# A genome-wide RNAi screen identifies determinants of human embryonic stem cells

Chia, Na Yu

2012

Chia, N. Y. (2012). A genome-wide RNAi screen identifies determinants of human embryonic stem cells. Doctoral thesis, Nanyang Technological University, Singapore.

https://hdl.handle.net/10356/50498

https://doi.org/10.32657/10356/50498

NANYANG TECHNOLOGICAL UNIVERSITY

# A GENOME-WIDE RNAI SCREEN IDENTIFIES DETERMINANTS OF HUMAN EMBRYONIC STEM CELLS


**CHIA NA YU (G0700012K)**

**School of Biological Sciences**


**2012**

# A GENOME-WIDE RNAI SCREEN IDENTIFIES DETERMINANTS OF HUMAN EMBRYONIC STEM CELLS

**CHIA NA YU (G0700012K)**

School of Biological Sciences

A thesis submitted to the Nanyang Technological University

in partial fulfilment of the requirement for the degree of

Doctor of Philosophy

**2012**

# ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr NG Huck Hui for his guidance and support during my PhD journey. He has moulded me to become a better PhD student and this will prepare me to become a better scientist. With his great acumen in science as well as tough training, I eventually managed to garner a publication in a very high impact journal Nature and I greatly appreciate his efforts.

I am thankful of Dr Frederic BARD, the collaborator of my project who shared his expertise in high throughput screening procedure as well as providing his facility as well as reagents for this whole genome wide screening to be carried out. I am also grateful to the members of his lab who has made this genome-wide RNAi screen a success.

I appreciate the efforts that my co-supervisor A/P Peter Dröge has contributed during my PhD journey.

I would like to thank the fellow colleagues of my lab and other labs for their assistance. They are essential to make this project a success. Their invaluable expertise has hastened the pace of this project to completion for the eventual publication in Nature.

Most importantly, I would like to thank my family for being there for me and giving me tremendous support at all times. They are the pillars of support for me and I appreciate their love and concern for me.

I am grateful to Agency for Science, Technology and Research (A*STAR) for sponsoring my scholarship as well as Nanyang Technological University, School of Biological Sciences, for this PhD candidature opportunity

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ABBREVIATIONS

| | |
|---|---|
| ANOVA | analysis of variance |
| AP | alkaline phosphatase |
| BAC | bacteria artificial chromosome |
| bFGF | basic fibroblast growth factor |
| BMP4 | bone morphogenetic protein 4 |
| cDNA | complementary DNA |
| ChIP | chromatin immunoprecipitation |
| ChIP-chip | ChIP-on-chip |
| ChIP-PET | ChIP-pair end tags |
| ChIP-seq | ChIP sequencing |
| CR | conserved region |
| CFP | cyan fluorescent protein |
| DF1 | DharmaFect 1 |
| DMSO | dimethly sulfoxide |
| DNA | Deoxyribonucleic acid |
| ds | double strand |
| EC | embryonal carcinoma |
| eGFP | enhanced green fluorescent protein |
| EpiSCs | epiblast stem cells |

| | |
|---|---|
| ESCs | embryonic stem cells |
| FACS | flurorescent activated cell sorting |
| FGF2 | fibroblast growth factor 2 |
| GFP | green fluorescent protein |
| GO | Gene Ontology |
| gp130 | glycoprotein 130 |
| GPCR | G-protein coupled receptor |
| H3K27ac | histone 3 acetylated at lysine 27 |
| H3K36me3 | histone 3 tri-methlated at lysine 36 |
| H3K4me1 | histone 3 mono-methlated at lysine 4 |
| H3K4me2 | histone 3 di-methlated at lysine 4 |
| H3K4me3 | histone 3 tri-methlated at lysine 4 |
| H3K9me3 | histone 3 tri-methlated at lysine 9 |
| H4K20me3 | histone 4 tri-methlated at lysine 20 |
| HCS | high content screening |
| hESCs | human embryonic stem cells |
| HR | homologous recombination |
| HT | high-throughput |
| ICM | inner cell mass |
| IGF | insulin growth factor |
| IL6 | interleukin 6 |

| | |
|---|---|
| IPSCs | induced pluripotent cells |
| JAK | Janus Kinase |
| LIF | leukemia inhibitory factor |
| LIF-R | leukemia inhibitory factor receptor |
| MeDIP | methylated DNA immunoprecipitation |
| MEFs | mouse embryonic fibroblast cells |
| mESCs | mouse embryonic stem cells |
| miRNA | microRNA |
| MRE | microRNA recognition element |
| mRNA | messanger RNA transcripts |
| MTL | multiple transcrption factor loci |
| NHEJ | non-homologous end joining |
| NT | non-targeting |
| OCT4 | octamer binding transcription factor 4 |
| PCR | polymerase chain reaction |
| PET | paired-end tags |
| PGCs | primordial germ cells |
| PRDM | PRDI-BF1 and RIZ domain proteins |
| qPCR | quantitative polymerase chain reaction |
| RA | retinoic acid |
| RFP | red fluorescent protein |

