767-806	yes
807	699-808	no

Figure 5-5 SARA mutants localisation and putative NLS site.

A. Plot showing the localisation of several SARA constructs (in percentage). Each construct has a different mutation (except for the full length). Localisation was categorised as vesicular when the staining is punctate in the cytoplasm, diffuse when it is not punctate (this can be either cytoplasmic or both cytoplasmic and nuclear) and nuclear. Approximately 100 transfected cells were counted for each construct (N=1) **B.** Table shows details regarding the putative NLS found in SARA sequence. The signal starts in position 711 and displays a medium score of 6.1, thus indicating that the protein could be present in both nucleus and cytoplasm. **C.** Table lists the construct number for SARA mutants, the deletions they harbour and the presence or absence of the NLS.

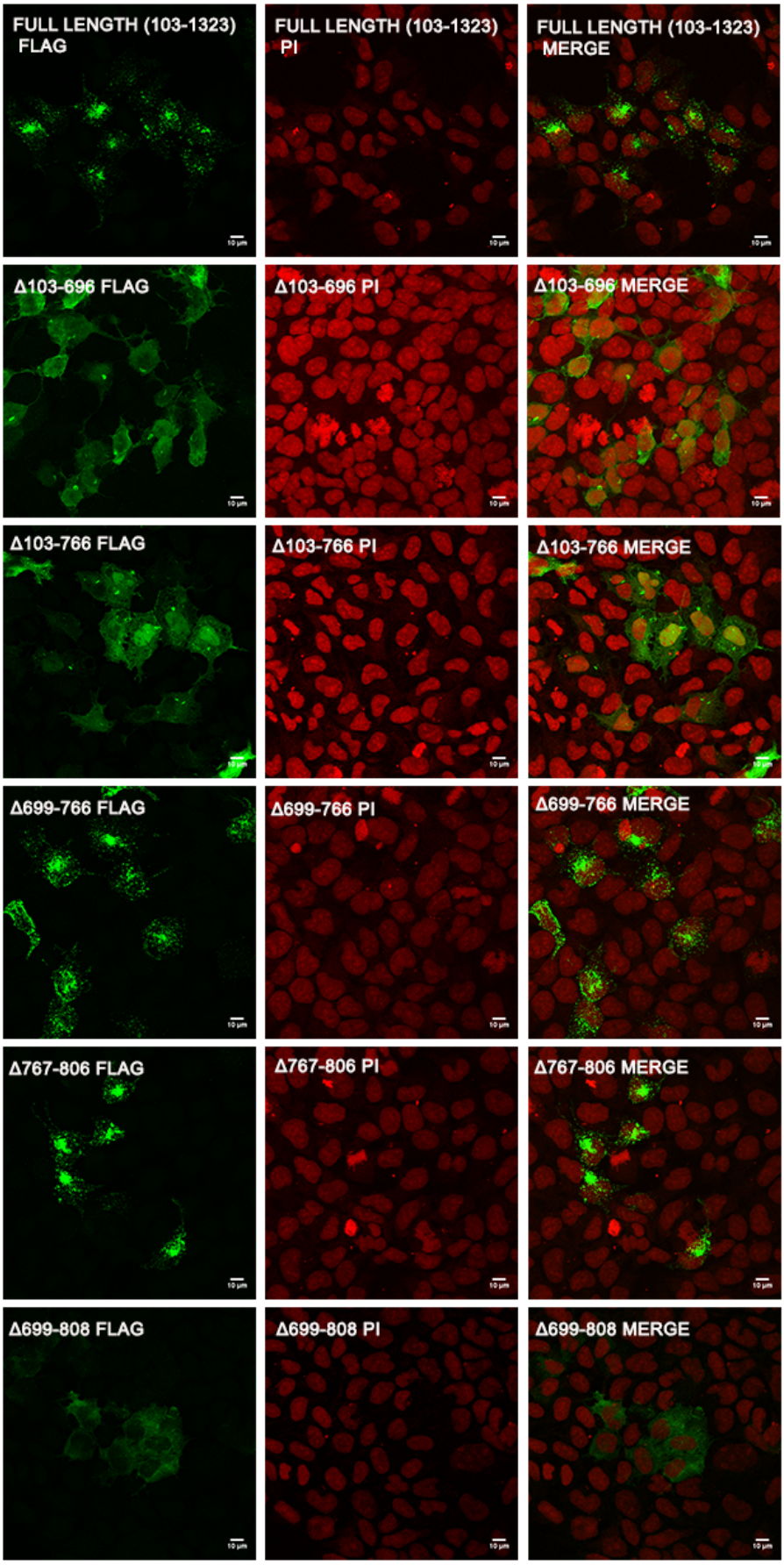

FULL LENGTH (103-1323) FLAG
FULL LENGTH (103-1323) PI
FULL LENGTH (103-1323) MERGE
Δ103-696 FLAG
Δ103-696 PI
Δ103-696 MERGE
Δ103-766 FLAG
Δ103-766 PI
Δ103-766 MERGE
Δ699-766 FLAG
Δ699-766 PI
Δ699-766 MERGE
Δ767-806 FLAG
Δ767-806 PI
Δ767-806 MERGE
Δ699-808 FLAG
Δ699-808 PI
Δ699-808 MERGE

Figure 5-6 SARA mutants' localisation.

HEK-293A cells were seeded on polylysine coated coverslips and when they reached approximately 50% confluence they were transfected with six SARA constructs. 36hrs later cells were fixed in 3.7% PFA and stained using anti-FLAG antibody, while PI was used to stain the nucleus. Approximately 100 transfected cells were imaged for every construct (N=1). The left panel shows FLAG staining; the central panel is the PI staining while the right panel is the merge of both panels. Samples were imaged using an inverted SP8 Leica confocal microscope, 40X objective with oil. Images are Z projections and the scale bar is set at 10μm.

This result can be interpreted in two ways. The straightforward interpretation is that SARA mutants do not exist and therefore their localisation could be considered an artefact which does not apply in culture. Alternatively, this result could imply that there are short SARA isoforms or that SARA under specific conditions is cleaved intracellularly and a functional part that derives from cleavage is free to enter the nucleus in order to exert novel functions. If cleavage events happen in culture, then using N-terminally tagged full length SARA constructs is not an efficient way to prove it. Therefore, C-terminally tagged constructs are needed.

For this reason, a SARA-MYC construct (kindly provided by Prof. Philip Woodman) was used to transiently transfect HEK-293A cells in order to investigate SARA localisation. Initial experiments showed that SARA-MYC displays a punctate staining without signs of presence in the nucleus **(Fig. A-6)**. Given the fact that some proteins can have a dynamic distribution and constantly shuttle between the nucleus and the cytoplasm, it seemed reasonable to test whether SARA accumulates upon inhibition of nuclear export. The majority of molecules are transported from the nucleus to the

cytoplasm with the help of a protein known as Chromosomal Maintenance 1 (CRM1) or EXPORTIN1 (Nguyen *et al,* 2012) and leptomycin B (LMB) inhibits CRM1 mediated nuclear export of proteins (Yashiroda & Yoshida, 2003). Therefore, HEK-293A cells were seeded on polylysine coated coverslips and one day later they were transfected with SARA-MYC. Next day they were treated overnight (18hrs) with either vehicle (ethanol-ETOH) or 20μM LMB. Next, cells were fixed in 3.7% PFA and stained with anti-MYC antibody, while HOECHST was used as a nuclear staining. Microscopy revealed that in both cases SARA displays punctate distribution in the cytoplasm without signs of nuclear accumulation **(Fig.5-7)**. There are two main explanations for the observed result. The simplest one is that SARA does not shuttle, at least under the conditions tested, between the two compartments. The second one is the chance that SARA uses a CRM1 independent mechanism for export and therefore LMB cannot induce accumulation in the nucleus.

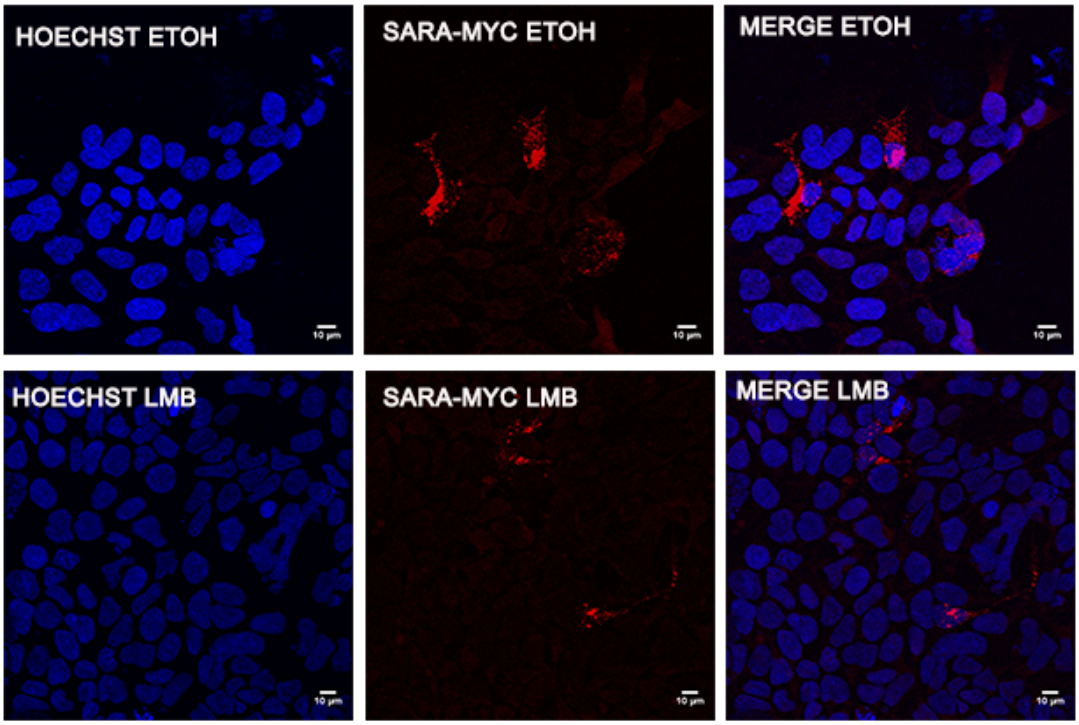


Figure 5-7 Leptomycin treatment does not alter SARA localisation.
HEK-293A cells were seeded on polylysine coated coverslips. Next day they were transfected with SARA-MYC and 24hrs later they were treated either with ethanol or with LMB overnight. Next, cells were fixed in 3.7% PFA and stained with anti-MYC antibody, while HOECHST was used for nuclear staining. Samples were imaged using an inverted SP8 Leica microscope, 40X objective with oil. Images are Z projections and the scale bar is set at 10μm. (N=3, 30-40 SARA-MYC positive cells were imaged for every replicate).

5.2.3 SARA is regulated during BMP4 induced differentiation in H1 and hiPSCs

In order to test the hypothesis that SARA is decreased during mesodermal differentiation, H1 cells were seeded on 6well dishes and two days after passaging they were induced with BMP4 (50ng/ml). Cells were lysed at 24, 48, 72 and 96hrs following the induction while control cells were left untreated and lysed at

approximately 70% confluence. Lysates were run on SDS-PAGE gel and immunoblotted with anti-SARA and anti-ACTIN antibodies. Since BMP4 can trigger several fates, lysates were also analysed for BRACHYURY expression in order to show that any observed differences occur in the context of mesodermal differentiation. According to the results, SARA (black arrow) levels remain stable for the first 24hrs of induction and then they are significantly downregulated for the next three days **(Fig. 5-8A,B)**. Approximately at the same period H1 cells upregulate BRACHYURY expression **(Fig. 5-8C).**

This finding prompted us to test whether SARA reduction is a feature unique to H1 cells or whether other pluripotent cell lines follow the same expression pattern. For this reason, hiPSCs were seeded on 6-well dishes and two days after passaging they were induced with BMP4 (50ng/ml) for 72hrs. Control cells were left untreated. Lysates were run on SDS-PAGE gel and immunoblotted with anti-SARA and anti-ACTIN antibodies and again BRACHYURY was used as a marker of efficient mesodermal differentiation. In accordance with results in H1 cells, hiPSCs present a significant decrease in SARA levels by 72hrs of differentiation **(Fig. 5-9A,C)**, while being positive for BRACHYURY **(Fig. 5-9B)**.

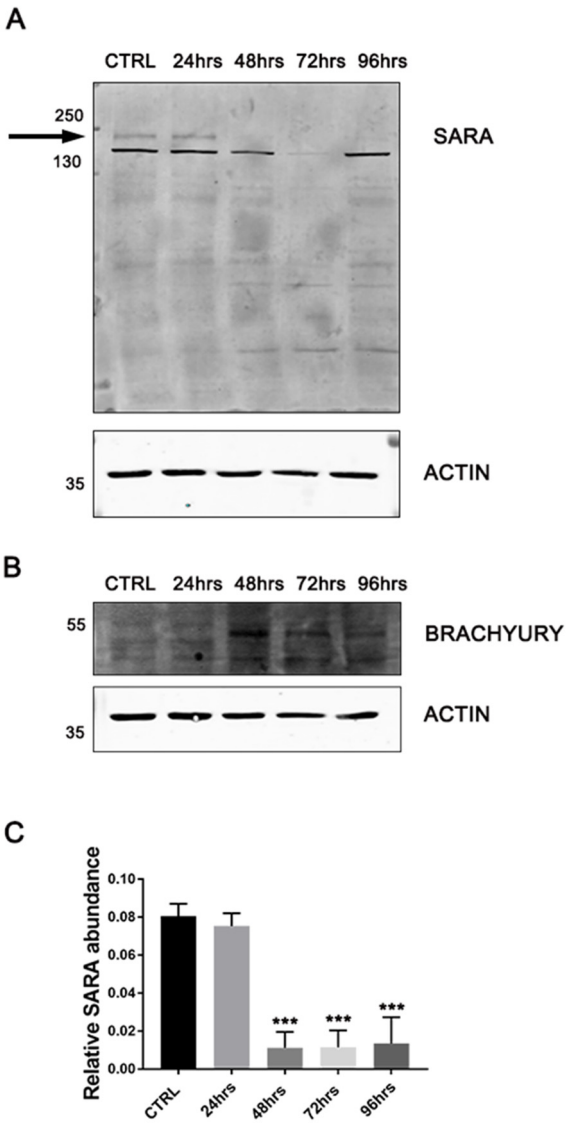


Figure 5-8 SARA regulation during BMP4 induced differentiation in H1 cells.
H1 cells were seeded on 6-well dishes and two days after passaging they were induced with BMP4 (50ng/ml) for 24, 48, 72 and 96hrs, while control cells were left untreated. Cells were lysed in 1% SDS and lysates were run on SDS-PAGE gels. **A**. Western blot showing SARA (black arrow) decrease during differentiation starting at 48hrs and continuing until 96hrs. **B.** Western blot showing BRACHYURY increase after 48hrs of differentiation. **C.** Plot showing normalised SARA intensity. Statistical significance was tested using Dunett's multiple comparisons test in Graph Pad (N=3, *: p-value<0.05, **:p-value<0.01, ***:p-value<0.001).

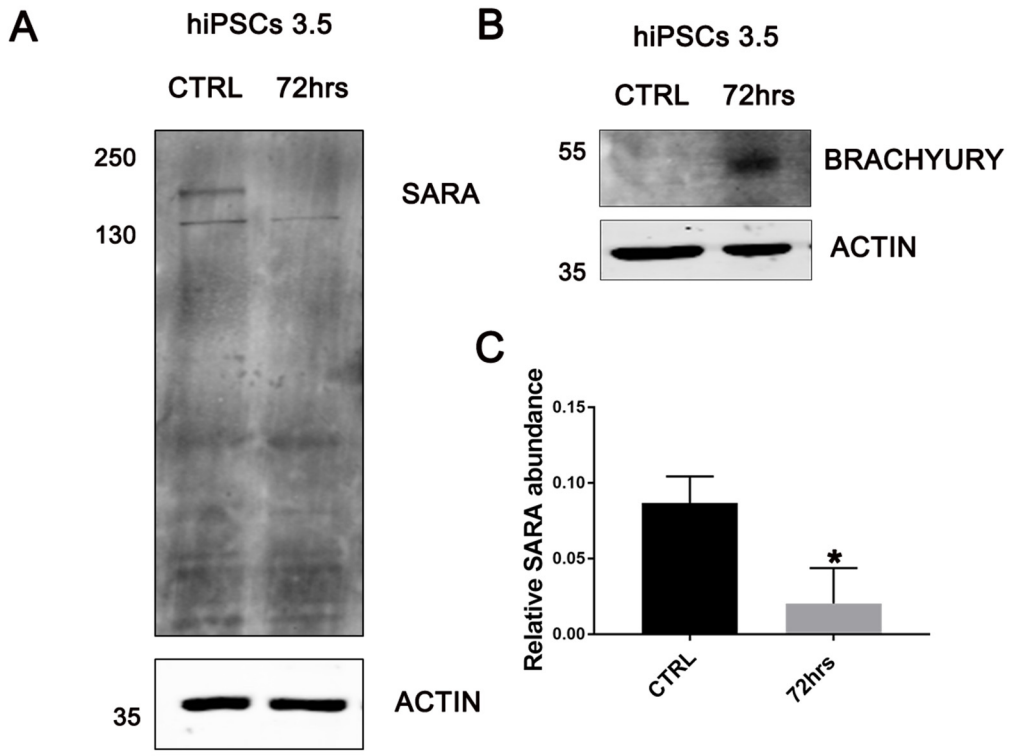


Figure 5-9 SARA regulation during BMP4 induced differentiation in hiPSCs. HiPSCs were seeded on 6-well dishes and two days after passaging they were induced with BMP4 (50ng/ml) for 72hrs while control cells were left untreated. Cells were lysed in 1% SDS and lysates were run on SDS-PAGE gel. **A.** Western blot showing SARA decrease after 72hrs of BMP4 induction. **B.** Western blot showing BRACHYURY increase after 72 hrs of BMP4 induction. **C.** Plot showing normalised SARA intensity in hiPSCs. Statistical significance was tested using unpaired t-test in Graph Pad (N=3, *: p-value<0.05, **:p-value<0.01, ***:p-value<0.001).

Given the known importance of SARA in TGFβ signalling we decided to investigate whether SARA reduction is accompanied by impaired phosphorylation of SMAD2. For this reason, the obtained lysates were run on SDS-PAGE gel and membranes were immunoblotted with anti-pSMAD2 and anti-Total SMAD2,3 antibodies. In both cases ACTIN was used as a loading control and relative pSMAD2 abundance was

calculated as a ratio of normalised pSMAD2 to normalised total SMAD2,3. According to the results, pSMAD2 is not significantly altered during differentiation **(Fig. 5-10A,B)** thus proposing that SARA might have a different role in differentiating stem cells.

