DISS. ETH NO 26431

Unravelling molecular drivers of human embryonic stem cell pluripotency and early cell fate specification

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

Panagiota Tsikrika

M.Sc., Karolinska Institute

born on 26.05.1987

citizen of Greece

accepted on the recommendation of

Prof. Anton Wutz
Prof. Ulrike Kutay
Prof. Laurent David
Dr. Matthias Gstaiger
Dr. Tobias Beyer

2019

To my family

1. Table of Contents

1   Summary

Human pluripotent stem cells (hPSCs) have the ability to self-renew indefinitely in culture, as well as differentiate towards somatic cells and extraembryonic tissues in appropriate culture conditions. These properties make stem cells (SCs) central to emerging concepts in modern medicine, by providing an invaluable source of cells for basic research, disease modelling, drug testing and ultimately cell replacement therapies. The core transcription factors (TF) OCT4, SOX2 and NANOG maintain the hPSC state by occupying specific regulatory regions to control the expression of both pluripotency and pro-differentiation genes. This delicate balance is maintained by the extracellular signaling cues of the Fibroblast Growth Factor (FGF) and TGF-β families. Their spatiotemporal pattern of activity controls the re-localization of downstream TF complexes to specific regulatory elements and enhancers. Such enhancers are called "poised" as these are bookmarked in pluripotency with a unique chromatin signature. To gain deeper mechanistic insights in binding events on those enhancers during cell fate decisions, we took advantage of mass spectrometry (MS) - coupled to DNA pull down. For this, enhancer regions of genes important for pluripotency or Primitive streak (PS) and Definitive Endoderm (DE) formation were labelled by PCR amplification using biotinylated oligonucleotides and the resulting biotinylated fragments were incubated with cell lysates from pluripotent human embryonic stem cells (hESCs) or differentiated cells, followed by DNA pull down. The differentially bound proteins between the two states were analysed by MS and classified according to their enrichment over a control sequence in the different conditions. In this manner, we identified Transcriptional Adaptor 2-beta (TADA2B), as a novel protein binding to the enhancer of *NANOG*. We uncovered an essential role of TADA2B in maintenance of pluripotency. hESCs with deletion of *TADA2B* exhibit reduced expression of pluripotency factors *OCT4* and *SOX2*, upregulation of Mesendoderm genes *GATA6, GSC, T, EOMES* and phenotypically resemble endoderm-like cells. TADA2B is a member of the STAGA complex which is involved in several processes and especially in development and cancer. Therefore, the dissection of the mechanism of action of this complex and identification of its partners could shed light on cancer research and embryonic development. In conclusion, in this study, we captured binding events on active and poised enhancers in hESCs and discovered a protein that has an essential

role in self-renewal of hESCs and possibly in embryo development.

2 Zusammenfassung

Humane pluripotente Stammzellen definieren sich durch ihre Fähigkeit, sich in Kultur auf unbestimmte Zeit selbst zu erneuern und sich bei geeigneten Kulturbedingungen in somatische Zellen und in extraembryonales Gewebe zu differenzieren. Diese Eigenschaften machen Stammzellen zum Dreh- und Angelpunkt neuer Ansätze in der modernen Medizin. Dies basiert auf der Tatsache, dass sie eine unversiegbare Quelle für Zellen für Grundlagenforschung, Krankheitsmodellierung, Arzneimitteltests und letztendlich für Zellersatztherapien bieten. Auf der molekularen Ebene, erhält ein Komplex, bestehend aus den Grundtranskriptionsfaktoren OCT4, SOX2 und NANOG den hPSC-Zustand aufrecht, indem er spezifische regulatorische Regionen bindet um die Expression von Pluripotenz- und Prodifferenzierungsgenen zu regulieren. Dieses empfindliche Gleichgewicht wird durch die extrazellulären Signalwege der Fibroblast Growth Factor (FGF) - und TGF-β-Familien aufrechterhalten. Ein komplexes räumliches und zeitliches Aktivitätsmuster dieser Wachstumsfaktoren steuert die Bindung von TF-Komplexen an spezifischen regulatorischen Elementen, welche als Enhancer bezeichnet werden. Diese Enhancer werden als "poised" bezeichnet, da diese im pluripotenten Zustand eine sowohl aktive wie auch repressive Chromatin-Zusammensetzung aufweisen. Um tiefere mechanistische Erkenntnisse über Bindungsereignisse an diesen Enhancern während der Differenzierung von pluripotenten SCs zu gewinnen, setzten wir Massenspektrometrie gekoppelt an DNS-Pulldown ein. Hierzu wurden Enhancerregionen von Genen, die für die Pluripotenz oder Primitive streak (PS) - und Definitive Endoderm (DE) -Bildung wichtig sind, mittels PCR-Amplifikation mit biotinylierten Oligonukleotiden markiert und die resultierenden biotinylierten Fragmente mit Zelllysaten aus hPSCs oder differenzierten Zellen inkubiert. Anschliessend wurden assoziierte Proteine präzipitiert und mittels MS differentiell gebundene Faktoren identifiziert. Die unterschiedlich gebundenen Proteine zwischen den beiden Konditionen wurden mit MS analysiert und entsprechend nach der Anreicherung gegenüber eine Kontrolsequenz klassifiziert Auf diese Weise identifizierten wir TADA2B als neues Protein, das an den Enhancer von NANOG bindet. Eine detaillierte Analyse der Funktion von TADA2B zeigte, dass TADA2B eine wesentliche Rolle bei der Aufrechterhaltung der Pluripotenz spielt. hPSCs mit einer Mutation im TADA2B Gens, zeigen eine verringerte Expression

der Pluripotenzfaktoren OCT4 und SOX2, eine Induktion der Mesendoderm-Gene GATA6, GSC, T, EOMES und ähneln phänotypisch endodermartigen Zellen. TADA2B ist Mitglied des STAGA-Komplexes, der an mehreren Prozessen beteiligt ist, insbesondere an der Entwicklung und an Krebs. Die Aufschlüsselung des molekularen Wirkmechanismus dieses Komplexes und die Identifizierung seiner Interaktionspartner könnten daher Aufschluss nicht nur über die Steuerung der humanen embryonalen Entwicklung, sondern auch in die Krebsforschung geben. Zusammenfassend haben wir in dieser Studie die Bindung von TF an einzelnen DNS loci auf aktiven und „poised" Enhancern in hPSCs systematisch erfasst und ein Protein entdeckt, das eine wesentliche Rolle bei der Selbsterneuerung von hESCs und möglicherweise bei der Embryonalentwicklung spielt.

3 List of Abbreviations

ICM	Inner cell mass
PrE	Primitive endoderm
TE	Trophectoderm
FGF	Fibroblast growth factor
ERK	Extracellular-signal-regulated kinase
iPSCs	Induced pluripotent stem cells
PSCs	Pluripotent stem cells
hPSCs	Human Pluripotent stem cells
hESCs	Human embryonic stem cells
mESCs	Mouse embryonic stem cells
hSPCs	Hematopoietic stem and progenitor cells
PS	Primitive streak
APS	Anterior primitive streak
DE	Definitive endoderm
ME	Mesendoderm
BMP	Bone morphogenetic protein
TGFβ1	Transforming growth factor β 1
MEFs	Mouse embryonic fibroblasts
LIF	Leukemia inhibitory factor
OSN	OCT4, SOX2 and NANOG
Chip-seq	Chromatin Immunoprecipitation-sequencing
ko	Knockout
SC	Stem cell
T	Brachyury
Gsc	Goosecoid
ZFN	Zinc finger nucleases
MS	Mass spectrometry
EMT	Epithelial-to-mesenchymal-transition
SBE	SMAD binding elements
R-SMAD	Receptor-activated SMAD
LOF	Loss-of-function
CRISPRi	CRISPR interference
siRNA	Small interfering RNA
IgG	Immunoglobulin G
PCR	Polymerase chain reaction
APS	Ammonium Persulfate
FA	Formic acid
FDR	False discovery rate
ABC	Ammonium Bicarbonate
dIdC	Poly (deoxyinosinic-deoxycytidylic) acid sodium salt
CRISPR	Clustered regularly interspaced palindromic repeats
H3K27ac	Histone 3 lysine 27 acetylation
H3K27me3	Histone 3 lysine 27 methylation
H3K9ac	Histone 3 lysine 9 acetylation
H3K4me3	Histone 3 lysine 4 tri- methylation
H3K4me1	Histone 3 lysine 4 mono methylation
WT	Wild type
TALEN	Transcription activator-like effector nuclease
LC-MS	Liquid chromatography–mass spectrometry
RT	Room temperature

4 Introduction

4.1 Early mammalian Embryogenesis

Since ancient times scientists and philosophers have been curious about embryonic development. One of the first documented descriptions of the formation of different organs was made by the Greek philosopher Aristotle (384–322 b.c.) in his book: On the generation of Animals [1]. Since then, humans have used a specific set of animals, such as *Caenorhabditis elegans*, *Drosophila melanogaster, Mus musculus, Xenopus laevis, Danio rerio* and *Gallus domesticus,* to experimentally study mammalian embryonic development. This indicates that most fundamental developmental processes remain impervious to evolution.

The embryonic development begins with the penetration of the zona pellucida of the oocyte by the sperm. This results in the fusion of the pronuclei of these two cells and merge of their chromosomal content in a diploid cell, the zygote. A series of (zygotic) cleavages start, leading to formation of new smaller cells, the blastomeres, without affecting the total volume of the embryo. At this point each blastomere is totipotent, able to generate an individual organism [2, 3]. Embryo compaction at the 16- to 32-cell stage (3-4 days after fertilization) generates a sphere of cells called morula [4], originating from the Latin word *morus* which means mulberry.

A Na+ gradient causes fluid influx inside the embryo leading to the formation of a cavity, the blastocoel [5]. At this stage the embryo is called blastocyst with a clear distinction of different cell structures and layers. The cells inside the sphere are arranging closer to each other on one side of the blastocyst and form the inner cell mass (ICM) [6]. The ICM is the source of embryonic stem cells which then can be isolated and transferred to an *in vitro* culture [7, 8]. They heterogeneously express the epiblast marker *Nanog* or primitive endoderm marker (PrE) *Gata6* (Figure 4-1) [2, 3].

Figure 4-1. Establishment of Trophectoderm (TE), Epiblast and Primitive endoderm layers during mouse development. ICM cells show mosaic expression of GATA6 and NANOG. GATA6 positive cells will give rise to PrE, whereas NANOG positive cells to epiblast. Modified from [9].

Negative feedback by Fibroblast Growth Factor (FGF) signaling between neighbouring ICM cells results in repression of *Nanog* by highly phosphorylated Extracellular-signal-Regulated Kinase (ERK) and activation of primitive endoderm markers. On the other hand, *Nanog* is de-repressed in cells with lower ERK activity leading to inhibition of PrE factors [2, 3].

The blastocyst is circulated by an outer polarized epithelial layer of cells, the trophoblast which will contribute to the placenta. The specification of the ICM and trophectoderm (TE), inner and outer identity, is regulated by the activity of the NOTCH and HIPPO pathways [3, 4, 10]. Inactive HIPPO signaling in outer cells of the embryo allows YAP to shuttle to the nucleus. There, its interaction with TEAD4 leads to transcriptional activation of outer cell determinants such as *Cdx2* and the repression of ICM factor *Oct4* [4]. In the cells of the ICM the HIPPO signaling is active resulting in phosphorylation of YAP and its sequestration in the cytoplasm. The interaction with TEAD4 is lost, resulting in repression of TE genes [3].

Before gastrulation the amniotic cavity has been formed leading to separation of ICM into epiblast and hypoblast or PrE [3]. The epiblast cells surrounded by TE cells on the outside and PrE cells on the inner part of the blastocyst cavity are pluripotent and express OCT4, SOX2 and NANOG, the core pluripotency factors [3, 11]. The primitive endoderm cells spread around the inner surface of the cavity facing the epiblast. They will later contribute to extraembryonic endoderm and the yolk sac.

During this time the embryo travels through the fallopian tubes until it reaches the uterus at the blastocyst stage (5-6 days after fertilization). The zona pellucida then breaks and the trophectoderm hatches in the uterus allowing the embryo to implant there. During gastrulation the epiblast will give rise to the primary germ layers. At first, the primitive streak (PS) emerges in the posterior part of the epiblast. This process is a hallmark of gastrulation and is regulated by signals from BMP, Nodal, WNT and FGF.

As the primitive streak elongates towards the distal end of the epiblast, cells undergo a coordinated epithelial-to-mesenchymal-transition (EMT) and start to migrate through the lateral portions of the primitive streak displacing cells from the primitive endoderm. These cells express *Foxa2* and give rise to definitive endoderm where gut, lungs, pancreas, liver and thyroid originate from. Cells migrating anteriorly between the newly formed definitive endoderm and the epiblast, form the mesoderm which will later give rise to tissues like heart, blood, endothelium, muscle cells [12]. Anterior epiblast cells that will not ingress through the primitive streak will form the ectoderm and later skin, brain and nervous tissues [4, 13, 14].

Figure 4-2. Timeline of mouse and human development after fertilization of the zygote. Modified from [15].

4.2 Human pluripotent stem cells (hPSCs)

4.2.1 History and derivation of hPSCs

Pluripotent stem cells (PSCs) are characterized by two hallmark features; self-renewal and differentiation into all cell types of the embryo. hPSCs can either be derived from ICMs of embryos which are then called human embryonic stem cells (hESCs), or recreated from somatic cells, by generating induced pluripotent stem cells (iPSCs) [16]. The derivation of the first hESC line was reported in 1998 by Dr. J. A Thomson. His team isolated ICMs from IVF-donated embryos and derived five hESCs lines, including the H9 cell line that has been extensively used

in research since then [8]. hESCs like their mouse counterpart cells, can form embryonic-like tumors, called teratomas, consisting of differentiating tissues of the three germ layers, upon transplantation to immunodeficient mice [17, 18]. Since injection of hESCs to human blastocyst and generation of chimeric embryos is impractical and hinders ethical limitations, the teratoma formation is an assessment test for differentiation potential of hESCs and induced pluripotent stem cells (iPSCs) *in vivo* [19-21]. However, although this assay is useful to evaluate hESC potency, it cannot prove the full definition of a stem cell, which is a cell contributing to all germ layers, including germ line in *vivo,* which so far has only been proven for mESCs [22].

The derivation of hESCs came almost two decades after the derivation of mouse ESCs [7, 23], mainly due to restrictive access of researchers to human embryos [24]. However, once it became obvious the potential clinical benefit of hESCs, there was huge progress in developing optimal culture conditions for the maintenance and expansion of those cells. Initially hESCs were cultured in bovine serum containing medium and grown on mitotically inactivated fibroblasts to support their growth. These culturing conditions were suboptimal for future clinical applications due to risk of infections coming from animal ingredients (zoonosis). Since then, improvements have been made to culture systems by transiting to xenofree cultures [25, 26]. In 2005, scientists succeeded in maintaining hESC in feeder-free culture by using extracellular matrix from mouse embryonic fibroblasts (MEFs) [27].

Figure 4-3. *In vitro* derivation of pluripotent cells. The totipotent zygote gives rise to the pluripotent blastocyst containing the epiblast cells of the ICM. Pluripotent ESCs can be derived from the ICM and maintained *in vitro* under defined conditions. Modified from [28]**.**

Soon after, T. Ludwig and colleagues discovered the basic ingredients essential for the maintenance of hESCs in a feeder-free culture; bFGF and transforming growth factor β 1 (TGFβ1) [29, 30]. This discovery was a hallmark in the hESC field and led to the formulation of mTeSR1, the most widely used hESC culturing media [30]. Nowadays there are several options for hESCs defined and xenofree culturing media such as Essential 8, mTeSR™ plus and feeder-free matrix for their support such as Geltrex, matrigel and laminin. Soon after the derivation of hESCs it became evident that hESCs rely on different signaling for the maintenance of their pluripotency state compared to mouse embryonic stem cells (mESCs) [31], since Bone morphogenetic protein (BMP), Leukemia inhibitory factor (LIF) and related cytokines that are essential for mouse ES, failed to support human ES self-renewal *in vitro*. This topic will be described thoroughly in the next chapter.

4.2.2   <u>Opportunities of human PSCs</u>

The fact that hESCs can be propagated indefinitely *in vitro* maintaining their pluripotency status [32], as well as differentiate towards somatic cells and extraembryonic tissues [33] in appropriate culture conditions, makes them developmentally and clinically valuable.

Investigation of mechanisms concerning pluripotency and human development *in vivo* is technically and ethically challenging due to unavailability of post-implantation embryos and is hindered by ethical limitations [34]. For many years developmental biologists have performed experiments on animal models such as frog, fish and mouse embryos to model human development. Especially, genetic ablations of key developmental genes on mouse embryos has given us powerful insights on mechanisms regulating development. However, even though many evolutionary features are conserved between species, the information captured from animal models should be carefully reviewed for their relevance on human development [35]. As an example, *OCT4*-targeted human embryos showed a role of OCT4 in specification of ICM and extraembryonic tissues [36] which does not coincide with its function in the mouse embryo where it is crucial only for ICM specification [37]. On the other hand, human PSCs (hPSCs) have their limitations when it comes to studying complex developmental processes. Although recent advances on development of 3D cultures, organoids, multi-organ generation from 3D hPSCs culture [38] and gastruloids [39], can offer insights into early human developmental processes, they still cannot capture the *in vivo* complexity. Therefore, findings carried out using hESCs should be complemented with mouse studies to validate their *in vivo* relevancy.

4.2.3   Insights into mechanisms of human development through hESCs

hPSCs represent an ideal, and so far, only model to study early human development as well as mechanisms controlling pluripotency and cell fate specification. The pluripotency state in the embryo is a transient state that lasts shortly, however it can be captured *in vitro* under defined culture conditions, making it possible to unlock the mysteries of self-renewal and differentiation [40]. Insights into these processes can be used to generate specialized cell types for clinical applications that could improve current therapies [34]. This idea has spurred efforts for development of efficient protocols to differentiate ESCs towards a plethora of specialized cells, representing all germ layers, mainly by mimicking bona fide embryonic developmental transitions in vitro [41]. Preliminary attempts to differentiate human or mouse embryonic stem cells required the preliminary step of three-dimensional embryoid body formation, with generation of the three germ layers. This differentiation process had limited efficiency in generating specific mature cell types. For more than 20 years, scientists have focused on developing and optimizing differentiation protocols to directly produce a pure homogenous population of specified cells (Figure 4-4). This is achieved mainly by defined perturbations of TGFβ, WNT, FGF and BMP pathways which impact on OSN stoichiometry tipping the balance towards the desired fate.

Nowadays, *in vitro* protocols can be used to derive a plethora of specialized tissues including primary neurons, astrocytes [42], pancreatic progenitors [43], hematopoietic progenitors [44, 45], myocardial cells [46], microglia [47], vocal fold mucosa [48] ,hepatic organoids [49] and cerebral organoids [50]. Most recently, we have witnessed the self-organization of hESCs into embryonic and extraembryonic tissues in vitro. hESCs grown in defined micropatterned colonies they can respond to BMP4 and self-organize to a PS-like structure generating embryonic germ layers [51, 52]. The response of these gastruloids to external BMP4 resembles the epiblast cells of the embryo that respond to BMP4 coming from extraembryonic ectoderm (ExE) by activating WNT and NODAL for PS formation [51, 53-55]. The authors showed that this PS-like structure is functional, because when grafted to chick embryos generated a secondary axis [52].

Several fluorescent reporter cell lines have been generated for key developmental genes to monitor their expression through lineage specification and to derive specialized populations of cells [56-60]. Genome-wide screens, proteomics, transcriptomics on hESCs and lineage derivatives have contributed to

build a roadmap of genes and extrinsic signals driving each step of differentiation as well as the epigenome landscape governing cell fate decisions.

Moreover, single cell transcriptomics in human embryos and lineage tracing have provided a rich characterization of the diversity of cell types and gene regulatory linkages controlling human embryonic development [61-65]. Rapidly evolving genome editing technologies have fostered the field of developmental biology research and clinical application potential of hPSCs. Zinc finger nucleases (ZFN), Transcription activator-like effector nucleases (TALENs), and especially CRISPR-Cas9 systems have progressed gene targeting in hPSCs and further advanced our knowledge of human development as well as human developmental biology.

Figure 4-4. Directed differentiation of hPSCs towards endoderm, mesoderm and ectoderm derivatives, using different combinations of growth factors and modulators of signaling pathways.

4.2.4 hPSCs and clinical applications

It has been more than a decade since the generation of iPSCs, a major breakthrough in science and medicine that would revolutionize the fields of drug testing, disease modelling and regenerative medicine [66, 67]. Forced expression of a cocktail of transcription factors (OCT4, SOX2, KLF4 and c-MYC) in human or mouse adult cells could reprogram them into a pluripotent state similar to ESCs [20, 66, 68]. The fact that pluripotent stem cells, with developmental potential similar to ESCs, can be derived from adult cells without interfering with the human embryo is a major breakthrough in science and received the Nobel Prize for Physiology or

Medicine in 2012. These cells can offer the possibility to generate patient-autologous or HLA-matched cells, providing a valuable platform for personalized medicine [69].

Figure 4-5. Clinical application potential of hPSCs. hPSCs derived from human blastocyst or by reprogramming patient's somatic cells can be cultured *in vitro,* expanded to generate an iPSC biobank, as well as differentiated to specialized cells. They can be used for disease modelling, drug discovery and transplantation for regenerative medicine. Modified from [70].

In such a short period after their discovery, remarkable progress has been made on the development of clinical application of hPSCs [71]. For example, human iPSCs have been used for brain organoid derivation to model life-threatening diseases such as Zika virus infection or Seckel syndrome [72, 73]. Patient-derived iPSCs are used in personalized medicine to model diseases for which mouse models cannot offer pathology recapitulation due to species differences, such as Alzheimer's disease [74]. Clinical trials with iPSCs are being conducted to treat macular degeneration [75] as well as with hESC-derived cardiac progenitors for transplantation to patients suffering from heart failure [76]. Currently in clinical trials are drugs derived from iPSC-based drug screening, including Ezogabine (GlaxoSmithKline) for Amyotrophic lateral sclerosis and RG7800 (Roche) for Spinal muscular atrophy [67]. Despite the advance on clinical applications of iPSCs, there are still some concerns and limitations that need to be taken into account for future applications. Not all iPSC lines perform the same when it comes to functional tissue differentiation potentially due to different genetic background and origin [40, 67]. Furthermore, there is the risk of teratoma formation upon transplantation of differentiated cells derived from iPSCs that still contain undifferentiated populations [77].

4.3 Regulatory networks controlling pluripotency

4.3.1 <u>Intrinsic regulation of pluripotent stem cells</u>

In vivo pluripotency is a transient state, lasting in the mouse embryo from E4.5 to E7.5, at which stage it is lost to allow further development. However, a pluripotent development potential can be maintained indefinitely *in vitro* under specific culture conditions. This allows us to dissect the mechanisms underlying pluripotency and exit from this state. The concept of pluripotency was first conceived in the 1950s and the first ES cell line were established in 1980s [78]. Since then a series of loss-of-function genetic studies in human and mouse stem cells or embryos have been performed to reveal essential genes and mechanisms governing pluripotency. However, little is known how these cells are endowed with such a diverse panel of developmental potential. These studies certainly revealed that the core players in regulation of pluripotency at the molecular level is a set of transcription factors: OCT4, SOX2 and NANOG (OSN) [79-82]. Genetic ablation of each one of those factors imperil the epiblast and leads to embryonic lethality [83-85]. Among them, *Oct4* is indispensable for stem cell identity. Its loss abolishes pluripotency in ICM directing towards a trophoblast fate [83]. ICM bereft of *Nanog* (named after Tir Na Nog: land of the ever young), lose pluripotency and differentiate towards an endoderm-like state [84]. However, *Nanog* null ESCs are still capable of self-renewal and show differentiation potential, but fail to form germ cells [86]. *Sox2* deficient blastocysts also exhibit defective ICM development [85]. Finally, inactivation of all three factors leads to the extraembryonic cell fate [87].

It is believed that pluripotency is a balanced state emerging from competition between rival lineage specifiers for the enforcement of reciprocally exclusive directions [78, 88]. Pluripotency factors unilaterally promote the commitment of ES cells to a desired cell fate and at the same time limit differentiation towards alternate lineages. However, this attempt for disruption of pluripotency and lineage commitment is counteracted by the combined actions of the OSN network that neutralize each other's actions to critical levels where pluripotency is maintained, but at the same time endows them with great developmental potential. Conversely, forced expression of combination of TFs can induce pluripotency in mature cell types and give rise to iPSCs.

Figure 4-6. Distinct roles of OCT4, SOX2 and NANOG in lineage specification. OCT4 is required for mesoderm formation, SOX2 for neurectoderm and NANOG is an Endoderm specifier.

We see that pluripotency is an unstable and agile state of affairs, that bestows stem cells with flexibility to rapidly move towards specific lineage programs. In concert with this hypothesis are findings from overexpression studies of pluripotency factors showing that even a small increase on the levels of these factors dominates distinct lineage fates. Overexpression of OCT4 in mouse ES cells promotes mesodermal differentiation [89], overexpression of NANOG in hESCs increases EOMES levels inducing primitive streak (PS) [90] and overexpression of SOX2 in mESCs favors neuroectodermal differentiation [91]. However, overexpression experiments are tricky to interpret and do not resemble physiological conditions. There is strong evidence that these factors are required for differentiation to different germ layers [90, 92-94]. Knockdown experiments in mouse blastocysts have shown that *OCT4* is required for cardiogenesis and proper mesoderm formation [95], whereas downregulation of *NANOG* in hESCs impaired mesendodermal (ME) differentiation [90] and *SOX2* suppression in hESCs disrupted ectodermal commitment. This suggests that these factors are required for proper lineage specification.

The role of OSN factors in controlling the balance between pluripotency and lineage specification is supported by Chromatin Immunoprecipitation-sequencing (Chip-seq) studies in ES cells, showing their genomic occupancy on active genes of the pluripotency network as well as inactive genes involved in differentiation programs [96, 97]. There are different opinions in the field why OSN factors occupy developmental genes. One explanation is the repression of those genes by OSN factors for maintenance of pluripotency [96-101] and a contrasting suggestion implies that each of the OSN factors promotes the expression of distinct and mutually exclusive lineages; and this competition leads to maintenance of

pluripotency [78, 88]. This genomic interaction has a functional role in regulation of transcription. For example, NANOG binds to *EOMES* and regulates its expression tipping differentiation of ES cells towards DE [90]. Several studies have shown the recruitment of repressor complexes like NuRD by OCT4 and NANOG to control gene transcription and regulate pluripotency [102, 103].

4.3.2 Extrinsic regulation of pluripotent stem cells

Continued extrinsic signaling controls expression of key pluripotency factors to levels necessary for maintenance of the pluripotent state. Without these signaling cues ES cells will eventually differentiate. Even though mESCs rely on LIF and BMP4 signaling for maintenance of the undifferentiated state, members of the fibroblast growth factor (FGF) family and transforming growth factor β (TGFβ) signaling support self-renewal in hESCs [104-108]. Despite this, both mESCs and hESCs differentiate towards the three germ layers in response to Nodal/Activin, BMP and WNT signaling [109, 110].

TGFβ, Activins and Nodal ligands act by binding to their corresponding transmembrane receptor heterodimer leading to activation of the receptor complex at the plasma membrane. This results in phosphorylation of receptor-activated SMAD (R-SMAD) proteins SMAD2/3 (TGFβ pathway) which in turn bind to SMAD4 and activate gene expression programs. In stem cell culture, recombinant TGFβ and bFGF2 are used for maintenance of self-renewal. Both pathways are essential for self-renewal of hESCs, as demonstrated by loss of pluripotency upon blockade of each or both of them. More specifically, inhibition of TGFβ/ACTIVIN signaling with SB431542 small inhibitor leads to neural differentiation [111, 112], resembling the deletion of Nodal at the mouse epiblast stage [113-115]. Similarly, blockade of FGF signaling with SU5402 inhibitor results in hESCs differentiation [116, 117]. However, it remains unclear how these signals regulate downstream gene expression for maintenance of pluripotency. It is proposed that exogenous FGF signaling acts autocrine upstream of TGFβ ligands to sustain expression of pluripotency factors that activate endogenous FGF2 expression [106]. This initiates a cascade of phosphorylation events of downstream effectors including pluripotency factors OCT4, SOX2, SALL4 [118]. Addition of ACTIVIN A to cultured hESCs can promote FGF2 production showing the synergizing effect between those signals [116, 119].

Several studies have attempted to identify the intersection of extrinsic signaling with intrinsic transcription programs to maintain pluripotency. A direct link between TGFβ/ACTIVIN signaling and pluripotency factors is demonstrated by the

binding of SMAD2/3 on *NANOG* and *OCT4* promoters on repetitive SMAD binding elements (SBE) to regulate their activity [117]. Inhibition of ACTIVIN/NODAL signaling downregulates expression of *OCT4* and *NANOG,* with most dramatic effect on *NANOG* expression, indicating that is a direct target of SMAD signaling [116]. ChIP-Seq data have revealed that SMAD2/3 genomic occupancy overlaps by 1/3 with those of pluripotency factors OCT4 and NANOG indicating a joint role in pluripotency networks [120, 121].

The role of WNT in hESC pluripotency is unclear. hESCs express low levels of WNT components possibly due to autocrine production. Inhibition of this minimal activity in hESCs suppresses background expression of developmental genes *BRACHYURY, GSC* and *EOMES* [122], showing a suppressive role of WNT in maintenance of pluripotency. In contrast, activation of WNT using a GSK3β inhibitor in combination with LIF or MEK inhibitors has positive effects on self-renewal of mESCs [123, 124]. This supportive action of WNT on inhibition of differentiation seems to happen through stabilization of β-CATENIN, the core component of WNT signaling which leads to de-repression of TCF3-target genes [125]. However is not very clear if WNT signaling is exclusively responsible for undifferentiation state or attributed partially to JAK-STAT signaling activation, and even to off target effects of the inhibitor since mice lacking *β-Catenin* maintain undisrupted pluripotency [126, 127].

Recent studies have revealed a role of HIPPO signaling pathway in ESC self-renewal and differentiation. In hESCs, TAZ not only facilitates the nuclear accumulation of SMADs promoting their transcriptional activity but is required for self-renewal, since its loss leads to TGFβ signaling blockade and neuroectoderm differentiation [128]. In mESCs association of YAP with SMAD1 is necessary for suppression of BMP mediated induction of neural fate [129], highlighting the important link of HIPPO and TGFβ signaling to control self-renewal. Proteomics coupled to genome wide ChIP-seq experiments have demonstrated that the transcriptional effector of HIPPO pathway, TEAD4 form a complex with OCT4 and SMADs to jointly control pluripotency elements in hESCs [102].

Downstream effectors of the main signaling pathways described above, such as TFCP2L1, KLF4, ZFP281, have been identified as OSN-interactome co-occupying common genomic regions [99, 100].

4.4 Signaling controlling early fate decisions of hPSCs

4.4.1 Opposing roles of TGFβ on pluripotency and primitive streak

The primitive streak (PS) is a fundamental structure in embryo development. It marks the onset of gastrulation and thereby induction of the three germ layers, as well as defines the anterior-posterior axis of the embryo. Gene-targeting studies have revealed the essentiality of TGFβ signaling members including Nodal and BMP4, as well as WNT family members for initiation and specification of PS [130-132]. During gastrulation, TGFβ ligands act as morphogens and spatiotemporal changes in Nodal signaling, drive specification of different cell fates. This asymmetry in Nodal signaling levels unbalances the pluripotency equilibrium of OSN factors via SMAD2 in the epiblast. Anterior epiblast cells experience low levels of Nodal-SMAD2, allowing SOX2 to dictate ectoderm differentiation. On the other hand, high Nodal-SMAD2 levels at the posterior epiblast promote upregulation of *OCT4* and *NANOG* allowing NANOG-SMAD2/3 to induce *EOMES* expression [90]. This results in the formation of the PS in the posterior epiblast at E6.5 via epithelial-to-mesenchymal transition of epiblast cells.

Figure 4-7. Primitive streak formation is a hallmark in embryonic development. Primitive streak formation in human embryo happens at around 16 days after fertilization. Due to ethical limitations, mechanisms regulating *in vivo* PS formation cannot be studied. Human embryonic stem cells isolated from ICM of embryos can be differentiated towards PS *in vitro,* offering unique insights into the molecular mechanisms governing this process. TGFβ, FGF and HIPPO signaling play crucial role in maintaining the pluripotent state of hESCs and the balanced expression of OCT4, SOX2 and NANOG. Spatiotemporal changes in TGFβ, as well as activation of WNT and BMP pathways, induce upregulation of *GSC, EOMES* and *FOXH1* leading to PS formation. Modified from [133]**.**

This cell movement is mainly regulated by FGF signaling as EMT is compromised in *Fgfr1* mutant embryos. This is mainly due to inability to downregulate *E-cadherin*, necessary for disruption of tight-junctions and failure to upregulate *Snail*, an *E-cadherin* repressor [134-136]. Expression of several markers coincides with PS formation, such as *Wnt3, Brachyury (T), Gsc* and *Nodal*. Cells migrating at the anterior edge of the PS are the Anterior primitive streak (APS) progenitors that will form different structures such as notochord, paraxial mesoderm and definitive endoderm (DE) depending on the intensity and duration of Nodal-SMAD2 and WNT signaling levels [54]. Increased Nodal/Activin signaling allows FoxH1 in association with SMAD2/4 to stimulate expression of *Goosecoid (Gsc)* at the Anterior PS [137-140]. *Gsc* expression marks the appearance of the node, a group of cells that can induce cell fates in surrounding cells. It has an important role in specification of dorsal regions along the A/P axis of the mouse embryo, demonstrated by formation of a secondary axis in Xenopus embryo after its ectopic expression [141]. Human embryos contain a structure similar to the mouse node [142], however, ethical limitations have restricted researchers from investigating the existence of an organizer. Preliminary evidence for the presence of a structure with organizer properties has been recently proved in humans [143]. Using a stem-cell based approach the authors differentiated hESCs to a node-like structure resembling primitive streak and expressing organizer-gene *GOOSECOID* [143]. After engraftment to chick embryos this structure gave rise to a notochord-like tissue forming a secondary body axis [143]. In hESCs the expression of *GSC* at the APS suppresses *BRACHYURY,* limiting its expression on the posterior PS where BMP signaling will induce formation of the mesodermal tissues [59, 130, 144]. APS cells will specify DE by partnering of SMAD2 with EOMES to activate DE genes *SOX17* and *FOXA2* and complete DE formation by E7-7.5 [13, 90, 145]. During PS formation expression of specific TFs and TGFβ signaling effectors is crucial for correct formation of this structure and its derivatives. Loss of either NODAL, FOXH1, SMAD2, SMAD4 or EOMES factors disturb DE specification in hESCs [146-151]. Mutant mouse embryos *Bmp, Nodal, Smad2* or *Smad3* fail to form mesoderm whereas loss of nodal antagonists *Cer1* and *Lefty1* results in multiple PS formation [9, 130, 152, 153].

4.5 Chromatin Controls the Embryonic Stem Cell State

4.5.1 TGFβ signals to chromatin to regulate transcription

ACTIVIN/Nodal signaling has a dual role in hPSCs: maintenance of pluripotency and early cell fate specification. For performing this role, SMAD2/3 switches partners to regulate transcription depending on cellular and context-depend TGFβ response [154]. Binding of SMAD complexes on enhancers and promoters of target genes happens at SMAD binding elements (SBE) and requires interaction with specific DNA binders on target genes.

SMAD2/3 co-occupy the same targets with OCT4 and NANOG (e.g. *OCT4, NANOG, GDF3, LIN28A*) in hESCs to control self-renewal [120] as well as PS associated genes for suppression of their expression during pluripotency (e.g. *FOXA2, EOMES, T*) [102]. One very well-known PS partner of SMAD2/3 is FoxH1 (forkhead box protein H1) [155]. *FoxH1* is expressed in hESCs and mouse epiblast and upon TGFβ signaling partners with SMAD2/3 to specify PS formation [102, 139, 155-157]. ChIP-seq studies have revealed a Smad2/3 and FoxH1 genome wide interaction on Nodal targets *nodal, cer1, hhex* and *gsc* in Xenopus embryos [155]. The partnering of TFs with SMADs depends on their abundance and dominance over other TFs [158] leading to TGFβ-dependent transcriptional regulation and cell type tailored response [120]. These studies reveal the collaboration of SMAD proteins with master transcription factors and DNA binding cofactors for cell type and context specific TGFβ responses and can explain how this signals switches from pluripotency promoter to cell fate specifier.

The regulation of transcription by SMADs relies on their ability to recruit co-activators or co-repressors with chromatin and histone modifying activities. Even though SMAD2 can directly interact with mediator components from the transcription machinery [159], the central mechanism of SMAD-dependent transcription regulation is through modification of chromatin. This is accomplished mainly by the interaction of SMADs with histone-modifying enzymes, such as histone acetyltransferases (HATs), histone deacetylases (HDACs) and chromatin-remodelers, such as SWI/SNF, NuRD, Ino80. Other unknown SMAD-cofactors are still to be discovered. Recruitment of HAT complexes by SMAD proteins leads to transcriptional activation through acetylation of histones and their dissociation from DNA, allowing the recruitment of the transcription machinery, specific TFs and "readers" on the accessible locus [160]. An example of SMAD-co-activators with HAT activity are p300 and CBP [161, 162]. In contrast, recruitment of HDACs by

SMAD proteins leads to transcription repression by removal of histone acetylation groups and creating a restrictive chromatin environment.