| | |
|---|---|
| RISC | RNA-induced silencing complex |
| RNA | ribonucleic acid |
| RNAi | RNA interference |
| RNA-seq | RNA-sequencing |
| SAGE | serial analysis of gene expression |
| SAM | significant analysis of microarray |
| SCID | severe combined immunodeficient |
| shRNA | short hairpin RNA |
| siRNA | small interfering RNA |
| SSEA-1 | stage specific antigen 1 |
| SSEA-3 | stage specific antigen 3 |
| SSEA-4 | stage specific antigen 4 |
| STAT3 | signal transducer and activator of transcription 3 |
| TSC | trophoblast stem cells |
| UTR | untranslated region |
| YFP | yellow fluorescent protein |
| ZFN | zinc finger nucleases |

# SUMMARY

**Background:** The derivation of human embryonic cells (hESCs) from human blastocysts represents one of the milestones in stem cell biology[1]. Exploitation of the full potential of hESCs for research and clinical applications will require a detailed understanding of the genetic network that underlies the unique properties of hESCs. hESCs are pluripotent where they have the capacity to differentiate into all derivatives of the three primary germ layers: ectoderm, mesoderm and endoderm. In addition, these cells are able to self-renew continuously. Pluripotent stem cells offer the possibility of a renewable source of replacement cells and tissues to treat a myriad of diseases, conditions, and disabilities including Parkinson's disease, amyotrophic lateral sclerosis, spinal cord injury, burns, heart disease, diabetes, and arthritis. Although hESCs are thought to offer potential cures and therapies for many devastating diseases, research using them is still in its early stages and the safety for the use of hESCs is a major concern. For instance, only fully differentiated progenies of hESCs that are devoid of undifferentiated hESCs could be used for cell replacement, otherwise, side-effects such as teratomas formation may ensue.

Likewise to hESCs, mESCs are pluripotent and can self-renew indefinitely. Mouse embryonic stem cells (mESCs) were derived prior to hESCs and were studied to a much greater extent than hESCs. hESCs and mESCs share the fundamental properties of stem cells but both have contrasting properties as well. Here, we are using hESCs as the model for our study.

**Aim:** To understand what are the factors that govern the unique properties of hESCs through a genome-wide siRNA screen.

**Results:** Here, we report a genome-wide RNA interference screen that has identified genes necessary for self-renewal and pluripotency of hESCs. This is attained through the generation of *OCT4-GFP* reporter hESC line and screening them against a library of about 21,000 siRNAs in 384-well plates. Interestingly, functionally distinct complexes such as INO80 chromatin remodelling complex[2], Mediator complex[3], COP9 signalosome[4], TAF complex[5], eukaryotic initiation factor complex[6] and spliceosome complex[7] were among the 566 "hits" identified in the primary screen. Further bioinformatics analyses on these 566 "hits" suggest that majority of them are associated with transcription factors and several components of integrin signalling pathways were enriched. 200 out of the 566 genes were subjected for further validation in the secondary screen and 93 of them were found to be required for continued expression of *OCT4* in three different hESC lines. Of these 93 "hit" genes, we intensely interrogated one gene specifically and its relevance to pluripotency—the transcription factor *PRDM14*. We found that *PRDM14* is unequivocally required for the undifferentiated self-renewal of hESCs. Taken together, our study uncovers a wealth of novel regulators that are essential and unique for hESCs.

# 1 INTRODUCTION

## 1.1 Introduction to embryonic stem cells (ESCs)

Tissue homeostasis is an essential compensatory event for multi-cellular organisms to replace tissue that are injured or aged. The bone marrow, skin and liver are some examples of the naturally occurring somatic stem-cell systems that maintain tissue replacement after birth but not all somatic cell types could be replenished. In addition to naturally occurring somatic stem-cell systems, stem cells can be obtained from peri- and early post-implantation-stage embryos. Retrospectively, from an ontogenic perspective, embryo-derived pluripotent stem cells might be considered to represent an archetypal genomic state from which other patterns of genome activity ensue later in development.