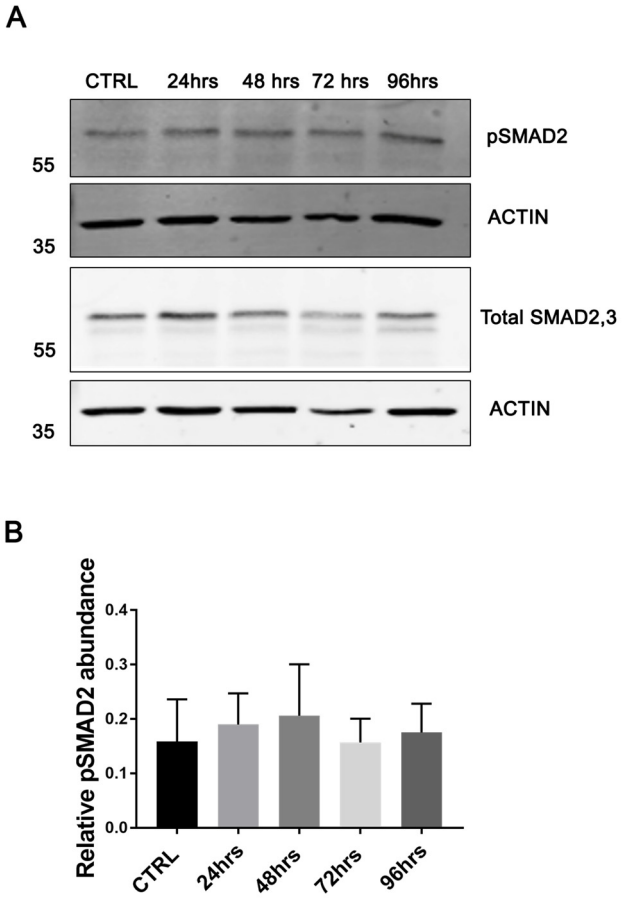


Figure 5-10 pSMAD2 levels during BMP4 induced differentiation in H1 cells.
H1 cells were seeded on 6-well dishes and two days after passaging they were induced with BMP4 (50ng/ml) for 24, 48, 72 and 96hrs, while control cells were left untreated. Cells were lysed in 1% SDS and lysates were run on SDS-PAGE gels. **A.** Western blots showing pSMAD2 and total SMAD2,3 levels during differentiation. **B.** Plot showing normalised pSMAD2 intensity during differentiation. Statistical significance was tested using Dunett's multiple comparisons test in Graph Pad (N=3, *: p-value<0.05, **:p-value<0.01, ***:p-value<0.001).

5.2.4 Elucidating the underlying mechanism for SARA decrease

The next question was which mechanism or mechanisms are used to lower SARA levels in differentiation. One possible explanation is the mRNA decrease. Therefore, H1 cells were seeded on 6-well dishes and two days after passaging they were induced with BMP4 (50ng/ml) while control cells were left untreated. RNA was extracted at 24, 48, 72 and 96 hrs following induction and control cells were lysed for RNA extraction when they reached 70% confluence. RNA was used for qRT-PCR experiments with primers against all SARA isoforms and the obtained CT values were normalised to GAPDH. According to the results, a decrease observed at 48hrs of differentiation but this was not statistically significant **(Fig. 5-11C)**. Therefore, although SARA could be regulated at the level of the transcription or before translation the data are not strong enough to prove it and imply that there is another mechanism responsible for the observed effect.

Another putative cause accounting for SARA decrease is degradation by the proteasome or the lysosome. Autophagy and the Ubiquitin Proteasome System (UPS) consist the degradation mechanisms of the cell (Korolchuk et al.,2010). In the final step of autophagy, lysosomes use their enzymes to break down molecules (Korolchuk et al.,2010). Chloroquine (CQ) was described as a lysomotropic molecule that neutralizes lysosomal pH (de Duve et al., 1974) and was found to inhibit protein degradation (Wibo & Poole, 1974). In order to study the hypothesis of lysosomal degradation, H1 cells were seeded on 6-well dishes and two days after passaging they were induced with BMP4 (50ng/ml) while control cells were left untreated. 26 hrs later cells were treated either with vehicle (H_20) or chloroquine (100μM). Cells were lysed 48hrs following BMP induction and lysates were run on SDS-PAGE gel.

Membranes were immunoblotted with anti-SARA and anti-ACTIN antibodies. According to the results SARA is significantly rescued by chloroquine treatment in differentiation **(Fig. 5-11A)**. An unexpected finding was that SARA is also significantly rescued in control cells thus implying that lysosomal degradation is routinely used for SARA destruction **(Fig. 5-11A)**. Furthermore, CQ treatment is not sufficient to restore control levels **(Fig. 5-11B)** thus implying that there is an additional mechanism leading to SARA decrease.

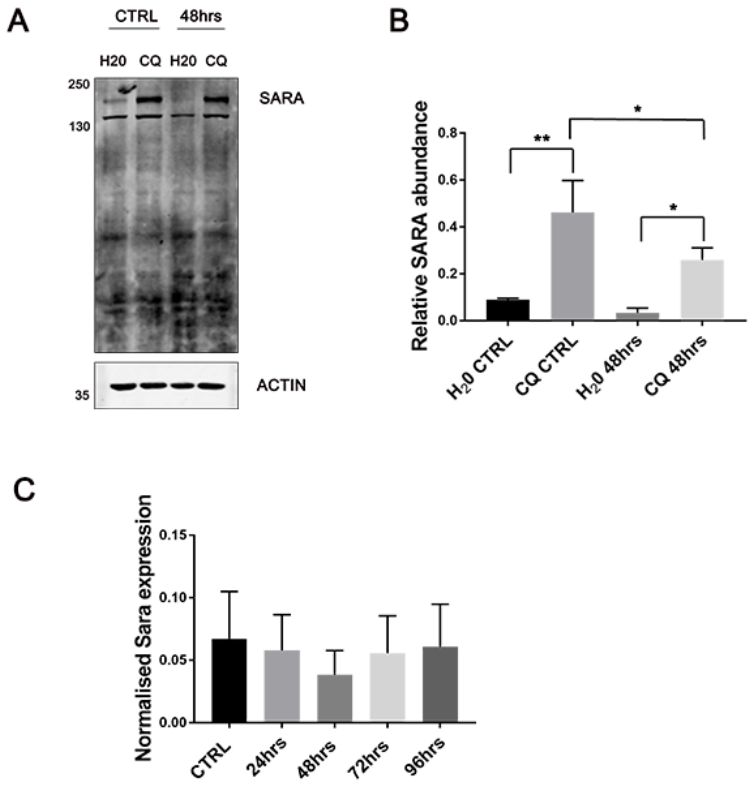


Figure 5-11 Mechanisms responsible for SARA decrease.

H1 cells were seeded on 6-well dishes and two days after passaging they were induced with BMP4. RNA was extracted 24, 48, 72 and 96hrs later for the induced cells while control cells were lysed when they reached 70% confluence. RNA was used to check mRNA levels of SARA using primers which bind all SARA isoforms. For the degradation experiments, H1 cells were seeded on 6-well dishes and two days later they were induced with BMP4 (50ng/ml). 26hrs later, cells were treated with either vehicle (H_20) or chloroquine (100μM) for 22hrs. Next, cells were lysed in 1% SDS and lysates were run on SDS-PAGE gels and immunoblotted with anti-SARA and anti-ACTIN antibodies. **A.** Western blot showing SARA rescue after chloroquine treatment in non-induced (control) H1 cells and differentiating H1 cells. **B.** Plots show normalised SARA intensity in water treated and chloroquine treated cells. Statistical significance was tested using Sidak's multiple comparisons test in Graph Pad (N=3, *: p-value<0.05, **:p-value<0.01, ***:p-value<0.001). **C.** Plot showing normalised Sara mRNA levels in differentiating H1 cells. Statistical significance was tested using Dunett's multiple comparisons test in Graph Pad (N=3, *: p-value<0.05, **:p-value<0.01, ***:p-value<0.001).

5.2.5 SARA levels before and after nuclear reprogramming

Taking into consideration the fact that SARA is decreased in EMT we hypothesized that SARA could increase in the inverse process of MET. MET is utilised at the initial stages of the reprogramming process and therefore it seemed reasonable that perhaps SARA could be one of the several proteins regulated during reprogramming. If that is true, then it is possible that SARA is differentially expressed between fibroblasts and iPSCs.

To test this hypothesis, human fibroblasts (used for the generation of hiPSCs) and the fibroblast derived hiPSCs were cultured and cells were lysed at approximately 70% confluence. Lysates were run on SDS-PAGE gel and immunoblotted with anti-SARA and anti-GAPDH antibodies. According to the results, SARA is not significantly different between the two cell types **(Fig. 5-12A,B)**. One of the factors contributing to this result could be the fact that there is variation on the levels of SARA between different passages of human fibroblasts. In fact, early passage fibroblasts have low levels of SARA while later passages show an increase. Moreover, SARA could indeed be regulated early in reprogramming but maybe later SARA levels recover. If that is the case, more detailed experiments are needed to follow SARA levels in the long reprogramming process.

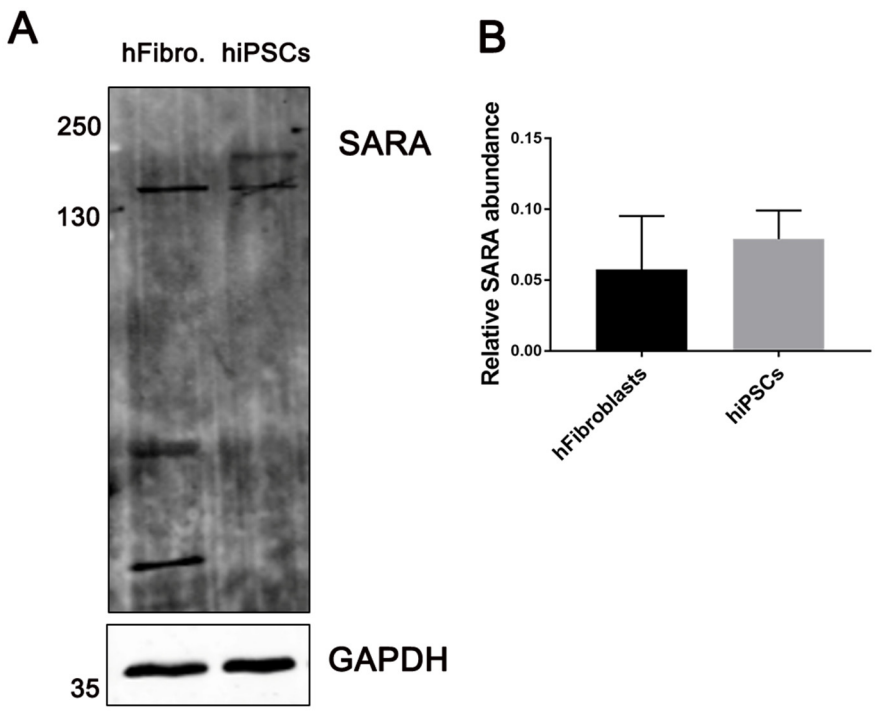


Figure 5-12 SARA levels in hFibroblasts and hiPSCs.

Human fibroblasts and hiPSCs were cultured on 6-well dishes and they were lysed in 1% SDS upon 70% confluence. Lysates were run on SDS-PAGE gel and immunoblotted with anti-SARA and anti-GAPDH antibodies. **A.** Western blot showing SARA in hfibroblasts and hiPSCs. **B.** Plot showing normalised SARA intensity in the two studied populations. Statistical significance was tested using unpaired t-test in Graph Pad (N=3, *: p-value<0.05, **:p-value<0.01, ***:p-value<0.001).

5.2.6 Generation of GFP-SARA stable cell lines in H1 cells.

Knowing that SARA is involved in asymmetric cell division prompted us to investigate whether SARA positive endosomes are unequally distributed between daughter cells in pluripotent and differentiating H1 cells. However, the low transfection efficiency combined with the long duration of differentiation protocols necessitated the generation of stable cell lines.

Accordingly, H1 cells were transfected with the GFP-SARA construct and selected with puromycin. At the end of the selection (when control cells died) 5 clones were obtained and they were further cultured for IF experiments and lysate extraction. When the clones were tested for GFP expression using a fluorescent microscope, they were found negative (data not shown). Lysates from the obtained clones as well as lysates from non-transfected (control) cells were run on SDS-PAGE gel and immunoblotted with anti-GFP antibody. Lysates form transiently transfected HEK-293A cells were used as a positive control. Additionally, the same lysates were run and immunoblotted with anti-SARA antibody. In all cases, ACTIN was used as a loading control.

According to the results, H1 clones show almost no expression of GFP-SARA compared to HEK-293A **(Fig. 5-13A)**. However, staining with an antibody recognising SARA reveals that all five clones express variable amounts of GFP-SARA while at the same time the presence of an intense smear below the GFP-SARA implies that the protein is markedly degraded **(Fig. 5-13B)**. The obtained cell lines express insufficient GFP-SARA levels and therefore they are not a useful tool for our experimental purposes.

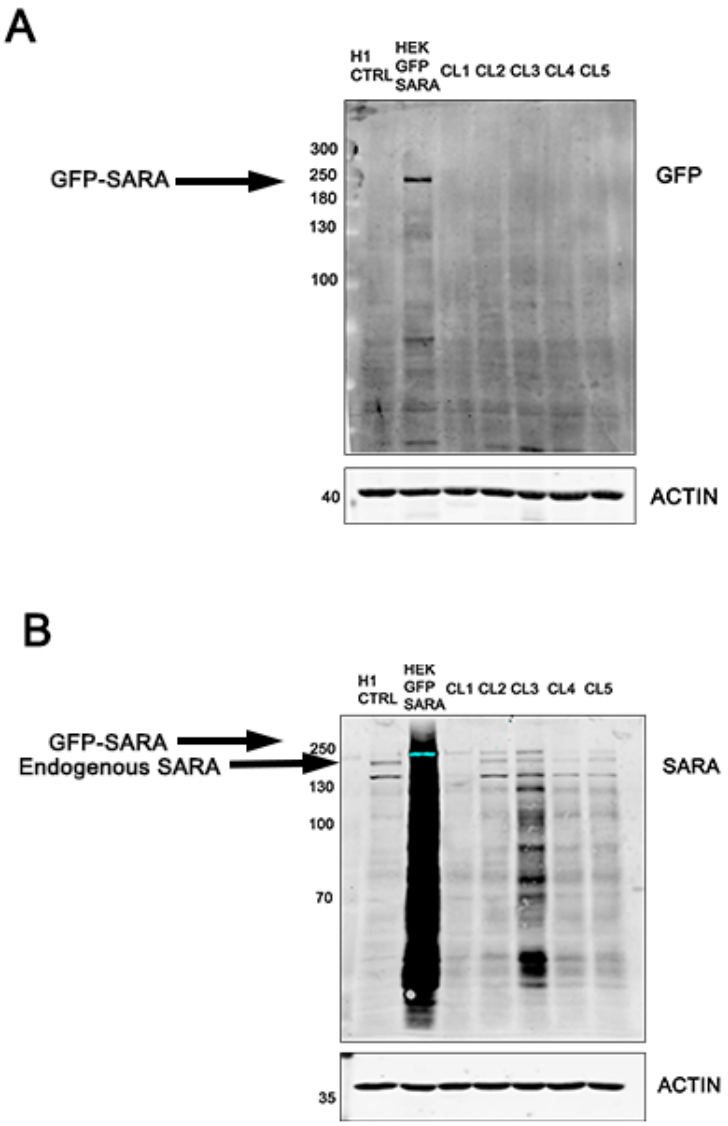


Figure 5-13 Generation of GFP-SARA stable cell lines.

H1 cells were transfected with GFP-SARA and selected in puromycin until control cells died. Five obtained clones were further cultured and cells were lysed in 1% SDS. Lysates were separately run on SDS-PAGE gels and immunoblotted with anti-GFP and anti-SARA antibodies. In both cases ACTIN was used as loading control. **A.** Western blot showing GFP expression in non-transfected H1 cells, transiently transfected HEK-293A cells and five H1 GFP-SARA clones. **B.** Western blot showing SARA expression in non-transfected H1 cells, transiently transfected HEK-293A cells and 5 H1 GFP-SARA clones.

5.2.7 CRISPR/Cas9 approach for SARA in H1 cells.

Data so far seem to indicate that SARA is regulated during differentiation. However, it is not known whether this effect is simply one of the many changes occurring during fate acquisition or if it is an essential step for the transition between the pluripotent state and the differentiated state. Loss of function experiments would be a direct way to address this possibility. Nevertheless, given the extreme difficulty to knockdown targets in H1 cells using siRNAs we decided to take advantage of the CRISPR/Cas9 technique.

For this reason, guide RNAs were designed to target regions (3rd and 4th exon) close to the start codon of the protein **(Table 5-1)**. The designing was performed according to the instructions of: SANGER CRISPR (Hodgkins et al., 2015). Guides were finally inserted in plasmids which encode for wild type Cas9 and the PAM they recognize is 5'NGG.

In all cases plasmids were transfected in HEK-293A and plasmids 2619 and 2620 **(Table 5-1)** were transfected in H1 cells. Both transfected and control cells were selected in puromycin and a few surviving clonal colonies were allowed to grow in the absence of antibiotic for one more week. Each clone was then transferred in a well of a 24 well plate with the help of cloning rings. Cells were further cultured for lysate extraction. Lysates were run on SDS-PAGE gel and immunoblotted with anti-SARA and anti-ACTIN antibodies. Although some clones showed decrease in SARA levels, none of the clones showed levels compatible with complete knockout but levels possibly compatible with monoallelic knock out **(Fig. 5-14).**

An additional approach was utilised to attempt complete SARA knock out. Two guides targeting the Transcription Start Site were designed using the MIT CRISPR

FINDER (http://crispr.mit.edu:8079/). After several cloning steps, both guides were integrated in a single vector which encoded for wild type Cas9 and recognised PAMs with 5'NGG sequence. Again both HEK-293A and H1 cells were transfected and selected and a few surviving colonies were further cultured for RNA and protein extraction. Again, no colonies with SARA complete knock out were found. Despite the unsuccessful application of CRISPR/Cas9 for SARA, it should be noted that the same experimental approach was used in parallel to knock out a different target, namely ARF6 and it was successful. This highlights that the approach works but other factors possibly related to the target of interest favour monoallelic instead of biallelic knockouts.

Table 5-1 Table depicts the guides used to target SARA, the targeted area and the plasmid number.