4.5.2   Enhancer bookmarking during early development

Embryonic stem cells and uncommitted progenitors are equipped with developmental potential. Upon defined signaling perturbations, specific lineage specifiers are becoming expressed and direct the cell to commit to a specific cell fate. However, those developmental genes are not expressed in the stem cell state but rapidly become expressed when the signaling is appropriate. How the expression of those genes is regulated, from being undetectable in pluripotency to active upon appropriate signaling cues? The answer lies at the chromatin level, on the regulatory regions of those genes. Experiments mapping the genome-wide histone modifications, uncovered which histone modifications are mostly found in actively transcribed genes and which in repressed-"poised" genes. In ESCs active promoters and especially unmethylated CpGs - containing promoters, are characterized by Histone 3 lysine 27 acetylation (H3K27ac) and Histone 3 lysine 4 trimethylation (H3K4me3) marks and a lower nucleosome occupancy [163-165]. In the same fashion, active enhancers show active features of promoters except that H3K4 is monomethylated [164]. On the other hand, enhancers of "poised" inactive developmental genes instead of acetylation, are bivalently flanked with repressive H3 lysine 27 trimethylation (H3K27me3) catalyzed by Polycomb repressor proteins, and active H3 lysine 4 mono-methylation (H3K4me1) histone marks [166-168]. The latter serves to prime developmental genes for rapid gene activation, depending on the lineage program. Other histone modifications correlated with active transcription are H3K9ac and H3K79me3, whereas H3K9me3 is linked to poised state [168-171]. Upon developmental signaling, poised genes associated to the induced fate, lose their repressive H3K27me3 mark, acquire H3K27ac mark and become activated [166]. Poised genes that will eventually not be expressed at the given developmental state, lose the H3K4me3 mark but maintain the repressive H3K27me3 [165, 166].

Advances in genomics and high-throughput sequencing; such as Chromosome conformation capture assay [172], ChIP-seq profiling of histone modifications [158, 168, 169], TFs [99, 173] and transcriptional coactivators [168, 169, 174-176] as well as FAIRE (Formaldehyde-assisted isolation of regulatory elements) and DNAse-seq [177, 178] has facilitated the discovery of regulatory regions in the genome. Around 9,000 enhancers have been identified in mESCs that commonly range between 200-500 bp in length [164]. Combination of all this

genomic and chromatin information together will uncover active enhancers. They can be occupied by several TFs forming enhanceosomes providing the template for combinatorial and co-operative interaction of those TFs on gene regulation [179] [180]. Pluripotency factors also play a role in this regulation by occupying these genes, before activation signals are received.

Figure 4-8. Schematic representation of chromatin landscape and protein complexes at active and poised enhancers. Enhancers of active genes in hESCs, such as *OCT4* and *NANOG*, are usually bookmarked with H3K27ac and H3K4me1 marks allowing gene transcription. Poised enhancers of inactive genes at the pluripotent state, such as *GSC* and *EOMES* are kept silent by PRC2-deposited H3K27me3 mark. At the same time H3K4me1 mark primes them for rapid activation upon appropriate signaling cues.

4.5.3 Modification of the chromatin landscape around enhancers regulates transcription programs

For a developmental program to be established, important role have the chromatin landscape and the binding of specific TFs and signaling co-effectors in regulatory regions of developmental genes. One very well studied example of poised enhancer regulation during differentiation is the enhancer of *EOMES*. In hESCs the poised enhancer of *EOMES* is bivalently marked with H3K27me3, H3K4me1 and lacks H3K27ac, keeping it silent [90]. The master pluripotency factors OCT4, SOX2 and NANOG co-occupy the enhancer together with SMAD2/3 and p300 [57, 90]. Upon high ACTIVIN/NODAL signaling for 24h, *SOX2* is downregulated allowing FOXH1, NANOG and SMAD2/3 to drive the activation of *EOMES* [90]. H3K27me3 rapidly gets depleted and H3K27ac mark is deposited by p300 [90]. Subsequently, EOMES upregulates DE genes *SOX17* and *FOXA2* [90].

An additional step for rapid gene activation and transcription upon differentiation signals is the pre-looping of enhancer with the target promoter [181, 182]. Promoters are also bivalently marked and occupied by RNA polymerase II which prepares them for rapid gene transcription [182]. As the differentiation program proceeds, pluripotency factors become silenced, mainly either by caspase-mediated degradation [183], miRNA-mediated blockade of mRNA translation [184] or chromatin repression [185]. For lineage programs to proceed successfully the silencing of pluripotency genes mainly by deposition of H3K9me3 mark on their promoter [185] or enhancer decommissioning is essential [186, 187]. For example, during differentiation of hESCs, LSD1, NuRD, or JARID1 - mediated demethylation of H3K4 or H3K9 results in silencing of pluripotency genes [186, 187]. Surprisingly loss of individual members of repressive complexes in ESCs, such as JARID1B, MBD3, EED, SUZ12 (PRC2) or EZH2 (PRC2), minimally affects self-renewal but instead has dramatic effect in differentiation mainly due to failure to silence pluripotency genes, and also due to redundancy mechanisms [188-193]. Demethylation is also employed for poised gene activation. For example, the SMAD2/3-dependent recruitment of JMJD3 demethylase removes the repressive H3K27me3 mark flanking the *Nodal* and *Eomes* enhancers within 24h upon ACTIVIN/Nodal signaling [194, 195].

Poised enhancer in hESC	
Repressive complex	Activation complex
Histone-modifying enzymes Polycomb repressive complex: PRC1 family: CBX2/4/6/7/8, RING1A/B, PCGF family PRC2: EZH2, EZH1, EED, SUZ12, RBB4/7 PRC2 interacting proteins JARID2, AEBP2, and PCL1-3	
SetDB1 HP1γ	TRIM33
Histone deacetylase complexes HDAC1 and HDAC2 SIN3: SIN3A/B, SDS3, SAP18, SAP30, MeCP2, RBP1, BRMS1, ING1/2, SAP25, SAP130, SAP180, RBP2/KDM5A NuRD: CHD3/4, GATAD2A/B, MBD2/3, MTA1/2/3 CoREST: ZNF217, BHC80, LSD1	H3K27me3 demethylases JMJD3 (Kdm6b) Utx (Kdm6a)
ATP-dependent Nucleosome remodelers SWI/SNF: BRG1, ARID1, BAF family, BRD CHD family ISWI family: NoRC, WICH, NURF, CERF complexes, INO80 family: p400-TIP60 complex	ATP-dependent chromatin remodelers SWI/SNF: BRG1, ARID1, BAF family, BRD CHD family ISWI family: NoRC, WICH, NURF, CERF complexes, INO80 family: p400-TIP60 complex
p300	p300 mediator

Table 4-1. Regulation of chromatin landscape around active and poised enhancers in hESC by chromatin and histone modifying complexes. Chromatin remodeling enzymes have emerging roles in pluripotency and differentiation. They are usually in complex with histone modifying enzymes and the different assembly of complexes defines the biological specificity in a given cell type. Polycomb complex mediates repression of target genes by histone ubiquitylation, catalysis of H3K27 methylation, chromatin compaction and condensation of nucleosomal arrays. It can also indirectly regulate RNA polymerase II activity through histone ubiquitylation. SIN3, NuRD and CoREST complexes mediate histone deacetylation activity. SetDB1 can be recruited by OCT4 through protein-protein interactions to catalyze H3K9me3 at poised genes for repression. CHD1 maintain ESC euchromatin state and TIP60-p400 fine-tunes appropriate gene expression. BAF and NuRD silence pluripotency genes in differentiating cells [195-209].

4.5.3.1 STAGA complex and TADA2B in development

A nucleosome consists of 146bp of DNA, wrapped around an octamer of histones: H3, H4, H2A, H2B. Nucleosomes around enhancers or promoters can be modified either by ATP-dependent chromatin remodeling or histone modifying enzymes. One of the first histone modifications to be described was acetylation of N-terminal histone tails. Histone acetyltransferases (HATs) catalyze this process at

the amino groups of lysine residues by using Acetyl-CoA [210]. This modification loosens the interaction of DNA with histones and positively promotes transcription by increasing the accessibility of TFs to DNA [211].

The first HAT to be identified and most well studied is GCN5 or KAT2A, a highly conserved protein from yeast to humans [212, 213]. It is essential for embryonic development since *Gcn5* null embryos die soon after implantation and exhibit defects in mesoderm formation [214, 215]. Human GCN5 has a close paralog with 70% similarity, PCAF or KAT2B and they can mutually exclusive be part of SAGA or ATAC complexes with distinct functions and genomic occupancy [216-218]. The human SAGA complex preferably acetylates H3/H4 on lysine 9 and 14 and has a role in assembly of transcription complexes by acting as a coactivator and recruiter of TATA-binding protein TBP [219]. It consists of more than 20 subunits that modulate the GCN5 catalytic HAT activity, including TADA2B, TADA3, TAF5L, TAF6L and SGF29 [220-222].

Figure 4-9. Composition of Human ATAC and STAGA complexes. Human STAGA and ATAC complexes evolved from an ancestral acetyltransferase complex and share some common subunits, such as KAT2A and KAT2B proteins with HAT activity. Adapted from [223].

The yeast and Drosophila Ada2 protein has two metazoan paralogues: ADA2a, which is part of the ATAC complex together with GCN5, ADA3, SGF29, and ADA2b belonging to SAGA complex [224, 225]. They both have SANT, SWIRM and ZnF domains important for contact with ADA3, GCN5 and for DNA binding activity [225-227].

Figure 4-10. Human TADA2B protein. The coloured boxes illustrate TADA2B protein domains. Protein domains have been identified using https://www.ebi.ac.uk/interpro/beta/, and http://www.genome3d.eu/.

The human protein is called Transcriptional adapter 2-beta (TADA2B) and is an approximately 50kDA protein translated from the *TADA2B* gene. It functions as a transcription adaptor that potentiates the catalytic activity of GCN5 within the STAGA complex [228]. ChIP - seq for Flag-HA tagged Ada2b in Drosophila embryos showed tissue-specific localization of Ada2b at H3K9 acetylated histones with preference on Pol II - bound regions at promoter-proximal pause sites [229]. However, the role of TADA2B in mammalian development is not described yet. Recently two studies have indicated a possible role in human pluripotency that remains to be addressed [230, 231]. Another study in hESCs showed that the mRNAs of *TADA2B, TAF5L* and *TAF6L* are targets of TGFβ - regulated m6A (N6-methyladenosine) methylation, which is essential for their rapid downregulation during exit from pluripotency [232]. The m6A modification on *TADA2B* mRNA was identified to have also a role in human erythropoiesis since its loss impaired erythroid specification of human hematopoietic stem and progenitor cells (hSPCs) [233]. *TADA2B* has also been implicated to have a role in cancer. More specifically, CRISPR ko and CRISPRi screen on melanoma cells showed that mutations of STAGA complex members *TAF6L, TAF5L, TADA1* or *TADA2B* and *MYC* lead to resistance to BRAF inhibitor, vemurafenib [234, 235]. These few, but solid studies have revealed a possibly essential role of *TADA2B* in human stem cell pluripotency and cancer which makes it interesting to investigate further in the future.

4.5.4 Transcription factors occupy specific enhancers to control cell fate choices

When a pluripotent stem cell receives the signal for differentiation, a series of events occur on chromatin level setting the stage for transcriptional response. Temporal changes of the epigenome affect chromatin accessibility and regulate dynamic cell response to extrinsic stimuli in collaboration with pioneer transcription factors during mammalian development [166, 236]. TF binding is guided and regulated by DNA sequence specificity, chromatin landscape, extrinsic signaling and cofactors availability [237, 238]. The enhancers are the signal integration hubs where specific combination of pioneer TFs and signaling effectors bind and activate the first line of response when the chromatin environment permits [239]. Therefore enhancers provide the template for spatiotemporal gene regulation during embryonic development, even if they are removed and positioned away from a given transcription start site [240]. OSN factors occupy regulatory regions in hESCs genome, 90% of them being at enhancer elements [96, 99, 158]. They tend to co-occupy enhancers of pluripotency genes including their own, maintaining a feed forward loop for maintenance of stem-self renewal. At the same time they occupy 50% of poised enhancers possibly for maintenance of their poised state in pluripotency and rapid activation upon differentiation signal as described above for *EOMES* enhancer [241].

10% of protein-coding genes are TFs [242, 243]. They activate gene expression by recruiting chromatin modifiers for increase of DNA accessibility [244], recruitment of transcription machinery or recruitment of cyclin-dependent kinases which phosphorylate and release of RNA polymerase II from pause sites [245]. OSN are involved on these mechanisms by interacting with RNA polymerase II coactivators [246] and selecting which genes will be activated to support the ESC state. A priming role of FOXA2 pioneer factor in enhancer activation during hESCs differentiation was recently described [247]. FOXA2 is necessary for deposition of H3K4me1 mark at developmental enhancers before their activation and is required for chromatin remodeling during differentiation [247].

Pioneer factors can therefore influence sequential cell type specific enhancer activation during development or reprogramming into iPSCs [247, 248]. OCT4, SOX2, and KLF4 are considered pioneer factors that can bind and activate target genes surrounded by nucleosomes during somatic reprogramming [66]. The mechanism however how this achieved is not well understood but it might involve recruitment of chromatin remodeling complexes such as BAF and BRG1 by pioneer factors to promote chromatin accessibility [249].

5 Aims of this thesis

During embryonic development the various embryonic cells commit to distinct directions, yet they are all equipped with identical genome originating from the zygote. Two milestones in embryo development are the establishment of pluripotency and formation of the primitive streak (PS) where the three germ layers will originate from. In hESCs the maintenance of pluripotency and initiation of PS is orchestrated mainly by TGFβ signaling. It strongly relies on the ability of SMADs to collaborate with different partners in order to regulate differential transcriptional responses leading either to maintenance of self-renewal or for PS induction. They mainly act by forming a regulatory network with specific set of transcription factors, cofactors, signaling effectors and chromatin modifiers on regulatory elements of target genes to drive cell- and stage- specific gene expression depending on the signaling.

However, the molecular mechanisms for signal integration and transcriptional regulation is incompletely understood. Little is known about the hierarchy of events leading to transcriptional regulation of different lineage programs. Are pioneer factors present at poised enhancers before the signal for differentiation is activated or chromatin remodelers facilitate recruitment of cell-specific TFs on open chromatin to activate transcriptional response?

To advance our understanding of the regulation of gene expression during the differentiation of hESCs, the thesis has the following aims:

- To establish a proteomics method that can be used to identify proteins binding to specific DNA regulatory elements
- To identify proteins occupying the enhancer regions of *NANOG* and *GSC* in hESCs and APS-differentiated cells using the proteomics method.
- To validate the endogenous binding of selected candidates and characterize their role in pluripotency and APS specification

In the first two aims we will establish a proteomics approach to identify potential candidate proteins. Subsequently, in the third aim we will functionally characterize factors identified using state of the art technology. Taken together, these experiments will uncover the underlying molecular circuitry responding to

extracellular signaling cues to elicit transcriptional changes during pluripotency and early cell fate specification. These results will help to understand the earliest events during human development a process so far difficult to study due to ethical concerns. In the future, these insights will help to develop new strategies to improve culturing conditions of hESCs and to define protocols for efficient generation of functional tissues for therapeutic purposes. In addition, the methodologies implemented here are expected to be of interest to other fields of life sciences (including cancer biology) to identify novel transcriptional complexes *in vivo* and potentially will result in collaborations.

6 Materials and Methods

6.1 DNA Pull Down assay for Western blot or Mass spectrometry

6.1.1 Buffers and Solutions

4X Biotin-Streptavidin Buffer		
Final concentration	Stock concentration	Amount
40 mM Tris pH 7.5	1 M	2 mL
4 mM EDTA	500 mM	400 µL
4 M NaCl	5 M	40 mL
0.012% Igepal	20%	30 µL
H$_2$O		7.6 mL
Final volume		**50 mL**

2X Binding Buffer		
Final concentration	Stock concentration	Amount
40 mM HEPES pH 7.5	1 M	2 mL
5 mM KCl	1 M	250 µL
2 mM DTT	1 M	100 mL
40% Glycerol	99%	20 mL
0.02% Igepal	20%	50 µL
H$_2$O		27.6 mL
Final volume		**50 mL**

Reaction Solution		
Components	Concentration	Amount
2X Binding Buffer		700 µL
Poly (deoxyinosinic-deoxycytidylic) acid sodium salt (dIdC)	0.5 µg/mL	20 µL
Nuclear cell lysate	100-300 µg	300 µL
H$_2$O		180 µL
Protein lysis buffer		200 µL
Final volume		**1400 µL**

DTT/ABC solution		
Final concentration	Stock concentration	Amount
5 mM DTT	1M DTT	5 µl
50 mM Ammonium bicarbonate		3.95 mg in 1 mL
		30 µl per sample

Iodoacetamide/ABC solution		
Final concentration	Stock preparation	Amount
1 M Iodoacetamide	1 mg in 5.4 µl H_2O	2.5 µl
50 mM Ammonium bicarbonate	3.95 mg in 1 mL	7.5 µl
		10 µl per sample

6.1.2 Experimental procedure

In a first step the promoter and enhancer regions of genes of interest were selected based on their epigenetic signature and TF binding profile using Integrative Genomics Viewer (IGV). The 300-400bp sequences were amplified by PCR from genomic DNA of hESCs and cloned into the pBluescript vector in between the T7 and T3 promoter sequences. PCR with the T7 primer holding a biotin modification at the 5' prime end together with the T3 primer resulted in biotinylation of the PCR fragments. Multiple PCR reactions were pooled together and the PCR product was isopropanol-ethanol precipitated and further purified using a Macherey-Nagel PCR clean-up kit (Cat. Number 740609.250). The DNA concentration was determined and 2 µg of biotinylated sequence were incubated in 1mL Biotin-Streptavidin buffer for 1 hour rotating at room temperature with 60µl of agarose beads (Pierce, 20353), prewashed twice with Biotin-Streptavidin Buffer. 50µM Biotin were added in the mixture and incubated for 15 minutes at room temperature on a rotor to block the excess of unbound beads. The supernatant was removed and the beads were washed twice with 1ml Biotin-Streptavidin Buffer. The reaction solution was added to the beads and was left incubating for 5 hours or overnight at 4°C on a rotor.

For Western blot analysis the beads were washed five times with 1mL Binding buffer each and boiled at 95°C for 5 minutes in 40 µl 2X Laemmli buffer to elute the bound proteins. For Mass spectrometry analysis the beads were washed twice with 1ml Binding Buffer and twice with 1ml 50mM Ammonium bicarbonate (ABC). After adding second volume of ABC a 100 µl aliquot was taken for Western Blot Analysis. Any remaining liquid was carefully removed using a gel loading tip. 30µl of DTT/ABC solution was added to the beads and incubated at 50°C for 30 minutes. The mixture was cooled on ice and then 10 µl Iodoacetamide/ABC solution was added per sample and incubated for 45 minutes at room temperature in the dark. The samples were then incubated with 2 µl of Trypsin (Promega, V5111) at 37°C overnight. The day after, 2 µl of trypsin were added and the samples were incubated for 1 hour at 37°C. After centrifuging briefly, the supernatant was

transferred to a new tube with addition of 3.5 μl 100% Formic acid (FA). The remaining beads were washed with 50 μl 5% FA and pooled with previously collected supernatant. The samples were then centrifuged again to remove any remaining beads and approximately 100 μl were transferred to two new tubes. Digested samples were cleaned-up using C18 MicroColumns (NestGroup) according to manufacturer instructions.

Proteomics data was acquired in data dependent acquisition mode on a LC-MS/MS system consisting of a Proxeon Ultra easy LC and an Orbitrap Elite (Thermo). Peptides were separated on a PepMap100 column (C18, 0.075 x 150 mm, 2 μm, 100A) with solvent A: 5% ACN, 0.1% formic acid (FA) in water and solvent B: 98% ACN, 0.1% FA. Gradient settings were: 0 – 60 min: 5%B-30%B at 300 nl/min. The Orbitrap Elite was run in data-dependent mode with parallel MS1/MS2 acquisition. Survey full scan MS1 spectra (from m/z 350 to 1600) were acquired in the Orbitrap with resolution R= 120000 at m/z 400. Up to 15 ions with charge state ≥ +2 were selected for fragmentation per cycle with a dynamic exclusion window of 30s. Peptides were identified using the Trans-Proteomic-Pipeline (TPP) v4.7 with search engines Comet, OMSSA, Myrimatch and XTandem with the parameters: precursor tolerance: 10 ppm; fragment tolerance: 0.5 Da; static modifications: iodoacetamide (C); dynamic modifications: Label:13C(6)15N(2) (K); Label:13C(6)15N(4) (R), phospho (STY), oxidation (M); enzyme: Trypsin; missed cleavages: 2. Posterior probabilities were assigned with PeptideProphet and iProphet and false discovery rate (FDR) cut-offs were determined using Mayu. Datasets were filtered to obtain a protein false discovery rate of 1%. Label-free quantification of identified proteins was performed in Progenesis QI for proteomics v3 based on peptide MS1 intensity. Proteins were quantified based on non-conflicting peptide features and total TIC normalization was enabled. Result tables were exported to Excel for further processing.

Figure 6-1. Schematic representation of the DNA pull-down method on *NANOG* enhancer.

<u>Enhancer Sequences cloned for DNA pull down</u>

Gene	Genome	Chromosome	Start	End
GSC	Human hg19	chr14	95,241,731	95,242,010
POU5F1	Human hg19	chr6	31,139,733	31,139,989
NANOG	Human hg19	chr12	7,940,230	7,940,617
EOMES	Human hg19	chr3	27,769,806	27,770,272
Control	Human hg19	chr12	7,938,699	7,939,056

GSC enhancer

```
AAtTCAACGTGTGGCCCGGTCTCGCACTTGGGAGGGCAACCTTATGGGACGCTTTGAATCCCCATTGGT
GACCCGCAGTCCAATCCACATTCCTTGGCTTATTTACGGGAGACGGATATATTAGCAACAAATGGGTTT
ATGCAGGTATGTATGTGTTTCAAAAACCCTGGCAGCCCAGGGAGGGCTCTGAACTGTGCCCTGCACGGA
CTTTGCAGTTCCCACGGAGAAAAACCCCCGTGCATGGCTCATATGAGCCATTTGTGTCCAGCCATTCTG
TGCCCTGGA
```

POU5F1 enhancer

```
TGAAGCTGGACGCCTGAGTcccccagaggaaggaggaactagatacctaggtccctgtaggggggccctt
ggtgcccgtctgaggctcagtcttttgaggggattgcagaggggggggttgctggagctccttttagcgtct
ctgaaggggattctgtgtgaggggattgggactggggggttggggagcaggaagcagtccccaggggag
ccatccaggcccattcaagggttgagcacttgtttagggttagagctgcccccctctggggaccgggatt
gtccagccaaggccattgtcctgccccccttcccccagtccctCCCAGGCTTCTTTGAACCA
```

NANOG enhancer

```
AAGCATAGGGGCATCAGGGAGCCGGTGAGTGGGGCTAGTACATCATGCTTCAGGCAGATAGTCTATGTC
TCGTAGGAAATAGGATGATTTCTTTAAAAGAAGGGAAAAACTTGCCCTCCTGTCTCTCTTAATCAGCAC
AGTTTGTTAAGTGGGGGCTGTCAAGGCCACCAAGACTCACCGAGGCTGAGCTTGCCCGCAGCTCTCCAA
AGGGCAGGTACCAGAAGCTTTGTTCTTTGCAGAAGGGGGTCTTTGCATAAAAGCCTGAGCTGAATTCCC
CCATCCCCCGCTCCCTGTCCCATTGTGTCTAGGGTAAGAGCCTCCGGAGTGAAAGACCAAAGGGAAGGG
GCTGGTGGCTGGAAGGCCGACTTACTACATTCTTCGCCAAGGA
```

EOMES enhancer

```
CTTAACTCCTTCGTGGTCAGGCGTTTAACCAAGTAGGTGGAGAACAAAAGCTGAGCTAAATACCCTATT
TTCTTTCTTTGACAAGTAATCGGCCGGAACGGCAAAATCCCAGCGGGTGGCGGGAGATACTGGAGGCCT
TGGGAAAAACAGAATGCTAATGGCCGCCTGGAGGGGGCGGAAGGCACAGCCTGGTTAGGGCGCCCCAGC
CGCCTCGCTTTGAAGTGCACCTCTCCAGCCGAGCTCACCGGGAGGGCACTTGATTTCCCCTTTCGGGGA
AATAAAGAAGAGCTGCTAGAGTAACTGGGGACTCTCTTTGGTTCTGACTGACTCCCTCCTGCTTCCTAC
CCCTGGTTCACGTTGCTTTTGACCTGTTTGCTCCTGGCCTCCCCCATCGGACTGGGAATCCCCGGGAGG
AAAAAGACTAAGGTTTATTCATCCTCATTTAGCTCCAGTATCTTAGTAAGAAC
```

6.2 Protein extraction from hESCs

6.2.1 Total protein extraction

6.2.1.1 Buffers and Solutions

10X TNT		
Final concentration	Stock concentration	Amount
500 mM Tris/HCl pH 7.6	1 M	8.3 mL
1.5 mM NaCl	5 M	15 mL
5% v/v Triton X-100	20%	12.5 mL
10 mM EDTA	50 mM	1 mL
H_2O		13.2 mL
Final volume		**50 mL**

1X TNT lysis buffer		
Final concentration	Stock concentration	Amount
1X TNT	10X	100 µl
25 mM NaF	500 mM	50 µl
5 mM Na_4PPi	100 mM	50 µl
1 mM Na_3VO_4	500 mM	2 µl
0.1 mg/ml Trypsin inhibitor (TI)	10 mg/ml	10 µl
varied	Proteinase inhibitors mix (PI): Antipain 1mg/ml; Aprotinin 5 mg/ml; Benzamidine hydrochloride 10 mg/ml; Leupeptin 1mg/ml	10 µl
10 µg/ml Pepstatin A	1 mg/ml	10 µl
1 mM AEBSF	400 mM	2.5 µl
H_2O		765.5 µl
Final volume		**1 mL**

6.2.1.2 Experimental procedure

Cells were washed twice in PBS and harvested in appropriate volume of TNT lysis buffer. After 30 minutes incubation on a rotor at 4°C the samples were centrifuged at maximum speed for 30 minutes. The supernatant was collected in new tubes, frozen down in liquid nitrogen and stored at -80°C for future use.

6.2.2   Nuclear and cytosolic protein fractionation

6.2.2.1   Buffers and Solutions

Nuclear Franctionation buffer A		
Final concentration	Stock concentration	Amount
10 mM HEPES	1 M	10 µl
10 mM KCl	1 M	10 µl
0.1 mM EDTA	500 mM	0.2 µl
0.1 mM EGTA	50 mM	2 µl
25 mM NaF	500 mM	5 µl
5 mM Na₄PPi	100 mM	5 µl
1 mM Na₃VO₄	500 mM	0.2 µl
0.1 mg/ml Trypsin inhibitor (TI)	10 mg/ml	10 µl
varied	Proteinase inhibitors mix (PI): Antipain 1mg/ml; Aprotinin 5 mg/ml; Benzamidine hydrochloride 10 mg/ml; Leupeptin 1mg/ml	10 µl
10 µg/ml Pepstatin A	1 mg/ml	10 µl
1 mM AEBSF	400 mM	2.5 µl
0.5 mM PMSF	1 M	0.5 µl
H₂O		930 µl
Final volume		**1 mL**

Nuclear Franctionation buffer B		
Final concentration	Stock concentration	Amount
10 mM HEPES	1 M	10 µl
0.4 M NaCl	5 M	80 µl
5 mM EDTA	500 mM	10 µl
0.1 mg/ml Trypsin inhibitor (TI)	10 mg/ml	10 µl
varied	Proteinase inhibitors mix (PI): Antipain 1 mg/ml; Aprotinin 5 mg/ml; Benzamidine hydrochloride 10 mg/ml; Leupeptin 1 mg/ml	10 µl
10 µg/ml Pepstatin A	1 mg/ml	10 µl
1 mM AEBSF	400 mM	2.5 µl
H₂O		867.5 µl
Final volume		**1 mL**

6.2.2.2   Experimental procedure

Cells were detached from the culture plate in PBS and centrifuged at 900rpm for 5 minutes. The cell pellet was resuspended in 500 µl (per 6 well) Nuclear Fractionation buffer A. After incubation for 15 minutes in the cold room on a rotor, 1% v/v Igepal (NP-40) was added and the samples were centrifuged at 11.400 g for 15 minutes. The supernatant containing the cytosolic fraction was kept and the pellet containing the nuclei was resuspended in 1 ml Nuclear Fractionation buffer

B. After 30 minutes incubation at 4°C on a rotor the samples were centrifuged at maximum speed for 15 minutes. The supernatant containing the nuclear protein fraction was transferred to a new tube.

6.3   Western Blot

6.3.1   Western blot with chemiluminescence detection

6.3.1.1   Buffers and Solutions

Stacking gel 4.5%		
Final concentration	**Stock concentration**	**Amount**
4.5% Acrylamide	30%	3.7 mL
0.0008% TEMED	99%	0.02 mL
0.1% Ammonium Persulfate (APS)	10%	0.25 mL
0.1% SDS	10%	0.25 mL
125 mM Tris/HCl, pH 6.8	0.5 M	6.25 mL
H$_2$O		14.5 mL
Final Volume		**25 mL**

Separating gel 10%		
Final concentration	**Stock concentration**	**Amount**
10% Acrylamide	30%	16.6 mL
375 mM Tris/HCl, pH 6.8	1.5 M	12.5 mL
0.1% Ammonium Persulfate (APS)	10%	0.5 mL
0.1% SDS	10%	0.5 mL
0.0004% TEMED	99%	0.02 mL
H$_2$O		19.9 mL
Final Volume		**50 mL**

Seperating gel 8%		
Final concentration	**Stock concentration**	**Amount**
8% Acrylamide	30%	13.3 mL
375 mM Tris/HCl, pH 6.8	1.5 M	12.5 mL
0.1% Ammonium Persulfate	10%	0.5 mL
0.1% SDS	10%	0.5 mL
0.0004% TEMED	99%	0.02 mL
H$_2$O		23.2 mL
Final Volume		**50 mL**

SDS-PAGE sample buffer		
Final concentration	**Stock concentration**	**Amount**
0.4M TRIS/HCl, pH 6.8	1 M	40 ml
40% Glycerol	99%	40 ml
8% SDS		8 g
0.4 M DTT		6.17 g
0.5% bromophenol blue		0.5 g
H$_2$O		**bring to 100 mL**

10X Running buffer		
Final concentration	Stock concentration	Amount
144 g/L L-Glycine		144 g
30.25 g/L Trizma Base		30.25 g
10 g/L SDS		10 g
H$_2$O		**bring to 1000 ml**

1X Transfer buffer		
Final concentration	Stock concentration	Amount
1x Transfer buffer	10x	100 ml
20% MetOH	100%	200 ml
H$_2$O		700 ml
Final Volume		**1000 mL**

10X Transfer buffer		
Final concentration	Stock concentration	Amount
L-Glycine	144 g/L	144 g
Trizma Base	30.25 g/L	30.25 g
H$_2$O		**bring to 1000 ml**

1X TBST		
Final concentration	Stock concentration	Amount
TBS 1X	10x	100 ml
0.05% Tween 20	20%	25 ml
H$_2$O		875 ml
Final Volume		**1000 mL**

10X TBS		
Final concentration	Stock concentration	Amount
NaCl	80 g/l	80 g
Trizma Base	24 g/L	24 g
to pH 7.5 HCl	37%	13.25 ml
H$_2$O		**bring to 1000 ml**

6.3.1.2 Experimental procedure

The concentration of protein samples was normalized to a reference sample using BCA protein assay kit (Pierce, 23227) in a final volume 100 µl. They were supplemented with Laemmli buffer at 4X dilution, boiled at 95°C for 5 minutes and centrifuged briefly for 1 minute. The samples were then loaded on 8% or 10% SDS-PAGE gels. Protein migration was initiated by applying a constant current of 20mA per gel. The protein gel was then placed on a nitrocellulose membrane and protein was transferred at 100V for 1.5 hours in the cold room using the Hoefer wet-transfer system. After that the membrane was stained with Ponceau S to verify the transfer of the proteins. After a brief wash with TBS-T the membrane was blocked in 5% milk w/v in TBST for 1 hour at room temperature. The membrane was later

incubated with the desired primary antibody diluted in 5% milk w/v-TBST in the cold room shaking overnight. The day after, the membrane was washed with TBST and the membrane was incubated with HRP-conjugated secondary antibody 25,000X diluted in 5% milk w/v-TBST at room temperature for 1 hour. The membrane was washed three times in TBST, for 15 minutes each at RT under constant shaking. The membrane was incubated with approximately 300 µl enhanced chemiluminescence (ECL) solution for 2 minutes wrapped in a plastic membrane and chemiluminescence was detected using BioRad ChemiDoc imaging system.

6.3.2   Western blot using fluorescence secondary antibodies

6.3.2.1   Reagents and Instruments

Reagent	Catalog number	Vendor
NuPAGE™ MOPS SDS Running Buffer (20X)	NP000102	Invitrogen
Transfer buffer NuPAGE™ Transfer Buffer (20X)	NP00061	Invitrogen
IRDye® 800CW Donkey anti-Rabbit IgG Secondary Antibody	926-32213	LI-COR Biosciences
IRDye® 680RD Donkey anti-Mouse IgG Secondary Antibody	926-68072	LI-COR Biosciences
Nitrocellulose Membrane, Roll, 0.2 µm, 30 cm x 3.5 m	1620112	Bio-Rad
NuPAGE™ 4-12% Bis-Tris Protein Gels, 1.0 mm, 12-well	NP0322BOX	Invitrogen
Mini Gel Tank	A25977	Invitrogen
Trans-Blot Turbo Transfer System	1704270	Bio-Rad

6.3.2.2 Experimental procedure

The protein samples were prepared as described above and the samples were supplemented with NuPAGE LDS Sample Buffer at 4X dilution, boiled at 70°C for 10 minutes and loaded on NuPAGE 4-12% Bis-Tris Protein Gels in 1X MOPS running buffer. Protein migration was initiated by applying 180 V for 35 minutes. The membrane was then placed on a nitrocellulose membrane pre-soaked in Transfer buffer and both were surrounded with filter paper. The blotting sandwich was then placed in a Trans-Blot Turbo Transfer System cassette and the proteins were transferred by applying 1.3A/25V for 15 minutes. After that the procedure was the same as described above. The membrane was incubated with IRDye 800CW and/or 680RD secondary antibodies in 10,000X dilution for 1 hour in the dark. The

fluorescent signal was detected using the Odyssey CLx imaging system and analyzed by using Image Studio Software.

Antibodies for Western blot			
Antibody	Supplier	Catalog No	Species
α-ADA2b	Novus Biologicals	#NBP1-91649	Rabbit
GAPDH	Sigma	G9545	rabbit
EOMES	Abcam	ab23345	rabbit
GATA6	Cell Signaling	5851	rabbit
GSC	Abnova	H00145258-M01	mouse
NANOG	BD	560482	mouse
FOXA2	Cell Signaling	8186	rabbit
Beta ACTIN	Sigma	A5316	mouse
OCT4	Santa Cruz	sc-9081	rabbit
OCT3/4	BD bioscience	611203	mouse
pSMAD2/3	Cell Signaling	8828	rabbit
pSMAD2	Cell Signaling	3101	rabbit
SF2	Santa Cruz	sc-33652	mouse
SMAD2	Cell Signaling	5339	rabbit
TUBULIN	Sigma	T8328	mouse
β-CATENIN	BD bioscience	610153	mouse
TEAD1	BD	610923	mouse
PBX1b	Santa Cruz	sc-101852	mouse
PREP-1	Santa Cruz	sc-25282	mouse

6.4 RNA

6.4.1 RNA isolation from hESCs

RNA was extracted from hESCs harvested using TRIzol or RLT (QIAGEN) followed by RNeasy columns and on-column DNase I digestion (QIAGEN) according to manufacturers' instructions. The purity and integrity of RNA was validated by a running a 2% agarose gel.

6.4.2 cDNA synthesis

500 ng to 1µg of RNA were reversed transcribed using the Promega GoScript Kit according to the manufacturer's protocol.