The differentiation potential of stem cells is one of the defining criteria for categorizing them into two main types. Embryonic stem cells (ESCs) are isolated from the epiblast of the inner cell mass (ICM) of preimplantation blastocysts. They can extensively self-renew in *vitro* while retaining the capacity to differentiate into every cell type present within the fetus, a capability known as pluripotency. On the other hand, adult stem cells/somatic stem cells are derived after embryonic development and they are more restricted than ESCs to form the different cell types.

These defining properties of ESCs; having an extensive self-renewal ability as well as pluripotency enable continual supply of ESCs for the derivation of the different cell types for replacement and this makes ESCs highly desirable for their application in the for regenerative medicine. This is especially applicable in cases such as Parkinson’s disease where even in normal healthy adults, dopaminergic neurons which are implicated in

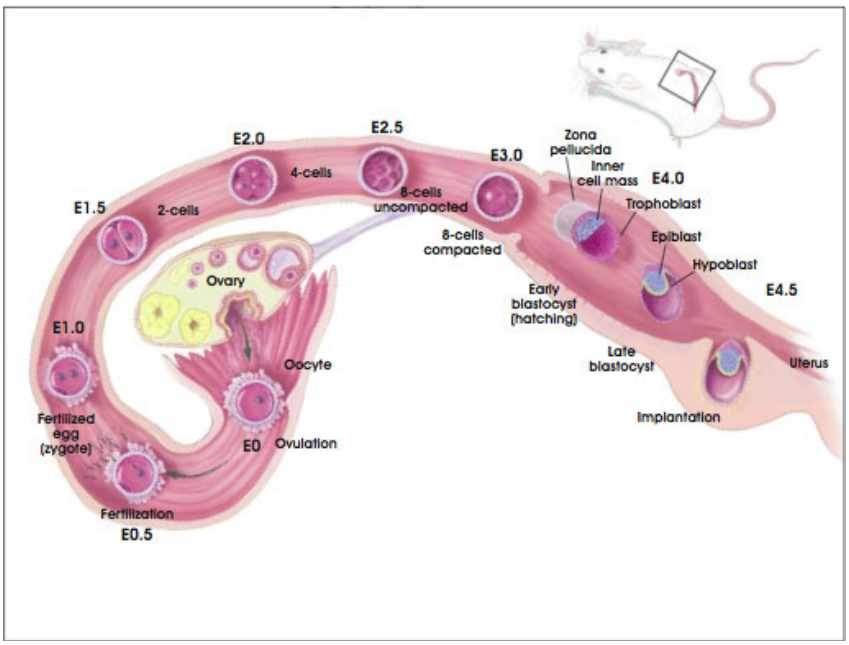
Parkinsons's Disease do not undergo self-renewal to replace damaged neurons[8]. The derivation of patient-specific replacement tissues for customized therapy of diseases could be the next revolutionary medicine in the future.

### 1.1.1 Derivation of ESCs

In mouse, embryo is the developing organism from the time of fertilization until E10.5, when it becomes known as a fetus. The stages are: pre-implantation embryo; post-implantation embryo and fetus[9]. The morula-stage mouse embryo (embryonic day 2.5; E2.5) comprises of a core of pre-ICM cells that turn into ICM cells at cavitation/blastulation (E3–E4). Pluripotent cells are first recognized in the ICM of the mouse blastocyst at embryonic day 3.5 (E3.5), and are present as late as in the pre-gastrulating embryo (Figure 1). Until ~E5.5, at least some primitive ectoderm cells remain pluripotent. In the rapidly changing embryo, these are not necessarily obvious as stem-cell populations, nor do they exist for very long. However, their stem cell capacity can be revealed by experimental intervention that essentially captures the cell in a "stem cell" state. In 1981, Martin explanted embryonic stem cells from the inner cell masses of late blastocysts and culture them in medium conditioned by an established teratocarcinoma stem cell line[10], whereas Evans and Kaufman used a co-culture system in which cells derived from the ICM of delayed mouse blastocysts were grown on a layer of mitotically inactivated mouse embryonic fibroblasts (MEFs) in the presence of blood serum[11]. In the absence of conditioned medium, the self-renewing and pluripotent ESC phenotype cannot be preserved, which leads to preferential neuronal differentiation. Pluripotency of these ESCs can be determined by their ability to integrate into the ICM of E3.5 blastocysts and produce a high rate of chimaerism in tissues of the developing fetus after intra-blastocyst

injection. And similar to the ICM, these ESCs express alkaline phosphatase (AP), E-cadherin, stage-specific embryonic antigen-1 (SSEA1) and octamer-binding transcription factor-4 (OCT4). In 1998, Thomson *et al* isolated human embryonic stem cell (hESC) lines from the blastocysts stage of donated human embryo that were produced by *in vitro* fertilization for clinical purposes[1]. In contrast to mouse, in which the blastocyst-derived ESC and trophoblast stem cells (TSCs) purported lack the ability to interconvert, human ESCs can be efficiently directed to differentiate into trophoblast cells.