Primer sequence	Targeted area	Plasmid No
CACCAAGTGGTGTTTCCTCACCGA, AAACTCGGTGAGGAAACACCACTT	EXON 3	2610
AAACGACCGGAGGGATTC TTCAGT, CACCACTGAAGAATCCCTC CGGTC	EXON 4	2619
AAACTCACAATTAAAGGATGACGG, CACCCCGTCATCCTTTAAT TGTGA	EXON 4	2620
CACCTTTGGTGGTGCAAGACCCAA, AAACTTGGGTCTTGCACCACCAAA	EXON 4	2632
CACCCTCTAATGTCGATACAAATG, AAACCATTTGTATCGACATTAGAG	EXON 4	2633
CACCATCTCTAATGTCGATACAAA, AAACTTTGTATCGACATTAGAGAT	EXON 4	2634
CACCGTATGGGTACCGGATTCTCA, AAACTGAGAATCCGGTACCCATAC	EXON 4	2635
CACCCCCACGTCCGGCTACGGCGT, AAACACGCCGTAGCCGGACGTGGG	TSS	2641
CACCGCGCTCCTCCGGGATCCGCG, AAACCGCGGATCCCGAGGAGCGC	TSS	2642

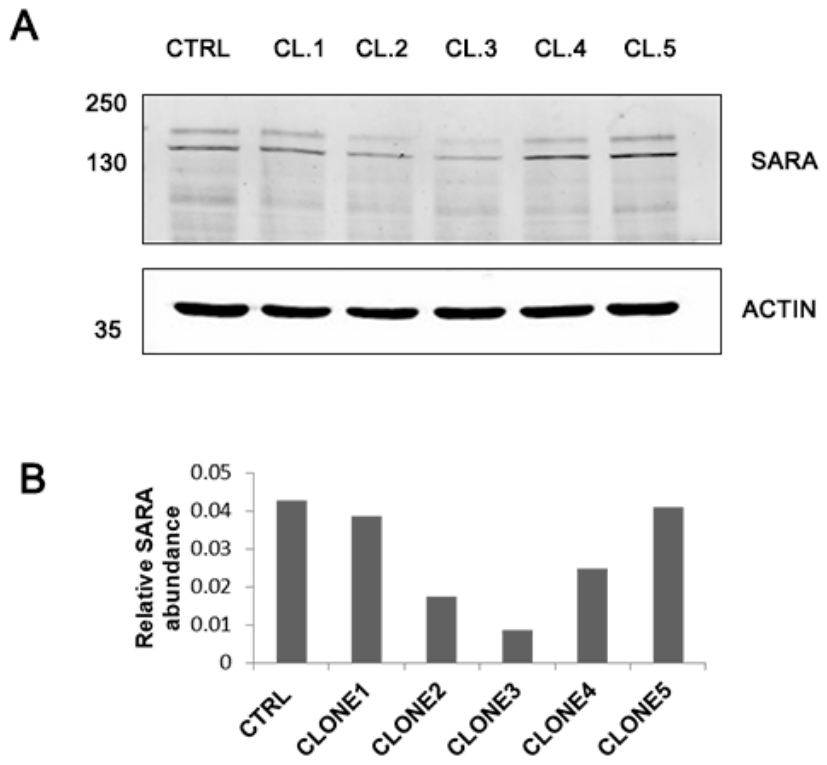


Figure 5-14 2619 plasmid in H1 cells.

H1 cells were transfected with 2619 while control cells were left untreated. 36hrs later cells were selected in puromycin (250ng/ml) until control cells died. Viable colonies were further cultured and lysates were extracted. Lysates were run on SDS-PAGE gel and immunoblotted against SARA and ACTIN. **A.** Western blot showing SARA in H1 control cells and five clones generated upon stable transfection with 2619. **B.** Plot showing normalised SARA intensity in control H1 cells and the five 2619 obtained H1 clones.

5.3 Discussion

In this chapter it is seen that HEK-293A express two SARA products, the higher being the full length and the lower one being either a novel isoform or a product of cleavage. Although, there is a difference between the predicted and the observed molecular weight, there are plausible reasons which could account for this. Full length SARA seems to migrate higher than expected and this could be due to PTMs. Indeed, SARA was proposed to be N-glycosylated when first described and this could increase the molecular weight (Meckelein et al., 1998). N- and O-glycosylation constitute the two types of glycosylation, a very common PTM which is estimated to affect half of the proteins (An *et al*, 2009). N-glycosylation in particular, occurs by the addition of an N-acetylglucosamine to an asparagine and recognizes the NXS/T motif in which X accounts for any amino acid, proline excluded (An *et al*, 2009). SARA's sequence seems to encompass 14 putative NXS/T sites; thus indicating that it could be, at least in some of them, N-glycosylated.

In addition, the presence of a band at around 100kDa, although not predicted, is in agreement with previous observations proposing that SARA is expressed in several isoforms and there is one band migrating at 97kDa (Meckelein et al., 1998). The function of this SARA product is unknown but it should be noted that according to its size it probably lacks a pig part of the C-terminal region. Mutated SARA lacking the last (852-1425) aminoacids fails to interact with the TGFβ receptor (Tsukazaki et al., 1998), therefore raising the possibility that this SARA version either is not involved in signalling or could even compete with full length SARA for binding to SMAD2. Further

experiments are needed in order to investigate its role as a potent signalling modulator.

Another, issue addressed in this chapter is whether SARA translocates to the nucleus. Despite the prediction of an NLS within the FYVE domain of SARA and the ability of two SARA mutants to accumulate in the nucleus, N- and C-terminally tagged full length SARA constructs do not display nuclear localisation. LMB treatment was utilised in order to reveal the possible shuttling of the protein between the cytoplasm and the nucleus but the results were negative. Nevertheless, it could be that SARA uses CRM1 independent machinery for nuclear export. CTNNB is a protein whose nuclear export is not impaired by LMB and is therefore CRM1 independent (Wiechens & Fagotto, 2001).

A novel finding of this chapter is the downregulation of SARA during BMP4 induced differentiation in PSCs (H1, hiPSCs). Taking into consideration the fact that mesodermal differentiation shares similarities with EMT, this finding is in agreement with previous results which show SARA decrease in the context of EMT in human kidney epithelial cells (Runyan et al., 2009). On the contrary to this publication however, our results do not seem to indicate a strong effect of SARA in the phosphorylation of SMADs and SARA is not, at least significantly, regulated at the mRNA level. The differences in the experimental context (cells, growth factors) could possibly explain the discrepancy. Besides, this is not the first time where it is noted that SMAD2 phosphorylation does not correlate with SARA levels (Bakkebø et al., 2012).

Another novel finding presented in this chapter is that SARA levels are regulated possibly by lysosomal degradation not only in the differentiation process but also in

normal pluripotent conditions. This could reflect the need for strict regulation of SARA levels in the cell, thus implying that SARA accumulation impairs some critical cellular functions. A previous study reported that SARA is degraded by the proteasome (Runyan *et al,* 2012), however, there are many reasons why this should not be considered a contradiction. First, the abovementioned paper did not check the effect of lysosomes on SARA levels; second in our case we lack data of proteasomal inhibition. Third, it has been shown that chloroquine can inhibit both lysosomal and proteasomal degradation (Myeku & Figueiredo-Pereira, 2011). Therefore, it is plausible that SARA follows both pathways for degradation and additional experiments are needed in order to dissect this and shed light in the role it plays in the function of the protein.

Last, gain and loss of function experiments would be the most straightforward approach to explore the importance of SARA in stem cells. The increased degradation of GFP-SARA in stable H1 cell lines combined with the severely low expression of GFP-SARA does not allow the efficient use of these cell lines for the study of division modes. Therefore, future work should focus in using alternative methods for the generation of the reporters.

Furthermore, the difficulty to obtain clones with complete SARA knockout directs future studies towards the use of alternative methods for genetic manipulation; such as inducible short hairpin RNA systems. An additional notion that can be proposed based on this difficulty is that SARA knockout could have adverse effects on cell proliferation or cell attachment. Such a scenario could explain the propensity for monoallelic knockouts. Despite the fact that there are no data suggestive of this effect in human cells; there is a report claiming that in Drosophila, SARA loss of

function leads to lethality since a very small percentage of mutant flies (5%) reach adulthood (Montagne & Gonzalez-Gaitan, 2014). Furthermore, in developing zebrafish, morpholino mediated zfyve9 knockdown impairs the proliferation of hepatoblasts (Liu et al., 2013).

6. GENERAL DISCUSSION AND FUTURE DIRECTIONS

The discovery of hESCs (Thomson et al., 1998) has enabled scientists to envision treating a plethora of conditions (Dupont et al., 2019). Understanding the mechanisms supporting their intrinsic properties (pluripotency and differentiation) sets the ground for improving their performance in regenerative medicine. This Thesis explores several aspects of the effect of BMP4 in hESCs from colony morphology and gene expression to phosphorylation status and effect on an endosomal protein involved in signalling.

Collectively, findings indicate that BMP4 induces a broad range of changes in H1 cells both in the short-term as well as in the long-term. In particular, the first hour of BMP4 treatment causes changes in the phosphoproteome which are compatible with preparation for migration while at the same time transcription is regulated possibly both negatively and positively. Meanwhile, 6 hrs of BMP4 exposure are sufficient to initiate mesendodermal differentiation. The above findings are suggestive of the notion that this growth factor induces irreversible key cellular responses which are responsible for fate acquisition. Indeed, 1hr of BMP4 treatment combined with removal of pluripotency sustaining factors changes almost half of the phosphoproteome in hESCs (Van Hoof et al., 2009).

Future work should try to enhance statistical significance of the obtained SILAC data by utilising new advantageous techniques such as the EasyPhos platform which offers less variation in results and downsizes the amount of cells required for phosphoproteomic analysis (Humphrey et al., 2018). Additionally, our attempts should focus on testing whether our SILAC candidates are necessary for fate determination. One possible approach would entail the generation of knockout H1

cells, for the protein of interest, which would then be compared with normal H1 cells for their responsiveness to BMP4. Migration assays could be used to decipher whether some of these proteins have a vital role in motility during differentiation (as expected). Another interesting approach would include the use of inhibitors targeting the kinases which are responsible for the observed phosphorylations. Finally, overexpression experiments would elucidate whether any of them are sufficient to cause loss of pluripotency.

In the long term, 48 hrs of exposure to BMP4 upregulates Msx2 and Hand1 while cells start expressing BRACHYURY. This indicates that the nature of these cells is mesendodermal but the presence of extraembryonic cells cannot be excluded. Moreover, 48-96hrs of BMP4 exposure leads to SARA loss possibly due to lysosomal degradation, which does not seem to impair significantly TGFβ/ACTIVIN A/NODAL signalling.

Further experiments are needed in order to elucidate the role of SARA in pluripotency and differentiation. Stable H1 cell lines with an inducible short hairpin RNA system for SARA would be useful in showing whether SARA decrease is one of several outcomes of differentiation or whether there is indeed a causative link between the two processes. Moreover, such cell lines could be tested not only in conventional cultures but also in micropatterned cultures (Warmflash et al., 2014), thus elucidating a more general role for SARA in conditions which mimic developmental events. The efficient generation of GFP-SARA stable H1 cells would enable us to reveal SARA's inheritance in dividing H1 cells under conditions which support pluripotency or cause differentiation. Live imaging experiments could be used to check for SARA

asymmetric segregation between daughter cells and whether such an asymmetry could cause asymmetry in fate acquisition.

SARA downregulation in the context of hESC differentiation towards mesendoderm seems plausible given the link between EMT and differentiation (Nakaya & Sheng, 2008). Although lysosomal function is not thoroughly investigated in differentiating hESCs, lysosomal degradation is essential for the differentiation of human keratinocytes (Monteleon et al., 2018). Moreover, in developing chick embryos autophagy seems to regulate EMT in gastrulation (W. Lu et al., 2014). Autophagy modulates the expression of molecules involved in cell adhesion and upon inhibition of autophagy the progression of the primitive streak and therefore the generation of germ layers is impaired (W. Lu et al., 2014).

E-CADHERIN downregulation is a well-known EMT feature and one of the possible mechanisms accounting for this is E-CADHERIN endocytosis and lysosomal degradation (Palacios et al., 2005). Interestingly, the activation of two RAB proteins (RAB5, RAB7) which results in endosomal enlargement, is necessary during this process thus implying that EMT regulates the endocytic network (Palacios et al., 2005). RAB5 is known to facilitate endosomal fusion, while RAB7 is involved in cargo delivery from endosomes to lysosomes (Pfeffer, 2001). It would be of interest to investigate whether mesodermal differentiation is accompanied by RAB5 activation. If that is true, then it is possible that our findings are related to a previous report demonstrating that PI3K inhibition or RAB5 activation enlarge endosomes and lead to SARA degradation via the proteasome (Runyan et al., 2012). This would also suggest that the endosomal network plays a vital role in developmental decisions.

In addition, SARA has been hypothesised to function as a RAB5 effector (Hu et al., 2002) but the feedback between SARA and RAB5 seems to be quite elaborate (Runyan et al., 2012). RAB5 activation is associated with metastasis and invasion across several cancer types (Igarashi et al., 2017), (P. Silva et al., 2016), (Mendoza et al., 2013), (Palamidessi et al., 2008). These observations combined with the protective role of SARA in sustaining an epithelial identity (Runyan et al., 2009) raises the question whether SARA could have an unknown role in cancer progression. Indeed, experiments in mice suggest that SARA acts preventively against the formation of skin cancer (Chang et al., 2014).

Another direction for future studies regards the effect of SARA on other signalling cascades. In particular, SARA is known to interact with SMAD3 (Tsukazaki et al., 1998) and could theoretically interact with CTNNB according to a yeast two-hybrid screening (Colland et al., 2004). These findings in combination with the protective role of SMAD3 against CTNNB degradation (M. Zhang et al., 2010), could indicate that SARA mediates crosstalk between WNT and TGFβ/ACTIVIN/NODAL signalling (Pálfy, Reményi, & Korcsmáros, 2012).

Recently it has been shown that SARA decrease affects L1 recycling thus impairing neuronal motility (Mestres et al., 2016). The hypothesised mechanism dictates that in the absence of SARA, L1 returns to the surface using the fast recycling pathway (Mestres et al., 2016). In vitro experiments revealed that when tight junctions are chemically disrupted, the pool of CTNNB which interacts with E-CADHERIN translocates to the ERC and can further contribute to WNT signalling (Kam & Quaranta, 2009). If similar events occur upon BMP4 induction in H1 cells, then CTNNB could follow the fast recycling pathway to the plasma membrane.

Surprisingly, in cells lacking E-CADHERIN, CTNNB is proposed to be activated at the cell membrane (Hendriksen et al., 2008). Despite the fact that this report addresses the activity of newly synthesised CTNNB, similar activation mechanism could be employed in migrating H1 cells in order to further modulate WNT signalling in differentiation.

APPENDIX A

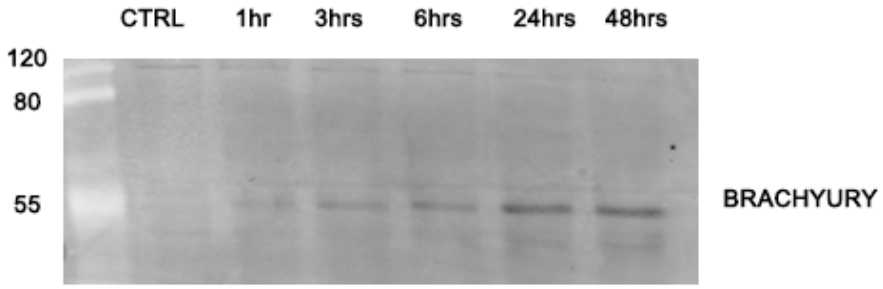


Figure A-1 Raw data from Fig. 3-3A

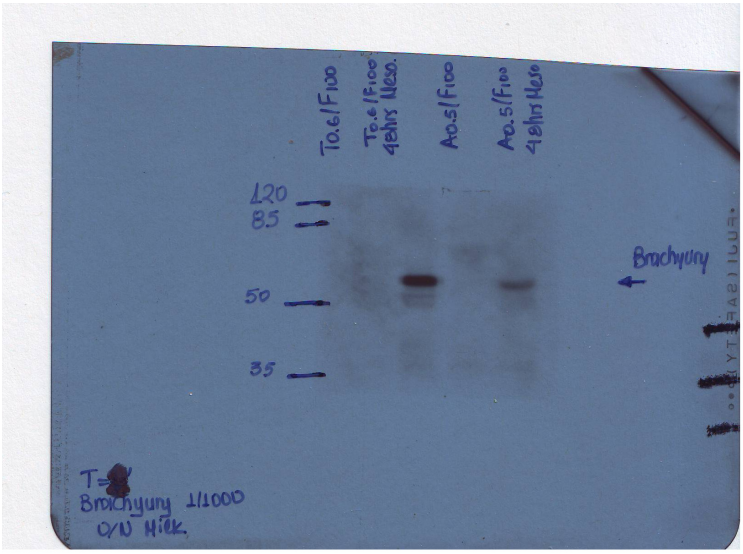


Figure A-2 Raw data from Fig. 3-7A

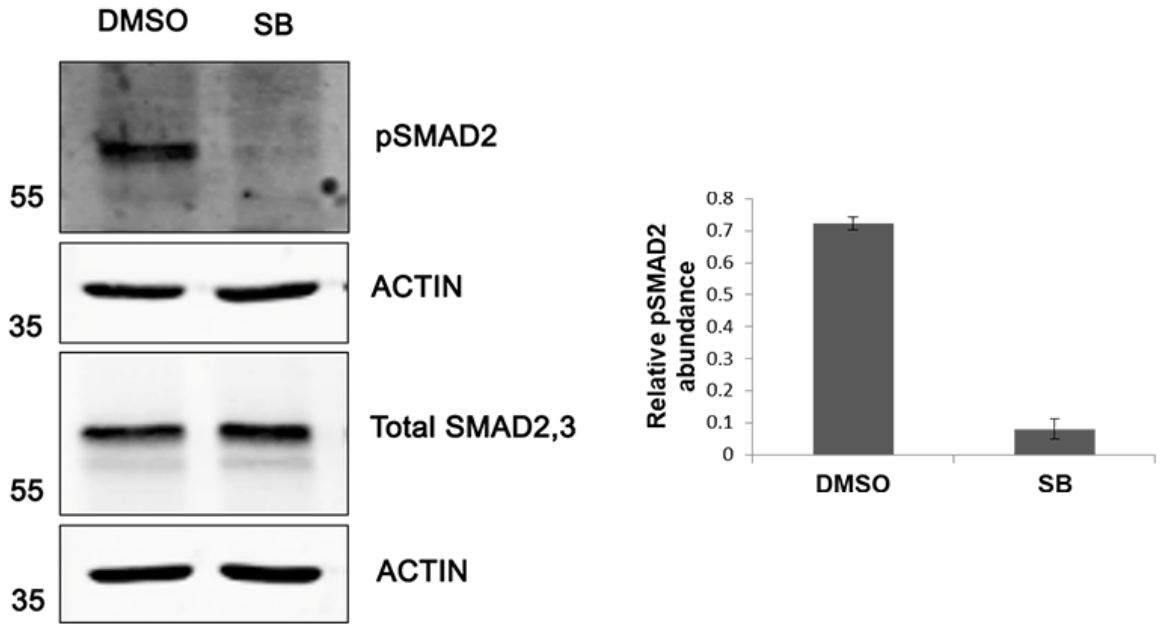


Figure A-3 SB (10μM) for 3hrs abrogates SMAD2 phosphorylation in H1 cells

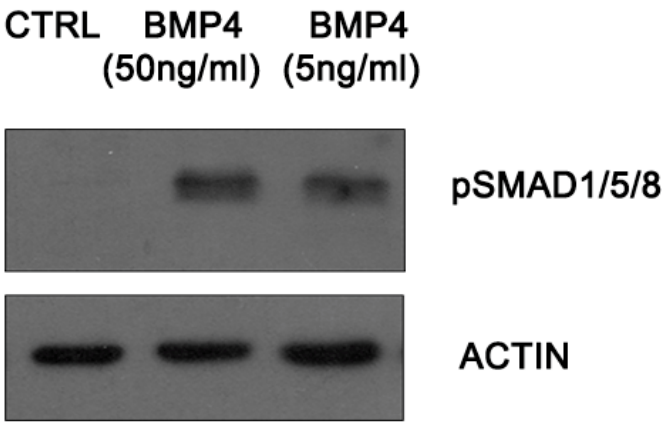


Figure A-4 BMP4 induction for 1hr increases SMAD1/5/8 phosphorylation in H1 cells