RNA amount	Random primer	ddH20
500 to 1000 ng	2.5 µl	up to 10 µl

PCR program RT1		PCR program RT2	
Temperature	Time	Temperature	Time
70°C	5 minutes	25°C	5 minutes
4°C	infinite	42°C	60 minutes
		70°C	15 minutes
		4°C	infinite

Reactions	Reagents				
	GoScript™ 5X Reaction Buffer	MgCl2 (25mM)	dNTP	ddH20	GoScript™ Reverse Transcriptase
1	4 µl	2 µl	1 µl	2 µl	1 µl

6.4.3 qRT-PCR

cDNA was diluted to 100 ng and 1.2 µl were used for qPCR with SYBR Fast 2x Master Mix (Kapa Biosystems) on a Roche 480 Lightcycler® or 2 µl were used for qPCR with SsoFast SYBR Green supermix (Bio-Rad) on a Applied Biosystems QuantStudio 6 Flex Real-Time PCR System. ΔCt was calculated by normalizing each sample to the average Ct of *HPRT*, *GAPDH* and *18srRNA* housekeeping genes. ΔΔCt was calculated by normalizing each condition to wild type (WT) or untreated expression levels. All graphs and statistical analyses were performed in Prism 8 (GraphPad). Primer sequences are available on Table 6-1. Fold change expression values were calculated according to the ΔΔCT method.

qPCR master mix with SYBR Fast mix (Kapa Biosystems)	Reaction 1
Reagent	Amount
SYBR Fast 2x Master	6 µl
5 µM Primer mix	0.6 µl
ddH$_2$O	4.2 µl
cDNA (added last separately)	1.2 µl
Total volume	12 µl

qPCR master mix with SsoFast SYBR Green supermix (Bio-Rad)	Reaction 1
Reagent	Amount
SYBR Fast 2x Master	6 µl
5µM Primer mix	0.5 µl
ddH$_2$O	2 µl
cDNA (added last separately)	2 µl

Total volume	10 µl

RT-qPCR Primers		
Gene	Forward primer	Reverse primer
TBX3	CGCTGTGACTGCATACCAGA	GTGTCCCGGAAACCTTTTGC
HAND1	TCAAGAAGGCGGATGGCGGC	TCCAGCGCCCAGACTTGCTG
MSGN1	CGGAATTACCTGCCACCTGT	GGTCTGTGAGTTCCCCGATG
TBX6	AAGTACCAACCCCGCATACA	TAGGCTGTCACGGAGATGAA
NODAL	AGAATGTGGGTGCCTCTGATGACA	AGCCCTTCATTTACAGAGTGGGCA
SOX17	GAGCCAAGGGCGAGTCCCGTA	CGTCAGCGCCTTCCACGACT
FOXA2	GATTGCTGGTCGTTTGTTGTGGC	GAGTAGCCCTCGGGCTCTGCAT
HHEX	CACCCGACGCCCTTTTACAT	GAAGGCTGGATGGATCGGC
HNF4A	CGTGCTGCTCCTAGGCAATGAC	ACGGACCTCCCAGCAGCATCT
MIXL1	GGTACCCCGACATCCACTTG	TAATCTCCGGCCTAGCCAAA
FOXH1	GGCCCCCAGAGGCAGAGTCG	ACCTGACGGATGATCTGGGCCA
GSC	TCAACCAGCTGCACTGTCGGC	TCCATTTGGCGCGGCGGTTC
T	TGCTTCCCTGAGACCCAGTT	GATCACTTCTTTCCTTTGCATCAAG
YEATS2	GAACACTGGGAGTCAGCACA	TCTCAGCATTGTGGTTCCAG
HOMEZ	CTGGACTGCGCTATCTCTGAA	CTGAAGGTTTTGAGCAGGTGT
PBX1	ACAGAGCCAAACTCTCACAGAT	TGACCATCCGCTCAATCTCC
ZNF281	GGAGCCAAAGCAGGACACTA	GCAGGTGATAGGAGCTTCGG
PKNOX1	ATCAGCCCTCAGGGAATTGTG	GACCGTGACAGGCTGATACA
PBX2	ATGGACGAACGGCTACTGG	CCCCGATGTCTTGCTTCC
NKX1-2	CAGGAAAAGTTTGGCGGAGG	TCAGGTGAGCGCGCTAGG
18S rRNA	GCTTAATTTGACTCAACACGGGA	AGCTATCAATCTGTCAATCCTGTC
GAPDH	TGGACTCCACGACGTACTCA	AATCCCATCACCATCTTCCA
GSC	GAGGAGAAAGTGGAGGTCTGGTT	CTCTGATGAGGACCGCTTCTG
HOMEZ	CTGGACTGCGCTATCTCTGAA	CTGAAGGTTTTGAGCAGGTGT
NANOG	AATGGTGTGACGCAGGGATG	TGCACCAGGTCTGAGTGTTC
OCT4	TGGGTGGAGGAAGCTGACAACAAT	TTCGGGCACTGCAGGAACAAATTC
TADA2B	GGTGATGGAGCATTACGTGAG	CCTGGTCATACTCGATCTCGT
EOMES	AGAGGGCTGTGCCTTCCGTTTC	AGCACACAGCAGAGGCCTAGCAAG
CDX2	CTGGAGCTGGAGAAGGAGTTTCAC	GACACTTCTCAGAGGACCTGGCTG
ASH2L	AGCAGCGGAAAAGGACGAG	GTTCCAATGGGTAGCCATGAG
GATA6	GAGGGTGAACCCGTGTGCAATG	TGGAAGTTGGAGTCATGGGAATGG

Table 6-1. List of primers used for RT-qPCR.

6.5 DNA

6.5.1 Genomic DNA isolation from hESCs

DNA was extracted with the QuickExtract DNA extraction solution following the manufacturer's instructions. 30 µl of QuickExtract solution was used for cells growing in a 24 well. DNA was diluted to 50 ng/µl and 2 µl were used for PCR amplification of the desired target site.

6.5.2 Validation of CRISPR-CAS9 mediated deletion by PCR

The successful deletion of the desired genomic region by paired sgRNAs was validated by genomic DNA PCR using flanking primer set. Primers used are listed in Table 6-2.

Genes amplified	Forward primers	Reverse primers
TADA2B	AGCATGTACATCCACGGGAAC	GCACACTGTCTGAGCATCAC
ASH2l	ATGGAGTACGGAAAGGTGCC	TAAGCAGAGGGTGAGCTTGC
MGA	TGGCTAATCAAGATGGTGGAACA	CACTAAGTGCTGAATCAGTATCCA

Table 6-2. Primers for genotyping

6.5.3 TOPO cloning and Sanger sequencing of CRISPR-Cas9 edited clones

The region spanning the exon 2 of TADA2B gene was amplified by PCR using Phusion polymerase (M0530L, NEB). PCR products were cloned into pCR™4Blunt-TOPO® vector (Invitrogen) and transformed into DH5α bacteria. 10 individual bacterial colonies per PCR (edited clone) were picked and plasmid DNA was isolated using QIAprep Spin Miniprep Kit (Qiagen) and sequenced using M13 forward primer. The acquired by capillary electrophoresis chromatograph was compared to the WT sequence.

6.5.4 PCR with biotinylated primers for DNA Pull Down

HPLC purified T7 with biotin modification at the 5' prime end and T3 reverse primers (Microsynth) were used to perform multiple PCRs with Taq polymerase (NEB) on the desired enhancer sequences (300-440bp) resulting in 5' biotinylated enhancer fragments. The PCR products were pooled together and purified using the PCR Clean-up kit (Macherey-Nagel).

6.6 Cloning

6.6.1 Cloning of Enhancer-sequences to pBlueScript for DNA- Pull Down

Genomic DNA from hESCs was isolated and PCR was performed with primers listed on (Table 6.3). The primers were designed with a restriction recognition site at the 5' prime end for EcoRI (forward primer) and SpeI (reverse primer). The PCR product of approximately 400bp was digested with indicated restriction enzymes and loaded on a 1.5% agarose w/v TAE gel. The expected band was excised from the gel and DNA was purified using QIAquick Gel Extraction Kit (Qiagen). The purified digested fragments were ligated into pre-digested with the same enzymes pBluescript vector using Quick Ligation Kit (NEB) and transformed into DH5α competent bacteria. The next day colonies were selected and Sanger sequenced to verify the correct sequence of the construct.

Regulatory regions amplified	Forward primers	Reverse primers
NANOG enhancer	AAAT**CTCGAG**TCAGATAGCTTCCTAAACC (***Xho*I**)	TCAG**ACTAGT**ATTCCGGAGAAGATTAAGC (***Spe*I**)
NANOG promoter	GTT**GGATCC**GTGGTGAACCTAGAAGTAT (***Bam*HI**)	AACAT**GAGCTC**ACCAGCTCAGTCCAGCAG (***Sac*I**)
GSC enhancer	GCAG**GAAtTC**AACGTGTGGCCCGGTCTCGC (***Eco*RI**)	GCAAA**ACTagT**CCAGGGCACAGAATGGCTGG (***Spe*I**)
EOMES enhancer	AACA**GAATTC**TAATGGCCGCCTGGAGGGG (***Eco*RI**)	TCCT**ACTAgt**CGCCATTACACTTAGACTGAGC (***Spe*I**)
Control sequence	aat**GGTAC**CCCATTCATACTCTTCTAGGGCTG (***Kpn*I**)	aac**GAGCTC**AGAATATAAGCTCTGTGAGGGC (***Sac*I**)

Table 6-3. Primers used for amplification of regulatory regions of *NANOG*, *OCT4*, *GSC* and *EOMES* from genomic DNA.

6.6.2 Cloning of single guide RNAs for genome editing with Cas9 or gene silencing with dCas9-KRAB

Oligos sequences were chosen according to highest efficiency ranking and lowest off-site effect by crispr.mit.edu. For cloning into pgRNA-CKB (Addgene #73501) the BsmBI restriction recognition site "TTGG" overhang was added at the 5' of the forward oligo and "AAAC" overhang was added at the 5' of the reverse oligo. For cloning into px458 (Addgene #48138) the BbsI restriction recognition site "CACCG" overhang was added at the 5' of the forward strand and "AAAC" overhang was added at the 5' of the reverse strand. Each pair of single stranded oligos was annealed at 95°C for 5 minutes and then incubated at room temperature for 1-2 hours.

Oligos annealing	
Reagent	Volume
Oligo forward	10 µl
Oligo reverse	10 µl
Tango Buffer 3.1	10 µl
H$_2$O	70 µl

The annealed oligos were ligated into the px458 or pgRNA-CKB vector following the reaction below:

Ligation of Oligos to Cas9 vector	
Reagent	Amount
Cas9 vector	150 ng
double stranded oligos	1 µl
T4 ligase buffer	2 µl
H$_2$O	to 18 µl
*Bsm*BI or *Bbs*I	1 µl
T4 ligase	1 µl

for *Bbs*I ligation

for *BsmB*I ligation

Temperature	Time	Cycles
37°C	5 minutes	
16°C	10 minutes	10
37°C	15 minutes	
80°C	5 minutes	

Temperature		Cycles
55°C	5 minutes	
16°C	10 minutes	10
55°C	15 minutes	
80°C	5 minutes	

Approximately 3µl of ligation were used to transform 50 µl of DH5α bacteria. The correct insertion of the desired sequence into the Cas9 vector was verified by bacterial colony picking and Colony PCR. For this the forward oligo and the StrepPyoCas9-5UTR-F sequencing primer were used. Correct colonies were analyzed by Sanger sequencing.

6.6.3   Bacterial Transformation

Approximately 4µl of ligation were used to transform 50 µl of DH5α bacteria. After 30 minutes on ice the mixture was heat-shocked at 42°C for 30 seconds. 100µl warm LB was added and plated on Petri dishes with selection antibody and bacteria culture was grown for 12-18 hours at 37°C.

6.7   Chromatin Immunoprecipitation ChIP

6.7.1   Buffers and Solutions

Wash buffer A		
Final concentration	Stock concentration	Amount
20mM Tris-HCl (pH=7.6)	1 M	10 ml
140mM NaCl	5 M	14 ml
1mM EDTA	500 mM	1 ml
0.1% NaDeoxycholate	10%	0.5 g
0.1% SDS	20% v/v	2.5 ml
1% Triton X	100% v/v	5 ml
H₂O		up to 500 ml

Wash buffer B		
Final concentration	Stock concentration	Amount
20 mM Tris-HCl (pH=7.6)	1 M	10 ml
500 mM NaCl	5 M	50 ml

Final concentration	Stock concentration	Amount
1 mM EDTA	500 mM	1 ml
0.5% NaDeoxycholate		2.5 g
1% Triton X	100% v/v	5 ml
H$_2$O		**up to 500 mL**

Wash buffer C		
Final concentration	Stock concentration	Amount
20 mM Tris-HCl (pH=7.6)	1 M	10 ml
1 mM EDTA	500 mM	1 ml
0.5% NaDeoxycholate		2.5 g
1% Triton X	100% v/v	5 ml
250 mM LiCl		5.3 g
H$_2$O		**up to 500 mL**

TBS Tris Buffered Saline		
Final concentration	Stock concentration	Amount
20 mM Tris-HCl (pH=7.6)	1 M	10 ml
50 mM NaCl	5 M	5 ml
1 mM EDTA	500 mM	1 ml
H$_2$O		**bring to 500 mL**

Hypotonic lysis buffer		
Final concentration	Stock concentration	Amount
25 mM HEPES pH=7.8	1 M	2.5 ml
1.5 mM MgCl$_2$	1 M	0.15 ml
10 mM KCl	1M	1 ml
Protease inhibitors (as described in Table 6.2.2.1)		
H$_2$O		**up to 50 mL**

5X Sonication buffer		
Final concentration	Stock concentration	Amount
50 mM HEPES, pH 7.9	1M	12.5 ml
140 mM NaCl	5M	7 ml
1 mM EDTA	0.5M	0.5 ml
1% Triton X-100	100% v/v	2.5 ml
0.1% NaDeoxycholate	5% v/v	5 ml
0.1% SDS	20% v/v	1.25 ml
Protease inhibitors (as described in Table 6.2.2.1)		
H$_2$O		**up to 50 mL**

6.7.2 Experimental procedure

1.6 * 10^7 H1 hESCs treated with DMSO or APS differentiated for 20h were used per IP. Briefly, 90% confluent cells growing onto Matrigel-coated 10 cm plates were washed with PBS and fixed with 1% Formaldehyde in PBS for 10 minutes at room temperature. After aspiration of the fixative solution excess of formaldehyde was quenched away by incubation with 0.125 M Glycine for 5 minutes. Cells were

then washed with PBS, scraped down in PBS containing 0.5 mM PMSF and centrifuged for 10 minutes at 720 RCF. The pellet was resuspended in 1 mL Hypotonic lysis buffer and incubated for 30 minutes on ice. Following centrifugation for 5 minutes at 5000 rpm, the pellet was resuspended in 800 µl of sonication buffer. The sample was transferred to a 1 ml milliTUBE (Covaris) avoiding bubbles. The chromatin was sheared by sonication to 200-500bp using on a S220 Focused-Ultrasonicator following the program: Peak Power: 140, Duty Factor 5.0, Cycles/Burst 200, time 960 seconds.

Sheared chromatin was centrifuged at 15,000 rpm for 10 minutes and supernatant was transferred to new tubes. A 50 µl aliquot was de-crosslinked by addition of 130 µl H_2O, 8 µl 5 M NaCl and 10 µl 1 M TRIS-HCl (pH8) and incubated for 4h at 65°C. After digestion with RNAse A an aliquot of 10 µl was run on an agarose gel to check the distribution of the sonicated chromatin. The rest of the 50 µl-aliquot was further digested with and Proteinase K and DNA was extracted using the MinElute PCR Purification Kit (Qiagen, 28004) to determine the concentration.10 µg of sheared chromatin were diluted in 400 µl sonication buffer. 10% of the volume was set aside and served as input and the rest samples were incubated with 2 µg of antibody of interest or IgG control and incubated rotating gently at 4 °C for 4 hours.

Samples were incubated with 25 µl protein G Dynabeads for 2 hours at 4°C. Afterwards beads were washed 2 times for 5 minutes each with Wash buffer A, B and C, followed by one washing step with TBS. Immunoprecipitated proteins were eluted in 100 µl freshly prepared 50 mM NaHCO3, 50 mM Tris-HCl pH8, 2 mM EDTA, 1%SDS for 30 minutes at 37°C. Supernatant was removed from beads on a magnetic rack and DNA was de-crosslinked by incubation with 4 µl 5 M NaCl per sample for 4 hours at 65°C. RNA was digested with 1 µl RNAseA (10 mg/ml) at 37°C for 20 minutes. Remaining protein was digested with 2 µl Proteinase K (10 mg/ml), 5 µl of 1 M TRIS/pH6.8, and 1 µl of 500 mM EDTA for 4h at 37°C. DNA was then purified using MinElute PCR Purification Kit (Qiagen) with addition of 3 M NaOAc and eluted in 50 µl elution buffer. 2 µl was used per qRT-PCR reaction. Quantitative PCR was performed as described above. Δct was calculated over the input value for the respective sequence of each sample. Fold enrichment over the IgG control was calculated using the ΔΔct method.

Antibodies used for Chromatin immunoprecipitation			
Antibody	Supplier/Catalog No.	Species	Amount per IP
α-H3K27ac	Abcam, #ab4729	Rabbit	1 µg
α-TADA2B	DSHB Hybridoma Product PCRP-TADA2B-1.1A3	Mouse	160 µl / 2 µg
α-TADA2B	Santa Cruz Biotechnology, #MB-56	Mouse	20 µl/2 µg
α-IgG	Cell Signaling Technology, #2729	Rabbit	2.5 µl
α-IgG	Santa Cruz Biotechnology	Mouse	sc-2025
α-H3K27me3	Millipore/#07-449	Rabbit	1 µg
α-Oct-4A	Cell Signaling Technology, #5677 - ChIP Formulated	Rabbit	8 µl

Genomic sequences used in Chromatin immunoprecipitation		
Target	Forward primer	Reverse primer
NANOG promoter	TGTCTTCAGGTTCTGTTGCTCGGT	TAACATGAGGCAACCAGCTCAGTC
EOMES enhancer 1	CCCGGGGATTCCCAGTCCGA	GCTCACCGGGAGGGCACTTG
POU5f1 enhancer	ACGGTAGGCCCCGTTCTCCC	TTGTCCTGCCCCCTTCCCCC
NANOG enhancer	AGCCACCAGCCCCTTCCCTT	CCACCGAGGCTGAGCTTGCC
GSC enhancer	TTCTCCGTGGGAACTGCAAA	GGAGGGCAACCTTATGGGAC
Control sequence from PD	ACCTGGTTTCAATGGTAGGCT	ACAGCCCTAGAAGAGTATGAATGG
TADA2B promoter	GCAGTTCGGCTTCGGAAACT	CGCGCCCGATTACAGGAG
GAPDH	CAGGCTGGATGGAATGAAAG	AAAGGCACTCCTGGAAACCT
EOMES enhancer 2	GGAGAAGGCACCCTTAACTGGATGT	GATCTTGCTCTGCACTTGCTCTGT

6.8   Loss of function approaches in hESCs

6.8.1   CRISPR-Cas9 mediated gene disruption in hESCs

H1 hESCs growing on Matrigel coated plates were maintained in mTeSR™1 media (Stem cell technologies) until they reach 90% confluency. 4 hours prior to electroporation, the media was supplemented with 10 µM of Y-27632. In total, 18 µg of 2 px458 plasmids carrying gRNAs targeting the exon 2 of the *TADA2B* gene were diluted in 100 µl Mirus ingenio solution and co-electroporated into $3*10^6$ cells in a 0.2 cm cuvette using B016 program in the Amaxa 2B nucleofector (Lonza). In

parallel $1*10^6$ cells were electroporated with 6 µg of a px458 plasmid carrying gRNAs targeting AAVS1 locus as a control. Electroporated cells were divided on 2 Matrigel-coated wells of a 6 well culture dish and maintained in mTeSR™1 supplemented with 10 µM Y-27632. GFP+ cells were FACS-sorted 36-48h after electroporation in 96 well plates coated with 1:20 diluted Matrigel. When colonies starting appearing cells were passaged to 24 well plates and maintained in regular hESCs culture conditions for expansion and genotyping.

6.8.2 siRNA mediated Knockdown in hESCs

siRNA mediated knockdown in hESCs was performed as described in [250] with modifications. Briefly 12-16h prior to transfection H1 or H9 hESCs were split with Accutase in a ratio of 1:9 and maintained in E8 media until transfection. 30 minutes prior to transfection 1.8 µl of siRNA (20 pmol/µl) was mixed with 2µl of RNAiMAX in 100 µl Opti-MEM and incubated at room temperature for 20 minutes. After that cell media was aspirated and replaced with 400 µl per well/12 well. The transfection mixture was added dropwise to the cells and media was replaced 8 hours after transfection. Differentiation of hESCs was initiated 30h after transfection. 48-72h after transfection cells were harvested for protein or RNA analysis. All siRNA knockdown experiments included not transfected hESCs as well as hESCs transfected with a scrambled siRNA (siControl). Sequences of siRNA duplexes used in this study are depicted in (Table 6.4).

6.8.3 Gene silencing in hIPSCs using CRISPR interference

CRISPRi human iPSCs cells stably expressing dCas9-KRAB upon doxycycline addition in the media were a kind gift from Bruce Conklin [251]. These cells were maintained under normal hESCs culture conditions. For specific gene silencing the top 4 gRNA sequences targeting close to the TSS of the gene of interest were chosen from [252] and cloned into the gRNA - expression vector (pgRNA-CKB-mKate2). 24 hours prior to transfection CRISPRi iPSCs were dissociated to single cells with Accutase. Following this, 2 µg of pgRNA-CKB carrying the gRNA of interest or empty vector were mixed with 6µl GeneJuice® transfection reagent and transfected into the cells that were pre-treated with 10µM of Y-27632 for 3-4 hours. Media was replaced to E8 without Y-27632 the day after. 48 hours after transfection Blasticidin selection was applied (5µg/ml) for 10-14 days until stable iPSCs colonies that express mKate2 were detected.

After successful integration 1.6 µM Doxycycline was added to cell culture medium for 4-5 days to induce the expression of dCas9-KRAB. Directed

differentiation was initiated 4 days after Doxycycline addition to the media. Efficient gene knockdown was verified by qRT-PCR and WB. All corresponding negative controls (minus doxycycline) were maintained in E8 for the same amount of time.

Integrated DNA technologies (IDT) DsiRNA Item name	Genes
hs.Ri.ZMYND8.13.6	ZMYND8
hs.Ri.ZMYM3.13.5	ZMYM3
hs.Ri.ZNF281.13.1	ZNF281
hs.Ri.YEATS2.13.7	YEATS2
hs.Ri.ZNF280C.13.1	ZNF280C
hs.Ri.POU5F1.13.1	POU5F1
hs.Ri.FAM208A.13.1	FAM208A
hs.Ri.FAM208A.13.2	FAM208A
hs.Ri.NOC2L.13.2	NOC2L
hs.Ri.TADA2B.13.1	TADA2B
hs.Ri.TADA2B.13.3	TADA2B
hs.Ri.MGA.13.1	MGA
hs.Ri.GSC.13.2	GSC
hs.Ri.GSC.13.3	GSC
hs.Ri.EOMES.13.2	EOMES
hs.Ri.EOMES.13.3	EOMES
hs.Ri.PKNOX1.13.1	PKNOX1
hs.Ri.PKNOX1.13.4	PKNOX1
hs.Ri.HOMEZ.13.3	HOMEZ
hs.Ri.HOMEZ.13.8	HOMEZ
hs.Ri.ZMYND8.13.3	ZMYND8
hs.Ri.ZMYND8.13.4	ZMYND8
hs.Ri.ASH2L.13.1	ASH2L
hs.Ri.ASH2L.13.4	ASH2L
#76982390	Negative Control
Horizon Discovery LTD	
siGENOME Human FAM208A (23272) siRNA - SMARTpool	FAM208A
siGENOME Human MGA (23269) siRNA - SMARTpool	MGA

Table 6-4. List of DsiRNAs used in this study.

gRNA Name	Genes	Targeted region	gRNA sequence **(PAM)**
g207-1	TADA2B	exon 2	TTCACAGTCACGTAGCGGGC**TGG**
g207-2	TADA2B	exon 2	CATCGCCCGTGACTACAATC**TGG**
g207-3	TADA2B	exon 2	GTGTGGTCTGTCACGCGGTT**GGG**
g211-1	MGA	exon 1	CACACTACTAGCCAAAGCAC**AGG**
g211-2	MGA	exon 1	CATCGGGTCCGTCTTACAGA**AGG**
g215-1	ASH2L	exon 11	CCAGATACCGCTGCCAGACT**GGG**
g215-2	ASH2L	exon 13	GCCATGACTCACCTCACTAT**GGG**
g215-3	ASH2L	exon 13	AGGCATGAGGAAACTTCCTG**AGG**

Table 6-5. gRNA sequences used for generation of generation of Knockout by CRISPR-Cas9.

gRNA Name	Genes	gRNA sequence
G71	NKX1-2	GGGTCTCCAGCTGTCGGACA
G72	NKX1-2	GCCGCCTCGGACTTGGATAG
G63	PBX1	GCTCCCGGCGCTTAAATCTG

G64	PBX1	GGCGGCGGCAGGCAAAGCAC
G65	PKNOX1	GCGGGCCGGTGTGATTGATA
G66	PKNOX1	GCGCCGCACTCCGAAAGGGA
G173	TADA2B	GACGGCGCCTGCGTACTGAG
G174	TADA2B	GGCCGGCCGAGAAGCACTCG
G175	TADA2B	GCAGCTGGTAGCCGTGGTAG
G176	TADA2B	GGCGCGCCCGATTACAGGAG
G54	YEATS2	GCAGGTTGCGGGGGTCGCTG
g117	NR6A1	GCCCGGCCGCGGCTCTCTCT
g118	NR6A1	GAGCGAGACCGGGGAGGAGA
g123	ZNF281	GCCATGCGTGCCGGTGCCGG
g124	ZNF281	GGACCCGTAAGTATTGCCGG
g150	HOMEZ	GGGAGGGTGAGTGTCTGTGT
g151	HOMEZ	GTAGGGGAGGGCAAGAAGGT
g110	TGIF1	GGAGCAGGAGCAGGGAACAA
g111	TGIF1	GGCAGGGCCAGTAGAGTTCG

Table 6-6. gRNA sequences used for generation of generation of knockdown CRISPRi.

6.9   Cell culture

6.9.1   Human pluripotent stem cell culture

6.9.1.1   Buffers and Solutions

Essential 8 hPSC medium			
Ingredient	Catalog number	Supplier	Amount
Gibco™ DMEM/F-12, HEPES	11330-057	Thermo Fischer Scientific	500 ml
Sodium Selenite 500x stock	S5261	Sigma	100 µl
Pen/Strep	151401-22	Thermo Fischer Scientific	5 ml
L-Ascorbic acid 2-phosphate Mg salt	A8960	Sigma	32 mg
Sodium bicarbonate	S6014	Sigma	271 mg
NaCl	S7653	Sigma	~250 mg
holo-Transferrin bovine	T1283	Sigma	5.3 mg
Insulin	I9278	Sigma	1 ml
TGFβ 5 µM	CA59	Novoprotein	7.5 µl
hFGF2 (40 ng/µl) aliquots	100-18B	Peprotech	10 µl
10 M NaOH	M6250	Sigma	to pH 7.4 (~250 µl)

6.9.1.2 Experimental procedure

H1 and H9 hESCs and CRISPRi iPSCs used in this study were routinely propagated in feeder-free in mTeSR1 (StemCell Technologies) or E8 media on cell culture plastics coated with 1:100 diluted Matrigel. Cells were maintained at 37°C with 5% CO2 and media was changed daily. Cells where either passaged to new plates using ReLesR (StemCell Technologies) in 1:10 ratio for maintenance or with Accutase (StemCell Technologies) for single cell dissociation and transfections. Experimental work on human embryonic stem cells was authorized by the Swiss Federal Office for Public Health, under the reference number R-FP-S-1-0008-0000.

6.9.2 Differentiation of hESCs

6.9.2.1 Buffers and Solutions

CDM2 Medium				
Ingredient	Catalog number	Supplier	Amount	Stock concentration
50% IMDM (+Glutamax, +HEPES, +Sodium bicarbonate;	31980-097	Gibco	44 mL	
50% F12 +Glutamax	31765-027	Gibco	44 mL	
1 mg/mL polyvinyl alcohol	P8136	Sigma	10 mL	10x stock
1% v/v concentrated lipids	11905-031	Gibco	1 mL	
450 µM monothioglycerol	M6145	Sigma	300 µL	150 mM
0.7 µg/mL insulin	I9278	Sigma	7 µl	10 mg/mL
15 µg/mL transferrin	T1283	Sigma	150 µl	10 mg/mL
1% v/v penicillin/streptomycin	151401-22	Gibco	1 mL	

Table 6-7. Composition of differentiation media CDM2.

6.9.2.2 Experimental procedure

50-60% confluent hESCs grown in E8 media were briefly washed with DMEM/F12 before adding the differentiation medium. All differentiation was conducted on matrigel-coated plates and monolayer conditions in chemically

defined CDM2 medium. The composition of CDM2 basal medium [57] is listed in Table 6-7.

hESCs were differentiated towards Anterior primitive streak for 24 hours (APS1) by supplementing CDM2 medium with 30 ng/mL ACTIVIN A, 4 µM CHIR99021, 20 ng/mL FGF2 and 100 nM PIK90. Subsequently APS1 cells were either differentiated towards Paraxial Mesoderm (APS2) for additional 24 hours by replacing APS1 media with CDM2 supplemented with 1 µM SB505124 + 3 µM CHIR99021 + 250 nM LDN-193189 [DM3189] + 20 ng/mL FGF2 or towards Definitive Endoderm by supplementing CDM2 with 30ng/mL ACTIVIN A and 250 nM LDN-193189 [DM3189] for 48 hours.

hESCs were differentiated towards Mid primitive streak (MPS1) in CDM2 with 30 mg/mL ACTIVIN A + 40 ng/mL BMP4 + 6 µM CHIR99021 + 20 ng/mL FGF2 + 100 nM PIK90 for 24 hours. MPS1 cells were further differentiated towards Lateral Mesoderm (MPS2) in CDM2 with 1 µM SB505124, 30 ng/ml BMP4 and 1 µM C59 for 24 hours.

Modulators of developmental pathways		
Item name	Catalog no.	Company
ACTIVIN A	120-14E	PeproTech
TGFβ	CA59	Novoprotein
FGF2	100-18B	Peprotech
CHIR99021	Axon 1386	Axon Medchem
C59	ab142216	ABCAM
PIK90	S1187	Selleckchem
BMP4	120-05	Peprotech
LDN-193189/DM3189	S2618	Selleckchem
SB 505124	3263	TOCRIS
Y27632	S1049	Selleckchem

Cell Culture Media and Reagents	Catalog Number	Supplier
mTeSR™1	85850	StemCell Technologies
mTeSR™ Plus	05825	StemCell Technologies

Essential 8			homemade Essential 8 human stem cell medium was prepared as previously described [253]
Accutase	07922		StemCell Technologies
ReLeSR	05873		StemCell Technologies
CDM2			homemade, prepared as previously described (Loh et al 2014)
Corning® Matrigel® Matrix, hESC-qualified	734-1440		Corning Life Science
CoolCell®	479-1842		Corning Life Science
Multiwell cell culture plastics, Falcon®	6 well 12 well 24 well 10cm dish	734-0019 391-0006 734-0020 734-0006	Corning Life Science

Cell lines	
H1 hESC	WiCell (WA01) human embryonic stem cells (XY)
H9 hESC	WiCell (WA09) human embryonic stem cells (XX)
CRISPRi iPSC	kind gift from Bruce Conklin described in [251]

2X Freeze media	Amount
FBS (FBS premium Pan bioTech)	60%
E8 media	20%
DMSO	20%
Mix 1:1 with stem cell media and store in CoolCell.	

6.10 Immunofluorescence (IF)

6.10.1 Buffers and Solutions

IF Blocking buffer		
Final concentration	Stock concentration	Amount
2% BSA		2 g
0.2% Gelatin		0.2 g
0.1% Triton X	100%	0.1 mL
0.012% Igepal	20%	30 µL
PBS		UP TO 100 mL
Final volume		**100 mL**

6.10.2 Experimental procedure

hESCs or *TADA2B* knockout cells were grown on Matrigel coated Nunc™ Thermanox™ Coverslips inserted in a 24well plate. Cells were briefly washed with PBS and fixed with 4% PFA in PBS for 15 minutes at room temperature. Cells were then washed three times with PBS and blocking buffer was applied for 1 hour at room temperature. Appropriate amounts of primaries antibodies (usually 1:100 dilution) were diluted in 80µl blocking buffer and the solution was pipetted on a parafilm slide placed in a humified chamber. The coverslips were removed from the plates and placed upside down on top of the solution and incubated overnight at 4°C. The coverslips were then placed back in the plate and washed 3 times with 0.1% Tween in PBS for 10 minutes each. The secondary antibodies together with DAPI were applied diluted 1:500 and 1:10.000 respectively in blocking buffer and incubated for 1 hour at room temperature in the dark. Cells were washed 3 times with PBS for 5 minutes each and fixed again with 4% PFA in PBS for 10 minutes. After 3 washes with PBS the coverslips were carefully mounted on microscope slides with Vectashield and pictures were acquired using a Zeiss Axio Observer.

Immunofluorescence			
Antibody	Supplier/Catalog No.	Species	Amount per IF
α-OCT4	BD Pharmingen /#611203	Mouse	1:100
α-GATA6	Cell Signaling/ #D6IE4	Rabbit	1:1600
α-EOMES	Abcam/ #23345	Rabbit	1:100
α-NANOG	BD Pharmingen/ #560482	Mouse	1:200
α-BRA	R&D Systems/ #AF2085	Goat	1:100

Additional material	Catalog number	Manufacturer
Iodoacetamide	I1149-5g	Sigma
Formic acid	94318	Sigma
Trypsin	U5113	Promega
BCA kit	23227	Pierce
Trizol	15596026	Invitrogen
RNeasy Mini Kit (50)	74104	Qiagen
SsoFast SYBR Green supermix	172-5270	Bio-Rad
GoScript Reverse Transcriptase Kit	A5003	Promega
Random Hexamer Primer	SO 142	Life Technologies Europe
QuickExtract DNA extraction solution	QE09050	Lucigen
Zero Blunt™ TOPO™ PCR Cloning Kit for Sequencing	450031	Invitrogen

Taq DNA Polymerase with ThermoPol® Buffer	M0267L	New England Biolabs
NucleoSpin® Gel and PCR Clean-up	740609.250	MACHEREY-NAGEL
QIAquick Gel Extraction Kit	28115	Qiagen
Quick Ligation Kit	M2200	New England Biolabs
milliTUBE 1 ml AFA Fiber	520135	Covaris
Poly (deoxyinosinic-deoxycytidylic) acid sodium salt (dIdC)	sc-286691	Santa Cruz Biotechnology
Pierce Streptavidin Agarose Resins	20353	Thermo Scientific
MinElute PCR Purification Kit	28004	Qiagen
Ingenio® Electroporation Solution	MIR 50117	Mirus
Ingenio® Cuvettes for the EZporator® Electroporation System and Lonza-Amaxa® Nucleofector® II/2b devices	MIR 50121	Mirus
GeneJuice® Transfection Reagent	70967	Merck Millipore

7 Results

7.1 Establishing a proteomics method to identify proteins binding to specific DNA regulatory elements

7.1.1 Differentiation of hESCs into Anterior primitive streak and Definitive Endoderm cells

During gastrulation, posterior epiblast cells will give rise to Primitive streak (PS), a hallmark in gastrulation which marks the onset of the three germ layers. At this stage, hESCs are in an intermediate stage before committing to a specific lineage. In order to identify proteins regulating pluripotency and/or PS specification, we first needed to efficiently differentiate hESCs towards PS. For this reason, we tested several published protocols by modulating TGFβ, WNT and FGF signaling in RPMI media [59]. Activation of WNT alone did not induce the expression of PS markers after 24h of differentiation (Figure 7-1A). However, the combination of WNT and ACTIVIN signaling activation resulted in induction of mRNA and protein levels of PS markers *BRACHYURY, EOMES, FOXA2* and *GSC* (Figure 7-1A, B). Although this protocol can be used for differentiation of hESCs towards PS, further differentiation towards definitive endoderm (DE) is insufficient (data not shown) and the media is supplemented with many unknown factors causing variability between batches. We therefore decided to use a newly published protocol by the Weissman laboratory [58] to differentiate hESCs in a chemically-defined media towards Anterior Primitive Streak (APS) cells and DE cells, which is the embryonic precursor to the liver and pancreas, among others. The protocol was tested by the authors for its efficiency in generating APS cells by acquiring a >98% pure MIXL1-GFP+ human PS population devoid of other undesired lineages [58]. Defined signaling perturbations yielded pure DE population from APS cells and later generated hepatic and liver progenitors [58]. Thus, we decided to apply this protocol in our culture system and efficiently differentiated hESCs towards APS (Day 1), which generated DE cells after additional 2 days of treatment (Day 3). 24h after initiation of APS differentiation, hESCs colonies, characteristic of the pluripotent state loosened up and gave rise to differentiating monolayer cells (Figure 7-1C), expressing high mRNA levels of APS markers *MIXL1, EOMES, GSC* and *TBX3* (Figure 7-1D). 48h later, extensive cell migration outside of the colonies was observed, marking the generation of DE cells expressing DE specific markers *SOX17, HNF4A and HHEX* (Figure 7-1C, D).

Figure 7-1. Efficient differentiation of hESCs towards APS and DE *in vitro*. A. Heatmap representation of quantitative RT–PCR analysis of the mRNA levels of *OCT4* and Primitive streak markers *BRACHYURY, EOMES, FOXA2* and *GSC* after 24h treatment of hESCs with ACTIVIN A, CHIR99021 or combination of both in RPMI medium. Data are presented as log2 fold change over untreated hESCs. B. Western blot depicting protein levels of OCT4, GOOSECOID, BRACHYURY and EOMES in untreated hESCs or differentiated cells treated with CHIR99021, ACTIVIN or both for 24 hours. GAPDH was used as a loading control. C. Morphology of hESCs colonies or hES cells treated with 30 ng/mL ACTIVIN A, 4 µM CHIR99021, 20 ng/mL FGF2 and 100 nM PIK90 in chemically defined media (CDM2) for 24h to induce Anterior Primitive streak (APS), followed by additional 48h treatment with ACTIVIN A and LDN-193189 to induce Definitive Endoderm (DE). Scale bars = 200µm. D. The upregulation of APS markers *MIXL1*, *TBX3, EOMES, GSC* and DE markers *HNF4A, HHEX* and *SOX17* were verified by RT–qPCR. The mean of 4 biological replicates was calculated and presented as log2 fold change over the untreated hESCs in a heatmap.