More recently, pluripotent cell lines were generated from postimplantation mouse embryos using the culture conditions employed for the derivation of stem cell lines from human blastocysts; specifically, explanting on to a substrate of feeder cells or fibronectin and supplementation of the medium with activin and Fgf2[12; 13]. These 'epiblast stem cells (EpiSCs)' cannot be propagated efficiently from single cells and do not readily contribute to chimaeras. They share many characteristics with the pluripotent lines derived from human embryos, suggesting that hESCs may actually represent a more advanced stage of development than the blastocyst[12; 13; 14].



(© 2001 Terese Winslow)

**Figure 1: Development of the Preimplantation Blastocyst in Mice.**
The diagram shows the different stages of development in mice.

### 1.1.2 Similarities and differences among mESCs, EpiSCs and hESCs

Akin to mESCs, hESCs could indefinitely self-renew *in vitro* and exhibited pluripotential characteristics. hESC lines expressed cell surface markers that characterize undifferentiated nonhuman primate ES and human embryonal carcinoma cells (EC) cells, including stage-specific embryonic antigen SSEA–3, SSEA-4, TRA-1-60 and TRA-1-81. hESCs also express high level of alkaline phosphatase (AP) and telomerase activity[1]. For ethical reasons, the ability of ESCs to contribute to the germ line in chimeras is not a testable property in hESCs, however, these hESC lines maintained the potential to form teratoma derivatives of all three embryonic germ layers after injection into severe combined immunodeficient (SCID) mice.

In addition to the ability to self-renew and differentiate, hESCs share many similarities with mESCs[15; 16]. Both of them express genes which are associated with pluripotency[17; 18;

[19]. *OCT4* (coding for the protein OCT4) and *NANOG*, both key components of the core transcriptional regulatory network[15; 20; 21] are highly expressed in undifferentiated ESCs [16; 22; 23; 24; 25; 26] and upon differentiation, the expression of these genes are reduced. However, further examinations by several groups revealed that both hESCs and mESCs are dissimilar in many other aspects and this is intriguing given that they share the main properties of ESCs. One of the key differences between them is the signalling pathway that governs them. ESCs can be propagated in the presence of serum and in co-culture with a layer of fibroblasts. Subsequent fractionation of the conditioned medium identified the active component as leukaemia inhibitory factor (LIF)[27]. LIF is a member of IL6 family of cytokines and binds to a receptor complex consisting of two transmembrane proteins, Leukemia inhibitory receptor (LIF-R) and gp130 which recruit JAK kinases and activate the STAT3 signalling pathway[28]. mESCs are sustained by growth factors like bone morphogenetic protein 4 (BMP4) and LIF which activate the downstream transcription factor SAMD1/5 and STAT3 respectively. On the contrary, LIF does not support hESCs and BMP4 induces hESCs to differentiate[17], hESCs are sustained by growth factors such as basic fibroblast growth factor (bFGF), activin and insulin growth factor (IGF) instead[29; 30]. Additionally, mESCs have a dome-like morphology whereas hESCs have a flat morphology. Furthermore, undifferentiated mESCs express SSEA1 but undifferentiated hESCs express SSEA3 and SSEA4. These differences could be due to species-specific differences in embryonic development or they could be derived from cells originating from different developmental stages. Consistent with this idea is the identification of post-implantation murine epiblast-derived stem cells. Epiblast stem cells (EpiSCs), which are derived from E5.5 or E6.5 mouse epiblasts and they exhibit features

of pluripotency where they express Oct4, Nanog and Sox2. EpiSCs are similar to hESCs in morphology and can be derived and maintained under the same growth conditions that support hESCs[13]. On the other hand, hESCs does not express epiblast-specific gene such as FGF5 and they differ in the regulation of Oct4[13]. Although Oct4 is similarly expressed in both cells, the Oct4 promoter distal enhancer region exhibits higher activity in hESCs, whereas the proximal enhancer exhibits higher activity in EpiSCs[13]. mESCs and EpiSCs are considered to represent defined states of pluripotency. These cellular states can be interconverted by changing the medium conditions. During conversion, each cell line/state loses its distinct colony morphology and acquires the growth characteristics and signaling dependence as well as the Oct4 gene expression regulation of other cell state.