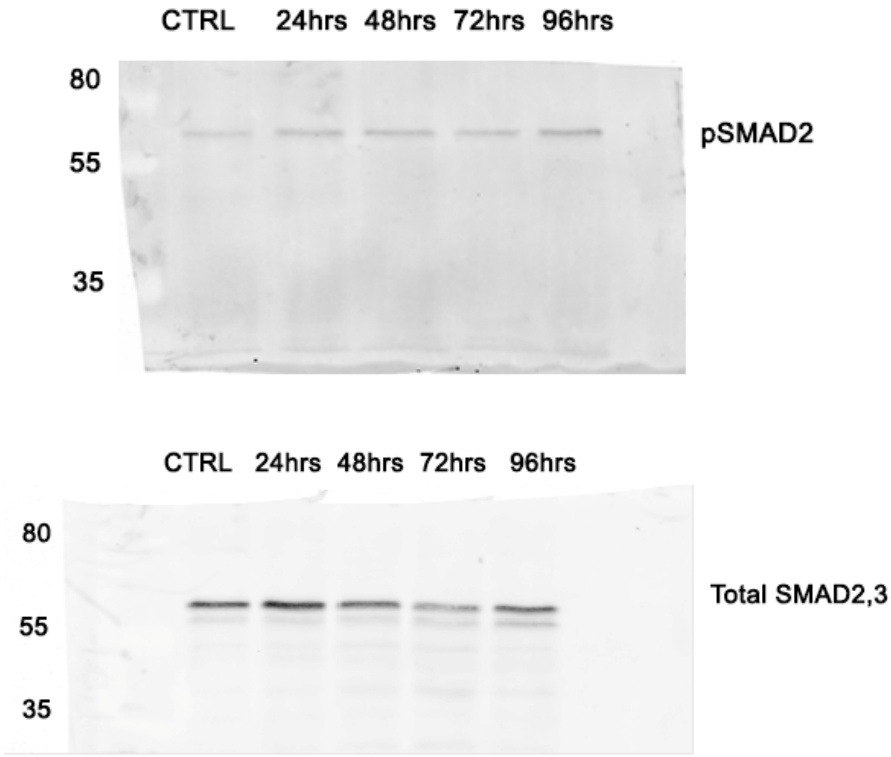


Figure A-5 Raw data from Fig. 5-10

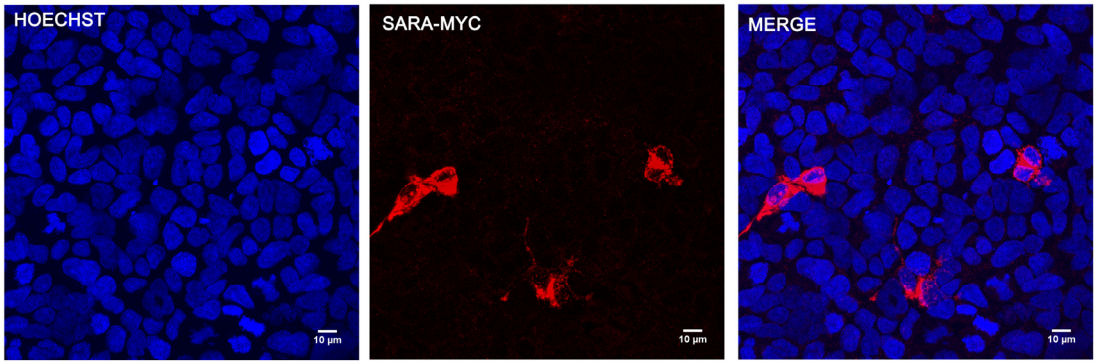


Figure A-6 SARA-MYC in untreated HEK-293A cells

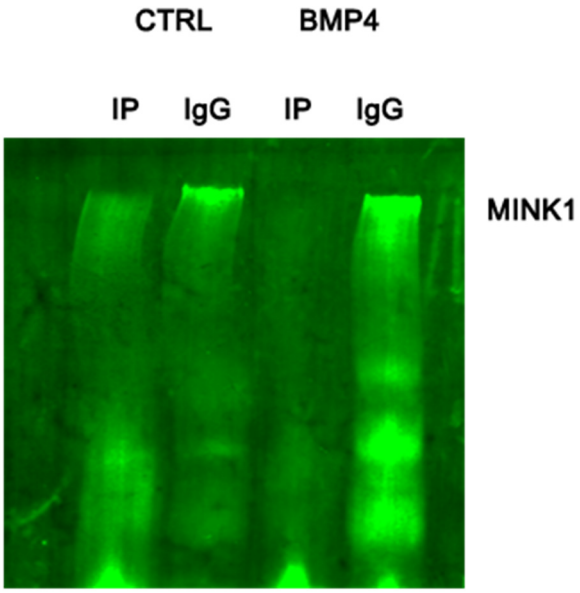


Figure A-7 MINK1 IP in H1 cells using phos-tag gels

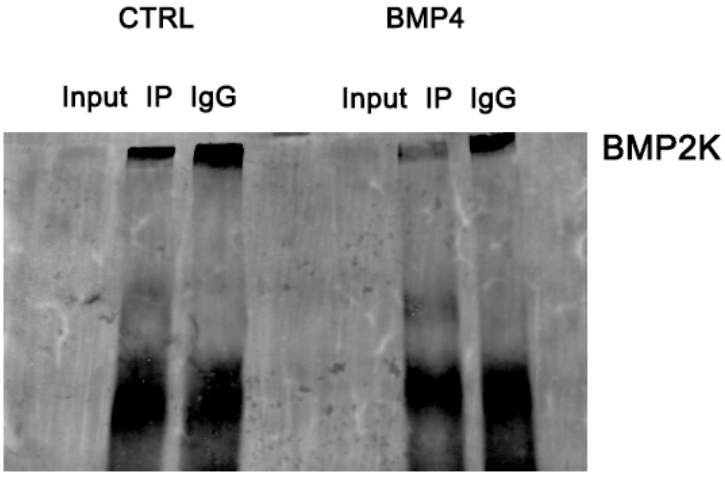


Figure A-8 BMP2K IP in H1 cells using phos-tag gels

REFERENCES

Abdollah, S., Macías-Silva, M., Tsukazaki, T., Hayashi, H., Attisano, L., & Wrana, J. L. (1997). TbetaRI phosphorylation of Smad2 on Ser465 and Ser467 is required for Smad2-Smad4 complex formation and signaling. *The Journal of Biological Chemistry*, *272*(44), 27678–27685. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9346908

Agarwal, S., Holton, K. L., & Lanza, R. (2008). Efficient differentiation of functional hepatocytes from human embryonic stem cells. *Stem Cells (Dayton, Ohio)*, *26*(5), 1117–1127. https://doi.org/10.1634/stemcells.2007-1102

Albano, R. M., Groome, N., & Smith, J. C. (1993). Activins are expressed in preimplantation mouse embryos and in ES and EC cells and are regulated on their differentiation. *Development (Cambridge, England)*, *117*(2), 711–723. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8330535

An, H. J., Froehlich, J. W., & Lebrilla, C. B. (2009). Determination of glycosylation sites and site-specific heterogeneity in glycoproteins. *Current Opinion in Chemical Biology*, *13*(4), 421–426. https://doi.org/10.1016/j.cbpa.2009.07.022

Andrews, P. W., Ben-David, U., Benvenisty, N., Coffey, P., Eggan, K., Knowles, B. B., … Stacey, G. N. (2017). Assessing the Safety of Human Pluripotent Stem Cells and Their Derivatives for Clinical Applications. *Stem Cell Reports*, *9*(1), 1–4. https://doi.org/10.1016/j.stemcr.2017.05.029

Arias, C. I., Siri, S. O., & Conde, C. (2015). Involvement of SARA in Axon and Dendrite Growth. *PloS One*, *10*(9), e0138792. https://doi.org/10.1371/journal.pone.0138792

Armstrong, L., Hughes, O., Yung, S., Hyslop, L., Stewart, R., Wappler, I., … 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. *Human Molecular Genetics*, *15*(11), 1894–1913. https://doi.org/10.1093/hmg/ddl112

Artus, J., Piliszek, A., & Hadjantonakis, A.-K. (2011). The primitive endoderm lineage of the mouse blastocyst: sequential transcription factor activation and regulation of differentiation by Sox17. *Developmental Biology*, *350*(2), 393–404. https://doi.org/10.1016/j.ydbio.2010.12.007

Atfi, A., Dumont, E., Colland, F., Bonnier, D., L'helgoualc'h, A., Prunier, C., … Théret, N. (2007). The disintegrin and metalloproteinase ADAM12 contributes to TGF-beta signaling through interaction with the type II receptor. *The Journal of Cell Biology*, *178*(2), 201–208. https://doi.org/10.1083/jcb.200612046

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 & Development*, *17*(1), 126–140. https://doi.org/10.1101/gad.224503

Baker, D. E. C., Harrison, N. J., Maltby, E., Smith, K., Moore, H. D., Shaw, P. J., … Andrews, P. W. (2007). Adaptation to culture of human embryonic stem cells and oncogenesis in vivo. *Nature Biotechnology*, *25*(2), 207–215. https://doi.org/10.1038/nbt1285

Bakkebø, M., Huse, K., Hilden, V. I., Forfang, L., Myklebust, J. H., Smeland, E. B., & Oksvold, M. P. (2012). SARA is dispensable for functional TGF-β signaling. *FEBS Letters*, *586*(19), 3367–3372. https://doi.org/10.1016/j.febslet.2012.07.027

Barford, D., Das, A. K., & Egloff, M. P. (1998). The structure and mechanism of protein phosphatases: insights into catalysis and regulation. *Annual Review of Biophysics and Biomolecular Structure*, *27*, 133–164.

https://doi.org/10.1146/annurev.biophys.27.1.133

Basta, J., & Rauchman, M. (2015). The nucleosome remodeling and deacetylase complex in development and disease. *Translational Research : The Journal of Laboratory and Clinical Medicine*, *165*(1), 36–47. https://doi.org/10.1016/j.trsl.2014.05.003

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*(4), 489–495. https://doi.org/10.1634/stemcells.2004-0279

Ben-Haim, N., Lu, C., Guzman-Ayala, M., Pescatore, L., Mesnard, D., Bischofberger, M., … Constam, D. B. (2006). The nodal precursor acting via activin receptors induces mesoderm by maintaining a source of its convertases and BMP4. *Developmental Cell*, *11*(3), 313–323. https://doi.org/10.1016/j.devcel.2006.07.005

Bennett, D., & Alphey, L. (2002). PP1 binds Sara and negatively regulates Dpp signaling in Drosophila melanogaster. *Nature Genetics*, *31*(4), 419–423. https://doi.org/10.1038/ng938

Bernardo, A. S., Faial, T., Gardner, L., Niakan, K. K., Ortmann, D., Senner, C. E., … Pedersen, R. A. (2011). BRACHYURY and CDX2 mediate BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and extraembryonic lineages. *Cell Stem Cell*, *9*(2), 144–155. https://doi.org/10.1016/j.stem.2011.06.015

Bertero, A., Brown, S., Madrigal, P., Osnato, A., Ortmann, D., Yiangou, L., … Vallier, L. (2018). The SMAD2/3 interactome reveals that TGFβ controls m6A mRNA methylation in pluripotency. *Nature*, *555*(7695), 256–259. https://doi.org/10.1038/nature25784

Bertero, A., Madrigal, P., Galli, A., Hubner, N. C., Moreno, I., Burks, D., … Vallier, L. (2015). Activin/nodal signaling and NANOG orchestrate human embryonic stem cell fate decisions by controlling the H3K4me3 chromatin mark. *Genes & Development*, *29*(7), 702–717. https://doi.org/10.1101/gad.255984.114

Beyer, T. A., Narimatsu, M., Weiss, A., David, L., & Wrana, J. L. (2013). The TGFβ superfamily in stem cell biology and early mammalian embryonic development. *Biochimica et Biophysica Acta*, *1830*(2), 2268–2279. https://doi.org/10.1016/j.bbagen.2012.08.025

Beyer, T. A., Weiss, A., Khomchuk, Y., Huang, K., Ogunjimi, A. A., Varelas, X., & Wrana, J. L. (2013). Switch enhancers interpret TGF-β and Hippo signaling to control cell fate in human embryonic stem cells. *Cell Reports*, *5*(6), 1611–1624. https://doi.org/10.1016/j.celrep.2013.11.021

Blauwkamp, T. A., Nigam, S., Ardehali, R., Weissman, I. L., & Nusse, R. (2012). Endogenous Wnt signalling in human embryonic stem cells generates an equilibrium of distinct lineage-specified progenitors. *Nature Communications*, *3*, 1070. https://doi.org/10.1038/ncomms2064

Bökel, C., Schwabedissen, A., Entchev, E., Renaud, O., & González-Gaitán, M. (2006). Sara endosomes and the maintenance of Dpp signaling levels across mitosis. *Science (New York, N.Y.)*, *314*(5802), 1135–1139. https://doi.org/10.1126/science.1132524

Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., … Young, R. A. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell*, *122*(6), 947–956. https://doi.org/10.1016/j.cell.2005.08.020

Braconi Quintaje, S., & Orchard, S. (2008). The annotation of both human and mouse kinomes in UniProtKB/Swiss-Prot: one small step in manual annotation, one giant

leap for full comprehension of genomes. *Molecular & Cellular Proteomics : MCP*, *7*(8), 1409–1419. https://doi.org/10.1074/mcp.R700001-MCP200

Brill, L. M., Xiong, W., Lee, K., Ficarro, S. B., Crain, A., Xu, Y., … Ding, S. (2009). Phosphoproteomic analysis of human embryonic stem cells. *Cell Stem Cell*, *5*(2), 204–213. https://doi.org/10.1016/j.stem.2009.06.002

Brons, I. G. M., Smithers, L. E., Trotter, M. W. B., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S. M., … Vallier, L. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature*, *448*(7150), 191–195. https://doi.org/10.1038/nature05950

Brown, S., Teo, A., Pauklin, S., Hannan, N., Cho, C. H.-H., Lim, B., … Vallier, L. (2011). Activin/Nodal signaling controls divergent transcriptional networks in human embryonic stem cells and in endoderm progenitors. *Stem Cells (Dayton, Ohio)*, *29*(8), 1176–1185. https://doi.org/10.1002/stem.666

Burrows, C., Abd Latip, N., Lam, S., Carpenter, L., Sawicka, K., Tzolovsky, G., … Blagden, S. P. (2010). The RNA binding protein Larp1 regulates cell division, apoptosis and cell migration. *Nucleic Acids Research*, *38*(16), 5542–5553. https://doi.org/10.1093/nar/gkq294

Campbell, K. H. S., McWhir, J., Ritchie, W. A., & Wilmut, I. (1996). Sheep cloned by nuclear transfer from a cultured cell line. *Nature*, *380*(6569), 64–66. https://doi.org/10.1038/380064a0

Carlson Scholz, J. A., Garg, R., Compton, S. R., Allore, H. G., Zeiss, C. J., & Uchio, E. M. (2011). Poliomyelitis in MuLV-infected ICR-SCID mice after injection of basement membrane matrix contaminated with lactate dehydrogenase-elevating virus. *Comparative Medicine*, *61*(5), 404–411. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22330347

Cha, Y., Kim, D.-K., Hyun, J., Kim, S., & Park, K. (2013). TCEA3 binds to TGF-beta receptor I and induces Smad-independent, JNK-dependent apoptosis in ovarian cancer cells. *Cellular Signalling*, *25*(5), 1245–1251. https://doi.org/10.1016/j.cellsig.2013.01.016

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*(5), 643–655. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12787505

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., … Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. *Nature*, *450*(7173), 1230–1234. https://doi.org/10.1038/nature06403

Chambers, I., & Smith, A. (2004). Self-renewal of teratocarcinoma and embryonic stem cells. *Oncogene*, *23*(43), 7150–7160. https://doi.org/10.1038/sj.onc.1207930

Chang, H., Lin, Y., Tsai, P., Liang, C., & Yan, Y. (2014). The FYVE domain of Smad Anchor for Receptor Activation (SARA) is required to prevent skin carcinogenesis, but not in mouse development. *PloS One*, *9*(8), e105299. https://doi.org/10.1371/journal.pone.0105299

Chazaud, C., Yamanaka, Y., Pawson, T., & Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. *Developmental Cell*, *10*(5), 615–624. https://doi.org/10.1016/j.devcel.2006.02.020

Chen, Y. (2009). Endocytic regulation of TGF-beta signaling. *Cell Research*, *19*(1), 58–70. https://doi.org/10.1038/cr.2008.315

Chen, Y. G., Hata, A., Lo, R. S., Wotton, D., Shi, Y., Pavletich, N., & Massagué, J.