7.1.2 <u>Refinement of the DNA - Pull down method coupled to Mass spectrometry</u>

7.1.2.1 Identification of the enhancer sequences used as bait in DNA-pull down

Next, we sought to establish the DNA-Pull down method. Initially, we selected for our studies the enhancer of *NANOG*, as it is active in the pluripotent state and because of its necessity for the specification of hESCs towards the endoderm lineage [90]. On the other hand, we selected the enhancer of *GSC* as an inactive "poised" enhancer in pluripotency. *GSC* becomes activated during PS differentiation and plays important role in the organizer during gastrulation. We used published ChIP-seq data to select the exact enhancer sequences of *NANOG* and *GSC* based on their classical epigenetic signature (active in hESCs: H3K27ac and inactive in APS and DE: H3K27me3) (Figure 7-2) [57]. Previous results from our and other laboratories established that *NANOG* is directly regulated by SMAD2/3 and the responsible SMAD binding sites lie 400bp and 1800bp, upstream of the transcriptional start sites [102, 120, 254]. Furthermore, these sites are also occupied by POU5F1 and NANOG indicating the importance of these genomic regulatory sequences [120, 255]. Hence, we chose a 300bp fragment, which is 1800bp upstream of *NANOG*. This served as model of a genomic regulatory element for pluripotency, which is inactivated during differentiation. The *GSC* enhancer is a well-established target of SMAD2/3 and FOXH1 in the stage of APS induction [256]. It is well-known to be bound and negatively regulated by OCT4 in pluripotency [102]. The location of the regulatory sequence was determined by examining the binding sites of SMAD2/3, FOXH1 and OCT4 in published ChIP-seq data sets [120, 145, 255, 256]. The identified site (6000bp upstream of the TSS of *GSC*) shows the typical histone modifications of poised enhancers in hESCs, and active histone marks in APS cells [57]. We additionally included a control non-specific sequence which we identified by analysing ChIP-seq data. This sequence lies 3,000 bp upstream of the promoter of *NANOG* and has no indication of protein binding (Figure 7-2). Thus, we selected this sequence as negative control to identify specific binders and calculate fold enrichments.

NANOG

GSC

Figure 7-2. ***NANOG* and *GSC* regulatory sequences show the classic epigenetic signature of active and poised enhancers respectively**. IGV genome browser captures show compiled ChIP-seq (histone modifications, POU5F1, NANOG), within the *NANOG* and *GSC* genomic loci in hESCs. The boxes highlight the control sequence used in the pull-down experiments (#1), the active enhancer of *NANOG* (#2) and poised inactive enhancer of *GSC* (#3) regions.

After determining the DNA sequences, we performed a round of technical optimization experiments in order to ensure that this method will help us identify specific proteins binding to our selected genomic loci by mass spectrometry. We took advantage of the mechanism by which TGFβ signaling regulates the subcellular localization of phosphorylated SMAD2 protein. SMAD2 is a direct target

of TGFβ signaling, and a short 4h-inhibition with 10μM TGFβ inhibitor (SB431542) leads to dephosphorylation of SMAD2 and loss of its DNA binding potential. Based on this mechanism we performed DNA pull down on biotinylated *NANOG* and *GSC* enhancers using lysates from hESCs treated with DMSO, SB or differentiated towards from APS. Proteins bound to the biotinylated DNA were captured with streptavidin agarose beads and analyzed by WB (Figure 7-3A, B). We observed pSMAD2 binding to the enhancer of *NANOG* only when TGFβ signaling was present, whereas the addition of excess of non-biotinylated *NANOG* enhancer sequence successfully competed with the binding to the biotinylated DNA fragment (Figure 7-3B).

Next, we focused on optimizing the method using the enhancer of *GSC*. *GSC* expression is maintained silent during pluripotency and its regulatory regions are occupied by the OCT4, NANOG, TEAD1 and SMAD2/3/4 in hESCs [90, 96, 145, 257]. We differentiated hESCs towards APS by treating them with ACTIVIN and CHIR99021 (CHIR) for 24h and performed pull down on biotinylated *GSC* enhancer (Figure 7-3C, D). We could not detect binding of pSMAD2 on *GSC* enhancer in hESCs, possibly due to its weaker binding on poised enhancers during pluripotency [256] which cannot be detected by western blot. We observed stronger binding of pSMAD2 on *GSC* enhancer on PS differentiated cells, which was abolished when we additionally treated cells with SB or when we added excess of non-biotinylated DNA (Figure 7-3D). OCT4 and TEAD1 binding was observed in both pluripotent and PS state and is not dependent on pSMAD2 binding (Figure 7-3D) [258]. Taken together, these experiments demonstrate that our approach is feasible to detect by western blot known binders to those DNA regulatory elements in a signaling dependent fashion.

Figure 7-3. DNA pull down can specifically detect the binding of OCT4, TEAD1 and pSMAD2 on NANOG and GSC enhancers. A. hESC were treated for 4 hours with DMSO or a chemical inhibitor to the TGFβ pathway (SB) and total protein was extracted. The lysates were incubated with biotinylated DNA fragments of 400bp corresponding to the enhancer of NANOG. The proteins binding to the DNA were purified using streptavidin beads and analyzed by western blot. B. pSMAD2 binds to the enhancer of NANOG in hESCs and treatment with SB abolished its binding. Furthermore, addition of non-biotinylated competitor DNA eliminated the detection of sequence-specific binding proteins from the precipitate. C. hESCs were differentiated towards APS for 24h and at 20h of differentiation they were treated for 4h with DMSO or SB. The lysates were subsequently incubated with biotinylated 300bp DNA fragment corresponding to GSC enhancer and recovered proteins were analyzed by western blot. D. OCT4 and TEAD1 bind to GSC enhancer in pluripotency and differentiation conditions. Addition of 15x non biotinylated DNA abolished their binding showing the specificity of the method. pSMAD2 binding was observed only on differentiated cells and was not detectable after SB treatment.

For the discovery of novel binders by Mass-spectrometry the sequence-bound proteins should be compared to a non-specific condition. All Liquid chromatography–mass spectrometry (LC-M) experiments were performed in a collaboration with Dr. Alexander Schäfer from the Gstaiger group at the IMSB, ETH Zurich. For this, we performed pull down on the enhancer of *NANOG* with total protein lysates and before elution we added Benzonase. This treatment served as a control to show that the proteins bind specifically to our DNA bait (Figure 7-4A). We then performed western blot for pSMAD2, OCT4 and TEAD1 and observed decrease or loss of the protein in benzonase-treated samples (Figure 7-4B). Recovered proteins were analyzed by mass spectrometry and 43 binding proteins were identified showing a 2-fold enrichment over the benzonase-treated samples (>4 peptides). Among these were expected proteins regulating NANOG and pluripotency state, such as SALL1, SALL2, SALL4, ZNF281, POU5F1, NuRD complex (CHD4, MBD3, RBBP4), as well a few novel binders LUZP1, GNL3, TAF15, EHMT2 and ZNF462 (Figure 7-4C). However, the benzonase treatment alone cannot exclude that the identified are general DNA binders rather than sequence specific binders. Hence, in the following experiments we decided not to include benzonase treatment but to use a non-specific control sequence of the same size and GC content as the enhancer sequence. In this experiment we also noticed that the discovery rate of nuclear proteins was quite low, mainly due to high cytosolic contaminants, pointing out the necessity to use of nuclear protein lysates in future pull down-MS experiments.

Figure 7-4. Benzonase-treated DNA pull down followed by MS analysis to identify specific DNA binding proteins. A. Total protein lysates from hESCs were incubated with the biotinylated DNA fragment corresponding to the enhancer of *NANOG* and DNA pull down was performed. Before elution from the agarose beads, recovered proteins were treated with benzonase and analyzed by western blot and Mass spectrometry. B. pSMAD2, TEAD1 and OCT4 were detected to bind the enhancer of *NANOG* in hESCs. Their binding is reduced after benzonase treatment or addition of 10X competitor DNA. C. Recovered proteins eluted from the DNA-pull down were analyzed by Mass spectrometry and 43 proteins showed more than 2-fold enrichment over the benzonase treated samples. Only proteins with more than 4 identified peptides were included in the analysis. The size of the circles is relative to number of identified peptides. CB= chromatin bound, TF= transcription factor, NS=not specified.

As a next step of optimization, we tried several protocols for nuclear protein extraction from hESCs that is compatible with our method. We selected the protocol from Vallier et. al [104], showing enriched detection of transcription factors by DNA-pull down. More specifically, we treated hESCs with DMSO or SB for 4h and proceeded with total protein extraction as well as, nuclear and cytosolic separation of protein lysates. We detected by Western blot phosphorylated SMAD2 only in the total and nuclear protein samples (Figure 7-5A). This is expected since as soon as SMAD2 gets phosphorylated, it shuttles from the cytoplasm to the nucleus. We also observed highest enrichment of OCT4 in the nuclear fraction, which was devoid of GAPDH indicating the correct fractionation of the proteins.

We performed DNA pull down on *NANOG* enhancer as well on the promoter as a positive control, since this method was previously performed using a shorter fragment of *NANOG* promoter [102]. We observed stronger detection of OCT4 on the *NANOG* enhancer and promoter after DNA pull down with nuclear rather than total lysates (Figure 7-5A). Subsequently, we performed pull down on the enhancer of *NANOG* and *GSC* and the control sequence. Western blot on the precipitated samples confirmed binding of TEAD1 and OCT4 on *NANOG* and *GSC* enhancers and to lesser extent to the control sequence (Figure 7-5B). With the nuclear protein separation, we were able to enrich for nuclear binders and to reduce the background in the non-specific control sequence.

Mass spectrometry analysis of the precipitated proteins showed higher detection ratio of nuclear proteins than previous attempts. Out of 640 total identified proteins, 175 were nuclear and 68 were nuclear enhancer binders to *NANOG* or *GSC* with enrichment score more than 1.5 log2 fold change over the control sequence (Figure 7-5C). Interestingly, TEAD1 was one of the top enriched proteins binding to both *NANOG* and *GSC* enhancers in pluripotent and differentiated cells (Figure 7-5D). However, even though with these optimization steps we managed to detect more nuclear proteins by MS, we did not identify many transcription factors including the master regulator of pluripotency OCT4. Therefore, as a last step of optimization, we increased the amount of nuclear protein lysates used per pull down in order to detect low abundant proteins. We performed DNA Pull down on *NANOG, EOMES, GSC* enhancers and the control sequence using nuclear lysates from hESCs or APS cells. This increase in protein amount did not result in non-specific binding of OCT4 on the control sequence of as seen by western blot (Figure 7-5E).

Figure 7-5. DNA-pull down using nuclear protein extracts enriches for nuclear protein detection by MS. A. On the left: hESCs were treated with DMSO or SB for 4h and total protein (TP), nuclear protein fraction (NF) and cytoplasmic protein fraction (CF) were extracted. Detection of OCT4 is increased in the nuclear fraction. GAPDH is detected only in the TP and CF. pSMAD2 is not detected in the CF since after its phosphorylation it shuttles in the nucleus. On the right: DNA pull down was performed on *NANOG* enhancer, promoter or the control sequence using the same lysates and the detection of OCT4 binding to *NANOG* enhancer and promoter was stronger in the NF as demonstrated by western blot. B. On the left: Western blot depicting protein levels of total SMAD2 and OCT4 in the input. SF2 was used a loading control. On the right: Pull down on *NANOG, GSC* enhancers and control sequence showed no or reduced detection signal of TEAD1 and OCT4 on the control sequence. C. The precipitated proteins were analyzed

by MS. Out of 640 total identified proteins, 175 were nuclear and 84 were nuclear enhancer-binding proteins that showed 1.2-fold enrichment over the control sequence. D. TEAD1 was identified by MS as a strong and specific binder on *NANOG* and *GSC* enhancers. (Mean ±SD; N=2); Y axis: MS1 intensity. E. On the left: Western blot depicting protein levels of OCT4 and total SMAD2 in the input. SF2 was used as a loading control for nuclear protein lysates. On the right: Western blot for OCT4 after a DNA pull down on *NANOG*, *GSC*, *EOMES* enhancers or a control sequence, in DMSO treated or APS differentiated hESCs.

7.2 Capturing enhancer binders using DNA-Pull Down coupled to Mass spectrometry

After optimizing the method, we used it as a tool to identify new factors associated with *NANOG* and *GSC* regulation. Recovered proteins from three biological replicates of DNA pull down experiments on the enhancers of *NANOG* and *GSC,* as well as on a control sequence, using DMSO-treated or APS-differentiated cells, were subjected to mass spectrometry. The proteins were analysed using Progenesis QI for proteomics v3 based on peptide MS1 intensity. Overall, we identified 1262 proteins, of which 505 were nuclear.

To determine binders to our bait sequences we divided the MS1 intensity of proteins precipitating with *NANOG* or *GSC* enhancer over the ones with the control sequence, in pluripotency or differentiation, and set as threshold the fold change to be equal or more than 1.5. Out of the nuclear proteins, 117 were enriched on *NANOG* enhancer and 128 were enriched on *GSC* enhancer in any of the two treatment conditions (FC>1.5) (Figure 7-6A and Appendix) and 43 were shared between the two enhancers. This group of proteins included transcription factors, chromatin regulators and RNA binders. OCT4 and TEAD1 were among the top binders to our sequences, validating our previous results obtained by DNA pull down followed by western blot (Figure 7-6B). Apart from those two, we also identified other well-known regulators of the pluripotency network (ZSCAN10, ZIC3, ZIC2, SALL2, SALL3, SOX2), Primitive streak (EVX1, MIXL1, FOXA2, GATA6, EOMES, LEF1), signaling effectors (TGIF1, SMAD2, LEF1, β-CATENIN) and histone/chromatin remodelling complexes (NuRD, SWI/SNF, STAGA, INO80, PrG) (Figure 7-6C).

Figure 7-6. Identification of proteins occupying the enhancers of *NANOG* and *GSC* in pluripotency and APS differentiation by DNA-pull down followed by label-free MS analysis. A. In total 1262 proteins were identified, out of which 505 were nuclear. 117 nuclear proteins showed more than 1.5-fold enrichment on NANOG enhancer sequence and 128 on *GSC* enhancer. B. Relative OCT4 and TEAD1 protein abundance changes as identified by label-free mass spectrometry. y axis represents the normalized peptide MS1 intensity. Data are represented as mean ± SD (N = 3). *P <0.05, **P <0.01, ***P <0.001, ****P <0.0001. B. Nuclear proteins binding to *NANOG* or *GSC* enhancers with more than 1.5-fold enrichment over the control sequence are depicted here. LEF1, PKNOX1, PATZ1, TEAD1, NR6A1, ABT1, POU5F1, POU5F1B were identified as binders to both enhancers in both conditions.

Additionally, we directly compared the changes in protein binding on *NANOG* and *GSC* enhancers between the two treatments (Figure 7-7). As

expected, APS markers EOMES, MIXL1 and FOXA2 together with TGIF1 were enriched in the differentiated samples. Interestingly, PKNOX1, PBX1, PBX2 showed stronger enrichment on NANOG, whereas RBPJL on GSC enhancer, in both conditions. In addition, NKX1-2, YEATS2, TFAP4 and DR1 showed stronger enrichment on *GSC* enhancer at the APS differentiated state. An interesting remark from this analysis is that some proteins shift binding from one enhancer to the other depending on the cell treatment. For example, FAM60A, TASOR, ASH2L bind the enhancer of *NANOG* in hESCs, whereas in differentiated cells we found them binding the enhancer of *GSC*.

Figure 7-7. Differential protein binding profile on the enhancers of NANOG and GSC during pluripotency and APS differentiation. The goal of this analysis is to identify proteins dynamically changing between the conditions and the two DNA regulatory sequences. A. The log2 fold ratio was calculated for nuclear proteins purified with *NANOG* or *GSC* enhancers in APS differentiation over pluripotency conditions. B. The log2 fold ratio was calculated for nuclear proteins purified with *NANOG* over *GSC* enhancer in pluripotency, as well in APS differentiation. Data are represented in a volcano plot with y axis: negative log10-transformed p values and x axis: log2 fold change ratio.

After combining the information we acquired from the two different analysis approaches shown in Figures 7-6 and 7-7, we selected a group of candidates based on additional criteria: the novelty of their function in pluripotency or differentiation of hESCs, enrichment over the control sequence, published ChIP-seq data (if available), data from protein-protein interactions and mouse *in vivo* data (if available). Finally, we classified them to two groups: top binders PKNOX1, TGIF1, PBX1, NR6A1, MBTD1, NKX1-2, PBX2, CUX1 and binders showing a lower enrichment over the control but a dynamic behavior between treatments: FAM60A, TASOR, TADA2B, ZNF280C, TFCP2L1, ASH2L, MGA, YEATS2, HOMEZ, ZMYM3.

7.3   Validation of candidates

7.3.1   Functional characterization of PKNOX1-PBX1 complex

Initially, we focused on PKNOX1, PBX1, PBX2 as these proteins showed binding to the *NANOG* enhancer with strongest occupancy in APS cells (Figure 7-8A), and because PBX1 was also described to act as pioneer factor [259]. They are transcription factors belonging to the TALE (three amino acids loop extension) superclass of proteins that include MEIS1-MEIS3, PKNOX1 and PBX1-PBX4 [260]. Direct binding of PBX1 on the promoter and enhancer of *NANOG* has been previously shown by ChIP and EMSA [261, 262]. Since PKNOX1 binds DNA cooperatively with PBX1 [263], we hypothesized that these proteins act as a complex providing an explanation why they were all identified in our proteomics screen. Moreover, both of them have important roles in embryogenesis. PKNOX1 activates anterior HOX genes and its loss leads to embryonic lethality of mice [264]. PBX1 is also implicated in axial skeletal development by controlling HOX and Polycomb genes in mesoderm and mice with homozygous deletion of PBX1 are lethal at late gestational stage [265]. This strong phenotype in embryo development intrigued us to investigate this complex further.

At first, we validated the binding of PKNOX1 on the *NANOG* enhancer in APS- differentiated cells by DNA-Pull down followed by Western blot (Figure 7-8B). We next sought to investigate the role of this complex in hESCs. We monitored the mRNA levels of *PKNOX1* and *PBX1* during differentiation towards DE and found that they decreased strongly after 1 day of differentiation (Figure 7-8C). Interestingly the protein levels of PKNOX1 (Figure 7-8B) and PBX1 (data not

shown) in pluripotent and APS differentiated cells remain the same suggesting that they are regulated at the post- transcriptional level.

Figure 7-8. PKNOX1-PBX1-PBX2 complex was identified by DNA pull down-MS to bind on the enhancer of *NANOG* during APS differentiation. A. PKNOX1, PBX1 and PBX2 show enriched abundance on the *NANOG* enhancer in APS cells over the control sequence. y axis= MS1 intensities. Each data point represents a biological replicate with the line drawn at the mean (Mean ±SD; N=3); *P <0.05, **P <0.01, ***P <0.001. B. DNA pull down followed by WB on *NANOG*, *GSC* enhancers and control sequence using lysates from DMSO-treated or APS differentiated cells, confirmed the binding of PKNOX1 protein on *NANOG* enhancer in APS differentiated cells. C. mRNA levels of *PKNOX1* and *PBX1* are decreased 24h after differentiation towards APS and remain decreased during the course of differentiation towards DE. mRNA levels are presented as percentage over the control untreated hESCs (Day 1); (N=1).

We therefore hypothesized that this protein complex binds to the enhancer of *NANOG* to negatively regulate its expression and induce differentiation. In order to investigate this hypothesis, we suppressed the mRNA levels of *PBX1* and *PKNOX1* by siRNA in hESCs. We used 2 individual siRNAs to downregulate *PKNOX1* in DMSO-treated and APS differentiated cells. We observed a decrease in protein levels of NANOG after downregulation of PKNOX1 in differentiated cells (Figure 7-9A). The mRNA levels of *PKNOX1* were downregulated by 60-80% 72h after transfection in both conditions (Figure 7-9B). However, the mRNA levels of NANOG remained the comparable. Additionally, there was no effect on differentiation (24-48h) as seen by the mRNA levels of APS marker *GSC* (Figure 7-9A), a well as of *FOXA2* and *T* (data not shown). We speculated that genetic compensation mechanisms between the different members of this complex are masking a possible phenotype caused by the down-regulation of *PKNOX1*. Moreover, the fact that mRNA levels of *NANOG* do not show similar reduction as its protein, suggest a post-translational regulation, which is out of this study's scope.

Figure 7-9. Downregulation of PKNOX1 in hESCs and APS differentiated cells reduced NANOG protein, but not mRNA levels. A. Western blot depicting PKNOX1, NANOG, OCT4 and EOMES protein levels 72h after transfection with siRNA against *PKNOX1* or a scrambled siRNA (siControl). B. RT-qPCR data showing mRNA expression levels of *PKNOX1, NANOG* and *GSC*, 72h after transfection of hESCs with siRNAs against *PKNOX1* or a scrambled siControl. 24h before lysis cells were treated with Activin A (100ng/ml), CHIR99021 (2 μM) and PIK-90 (50 nM) in CDM2 medium for 24 hr to specify APS or with DMSO. *HPRT* and *GAPDH* were used as housekeeping controls for normalization of qPCR data. mRNA levels are presented as fold change over the siControl-transfected hESCs. Data are represented as Mean ±SD; (N=3).

Similarly, downregulation of *PBX1* by CRISPRi did not have an effect on self-renewal or APS differentiation (Figure 7-10). Since mRNA levels of *PKNOX1* and *PBX1* are decreased by 60-70% in PS and DE differentiated cells it would be interesting to overexpress these genes and study their effect on *NANOG* and their role in hESCs.

Figure 7-10. Downregulation of *PBX1* in hESCs does not influence self-renewal or APS differentiation. RT-qPCR data showing mRNA expression levels of *PBX1*, *NANOG,* as well as APS markers: *GSC, EOMES* and *FOXA2* in DMSO-treated or APS- differentiated cells. CRISPRi-iPSCs stably expressing a gRNA against *PBX1* were treated for 5 days with doxycycline to induce expression of dCAS9-KRAB and repress transcription of *PBX1*. On the 4th day of doxycycline treatment, cells were either treated

with DMSO or with 30 ng/mL ACTIVIN A, 4 µM CHIR99021, 20 ng/mL FGF2 and 100 nM PIK90 in CDM2 medium to differentiate them towards Anterior primitive streak for 24 hours (APS1). *GAPDH* was used as housekeeping gene for normalization of qPCR data. Data are represented as mean ± SD (N = 2).

These experiments would align with results from previous studies showing that overexpression of KLF4 and PBX1 upregulated *NANOG* promoter activity and also the endogenous NANOG protein expression in hESCs [261]. In summary, these data indicate that our established DNA pull-down system is able to capture binders to either *NANOG* or *GSC* enhancers. However, the fact that we could not reproduce the embryonic phenotype of PKNOX1 and PBX1 in our setting possibly is due to the limitations of *in vitro* cultured stem cells that lack the 3-dimensional structure as well possible compensation mechanism by PBX2 [265].

7.3.2   Functional characterization of *ASH2L, MGA, TADA2B*

We performed additional loss function experiments by siRNA or CRISPRi to evaluate the biological impact of the remaining candidates (see Table 7-1). Interestingly, siRNA-mediated downregulation of 3 candidates (*ASH2L, MGA* and *TADA2B*) showed phenotypical effect in APS differentiation (Figure 7-11). More specifically we transfected siRNAs against *MGA, TADA2B* and *AHS2L* and after 30h we induced APS differentiation for 18h (APS1) with further differentiation towards Paraxial mesoderm (APS2) for 24h. We observed morphological differences during differentiation of si*MGA*, si*TADA2B* and si*ASH2l* -transfected cells compared to the cells transfected with the control siRNA. This observation leads us to assess the expression of Primitive streak and Paraxial mesoderm markers. *ASH2L* and *MGA* knockdown cells showed defects in expression of Paraxial mesoderm (APS2) markers (*MIXL1, TBX6, GSC*) during differentiation (Figure 7-11). On the other hand, cells transfected with siRNAs against *TADA2B* showed morphological advanced differentiation when switched to APS media compared to the untreated and siControl transfected cells (data not shown). This phenotypic difference was accompanied by transcriptional upregulation of PS marker *GSC* (Figure 7-11). This enhanced differentiation upon loss of *TADA2B* could indicate a role of *TADA2B* in controlling pluripotency. Thus, we proceeded in functional characterization of *TADA2B* in hESCs as well as validation of its genomic binding to the enhancer.

Figure 7-11. Downregulation of MGA, ASH2L and TADA2B results in deregulation of PS markers.
hESCs were transfected with siRNAs against *MGA* (pool of siRNAs), *ASH2L*, *TADA2B* or a scrambled siControl sequence and differentiated for 18h towards APS and for additional 24h towards APS2. 72h after transfection RNA was harvested. A. RT-qPCR data showing mRNA expression levels of *MGA* as well as Anterior primitive streak (APS) and Paraxial mesoderm (APS2) markers *MSGN1, MIXL1, TBX6*. B. RT-qPCR data showing mRNA expression levels of *ASH2L*, APS and APS2 markers *GSC, MIXL1* and *TBX6*. C. RT-qPCR data showing mRNA expression levels of *TADA2B*, APS and APS2 markers *MIXL1* and *GSC*. *HPRT* was used as housekeeping control for normalization of qPCR data. mRNA levels are presented as fold change over the untreated APS2 differentiated cells. Data are represented as mean ± SD (N = 4). *P <0.05, **P <0.01, ***P <0.001, ****P <0.0001.

Summary of validation results of candidates by LOF studies				
Candidates	**Performed knockdown by CRISPRi**	**Performed knockdown by siRNA**	**Generated knockout cells by CRISPR**	**Observed phenotype**
PKNOX1	YES	YES	NO	NO
PBX2	YES	YES	NO	NO
TGIF1	YES	YES	NO	NO
CUX1	YES	YES	NO	NO
PBX1	YES	YES	NO	NO
ZNF280C	NO	NO	not genotyped	NO
TASOR	NO	YES	NO	NO
TADA2B	NO	YES	YES	YES
MGA	NO	YES	YES	YES
ASH2L	NO	YES	not genotyped	YES
ZMYM3	YES	YES	NO	NO
YEATS2	YES	YES	NO	NO
NKX1-2	YES	NO	NO	NO
HOMEZ	YES	NO	NO	NO

Table 7-1. Summary of LOF experiments to dissect the phenotype of the selected candidates.

7.3.2.1 TADA2B binds regulatory regions of *NANOG* and *GSC* in hESCs

In order to determine the genomic binding of TADA2B to enhancers of pluripotency genes or APS genes we performed ChIP for TADA2B. This revealed strong binding of TADA2B on the promoter (4-fold enrichment over IgG) and enhancer (6-fold enrichment over IgG) of *NANOG,* as well as on the enhancer of *GSC* (3-fold enrichment over IgG control) in the pluripotent state (Figure 7-12). We did not observe binding of TADA2B on the enhancer of another poised gene *EOMES* nor on the pluripotency factor *OCT4*. ChIP-seq for TADA2B would be more informative about the binding pattern of TADA2B on regulatory regions of pluripotency or early cell fate genes. As control sequence we used *GAPDH* as well as the control sequence we used in our pull-down experiments (PD control).

Figure 7-12. TADA2B occupies regulatory regions of NANOG and GSC in pluripotency. A. ChIP-qPCR for TADA2B in hESCs (blue) and APS differentiated cells (red). Signal was normalized to input DNA and calculated as fold enrichment over IgG control. Data points represent two independent experiments. GAPDH and PD control sequences were used as negative control sequences. PD control: pull down control sequence that was used in the Pull down-MS experiments. Each data point represents one biological replicate with the line drawn at the mean (Mean ±SD; N=2).

7.3.2.2   Generation of *TADA2B* knockout hESCs by CRISPR-Cas9 genome editing

We next sought to investigate role of TADA2B in hESCs. We employed a CRISPR-CAS9 approach to generate *TADA2B* gene knockout (KO) in hESCs. Briefly, we nucleofected hESCs with a plasmid carrying a CAS9-GFP construct together with a pair of sgRNAs targeting the exon 2 region of *TADA2B* gene. We sorted targeted cells expressing GFP, derived single cell clones and verified the desired deletion by PCR on the isolated genomic DNA. We performed TOPO cloning and sequenced the alleles of clones #2, #3, and WT1. Both alleles of clone #3 show deletion at the exon 2 of the gene, whereas one allele of clone #2, displayed an inversion of the targeted region while the 2nd allele a deletion. The non-targeted clone was not edited and we therefore used it as additional control (#1) for subsequent experiments (Figure 7-13).

Figure 7-13. Schematic representation of the paired-KO strategy for deletion of the exon 2 of human *TADA2B* gene and genotyping of targeted alleles of clone #2, #3 and WT.

7.3.2.3   Loss of TADA2B leads to Mesendoderm differentiation of hESCs

The loss of TADA2B protein on the edited clones was confirmed by western blot (Figure 7-14A). Remarkably, after passaging these cells for 1-2 times under defined conditions suitable for hPSCs, we observed a switch from an hESCs to meso-endodermal–like morphology (Figure 7-14B).

Figure 7-14. TADA2B KO hESCs show mesendoderm-like morphology. A. Western blot depicting loss of TADA2B protein in *TADA2B* knockout clones #2, #3 and #8. Control 1 (CT1) and Control 2 (CT2) are clones were derived from the same CRISPR targeting and sorting experiment but were not edited and thus served as control. B. Representative microscopy images of a WT hESC and a *TADA2B* KO clone. Scale bas = 200µm.

To systematically dissect the differentiation status of *TADA2B* KO cells we examined the expression of lineage-specific markers. *TADA2B* deletion led to a corresponding decrease in the expression of pluripotency genes (*OCT4* and *SOX2*), while resulting in upregulation of a subset of endoderm (*SOX17, GATA6, HHEX* and *HNF4A*), Primitive streak [*BRACHYURY (T), EOMES*] and trophectoderm (*CDX2*) markers (Figure 7-15A). Similar results were observed in immunofluorescence (IF) analyses (Figure 7-15B). 50-80% of *TADA2B* KO cells do not express OCT4 and a 10-30% fraction of those express the endoderm marker GATA6 (Figure 7-15C) indicating that these cells are in the intermediate between exiting pluripotency and early lineage commitment.

Figure 7-15. *TADA2B−/−* hESCs exhibit spontaneous differentiation to meso-endoderm fate. A. RT-qPCR data showing mRNA expression levels of markers of Primitive streak (*T, EOMES, GSC*), Endoderm (*SOX17, GATA6, HNF4A, HHEX*), trophectoderm (*CDX2*), pluripotency (*OCT4, NANOG, SOX2*) and *TADA2B*. *HPRT* and *18srRNA* were used as housekeeping control for normalization of qPCR data. The fold change over the WT hESCs was calculated and data are represented as mean ± SD (N = 4). *P <0.05,

**P <0.01, ***P <0.001, ****P <0.0001. B. Immunofluorescence for OCT4, BRACHYURY, GATA6 and DAPI in WT hESCs and *TADA2B* KO cells. Scale bar = 40 μm. C. Quantification of percentage of OCT4- or GATA6-positive cells normalized to DAPI-stained nuclei counted using ImageJ. Each data point in the scatter plot represents an independent biological replicate with the line drawn at the mean (Mean ±SD; N=2). Control 1 (CT1) and Control 2 (CT2) are clones were derived from the same CRISPR targeting and sorting experiment but were not edited and thus served as control.

In conclusion, these data suggest that our DNA-pull down method coupled to Mass spectrometry led to the discovery of TADA2B, a novel factor crucial for maintenance of hESC self-renewal.

8   Discussion and Outlook

The various cells of the developing embryo commit to distinct fate directions, yet they all originate from the same zygote, sharing identical genomes. During this journey of embryonic development, TFs, chromatin modifiers and signaling effectors play a fundamental role on regulating specification to different lineages. They mainly act on regulation of transcriptional outputs by collaborative or competitive binding on regulatory regions of genes driving pluripotency or differentiation programs. However, it is still unclear which is the hierarchy of events leading to transcriptional regulation and response to different lineage programs. Are pioneer factors present at poised enhancers before the signal for differentiation is activated? Do they recruit chromatin modifying enzymes increasing DNA accessibility for other factors essential for execution of the differentiation program? Or chromatin remodelers are the ones who modify the chromatin accordingly to different cell context and signaling, setting the stage for TFs to occupy the open chromatin and activate the transcriptional response? FOXH1, a master regulator of PS formation, is present in both pluripotent and PS state [57], however only upon appropriate signaling, a cell type specific complex is formed with other chromatin modulators and signaling effectors on specific genes driving the differentiation program towards PS. Important roles on this differential response play post-translational modifications that regulate protein-protein interactions or DNA affinity [266]. This can explain why TGFβ signaling has different phenotypic effects on hESCs or PS cells. It strongly relies on the ability of SMADs to partner with different cofactors in order to regulate differential transcriptional response [267-269]. The identification of those cofactors would enlighten the mechanism of TGFβ interplay between pluripotency and early cell fate decisions.

Ultimately TGFβ- and SMADs - cofactors regulate transcription by modulating the chromatin state at specific target genes. Poised genes necessary for PS induction will be released from repressive H3K27me3 mark and become activated. This unbalances the expression of pluripotency factors and eventually makes it possible for the cell to exit pluripotency and commit to a lineage. Other genes not necessary for that lineage, will retain the repressive H3K27me3 on their regulatory regions and remain silent [165, 166]. This pattern of activation and repression of specific genes at a specific time throughout development is essential for a stem cell to generate more than 200 tissues of an organism. It is therefore

becoming clear that perturbations on any of those players regulating this process can have deleterious effects on self-renewal or cell fate specification. Identifying the composition of complexes occupying these enhancers at different stages of development is crucial in order to elucidate the mechanism regulating the transition from self-renewal state to early cell fate choice. Many experiments have been conducted to unravel the chromatin environment around those enhancers, how it changes during different developmental stages and which components are essential for every process. Several TFs and cofactors were identified by genetic means as essential for embryonic development, and their genomic occupancy on enhancers was validated by ChIP - seq experiments. However, ChIP-seq experiments require *a priori* knowledge of the candidate and do not contribute in identification of new factors of these regulatory complexes. One way to uncover molecular players occupying specific genome loci, is *in vitro* DNA Pull down followed by mass spectrometry. We demonstrated in this study that this method can be used as a discovery tool for identification of proteins occupying active and poised enhancers in hESCs and differentiated cells. In order to distinguish real and novel candidates with a functional role in the hESCs state among the long list of acquired proteins, we used information from online databases such as protein-protein interactions (e.g. Biogrid), functional protein association networks (e.g. String) and Mouse genome informatics (Jackson laboratory). Chromatin Immunoprecipitation (ChIP) can then be used to verify the genomic occupancy of the candidates in a physiological context.

Among the identified proteins were ATP-dependent chromatin remodelers, histone-modifying enzymes and transcription factors. Many of them have been described to occupy enhancers of pluripotency or developmental genes. For example, we identified members of the SWI/SNF (EP400 and SMARCA5), NuRD (ZMYND8), INO80 (p400-TIP60, INO80D, TCF3 fusion partner), and PcG (L3MBTL3, L3MBTL2) complexes. We also identified proteins that are part of the pluripotency network (TEAD1, OCT4, ZIC3, ZSCAN10 and ZIC2), developmental programs (EVX1, MIXL1, FOXA2, GATA6, EOMES) or signaling effectors (TGIF1, SMAD2, LEF1, β-CATENIN). To our surprise, we could not identify FOXH1, the master regulator of PS formation, occupying enhancers of developmental genes such as *GSC* to activate their expression. Another factor that was missing from our dataset was p300 protein that occupies enhancers and regulates gene transcription [164, 270]. However, we identified many novel and interesting candidates (e.g.

ASH2L, MGAP, FAM60A, FAM208A, ZNF281) that are worth investigating further for their role in hESCs.

We noticed that most of the identified binders show a weak binding score, with homeobox-domain transcription factors showing strongest affinity to our bait sequences. This is expected since most chromatin remodelers lack sequence-specific DNA binding, whereas TFs also tend to occupy regions in a non-specific manner that does not involve recognition of a specific consensus but their binding to sequences is supported by other TFs or chromatin-bound proteins [266]. DNA pull down is an *in vitro* method in which proteins interact with the bait DNA through electrostatic interactions, hydrogen bonds, dispersion forces or hydrophobic interactions. In this system, high concentrations of proteins are "forced" to interact with large amounts of a long double stranded oligo that can lead to interactions that would not occur in living cells. This causes different levels of specificity of binding to the bait and high degree of non - sequence specific interactions. Proteins that interact with DNA not by recognizing specific sequence but by recognizing specific DNA topology and conformation, will not be possible to be captured using *in vitro* methods.

On the other hand, this method can efficiently identify sequence specific proteins that directly interact with DNA. This highlights the necessity not to select candidate DNA binders only based on mathematical computations but include criteria from biological tools we mentioned above. Moreover, the control sequence we have used in our experiments is not studied for its chromatin occupancy profile. Analysis of ChIP-seq data for OSN factors as well as active and repressive histone marks showed no enrichment on this control sequence. However, we cannot exclude that repressive or active chromatin modifying complexes are not occupying this sequence keeping it silent or TFs are bound there waiting for their recruitment to specific sequences. In retrospective, a scrambled sequence would have served as a better control.