### 1.1.3 Transcription factors that are essential for the maintenance of ESCs

Homeodomain transcription factors are evolutionarily conserved and are involved in morphogenesis and organogenesis in many organisms[31]. *Oct4* and *Nanog* are two of the essential regulators of early development and ESC identity[24; 26; 32; 33] that were identified earlier. Oct4 is a POU homeodomain transcription factor that is encoded by *Oct4*. It is expressed in the pluripotent cells of the ICM, epiblasts and primordial germ cells (PGCs). Oct4 knockout embryos are not viable as it is unable to form the ICM and the embryos result in trophectoderm cells. In mouse, the depletion of *Oct4* induces inappropriate differentiation of ICM and ESCs to trophectoderm while an overexpression of Oct4 leads to differentiation into endodermal and mesodermal lineages[25]. Similarly, in hESCs, the manipulation of OCT4 levels induces differentiation. A reduction of OCT4 promotes upregulation of mesodermal and endodermal markers while an upregulation of OCT4 promotes endodermal derivatives[33]. Additionally, Oct4 has been reported to

heterodimerize with another key transcription factor Sox2 to repress and activate gene expression in mESCs[34; 35]. Sox2 belongs to the high mobility group (HMG)-box transcription factor. Although Sox2 expression is not restricted to the ICM, the depletion of Sox2 also results in defective epiblast and differentiation into multiple lineages[36].

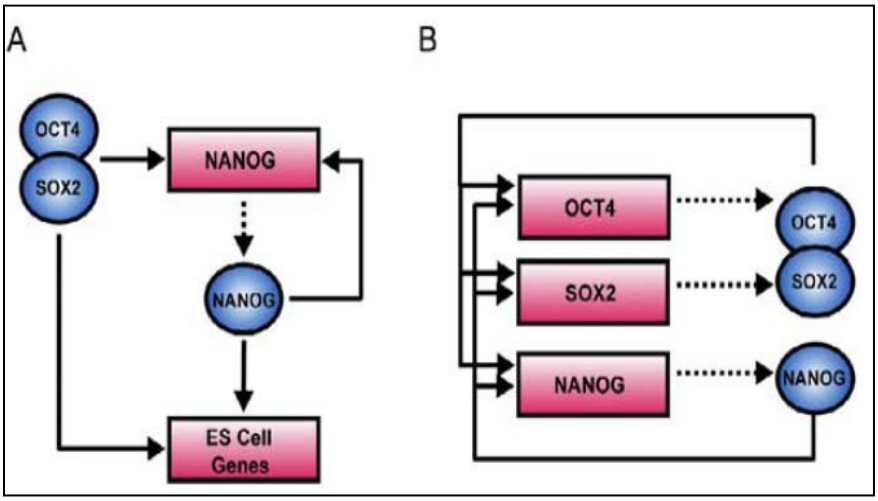
Nanog belongs to the homeobox-containing gene with homology to members of the natural killer gene family. It is preferentially expressed in ESCs and the identification of Nanog was a breakthrough as it is able to maintain mESCs in the absence of LIF[24; 26]. In mouse, Nanog is expressed in the ICM and epiblast and Nanog knockout embryo is unable to develop an epiblast and differentiates into the extraembryonic lineage. Nanog plays an important role in the maintenance of epiblast by repressing differentiation along the primitive endoderm lineage. Similar to Oct4, downregulation of Nanog induces differentiation in both mESCs and hESCs[37].

### 1.1.4 Molecular basis of embryonic stem cells

#### 1.1.4.1 Transcriptional regulation

The molecular basis of self-renewing and pluripotentiality of ESCs is relevant to the understanding of stem cell biology. Transcription factors such as *Oct4*, *Sox2*, *Nanog* and many other essential regulators[20] that underlie pluripotential character defines the determinants that are governing the ESC state. Evaluating how multiple transcription factors might integratively regulate its target genes may be attained by leveraging on the genome-wide DNA binding data to construct a transcriptional regulatory network. For instance, to gain insights into the transcriptional regulation of hESCs, Boyer *et al* performed genome-scale location analysis on OCT4, SOX2, and NANOG. They