(1998). Determinants of specificity in TGF-beta signal transduction. *Genes & Development*, *12*(14), 2144–2152. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9679059

Chen, Y., Wang, Z., Ma, J., Zhang, L., & Lu, Z. (2007). Endofin, a FYVE domain protein, interacts with Smad4 and facilitates transforming growth factor-beta signaling. *The Journal of Biological Chemistry*, *282*(13), 9688–9695. https://doi.org/10.1074/jbc.M611704200

Cheng, L., Hammond, H., Ye, Z., Zhan, X., & Dravid, G. (2003). Human adult marrow cells support prolonged expansion of human embryonic stem cells in culture. *Stem Cells (Dayton, Ohio)*, *21*(2), 131–142. https://doi.org/10.1634/stemcells.21-2-131

Clay, M. R., & Halloran, M. C. (2013). Rho activation is apically restricted by Arhgap1 in neural crest cells and drives epithelial-to-mesenchymal transition. *Development (Cambridge, England)*, *140*(15), 3198–3209. https://doi.org/10.1242/dev.095448

Cobo, F., Navarro, J. M., Herrera, M. I., Vivo, A., Porcel, D., Hernández, C., … Menendez, P. (2008). Electron microscopy reveals the presence of viruses in mouse embryonic fibroblasts but neither in human embryonic fibroblasts nor in human mesenchymal cells used for hESC maintenance: toward an implementation of microbiological quality assurance program i. *Cloning and Stem Cells*, *10*(1), 65–74. https://doi.org/10.1089/clo.2007.0020

Cole, A., Frame, S., & Cohen, P. (2004). Further evidence that the tyrosine phosphorylation of glycogen synthase kinase-3 (GSK3) in mammalian cells is an autophosphorylation event. *The Biochemical Journal*, *377*(Pt 1), 249–255. https://doi.org/10.1042/BJ20031259

Colland, F., Jacq, X., Trouplin, V., Mougin, C., Groizeleau, C., Hamburger, A., … Gauthier, J. (2004). Functional proteomics mapping of a human signaling pathway.

Genome Research, *14*(7), 1324–1332. https://doi.org/10.1101/gr.2334104

Coumailleau, F., Fürthauer, M., Knoblich, J. A., & González-Gaitán, M. (2009). Directional Delta and Notch trafficking in Sara endosomes during asymmetric cell division. *Nature*, *458*(7241), 1051–1055. https://doi.org/10.1038/nature07854

Cowan, C. A., Klimanskaya, I., McMahon, J., Atienza, J., Witmyer, J., Zucker, J. P., … Melton, D. A. (2004). Derivation of embryonic stem-cell lines from human blastocysts. *The New England Journal of Medicine*, *350*(13), 1353–1356. https://doi.org/10.1056/NEJMsr040330

Crease, D. J., Dyson, S., & Gurdon, J. B. (1998). Cooperation between the activin and Wnt pathways in the spatial control of organizer gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(8), 4398–4403. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9539748

Dailey, L., Ambrosetti, D., Mansukhani, A., & Basilico, C. (2005). Mechanisms underlying differential responses to FGF signaling. *Cytokine & Growth Factor Reviews*, *16*(2), 233–247. https://doi.org/10.1016/j.cytogfr.2005.01.007

Daley, G. Q. (2015). Stem cells and the evolving notion of cellular identity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1680), 20140376. https://doi.org/10.1098/rstb.2014.0376

Daugherty, R. L., & Gottardi, C. J. (2007). Phospho-regulation of Beta-catenin adhesion and signaling functions. *Physiology (Bethesda, Md.)*, *22*(14), 303–309. https://doi.org/10.1152/physiol.00020.2007

Daulat, A. M., Luu, O., Sing, A., Zhang, L., Wrana, J. L., McNeill, H., … Angers, S. (2012). Mink1 regulates β-catenin-independent Wnt signaling via Prickle phosphorylation. *Molecular and Cellular Biology*, *32*(1), 173–185.

https://doi.org/10.1128/MCB.06320-11

de Duve, C., de Barsy, T., Poole, B., Trouet, A., Tulkens, P., & Van Hoof, F. (1974). Commentary. Lysosomotropic agents. *Biochemical Pharmacology*, *23*(18), 2495–2531. https://doi.org/10.1016/0006-2952(74)90174-9

De Los Angeles, A., Loh, Y.-H., Tesar, P. J., & Daley, G. Q. (2012). Accessing naïve human pluripotency. *Current Opinion in Genetics & Development*, *22*(3), 272–282. https://doi.org/10.1016/j.gde.2012.03.001

De Paepe, C., Krivega, M., Cauffman, G., Geens, M., & Van de Velde, H. (2014). Totipotency and lineage segregation in the human embryo. *Molecular Human Reproduction*, *20*(7), 599–618. https://doi.org/10.1093/molehr/gau027

Delcuve, G. P., Khan, D. H., & Davie, J. R. (2012). Roles of histone deacetylases in epigenetic regulation: emerging paradigms from studies with inhibitors. *Clinical Epigenetics*, *4*(1), 5. https://doi.org/10.1186/1868-7083-4-5

Delevoye, C., & Goud, B. (2015). Rab GTPases and kinesin motors in endosomal trafficking. *Methods in Cell Biology*, *130*, 235–246. https://doi.org/10.1016/bs.mcb.2015.05.004

Derivery, E., Seum, C., Daeden, A., Loubéry, S., Holtzer, L., Jülicher, F., & Gonzalez-Gaitan, M. (2015). Polarized endosome dynamics by spindle asymmetry during asymmetric cell division. *Nature*, *528*(7581), 280–285. https://doi.org/10.1038/nature16443

Derynck, R., Gelbart, W. M., Harland, R. M., Heldin, C. H., Kern, S. E., Massagué, J., … Wang, X. F. (1996). Nomenclature: vertebrate mediators of TGFbeta family signals. *Cell*, *87*(2), 173. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8861901

Desai, N., Rambhia, P., & Gishto, A. (2015). Human embryonic stem cell cultivation:

historical perspective and evolution of xeno-free culture systems. *Reproductive Biology and Endocrinology : RB&E*, *13*(1), 9. https://doi.org/10.1186/s12958-015-0005-4

Di Guglielmo, G. M., Le Roy, C., Goodfellow, A. F., & Wrana, J. L. (2003). Distinct endocytic pathways regulate TGF-beta receptor signalling and turnover. *Nature Cell Biology*, *5*(5), 410–421. https://doi.org/10.1038/ncb975

Diaz, R., Mayorga, L., & Stahl, P. (1988). In vitro fusion of endosomes following receptor-mediated endocytosis. *The Journal of Biological Chemistry*, *263*(13), 6093–6100. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3360775

Dobson, A. T., Raja, R., Abeyta, M. J., Taylor, T., Shen, S., Haqq, C., & Pera, R. A. R. (2004). The unique transcriptome through day 3 of human preimplantation development. *Human Molecular Genetics*, *13*(14), 1461–1470. https://doi.org/10.1093/hmg/ddh157

Doherty, G. J., & McMahon, H. T. (2009). Mechanisms of endocytosis. *Annual Review of Biochemistry*, *78*, 857–902. https://doi.org/10.1146/annurev.biochem.78.081307.110540

Dowler, S., Currie, R. A., Campbell, D. G., Deak, M., Kular, G., Downes, C. P., & Alessi, D. R. (2000). Identification of pleckstrin-homology-domain-containing proteins with novel phosphoinositide-binding specificities. *The Biochemical Journal*, *351*(Pt 1), 19–31. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11001876

Dravid, G., Ye, Z., Hammond, H., Chen, G., Pyle, A., Donovan, P., … Cheng, L. (2005). Defining the role of Wnt/beta-catenin signaling in the survival, proliferation, and self-renewal of human embryonic stem cells. *Stem Cells (Dayton, Ohio)*, *23*(10), 1489–1501. https://doi.org/10.1634/stemcells.2005-0034

Dulak, J., Szade, K., Szade, A., Nowak, W., & Józkowicz, A. (2015). Adult stem cells: hopes and hypes of regenerative medicine. *Acta Biochimica Polonica*, *62*(3), 329–337. https://doi.org/10.18388/abp.2015_1023

Dupont, G., Yilmaz, E., Loukas, M., Macchi, V., De Caro, R., & Tubbs, R. S. (2019). Human embryonic stem cells: Distinct molecular personalities and applications in regenerative medicine. *Clinical Anatomy (New York, N.Y.)*, *32*(3), 354–360. https://doi.org/10.1002/ca.23318

Eastham, A. M., Spencer, H., Soncin, F., Ritson, S., Merry, C. L. R., Stern, P. L., & Ward, C. M. (2007). Epithelial-mesenchymal transition events during human embryonic stem cell differentiation. *Cancer Research*, *67*(23), 11254–11262. https://doi.org/10.1158/0008-5472.CAN-07-2253

Ehrlich, M. (2016). Endocytosis and trafficking of BMP receptors: Regulatory mechanisms for fine-tuning the signaling response in different cellular contexts. *Cytokine & Growth Factor Reviews*, *27*, 35–42. https://doi.org/10.1016/j.cytogfr.2015.12.008

Eisele, S. C., Gill, C. M., Shankar, G. M., & Brastianos, P. K. (2015). PLEKHA5: A Key to Unlock the Blood-Brain Barrier? *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, *21*(9), 1978–1980. https://doi.org/10.1158/1078-0432.CCR-14-2604

Eiselleova, L., Matulka, K., Kriz, V., Kunova, M., Schmidtova, Z., Neradil, J., … Dvorak, P. (2009). A complex role for FGF-2 in self-renewal, survival, and adhesion of human embryonic stem cells. *Stem Cells (Dayton, Ohio)*, *27*(8), 1847–1857. https://doi.org/10.1002/stem.128

Evseenko, D., Zhu, Y., Schenke-Layland, K., Kuo, J., Latour, B., Ge, S., … Crooks, G. M. (2010). Mapping the first stages of mesoderm commitment during differentiation

of human embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(31), 13742–13747. https://doi.org/10.1073/pnas.1002077107

Fang, D., Hawke, D., Zheng, Y., Xia, Y., Meisenhelder, J., Nika, H., … Lu, Z. (2007). Phosphorylation of beta-catenin by AKT promotes beta-catenin transcriptional activity. *The Journal of Biological Chemistry*, *282*(15), 11221–11229. https://doi.org/10.1074/jbc.M611871200

Gafni, O., Weinberger, L., Mansour, A. A., Manor, Y. S., Chomsky, E., Ben-Yosef, D., … Hanna, J. H. (2013). Derivation of novel human ground state naive pluripotent stem cells. *Nature*, *504*(7479), 282–286. https://doi.org/10.1038/nature12745

Gaj, T., Gersbach, C. A., & Barbas, C. F. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology*, *31*(7), 397–405. https://doi.org/10.1016/j.tibtech.2013.04.004

Geer, L. Y., Marchler-Bauer, A., Geer, R. C., Han, L., He, J., He, S., … Bryant, S. H. (2010). The NCBI BioSystems database. *Nucleic Acids Research*, *38*(Database issue), D492-6. https://doi.org/10.1093/nar/gkp858

Gnanapragasam, M. N., Scarsdale, J. N., Amaya, M. L., Webb, H. D., Desai, M. A., Walavalkar, N. M., … Williams, D. C. (2011). p66Alpha-MBD2 coiled-coil interaction and recruitment of Mi-2 are critical for globin gene silencing by the MBD2-NuRD complex. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(18), 7487–7492. https://doi.org/10.1073/pnas.1015341108

Goy, I. (2018). Was Aristotle the ‘father’ of the epigenesis doctrine? *History and Philosophy of the Life Sciences*, *40*(2), 28. https://doi.org/10.1007/s40656-018-0193-2

Greber, B., Coulon, P., Zhang, M., Moritz, S., Frank, S., Müller-Molina, A. J., … Schöler, H. R. (2011). FGF signalling inhibits neural induction in human embryonic stem cells. *The EMBO Journal*, *30*(24), 4874–4884. https://doi.org/10.1038/emboj.2011.407

Greber, B., Lehrach, H., & Adjaye, J. (2007). Fibroblast growth factor 2 modulates transforming growth factor beta signaling in mouse embryonic fibroblasts and human ESCs (hESCs) to support hESC self-renewal. *Stem Cells (Dayton, Ohio)*, *25*(2), 455–464. https://doi.org/10.1634/stemcells.2006-0476

Guo, D., Huang, J., & Gong, J. (2012). Bone morphogenetic protein 4 (BMP4) is required for migration and invasion of breast cancer. *Molecular and Cellular Biochemistry*, *363*(1–2), 179–190. https://doi.org/10.1007/s11010-011-1170-1

GURDON, J. B. (1962). The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. *Journal of Embryology and Experimental Morphology*, *10*(December), 622–640. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13951335

Hamada, S., Satoh, K., Hirota, M., Kimura, K., Kanno, A., Masamune, A., & Shimosegawa, T. (2007). Bone morphogenetic protein 4 induces epithelial-mesenchymal transition through MSX2 induction on pancreatic cancer cell line. *Journal of Cellular Physiology*, *213*(3), 768–774. https://doi.org/10.1002/jcp.21148

Hartung, A., Bitton-Worms, K., Rechtman, M. M., Wenzel, V., Boergermann, J. H., Hassel, S., … Knaus, P. (2006). Different Routes of Bone Morphogenic Protein (BMP) Receptor Endocytosis Influence BMP Signaling. *Molecular and Cellular Biology*, *26*(20), 7791–7805. https://doi.org/10.1128/mcb.00022-06

Heard, E. (2004). Recent advances in X-chromosome inactivation. *Current Opinion in Cell Biology*, *16*(3), 247–255. https://doi.org/10.1016/j.ceb.2004.03.005

HEARD, E., CHAUMEIL, J., MASUI, O., & OKAMOTO, I. (2004). Mammalian X-Chromosome Inactivation: An Epigenetics Paradigm. *Cold Spring Harbor Symposia on Quantitative Biology*, *69*, 89–102. https://doi.org/10.1101/sqb.2004.69.89

Heldin, C., Vanlandewijck, M., & Moustakas, A. (2012). Regulation of EMT by TGFβ in cancer. *FEBS Letters*, *586*(14), 1959–1970. https://doi.org/10.1016/j.febslet.2012.02.037

Hendriksen, J., Jansen, M., Brown, C. M., van der Velde, H., van Ham, M., Galjart, N., … Fornerod, M. (2008). Plasma membrane recruitment of dephosphorylated beta-catenin upon activation of the Wnt pathway. *Journal of Cell Science*, *121*(11), 1793–1802. https://doi.org/10.1242/jcs.025536

Hill, C. S. (2009). Nucleocytoplasmic shuttling of Smad proteins. *Cell Research*, *19*(1), 36–46. https://doi.org/10.1038/cr.2008.325

Hochedlinger, K., & Jaenisch, R. (2006). Nuclear reprogramming and pluripotency. *Nature*, *441*(7097), 1061–1067. https://doi.org/10.1038/nature04955

Hodgkins, A., Farne, A., Perera, S., Grego, T., Parry-Smith, D. J., Skarnes, W. C., & Iyer, V. (2015). WGE: a CRISPR database for genome engineering. *Bioinformatics (Oxford, England)*, *31*(18), 3078–3080. https://doi.org/10.1093/bioinformatics/btv308

Horn, H., Schoof, E. M., Kim, J., Robin, X., Miller, M. L., Diella, F., … Linding, R. (2014). KinomeXplorer: an integrated platform for kinome biology studies. *Nature Methods*, *11*(6), 603–604. https://doi.org/10.1038/nmeth.2968

Hornbeck, P. V, Zhang, B., Murray, B., Kornhauser, J. M., Latham, V., & Skrzypek, E. (2015). PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Research*, *43*(Database issue), D512-20. https://doi.org/10.1093/nar/gku1267

Hu, Y., Chuang, J.-Z., Xu, K., McGraw, T. G., & Sung, C.-H. (2002). SARA, a FYVE domain protein, affects Rab5-mediated endocytosis. *Journal of Cell Science*, *115*(Pt 24), 4755–4763. https://doi.org/10.1021/nl061786n.Core-Shell

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, *37*(1), 1–13. https://doi.org/10.1093/nar/gkn923

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4*(1), 44–57. https://doi.org/10.1038/nprot.2008.211

Huang, F., & Chen, Y. (2012). Regulation of TGF-β receptor activity. *Cell & Bioscience*, *2*(1), 9. https://doi.org/10.1186/2045-3701-2-9

Huang, T.-S., Li, L., Moalim-Nour, L., Jia, D., Bai, J., Yao, Z., … Wang, L. (2015). A Regulatory Network Involving β-Catenin, e-Cadherin, PI3k/Akt, and Slug Balances Self-Renewal and Differentiation of Human Pluripotent Stem Cells In Response to Wnt Signaling. *Stem Cells (Dayton, Ohio)*, *33*(5), 1419–1433. https://doi.org/10.1002/stem.1944

Hughes, K., Nikolakaki, E., Plyte, S. E., Totty, N. F., & Woodgett, J. R. (1993). Modulation of the glycogen synthase kinase-3 family by tyrosine phosphorylation. *The EMBO Journal*, *12*(2), 803–808. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8382613

Humphrey, S. J., Karayel, O., James, D. E., & Mann, M. (2018). High-throughput and high-sensitivity phosphoproteomics with the EasyPhos platform. *Nature Protocols*, *13*(9), 1897–1916. https://doi.org/10.1038/s41596-018-0014-9

Igarashi, T., Araki, K., Yokobori, T., Altan, B., Yamanaka, T., Ishii, N., … Kuwano, H.