Many proteins show dynamic binding changes from ESC to early cell fate choices. They can bind to regulatory sequences to activate gene expression but may not be necessary for the maintenance of their expression and get released from the relevant DNA. It can be therefore challenging to capture those dynamics using only one timepoint. Furthermore, this bait lacks the three-dimensional nucleosome landscape that surrounds the enhancers in their natural environment. Histone methylation or acetylation marks that maintain the active or repressed state of the gene, are therefore absent in this system. This chromatin signature changes

during differentiation leading to activation of specific regulators of cell lineages. However, with the DNA-pull down method we expect to find pioneer factors, such as FOXA2 and OCT4, that can occupy genomic sequences through sequence-specific binding, without need for open chromatin. Usually those pioneer factors recruit on the sequences they occupy, other cofactors including chromatin remodels and histone modifying enzymes for modification of the chromatin. One such example is recruitment of BRG1 by OCT4 [205].

We therefore concluded that it is possible to identify direct and indirect binders to our sequence with the method used. Indirect binders can have important roles in pluripotency and cell fate regulation by modifying the chromatin environment. Our method allows to capture those proteins on the enhancers of active and poised genes during different cell fates. Their low abundance in the cell and indirect binding on DNA can make it challenging to capture them by MS. Another method that could be used for identification of novel binders in their natural environment would be a method combining CAS9 recruitment to specific sequences followed by biotinylation of proximal proteins by BioID [271] and mass spectrometry. In this way proteins occupying genomic loci of interest can be identified in their chromatin environment.

Functional validation of candidates

In our study using the DNA pull down-MS method, we captured proteins binding to our enhancers for which we could not identify a functional role in hESC or early cell fate (e.g. PKNOX1, PBX1, HOMEZ, CDCDA7L, NKX1-2). This can be attributed to multiple reasons: many repressive complexes do not show a phenotype in *in vitro* pluripotency but show defects in differentiation or *in vivo* development, suggesting that molecular signatures are not always correlated with functional states [188-193]. Additionally, genetic redundancy can mask the phenotypic effect of an individual protein. Moreover, we conducted loss-of-function (LOF) studies by siRNA-mediated downregulation of gene expression in order to study the role of the candidates in hESCs or PS state. A more robust and efficient way to address this issue would be CRISPRi or CRISPRko which present less off target effects compared to siRNA targeting. In our case this would have been very challenging due to the high number of selected candidates and also the difficulty to genetically modifying genes in hESCs. Moreover, it is well known that protein binding on a genomic locus does not necessarily lead to control of transcription. There are three types of binding interactions between proteins and DNA: a) specific

binding with impact on transcriptional regulation, b) specific binding without functional role, and c) non-specific non-functional binding [266]. Therefore, it is expected the identification of DNA binding proteins with no functional role in transcriptional regulation.

Identification of TADA2B as a regulator of self-renewal in hESCs

It has been shown that TFs can regulate transcription by recruiting adaptor proteins and coactivators that have weak or non-existent DNA binding affinity, such as acetyltransferases-containing complexes on regulatory regions of target genes [239, 272, 273]. This certainly explains the identification of TADA2B as a binder to our enhancer sequences, even though TADA2B might also interact directly with DNA via its SWIRM domain, as it has been shown for TADA2A [226].

Role of STAGA members in embryo development

TADA2B is a member of the GCN5-STAGA complex together with GCN5 (KAT2A), TAF5L, TAF6L and TADA3. Some of these members are studied for their role in development, and particularly GCN5 and TADA3 are essential for mouse embryogenesis. *Gcn5*-null embryos fail to form mesoderm lineages including paraxial mesoderm and somites, have defects on neural tube closure and anterior-posterior patterning and die between E9.5-11.5 [274-276].

Tada3 null embryos present an earlier phenotype with embryonic lethality at E3.5 and strikingly, *Tada3* null blastocysts show no inner cell mass formation [277]. This urges for characterization of the role of *TADA2B* in development by genetic studies in mESCs and mouse embryos. Since deletion of *TADA2B* in hESCs has a tremendous effect in pluripotency it would be very interesting if *Tada2b* null blastocysts can form ICM or germ layer structures. This will give insights on the role of a novel protein in embryonic development.

Role of TADA2B and STAGA members in pluripotency

On the opposite side, loss of STAGA members in PSCs has distinct effects of the *in vivo* studies. *Gcn5*$^{-/-}$ mESCs show no aberrant growth or morphological defects and can be maintained pluripotent *in vitro* [214]. One explanation can be compensation mechanisms between GCN5 and PCAF [214]. Moreover, *Taf5l*$^{-/-}$ and *Taf6l*$^{-/-}$ mESCs show decreased levels of ESC-specific genes (*Oct4, c-Myc, Nanog*)

[278]. Although, this downregulation did not lead to loss of the pluripotent state or to differentiation, it slightly induced misexpression of several lineage-specific genes [278].

The role of TADA3 or TADA2B has not been investigated before in PSCs. In our study, loss of TADA2B apart from downregulation of *OCT4/SOX2*, also resulted in approximately 3- to 8-fold increase in expression of ME markers (*EOMES, GSC, GATA6, T*), TE marker *CDX2* and change of cell morphology towards endoderm-like cells. This indicates a more direct role of TADA2B in regulation of pluripotency compared to the other two members. Transcriptome analysis would help to characterize more thoroughly the differentiation status of *TADA2B* KO cells, as well as unravel whether other STAGA members, PRC2 or other complexes are deregulated. Generation of a doxycycline - inducible system to overexpress TADA2B and subsequent mutation of the endogenous gene, or inducible downregulation of *TADA2B* by CRISPRi and simultaneous overexpression of the protein, could show that loss of TADA2B is responsible for the observed phenotype. Moreover, whole genome transcriptome analysis of those cells at different timepoints, will unravel the sequence of events leading to the differentiation of the cells as well as characterize better their differentiation state.

Role of STAGA complex and TADA2B in reprogramming of somatic cells

Another interesting remark is that many members of STAGA complex are required for somatic cell reprogramming. *Taf5l-, Taf6l-, Trrap-* and *Gcn5-* deficient MEFs cannot generate pluripotent iPSCs [278, 279]. This shows that the STAGA complex (at least the above-mentioned members) is required to establish, and to lesser extent maintain pluripotency. An essential experiment would be to do somatic cell reprogramming of mouse embryonic fibroblasts (MEFs) targeted with *Tada2b*-gRNAs and uncover a possible role of TADA2B in iPSC generation. This can possibly reveal a requirement of TADA2B in iPSCs generation and can have an impact on regenerative medicine.

Genomic occupancy of STAGA members and TADA2B in PSCs

The fact that other STAGA members such as TAF5L and TAF6L were shown to bind on enhancers and promoters of ESC network genes such as *Oct4, Nanog, Klf4* and *c-Myc* [278], is in agreement with our finding that TADA2B occupies the enhancer and promoter of *NANOG*. Additionally, considering that

SMAD2 regulates transcription by interacting with chromatin remodelers to modulate chromatin environment [280], it would be interesting to identify whether TADA2B interacts with SMAD2 on enhancers of *NANOG* and other ESC factors for maintenance of pluripotency by performing immunoprecipitation (IP) experiments [280]. It is well known that the STAGA complex is related to H3 acetylation, an active gene transcription chromatin mark, and loss of STAGA members including proteins without HAT activity, results in a decrease on the global levels of histone H3K9 acetylation [228, 277, 281, 282]. This shows the interplay and dependence between structural subunits and HAT modules of the STAGA complex for its efficient catalytic activity on histones. It would be interesting to explore the genomic binding of TADA2B in hESCs and identify possible target genes, as well as correlate the finding with the TAF5L and TAF6L binding profile to uncover a common or distinct mechanism of action.

Moreover ChIP - seq experiments for active and repressed methylation marks such as H3K9ac, H3K27ac, H3K4me1 and H3K27me3 will unravel mechanisms of regulation of gene expression by TADA2B-mediated chromatin modification. ChIP or ChIP - seq for OSN factors will reveal how these changes in chromatin affect their binding on their target genes. Preliminary results indicate reduction of H3K27me3 on poised genes upon loss of TADA2B. This loss of repressive chromatin marks correlates with the increased expression of ME and TE genes (*GATA6, GSC, T, EOMES, CDX2*) in *TADA2B* knockout cells. This can possibly explain the differentiation phenotype of those cells. The observed phenotype and gene expression of those cells in combination with reduction of H3K27me3, resembles loss of c-MYC or PRC2 (*EED$^{-/-}$, SUZ12$^{-/-}$*) in hESCs [283-287]. c-MYC maintains self-renewal by recruiting HAT complexes on pluripotency factors [288] and repression of developmental genes by PRC2 recruitment on their promoters for deposition of H3K27me3 [287, 289]. Loss of MYC results in hypoacetylation due to inactivation of *Gcn5* [290], and downregulation of PRC2 complex members, leading to decreased H3K27me3 on bivalent genes and induction of primitive endoderm genes in mESCs (*Gata6, Sox17, Foxa2*) [287]. Since GCN5 and TADA3 belong to the MYC module in mESCs [288], and also GCN5 stabilizes MYC protein through acetylation [291], we can hypothesize that their loss deregulates the same subset of genes directly, or through modulation of each other's activities. If MYC is deregulated upon loss of TADA2B would be enlightening to study any cell cycle defects, considering MYC has a major role in control of the cell cycle.

GCN5 acts always in a complex [217] and deregulation of the complex affects its activity. We therefore hypothesize that loss of TADA2B dissociates the STAGA complex, mainly because TADA2B as a structural subunit of this complex, recognizes and binds chromatin via its SWIRM domain [217, 226, 292]. Through its SANT domain it interacts with GCN5 and recruits it to chromatin [217]. Indeed deletion of SANT domain of TADA2B leads to exclusion of GCN5 from STAGA complex [217]. TADA2B is not only an adaptor protein facilitating recruitment of GCN5 to chromatin, but is required for the catalytic activity of GCN5 [217]. Moreover, TADA2B interacts with TADA3 via its SANT domain, acting as a bridge between GCN5 and TADA3 [217], with the latter not being exclusive to STAGA complex but also member of the ATAC complex [293]; an essential complex for embryonic development [294]. In conclusion, TADA2B is essential for the integrity of the STAGA complex, but also probably for the interaction of the STAGA with the ATAC and with other complexes.

Figure 8-1. Proposed model for TADA2B function in hESCs. The proposed model describes how TADA2B links STAGA and ATAC complex for efficient activation of target genes by GNC5-mediated acetylation. Upon loss of TADA2B the STAGA complex is dissociated and GCN5 fails to be recruited to the chromatin to acetylate target genes, resulting in downregulation of pluripotency factors. Because of the close relationship of STAGA complex with MYC, dissociation of STAGA complex destabilizes MYC, conferring reduced MYC-mediated recruitment of PRC2 on poised genes. Loss of H3K27me3 de-represses developmental genes of ME and TE programs leading to differentiation of *TADA2B* KO hESCs.

Role of TADA2B in cancer

Similarities between MYC - associated transcription programs in ES and cancer cells has been previously described [288], and the connection of STAGA complex with the oncoprotein MYC, could indicate a role of *TADA2B* in cancer. Indeed, recent studies have implicated TADA2B in cancer, and more specifically,

its loss conferred resistance to Vemurafenib in melanoma cells [295]. It was also identified as an essential gene in Acute Myeloid Leukemia (AML), together with *KAT2A (GCN5), TAF5L* and *TAF6L* from STAGA complex, but its function was not addressed [296]. Instead the authors investigated the role of *KAT2A* in AML and demonstrated that inhibition of KAT2A with MB-3 could potentially have clinical applications [296].

The STAGA complex is involved in several processes and especially in development and cancer. Therefore, the dissection of the mechanism of action of this complex and identification of its partners could shed light on cancer research and embryonic development. Identification of cellular processes regulated by TADA2B in normal and pathological conditions could possibly show clinical advantage of targeting TADA2B instead of GCN5, which regulates many different transcription programs. Considering that chromatin modification is reversible, targeting chromatin regulators as a drug treatment could have clinical advantage over genetic modifications. Further research needs to be conducted to characterize this small protein and understand its potential in clinical settings.

Conclusions

In this study, we captured binding events in hESCs and differentiated cells on active and poised enhancers. We identified several proteins that occupy the *NANOG* active and poised *GSC* enhancers in hESCs and early differentiation. A role in embryo development is indicated by studies on some of the identified proteins [297, 298].

We described TADA2B as a novel protein binding on the enhancer of *NANOG* and its functional role in pluripotency. By generating for the first time knockout hESCs for TADA2B we uncovered an essential role of TADA2B in self-renewal that was not described before. TADA2B is implicated in hESCs self-renewal, erythropoiesis of hSPCs and cancer, therefore, its function should be further investigated.

9 Bibliography

1. Aristoteles and A.L. Peck, *Generation of animals*. [Rev. and reprinted] ed. The Loeb classical library. 1953, London: Heinemann. 607 S.
2. Rossant, J. and P.P. Tam, *Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse.* Development, 2009. 136(5): p. 701-13.
3. Saiz, N. and B. Plusa, *Early cell fate decisions in the mouse embryo.* Reproduction, 2013. 145(3): p. R65-80.
4. Wilson, V. and K.A. Lawson, *4 - Development of Early Embryonic Lineages*, in *Kaufman's Atlas of Mouse Development Supplement*, R. Baldock, et al., Editors. 2016, Academic Press: Boston. p. 65-76.
5. Barr, K.J., et al., *Contributions of Na+/H+ exchanger isoforms to preimplantation development of the mouse.* Mol Reprod Dev, 1998. 50(2): p. 146-53.
6. Tarkowski, A.K. and J. Wroblewska, *Development of blastomeres of mouse eggs isolated at the 4- and 8-cell stage.* J Embryol Exp Morphol, 1967. 18(1): p. 155-80.
7. Evans, M.J. and M.H. Kaufman, *Establishment in culture of pluripotential cells from mouse embryos.* Nature, 1981. 292(5819): p. 154-6.
8. Thomson, J.A., et al., *Embryonic stem cell lines derived from human blastocysts.* Science, 1998. 282(5391): p. 1145-7.
9. Arnold, S.J. and E.J. Robertson, *Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo.* Nat Rev Mol Cell Biol, 2009. 10(2): p. 91-103.
10. Manzanares, M. and T.A. Rodriguez, *Development: Hippo signalling turns the embryo inside out.* Curr Biol, 2013. 23(13): p. R559-61.
11. Gardner, R.L. and J. Rossant, *Investigation of the fate of 4-5 day post-coitum mouse inner cell mass cells by blastocyst injection.* J Embryol Exp Morphol, 1979. 52: p. 141-52.
12. Tam, P.P.L. and R.S.P. Beddington, *The Formation of Mesodermal Tissues in the Mouse Embryo during Gastrulation and Early Organogenesis.* Development, 1987. 99(1): p. 109-126.
13. Lawson, K.A., J.J. Meneses, and R.A. Pedersen, *Clonal analysis of epiblast fate during germ layer formation in the mouse embryo.* Development, 1991. 113(3): p. 891-911.
14. Tam, P.P., *Regionalisation of the mouse embryonic ectoderm: allocation of prospective ectodermal tissues during gastrulation.* Development, 1989. 107(1): p. 55-67.
15. Rossant, J. and P.P.L. Tam, *New Insights into Early Human Development: Lessons for Stem Cell Derivation and Differentiation.* Cell Stem Cell, 2017. 20(1): p. 18-28.
16. De Los Angeles, A., et al., *Hallmarks of pluripotency.* Nature, 2015. 525(7570): p. 469-78.
17. Przyborski, S.A., *Differentiation of human embryonic stem cells after transplantation in immune-deficient mice.* Stem Cells, 2005. 23(9): p. 1242-50.
18. Blum, B. and N. Benvenisty, *The tumorigenicity of human embryonic stem cells.* Adv Cancer Res, 2008. 100: p. 133-58.
19. International Stem Cell, I., et al., *Characterization of human embryonic stem cell lines by the International Stem Cell Initiative.* Nat Biotechnol, 2007. 25(7): p. 803-16.
20. Yu, J., et al., *Induced pluripotent stem cell lines derived from human somatic cells.* Science, 2007. 318(5858): p. 1917-20.
21. Takahashi, K., et al., *Induction of pluripotent stem cells from adult human fibroblasts by defined factors.* Cell, 2007. 131(5): p. 861-72.
22. Smith, A.G., *Embryo-derived stem cells: of mice and men.* Annu Rev Cell Dev Biol, 2001. 17: p. 435-62.
23. Martin, G.R., *Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells.* Proc Natl Acad Sci U S A, 1981. 78(12): p. 7634-8.
24. Solter, D., *From teratocarcinomas to embryonic stem cells and beyond: a history of embryonic stem cell research.* Nat Rev Genet, 2006. 7(4): p. 319-27.
25. Richards, M., et al., *Human feeders support prolonged undifferentiated growth of human inner cell masses and embryonic stem cells.* Nature Biotechnology, 2002. 20(9): p. 933-936.

26. Hovatta, O., et al., *A culture system using human foreskin fibroblasts as feeder cells allows production of human embryonic stem cells.* Hum Reprod, 2003. 18(7): p. 1404-9.

27. Klimanskaya, I., et al., *Human embryonic stem cells derived without feeder cells.* Lancet, 2005. 365(9471): p. 1636-41.

28. Tewary, M., N. Shakiba, and P.W. Zandstra, *Stem cell bioengineering: building from stem cell biology.* Nat Rev Genet, 2018. 19(10): p. 595-614.

29. Ludwig, T.E., et al., *Derivation of human embryonic stem cells in defined conditions.* Nat Biotechnol, 2006. 24(2): p. 185-7.

30. Ludwig, T.E., et al., *Feeder-independent culture of human embryonic stem cells.* Nat Methods, 2006. 3(8): p. 637-46.

31. Nichols, J. and A. Smith, *Naive and primed pluripotent states.* Cell Stem Cell, 2009. 4(6): p. 487-92.

32. Amit, M., et al., *Clonally derived human embryonic stem cell lines maintain pluripotency and proliferative potential for prolonged periods of culture.* Dev Biol, 2000. 227(2): p. 271-8.

33. Xu, R.H., et al., *BMP4 initiates human embryonic stem cell differentiation to trophoblast.* Nat Biotechnol, 2002. 20(12): p. 1261-4.

34. Vallier, L. and R.A. Pedersen, *Human embryonic stem cells: an in vitro model to study mechanisms controlling pluripotency in early mammalian development.* Stem Cell Rev, 2005. 1(2): p. 119-30.

35. Ruzo, A. and A.H. Brivanlou, *At Last: Gene Editing in Human Embryos to Understand Human Development.* Cell Stem Cell, 2017. 21(5): p. 564-565.

36. Fogarty, N.M.E., et al., *Genome editing reveals a role for OCT4 in human embryogenesis.* Nature, 2017. 550(7674): p. 67-73.

37. Frum, T., et al., *Oct4 cell-autonomously promotes primitive endoderm development in the mouse blastocyst.* Dev Cell, 2013. 25(6): p. 610-22.

38. Koike, H., et al., *Modelling human hepato-biliary-pancreatic organogenesis from the foregut-midgut boundary.* Nature, 2019.

39. Deglincerti, A., et al., *Self-organization of human embryonic stem cells on micropatterns.* Nat Protoc, 2016. 11(11): p. 2223-2232.

40. Wu, J. and J.C. Izpisua Belmonte, *Dynamic Pluripotent Stem Cell States and Their Applications.* Cell Stem Cell, 2015. 17(5): p. 509-25.

41. Trounson, A., *The production and directed differentiation of human embryonic stem cells.* Endocr Rev, 2006. 27(2): p. 208-19.

42. Canals, I., et al., *Rapid and efficient induction of functional astrocytes from human pluripotent stem cells.* Nat Methods, 2018. 15(9): p. 693-696.

43. Rezania, A., et al., *Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells.* Nat Biotechnol, 2014. 32(11): p. 1121-33.

44. Mills, J.A., et al., *Hematopoietic differentiation of pluripotent stem cells in culture.* Methods Mol Biol, 2014. 1185: p. 181-94.

45. Ma, F., et al., *Novel method for efficient production of multipotential hematopoietic progenitors from human embryonic stem cells.* Int J Hematol, 2007. 85(5): p. 371-9.

46. Aikawa, N., Y. Suzuki, and K. Takaba, *A Simple Protocol for the Myocardial Differentiation of Human iPS Cells.* Biol Pharm Bull, 2015. 38(7): p. 1070-5.

47. McQuade, A., et al., *Development and validation of a simplified method to generate human microglia from pluripotent stem cells.* Mol Neurodegener, 2018. 13(1): p. 67.

48. Lungova, V., et al., *Human induced pluripotent stem cell-derived vocal fold mucosa mimics development and responses to smoke exposure.* Nat Commun, 2019. 10(1): p. 4161.

49. Wang, S., et al., *Human ESC-derived expandable hepatic organoids enable therapeutic liver repopulation and pathophysiological modeling of alcoholic liver injury.* Cell Res, 2019.

50. Luo, C., et al., *Cerebral Organoids Recapitulate Epigenomic Signatures of the Human Fetal Brain.* Cell Rep, 2016. 17(12): p. 3369-3384.

51. Yoney, A., et al., *WNT signaling memory is required for ACTIVIN to function as a morphogen in humam gastruloids.* Elife, 2018. 7.

52. Martyn, I., et al., *Self-organization of a human organizer by combined Wnt and Nodal signalling (vol 558, pg 132, 2018).* Nature, 2018. 564(7735): p. E10-E10.

53. Tam, P.P.L. and D.A.F. Loebel, *Gene function in mouse embryogenesis: get set for gastrulation.* Nature Reviews Genetics, 2007. 8(5): p. 368-381.
54. Ben-Haim, N., et al., *The nodal precursor acting via activin receptors induces mesoderm by maintaining a source of its convertases and BMP4.* Developmental Cell, 2006. 11(3): p. 313-323.
55. Winnier, G., et al., *Bone Morphogenetic Protein-4 Is Required for Mesoderm Formation and Patterning in the Mouse.* Genes & Development, 1995. 9(17): p. 2105-2116.
56. Lavon, N., O. Yanuka, and N. Benvenisty, *Differentiation and isolation of hepatic-like cells from human embryonic stem cells.* Differentiation, 2004. 72(5): p. 230-8.
57. Loh, K.M., et al., *Efficient endoderm induction from human pluripotent stem cells by logically directing signals controlling lineage bifurcations.* Cell Stem Cell, 2014. 14(2): p. 237-52.
58. Loh, K.M., et al., *Mapping the Pairwise Choices Leading from Pluripotency to Human Bone, Heart, and Other Mesoderm Cell Types.* Cell, 2016. 166(2): p. 451-67.
59. D'Amour, K.A., et al., *Efficient differentiation of human embryonic stem cells to definitive endoderm.* Nat Biotechnol, 2005. 23(12): p. 1534-41.
60. Huber, I., et al., *Identification and selection of cardiomyocytes during human embryonic stem cell differentiation.* Faseb Journal, 2007. 21(10): p. 2551-2563.
61. Klein, A.M., et al., *Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells.* Cell, 2015. 161(5): p. 1187-1201.
62. Kester, L. and A. van Oudenaarden, *Single-Cell Transcriptomics Meets Lineage Tracing.* Cell Stem Cell, 2018. 23(2): p. 166-179.
63. Yao, Z., et al., *A Single-Cell Roadmap of Lineage Bifurcation in Human ESC Models of Embryonic Brain Development.* Cell Stem Cell, 2017. 20(1): p. 120-134.
64. Yan, L., et al., *Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells.* Nat Struct Mol Biol, 2013. 20(9): p. 1131-9.
65. Petropoulos, S., et al., *Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos.* Cell, 2016. 167(1): p. 285.
66. Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.* Cell, 2006. 126(4): p. 663-76.
67. Shi, Y., et al., *Induced pluripotent stem cell technology: a decade of progress.* Nat Rev Drug Discov, 2017. 16(2): p. 115-130.
68. Takahashi, K., et al., *Induction of pluripotent stem cells from fibroblast cultures.* Nat Protoc, 2007. 2(12): p. 3081-9.
69. Taylor, C.J., et al., *Generating an iPSC bank for HLA-matched tissue transplantation based on known donor and recipient HLA types.* Cell Stem Cell, 2012. 11(2): p. 147-52.
70. Bellin, M., et al., *Induced pluripotent stem cells: the new patient?* Nat Rev Mol Cell Biol, 2012. 13(11): p. 713-26.
71. Murry, C.E. and G. Keller, *Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development.* Cell, 2008. 132(4): p. 661-80.
72. Garcez, P.P., et al., *Zika virus impairs growth in human neurospheres and brain organoids.* Science, 2016. 352(6287): p. 816-8.
73. Gabriel, E., et al., *CPAP promotes timely cilium disassembly to maintain neural progenitor pool.* EMBO J, 2016. 35(8): p. 803-19.
74. Kondo, T., et al., *Modeling Alzheimer's disease with iPSCs reveals stress phenotypes associated with intracellular Abeta and differential drug responsiveness.* Cell Stem Cell, 2013. 12(4): p. 487-96.
75. Schwartz, S.D., et al., *Embryonic stem cell trials for macular degeneration: a preliminary report.* Lancet, 2012. 379(9817): p. 713-20.
76. Menasche, P., et al., *Human embryonic stem cell-derived cardiac progenitors for severe heart failure treatment: first clinical case report.* Eur Heart J, 2015. 36(30): p. 2011-7.
77. Lee, A.S., et al., *Tumorigenicity as a clinical hurdle for pluripotent stem cell therapies.* Nature Medicine, 2013. 19(8): p. 998-1004.
78. Loh, K.M., B. Lim, and L.T. Ang, *Ex uno plures: molecular designs for embryonic pluripotency.* Physiol Rev, 2015. 95(1): p. 245-95.
79. Silva, J., et al., *Nanog Is the Gateway to the Pluripotent Ground State.* Cell, 2009. 138(4): p. 722-737.
80. Chambers, I. and A. Smith, *Self-renewal of teratocarcinoma and embryonic stem cells.* Oncogene, 2004. 23(43): p. 7150-60.

81. Niwa, H., *How is pluripotency determined and maintained?* Development, 2007. 134(4): p. 635-46.
82. Silva, J. and A. Smith, *Capturing pluripotency.* Cell, 2008. 132(4): p. 532-6.
83. Nichols, J., et al., *Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4.* Cell, 1998. 95(3): p. 379-91.
84. Mitsui, K., et al., *The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells.* Cell, 2003. 113(5): p. 631-642.
85. Avilion, A.A., et al., *Multipotent cell lineages in early mouse development depend on SOX2 function.* Genes & Development, 2003. 17(1): p. 126-140.
86. Chambers, I., et al., *Nanog safeguards pluripotency and mediates germline development.* Nature, 2007. 450(7173): p. 1230-U8.
87. Babaie, Y., et al., *Analysis of Oct4-dependent transcriptional networks regulating self-renewal and pluripotency in human embryonic stem cells.* Stem Cells, 2007. 25(2): p. 500-10.
88. Loh, K.M. and B. Lim, *A Precarious Balance: Pluripotency Factors as Lineage Specifiers.* Cell Stem Cell, 2011. 8(4): p. 363-369.
89. Niwa, H., J. Miyazaki, and A.G. Smith, *Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells.* Nat Genet, 2000. 24(4): p. 372-6.
90. Teo, A.K., et al., *Pluripotency factors regulate definitive endoderm specification through eomesodermin.* Genes Dev, 2011. 25(3): p. 238-50.
91. Kopp, J.L., et al., *Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells.* Stem Cells, 2008. 26(4): p. 903-11.
92. Wang, Z., et al., *Distinct Lineage Specification Roles for NANOG, OCT4, and SOX2 in Human Embryonic Stem Cells.* Cell Stem Cell, 2012. 10(4): p. 440-454.
93. Thomson, M., et al., *Pluripotency factors in embryonic stem cells regulate differentiation into germ layers.* Cell, 2011. 145(6): p. 875-89.
94. Greber, B., H. Lehrach, and J. Adjaye, *Control of early fate decisions in human ES cells by distinct states of TGFbeta pathway activity.* Stem Cells Dev, 2008. 17(6): p. 1065-77.
95. Zeineddine, D., et al., *Oct-3/4 dose dependently regulates specification of embryonic stem cells toward a cardiac lineage and early heart development.* Dev Cell, 2006. 11(4): p. 535-46.
96. Boyer, L.A., et al., *Core transcriptional regulatory circuitry in human embryonic stem cells.* Cell, 2005. 122(6): p. 947-56.
97. Loh, Y.H., et al., *The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells.* Nat Genet, 2006. 38(4): p. 431-40.
98. Boyer, L.A., et al., *Polycomb complexes repress developmental regulators in murine embryonic stem cells.* Nature, 2006. 441(7091): p. 349-53.
99. Chen, X., et al., *Integration of external signaling pathways with the core transcriptional network in embryonic stem cells.* Cell, 2008. 133(6): p. 1106-17.
100. Kim, J., et al., *An extended transcriptional network for pluripotency of embryonic stem cells.* Cell, 2008. 132(6): p. 1049-61.
101. Pasini, D., et al., *JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells.* Nature, 2010. 464(7286): p. 306-10.
102. Beyer, T.A., et al., *Switch enhancers interpret TGF-beta and Hippo signaling to control cell fate in human embryonic stem cells.* Cell Rep, 2013. 5(6): p. 1611-24.
103. Liang, J., et al., *Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells.* Nat Cell Biol, 2008. 10(6): p. 731-9.
104. Vallier, L., M. Alexander, and R.A. Pedersen, *Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells.* J Cell Sci, 2005. 118(Pt 19): p. 4495-509.
105. Levenstein, M.E., et al., *Basic fibroblast growth factor support of human embryonic stem cell self-renewal.* Stem Cells, 2006. 24(3): p. 568-74.
106. Greber, B., H. Lehrach, and J. Adjaye, *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, 2007. 25(2): p. 455-64.
107. Besser, D., *Expression of nodal, lefty-a, and lefty-B in undifferentiated human embryonic stem cells requires activation of Smad2/3.* J Biol Chem, 2004. 279(43): p. 45076-84.
108. Beattie, G.M., et al., *Activin A maintains pluripotency of human embryonic stem cells in the absence of feeder layers.* Stem Cells, 2005. 23(4): p. 489-95.

109.	Sui, L., L. Bouwens, and J.K. Mfopou, *Signaling pathways during maintenance and definitive endoderm differentiation of embryonic stem cells.* Int J Dev Biol, 2013. 57(1): p. 1-12.

110.	Willems, E. and L. Leyns, *Patterning of mouse embryonic stem cell-derived pan-mesoderm by Activin A/Nodal and Bmp4 signaling requires Fibroblast Growth Factor activity.* Differentiation, 2008. 76(7): p. 745-59.

111.	Smith, J.R., et al., *Inhibition of Activin/Nodal signaling promotes specification of human embryonic stem cells into neuroectoderm.* Dev Biol, 2008. 313(1): p. 107-17.

112.	Chambers, S.M., et al., *Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling.* Nat Biotechnol, 2009. 27(3): p. 275-80.

113.	Camus, A., et al., *Absence of Nodal signaling promotes precocious neural differentiation in the mouse embryo.* Dev Biol, 2006. 295(2): p. 743-55.

114.	Brennan, J., et al., *Nodal signalling in the epiblast patterns the early mouse embryo.* Nature, 2001. 411(6840): p. 965-9.

115.	Mesnard, D., M. Guzman-Ayala, and D.B. Constam, *Nodal specifies embryonic visceral endoderm and sustains pluripotent cells in the epiblast before overt axial patterning.* Development, 2006. 133(13): p. 2497-505.

116.	Vallier, L., M. Alexander, and R.A. Pedersen, *Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells.* Journal of Cell Science, 2005. 118(19): p. 4495-4509.

117.	Xu, R.H., et al., *NANOG is a direct target of TGF beta/Activin-mediated SMAD signaling in human ESCs.* Cell Stem Cell, 2008. 3(2): p. 196-206.

118.	Zoumaro-Djayoon, A.D., et al., *Investigating the role of FGF-2 in stem cell maintenance by global phosphoproteomics profiling.* Proteomics, 2011. 11(20): p. 3962-71.

119.	Xiao, L., X. Yuan, and S.J. Sharkis, *Activin A maintains self-renewal and regulates fibroblast growth factor, Wnt, and bone morphogenic protein pathways in human embryonic stem cells.* Stem Cells, 2006. 24(6): p. 1476-86.

120.	Mullen, A.C., et al., *Master transcription factors determine cell-type-specific responses to TGF-beta signaling.* Cell, 2011. 147(3): p. 565-76.

121.	Brown, S., et al., *Activin/Nodal signaling controls divergent transcriptional networks in human embryonic stem cells and in endoderm progenitors.* Stem Cells, 2011. 29(8): p. 1176-85.

122.	Singh, A.M., et al., *Signaling Network Crosstalk in Human Pluripotent Cells: A Smad2/3-Regulated Switch that Controls the Balance between Self-Renewal and Differentiation.* Cell Stem Cell, 2012. 10(3): p. 312-326.

123.	Ogawa, K., et al., *Synergistic action of Wnt and LIF in maintaining pluripotency of mouse ES cells.* Biochem Biophys Res Commun, 2006. 343(1): p. 159-66.

124.	Hao, J., et al., *WNT/beta-catenin pathway up-regulates Stat3 and converges on LIF to prevent differentiation of mouse embryonic stem cells.* Dev Biol, 2006. 290(1): p. 81-91.

125.	Kelly, K.F., et al., *beta-catenin enhances Oct-4 activity and reinforces pluripotency through a TCF-independent mechanism.* Cell Stem Cell, 2011. 8(2): p. 214-27.

126.	Huelsken, J., et al., *Requirement for beta-catenin in anterior-posterior axis formation in mice.* J Cell Biol, 2000. 148(3): p. 567-78.

127.	Lyashenko, N., et al., *Differential requirement for the dual functions of beta-catenin in embryonic stem cell self-renewal and germ layer formation.* Nat Cell Biol, 2011. 13(7): p. 753-61.

128.	Varelas, X., et al., *TAZ controls Smad nucleocytoplasmic shuttling and regulates human embryonic stem-cell self-renewal.* Nat Cell Biol, 2008. 10(7): p. 837-48.

129.	Alarcon, C., et al., *Nuclear CDKs drive Smad transcriptional activation and turnover in BMP and TGF-beta pathways.* Cell, 2009. 139(4): p. 757-69.

130.	Conlon, F.L., et al., *A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse.* Development, 1994. 120(7): p. 1919-28.

131.	Hogan, B.L., *Bone morphogenetic proteins in development.* Curr Opin Genet Dev, 1996. 6(4): p. 432-8.

132.	Yamaguchi, T.P., *Heads or tails: Wnts and anterior-posterior patterning.* Curr Biol, 2001. 11(17): p. R713-24.

133.	Chial, H., *Gene-based therapeutic approaches.* Nature Education, 2008. 1(1):210.

134.	Ciruna, B. and J. Rossant, *FGF signaling regulates mesoderm cell fate specification and morphogenetic movement at the primitive streak.* Developmental Cell, 2001. 1(1): p. 37-49.

135. Cano, A., et al., *The transcription factor Snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression.* Nature Cell Biology, 2000. 2(2): p. 76-83.

136. Batlle, E., et al., *The transcription factor Snail is a repressor of E-cadherin gene expression in epithelial tumour cells.* Nature Cell Biology, 2000. 2(2): p. 84-89.

137. Norris, D.P., et al., *The Foxh1-dependent autoregulatory enhancer controls the level of Nodal signals in the mouse embryo.* Development, 2002. 129(14): p. 3455-68.

138. Izzi, L., et al., *Foxh1 recruits Gsc to negatively regulate Mixl1 expression during early mouse development.* EMBO J, 2007. 26(13): p. 3132-43.

139. Yamamoto, M., et al., *The transcription factor FoxH1 (FAST) mediates Nodal signaling during anterior-posterior patterning and node formation in the mouse.* Genes Dev, 2001. 15(10): p. 1242-56.

140. Heasman, J., *Patterning the early Xenopus embryo.* Development, 2006. 133(7): p. 1205-17.

141. Blum, M., S.J. Gaunt, and E.M. Derobertis, *Gastrulation in the Mouse - the Role of the Homeobox Gene Goosecoid.* Molecular Biology of the Cell, 1992. 3: p. A98-A98.

142. Warmflash, A., et al., *A method to recapitulate early embryonic spatial patterning in human embryonic stem cells.* Nat Methods, 2014. 11(8): p. 847-54.

143. Martyn, I., et al., *Self-organization of a human organizer by combined Wnt and Nodal signalling.* Nature, 2018. 558(7708): p. 132-135.

144. Bernardo, A.S., et al., *BRACHYURY and CDX2 mediate BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and extraembryonic lineages.* Cell Stem Cell, 2011. 9(2): p. 144-55.

145. Brown, S., et al., *Activin/Nodal Signaling Controls Divergent Transcriptional Networks in Human Embryonic Stem Cells and in Endoderm Progenitors.* Stem Cells, 2011. 29(8): p. 1176-1185.

146. Arnold, S.J., et al., *Pivotal roles for eomesodermin during axis formation, epithelium-to-mesenchyme transition and endoderm specification in the mouse.* Development, 2008. 135(3): p. 501-511.