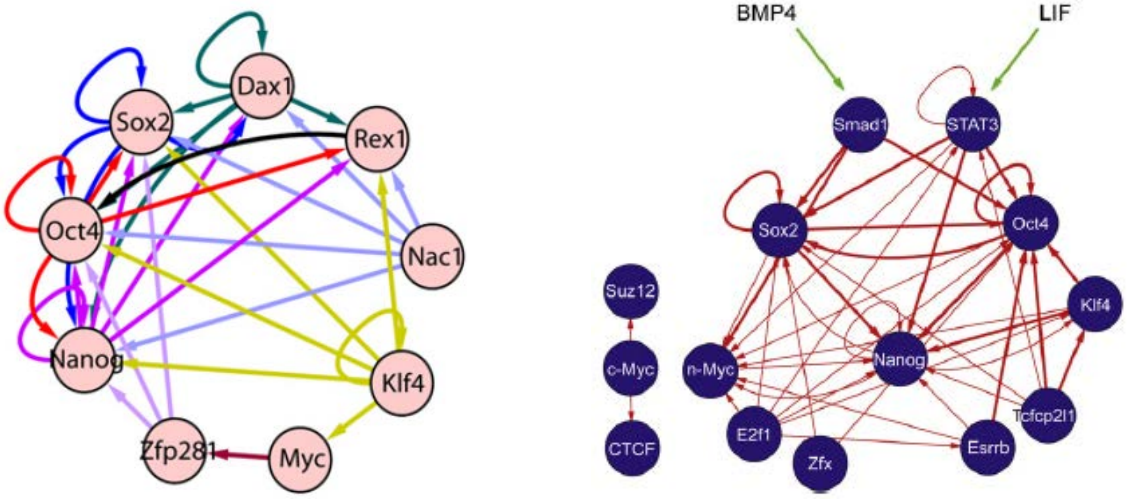
discovered a substantial portion of co-occupancy on each other's target genes[15]. These target genes frequently encode transcription factors, many of which are developmentally important homeodomain proteins. They showed that OCT4, SOX2, and NANOG collaborate to form regulatory circuitry in ES cells consisting of autoregulatory and feedforward loops (Figure 2). These results provide new insights into the transcriptional regulation of stem cells and reveal how OCT4, SOX2, and NANOG contribute to pluripotency and self-renewal. The conservation of the autoregulatory loops among OCT4, SOX2 and NANOG is also conserved in the Oct4 and Nanog ChIP-PET study performed in mESCs[38], suggesting the functional relevance of this network motif in the maintenance of stemness in ESCs.




Cell. 2005 Sep 23;122(6):947-56.

**Figure 2: Transcriptional regulatory motifs in hESCs**
**A)** Feedforward transcriptional regulatory circuitry in hESCs. Regulators are represented by blue circles; gene promoters are represented by red rectangles. Binding of a regulator to a promoter is indicated by a solid arrow. Genes encoding regulators are linked to their respective regulators by dashed arrows.
**B)** The interconnected autoregulatory loop formed by OCT4, SOX2, and NANOG

With the discovery of more ESC transcription factors that are essential for the maintenance of ESCs through RNAi screen and other studies, it is interesting to study how these additional factors participate in the transcriptional regulatory network. Kim *et al* performed a genome-wide ChIP- chip mapping study on nine biotin-tagged ESC transcription factors (Oct4, Sox2, Nanog, Klf4, c-Myc, Dax1, Rex1, Zfp281 and Nac1)[20] while Xi *et al* studied 13 sequence-specific transcription factors (Nanog, Oct4, STAT3, Smad1, Sox2, Zfx, c-Myc, n-Myc, Klf4, Esrrb, Tcfcp2l1, E2f1, and CTCF) and two transcriptional regulators (p300 and Suz12) using the ChIP-seq platform for genome-wide location analysis in mESCs[21]. Interestingly, both groups concertedly observed two distinct clusters of binding targets and they showed extensive co-regulatory network (Figure 3) in mESCs with high density binding of transcription factors on the same loci and associated hierarchical relevance in a transcriptional regulatory network.