(2017). Association of RAB5 overexpression in pancreatic cancer with cancer progression and poor prognosis via E-cadherin suppression. *Oncotarget*, *8*(7), 12290–12300. https://doi.org/10.18632/oncotarget.14703

Ilic, D., Devito, L., Miere, C., & Codognotto, S. (2015). Human embryonic and induced pluripotent stem cells in clinical trials. *British Medical Bulletin*, *116*(July 2018), 19–27. https://doi.org/10.1093/bmb/ldv045

Ilic, D., & Ogilvie, C. (2017). Concise Review: Human Embryonic Stem Cells-What Have We Done? What Are We Doing? Where Are We Going? *Stem Cells (Dayton, Ohio)*, *35*(1), 17–25. https://doi.org/10.1002/stem.2450

Inman, G. J., Nicolás, F. J., Callahan, J. F., Harling, J. D., Gaster, L. M., Reith, A. D., … Hill, C. S. (2002). SB-431542 is a potent and specific inhibitor of transforming growth factor-beta superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7. *Molecular Pharmacology*, *62*(1), 65–74. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12065756

International Stem Cell Initiative, Adewumi, O., Aflatoonian, B., Ahrlund-Richter, L., Amit, M., Andrews, P. W., … Zhang, W. (2007). Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. *Nature Biotechnology*, *25*(7), 803–816. https://doi.org/10.1038/nbt1318

International Stem Cell Initiative, Amps, K., Andrews, P. W., Anyfantis, G., Armstrong, L., Avery, S., … Zhou, Q. (2011). Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage. *Nature Biotechnology*, *29*(12), 1132–1144. https://doi.org/10.1038/nbt.2051

Invergo, B. M., & Beltrao, P. (2018). Reconstructing phosphorylation signalling networks from quantitative phosphoproteomic data. *Essays in Biochemistry*, *62*(4), 525–534.

https://doi.org/10.1042/EBC20180019

Ito, T., Williams, J. D., Fraser, D. J., & Phillips, A. O. (2004). Hyaluronan regulates transforming growth factor-beta1 receptor compartmentalization. *The Journal of Biological Chemistry*, *279*(24), 25326–25332. https://doi.org/10.1074/jbc.M403135200

Itoh, F., Divecha, N., Brocks, L., Oomen, L., Janssen, H., Calafat, J., … Dijke, P. Ten. (2002). The FYVE domain in Smad anchor for receptor activation (SARA) is sufficient for localization of SARA in early endosomes and regulates TGF-β/Smad signalling. *Genes to Cells*, *7*(3), 321–331. https://doi.org/10.1046/j.1365-2443.2002.00519.x

Iwata, K., Yumoto, K., Sugishima, M., Mizoguchi, C., Kai, Y., Iba, Y., & Mio, Y. (2014). Analysis of compaction initiation in human embryos by using time-lapse cinematography. *Journal of Assisted Reproduction and Genetics*, *31*(4), 421–426. https://doi.org/10.1007/s10815-014-0195-2

James, D., Levine, A. J., Besser, D., & Hemmati-Brivanlou, A. (2005). TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development (Cambridge, England)*, *132*(6), 1273–1282. https://doi.org/10.1242/dev.01706

Joubin, K., & Stern, C. D. (1999). Molecular interactions continuously define the organizer during the cell movements of gastrulation. *Cell*, *98*(5), 559–571. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10490096

Jovic, M., Sharma, M., Rahajeng, J., & Caplan, S. (2010). The early endosome: a busy sorting station for proteins at the crossroads. *Histology and Histopathology*, *25*(1), 99–112. https://doi.org/10.14670/HH-25.99

Jullien, J., & Gurdon, J. (2005). Morphogen gradient interpretation by a regulated trafficking step during ligand-receptor transduction. *Genes & Development*, *19*(22), 2682–2694. https://doi.org/10.1101/gad.341605

Kalluri, R., & Neilson, E. G. (2003). Epithelial-mesenchymal transition and its implications for fibrosis. *The Journal of Clinical Investigation*, *112*(12), 1776–1784. https://doi.org/10.1172/JCI20530

Kalluri, R., & Weinberg, R. A. (2009). The basics of epithelial-mesenchymal transition. *The Journal of Clinical Investigation*, *119*(6), 1420–1428. https://doi.org/10.1172/JCI39104

Kam, Y., & Quaranta, V. (2009). Cadherin-bound beta-catenin feeds into the Wnt pathway upon adherens junctions dissociation: evidence for an intersection between beta-catenin pools. *PloS One*, *4*(2), e4580. https://doi.org/10.1371/journal.pone.0004580

Kamachi, Y., Uchikawa, M., & Kondoh, H. (2000). Pairing SOX off: with partners in the regulation of embryonic development. *Trends in Genetics : TIG*, *16*(4), 182–187. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10729834

Kaneko, S., Chen, X., Lu, P., Yao, X., Wright, T. G., Rajurkar, M., … Xu, L. (2011). Smad inhibition by the Ste20 kinase Misshapen. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(27), 11127–11132. https://doi.org/10.1073/pnas.1104128108

Kavsak, P., Rasmussen, R. K., Causing, C. G., Bonni, S., Zhu, H., Thomsen, G. H., & Wrana, J. L. (2000). Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGFβ receptor for degradation. *Molecular Cell*, *6*(6), 1365–1375. https://doi.org/10.1016/S1097-2765(00)00134-9

Keirstead, H. S., Nistor, G., Bernal, G., Totoiu, M., Cloutier, F., Sharp, K., & Steward, O. (2005). Human embryonic stem cell-derived oligodendrocyte progenitor cell transplants remyelinate and restore locomotion after spinal cord injury. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *25*(19), 4694–4705. https://doi.org/10.1523/JNEUROSCI.0311-05.2005

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., & Haussler, D. (2002). The human genome browser at UCSC. *Genome Research*, *12*(6), 996–1006. https://doi.org/10.1101/gr.229102

Kilpinen, H., Goncalves, A., Leha, A., Afzal, V., Alasoo, K., Ashford, S., … Gaffney, D. J. (2017). Common genetic variation drives molecular heterogeneity in human iPSCs. *Nature*, *546*(7658), 370–375. https://doi.org/10.1038/nature22403

Kim, H., Wu, J., Ye, S., Tai, C., Zhou, X., Yan, H., … Ying, Q. (2013). Modulation of β-catenin function maintains mouse epiblast stem cell and human embryonic stem cell self-renewal. *Nature Communications*, *4*, 2403. https://doi.org/10.1038/ncomms3403

Kinoshita, E., & Kinoshita-Kikuta, E. (2011). Improved Phos-tag SDS-PAGE under neutral pH conditions for advanced protein phosphorylation profiling. *Proteomics*, *11*(2), 319–323. https://doi.org/10.1002/pmic.201000472

Kinoshita, E., Kinoshita-Kikuta, E., Takiyama, K., & Koike, T. (2006). Phosphate-binding tag, a new tool to visualize phosphorylated proteins. *Molecular & Cellular Proteomics : MCP*, *5*(4), 749–757. https://doi.org/10.1074/mcp.T500024-MCP200

Kinoshita, E., Takahashi, M., Takeda, H., Shiro, M., & Koike, T. (2004). Recognition of phosphate monoester dianion by an alkoxide-bridged dinuclear zinc(II) complex. *Dalton Transactions (Cambridge, England : 2003)*, (8), 1189–1193. https://doi.org/10.1039/b400269e

Kirchhausen, T. (2000). Clathrin. *Annual Review of Biochemistry*, *69*, 699–727. https://doi.org/10.1146/annurev.biochem.69.1.699

Kishigami, S., & Mishina, Y. (2005). BMP signaling and early embryonic patterning. *Cytokine & Growth Factor Reviews*, *16*(3), 265–278. https://doi.org/10.1016/j.cytogfr.2005.04.002

Kleinman, H. K., & Martin, G. R. (2005). Matrigel: basement membrane matrix with biological activity. *Seminars in Cancer Biology*, *15*(5), 378–386. https://doi.org/10.1016/j.semcancer.2005.05.004

Klimanskaya, I., Hipp, J., Rezai, K. A., West, M., Atala, A., & Lanza, R. (2004). Derivation and Comparative Assessment of Retinal Stem Cells Using Transcriptomics, *6*(3).

Knoblich, J. A. (2006). Sara splits the signal. *Science (New York, N.Y.)*, *314*(5802), 1094–1096. https://doi.org/10.1126/science.1135825

Kobayashi, T., Gu, F., & Gruenberg, J. (1998). Lipids, lipid domains and lipid-protein interactions in endocytic membrane traffic. *Seminars in Cell & Developmental Biology*, *9*(5), 517–526. https://doi.org/10.1006/scdb.1998.0257

Korolchuk, V. I., Menzies, F. M., & Rubinsztein, D. C. (2010). Mechanisms of cross-talk between the ubiquitin-proteasome and autophagy-lysosome systems. *FEBS Letters*, *584*(7), 1393–1398. https://doi.org/10.1016/j.febslet.2009.12.047

Kostaras, E., Sflomos, G., Pedersen, N. M., Stenmark, H., Fotsis, T., & Murphy, C. (2013). SARA and RNF11 interact with each other and ESCRT-0 core proteins and regulate degradative EGFR trafficking. *Oncogene*, *32*(44), 5220–5232. https://doi.org/10.1038/onc.2012.554

Kosugi, S., Hasebe, M., Entani, T., Takayama, S., Tomita, M., & Yanagawa, H. (2008).

Design of peptide inhibitors for the importin alpha/beta nuclear import pathway by activity-based profiling. *Chemistry & Biology*, *15*(9), 940–949. https://doi.org/10.1016/j.chembiol.2008.07.019

Kosugi, S., Hasebe, M., Matsumura, N., Takashima, H., Miyamoto-Sato, E., Tomita, M., & Yanagawa, H. (2009). Six classes of nuclear localization signals specific to different binding grooves of importin alpha. *The Journal of Biological Chemistry*, *284*(1), 478–485. https://doi.org/10.1074/jbc.M807017200

Kosugi, S., Hasebe, M., Tomita, M., & Yanagawa, H. (2009). Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(25), 10171–10176. https://doi.org/10.1073/pnas.0900604106

Koyanagi-Aoi, M., Ohnuki, M., Takahashi, K., Okita, K., Noma, H., Sawamura, Y., … Yamanaka, S. (2013). Differentiation-defective phenotypes revealed by large-scale analyses of human pluripotent stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(51), 20569–20574. https://doi.org/10.1073/pnas.1319061110

Kramps, T., Peter, O., Brunner, E., Nellen, D., Froesch, B., Chatterjee, S., … Basler, K. (2002). Wnt/wingless signaling requires BCL9/legless-mediated recruitment of pygopus to the nuclear beta-catenin-TCF complex. *Cell*, *109*(1), 47–60. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11955446

Kressmann, S., Campos, C., Castanon, I., Fürthauer, M., & González-Gaitán, M. (2015). Directional Notch trafficking in Sara endosomes during asymmetric cell division in the spinal cord. *Nature Cell Biology*, *17*(3), 333–339. https://doi.org/10.1038/ncb3119

Kroon, E., Martinson, L. A., Kadoya, K., Bang, A. G., Kelly, O. G., Eliazer, S., … Baetge, E. E. (2008). Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. *Nature Biotechnology*, *26*(4), 443–452. https://doi.org/10.1038/nbt1393

Kumari, S., Mg, S., & Mayor, S. (2010). Endocytosis unplugged: multiple ways to enter the cell. *Cell Research*, *20*(3), 256–275. https://doi.org/10.1038/cr.2010.19

Kurek, D., Neagu, A., Tastemel, M., Tüysüz, N., Lehmann, J., van de Werken, H. J. G., … Ten Berge, D. (2015). Endogenous WNT signals mediate BMP-induced and spontaneous differentiation of epiblast stem cells and human embryonic stem cells. *Stem Cell Reports*, *4*(1), 114–128. https://doi.org/10.1016/j.stemcr.2014.11.007

Kyrkou, A., Stellas, D., Syrrou, M., Klinakis, A., Fotsis, T., & Murphy, C. (2016). Generation of human induced pluripotent stem cells in defined, feeder-free conditions. *Stem Cell Research*, *17*(2), 458–460. https://doi.org/10.1016/j.scr.2016.05.006

Lachmann, A., & Ma'ayan, A. (2009). KEA: kinase enrichment analysis. *Bioinformatics (Oxford, England)*, *25*(5), 684–686. https://doi.org/10.1093/bioinformatics/btp026

Lemmon, M. A. (2004). Pleckstrin homology domains: not just for phosphoinositides. *Biochemical Society Transactions*, *32*(Pt 5), 707–711. https://doi.org/10.1042/BST0320707

Levanon, D., Goldstein, R. E., Bernstein, Y., Tang, H., Goldenberg, D., Stifani, S., … Groner, Y. (1998). Transcriptional repression by AML1 and LEF-1 is mediated by the TLE/Groucho corepressors. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(20), 11590–11595. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9751710

Li, J., Wang, G., Wang, C., Zhao, Y., Zhang, H., Tan, Z., … Deng, H. (2007). MEK/ERK signaling contributes to the maintenance of human embryonic stem cell self-renewal. *Differentiation; Research in Biological Diversity*, *75*(4), 299–307. https://doi.org/10.1111/j.1432-0436.2006.00143.x

Li, R., Liang, J., Ni, S., Zhou, T., Qing, X., Li, H., … Pei, D. (2010). A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell*, *7*(1), 51–63. https://doi.org/10.1016/j.stem.2010.04.014

Li, Z., & Chen, Y.-G. (2013). Functions of BMP signaling in embryonic stem cell fate determination. *Experimental Cell Research*, *319*(2), 113–119. https://doi.org/10.1016/j.yexcr.2012.09.016

Li, Z., Fei, T., Zhang, J., Zhu, G., Wang, L., Lu, D., … Chen, Y. (2012). BMP4 Signaling Acts via dual-specificity phosphatase 9 to control ERK activity in mouse embryonic stem cells. *Cell Stem Cell*, *10*(2), 171–182. https://doi.org/10.1016/j.stem.2011.12.016

Liang, J., Wan, M., Zhang, Y., Gu, P., Xin, H., Jung, S. Y., … Songyang, Z. (2008). Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. *Nature Cell Biology*, *10*(6), 731–739. https://doi.org/10.1038/ncb1736

Ling, N., Ying, S. Y., Ueno, N., Shimasaki, S., Esch, F., Hotta, M., & Guillemin, R. (1986). Pituitary FSH is released by a heterodimer of the beta-subunits from the two forms of inhibin. *Nature*, *321*(6072), 779–782. https://doi.org/10.1038/321779a0

Liu, C., Kato, Y., Zhang, Z., Do, V. M., Yankner, B. a, & He, X. (1999). beta-Trcp couples beta-catenin phosphorylation-degradation and regulates Xenopus axis formation. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(11), 6273–6278. https://doi.org/10.1073/pnas.96.11.6273

Liu, C., Li, Y., Semenov, M., Han, C., Baeg, G. H., Tan, Y., … He, X. (2002). Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism. *Cell*, *108*(6), 837–847. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11955436

Liu, N., Li, Z., Pei, D., & Shu, X. (2013). Zfyve9a regulates the proliferation of hepatic cells during zebrafish embryogenesis. *The International Journal of Developmental Biology*, *57*(9–10), 773–778. https://doi.org/10.1387/ijdb.130065xs

Logan, C. Y., & Nusse, R. (2004). The Wnt signaling pathway in development and disease. *Annual Review of Cell and Developmental Biology*, *20*, 781–810. https://doi.org/10.1146/annurev.cellbio.20.010403.113126

Loh, K. M., & Lim, B. (2011). A precarious balance: pluripotency factors as lineage specifiers. *Cell Stem Cell*, *8*(4), 363–369. https://doi.org/10.1016/j.stem.2011.03.013

López, E., Wang, X., Madero, L., López-Pascual, J., & Latterich, M. (2012). Functional phosphoproteomic mass spectrometry-based approaches. *Clinical and Translational Medicine*, *1*(1), 20. https://doi.org/10.1186/2001-1326-1-20

Loubéry, S., Daeden, A., Seum, C., Holtzer, L., Moraleda, A., Damond, N., … Gonzalez-Gaitan, M. (2017). Sara phosphorylation state controls the dispatch of endosomes from the central spindle during asymmetric division. *Nature Communications*, *8*, 15285. https://doi.org/10.1038/ncomms15285

Lu, W., Wang, G., Li, Y., Li, S., Song, X., Wang, X., … Yang, X. (2014). Autophagy functions on EMT in gastrulation of avian embryo. *Cell Cycle (Georgetown, Tex.)*, *13*(17), 2752–2764. https://doi.org/10.4161/15384101.2015.945850

Lu, Z., Murray, J. T., Luo, W., Li, H., Wu, X., Xu, H., … Chen, Y. (2002). Transforming growth factor beta activates Smad2 in the absence of receptor endocytosis. *The*

Journal of Biological Chemistry, *277*(33), 29363–29368. https://doi.org/10.1074/jbc.M203495200

Ludwig, T. E., Bergendahl, V., Levenstein, M. E., Yu, J., Probasco, M. D., & Thomson, J. A. (2006). Feeder-independent culture of human embryonic stem cells. *Nature Methods*, *3*(8), 637–646. https://doi.org/10.1038/nmeth902

Lv, W., Guo, X., Wang, G., Xu, Y., & Kang, J. (2014). Histone deacetylase 1 and 3 regulate the mesodermal lineage commitment of mouse embryonic stem cells. *PloS One*, *9*(11), e113262. https://doi.org/10.1371/journal.pone.0113262

LYON, M. F. (1961). Gene action in the X-chromosome of the mouse (Mus musculus L.). *Nature*, *190*, 372–373. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13764598

Macek, B., Mann, M., & Olsen, J. V. (2009). Global and site-specific quantitative phosphoproteomics: principles and applications. *Annual Review of Pharmacology and Toxicology*, *49*, 199–221. https://doi.org/10.1146/annurev.pharmtox.011008.145606

Maitra, A., Arking, D. E., Shivapurkar, N., Ikeda, M., Stastny, V., Kassauei, K., … Chakravarti, A. (2005). Genomic alterations in cultured human embryonic stem cells. *Nature Genetics*, *37*(10), 1099–1103. https://doi.org/10.1038/ng1631

Manning, G., Whyte, D. B., Martinez, R., Hunter, T., & Sudarsanam, S. (2002). The protein kinase complement of the human genome. *Science (New York, N.Y.)*, *298*(5600), 1912–1934. https://doi.org/10.1126/science.1075762

Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, *78*(12), 7634–7638.

Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6950406

Martinez-Quiles, N., Ho, H. H., Kirschner, M. W., Ramesh, N., & Geha, R. S. (2004). Erk/Src phosphorylation of cortactin acts as a switch on-switch off mechanism that controls its ability to activate N-WASP. *Molecular and Cellular Biology*, *24*(12), 5269–5280. https://doi.org/10.1128/MCB.24.12.5269-5280.2004

Martyn, I., Kanno, T. Y., Ruzo, A., Siggia, E. D., & Brivanlou, A. H. (2018). Self-organization of a human organizer by combined Wnt and Nodal signalling. *Nature*, *558*(7708), 132–135. https://doi.org/10.1038/s41586-018-0150-y

Massagué, J. (1998). TGF-beta signal transduction. *Annual Review of Biochemistry*, *67*, 753–791. https://doi.org/10.1146/annurev.biochem.67.1.753

Massagué, J. (2012). TGFβ signalling in context. *Nature Reviews. Molecular Cell Biology*, *13*(10), 616–630. https://doi.org/10.1038/nrm3434

Massagué, J., Seoane, J., & Wotton, D. (2005). Smad transcription factors. *Genes & Development*, *19*(23), 2783–2810. https://doi.org/10.1101/gad.1350705

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., … Niwa, H. (2007). Pluripotency governed by Sox2 via regulation of Oct3 / 4 expression in mouse embryonic stem cells, *9*(6). https://doi.org/10.1038/ncb1589

MCCULLOCH, E. A., & TILL, J. E. (1960). The radiation sensitivity of normal mouse bone marrow cells, determined by quantitative marrow transplantation into irradiated mice. *Radiation Research*, *13*, 115–125. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13858509

McManus, M. T., & Sharp, P. A. (2002). Gene silencing in mammals by small interfering RNAs. *Nature Reviews. Genetics*, *3*(10), 737–747. https://doi.org/10.1038/nrg908

Meckelein, B., Marshall, D. C., Conn, K. J., Pietropaolo, M., Van Nostrand, W., &

Abraham, C. R. (1998). Identification of a novel serine protease-like molecule in human brain. *Brain Research. Molecular Brain Research*, *55*(2), 181–197. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9582421

Melchior, K., Weiss, J., Zaehres, H., Kim, Y., Lutzko, C., Roosta, N., … Müschen, M. (2008). The WNT receptor FZD7 contributes to self-renewal signaling of human embryonic stem cells. *Biological Chemistry*, *389*(7), 897–903. https://doi.org/10.1515/BC.2008.108

Menasché, P., Vanneaux, V., Hagège, A., Bel, A., Cholley, B., Cacciapuoti, I., … Larghero, J. (2015). Human embryonic stem cell-derived cardiac progenitors for severe heart failure treatment: first clinical case report. *European Heart Journal*, *36*(30), 2011–2017. https://doi.org/10.1093/eurheartj/ehv189

Mendoza, P., Ortiz, R., Díaz, J., Quest, A. F. G., Leyton, L., Stupack, D., & Torres, V. A. (2013). Rab5 activation promotes focal adhesion disassembly, migration and invasiveness in tumor cells. *Journal of Cell Science*, *126*(Pt 17), 3835–3847. https://doi.org/10.1242/jcs.119727

Mestres, I., Chuang, J., Calegari, F., Conde, C., & Sung, C. (2016). SARA regulates neuronal migration during neocortical development through L1 trafficking. *Development (Cambridge, England)*, *143*(17), 3143–3153. https://doi.org/10.1242/dev.129338

Meyer, C., Liu, Y., Kaul, A., Peipe, I., & Dooley, S. (2013). Caveolin-1 abrogates TGF-β mediated hepatocyte apoptosis. *Cell Death & Disease*, *4*(1), e466. https://doi.org/10.1038/cddis.2012.204

Mitalipov, S., & Wolf, D. (2009). Totipotency, Pluripotency and Nuclear Reprogramming. In *Engineering of Stem Cells* (pp. 185–199). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/10_2008_45

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., … Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell*, *113*(5), 631–642. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12787504

Miura, S., Takeshita, T., Asao, H., Kimura, Y., Murata, K., Sasaki, Y., … Sugamura, K. (2000). Hgs (Hrs), a FYVE domain protein, is involved in Smad signaling through cooperation with SARA. *Molecular and Cellular Biology*, *20*(24), 9346–9355. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11094085

Miyazono, K. (2000). Positive and negative regulation of TGF-beta signaling. *Journal of Cell Science*, *113 (Pt 7*, 1101–1109.