147. Yamamoto, M., et al., *The transcription factor FoxH1 (FAST) mediates Nodal signaling during anterior-posterior patterning and node formation in the mouse.* Genes & Development, 2001. 15(10): p. 1242-1256.

148. Hoodless, P.A., et al., *FoxH1 (Fast) functions to specify the anterior primitive streak in the mouse.* Genes & Development, 2001. 15(10): p. 1257-1271.

149. Chu, G.C., et al., *Differential requirements for Smad4 in TGF beta-dependent patterning of the early mouse embryo.* Development, 2004. 131(15): p. 3501-3512.

150. Tremblay, K.D., et al., *Formation of the definitive endoderm in mouse is a Smad2-dependent process.* Development, 2000. 127(14): p. 3079-3090.

151. Chu, G.C., et al., *Differential requirements for Smad4 in TGFbeta-dependent patterning of the early mouse embryo.* Development, 2004. 131(15): p. 3501-12.

152. Perea-Gomez, A., et al., *Nodal antagonists in the anterior visceral endoderm prevent the formation of multiple primitive streaks.* Dev Cell, 2002. 3(5): p. 745-56.

153. Dunn, N.R., et al., *Combinatorial activities of Smad2 and Smad3 regulate mesoderm formation and patterning in the mouse embryo.* Development, 2004. 131(8): p. 1717-28.

154. Schmierer, B. and C.S. Hill, *TGFbeta-SMAD signal transduction: molecular specificity and functional flexibility.* Nat Rev Mol Cell Biol, 2007. 8(12): p. 970-82.

155. Chiu, W.T., et al., *Genome-wide view of TGFbeta/Foxh1 regulation of the early mesendoderm program.* Development, 2014. 141(23): p. 4537-47.

156. Labbe, E., et al., *Smad2 and Smad3 positively and negatively regulate TGF beta-dependent transcription through the forkhead DNA-binding protein FAST2.* Mol Cell, 1998. 2(1): p. 109-20.

157. Hoodless, P.A., et al., *FoxH1 (Fast) functions to specify the anterior primitive streak in the mouse.* Genes Dev, 2001. 15(10): p. 1257-71.

158. Young, R.A., *Control of the embryonic stem cell state.* Cell, 2011. 144(6): p. 940-54.

159. Kato, Y., et al., *A component of the ARC/Mediator complex required for TGF beta/Nodal signalling.* Nature, 2002. 418(6898): p. 641-6.

160. Ross, S. and C.S. Hill, *How the Smads regulate transcription.* Int J Biochem Cell Biol, 2008. 40(3): p. 383-408.

161. Feng, X.H., et al., *The tumor suppressor Smad4/DPC4 and transcriptional adaptor CBP/p300 are coactivators for smad3 in TGF-beta-induced transcriptional activation.* Genes Dev, 1998. 12(14): p. 2153-63.

162. Janknecht, R., N.J. Wells, and T. Hunter, *TGF-beta-stimulated cooperation of smad proteins with the coactivators CBP/p300.* Genes Dev, 1998. 12(14): p. 2114-9.

163. Zhou, V.W., A. Goren, and B.E. Bernstein, *Charting histone modifications and the functional organization of mammalian genomes.* Nat Rev Genet, 2011. 12(1): p. 7-18.

164. Calo, E. and J. Wysocka, *Modification of Enhancer Chromatin: What, How, and Why?* Molecular Cell, 2013. 49(5): p. 825-837.

165. Voigt, P., W.W. Tee, and D. Reinberg, *A double take on bivalent promoters.* Genes Dev, 2013. 27(12): p. 1318-38.

166. Bernstein, B.E., et al., *A bivalent chromatin structure marks key developmental genes in embryonic stem cells.* Cell, 2006. 125(2): p. 315-26.

167. Azuara, V., et al., *Chromatin signatures of pluripotent cell lines.* Nat Cell Biol, 2006. 8(5): p. 532-8.

168. Rada-Iglesias, A., et al., *A unique chromatin signature uncovers early developmental enhancers in humans.* Nature, 2011. 470(7333): p. 279-83.

169. Zentner, G.E., P.J. Tesar, and P.C. Scacheri, *Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions.* Genome Research, 2011. 21(8): p. 1273-1283.

170. Bonn, S., et al., *Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development.* Nat Genet, 2012. 44(2): p. 148-56.

171. Ernst, J., et al., *Mapping and analysis of chromatin state dynamics in nine human cell types.* Nature, 2011. 473(7345): p. 43-9.

172. van Steensel, B. and J. Dekker, *Genomics tools for unraveling chromosome architecture.* Nat Biotechnol, 2010. 28(10): p. 1089-95.

173. Zinzen, R.P., et al., *Combinatorial binding predicts spatio-temporal cis-regulatory activity.* Nature, 2009. 462(7269): p. 65-U72.

174. May, D., et al., *Large-scale discovery of enhancers from human heart tissue.* Nature Genetics, 2012. 44(1): p. 89-U125.

175. Ram, O., et al., *Combinatorial Patterning of Chromatin Regulators Uncovered by Genome-wide Location Analysis in Human Cells.* Cell, 2011. 147(7): p. 1628-1639.

176. Krebs, A.R., et al., *SAGA and ATAC Histone Acetyl Transferase Complexes Regulate Distinct Sets of Genes and ATAC Defines a Class of p300-Independent Enhancers.* Molecular Cell, 2011. 44(3): p. 410-423.

177. Song, L., et al., *Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity.* Genome Res, 2011. 21(10): p. 1757-67.

178. Giresi, P.G., et al., *FAIRE ((F)under-barormaldehyde-(A)under-barssisted (I)under-barsolation of (R)under-baregulatory (E)under-barlements) isolates active regulatory elements from human chromatin.* Genome Research, 2007. 17(6): p. 877-885.

179. Maniatis, T., et al., *Structure and function of the interferon-beta enhanceosome.* Cold Spring Harbor Symposia on Quantitative Biology, 1998. 63: p. 609-620.

180. Whyte, W.A., et al., *Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes.* Cell, 2013. 153(2): p. 307-319.

181. Jin, F., et al., *A high-resolution map of the three-dimensional chromatin interactome in human cells.* Nature, 2013. 503(7475): p. 290-4.

182. Zhang, Y., et al., *Chromatin connectivity maps reveal dynamic promoter-enhancer long-range associations.* Nature, 2013. 504(7479): p. 306-310.

183. Fujita, J., et al., *Caspase activity mediates the differentiation of embryonic stem cells.* Cell Stem Cell, 2008. 2(6): p. 595-601.

184. Tay, Y., et al., *MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation.* Nature, 2008. 455(7216): p. 1124-8.

185. Feldman, N., et al., *G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis.* Nat Cell Biol, 2006. 8(2): p. 188-94.

186. Whyte, W.A., et al., *Enhancer decommissioning by LSD1 during embryonic stem cell differentiation.* Nature, 2012. 482(7384): p. 221-5.

187. Cloos, P.A., et al., *Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease.* Genes Dev, 2008. 22(9): p. 1115-40.

188. Schmitz, S.U., et al., *Jarid1b targets genes regulating development and is involved in neural differentiation.* Embo Journal, 2011. 30(22): p. 4586-4600.

189. Pasini, D., et al., *The polycomb group protein Suz12 is required for embryonic stem cell differentiation.* Molecular and Cellular Biology, 2007. 27(10): p. 3769-3779.

190. Jackson, M., et al., *Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells.* Molecular and Cellular Biology, 2004. 24(20): p. 8862-8871.

191. Lei, H., et al., *De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells.* Development, 1996. 122(10): p. 3195-3205.

192. O'Carroll, D., et al., *The polycomb-group gene Ezh2 is required for early mouse development.* Mol Cell Biol, 2001. 21(13): p. 4330-6.

193. Ura, H., et al., *Eed/Sox2 regulatory loop controls ES cell self-renewal through histone methylation and acetylation.* EMBO J, 2011. 30(11): p. 2190-204.

194. Dahle, O., A. Kumar, and M.R. Kuehn, *Nodal signaling recruits the histone demethylase Jmjd3 to counteract polycomb-mediated repression at target genes.* Sci Signal, 2010. 3(127): p. ra48.

195. Kartikasari, A.E., et al., *The histone demethylase Jmjd3 sequentially associates with the transcription factors Tbx3 and Eomes to drive endoderm differentiation.* EMBO J, 2013. 32(10): p. 1393-408.

196. Laugesen, A. and K. Helin, *Chromatin repressive complexes in stem cells, development, and cancer.* Cell Stem Cell, 2014. 14(6): p. 735-51.

197. Di Croce, L. and K. Helin, *Transcriptional regulation by Polycomb group proteins.* Nat Struct Mol Biol, 2013. 20(10): p. 1147-55.

198. Margueron, R. and D. Reinberg, *The Polycomb complex PRC2 and its mark in life.* Nature, 2011. 469(7330): p. 343-9.

199. Pengelly, A.R., et al., *A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb.* Science, 2013. 339(6120): p. 698-9.

200. Francis, N.J., R.E. Kingston, and C.L. Woodcock, *Chromatin compaction by a polycomb group protein complex.* Science, 2004. 306(5701): p. 1574-7.

201. Hayakawa, T. and J. Nakayama, *Physiological roles of class I HDAC complex and histone demethylase.* J Biomed Biotechnol, 2011. 2011: p. 129383.

202. Gaspar-Maia, A., et al., *Chd1 regulates open chromatin and pluripotency of embryonic stem cells.* Nature, 2009. 460(7257): p. 863-8.

203. Ho, L., et al., *An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency.* Proc Natl Acad Sci U S A, 2009. 106(13): p. 5181-6.

204. Kidder, B.L., S. Palmer, and J.G. Knott, *SWI/SNF-Brg1 regulates self-renewal and occupies core pluripotency-related genes in embryonic stem cells.* Stem Cells, 2009. 27(2): p. 317-28.

205. Ho, L. and G.R. Crabtree, *Chromatin remodelling during development.* Nature, 2010. 463(7280): p. 474-84.

206. Gao, F., et al., *The interactomes of POU5F1 and SOX2 enhancers in human embryonic stem cells.* Sci Rep, 2013. 3: p. 1588.

207. Agger, K., et al., *UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development.* Nature, 2007. 449(7163): p. 731-4.

208. Bilodeau, S., et al., *SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state.* Genes & Development, 2009. 23(21): p. 2484-2489.

209. Xi, Q., et al., *A poised chromatin platform for TGF-beta access to master regulators.* Cell, 2011. 147(7): p. 1511-24.

210. Nagy, Z. and L. Tora, *Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation.* Oncogene, 2007. 26(37): p. 5341-5357.

211. Sterner, D.E. and S.L. Berger, *Acetylation of histones and transcription-related factors.* Microbiology and Molecular Biology Reviews, 2000. 64(2): p. 435-+.

212. Brownell, J.E. and C.D. Allis, *Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation.* Curr Opin Genet Dev, 1996. 6(2): p. 176-84.

213. Candau, R., et al., *Identification of human proteins functionally conserved with the yeast putative adaptors ADA2 and GCN5.* Mol Cell Biol, 1996. 16(2): p. 593-602.

214. Lin, W., et al., *Developmental potential of Gcn5(-/-) embryonic stem cells in vivo and in vitro.* Dev Dyn, 2007. 236(6): p. 1547-57.

215. Wang, L., et al., *GCN5 Regulates FGF Signaling and Activates Selective MYC Target Genes during Early Embryoid Body Differentiation.* Stem Cell Reports, 2018. 10(1): p. 287-299.

216. Ogryzko, V.V., *Hot papers - Genetics - The transcriptional coactivators p300 and CBP are histone acetyltransferases by V.V. Ogryzko, R.L. Schiltz, V. Russanova, B.H. Howard, Y. Nakatani - Comments.* Scientist, 1998. 12(9): p. 11-11.

217. Gamper, A.M., J. Kim, and R.G. Roeder, *The STAGA Subunit ADA2b Is an Important Regulator of Human GCN5 Catalysis.* Molecular and Cellular Biology, 2009. 29(1): p. 266-280.

218. Krebs, A.R., et al., *SAGA and ATAC histone acetyl transferase complexes regulate distinct sets of genes and ATAC defines a class of p300-independent enhancers.* Mol Cell, 2011. 44(3): p. 410-423.

219. Baker, S.P. and P.A. Grant, *The SAGA continues: expanding the cellular role of a transcriptional co-activator complex.* Oncogene, 2007. 26(37): p. 5329-40.

220. Riss, A., et al., *Subunits of ADA-two-A-containing (ATAC) or Spt-Ada-Gcn5-acetyltrasferase (SAGA) Coactivator Complexes Enhance the Acetyltransferase Activity of GCN5.* J Biol Chem, 2015. 290(48): p. 28997-9009.

221. Grant, P.A., et al., *Expanded lysine acetylation specificity of Gcn5 in native complexes.* Journal of Biological Chemistry, 1999. 274(9): p. 5895-5900.

222. Koutelou, E., C.L. Hirsch, and S.Y. Dent, *Multiple faces of the SAGA complex.* Curr Opin Cell Biol, 2010. 22(3): p. 374-82.

223. Pan, J., et al., *Interrogation of Mammalian Protein Complex Structure, Function, and Membership Using Genome-Scale Fitness Screens.* Cell Systems, 2018. 6(5): p. 555-+.

224. Kusch, T., et al., *Two Drosophila Ada2 homologues function in different multiprotein complexes.* Mol Cell Biol, 2003. 23(9): p. 3305-19.

225. Spedale, G., H.T. Timmers, and W.W. Pijnappel, *ATAC-king the complexity of SAGA during evolution.* Genes Dev, 2012. 26(6): p. 527-41.

226. Qian, C., et al., *Structure and chromosomal DNA binding of the SWIRM domain.* Nat Struct Mol Biol, 2005. 12(12): p. 1078-85.

227. Muratoglu, S., et al., *Two different Drosophila ADA2 homologues are present in distinct GCN5 histone acetyltransferase-containing complexes.* Mol Cell Biol, 2003. 23(1): p. 306-21.

228. Gamper, A.M., J. Kim, and R.G. Roeder, *The STAGA subunit ADA2b is an important regulator of human GCN5 catalysis.* Mol Cell Biol, 2009. 29(1): p. 266-80.

229. Weake, V.M., et al., *Post-transcription initiation function of the ubiquitous SAGA complex in tissue-specific gene activation.* Genes Dev, 2011. 25(14): p. 1499-509.

230. Mair, B., et al., *Essential Gene Profiles for Human Pluripotent Stem Cells Identify Uncharacterized Genes and Substrate Dependencies.* Cell Rep, 2019. 27(2): p. 599-615 e12.

231. Ihry, R.J., et al., *Genome-Scale CRISPR Screens Identify Human Pluripotency-Specific Genes.* Cell Rep, 2019. 27(2): p. 616-630 e6.

232. Bertero, A., et al., *The SMAD2/3 interactome reveals that TGFbeta controls m(6)A mRNA methylation in pluripotency.* Nature, 2018. 555(7695): p. 256-259.

233. Daniel A. Kuppers, S.A., Yiting Lim, Andrea Lim, Lucas Carter, Philip Corrin, Christopher L. Plaisier, Ryan Basom, Jeffrey J. Delrow, Shiyan Wang, Housheng Hansen He, Beverly Torok-Storb, Andrew C. Hsieh, Patrick J. Paddison, *N6-methyladenosine mRNA marking promotes selective translation of regulons required for human erythropoiesis.* 2018.

234. le Sage, C., et al., *Dual direction CRISPR transcriptional regulation screening uncovers gene networks driving drug resistance.* Sci Rep, 2017. 7(1): p. 17693.

235. Shalem, O., et al., *Genome-scale CRISPR-Cas9 knockout screening in human cells.* Science, 2014. 343(6166): p. 84-87.

236. Delaval, K. and R. Feil, *Epigenetic regulation of mammalian genomic imprinting.* Curr Opin Genet Dev, 2004. 14(2): p. 188-95.

237. Gaarenstroom, T. and C.S. Hill, *TGF-beta signaling to chromatin: How Smads regulate transcription during self-renewal and differentiation.* Seminars in Cell & Developmental Biology, 2014. 32: p. 107-118.

238. Spitz, F. and E.E. Furlong, *Transcription factors: from enhancer binding to developmental control.* Nat Rev Genet, 2012. 13(9): p. 613-26.
239. Buecker, C. and J. Wysocka, *Enhancers as information integration hubs in development: lessons from genomics.* Trends in Genetics, 2012. 28(6): p. 276-284.
240. Uchikawa, M., et al., *Functional analysis of chicken Sox2 enhancers highlights an array of diverse regulatory elements that are conserved in mammals.* Dev Cell, 2003. 4(4): p. 509-19.
241. Fisher, C.L. and A.G. Fisher, *Chromatin states in pluripotent, differentiated, and reprogrammed cells.* Curr Opin Genet Dev, 2011. 21(2): p. 140-6.
242. Levine, M. and R. Tjian, *Transcription regulation and animal diversity.* Nature, 2003. 424(6945): p. 147-51.
243. Vaquerizas, J.M., et al., *A census of human transcription factors: function, expression and evolution.* Nat Rev Genet, 2009. 10(4): p. 252-63.
244. Li, B., M. Carey, and J.L. Workman, *The role of chromatin during transcription.* Cell, 2007. 128(4): p. 707-19.
245. Fuda, N.J., M.B. Ardehali, and J.T. Lis, *Defining mechanisms that regulate RNA polymerase II transcription in vivo.* Nature, 2009. 461(7261): p. 186-92.
246. Kagey, M.H., et al., *Mediator and cohesin connect gene expression and chromatin architecture.* Nature, 2010. 467(7314): p. 430-5.
247. Lee, K., et al., *FOXA2 Is Required for Enhancer Priming during Pancreatic Differentiation.* Cell Reports, 2019. 28(2): p. 382-+.
248. Cirillo, L.A., et al., *Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4.* Mol Cell, 2002. 9(2): p. 279-89.
249. King, H.W. and R.J. Klose, *The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support gene regulatory element function in mouse embryonic stem cells.* Elife, 2017. 6.
250. Renz, P.F. and T.A. Beyer, *A Concise Protocol for siRNA-Mediated Gene Suppression in Human Embryonic Stem Cells.* Methods Mol Biol, 2016. 1341: p. 369-76.
251. Mandegar, M.A., et al., *CRISPR Interference Efficiently Induces Specific and Reversible Gene Silencing in Human iPSCs.* Cell Stem Cell, 2016. 18(4): p. 541-53.
252. Gilbert, L.A., et al., *Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation.* Cell, 2014. 159(3): p. 647-61.
253. Chen, G., et al., *Chemically defined conditions for human iPSC derivation and culture.* Nat Methods, 2011. 8(5): p. 424-9.
254. Xu, R.H., et al., *NANOG is a direct target of TGFbeta/activin-mediated SMAD signaling in human ESCs.* Cell Stem Cell, 2008. 3(2): p. 196-206.
255. Kunarso, G., et al., *Transposable elements have rewired the core regulatory network of human embryonic stem cells.* Nat Genet, 2010. 42(7): p. 631-4.
256. Kim, S.W., et al., *Chromatin and transcriptional signatures for Nodal signaling during endoderm formation in hESCs.* Dev Biol, 2011. 357(2): p. 492-504.
257. Martin-Malpartida, P., et al., *Structural basis for genome wide recognition of 5-bp GC motifs by SMAD transcription factors.* Nat Commun, 2017. 8(1): p. 2070.
258. Sawada, A., et al., *Tead proteins activate the Foxa2 enhancer in the node in cooperation with a second factor.* Development, 2005. 132(21): p. 4719-29.
259. Grebbin, B.M. and D. Schulte, *PBX1 as Pioneer Factor: A Case Still Open.* Front Cell Dev Biol, 2017. 5: p. 9.
260. Burglin, T.R., *Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals.* Nucleic Acids Res, 1997. 25(21): p. 4173-80.
261. Chan, K.K., et al., *KLF4 and PBX1 directly regulate NANOG expression in human embryonic stem cells.* Stem Cells, 2009. 27(9): p. 2114-25.
262. Das, S., et al., *Transcriptional Regulation of Human NANOG by Alternate Promoters in Embryonic Stem Cells.* J Stem Cell Res Ther, 2012. Suppl 10: p. 009.
263. Knoepfler, P.S., et al., *Meis1 and pKnox1 bind DNA cooperatively with Pbx1 utilizing an interaction surface disrupted in oncoprotein E2a-Pbx1.* Proc Natl Acad Sci U S A, 1997. 94(26): p. 14553-8.
264. Fernandez-Diaz, L.C., et al., *The absence of Prep1 causes p53-dependent apoptosis of mouse pluripotent epiblast cells.* Development, 2010. 137(20): p. 3393-403.

265. Capellini, T.D., et al., *Pbx1/Pbx2 govern axial skeletal development by controlling Polycomb and Hox in mesoderm and Pax1/Pax9 in sclerotome.* Dev Biol, 2008. 321(2): p. 500-14.

266. Todeschini, A.L., A. Georges, and R.A. Veitia, *Transcription factors: specific DNA binding and specific gene regulation.* Trends Genet, 2014. 30(6): p. 211-9.

267. Fei, T., et al., *Smad2 mediates Activin/Nodal signaling in mesendoderm differentiation of mouse embryonic stem cells.* Cell Research, 2010. 20(12): p. 1306-1318.

268. Fei, T., et al., *Smad2 mediates Activin/Nodal signaling in mesendoderm differentiation of mouse embryonic stem cells.* Cell Res, 2010. 20(12): p. 1306-18.

269. Gaarenstroom, T. and C.S. Hill, *TGF-beta signaling to chromatin: how Smads regulate transcription during self-renewal and differentiation.* Semin Cell Dev Biol, 2014. 32: p. 107-18.

270. Raisner, R., et al., *Enhancer Activity Requires CBP/P300 Bromodomain-Dependent Histone H3K27 Acetylation.* Cell Rep, 2018. 24(7): p. 1722-1729.

271. Roux, K.J., et al., *BioID: A Screen for Protein-Protein Interactions.* Curr Protoc Protein Sci, 2018. 91: p. 19 23 1-19 23 15.

272. Roeder, R.G., *Transcriptional regulation and the role of diverse coactivators in animal cells.* Febs Letters, 2005. 579(4): p. 909-915.

273. D'Alessio, J.A., K.J. Wright, and R. Tjian, *Shifting players and paradigms in cell-specific transcription.* Mol Cell, 2009. 36(6): p. 924-31.

274. Xu, W., et al., *Loss of Gcn5l2 leads to increased apoptosis and mesodermal defects during mouse development.* Nat Genet, 2000. 26(2): p. 229-32.

275. Lin, W.C., et al., *Proper expression of the Gcn5 histone acetyltransferase is required for neural tube closure in mouse embryos.* Developmental Dynamics, 2008. 237(4): p. 928-940.

276. Lin, W.C., et al., *Proper Gcn5 histone acetyltransferase expression is required for normal anteroposterior patterning of the mouse skeleton.* Development Growth & Differentiation, 2008. 50(5): p. 321-330.

277. Mohibi, S., et al., *Mammalian Alteration/Deficiency in Activation 3 (Ada3) Is Essential for Embryonic Development and Cell Cycle Progression.* Journal of Biological Chemistry, 2012. 287(35): p. 29442-29456.

278. Seruggia, D., et al., *TAF5L and TAF6L Maintain Self-Renewal of Embryonic Stem Cells via the MYC Regulatory Network.* Mol Cell, 2019. 74(6): p. 1148-1163 e7.

279. Hirsch, C.L., et al., *Myc and SAGA rewire an alternative splicing network during early somatic cell reprogramming.* Genes Dev, 2015. 29(8): p. 803-16.

280. Ross, S., et al., *Smads orchestrate specific histone modifications and chromatin remodeling to activate transcription.* EMBO J, 2006. 25(19): p. 4490-502.

281. Jin, Q., et al., *Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation.* EMBO J, 2011. 30(2): p. 249-62.

282. Bonnet, J., et al., *The SAGA coactivator complex acts on the whole transcribed genome and is required for RNA polymerase II transcription.* Genes Dev, 2014. 28(18): p. 1999-2012.

283. Neri, F., et al., *Myc regulates the transcription of the PRC2 gene to control the expression of developmental genes in embryonic stem cells.* Mol Cell Biol, 2012. 32(4): p. 840-51.

284. Yu, Y., et al., *Inhibition of EZH2 Promotes Human Embryonic Stem Cell Differentiation into Mesoderm by Reducing H3K27me3.* Stem Cell Reports, 2017. 9(3): p. 752-761.

285. Collinson, A., et al., *Deletion of the Polycomb-Group Protein EZH2 Leads to Compromised Self-Renewal and Differentiation Defects in Human Embryonic Stem Cells.* Cell Rep, 2016. 17(10): p. 2700-2714.

286. Shan, Y., et al., *PRC2 specifies ectoderm lineages and maintains pluripotency in primed but not naive ESCs.* Nat Commun, 2017. 8(1): p. 672.

287. Fagnocchi, L., et al., *A Myc-driven self-reinforcing regulatory network maintains mouse embryonic stem cell identity.* Nat Commun, 2016. 7: p. 11903.

288. Kim, J., et al., *A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs.* Cell, 2010. 143(2): p. 313-24.

289. Varlakhanova, N., et al., *Myc and Miz-1 have coordinate genomic functions including targeting Hox genes in human embryonic stem cells.* Epigenetics Chromatin, 2011. 4: p. 20.

290. Knoepfler, P.S., et al., *Myc influences global chromatin structure.* EMBO J, 2006. 25(12): p. 2723-34.
291. Faiola, F., et al., *Dual regulation of c-Myc by p300 via acetylation-dependent control of Myc protein turnover and coactivation of Myc-induced transcription.* Mol Cell Biol, 2005. 25(23): p. 10220-34.
292. Marmorstein, R., et al., *Structure and function of the SWIRM domain, a conserved protein module found in chromatin regulatory complexes.* Faseb Journal, 2006. 20(4): p. A34-A35.
293. Helmlinger, D. and L. Tora, *Sharing the SAGA.* Trends Biochem Sci, 2017. 42(11): p. 850-861.
294. Guelman, S., et al., *The double-histone-acetyltransferase complex ATAC is essential for mammalian development.* Mol Cell Biol, 2009. 29(5): p. 1176-88.
295. Wang, T., et al., *Genetic Screens in Human Cells Using the CRISPR-Cas9 System.* Science, 2014. 343(6166): p. 80-84.
296. Tzelepis, K., et al., *A CRISPR Dropout Screen Identifies Genetic Vulnerabilities and Therapeutic Targets in Acute Myeloid Leukemia.* Cell Rep, 2016. 17(4): p. 1193-1205.
297. Tsai, P.H., et al., *Ash2l interacts with Oct4-stemness circuitry to promote super-enhancer-driven pluripotency network.* Nucleic Acids Res, 2019.
298. Gresakova, V., et al., *Fam208a orchestrates interaction protein network essential for early embryonic development and cell division.* Exp Cell Res, 2019. 382(1): p. 111437.

10 Appendix

		MS1 replicates																	
Accession	Uniprot	NANOG plurip 1	NANOG plurip 2	NANOG plurip 3	NANOG diff 1	NANOG diff 2	NANOG diff 3	GSC plurip 1	GSC plurip 2	GSC plurip 3	GSC diff 1	GSC diff 2	GSC diff 3	Control plurip 1	Control plurip 2	Control plurip 3	Control diff 1	Control diff 2	Control diff 3
P46937	YAP1_HUMAN	122.41	0.56	630.57	3848.56	2961.09	2355.68	49.49	205.64	2134.51	6451.41	8502.54	3551.82	44.94	205.55	352.51	806.72	511.85	1052.13
P28347	TEAD1_HUMAN	2506.99	2151.98	1532.27	1732.54	1840.50	2682.07	2657.53	1661.28	2367.23	3052.38	3848.64	2296.45	0.88	384.23	259.35	210.43	311.98	408.52
P41212	ETV6_HUMAN	101.70	0.00	1161.33	805.63	0.00	551.83	0.00	0.00	0.00	0.00	0.00	115.20	19.34	0.00	0.00	1.00	0.00	0.00
P39880	CUX1_HUMAN	232.33	324.47	900.72	796.61	1241.03	973.16	2915.67	3005.15	3285.00	6500.70	5719.27	4713.76	308.15	730.02	600.67	2367.05	730.98	2942.07
Q9UD57	NKX12_HUMAN	0.26		51.72	501.33	99.15	417.85	3.94	4.30	14.53	746.48	1018.33	1941.79	5.88	30.42	17.96	584.43	713.80	489.59
P55347	PKNX1_HUMAN	1218.11	1381.87	1555.44	4258.30	1576.93	3232.43	870.03	1195.82	41.54	554.40	739.97	880.04	324.64	340.24		75.38	79.90	3.96
Q06330	SUH_HUMAN	64420.30	65682.29	67753.15	33051.69	40694.55	58243.64	431004.53	446007.41	179722.30	208756.77	260526.26	271477.30	67379.89	85002.16	50669.84	96180.10	53358.97	62784.37
Q9UBG7	RBPJL_HUMAN	1532.40	1408.93	2286.06	1655.40	1368.82	2765.58	12286.84	21889.31	3113.88	5468.31	8793.31	5291.71	1696.91	2277.66	1293.45	3429.08	1887.04	395.64
Q9Y261	FOXA2_HUMAN	1264.98	1351.19	247.82	3972.69	4860.95	646.56	330.42	433.24	181.84	4969.81	5961.67	1291.66	616.98	71.13	264.36	1118.45	3983.03	1339.19
Q15583	TGIF1_HUMAN	1.93	24.58	1620.07	3716.14	567.41	2105.29	151.79		339.20	743.15	741.70	671.39	60.67	449.70	167.72	276.95	353.99	60.57
Q13948	CASP_HUMAN	448.25	359.70	566.86	1659.69	2943.98	1905.97	441.73	1059.88	2305.17	3927.04	4592.23	5030.91	531.79	923.48	536.55	1859.63	2110.46	3127.09
Q9UNS1	TIM_HUMAN	8.39	0.01	67.09	15.68	8.90	11.13	27.39	16.07	45.38	57.20	310.03	362.61	6.80	131.63	161.98	107.29	60.83	21.99
Q01860	PO5F1_HUMAN	255666.29	269924.95	163820.90	262085.90	277572.68	170941.10	245068.15	249444.09	100585.09	321827.10	354912.92	221795.68	102809.87	76819.78	49435.69	101644.18	166561.06	119058.05
Q9UK61	TASOR_HUMAN	8.56		300.52	12.99	7.46	50.86	0.05		39.83		37.16	278.64	4.58	166.22	144.51	17.36		7.37
Q15014	MO4L2_HUMAN	16.51	7.26	342.75	12.80	20.72	144.78	10.82	16.67	151.60	28.02	168.51	613.68	18.38	185.57	185.81	101.12	105.86	8.94
Q5TKA1	LIN9_HUMAN	82.77	34.93	699.19	506.82	1029.65	2527.03	105.81	66.74	1354.11	2185.94	3140.90	3417.08	80.66	1053.71	568.81	1667.67	1048.51	2435.09
P05204	HMGN2_HUMAN	63532.88	122643.58	30531.50	94159.02	44229.12	31545.89	77655.39	254129.83	48311.75	33544.42	269764.89	51086.71	83234.42	30748.85	32712.18	50640.81	57679.89	50611.06
Q15406	NR6A1_HUMAN	25047.42	30389.09	23150.66	50264.29	46891.31	40571.94	26158.30	26845.24	19981.84	36802.65	36445.32	34665.34	12314.06	16048.79	10455.42	18985.77	18425.71	18350.48
P63104	1433Z_HUMAN	458.47	99.07	1451.11	1750.28	478.57	1058.98	600.55	934.08	1551.58	1473.88	2637.89	4527.69	591.83	1318.20	1915.10	1276.22	1374.35	1134.00
Q06416;P20264	P5F1B_HUMAN	9507.84	7813.52	8000.71	15660.40	18937.40	6439.77	6977.37	11312.92	4458.59	15339.54	18298.63	6406.86	4484.71	5456.66	2325.11	4670.31	13464.69	5052.16
P49640	EVX1_HUMAN	388.19	836.13	498.22	1644.39	3111.45	699.40	593.02	304.64	506.65	1791.43	1928.56	1432.07	417.61	648.01	587.02	757.17	1463.10	805.37
Q96JM7	LMBL3_HUMAN	128.77	200.55	530.40	556.12	507.65	1498.97	124.75	76.89	1159.94	1030.69	1688.22	1276.83	36.86	765.88	795.61	995.97	1167.22	555.45
Q7Z2T5	TRM1L_HUMAN	850.12	967.19	491.51	470.22	1330.31	670.11	2244.90	3260.78	771.86	2381.90	2378.72	1867.84	1836.76	731.16	497.34	1847.78	316.07	1046.73
O75461	E2F6_HUMAN	86.50	1.32	665.83	268.29	1004.33	101.19	61.97	148.94	293.49	277.32	228.78	1824.29	77.53	634.08	509.84	362.39	218.96	435.00
Q14562	DHX8_HUMAN	40.46	20.15	82.73	9.02	2.87	1.06	1.68	3.85	21.21	15.59	35.52	177.80	7.78	96.66	62.02	15.34	31.47	7.58
P62834	RAP1A_HUMAN	7.88	1.39	100.11	27.67	0.00	0.11	18.69		12.76	6.65	78.00	244.62	0.47	41.91	11.03	11.03	150.88	8.63
P31941	ABC3A_HUMAN	323.53	5.31	899.12	789.87	22.48	649.68	80.70	403.18	437.47	462.87	952.68	1472.48	36.06	586.08	988.70	475.03	307.93	509.69
P23511	NFYA_HUMAN	9208.87	55654.49	6931.40	24512.39	27783.10	18874.63	40123.51	56645.06	10414.69	31513.63	61819.00	32369.99	19836.68	12388.41	9233.33	61193.66	14321.71	10159.64
Q9H2W2;O75360	MIXL1_HUMAN	34822.16	103207.22	27419.90	778450.63	759115.06	361476.15	27686.29	45392.22	28064.47	744392.82	688382.11	283211.17	51729.62	26609.97	27441.05	486435.14	788161.19	375311.59
Q05BQ5	MBTD1_HUMAN	256.22		886.35	413.13	47.35	667.76	362.71	89.48	674.50	972.09	924.42	1435.37	378.27	1243.22	570.76	333.14	292.29	627.56
Q9UPT8	ZC3H4_HUMAN			236.96	295.34					181.35	312.82	384.74	435.40	4.69	177.52	311.83	269.51	274.13	140.38
Q2NKX8	ERC6L_HUMAN	449.78	757.40	412.79	1581.56	1628.77	1039.50	955.66	901.53	264.54	2342.74	1153.94	378.67	1204.76	190.30	223.95	553.23	1567.14	323.62
Q9H8Y1	VRTN_HUMAN	110682.65	148848.84	74415.51	245321.57	245079.34	161477.11	228479.98	270323.02	117249.81	351181.08	402044.17	240627.34	149973.00	114539.36	102802.83	214155.88	313555.21	150411.32
Q8IZX4	TAF1L_HUMAN	513.48	667.58	348.44	882.48	633.80	543.43	715.17	1355.02	448.80	680.25	766.52	725.65	443.39	412.02	288.58	308.99	544.51	298.50
Q8WUY9	DEP1B_HUMAN			67.74			24.77		29.24	19.13	2.78	293.83	121.14		59.26	106.88	65.29	129.04	12.36
Q92908	GATA6_HUMAN	27691.65	14606.75	14006.06	43005.86	44849.89	11608.79	20716.81	16866.43	15250.94	49413.78	36573.37	17007.21	18294.02	9256.48	9948.09	13522.90	44844.62	16653.37
Q12800	TFCP2_HUMAN	406189.07	463728.57	189169.72	644256.04	738889.42	480480.36	610168.55	596039.08	294448.25	770941.19	836176.51	567811.85	401228.39	237493.86	203731.09	453894.58	654350.12	429969.66
P61981	1433G_HUMAN	570.36	294.19	1183.90	1344.25	567.42	610.71	481.66	414.24	781.63	1782.52	1009.04	2025.70	488.20	876.00	984.51	1070.99	1167.22	722.33
Q5C9Z4	NOM1_HUMAN	389.25	162.69	565.31	651.95	204.28	463.23	542.05	867.97	493.24	809.46	966.16	1021.37	412.35	495.23	702.15	318.53	772.04	382.82
Q7Z460	CLAP1_HUMAN	140.06	177.86	202.83	205.86	204.27	253.87	186.11	94.23	433.98	535.12	524.55	822.24	158.14	612.80	362.01	283.32	253.32	441.49
O60293	ZC3H1_HUMAN	37.70	48.00	596.43	340.87	131.21	256.86	59.30	78.88	193.88	347.46	417.47	1074.65	59.80	415.21	405.00	610.63	332.94	251.49
Q9NP50	FA60A_HUMAN			281.72		14.68	70.84		4.22	66.76	10.59	52.87	459.23		150.66	188.33	69.55	14.11	6.43
Q13952	NFYC_HUMAN	138696.77	162145.46	83568.27	164257.04	154949.36	130269.84	189344.63	213804.60	114634.95	233259.84	232315.17	168011.41	126872.66	101101.38	70862.81	167107.50	159409.75	100402.87
Q07864	DPOE1_HUMAN	9.80	13.62	347.22	208.15	5.07	5.46	100.58	57.22	40.88	525.21	698.57	561.11		445.20	380.01	109.16	585.27	154.22
P40425	PBX2_HUMAN	467.68	1573.92	262.78	9085.43	6258.22	2209.88	2018.49	879.62	151.63	1595.20	1758.07	704.06	397.06	78.94	0.00	1663.93	1968.04	642.25
Q9HBE1	PATZ1_HUMAN	25941.10	48857.57	11626.36	23535.56	19671.01	14786.95	39973.40	36252.60	14685.68	20883.17	21087.27	13403.35	14675.12	8407.16	9320.09	10022.92	14740.19	7824.52
Q9UBL3	ASH2L_HUMAN			344.49	20.83		75.79	6.32		129.97	62.43	173.69	423.24	2.39	219.80	312.60	66.48	102.54	82.23
Q9NZI6	TF2L1_HUMAN	257.56	176.98	33.83	182.54	452.74	142.87	334.75	362.56	134.01	213.71	350.35	127.96	270.93	52.71	49.59	33.29	255.49	170.40
P26232	CTNA2_HUMAN	232.11	103.26	1064.83	526.34	204.27	2222.78	106.44	195.30	594.34	431.41	1068.44	2207.52	614.84	703.92	1291.33	550.42	462.82	977.37
Q9NVN8	GNL3L_HUMAN	462.05	853.97	678.80	2335.37	8507.12	530.22	280.05	428.89	630.29	1051.28	1752.12	1706.41	1115.50	737.04	918.81	918.26	1430.39	498.91
Q9BQ67	GRWD1_HUMAN	5.64		301.37	9.44		184.81	18.18	4.50	48.00	70.04	4.99	266.50	267.62	59.82	158.53	52.24		
P29692	EF1D_HUMAN	4671.61	7369.52	1141.62	373.71	2240.05	356.75	2815.16	1899.74	1055.55	1823.01	5271.49	537.59	2190.51	2072.61	830.45	3801.32	242.92	508.99
Q8NEG4	FA83F_HUMAN	20.66	8.27	356.51	341.44	142.26	232.72	24.62	48.00	165.03	174.45	542.86	584.06	22.12	537.33	335.72	215.57	261.29	297.40
Q9NZI7	UBIP1_HUMAN	56854.03	70641.94	17880.89	93620.39	102445.61	64623.11	67110.47	74798.24	36387.12	82489.93	97395.97	56701.67	46299.56	32415.02	30911.42	57188.88	83023.48	54881.72
Q9NSI2	F207A_HUMAN	1779.96	2122.86	748.70	4941.53	5628.47	1495.82	1208.00	1623.22	3285.60	2789.92	1584.69	1361.81	640.91	981.62	2188.77	3214.70	1516.46	
Q9Y2X9	ZN281_HUMAN	343.00	605.60	1760.61	1573.83	1282.50	2024.20	654.98	642.60	1421.16	1604.45	2353.11	2537.31	818.52	1961.44	1696.34	1194.39	1567.27	1214.86
Q96QE3	ATAD5_HUMAN	66.87	63.21	437.61	43.05	130.97	226.93	31.09	24.95	199.60	125.54	479.59	621.67	12.27	564.71	413.89	284.14	212.81	109.16
Q9UN81	LORF1_HUMAN	428.5807654	574.75	776.87	903.24	936.75	633.14	722.54	538.48	789.76	901.41	857.76	536.22	614.84	409.38	572.89	560.48		
P11166	GTR1_HUMAN	728.17	1020.01	611.36	654.10	635.09	693.24	1129.34	576.40	636.76	1559.70	707.97	943.23	642.77	1067.85	640.29	513.23	669.77	704.90
Q9UJU2	LEF1_HUMAN	15760.88	24158.85	8496.00	11116.03	7191.79	12513.69	10635.81	8360.18	5017.65	7057.54	6203.40	7737.08	3869.18	5268.00	3210.99			
Q8NBU5	ATAD1_HUMAN	1027.98	1083.33	482.57	971.97	1223.76	364.98	1140.87	870.09	399.00	1073.60	1140.59	688.98	855.86	653.18	483.54	248.89	1371.42	259.67
Q15723	ELF2_HUMAN	4566.52	7270.26	1374.80	3840.86	3874.16	5017.47	4769.70	4998.70	1771.52	6435.05	4538.75	5058.26	3217.76	3945.31	2387.69	4644.65	3926.33	2135.09
Q01831	XPC_HUMAN	36.55	866.47	492.80	745.69	2322.06	873.22	1026.56	774.36	463.95	1387.11	1103.58	621.33	768.87	304.06	593.22	350.01	1386.45	768.89
Q8TD26	CHD6_HUMAN	6381.25	5029.14	4758.86	3840.86	3874.16	5404.77	4999.85	3508.62	3217.76	3945.31	2387.69	4440.60	4938.82	3003.91				
O95347	SMC2_HUMAN	96.11	29.95	372.71	163.05	65.10	225.08	191.12	86.24	229.47	201.63	466.01	689.39	324.38	345.89	336.65	126.62	466.78	229.15
Q86U86	PB1_HUMAN	0.17		18.79	4.52	23.88	4.49	16.90	3.39	79.87	24.48	58.77	219.81	6.02	116.21	58.70	44.54	9.41	23.07
Q04917	1433F_HUMAN	77.79	50.04	406.05	244.00	118.31	239.91	94.64	32.52	219.85	440.90	258.05	409.15	150.05	256.67	458.19	68.16	488.88	90.04
Q9ULW3	ABT1_HUMAN	2762.85	3169.25	1834.08	2527.97	2255.57	1780.00	930.06	1433.22	1093.84	1500.21	2044.84	1782.22	596.47	1372.76	2122.43	396.35	1256.71	1789.66
P16403	H12_HUMAN	679.60	7510.26	2420.99	3720.52	5470.86	2895.47	11403.41	10533.41	5337.14	3217.14	3655.47	3742.38	4284.72	2704.45	661.06	2223.69	5482.93	3105.60
Q14119	VEZF1_HUMAN	14116.96	17112.64	9705.89	9577.35	10772.51	14307.62	10277.26	14331.05	11025.88	10912.87	13486.23	11609.87	10150.39	6637.51	8370.88	10783.22	8784.87	6480.30
P14859;P09086;Q9UKI9	PO2F1_HUMAN	69479.25	106873.21	84042.63	220010.02	197549.56	144726.17	49281.19	45473.67	29937.82	128507.19	135578.05	115703.88	61269.99	65546.33	30528.30	147627.39	135638.92	106784.71
P60953	CDC42_HUMAN	333.55	469.41	493.90	768.59	422.46	418.70	428.83	728.96	423.41	508.27	924.88	759.44	653.00	568.34	489.86	433.23	649.11	372.74
Q99959	PKP2_HUMAN	144.53	12.44	241.10	85.42	27.02	1206.77	101.85	248.72	199.05	81.12	219.03	455.14	11.64	182.51	271.52	348.91	230.74	52.38
P17480	UBF1_HUMAN	517905.59	580322.18	203295.11	599578.91	503767.69	418736.92	582686.88	551287.15	332030.36	495123.10	601791.27	287922.57	405915.47	186006.74	308646.89	362877.49	489123.77	261618.53
Q5T5X7	BEND3_HUMAN	64643.54	90901.27	61191.44	66947.41	62500.57	109851.05	97221.99	93869.60	107259.74	86081.66	87086.03	110348.87	56739.39	58564.92	74250.90	94493.19	56925.94	77194.49
Q86TJ2	TAD2B_HUMAN	7.00		355.38	0.19	8.90		4.27	2.45	141.79	0.80	83.71	506.32	1.66	353.02	309.17	176.91	17.68	18.08
Q96GN5	CDA7L_HUMAN	18593.96	14957.85	12000.19	36912.51	32607.39	20937.83	11538.79	15244.43	22400.15	28496.25	23745.76	11933.10	12409.62	9288.88	17755.86	18248.60	22755.57	14834.99
Q9GZR2	REXO4_HUMAN	18.20	2.03	1104.11	633.25	24.03	529.29		88.41	808.74	599.75	855.65	1263.60	108.49	1498.02	1343.47	558.45	427.39	115.03