Cell. 2008 Mar 21;132(6):1049-61. Cell. 2008 Jun 13;133(6):1106-17

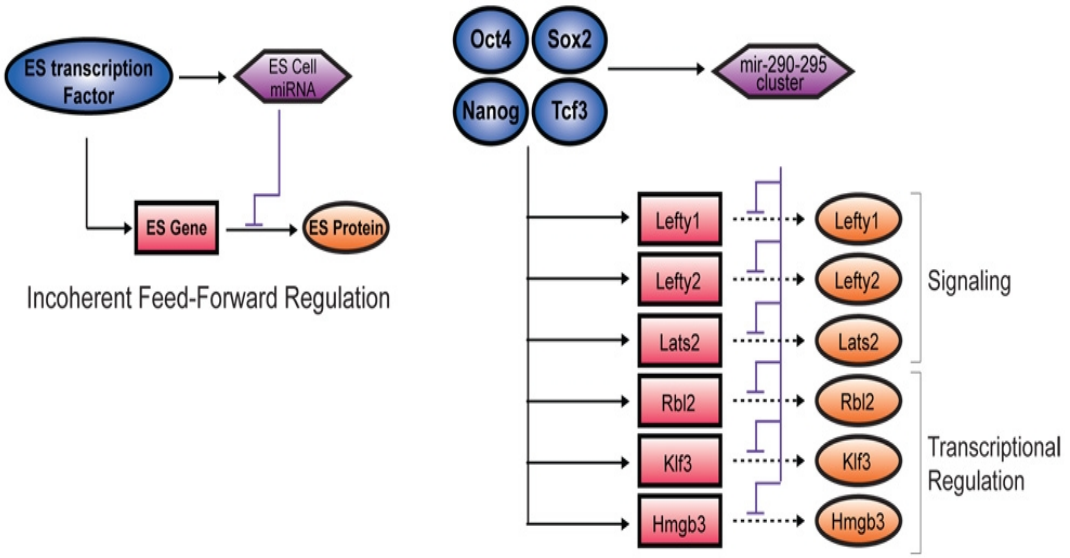
**Figure 3: Transcriptional regulatory network in mESCs**
Transcriptional regulatory circuit within nine factors and five factors (Nanog, Oct4, Sox2, Dax1, and Klf4) show autoregulatory mechanism (left from Kim et al). Network of regulatory interactions inferred from ChIP-seq binding assays and from gene-expression changes during differentiation (right from Xi et al).

Hence, an understanding in the transcriptional regulatory circuitry that is responsible for pluripotency and self-renewal in ESCs is crucial to gain insights into how the key ES cell regulators cooperate to espouse hESCs with their quintessential properties.

#### 1.1.4.2 Transcriptional regulation of micro-RNAs (miRNAs)

Interestingly, the genome also encodes a group of non-protein coding transcripts known as miRNAs that are proven to be essential for ESCs wherein they function to fine-tune the transcriptional regulatory network, buffering it from the direct effect of upstream elements[39]. MicroRNAs (miRNAs) are a group of non-protein encoding transcripts of ~20-25 nucleotides that are known to direct mRNAs for degradation or repress mRNA translation in plants and animals[40]. Interestingly, a subset of miRNAs is preferentially expressed in both human and mouse ESCs[41; 42; 43] and a deficiency in the components of the miRNA processing apparatus compromised their ability to differentiate as well as their ability to proliferate[44; 45; 46], indicating the importance of miRNAs in ESCs. To understand how the miRNA genes are wired into the core transcriptional regulatory circuitry to provide a different level of transcription regulation, Marson *et al* examined the binding of Oct4, Sox2, Nanog and Tcf3 at the transcriptional start sites of miRNA transcripts and performed quantitative sequencing of short transcripts in ESCs, neural precursor cells and mouse embryonic fibroblasts[39]. Interestingly, they identified two groups of miRNAs that are bound by Oct4/Sox2/Nanog/Tcf3; one group of miRNAs that are preferentially expressed in pluripotent cells and a second; which are silenced in ESCs by polycomb group proteins but are upregulated when the cells differentiate. miRNAs provide a distinct level of transcriptional regulation by modulating the direct effect of transcription factors through different network motifs. One such example is evident in the binding of

Oct4/Sox2/Nanog/Tcf3 on signalling protein genes such as *Lefty1*, *Lefty2* and the microRNA gene *mi-290-295*. Oct4/Sox2/Nanog/Tcf3 promotes the expression of *Lefty1*, *Lefty2* and *mir-290-295* although *mir-290-295* inhibits the expression of *Lefty1* and *Lefty2*. They propose that *mir-290-295* functions to fine tune the expression of these signaling proteins and this type of network motif that exerts both positive and negative effects on its target is known as "incoherent feed-forward" regulation (Figure 4).




Cell. 2008 Aug 8;134(3):521-33.