Miyazono, K., Maeda, S., & Imamura, T. (2005). BMP receptor signaling: transcriptional targets, regulation of signals, and signaling cross-talk. *Cytokine & Growth Factor Reviews*, *16*(3), 251–263. https://doi.org/10.1016/j.cytogfr.2005.01.009

Montagne, C., & Gonzalez-Gaitan, M. (2014). Sara endosomes and the asymmetric division of intestinal stem cells. *Development (Cambridge, England)*, *141*(10), 2014–2023. https://doi.org/10.1242/dev.104240

Monteleon, C. L., Agnihotri, T., Dahal, A., Liu, M., Rebecca, V. W., Beatty, G. L., … Ridky, T. W. (2018). Lysosomes Support the Degradation, Signaling, and Mitochondrial Metabolism Necessary for Human Epidermal Differentiation. *The Journal of Investigative Dermatology*, *138*(9), 1945–1954. https://doi.org/10.1016/j.jid.2018.02.035

Moustakas, A., & Heldin, C. (2005). Non-Smad TGF-beta signals. *Journal of Cell Science*, *118*(Pt 16), 3573–3584. https://doi.org/10.1242/jcs.02554

Moustakas, A., Souchelnytskyi, S., & Heldin, C. H. (2001). Smad regulation in TGF-beta

signal transduction. *Journal of Cell Science*, *114*(Pt 24), 4359–4369. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11792802

Mu, F. T., Callaghan, J. M., Steele-Mortimer, O., Stenmark, H., Parton, R. G., Campbell, P. L., … Toh, B. H. (1995). EEA1, an early endosome-associated protein. EEA1 is a conserved alpha-helical peripheral membrane protein flanked by cysteine “fingers” and contains a calmodulin-binding IQ motif. *The Journal of Biological Chemistry*, *270*(22), 13503–13511. https://doi.org/10.1074/jbc.270.22.13503

Mullen, A. C., & Wrana, J. L. (2017). TGF-β Family Signaling in Embryonic and Somatic Stem-Cell Renewal and Differentiation. *Cold Spring Harbor Perspectives in Biology*, *9*(7), 1–45. https://doi.org/10.1101/cshperspect.a022186

Muñoz, J. P., Huichalaf, C. H., Orellana, D., & Maccioni, R. B. (2007). cdk5 modulates beta- and delta-catenin/Pin1 interactions in neuronal cells. *Journal of Cellular Biochemistry*, *100*(3), 738–749. https://doi.org/10.1002/jcb.21041

Murphy, R. F., Powers, S., & Cantor, C. R. (1984). Endosome pH measured in single cells by dual fluorescence flow cytometry: rapid acidification of insulin to pH 6. *The Journal of Cell Biology*, *98*(5), 1757–1762. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6144684

Murphy, S. J., Doré, J. J. E., Edens, M., Coffey, R. J., Barnard, J. A., Mitchell, H., … Leof, E. B. (2004). Differential trafficking of transforming growth factor-beta receptors and ligand in polarized epithelial cells. *Molecular Biology of the Cell*, *15*(6), 2853–2862. https://doi.org/10.1091/mbc.e04-02-0097

Myeku, N., & Figueiredo-Pereira, M. E. (2011). Dynamics of the degradation of ubiquitinated proteins by proteasomes and autophagy: association with sequestosome 1/p62. *The Journal of Biological Chemistry*, *286*(25), 22426–22440. https://doi.org/10.1074/jbc.M110.149252

Nagy, Z., Comer, S., & Smolenski, A. (2018). Analysis of Protein Phosphorylation Using Phos-Tag Gels. *Current Protocols in Protein Science*, *93*(1), e64. https://doi.org/10.1002/cpps.64

Nakaya, Y., & Sheng, G. (2008). Epithelial to mesenchymal transition during gastrulation: an embryological view. *Development, Growth & Differentiation*, *50*(9), 755–766. https://doi.org/10.1111/j.1440-169X.2008.01070.x

Naslavsky, N., & Caplan, S. (2018). The enigmatic endosome - sorting the ins and outs of endocytic trafficking. *Journal of Cell Science*, *131*(13). https://doi.org/10.1242/jcs.216499

Nguyen, K. T., Holloway, M. P., & Altura, R. A. (2012). The CRM1 nuclear export protein in normal development and disease. *International Journal of Biochemistry and Molecular Biology*, *3*(2), 137–151. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22773955

Niakan, K. K., & Eggan, K. (2013). Analysis of human embryos from zygote to blastocyst reveals distinct gene expression patterns relative to the mouse. *Developmental Biology*, *375*(1), 54–64. https://doi.org/10.1016/j.ydbio.2012.12.008

Niakan, K. K., Han, J., Pedersen, R. A., Simon, C., & Pera, R. A. R. (2012). Human pre-implantation embryo development. *Development (Cambridge, England)*, *139*(5), 829–841. https://doi.org/10.1242/dev.060426

Nichols, J., & Smith, A. (2009). Naive and primed pluripotent states. *Cell Stem Cell*, *4*(6), 487–492. https://doi.org/10.1016/j.stem.2009.05.015

Niwa, H., Burdon, T., Chambers, I., & Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes & Development*, *12*(13), 2048–2060. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9649508

Niwa, H., Miyazaki, J., & Smith, A. G. (2000). Quantitative expression of OCT3 / 4 defines differentiation , dedifferentiation or self-renewal of ES cells . Nat . Quantitative expression of Oct-3 / 4 defines differentiation , dedifferentiation or self-renewal of ES cells. *Nat Genet*, *24*(July), 2–6. https://doi.org/10.1038/74199

Noggle, S. A., Weiler, D., & Condie, B. G. (2006). Notch signaling is inactive but inducible in human embryonic stem cells. *Stem Cells (Dayton, Ohio)*, *24*(7), 1646–1653. https://doi.org/10.1634/stemcells.2005-0314

Norrman, K., Fischer, Y., Bonnamy, B., Wolfhagen Sand, F., Ravassard, P., & Semb, H. (2010). Quantitative comparison of constitutive promoters in human ES cells. *PloS One*, *5*(8), e12413. https://doi.org/10.1371/journal.pone.0012413

Norwitz, E. R., Schust, D. J., & Fisher, S. J. (2001). Implantation and the survival of early pregnancy. *The New England Journal of Medicine*, *345*(19), 1400–1408. https://doi.org/10.1056/NEJMra000763

Ong, S., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., & Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & Cellular Proteomics : MCP*, *1*(5), 376–386. https://doi.org/10.1074/mcp.M200025-MCP200

Ornitz, D. M., & Itoh, N. (2001). Fibroblast growth factors. *Genome Biology*, *2*(3), REVIEWS3005. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11276432

Palacios, F., Tushir, J. S., Fujita, Y., & D'Souza-Schorey, C. (2005). Lysosomal targeting of E-cadherin: a unique mechanism for the down-regulation of cell-cell adhesion during epithelial to mesenchymal transitions. *Molecular and Cellular Biology*, *25*(1), 389–402. https://doi.org/10.1128/MCB.25.1.389-402.2005

Palamidessi, A., Frittoli, E., Garré, M., Faretta, M., Mione, M., Testa, I., … Di Fiore, P.

P. (2008). Endocytic trafficking of Rac is required for the spatial restriction of signaling in cell migration. *Cell*, *134*(1), 135–147. https://doi.org/10.1016/j.cell.2008.05.034

Pálfy, M., Reményi, A., & Korcsmáros, T. (2012). Endosomal crosstalk: meeting points for signaling pathways. *Trends in Cell Biology*, *22*(9), 447–456. https://doi.org/10.1016/j.tcb.2012.06.004

Pan, G., Tian, S., Nie, J., Yang, C., Ruotti, V., Wei, H., … Thomson, J. A. (2007). Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell Stem Cell*, *1*(3), 299–312. https://doi.org/10.1016/j.stem.2007.08.003

Panopoulou, E., Gillooly, D. J., Wrana, J. L., Zerial, M., Stenmark, H., Murphy, C., & Fotsis, T. (2002). Early endosomal regulation of Smad-dependent signaling in endothelial cells. *The Journal of Biological Chemistry*, *277*(20), 18046–18052. https://doi.org/10.1074/jbc.M107983200

Papanayotou, C., & Collignon, J. (2014). Activin/Nodal signalling before implantation: setting the stage for embryo patterning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1657). https://doi.org/10.1098/rstb.2013.0539

Pauklin, S., & Vallier, L. (2015). Activin/Nodal signalling in stem cells. *Development (Cambridge, England)*, *142*(4), 607–619. https://doi.org/10.1242/dev.091769

Paule, M. R. (2000). SURVEY AND SUMMARY Transcription by RNA polymerases I and III. *Nucleic Acids Research*, *28*(6), 1283–1298. https://doi.org/10.1093/nar/28.6.1283

Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S., & Brockdorff, N. (1996).

Requirement for Xist in X chromosome inactivation. *Nature*, *379*(6561), 131–137. https://doi.org/10.1038/379131a0

Petropoulos, S., Edsgärd, D., Reinius, B., Deng, Q., Panula, S. P., Codeluppi, S., … Lanner, F. (2016). Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell*, *165*(4), 1012–1026. https://doi.org/10.1016/j.cell.2016.03.023

Pfeffer, S. R. (2001). Rab GTPases: specifying and deciphering organelle identity and function. *Trends in Cell Biology*, *11*(12), 487–491. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11719054

Pike, L. J. (2003). Lipid rafts: bringing order to chaos. *Journal of Lipid Research*, *44*(4), 655–667. https://doi.org/10.1194/jlr.R200021-JLR200

Piper, R. C., & Katzmann, D. J. (2007). Biogenesis and function of multivesicular bodies. *Annual Review of Cell and Developmental Biology*, *23*, 519–547. https://doi.org/10.1146/annurev.cellbio.23.090506.123319

Poteryaev, D., Datta, S., Ackema, K., Zerial, M., & Spang, A. (2010). Identification of the switch in early-to-late endosome transition. *Cell*, *141*(3), 497–508. https://doi.org/10.1016/j.cell.2010.03.011

Pyrzynska, B., Pilecka, I., & Miaczynska, M. (2009). Endocytic proteins in the regulation of nuclear signaling, transcription and tumorigenesis. *Molecular Oncology*, *3*(4), 321–338. https://doi.org/10.1016/j.molonc.2009.06.001

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. a, & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, *8*(11), 2281–2308. https://doi.org/10.1038/nprot.2013.143

Reines, D., Conaway, J. W., & Conaway, R. C. (1996). The RNA polymerase II general

elongation factors. *Trends in Biochemical Sciences*, *21*(9), 351–355. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8870500

Revach, O., & Geiger, B. (2014). The interplay between the proteolytic, invasive, and adhesive domains of invadopodia and their roles in cancer invasion. *Cell Adhesion & Migration*, *8*(3), 215–225. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24714132

Richter, A., Valdimarsdottir, L., Hrafnkelsdottir, H. E., Runarsson, J. F., Omarsdottir, A. R., Ward-van Oostwaard, D., … Valdimarsdottir, G. (2014). BMP4 promotes EMT and mesodermal commitment in human embryonic stem cells via SLUG and MSX2. *Stem Cells (Dayton, Ohio)*, *32*(3), 636–648. https://doi.org/10.1002/stem.1592

Runyan, C. E., Hayashida, T., Hubchak, S., Curley, J. F., & Schnaper, H. W. (2009). Role of SARA (SMAD anchor for receptor activation) in maintenance of epithelial cell phenotype. *The Journal of Biological Chemistry*, *284*(37), 25181–25189. https://doi.org/10.1074/jbc.M109.032847

Runyan, C. E., Liu, Z., & Schnaper, H. W. (2012). Phosphatidylinositol 3-kinase and Rab5 GTPase inversely regulate the Smad anchor for receptor activation (SARA) protein independently of transforming growth factor-β1. *The Journal of Biological Chemistry*, *287*(43), 35815–35824. https://doi.org/10.1074/jbc.M112.380493

Runyan, C. E., Schnaper, H. W., & Poncelet, A. (2005). The role of internalization in transforming growth factor beta1-induced Smad2 association with Smad anchor for receptor activation (SARA) and Smad2-dependent signaling in human mesangial cells. *The Journal of Biological Chemistry*, *280*(9), 8300–8308. https://doi.org/10.1074/jbc.M407939200

Saftig, P., & Klumperman, J. (2009). Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function. *Nature Reviews. Molecular Cell Biology*, *10*(9),

623–635. https://doi.org/10.1038/nrm2745

Saharinen, J., Hyytiäinen, M., Taipale, J., & Keski-Oja, J. (1999). Latent transforming growth factor-beta binding proteins (LTBPs)--structural extracellular matrix proteins for targeting TGF-beta action. *Cytokine & Growth Factor Reviews*, *10*(2), 99–117. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10743502

Sakaki-Yumoto, M., Liu, J., Ramalho-Santos, M., Yoshida, N., & Derynck, R. (2013). Smad2 is essential for maintenance of the human and mouse primed pluripotent stem cell state. *The Journal of Biological Chemistry*, *288*(25), 18546–18560. https://doi.org/10.1074/jbc.M112.446591

Samavarchi-Tehrani, P., Golipour, A., David, L., Sung, H., Beyer, T. A., Datti, A., … Wrana, J. L. (2010). Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell*, *7*(1), 64–77. https://doi.org/10.1016/j.stem.2010.04.015

Sann, S., Wang, Z., Brown, H., & Jin, Y. (2009). Roles of endosomal trafficking in neurite outgrowth and guidance. *Trends in Cell Biology*, *19*(7), 317–324. https://doi.org/10.1016/j.tcb.2009.05.001

Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., & Brivanlou, A. H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nature Medicine*, *10*(1), 55–63. https://doi.org/10.1038/nm979

Schmidt, O., & Teis, D. (2012). The ESCRT machinery. *Current Biology : CB*, *22*(4), R116-20. https://doi.org/10.1016/j.cub.2012.01.028

Schöler, H. R., Hatzopoulos, A. K., Balling, R., Suzuki, N., & Gruss, P. (1989). A family of octamer-specific proteins present during mouse embryogenesis: evidence for

germline-specific expression of an Oct factor. *The EMBO Journal*, *8*(9), 2543–2550. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2573523

Schwartz, S. D., Hubschman, J., Heilwell, G., Franco-Cardenas, V., Pan, C. K., Ostrick, R. M., … Lanza, R. (2012). Embryonic stem cell trials for macular degeneration: a preliminary report. *Lancet (London, England)*, *379*(9817), 713–720. https://doi.org/10.1016/S0140-6736(12)60028-2

Schwartz, S. D., Regillo, C. D., Lam, B. L., Eliott, D., Rosenfeld, P. J., Gregori, N. Z., … Lanza, R. (2015). Human embryonic stem cell-derived retinal pigment epithelium in patients with age-related macular degeneration and Stargardt's macular dystrophy: follow-up of two open-label phase 1/2 studies. *Lancet (London, England)*, *385*(9967), 509–516. https://doi.org/10.1016/S0140-6736(14)61376-3

Scita, G., & Di Fiore, P. P. (2010). The endocytic matrix. *Nature*, *463*(7280), 464–473. https://doi.org/10.1038/nature08910

Scott, C. C., Vacca, F., & Gruenberg, J. (2014). Endosome maturation, transport and functions. *Seminars in Cell & Developmental Biology*, *31*, 2–10. https://doi.org/10.1016/j.semcdb.2014.03.034

Shi, W., Sun, C., He, B., Xiong, W., Shi, X., Yao, D., & Cao, X. (2004). GADD34-PP1c recruited by Smad7 dephosphorylates TGFbeta type I receptor. *The Journal of Cell Biology*, *164*(2), 291–300. https://doi.org/10.1083/jcb.200307151

Shi, Y., & Massagué, J. (2003). Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell*, *113*(6), 685–700. Retrieved from https://www.cell.com/action/showPdf?pii=S0092-8674%2803%2900432-X

Shi, Y., Wang, Y. F., Jayaraman, L., Yang, H., Massagué, J., & Pavletich, N. P. (1998). Crystal structure of a Smad MH1 domain bound to DNA: insights on DNA binding in

TGF-beta signaling. *Cell*, *94*(5), 585–594. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9741623

Silva, J., & Smith, A. (2008). Capturing pluripotency. *Cell*, *132*(4), 532–536. https://doi.org/10.1016/j.cell.2008.02.006

Silva, P., Mendoza, P., Rivas, S., Díaz, J., Moraga, C., Quest, A. F. G., & Torres, V. A. (2016). Hypoxia promotes Rab5 activation, leading to tumor cell migration, invasion and metastasis. *Oncotarget*, *7*(20), 29548–29562. https://doi.org/10.18632/oncotarget.8794