		MS1 replicates																	
Accession	Uniprot	NANOG plurip 1	NANOG plurip 2	NANOG plurip 3	NANOG diff 1	NANOG diff 2	NANOG diff 3	GSC plurip 1	GSC plurip 2	GSC plurip 3	GSC diff 1	GSC diff 2	GSC diff 3	Control plurip 1	Control plurip 2	Control plurip 3	Control diff 1	Control diff 2	Control diff 3
Q92616	GCN1_HUMAN	533.19	239.93	760.31	661.16	428.67	280.67	758.39	814.99	486.48	917.72	1117.63	819.44	914.30	886.93	943.56	336.07	870.42	312.98
O14880	MGST3_HUMAN	433.90	87.50	729.63	375.24	124.36	198.81	300.01	364.62	699.78	407.37	764.19	438.75	389.96	582.69	483.25	189.68	786.67	43.09
P33993	MCM7_HUMAN	2639.87	2399.68	2839.01	8297.54	5213.41	3191.63	3432.87	4232.20	2966.61	4852.83	4513.94	3573.98	3437.06	3796.68	2948.61	3212.35	3265.91	3445.56
P13674	P4HA1_HUMAN		2.13	170.67	310.47	13.70	214.59		104.15	52.36	537.34	622.97	325.38	64.99	794.14	636.34	377.53	413.09	33.14
Q3YBR2	TBRG1_HUMAN	3709.15	4289.36	1701.85	4727.79	3487.73	2243.36	1949.89	4601.24	1746.04	3064.67	2834.48	941.66	2498.79	2133.68	750.06	1635.35	2834.27	862.32
Q8N9M1	CS047_HUMAN	123.02	37.85	429.27	1277.81	312.14	450.24	112.84	173.98	210.64	589.02	519.85	612.57	366.65	561.74	428.52	689.17	415.08	236.28
Q9H160	ING2_HUMAN	1764.98	1951.48	1744.85	2084.69	1660.78	2198.66	3059.53	1899.20	2498.48	1912.14	1723.92	2092.36	1415.78	1437.39	1532.26	2048.11	1438.08	1317.07
Q5M775	CYTSB_HUMAN	23.43		505.11	311.80	52.49	339.36	40.79	126.08	227.61	208.48	206.46	1058.31	3.92	316.74	1254.88	271.17	96.89	474.93
P16401	H15_HUMAN	1866026.92	1878135.01	231740.62	1291224.15	1210773.25	719549.22	1639336.62	1413685.11	957713.68	1168970.55	1288357.44	663511.35	1401610.05	226195.42	994287.34	913710.07	1058906.89	610813.03
P25208	NFYB_HUMAN	77361.36	92814.79	21564.56	82927.96	127546.92	41655.62	76775.94	137803.20	39467.77	90638.05	80402.54	36266.53	65094.65	40455.04	25016.44	68884.86	97263.13	49450.08
Q86VM9	ZCH18_HUMAN	1350.91	8324.19	2746.29	9241.93	6243.15	3950.58	4840.72	6617.49	1855.70	5404.81	4174.53	2648.45	5177.48	1455.75	2015.33	3234.87	5136.48	3619.58
Q9UHX1	PUF60_HUMAN	5222.40	7630.98	656.81	2923.31	4676.24	1165.25	3619.31	3913.26	395.40	3333.61	3091.90	743.98	3756.75	1210.28	395.38	1861.01	3428.86	1565.38
Q02447	SP3_HUMAN	14096.05	13102.31	6153.27	7120.51	4185.11	5025.42	14000.38	13670.22	5184.63	7253.13	7303.26	5755.72	7745.06	5924.86	3679.88	6287.64	6165.05	4934.36
Q7RTV0	PHF5A_HUMAN	49736.27	60775.05	29404.61	50699.08	35065.45	27563.32	55883.26	44030.36	27856.22	36391.94	37790.83	25360.88	33161.78	24064.64	23215.26	35617.28	31876.58	23741.61
P15924	DESP_HUMAN	18100.04	13862.40	3764.57	12493.52	14865.63	14806.54	10668.56	9865.28	7818.79	11266.06	11706.07	10270.46	9173.73	4806.19	5104.08	11954.14	14337.91	14482.15
P11388	TOP2A_HUMAN	3042313.06	3760277.27	1307020.12	2589652.57	2500849.09	1607922.13	3663814.32	3698269.49	1806426.19	2318919.05	2471484.16	1292639.42	3142090.47	1247653.31	1608529.84	1469986.66	2320582.87	1258995.63
Q8N3C0	ASCC3_HUMAN	252.11	582.32	154.27	135.67	242.54	63.96	305.39	309.35	34.17	171.28	182.47	122.85	290.53	91.25	53.36	119.23	198.50	115.18
Q92522	H1X_HUMAN	213587.73	255558.11	57439.61	91771.39	92215.53	67777.74	229313.13	211872.97	131925.54	86025.67	105163.00	89474.36	117222.56	63792.71	117225.69	94900.75	67848.32	52075.70
P50402	EMD_HUMAN	42511.16	46271.89	20034.80	26085.26	23435.84	33103.32	41273.64	40048.01	27977.05	23733.20	25500.47	24218.00	31915.31	15523.20	21994.78	17123.07	27755.89	20788.06
P50151	GBG10_HUMAN	840.43	1008.62	513.23	3218.32	3675.61	3365.11	818.86	658.06	971.15	1568.51	1876.62	1164.43	494.51	449.48	891.67	1783.60	2385.99	2478.21
P40424	PBX1_HUMAN	8637.13	14779.51	8577.00	25723.67	15624.19	15459.57	7519.95	8481.31	3456.27	6294.97	6581.44	1821.01	5470.95	4525.40	2001.07	6893.56	5680.88	2565.38
Q15717	ELAV1_HUMAN	37492.07	56760.14	17760.33	31183.78	30173.61	26431.42	53807.02	60873.37	25269.79	28880.23	34694.25	19746.95	36709.35	16474.71	22630.93	29472.13	27075.02	21669.38
Q9UIF9	BAZ2A_HUMAN	3664.02	3745.20	3666.50	2265.12	1788.37	1388.23	9666.19	4419.44	4471.95	4444.90	3312.51	2809.25	5826.10	3346.84	2784.55	2311.33	3327.37	2085.40
Q9UIG0	BAZ1B_HUMAN	102779.19	126792.21	34431.67	69756.53	66672.49	51122.26	141164.42	127584.38	68021.57	75545.45	73540.11	41372.26	103695.97	40955.45	55072.62	41535.35	73212.11	41228.86
Q86U42	PABP2_HUMAN	1398.01	1501.11	878.80	1490.62	1234.95	574.47	976.49	900.02	665.61	1180.23	987.44	428.99	924.67	849.50	476.42	398.16	779.30	1441.37
Q5SSJ5	HP1B3_HUMAN	53091.38	53777.38	9745.69	29557.91	37988.20	16679.54	42399.99	42607.16	23212.39	27636.27	24556.14	13103.18	33833.82	9760.93	21323.74	14020.96	29894.70	16327.82
Q14674	ESPL1_HUMAN			130.13	8.12		40.33			48.32	15.02	41.32	151.88		118.16	104.45	30.19	23.23	12.09
P84090	ERH_HUMAN	309853.78	466575.68	148158.86	294383.56	213279.39	263698.09	326862.50	365649.47	191915.89	218302.36	225196.61	229915.24	238956.82	143740.66	181115.73	394834.14	187755.39	167554.36
Q9GZL7	WDR12_HUMAN	2433.74	1923.40	2113.73	4706.90	2041.18	3190.85	2277.59	2515.90	3093.03	2784.01	2986.90	1250.72	2042.76	2322.28	2760.95	2375.07	2719.78	1518.38
Q9Y3Y2	CHTOP_HUMAN	77697.88	117709.96	48877.87	62860.26	44572.51	102492.76	96368.72	86795.97	45295.64	49624.89	71466.46	99170.35	58811.22	31805.43	49691.40	182224.09	36786.80	35538.01
A6NHQ2	FBL1_HUMAN	5568.48	3776.94	4794.07	14419.04	9892.67	5691.87	6158.58	7682.59	9190.08	6859.73	5132.47	5610.48	6143.29	6622.89	6598.03	3715.63	4626.27	7002.18
Q8IWS0	PHF6_HUMAN	68142.42	89672.68	47700.58	57252.67	49150.65	60635.34	70521.62	69738.38	46717.56	45457.72	50756.27	42880.19	48464.71	39324.37	46481.11	55835.39	43508.57	40657.66
P18754	RCC1_HUMAN	120345.09	123482.20	49253.08	65778.31	63629.32	58015.07	97783.57	96600.25	71340.82	60137.06	64708.08	57884.16	80521.63	46905.27	66953.98	59182.44	63285.37	48713.40
O15160	RPAC1_HUMAN			933.94	41.23	11.94	15.08	1.24	227.00	78.20	32.52	107.85	518.28	17.66	753.77	346.15	107.94	89.01	3.64
O95400	CD2B2_HUMAN	402.97	488.11	641.32	863.54	1387.13	672.97	786.55	648.63	346.59	827.21	433.84	527.43	790.17	477.43	744.83	515.64	827.21	243.37
P06493	CDK1_HUMAN	1449.37	1948.42	954.71	3713.01	3178.13	1095.09	1876.33	1254.25	1728.32	1748.64	1740.59	1248.89	1094.61	2103.82	1144.45	817.75	2428.33	1962.74
Q6RFH5	WDR74_HUMAN	2283.44	4707.24	2861.52	4246.90	5374.87	5443.66	4005.13	3965.08	3451.76	3419.81	2643.16	3280.56	3232.97	2908.97	3105.51	3111.46	3324.75	3258.43
Q53F19	NCBP3_HUMAN	3806.41	10538.23	5673.87	7305.77	6288.25	3402.27	11913.12	9739.28	5626.46	5815.52	5488.27	3325.43	7496.49	4650.18	4590.86	3313.86	6845.10	3210.03
DECOY_P29353	SHC1_HUMAN	1724744.42	2004913.87	677198.21	1016750.59	945932.19	929625.92	1666067.40	1449076.70	640754.47	972922.47	811206.59	912111.80	1430594.42	785161.21	720268.47	1073173.31	890449.44	663764.36
Q9ULU4	PKCB1_HUMAN	135851.76	200251.83	125851.21	118589.88	131827.95	148063.55	242320.58	251293.73	127087.99	135618.79	116154.59	119822.35	145449.65	137372.39	107797.25	190048.27	79290.93	116394.84
				71.26				33.68	29.71		25.72	97.10	168.52		229.71		72.03		52.50
O60828	PQBP1_HUMAN	9232.32	17149.37	12208.68	22833.69	13979.94	11206.09	11943.75	17902.34	7672.11	13753.45	11253.57	6220.32	14843.29	11635.73	9117.33	12061.93	10129.67	7715.93
Q7L7X3	TAOK1_HUMAN	5347.58	6199.46	1589.24	2887.56	4072.88	880.48	8166.49	6132.65	1852.86	3954.31	3931.84	998.01	6004.26	1796.18	1769.62	3609.00	4091.68	1868.45
Q8N567	ZCHC9_HUMAN	3585.96	4697.96	2080.79	4315.60	5250.15	3255.20	3427.26	4958.74	2148.29	3118.16	2894.90	1872.80	3785.29	2474.58	2438.30	2507.24	3119.42	2745.09
P06239;O94921;P51451	LCK_HUMAN	1334.69	1084.53	936.16	1084.42	1403.22	1428.38	1202.16	2975.80	1129.35	1336.38	1210.42	1498.84	974.67	897.81	1433.47	2172.34	1676.24	1669.14
Q13601	KRR1_HUMAN	25624.41	26492.06	19853.42	31420.19	28739.38	41489.78	26050.64	25538.04	30291.07	25146.59	22527.73	25628.46	24741.50	30559.06	37076.28	23507.53	21721.14	22477.05
O60481	ZIC3_HUMAN	4057.73	4709.20	1267.93	2477.39	2334.45	1787.47	5602.77	4213.39	1501.81	2116.79	1624.70	1526.13	3510.73	1406.28	1403.41	1670.03	1911.30	1622.57
P18583	SON_HUMAN	43355.85	50586.90	21562.68	24335.75	21735.86	17818.03	30842.68	37300.64	17683.82	17757.79	19726.32	22762.97	32876.25	20927.77	19577.31	21973.36	19330.30	17707.65
Q7Z7K6	CENPV_HUMAN	27935.17	79628.84	7380.20	28422.64	13437.00	13073.98	47246.71	60231.60	20031.58	21287.81	20531.35	10195.98	47736.48	11410.26	20916.16	9975.82	17638.78	7167.57
O14776	TCRG1_HUMAN	47179.47	59327.23	82452.82	36061.88	38174.17	34868.96	53754.27	63926.45	35172.69	33555.69	39022.31	35232.66	43824.12	38762.67	37322.20	41919.81	37966.26	39302.84
Q15361	TTF1_HUMAN	89613.01	109144.74	23300.24	57532.01	63881.46	31733.67	54803.07	56398.98	21874.69	35938.04	36329.28	24823.41	56086.56	19406.34	30927.86	35434.85	47133.64	26846.86
Q9Y6A4	CFA20_HUMAN	3068.03	3932.17	2006.87	5886.95	9075.57	4089.76	2237.63	3859.98	2507.63	3179.73	3692.85	3042.33	5345.54	2372.39	2994.33	3447.76	4420.68	3476.67
Q5BKZ1	ZN326_HUMAN	70792.57	72934.84	31278.11	43369.73	45102.87	37863.43	62777.23	57464.18	36256.34	43156.71	53496.81	30850.22	47220.72	22261.93	35137.07	106656.64	43744.52	29453.75
P67809	YBOX1_HUMAN	148879.69	183644.16	52751.57	105311.84	74613.09	80899.63	129344.75	121714.88	69606.95	65048.67	65604.41	53601.73	95795.95	49761.78	58158.66	85758.87	74887.47	48705.75
P49711;Q8NI51	CTCF_HUMAN	117890.50	177637.76	55677.73	72122.99	50790.67	58571.14	119549.57	116600.67	74772.68	62371.03	60383.04	49849.07	98815.87	67061.78	68015.00	68379.39	52428.03	34343.19
Q96NC0	ZMAT2_HUMAN	2587.76	3925.35	910.03	2059.07	1522.08	1323.38	2336.38	2415.81	1579.87	1324.50	1206.69	1116.18	1406.10	1756.78	1346.62	1506.07	1135.94	1091.36
P55081	MFAP1_HUMAN	131674.18	193559.91	92621.91	126462.34	80382.93	86859.16	140910.62	138155.91	85023.20	79139.45	78246.62	67562.73	104296.49	83427.48	76176.61	106108.60	79168.10	63420.45
DECOY_Q9P219	DAPLE_HUMAN	16854.45	12280.32	6806.34	9792.52	12516.47	5300.66	6764.93	9137.82	7844.04	7091.02	4474.59	4704.98	11203.51	4395.43	6584.50	3504.76	6045.54	5630.89
O00541	PESC_HUMAN	9527.75	7836.35	11267.64	6541.02	4529.46	7966.41	7191.43	7806.91	7893.85	6159.20	5155.63	5956.30	7007.67	6427.10	5246.91	8044.97	4872.72	8639.66
Q9UNQ2	DIM1_HUMAN	4210.63	2315.36	2668.57	1845.22	3660.67	4507.26	2856.08	2330.00	2147.90	2000.32	1866.18	2208.31	3222.13	2183.86	2393.64	1967.66	1859.75	2682.59
Q86Y07	VRK2_HUMAN	368.21	963.13	638.32	598.37	467.20	406.66	2917.04	603.49	844.11	480.70	604.35	629.29	995.71	747.94	669.49	403.39	828.96	411.56
Q15287	RNPS1_HUMAN	42710.67	68189.81	14362.38	39163.31	24686.68	17813.47	47997.01	48941.52	19086.71	18059.81	22233.86	16929.63	36675.39	22929.17	20133.73	22187.95	21716.89	12906.60
Q8N5F7	NKAP_HUMAN	1117.73	2210.22	754.90	1365.36	2343.62	410.85	1742.36	1640.27	607.99	950.90	729.34	539.77	1381.87	885.31	814.62	684.72	946.43	606.71
Q9UQ35	SRRM2_HUMAN	4078.42	9056.89	4314.96	4408.76	3555.76	1658.65	5096.25	6107.77	2328.12	2613.77	2733.33	2302.64	4253.93	3416.91	3380.40	3349.57	2179.15	1793.51
O43251	RFOX2_HUMAN	3459.29	6562.82	3579.39	3901.63	6104.68	5125.21	5088.57	5354.03	2116.47	3347.54	3283.24	2335.22	5368.71	4191.06	2556.97	2345.04	3676.51	3545.58
Q8ND82	Z280C_HUMAN	9843.41	16078.83	2588.01	6658.22	5488.68	1840.94	16073.21	13905.46	2432.57	3318.60	4536.59	1605.33	8719.34	2356.91	2374.14	2714.42	5021.36	1762.97
P0C1Z6	TFPT_HUMAN	5246.82	7530.61	2404.29	5984.89	6348.19	4692.68	5092.92	6310.50	3009.75	4001.44	3376.74	2314.59	5021.95	3480.21	4384.18	5231.71	3862.12	1761.79
P07305	H10_HUMAN	16022.59	15496.41	1886.26	6676.99	8067.29	6323.87	19236.11	14834.68	11143.40	7284.22	6775.11	4464.25	13250.15	3174.75	11064.78	7384.66	6792.56	5060.34
P62807;O60814;Q96A08	H2B1C_HUMAN	524856.26	508549.36	229986.59	188675.97	197032.67	186208.93	410695.38	430648.70	340413.37	176064.29	187332.00	165791.31	308805.33	200531.69	330382.07	205985.65	151066.17	142106.10
P52747	ZN143_HUMAN	8152.00	10266.28	4730.00	10243.60	5473.10	10920.90	8466.85	15202.92	3232.63	8213.80	8643.11	6776.91	8721.85	3992.77	3395.76	33719.23	6448.83	3959.73
Q16650	TBR1_HUMAN	842.82	3096.92	1265.91	1872.80	2346.56	671.30	1279.12	1657.20	2425.75	1015.38	1458.66	961.55	2291.64	1867.55	2681.51	353.14	1166.72	443.90
P84077	ARF1_HUMAN	1963.30	2201.56	473.44	610.04	435.40	640.66	1820.37	1920.62	481.65	617.49	669.30	405.00	1990.38	473.29	402.74	466.50	537.71	509.30
Q96SZ4	ZSC10_HUMAN	14436.58	23318.20	10475.14	8038.37	3920.22	7189.88	17807.40	14443.99	10596.43	4565.87	6298.49	5348.34	12182.65	7452.14	8590.18	6058.59	5001.64	3598.14
Q9Y3E1	HDGR3_HUMAN	2156.50	2554.05	1233.63	1535.98	1515.93	1390.21	2970.87	3762.71	1609.57	1390.16	900.97	973.17	2085.11	2130.81	1052.09	1431.88	1071.24	944.74
Q9P031	TAP26_HUMAN	3176.83	3130.76	1669.44	2672.91	2966.70	2354.79	5413.09	2407.38	3241.11	1889.34	2106.01	1220.84	2572.96	1665.86	2644.43	2331.21	2426.64	2385.92
P35659	DEK_HUMAN	443371.58	508621.43	246688.25	181722.04	155524.88	161165.07	439072.01	380575.31	261706.75	159473.72	153427.54	136408.49	339441.94	194831.95	244283.23	176475.06	147140.64	122100.35
P46940	IQGA1_HUMAN	374.21	29.04	1277.74	104.21	144.95	385.01	221.41	104.26	944.52	256.64	430.85	744.66	401.03	1136.31	1657.34	286.82	260.36	175.54
Q9Y5S9	RBM8A_HUMAN	12956.80	10326.68	6293.42	5981.37	2647.95	4682.02	9811.52	7403.02	4335.35	3435.86	3249.03	2493.94	6544.71	4017.53	3772.04	4626.10	3918.64	2309.90
Q5T280	C114A_HUMAN	2773.55	1266.98	1676.39	863.76	997.80	2760.81	2455.66	1572.46	2134.34	591.89	1532.28	575.96	888.03	903.99	1704.48	1771.70	979.36	1310.75
Q96EP5	DAZP1_HUMAN	19306.72	10130.77	2670.30	2977.35	1312.23	3601.34	10374.78	6037.28	2693.21	3030.66	1771.76	2601.78	8736.59	2127.71	3066.51	3882.33	1995.64	1541.98
DECOY_A6NK89	RASFA_HUMAN	146460.71	28336.48	22438.09	21255.93	16737.40	12447.99	30472.77	20816.12	17635.73	23908.25	12404.35	13675.89	53834.71	18087.02	16726.14	14255.96	15146.28	29739.88
P54274	TERF1_HUMAN	934.11	1925.24	77.25	2055.65	2719.60	4890.56	2084.31	1640.50	517.62	1623.18	490.05	1612.92	2641.72	4698.38	1003.60	915.16	389.66	1502.28
P14136	GFAP_HUMAN	55927.04	113156.39	77482.47	38178.88	33521.13	96818.01	32854.43	30864.75	72990.65	51463.03	55890.79	102510.45	36525.20	47485.23	52784.05	381028.31	45957.79	91837.16
P22626	ROA2_HUMAN	4460119.17	6807730.28	1887952.13	2742431.22	2904626.39	2260370.26	2617145.12	2620918.39	1375248.52	1215173.61	1470790.05	1453763.67	3044624.20	1934425.38	1702953.85	2735084.85	1998663.31	1704177.82
Q15024	EXOS7_HUMAN	3450.17	5053.47	1927.81	4807.96	3612.53	2832.16	3425.55	5473.16	1988.80	1841.97	2475.17	1044.41	5711.16	2212.94	2284.97	2384.15	2845.60	1776.95
DECOY_Q9C0B2	CFA74_HUMAN	24018.69	2011.05	3210.25	5089.79	2143.68	914.03	3797.92	3388.87	448.81	7168.46	3055.78	259.34	2623.61	5509.38	2047.66	8717.43	9496.15	5429.66
Q99623	PHB2_HUMAN	58618.45	48136.66	29059.28	124342.53	31407.37	26613.72	46943.41	55845.53	27423.45	34685.42	32367.55	23866.54	44099.37	189746.77	30449.29	28746.03	31459.46	28449.07
Q99590	SCAFB_HUMAN	188.31		807.51	135.52	350.46	2192.74		1657.65	255.56	276.79	1401.32		1158.69	970.36	2132.74			1261.80
Q95359	TACC2_HUMAN			2289.40	119.48		515.02	0.06		256.91	101.54	115.18	680.98	85.30	1986.12	315.11	1008.36	145.41	698.40

Table 10-1. Mass spectrometry data related to figures 7-6 and 7-7. Depicted are the normalized MS1 intensities for proteins showing more than 1.5-fold enrichment over the control sequence in any of the two conditions (N=3).