Simonsen, A., Lippé, R., Christoforidis, S., Gaullier, J. M., Brech, A., Callaghan, J., … Stenmark, H. (1998). EEA1 links PI(3)K function to Rab5 regulation of endosome fusion. *Nature*, *394*(6692), 494–498. https://doi.org/10.1038/28879

Singh, A. M., Reynolds, D., Cliff, T., Ohtsuka, S., Mattheyses, A. L., Sun, Y., … 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*(3), 312–326. https://doi.org/10.1016/j.stem.2012.01.014

Song, W. K., Park, K., Kim, H., Lee, J. H., Choi, J., Chong, S. Y., … Lanza, R. (2015). Stem Cell Reports. *Stem Cell Reports*, *4*(5), 860–872. https://doi.org/10.1016/j.stemcr.2015.04.005

Souchelnytskyi, S., Tamaki, K., Engström, U., Wernstedt, C., ten Dijke, P., & Heldin, C. H. (1997). Phosphorylation of Ser465 and Ser467 in the C terminus of Smad2 mediates interaction with Smad4 and is required for transforming growth factor-beta signaling. *The Journal of Biological Chemistry*, *272*(44), 28107–28115. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9346966

Sozen, B., Can, A., & Demir, N. (2014). Cell fate regulation during preimplantation

development: A view of adhesion-linked molecular interactions. *Developmental Biology*, *395*(1), 73–83. https://doi.org/10.1016/j.ydbio.2014.08.028

Spemann, H., & Mangold, H. (2001). Induction of embryonic primordia by implantation of organizers from a different species. 1923. *The International Journal of Developmental Biology*, *45*(1), 13–38. https://doi.org/10.1007/BF02108133

Spits, C., Mateizel, I., Geens, M., Mertzanidou, A., Staessen, C., Vandeskelde, Y., … Sermon, K. (2008). Recurrent chromosomal abnormalities in human embryonic stem cells. *Nature Biotechnology*, *26*(12), 1361–1363. https://doi.org/10.1038/nbt.1510

Stenmark, H., Aasland, R., & Driscoll, P. C. (2002). The phosphatidylinositol 3-phosphate-binding FYVE finger. *FEBS Letters*, *513*(1), 77–84. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11911884

Strumpf, D., Mao, C.-A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., & Rossant, J. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. *Development (Cambridge, England)*, *132*(9), 2093–2102. https://doi.org/10.1242/dev.01801

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*(5), 861–872. https://doi.org/10.1016/j.cell.2007.11.019

Takahashi, K., & Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, *126*(4), 663–676. https://doi.org/10.1016/j.cell.2006.07.024

Tam, P. P. L., & Loebel, D. A. F. (2007). Gene function in mouse embryogenesis: get

set for gastrulation. *Nature Reviews. Genetics*, *8*(5), 368–381. https://doi.org/10.1038/nrg2084

Tam, P. P., Williams, E. A., & Chan, W. Y. (1993). Gastrulation in the mouse embryo: ultrastructural and molecular aspects of germ layer morphogenesis. *Microscopy Research and Technique*, *26*(4), 301–328. https://doi.org/10.1002/jemt.1070260405

Tang, W., Ling, G., Sun, L., Zhang, K., Zhu, X., Zhou, X., & Liu, F. (2015). Smad Anchor for Receptor Activation Regulates High Glucose-Induced EMT via Modulation of Smad2 and Smad3 Activities in Renal Tubular Epithelial Cells. *Nephron*, *130*(3), 213–220. https://doi.org/10.1159/000431105

Taurin, S., Sandbo, N., Qin, Y., Browning, D., & Dulin, N. O. (2006). Phosphorylation of beta-catenin by cyclic AMP-dependent protein kinase. *The Journal of Biological Chemistry*, *281*(15), 9971–9976. https://doi.org/10.1074/jbc.M508778200

Tesar, P. J., Chenoweth, J. G., Brook, F. A., Davies, T. J., Evans, E. P., Mack, D. L., … McKay, R. D. G. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature*, *448*(7150), 196–199. https://doi.org/10.1038/nature05972

Theunissen, T. W., Friedli, M., He, Y., Planet, E., O'Neil, R. C., Markoulaki, S., … Jaenisch, R. (2016). Molecular Criteria for Defining the Naive Human Pluripotent State. *Cell Stem Cell*, *19*(4), 502–515. https://doi.org/10.1016/j.stem.2016.06.011

Theveneau, E., & Mayor, R. (2012). Neural crest delamination and migration: from epithelium-to-mesenchyme transition to collective cell migration. *Developmental Biology*, *366*(1), 34–54. https://doi.org/10.1016/j.ydbio.2011.12.041

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 (New York, N.Y.)*, *282*(5391), 1145–1147. Retrieved from http://www.sciencemag.org/cgi/doi/10.1126/science.282.5391.1145

Thomson, J. A., & Marshall, V. S. (1998). Primate embryonic stem cells. *Current Topics in Developmental Biology*, *38*, 133–165. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9399078

TILL, J. E., & McCULLOCH, E. A. (1961). A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiation Research*, *14*, 213–222. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13776896

Tobe, B. T. D., Hou, J., Crain, A. M., Singec, I., Snyder, E. Y., & Brill, L. M. (2012). Phosphoproteomic analysis: an emerging role in deciphering cellular signaling in human embryonic stem cells and their differentiated derivatives. *Stem Cell Reviews*, *8*(1), 16–31. https://doi.org/10.1007/s12015-011-9317-8

Tsukazaki, T., Chiang, T. A., Davison, A. F., Attisano, L., & Wrana, J. L. (1998). SARA, a FYVE domain protein that recruits Smad2 to the TGFbeta receptor. *Cell*, *95*(6), 779–791. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9865696

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., … Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods*, *13*(9), 731–740. https://doi.org/10.1038/nmeth.3901

Ullmann, U., In't Veld, P., Gilles, C., Sermon, K., De Rycke, M., Van de Velde, H., … Liebaers, I. (2007). Epithelial-mesenchymal transition process in human embryonic stem cells cultured in feeder-free conditions. *Molecular Human Reproduction*, *13*(1), 21–32. https://doi.org/10.1093/molehr/gal091

Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S., & Gregory, P. D. (2010). Genome editing with engineered zinc finger nucleases. *Nature Reviews. Genetics*, *11*(9), 636–646. https://doi.org/10.1038/nrg2842

Vale, W., Rivier, J., Vaughan, J., McClintock, R., Corrigan, A., Woo, W., … Spiess, J. (1986). Purification and characterization of an FSH releasing protein from porcine ovarian follicular fluid. *Nature*, *321*(6072), 776–779. https://doi.org/10.1038/321776a0

Vallier, L., Alexander, M., & Pedersen, R. A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *Journal of Cell Science*, *118*(Pt 19), 4495–4509. https://doi.org/10.1242/jcs.02553

Vallier, L., Mendjan, S., Brown, S., Chng, Z., Teo, A., Smithers, L. E., … Pedersen, R. A. (2009). Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. *Development*, *136*(8), 1339–1349. https://doi.org/10.1242/dev.033951

Vallier, L., Reynolds, D., & Pedersen, R. A. (2004). Nodal inhibits differentiation of human embryonic stem cells along the neuroectodermal default pathway. *Developmental Biology*, *275*(2), 403–421. https://doi.org/10.1016/j.ydbio.2004.08.031

Vallier, L., Touboul, T., Chng, Z., Brimpari, M., Hannan, N., Millan, E., … Pedersen, R. A. (2009). Early cell fate decisions of human embryonic stem cells and mouse epiblast stem cells are controlled by the same signalling pathways. *PloS One*, *4*(6), e6082. https://doi.org/10.1371/journal.pone.0006082

Van Hoof, D., Muñoz, J., Braam, S. R., Pinkse, M. W. H., Linding, R., Heck, A. J. R., … Krijgsveld, J. (2009). Phosphorylation dynamics during early differentiation of human embryonic stem cells. *Cell Stem Cell*, *5*(2), 214–226. https://doi.org/10.1016/j.stem.2009.05.021

Waddington, S. N., Privolizzi, R., Karda, R., & O'Neill, H. C. (2016). A Broad Overview and Review of CRISPR-Cas Technology and Stem Cells. *Current Stem Cell Reports*, *2*, 9–20. https://doi.org/10.1007/s40778-016-0037-5

Wagner, D. O., Sieber, C., Bhushan, R., Börgermann, J. H., Graf, D., & Knaus, P. (2010). BMPs: from bone to body morphogenetic proteins. *Science Signaling*, *3*(107), mr1. https://doi.org/10.1126/scisignal.3107mr1

Wang, L., Schulz, T. C., Sherrer, E. S., Dauphin, D. S., Shin, S., Nelson, A. M., … Robins, A. J. (2007). Self-renewal of human embryonic stem cells requires insulin-like growth factor-1 receptor and ERBB2 receptor signaling. *Blood*, *110*(12), 4111–4119. https://doi.org/10.1182/blood-2007-03-082586

Wang, R. N., Green, J., Wang, Z., Ye, J., Yan, Z., Denduluri, S., … Shi, L. L. (2014). ScienceDirect Bone Morphogenetic Protein (BMP) signaling in development and human diseases. *Genes & Diseases*, *1*(1), 87–105. https://doi.org/10.1016/j.gendis.2014.07.005

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*(4), 440–454. https://doi.org/10.1016/j.stem.2012.02.016

Warmflash, A., Sorre, B., Etoc, F., Siggia, E. D., & Brivanlou, A. H. (2014). A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. *Nature Methods*, *11*(8), 847–854. https://doi.org/10.1038/nmeth.3016

Watson, A. J., & Barcroft, L. C. (2001). Regulation of blastocyst formation. *Frontiers in Bioscience : A Journal and Virtual Library*, *6*, D708-30. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11333210

Weiss, A., & Attisano, L. (2013). The TGFbeta superfamily signaling pathway. *Wiley*

Interdisciplinary Reviews. Developmental Biology, *2*(1), 47–63. https://doi.org/10.1002/wdev.86

Wibo, M., & Poole, B. (1974). Protein degradation in cultured cells. II. The uptake of chloroquine by rat fibroblasts and the inhibition of cellular protein degradation and cathepsin B1. *The Journal of Cell Biology*, *63*(2 Pt 1), 430–440. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4607946

Wiechens, N., & Fagotto, F. (2001). CRM1- and Ran-independent nuclear export of β-catenin. *Current Biology*, *11*(1), 18–27. https://doi.org/10.1016/S0960-9822(00)00045-2

Wilkinson, D. G., Bhatt, S., & Herrmann, B. G. (1990). Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature*, *343*(6259), 657–659. https://doi.org/10.1038/343657a0

Williams, R. L., Hilton, D. J., Pease, S., Willson, T. a, Stewart, C. L., Gearing, D. P., … Gough, N. M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature*, *336*(6200), 684–687. https://doi.org/10.1038/336684a0

Winnier, G., Blessing, M., Labosky, P. A., & Hogan, B. L. (1995). Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. *Genes & Development*, *9*(17), 2105–2116. https://doi.org/10.1101/gad.9.17.2105

Wu, Q., Zhang, L., Su, P., Lei, X., Liu, X., Wang, H., … Zhou, J. (2015). MSX2 mediates entry of human pluripotent stem cells into mesendoderm by simultaneously suppressing SOX2 and activating NODAL signaling. *Cell Research*, *25*(12), 1314–1332. https://doi.org/10.1038/cr.2015.118

Wu, X., Tu, X., Joeng, K. S., Hilton, M. J., Williams, D. A., & Long, F. (2008). Rac1

activation controls nuclear localization of beta-catenin during canonical Wnt signaling. *Cell*, *133*(2), 340–353. https://doi.org/10.1016/j.cell.2008.01.052

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*(6), 1476–1486. https://doi.org/10.1634/stemcells.2005-0299

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. *Nature Biotechnology*, *19*(10), 971–974. https://doi.org/10.1038/nbt1001-971

Xu, P., Liu, J., & Derynck, R. (2012). Post-translational regulation of TGF-β receptor and Smad signaling. *FEBS Letters*, *586*(14), 1871–1884. https://doi.org/10.1016/j.febslet.2012.05.010

Xu, R., Chen, X., Li, D. S., Li, R., Addicks, G. C., Glennon, C., … Thomson, J. A. (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nature Biotechnology*, *20*(12), 1261–1264. https://doi.org/10.1038/nbt761

Xu, R. H., Sampsell-Barron, T. L., Gu, F., Root, S., Peck, R. M., Pan, G., … Thomson, J. A. (2008). NANOG Is a Direct Target of TGFβ/Activin-Mediated SMAD Signaling in Human ESCs. *Cell Stem Cell*, *3*(2), 196–206. https://doi.org/10.1016/j.stem.2008.07.001

Xu, R., 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, *2*(3), 185–190. https://doi.org/10.1038/NMETH744

Xu, Z., Robitaille, A. M., Berndt, J. D., Davidson, K. C., Fischer, K. A., Mathieu, J., … Moon, R. T. (2016). Wnt/β-catenin signaling promotes self-renewal and inhibits the

primed state transition in naïve human embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(42), E6382–E6390. https://doi.org/10.1073/pnas.1613849113

Yakymovych, I., Yakymovych, M., & Heldin, C. (2018). Intracellular trafficking of transforming growth factor β receptors. *Acta Biochimica et Biophysica Sinica*, *50*(1), 3–11. https://doi.org/10.1093/abbs/gmx119

Yamada, K., Nomura, N., Yamano, A., Yamada, Y., & Wakamatsu, N. (2012). Identification and characterization of splicing variants of PLEKHA5 (Plekha5) during brain development. *Gene*, *492*(1), 270–275. https://doi.org/10.1016/j.gene.2011.10.018

Yamada, S., Nakamura, H., Kinoshita, E., Kinoshita-Kikuta, E., Koike, T., & Shiro, Y. (2007). Separation of a phosphorylated histidine protein using phosphate affinity polyacrylamide gel electrophoresis. *Analytical Biochemistry*, *360*(1), 160–162. https://doi.org/10.1016/j.ab.2006.10.005

Yang, T., Cui, H., Wen, M., Zuber, J., Kogan, S. C., & Wei, G. (2018). TCEA1 regulates the proliferative potential of mouse myeloid cells. *Experimental Cell Research*, *370*(2), 551–560. https://doi.org/10.1016/j.yexcr.2018.07.020

Yang, Y., Adachi, K., Sheridan, M. A., Alexenko, A. P., Schust, D. J., Schulz, L. C., … Roberts, R. M. (2015). Heightened potency of human pluripotent stem cell lines created by transient BMP4 exposure. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(18), E2337-46. https://doi.org/10.1073/pnas.1504778112

Yashiroda, Y., & Yoshida, M. (2003). Nucleo-cytoplasmic transport of proteins as a target for therapeutic drugs. *Current Medicinal Chemistry*, *10*(9), 741–748. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12678777

Yin, M., Ma, W., & An, L. (2017). Cortactin in cancer cell migration and invasion. *Oncotarget*, *8*(50), 88232–88243. https://doi.org/10.18632/oncotarget.21088

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*(3), 281–292. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14636556

Yoo, Y. D., Huang, C. T., Zhang, X., Lavaute, T. M., & Zhang, S.-C. (2011). Fibroblast growth factor regulates human neuroectoderm specification through ERK1/2-PARP-1 pathway. *Stem Cells (Dayton, Ohio)*, *29*(12), 1975–1982. https://doi.org/10.1002/stem.758

Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., … Thomson, J. A. (2007). Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. *Science*, *318*(5858), 1917–1920. https://doi.org/10.1126/science.1151526

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*(3), 326–334. https://doi.org/10.1016/j.stem.2011.01.001

Yu, X., Zou, J., Ye, Z., Hammond, H., Chen, G., Tokunaga, A., … Cheng, L. (2008). Notch signaling activation in human embryonic stem cells is required for embryonic, but not trophoblastic, lineage commitment. *Cell Stem Cell*, *2*(5), 461–471. https://doi.org/10.1016/j.stem.2008.03.001

Zeineddine, D., Hammoud, A. A., Mortada, M., & Boeuf, H. (2014). The Oct4 protein: more than a magic stemness marker. *American Journal of Stem Cells*, *3*(2), 74–82. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/25232507

Zeisberg, M., & Neilson, E. G. (2009). Biomarkers for epithelial-mesenchymal transitions. *The Journal of Clinical Investigation*, *119*(6), 1429–1437. https://doi.org/10.1172/JCI36183

Zhang, M., Wang, M., Tan, X., Li, T.-F., Zhang, Y. E., & Chen, D. (2010). Smad3 prevents beta-catenin degradation and facilitates beta-catenin nuclear translocation in chondrocytes. *The Journal of Biological Chemistry*, *285*(12), 8703–8710. https://doi.org/10.1074/jbc.M109.093526

Zhang, P., Li, J., Tan, Z., Wang, C., Liu, T., Chen, L., … Deng, H. (2008). Short-term BMP-4 treatment initiates mesoderm induction in human embryonic stem cells. *Blood*, *111*(4), 1933–1941. https://doi.org/10.1182/blood-2007-02-074120

Zhang, S., Fei, T., Zhang, L., Zhang, R., Chen, F., Ning, Y., … Chen, Y. (2007). Smad7 antagonizes transforming growth factor beta signaling in the nucleus by interfering with functional Smad-DNA complex formation. *Molecular and Cellular Biology*, *27*(12), 4488–4499. https://doi.org/10.1128/MCB.01636-06

Zhang, X., Fang, A., Riley, C. P., Wang, M., Regnier, F. E., & Buck, C. (2010). Multi-dimensional liquid chromatography in proteomics--a review. *Analytica Chimica Acta*, *664*(2), 101–113. https://doi.org/10.1016/j.aca.2010.02.001

Zhang, X. L., Topley, N., Ito, T., & Phillips, A. (2005). Interleukin-6 regulation of transforming growth factor (TGF)-beta receptor compartmentalization and turnover enhances TGF-beta1 signaling. *The Journal of Biological Chemistry*, *280*(13), 12239–12245. https://doi.org/10.1074/jbc.M413284200

Zhang, Y. E. (2017). Non-Smad Signaling Pathways of the TGF-β Family. *Cold Spring Harbor Perspectives in Biology*, *9*(2). https://doi.org/10.1101/cshperspect.a022129

Zhao, H., & Jin, Y. (2017). Signaling networks in the control of pluripotency. *Current*

Opinion in Genetics & Development, *46*, 141–148. https://doi.org/10.1016/j.gde.2017.07.013

Zhen, Y., & Stenmark, H. (2015). Cellular functions of Rab GTPases at a glance. *Journal of Cell Science*, *128*(17), 3171–3176. https://doi.org/10.1242/jcs.166074

Zuo, W., & Chen, Y. (2009). Specific activation of mitogen-activated protein kinase by transforming growth factor-beta receptors in lipid rafts is required for epithelial cell plasticity. *Molecular Biology of the Cell*, *20*(3), 1020–1029. https://doi.org/10.1091/mbc.e08-09-0898