Accession	Uniprot	Fold change				p value				Fold change				log2 Fold change				-LOG10 p value			
		plurip vs Control	diff vs Control	plurip vs Control	vs Control	plurip vs Control	diff vs Control	plurip vs Control	diff vs Control	diff vs plurip	GSC diff vs plurip	vs GSC plurip	vs GSC diff	diff vs plurip	GSC diff vs plurip	vs GSC plurip	vs GSC diff	diff vs plurip	GSC diff vs plurip	vs GSC plurip	vs GSC diff
P46937	YAP1_HUMAN	1.250	3.866	3.963	7.806	0.825	0.008	0.428	0.020	12.16	7.74	0.32	0.50	3.60	2.95	-1.66	-1.01	2.69	1.86	0.32	0.97
P28347	TEAD1_HUMAN	9.607	6.719	10.375	9.880	0.004	0.004	0.003	0.004	1.01	1.38	0.93	0.68	0.01	0.46	-0.11	-0.56	0.32	1.01	0.15	0.84
P41212	ETV6_HUMAN	65.305	1357.46	0.000	115.196	0.327	0.130	0.374	0.378	1.07	#DIV/0!	#DIV/0!	11.78	0.10	#DIV/0!	#DIV/0!	3.56	0.32	0.73	0.49	0.79
P39880	CUX1_HUMAN	0.889	0.498	5.617	2.804	0.816	0.209	0.000	0.012	2.07	1.84	0.16	0.18	1.05	0.88	-2.66	-2.49	1.29	2.39	3.40	3.02
Q9UD57	NKX12_HUMAN	1.437	0.570	0.420	2.073	0.735	0.138	0.254	0.157	13.06	162.85	3.43	0.27	3.71	7.35	1.78	-1.86	1.14	1.86	0.38	1.10
P55347	PKNX1_HUMAN	4.167	56.947	2.113	13.656	0.004	0.019	0.465	0.002	2.18	1.03	1.97	4.17	1.13	0.05	0.98	2.06	1.28	0.32	0.89	1.36
Q06330	SUH_HUMAN	0.974	0.622	5.204	3.489	0.870	0.148	0.031	0.002	0.67	0.70	0.19	0.18	-0.58	-0.51	-2.42	-2.49	1.66	0.82	1.53	3.21
Q9UBG7	RBPJL_HUMAN	0.992	1.014	7.079	3.423	0.974	0.980	0.121	0.033	1.11	0.52	0.14	0.30	0.15	-0.93	-2.83	-1.76	0.44	0.76	0.92	1.71
Q9Y261	FOXA2_HUMAN	3.007	1.472	0.993	1.898	0.176	0.556	0.990	0.318	3.31	12.93	3.03	0.78	1.73	3.69	1.60	-0.37	1.06	1.54	0.82	0.18
Q15583	TGIF1_HUMAN	2.428	9.239	1.086	3.118	0.587	0.106	0.914	0.006	3.88	2.93	2.24	2.96	1.96	1.55	1.16	1.57	0.98	2.38	0.16	0.71
Q13948	CASP_HUMAN	0.690	0.917	1.911	1.909	0.224	0.741	0.343	0.013	4.73	3.56	0.36	0.48	2.24	1.83	-1.47	-1.06	2.20	2.46	0.67	2.01
Q9UNS1	TIM_HUMAN	0.251	0.188	0.296	3.839	0.223	0.106	0.218	0.139	0.47	8.22	0.85	0.05	-1.08	3.04	-0.23	-4.35	0.55	1.36	0.07	1.15
Q01860	PO5F1_HUMAN	3.010	1.835	2.598	2.320	0.014	0.049	0.076	0.019	1.03	1.51	1.16	0.79	0.04	0.59	0.21	-0.34	0.35	1.03	0.21	0.53
Q9UK61	TASOR_HUMAN	1.470	1.922	0.190	12.769	0.723	0.571	0.294	0.352	0.15	7.92	7.75	0.15	-2.70	2.99	2.95	-2.73	0.79	0.73	0.34	0.62
Q15014	MO4L2_HUMAN	0.940	0.826	0.459	3.752	0.953	0.825	0.386	0.331	0.49	4.52	2.05	0.22	-1.04	2.18	1.03	-2.18	0.51	0.81	0.20	0.51
Q5TKA1	LIN9_HUMAN	0.480	0.789	0.896	1.697	0.450	0.644	0.913	0.094	4.97	5.73	0.54	0.46	2.31	2.52	-0.90	-1.11	1.08	2.19	0.19	1.03
P05204	HMGN2_HUMAN	1.477	1.069	2.591	2.230	0.506	0.858	0.307	0.440	0.78	0.93	0.57	0.48	-0.35	-0.10	-0.81	-1.06	0.48	0.33	0.32	0.32
Q15406	NR6A1_HUMAN	2.024	2.470	1.880	1.935	0.008	0.008	0.014	0.000	1.75	1.48	1.08	1.28	0.81	0.56	0.11	0.35	2.58	2.46	0.24	1.57
P63104	1433Z_HUMAN	0.525	0.869	0.807	2.283	0.338	0.681	0.630	0.144	1.64	2.80	0.65	0.38	0.71	1.49	-0.62	-1.39	0.62	1.23	0.30	0.86
Q06416;P20264	P5F1B_HUMAN	2.064	1.770	1.855	1.727	0.015	0.276	0.188	0.287	1.62	1.76	1.11	1.02	0.70	0.82	0.15	0.04	0.92	0.94	0.15	0.02
P49640	EVX1_HUMAN	1.042	1.803	0.850	1.703	0.885	0.334	0.493	0.059	3.17	3.67	1.23	1.06	1.66	1.88	0.29	0.08	1.11	3.03	0.27	0.05
Q96JM7	LMBL3_HUMAN	0.538	1.222	0.852	1.905	0.425	0.685	0.864	0.059	2.98	2.93	0.63	0.64	1.58	1.55	-0.66	-0.64	1.06	1.33	0.17	0.57
Q7Z2T5	TRM1L_HUMAN	0.753	0.770	2.048	2.065	0.595	0.656	0.268	0.074	1.07	1.06	0.37	0.37	0.10	0.08	-1.44	-1.42	0.36	0.36	0.83	1.95
O75461	E2F6_HUMAN	0.617	1.352	0.413	2.293	0.592	0.697	0.259	0.453	1.82	4.62	1.49	0.59	0.87	2.21	0.58	-0.76	0.53	0.80	0.14	0.21
Q14562	DHX8_HUMAN	0.861	0.238	0.161	4.209	0.820	0.137	0.155	0.322	0.09	8.56	5.36	0.06	-3.47	3.10	2.42	-4.14	1.40	0.89	0.93	0.63
P62834	RAP1A_HUMAN	0.704	0.163	0.304	1.931	0.754	0.457	0.458	0.566	0.25	6.98	2.32	0.08	-1.98	2.80	1.21	-3.57	0.64	0.72	0.19	0.64
P31941	ABC3A_HUMAN	0.762	1.131	0.572	2.234	0.754	0.828	0.484	0.149	1.19	3.13	1.33	0.51	0.25	1.65	0.41	-0.98	0.38	1.28	0.13	0.56
P23511	NFYA_HUMAN	1.732	0.831	2.585	1.467	0.566	0.785	0.190	0.524	0.99	1.17	0.67	0.57	-0.01	0.23	-0.58	-0.82	0.31	0.44	0.22	0.82
Q9H2W2;O75360	MIXL1_HUMAN	1.564	1.151	0.956	1.040	0.479	0.674	0.886	0.914	11.48	16.97	1.64	1.11	3.52	4.08	0.71	0.15	2.16	1.98	0.36	0.11
Q05BQ5	MBTD1_HUMAN	0.782	0.900	0.514	2.659	0.725	0.852	0.318	0.023	0.66	2.96	1.52	0.34	-0.60	1.56	0.61	-1.56	0.53	1.75	0.23	1.41
Q9UPT8	ZC3H4_HUMAN	1.439	1.295	1.101	1.656	#DIV/0!	#DIV/0!	#DIV/0!	0.057	1.25	2.08	1.31	0.78	0.32	1.06	0.39	-0.35	#DIV/0!	#DIV/0!	#DIV/0!	#DIV/0!
Q2NKX8	ERC6L_HUMAN	1.001	1.739	1.311	1.586	0.999	0.231	0.697	0.526	2.62	1.83	0.76	1.10	1.39	0.87	-0.39	0.13	2.10	0.71	0.27	0.07
Q9H8Y1	VRTN_HUMAN	0.909	0.961	1.677	1.466	0.688	0.881	0.158	0.193	1.95	1.61	0.54	0.66	0.96	0.69	-0.88	-0.61	1.70	1.19	0.87	0.97
Q8IZX4	TAF1L_HUMAN	1.337	1.788	2.202	1.886	0.282	0.079	0.169	0.016	1.35	0.86	0.61	0.95	0.43	-0.21	-0.72	-0.08	0.88	0.46	0.51	0.13
Q8WUY9	DEP1B_HUMAN	0.815	0.360	0.291	2.021	#DIV/0!	#DIV/0!	0.137	0.483	0.37	5.76	2.80	0.18	-1.45	2.53	1.49	-2.49	#DIV/0!	0.73	#DIV/0!	#DIV/0!
Q92908	GATA6_HUMAN	1.502	1.326	1.409	1.373	0.304	0.608	0.199	0.534	1.77	1.95	1.07	0.97	0.82	0.96	0.09	-0.05	0.85	1.11	0.09	0.03
Q12800	TFCP2_HUMAN	1.257	1.212	1.781	1.414	0.524	0.355	0.141	0.120	1.76	1.45	0.71	0.86	0.82	0.54	-0.50	-0.22	1.42	1.09	0.48	0.40
P61981	1433G_HUMAN	0.872	0.852	0.714	1.627	0.758	0.636	0.300	0.138	1.23	2.87	1.22	0.52	0.30	1.52	0.29	-0.93	0.46	1.79	0.16	0.90
Q5C9Z4	NOM1_HUMAN	0.694	0.896	1.182	1.898	0.321	0.803	0.539	0.047	1.18	1.47	0.59	0.47	0.24	0.56	-0.77	-1.08	0.44	1.35	0.72	1.57
Q7Z460	CLAP1_HUMAN	0.460	0.659	0.631	1.923	0.199	0.151	0.448	0.057	1.24	2.63	0.73	0.34	0.31	1.40	-0.46	-1.54	0.95	1.60	0.25	1.84
O60293	ZC3H1_HUMAN	0.775	0.610	0.377	1.539	0.778	0.280	0.215	0.448	1.07	5.54	2.05	0.40	0.10	2.47	1.04	-1.34	0.33	1.30	0.24	0.71
Q9NP50	FA60A_HUMAN	1.662	1.424	0.209	5.802	#DIV/0!	0.726	0.067	0.374	0.15	4.91	7.94	0.25	-2.72	2.30	2.99	-2.03	#DIV/0!	0.59	#DIV/0!	0.28
Q13952	NFYC_HUMAN	1.286	1.053	1.733	1.484	0.371	0.764	0.098	0.084	1.17	1.22	0.74	0.71	0.23	0.29	-0.43	-0.50	0.66	0.75	0.52	1.21
Q07864	DPOE1_HUMAN	0.299	0.258	0.160	2.103	0.143	0.275	0.002	0.124	0.59	8.98	1.87	0.12	-0.76	3.17	0.90	-3.03	0.44	3.46	0.19	2.43
P40425	PBX2_HUMAN	4.841	4.107	6.407	0.949	0.225	0.095	0.198	0.896	7.62	1.33	0.76	4.33	2.93	0.41	-0.40	2.11	1.47	0.51	0.13	1.05
Q9HBE1	PATZ1_HUMAN	2.667	1.780	2.806	1.699	0.060	0.074	0.080	0.070	0.67	0.61	0.95	1.05	-0.58	-0.72	-0.07	0.07	0.65	0.95	0.04	0.09
Q9UBL3	ASH2L_HUMAN	1.932	0.577	0.382	2.624	#DIV/0!	0.244	0.451	0.273	0.14	3.23	5.06	0.22	-2.83	1.69	2.34	-2.19	#DIV/0!	0.73	#DIV/0!	0.51
Q9NZI6	TF2L1_HUMAN	1.255	1.695	2.227	1.507	0.763	0.415	0.211	0.444	1.66	0.83	0.56	1.12	0.73	-0.26	-0.83	0.17	0.67	0.48	0.55	0.09
P26232	CTNA2_HUMAN	0.536	1.383	0.345	1.862	0.335	0.725	0.094	0.352	1.97	4.11	1.55	0.74	0.98	2.04	0.64	-0.43	0.55	1.10	0.19	0.14
Q9NVN8	GNL3L_HUMAN	0.720	3.994	0.483	1.584	0.176	0.307	0.033	0.190	5.70	3.37	1.49	2.52	2.51	1.75	0.57	1.33	0.88	2.19	0.65	0.40
Q9BQ67	GRWD1_HUMAN	0.791	1.077	0.136	2.524	0.798	0.936	0.088	0.255	0.63	8.63	5.82	0.43	-0.66	3.11	2.54	-1.23	0.41	1.25	0.47	0.37
P29692	EF1D_HUMAN	2.588	0.653	1.133	1.676	0.220	0.287	0.707	0.603	0.23	1.32	2.28	0.39	-2.15	0.40	1.19	-1.36	1.13	0.46	0.59	0.43
Q8NEG4	FA83F_HUMAN	0.431	0.925	0.265	1.681	0.418	0.772	0.233	0.255	1.86	5.48	1.62	0.55	0.89	2.45	0.70	-0.86	0.66	1.51	0.15	0.61
Q9NZI7	UBIP1_HUMAN	1.326	1.494	1.626	1.213	0.511	0.207	0.146	0.406	1.79	1.33	0.82	1.10	0.84	0.41	-0.29	0.14	1.22	0.81	0.22	0.19
Q9NSI2	F207A_HUMAN	1.559	1.744	1.208	1.107	0.296	0.279	0.556	0.744	2.59	2.13	1.29	1.58	1.38	1.09	0.37	0.66	1.16	1.43	0.29	0.46
Q9Y2X9	ZN281_HUMAN	0.605	1.227	0.607	1.633	0.349	0.290	0.245	0.054	1.80	2.39	1.00	0.75	0.85	1.26	-0.01	-0.41	0.98	1.81	0.00	0.68
Q96QE3	ATAD5_HUMAN	0.573	0.662	0.258	2.024	0.532	0.405	0.233	0.255	0.71	4.80	2.22	0.33	-0.50	2.26	1.15	-1.61	0.45	1.36	0.31	0.81
Q9UN81	LORF1_HUMAN	1.010	1.603	1.022	1.747	0.979	0.047	0.943	0.003	1.22	1.31	0.99	0.92	0.29	0.39	-0.02	-0.12	0.72	1.59	0.02	0.31
P11166	GTR1_HUMAN	1.004	1.050	0.996	1.701	0.988	0.635	0.991	0.166	0.84	1.37	1.01	0.62	-0.25	0.45	0.01	-0.70	0.74	0.70	0.01	0.74
Q9UJU2	LEF1_HUMAN	3.161	2.496	1.568	1.684	0.075	0.023	0.192	0.017	0.64	0.87	2.02	1.48	-0.65	-0.21	1.01	0.57	0.84	0.55	0.78	0.95
Q8NBU5	ATAD1_HUMAN	1.302	1.362	1.209	1.544	0.413	0.642	0.596	0.440	0.99	1.20	1.08	0.88	-0.02	0.11	0.11	-0.18	0.31	0.55	0.07	0.15
Q15723	ELF2_HUMAN	1.885	1.444	1.668	1.194	0.371	0.047	0.378	0.563	1.29	1.21	1.13	1.21	0.37	0.27	0.18	0.27	0.60	0.50	0.09	0.33
Q01831	XPC_HUMAN	0.838	1.573	1.359	1.242	0.760	0.042	0.399	0.618	2.84	1.37	0.62	1.27	1.50	0.46	-0.70	0.34	0.99	0.74	0.43	0.19
Q8TD26	CHD6_HUMAN	1.693	1.098	1.435	1.337	0.031	0.637	0.067	0.165	0.81	1.17	1.18	0.82	-0.30	0.23	0.24	-0.28	0.92	0.82	0.62	0.54
O95347	SMC2_HUMAN	0.495	0.551	0.503	1.650	0.183	0.330	0.018	0.362	0.91	2.68	0.98	0.33	-0.14	1.42	-0.02	-1.58	0.35	1.20	0.01	0.95
Q86U86	PB1_HUMAN	0.806	0.427	0.300	3.935	0.842	0.291	0.275	0.285	0.12	3.03	2.68	0.11	-3.03	1.60	1.42	-3.20	0.79	0.75	0.30	0.68
Q04917	1433F_HUMAN	0.617	0.931	0.401	1.712	0.491	0.922	0.178	0.358	1.13	3.19	1.54	0.54	0.17	1.68	0.62	-0.88	0.37	1.79	0.19	1.14
Q9ULW3	ABT1_HUMAN	1.898	1.907	0.845	1.547	0.107	0.087	0.673	0.222	0.85	1.54	2.25	1.23	-0.24	0.62	1.17	0.30	0.67	1.65	1.57	0.70
P16403	H12_HUMAN	1.898	1.006	5.030	1.060	0.478	0.988	0.087	0.826	1.04	0.42	0.38	0.95	0.06	-1.27	-1.41	-0.08	0.32	1.07	0.69	0.07
Q14119	VEZF1_HUMAN	1.627	1.331	1.416	1.382	0.091	0.203	0.095	0.086	0.85	1.01	1.15	0.96	-0.24	0.02	0.20	-0.06	0.64	0.33	0.29	0.10
P14859;P09086;Q9UKI9	PO2F1_HUMAN	1.655	1.442	0.792	0.974	0.091	0.087	0.434	0.812	2.16	3.05	2.09	1.48	1.11	1.61	1.06	0.57	2.11	3.59	1.66	1.24
P60953	CDC42_HUMAN	0.758	1.106	0.924	1.507	0.114	0.737	0.717	0.171	1.24	1.39	0.82	0.73	0.31	0.47	-0.29	-0.45	0.64	0.88	0.35	0.51
Q99959	PKP2_HUMAN	0.855	2.087	1.180	1.195	0.834	0.592	0.765	0.782	3.31	1.37	0.72	1.75	1.73	0.46	-0.47	0.80	0.62	0.53	0.25	0.18
P17480	UBF1_HUMAN	1.445	1.367	1.628	1.244	0.371	0.180	0.136	0.469	1.17	0.94	0.89	1.10	0.23	-0.08	-0.17	0.14	0.53	0.38	0.14	0.16
Q5T5X7	BEND3_HUMAN	1.143	1.047	1.574	1.240	0.453	0.857	0.006	0.245	1.10	0.95	0.73	0.84	0.14	-0.07	-0.46	-0.24	0.46	0.52	1.25	0.36
Q86TJ2	TAD2B_HUMAN	0.819	0.064	0.224	2.778	0.849	0.404	0.225	0.488	0.03	3.98	3.66	0.02	-5.32	1.99	1.87	-5.44	0.68	0.68	0.37	0.39
Q96GN5	CDA7L_HUMAN	1.155	1.620	1.247	1.149	0.551	0.095	0.467	0.636	1.99	1.30	0.93	1.41	0.99	0.38	-0.11	0.50	1.66	0.66	0.12	0.57
Q9GZR2	REXO4_HUMAN	0.381	1.078	0.456	2.470	0.347	0.907	0.457	0.082	1.06	2.02	0.84	0.44	0.08	1.01	-0.26	-1.20	0.32	0.82	0.05	0.88

Continued from page 103.

Accession	Uniprot	Fold change				p value				Fold change				log2 Fold change				- LOG10 p value			
		NANOG plurip vs Control	NANOG diff vs Control	GSC plurip vs Control	GSC diff vs Control	NANOG plurip vs Control	NANOG diff vs Control	GSC plurip vs Control	GSC diff vs Control	NANOG diff vs plurip	GSC diff vs plurip	NANOG vs GSC plurip	NANOG vs GSC diff	NANOG diff vs plurip	GSC diff vs plurip	NANOG vs GSC plurip	NANOG vs GSC diff	NANOG diff vs plurip	GSC diff vs plurip	NANOG vs GSC plurip	NANOG vs GSC diff
Q92616	GCN1_HUMAN	0.559	0.902	0.750	1.879	0.056	0.827	0.090	0.092	0.89	1.39	0.74	0.48	-0.16	0.47	-0.43	-1.06	0.41	1.22	0.41	1.60
O14880	MGST3_HUMAN	0.859	0.685	0.937	1.580	0.742	0.678	0.833	0.482	0.56	1.18	0.92	0.43	-0.84	0.24	-0.13	-1.21	0.69	0.49	0.06	1.05
P33993	MCM7_HUMAN	0.774	1.683	1.044	1.304	0.050	0.203	0.753	0.061	2.12	1.22	0.74	1.29	1.08	0.28	-0.43	0.37	1.22	0.96	1.10	0.34
P13674	P4HA1_HUMAN	0.173	0.654	0.157	1.804	0.254	0.559	0.239	0.215	2.08	6.33	1.10	0.36	1.06	2.66	0.14	-1.46	0.58	1.74	0.03	1.19
Q3YBR2	TBRG1_HUMAN	1.802	1.962	1.542	1.283	0.203	0.136	0.412	0.600	1.08	0.82	1.17	1.53	0.11	-0.28	0.23	0.61	0.39	0.46	0.14	0.54
Q8N9M1	CS047_HUMAN	0.435	1.522	0.367	1.284	0.125	0.518	0.011	0.399	3.46	3.46	1.19	0.25	1.79	1.79	0.25	0.25	0.98	3.59	0.09	0.13
Q9H160	ING2_HUMAN	1.245	1.238	1.700	1.193	0.009	0.245	0.038	0.285	1.09	0.77	0.73	1.04	0.12	-0.38	-0.45	0.05	0.68	1.05	0.91	0.14
Q5M775	CYTSB_HUMAN	0.503	0.835	0.250	1.748	0.649	0.761	0.358	0.527	0.89	3.73	2.01	0.48	-0.17	1.90	1.01	-1.07	0.35	0.85	0.27	0.36
P16401	H15_HUMAN	1.516	1.247	1.530	1.208	0.524	0.393	0.310	0.484	0.81	0.78	0.99	1.03	-0.30	-0.36	-0.01	0.05	0.47	0.76	0.01	0.04
P25208	NFYB_HUMAN	1.469	1.169	1.946	0.962	0.454	0.690	0.254	0.905	1.31	0.82	0.75	1.22	0.40	-0.29	-0.41	0.28	0.54	0.48	0.23	0.19
Q86VM9	ZCH18_HUMAN	1.436	1.621	1.539	1.020	0.631	0.204	0.438	0.940	1.56	0.92	0.93	1.59	0.65	-0.12	-0.10	0.67	0.68	0.38	0.04	0.63
Q9UHX1	PUF60_HUMAN	2.519	1.279	1.478	1.046	0.300	0.614	0.603	0.922	0.65	0.90	1.70	1.22	-0.62	-0.15	0.77	0.29	0.58	0.36	0.33	0.15
Q02447	SP3_HUMAN	1.922	0.939	1.894	1.168	0.126	0.736	0.172	0.217	0.49	0.62	1.02	0.80	-1.03	-0.69	0.02	-0.31	1.31	0.95	0.01	0.59
Q7RTV0	PHF5A_HUMAN	1.739	1.242	1.588	1.091	0.111	0.391	0.145	0.627	0.81	0.78	1.10	1.14	-0.30	-0.36	0.13	0.19	0.62	0.75	0.12	0.23
P15924	DESP_HUMAN	1.872	1.034	1.486	0.815	0.283	0.703	0.133	0.053	1.18	1.17	1.26	1.27	0.24	0.23	0.33	0.34	0.49	1.09	0.22	1.54
P11388	TOP2A_HUMAN	1.352	1.327	1.529	1.205	0.492	0.290	0.283	0.523	0.83	0.66	0.88	1.10	-0.28	-0.59	-0.18	0.14	0.53	0.94	0.14	0.16
Q8N3C0	ASCC3_HUMAN	2.272	1.021	1.491	1.101	0.283	0.960	0.576	0.679	0.45	0.73	1.52	0.93	-1.16	-0.45	0.61	-0.11	0.88	0.55	0.29	0.07
Q92522	H1X_HUMAN	1.766	1.172	1.922	1.306	0.293	0.455	0.058	0.188	0.48	0.49	0.92	0.90	-1.06	-1.03	-0.12	-0.16	0.99	1.78	0.08	0.41
P50402	EMD_HUMAN	1.567	1.258	1.574	1.119	0.238	0.254	0.106	0.458	0.76	0.67	1.00	1.12	-0.40	-0.57	-0.01	0.17	0.73	1.61	0.01	0.45
P50151	GBG10_HUMAN	1.287	1.543	1.334	0.693	0.434	0.009	0.289	0.086	4.34	1.88	0.96	2.23	2.12	0.91	-0.05	1.15	4.03	1.78	0.06	2.80
P40424	PBX1_HUMAN	2.667	3.752	1.622	0.971	0.044	0.019	0.251	0.945	1.78	0.76	1.64	3.87	0.83	-0.40	0.72	1.95	1.28	0.60	0.75	1.71
Q15717	ELAV1_HUMAN	1.477	1.122	1.846	1.065	0.398	0.306	0.160	0.747	0.78	0.60	0.80	1.05	-0.35	-0.75	-0.32	0.08	0.59	1.04	0.23	0.12
Q9UIF9	BAZ2A_HUMAN	0.926	0.705	1.552	1.368	0.769	0.172	0.328	0.199	0.49	0.57	0.60	0.51	-1.03	-0.81	-0.74	-0.96	3.04	0.97	0.65	1.45
Q9UIG0	BAZ1B_HUMAN	1.322	1.202	1.686	1.221	0.558	0.433	0.195	0.495	0.71	0.57	0.78	0.98	-0.49	-0.82	-0.35	-0.02	0.68	1.21	0.27	0.03
Q86U42	PABP2_HUMAN	1.679	1.260	1.130	0.992	0.098	0.608	0.592	0.985	0.87	1.02	1.49	1.27	-0.20	0.03	0.57	0.35	0.48	0.33	0.90	0.26
Q5SSJ5	HP1B3_HUMAN	1.796	1.398	1.667	1.084	0.346	0.371	0.202	0.812	0.72	0.60	1.08	1.29	-0.47	-0.73	0.11	0.37	0.57	1.15	0.06	0.34
Q14674	ESPL1_HUMAN	1.169	1.109	0.434	3.179	#DIV/0!	0.874	#DIV/0!	0.323	0.19	1.44	2.69	0.35	-2.43	0.52	1.43	-1.52	#DIV/0!	#DIV/0!	#DIV/0!	0.32
P84090	ERH_HUMAN	1.640	1.028	1.569	0.898	0.279	0.931	0.147	0.743	0.83	0.76	1.05	1.15	-0.26	-0.39	0.06	0.20	0.51	0.90	0.04	0.61
Q9GZL7	WDR12_HUMAN	0.908	1.503	1.107	1.062	0.443	0.262	0.472	0.845	1.54	0.89	0.82	1.42	0.62	-0.17	-0.29	0.50	0.97	0.48	0.76	0.44
Q9Y3Y2	CHTOP_HUMAN	1.741	0.825	1.628	0.787	0.182	0.787	0.170	0.733	0.86	0.88	1.07	1.05	-0.22	-0.19	0.10	0.07	0.47	0.50	0.07	0.06
A6NHQ2	FBLL1_HUMAN	0.730	1.955	1.189	1.147	0.032	0.145	0.241	0.534	2.12	0.76	0.61	1.70	1.09	-0.39	-0.70	0.77	1.26	1.13	1.36	0.74
Q8IWS0	PHF6_HUMAN	1.531	1.193	1.393	0.994	0.129	0.193	0.101	0.956	0.81	0.74	1.10	1.20	-0.30	-0.43	0.14	0.26	0.74	1.22	0.16	1.06
P18754	RCC1_HUMAN	1.508	1.095	1.367	1.067	0.276	0.333	0.142	0.466	0.64	0.69	1.10	1.03	-0.65	-0.54	0.14	0.04	0.96	1.75	0.13	0.20
O15160	RPAC1_HUMAN	2.507	0.340	0.274	3.283	#DIV/0!	0.257	0.292	0.378	0.02	2.15	9.14	0.10	-5.36	1.10	3.19	-3.27	#DIV/0!	0.59	#DIV/0!	0.58
O95400	CD2B2_HUMAN	0.761	1.843	0.885	1.128	0.253	0.177	0.661	0.760	1.91	1.00	0.86	1.63	0.93	0.01	-0.22	0.71	1.27	0.31	0.22	0.71
P06493	CDK1_HUMAN	1.002	1.533	1.119	0.910	0.994	0.376	0.673	0.772	1.83	0.98	0.90	1.69	0.88	-0.04	-0.16	0.75	0.95	0.36	0.19	0.59
Q6RFH5	WDR74_HUMAN	1.065	1.554	1.235	0.964	0.798	0.010	0.023	0.661	1.53	0.82	0.86	1.61	0.61	-0.29	-0.21	0.69	1.29	1.39	0.28	1.86
Q53F19	NCBP3_HUMAN	1.196	1.271	1.630	1.094	0.649	0.510	0.166	0.783	0.85	0.54	0.73	1.16	-0.24	-0.90	-0.45	0.22	0.46	1.29	0.37	0.22
DECOY_P 29353	SHC1_HUMAN	1.501	1.101	1.279	1.026	0.350	0.507	0.518	0.866	0.66	0.72	1.17	1.07	-0.61	-0.48	0.23	0.10	0.85	0.79	0.16	0.53
Q9ULU4	PKCB1_HUMAN	1.183	1.033	1.589	0.963	0.412	0.906	0.139	0.894	0.86	0.60	0.74	1.07	-0.21	-0.74	-0.43	0.10	0.66	1.26	0.50	0.36
		0.472	0.000	0.210	1.855	#DIV/0!	#DIV/0!	0.270	0.344	0.00	3.06	2.25	0.00	#NUM!	1.62	1.17	#NUM!	#DIV/0!	0.81	#NUM!	#NUM!
O60828	PQBP1_HUMAN	1.084	1.606	1.054	1.044	0.743	0.180	0.860	0.871	1.24	0.83	1.03	1.54	0.32	-0.26	0.04	0.62	0.61	0.52	0.03	0.60
Q7L7X3	TAOK1_HUMAN	1.373	0.819	1.688	0.928	0.584	0.643	0.400	0.857	0.60	0.55	0.81	0.88	-0.74	-0.86	-0.30	-0.18	0.75	0.80	0.16	0.09
Q8N567	ZCHC9_HUMAN	1.192	1.531	1.211	0.942	0.561	0.070	0.545	0.721	1.24	0.75	0.98	1.63	0.31	-0.42	-0.02	0.70	0.66	0.72	0.02	1.12
P06239;O 94921;P5 1451	LCK_HUMAN	1.015	0.710	1.605	0.733	0.939	0.056	0.347	0.058	1.17	0.76	0.63	0.97	0.22	-0.39	-0.66	-0.05	0.81	0.58	0.46	0.11
Q13601	KRR1_HUMAN	0.779	1.501	0.886	1.083	0.175	0.045	0.417	0.163	1.41	0.90	0.88	1.39	0.50	-0.16	-0.19	0.47	1.36	1.03	0.57	1.11
O60481	ZIC3_HUMAN	1.588	1.268	1.791	1.012	0.384	0.111	0.298	0.922	0.66	0.47	0.89	1.25	-0.60	-1.10	-0.17	0.33	0.76	1.06	0.10	0.73
P18583	SON_HUMAN	1.574	1.083	1.170	1.021	0.221	0.513	0.593	0.840	0.55	0.70	1.35	1.06	-0.85	-0.51	0.43	0.08	1.20	0.95	0.40	0.20
Q7Z7K6	CENPV_HUMAN	1.436	1.579	1.593	1.495	0.655	0.322	0.381	0.293	0.48	0.41	0.90	1.06	-1.07	-1.29	-0.15	0.08	0.68	1.25	0.06	0.08
O14776	TCRG1_HUMAN	1.576	0.915	1.275	0.905	0.094	0.090	0.273	0.129	0.58	0.71	1.24	1.01	-0.79	-0.50	0.31	0.02	1.50	1.11	0.38	0.08
Q15361	TTF1_HUMAN	2.087	1.418	1.250	0.887	0.243	0.267	0.600	0.588	0.70	0.73	1.67	1.60	-0.52	-0.45	0.74	0.68	0.63	0.73	0.45	0.82
Q9Y6A4	CFA20_HUMAN	0.841	1.679	0.803	0.874	0.621	0.160	0.535	0.274	2.12	1.15	1.05	1.92	1.08	0.20	0.07	0.94	1.31	0.63	0.06	0.97
Q5BKZ1	ZN326_HUMAN	1.673	0.702	1.496	0.709	0.201	0.495	0.186	0.517	0.72	0.81	1.12	0.99	-0.47	-0.30	0.16	-0.01	0.82	0.69	0.15	0.02
P67809	YBOX1_HUMAN	1.891	1.246	1.574	0.880	0.220	0.301	0.173	0.513	0.68	0.57	1.20	1.42	-0.56	-0.80	0.26	0.50	0.74	1.42	0.19	1.18
P49711;Q 8NI51	CTCF_HUMAN	1.502	1.170	1.329	1.112	0.347	0.493	0.223	0.611	0.52	0.56	1.13	1.05	-0.95	-0.85	0.18	0.07	1.03	1.73	0.13	0.15
Q96NC0	ZMAT2_HUMAN	1.646	1.314	1.404	0.977	0.332	0.202	0.109	0.853	0.66	0.58	1.17	1.34	-0.60	-0.80	0.23	0.43	0.70	1.82	0.15	0.85
P55081	MFAP1_HUMAN	1.583	1.181	1.380	0.905	0.169	0.475	0.171	0.576	0.70	0.62	1.15	1.31	-0.51	-0.69	0.20	0.38	0.86	1.48	0.20	0.70
DECOY_Q 9P219	DAPLE_HUMAN	1.620	1.819	1.070	1.072	0.264	0.139	0.818	0.768	0.77	0.69	1.51	1.70	-0.38	-0.55	0.60	0.76	0.62	1.38	0.61	0.77
O00541	PESC_HUMAN	1.533	0.883	1.225	0.801	0.041	0.614	0.600	0.303	0.66	0.75	1.25	1.10	-0.59	-0.41	0.32	0.14	1.37	0.24	0.88	0.22
Q9UNQ2	DIM1_HUMAN	1.179	1.538	0.940	0.933	0.522	0.231	0.705	0.628	1.09	0.83	1.25	1.65	0.12	-0.27	0.33	0.72	0.40	1.13	0.43	0.76
Q86Y07	VRK2_HUMAN	0.816	0.896	1.809	1.043	0.497	0.725	0.430	0.882	0.75	0.39	0.45	0.86	-0.42	-1.35	-1.15	-0.22	0.69	0.83	0.46	0.48
Q15287	RNPS1_HUMAN	1.571	1.437	1.455	1.007	0.406	0.301	0.061	0.970	0.65	0.49	1.08	1.43	-0.62	-1.02	0.11	0.51	0.66	1.22	0.06	0.56
Q8N5F7	NKAP_HUMAN	1.361	1.841	1.295	0.992	0.504	0.331	0.495	0.972	0.98	0.56	1.05	1.86	-0.03	-0.85	0.07	0.89	0.31	1.01	0.04	0.48
Q9UQ35	SRRM2_HUMAN	1.579	1.314	1.224	1.045	0.265	0.459	0.517	0.833	0.55	0.57	1.29	1.26	-0.86	-0.82	0.37	0.33	0.95	1.10	0.26	0.33
O43251	RFOX2_HUMAN	1.123	1.582	1.037	0.937	0.723	0.072	0.916	0.727	1.11	0.71	1.08	1.69	0.15	-0.49	0.12	0.76	0.46	0.78	0.08	1.34
Q8ND82	Z280C_HUMAN	2.120	1.473	2.410	0.996	0.321	0.439	0.253	0.993	0.49	0.29	0.88	1.48	-1.03	-1.78	-0.19	0.56	0.81	1.12	0.08	0.38
P0C1Z6	TFPT_HUMAN	1.178	1.568	1.118	0.893	0.647	0.142	0.657	0.747	1.12	0.67	1.05	1.76	0.17	-0.57	0.07	0.81	0.45	0.96	0.05	1.59
P07305	H10_HUMAN	1.215	1.095	1.645	0.963	0.740	0.525	0.200	0.841	0.63	0.41	0.74	1.14	-0.67	-1.29	-0.44	0.19	0.67	1.93	0.31	0.35
P62807;O 60814;Q9 6A08	H2B1C_HUMAN	1.505	1.146	1.407	1.060	0.245	0.297	0.079	0.657	0.45	0.45	1.07	1.08	-1.14	-1.16	0.10	0.11	1.43	3.13	0.10	0.95
P52747	ZN143_HUMAN	1.437	0.604	1.670	0.536	0.371	0.580	0.403	0.514	1.15	1.88	0.86	0.53	0.20	-0.19	-0.22	0.17	0.49	0.41	0.12	0.22
Q16650	TBR1_HUMAN	0.761	2.490	0.784	1.749	0.497	0.157	0.297	0.179	0.94	0.64	0.97	1.42	-0.09	-0.64	-0.04	0.51	0.34	1.10	0.02	0.39
P84077	ARF1_HUMAN	1.618	1.114	1.473	1.118	0.474	0.440	0.551	0.516	0.36	0.40	1.10	1.00	-1.46	-1.32	0.14	0.00	1.14	1.13	0.07	0.01
Q96SZ4	ZSC10_HUMAN	1.709	1.306	1.518	1.106	0.175	0.358	0.126	0.584	0.40	0.38	1.13	1.18	-1.33	-1.40	0.17	0.24	1.44	2.14	0.15	0.29
Q9Y3E1	HDGR3_HUMAN	1.128	1.288	1.584	0.947	0.691	0.096	0.228	0.786	0.75	0.39	0.71	1.36	-0.42	-1.35	-0.49	0.44	0.87	1.53	0.47	1.16
Q9P031	TAP26_HUMAN	1.159	1.119	1.607	0.730	0.568	0.188	0.216	0.074	1.00	0.47	0.72	1.53	0.00	-1.08	-0.47	0.62	0.30	1.28	0.43	1.35
P35659	DEK_HUMAN	1.540	1.118	1.389	1.008	0.192	0.375	0.208	0.948	0.42	0.42	1.11	1.11	-1.27	-1.27	0.15	0.15	1.68	2.09	0.15	0.71
P46940	IQGA1_HUMAN	0.526	0.877	0.398	1.982	0.388	0.769	0.227	0.182	0.38	1.13	1.32	0.44	-1.41	0.17	0.40	-1.18	0.68	0.36	0.11	0.73
Q9Y5S9	RBM8A_HUMAN	2.063	1.226	1.503	0.846	0.076	0.528	0.256	0.494	0.45	0.43	1.37	1.45	-1.15	-1.23	0.46	0.54	1.48	1.50	0.46	0.61
Q5T280	CI114_HUMAN	1.635	1.138	1.762	0.665	0.231	0.789	0.076	0.310	0.81	0.44	0.93	1.71	-0.31	-1.19	-0.11	0.78	0.48	1.62	0.10	0.39
Q96EP5	DAZP1_HUMAN	2.305	1.063	1.371	0.998	0.312	0.882	0.600	0.995	0.25	0.39	1.68	1.07	-2.02	-1.37	0.75	0.09	1.07	1.10	0.34	0.07
DECOY_A 6NK89	RASFA_HUMAN	2.225	0.853	0.778	0.845	0.439	0.633	0.633	0.648	0.26	0.73	2.86	1.01	-1.97	-0.46	1.52	0.01	0.83	0.82	0.45	0.01
P54274	TERF1_HUMAN	0.352	3.443	0.508	1.327	0.206	0.067	0.365	0.569	3.28	0.69	0.69	2.59	1.72	-0.19	-0.53	1.38	1.35	0.40	0.24	0.99
P14136	GFAP_HUMAN	1.788	0.325	0.999	0.405	0.109	0.336	0.999	0.387	0.69	1.54	1.79	0.80	-0.54	0.62	0.84	-0.32	0.71	0.80	0.76	0.20
P22626	ROA2_HUMAN	1.969	1.228	0.990	0.643	0.219	0.248	0.971	0.073	0.60	0.63	1.99	1.91	-0.73	-0.68	0.99	0.93	0.84	1.21	0.67	2.40
Q15024	EXOS7_HUMAN	1.022	1.606	1.066	0.765	0.962	0.096	0.890	0.348	1.08	0.49	0.96	2.10	0.11	-1.02	-0.06	1.07	0.39	1.08	0.04	1.30
DECOY_Q 9C0B2	CFA74_HUMAN	2.872	0.345	0.750	0.443	0.429	0.042	0.603	0.137	0.28	1.37	3.83	0.78	-1.84	0.46	1.94	-0.36	0.71	0.46	0.43	0.12
Q99623	PHB2_HUMAN	0.514	2.057	0.493	1.026	0.454	0.382	0.436	0.836	1.34	0.70	1.04	2.01	0.43	-0.52	0.06	1.00	0.48	0.96	0.05	0.40
Q99590	SCAFB_HUMAN	0.468	0.526	1.557	0.380	0.222	0.439	#DIV/0!	0.171	1.79	0.39	0.30	1.39	0.84	-1.36	-1.74	0.47	0.47	#DIV/0!	#DIV/0!	0.12
O95359	TACC2_HUMAN	2.878	0.514	0.162	0.485	#DIV/0!	0.472	0.463	0.455	0.14	2.33	17.82	1.06	-2.85	1.22	4.16	0.08	#DIV/0!	0.55	#DIV/0!	0.02

Table 10-2. Analysis of Mass spectrometry data related to figures 7-6 and 7-7. Listed are only the proteins from Table 10-1 with more than 1.5-fold enrichment over the control sequence in any of the two conditions.

11 Acknowledgements

I would like to thank Professor Anton Wutz for giving me the opportunity to be part of his lab and for the helpful suggestions and precious feedback on my thesis.

I would like to thank Prof. Jacob Corn for giving me the opportunity to continue my PhD in his lab and for the tremendous support on my project. I am delayed to have met such a great mentor.

I would like to express my gratitude to Dr. Tobias Beyer for his great supervision, trust in my work, for believing in me and for giving me the freedom to explore new ideas. I am very grateful to him. I would also like to thank Prof. Constance Ciaudo for her invaluable suggestions on my project.

I would also like to thank my committee members Prof. Kutay, Dr. Gstaiger, Prof. David for all the fruitful discussions we had during our meetings.

I am very grateful to the amazing people I had the fortune to meet in the Wutz and the Corn lab. Each of them made my PhD journey very enjoyable. A special thanks goes to Asun, Eishi, Ferdi, Giulio and Charlie for all the fun we had during this time, to Peter my desk neighbour, for introducing me to the lab, and for "loving my music selection". Despite his questionable "temperature taste", he has been a great lab mate.

I am also very thankful to all the people from the Corn lab Jenny, Marija, Amos, Mandy, Matthias, Charles, Erman, Zac, Matteo, who taught me many new methods after moving to the new lab, and showed me their support during the last and most difficult period of my PhD journey. I am grateful to Lily and Markus for their support on my project.

I would also like to thank past colleagues and especially my friend Paul, Hasan and Max who always gave valuable feedback on my project over a drink or two.

I am very grateful to my friends that have been by my side throughout this long and difficult journey, and especially Dimitra, Madeleine, Gabi, Uli, Simona, Marco, Spyros S., Giannis, Gigi, Nikolett, Stefania, Karun and Spyros P.

Acknowledgements

I am extremely thankful to my parents Eleni and Grigoris and my sister Georgia that have always believed in me and supported me in every aspect of my educational journey, from financial to emotional perspective. They always encourage me to do my best even when I fail, and that is OK to fail but important is to come back stronger. They always supported my decision to work and study abroad away from them if the opportunities are better for me.

A very special thanks goes to Gabriele. I have no words to explain how much he contributed on helping me go through the PhD adventure. His scientific input and criticism whenever I was rehearsing my lab meeting presentations at home, was fundamental for my project's progress. Apart from that, he encouraged my idea and motivation to start the PhD in a first place, and supported me throughout the whole duration. I could not be writing this thesis now without him by my side all these years.