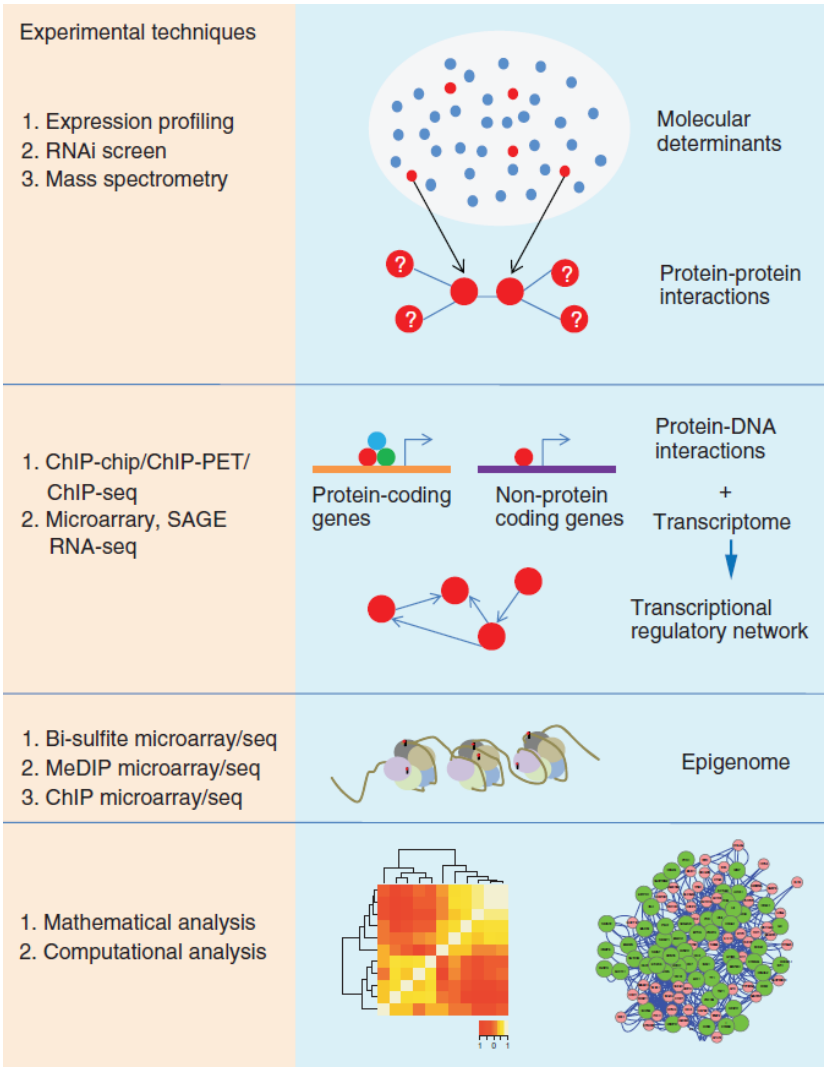
**Figure 4: An incoherent feed-forward motif**
An incoherent feed-forward motif (Alon, 2007) involving miRNA repression of a transcription factor target gene is illustrated (left). Oct4/Sox2/Nanog/Tcf3 promotes the expression of the miRNA cluster *mir-290-295* as well as its downstream targets such as *Lefty1*, *Lefty2* and the rest of the other genes. However, *mir-290-295* inhibits the expression of Oct4/Sox2/Nanog/Tcf3's downstream target genes. Transcription factors are represented by dark blue circles, miRNAs in purple hexagons, protein-coding gene in pink rectangles, and proteins in orange ovals.

### 1.1.4.3 Epigenetic regulation of embryonic stem cells

Additionally, epigenomic studies define the *cis*-regulatory elements; for instance, active/inactive/poised promoters and enhancers that help to predict the transcriptional

status of the genes that are necessary to maintain the "stemness" of stem cells. This can be achieved by analyzing the histone marks as well as DNA methylation at a genome-wide level, additionally revealing insights into the epigenetic architecture of stem cells[47]. Hawkins *et al* carried out ChIP-seq was conducted on 11 histone marks to identify the global histone modifications patterns for both H1 hESCs and IMR90 fibroblasts. In combination with RNA-seq, chromatin architecture and gene expression were analyzed and they observed an apparent expansion of repressive domains marked by H3K27me3 on genes related to pluripotency, development and lineage-specific functions in differentiated cells, indicating differentiated cells have an increased repressive chromatic structure as compared to hESCs. On the contrary, histone modifications associated with gene activity like H3K4me1, H3K4me2, H3K4me3, H3K27ac, and H3K36me3 do not expand in differentiated cells when compared to hESCs. Unlike enhancers, promoters and ChIP-rich regions (regions of ChIP-Seq enriched sequences) showed little variation in DNA methylation across cell types. Hence, in combination of the different high-throughput platforms and the integration of the bioinformatics analysis, new insights into the differential genome regulation between hESCs and differentiated cells can be gained.

Given that the transcriptional programs that underlie mESC and hESC characteristics are governed by transcription factors, epigenetic regulators and miRNAs, an understanding of the epigenetic landscape and transcriptional regulation of ESCs is required for a holistic understanding of the ESC transcriptional program[48] (Figure 5) and to achieve this knowledge of ESCs, it is necessary to uncover the determinants that are governing ESC properties.




Wiley Interdiscip Rev Syst Biol Med. 2012 Jan-Feb;4(1):39-49. doi: 10.1002/wsbm.151.

**Figure 5: High-throughput experimental methodologies**
Building the network model for pluripotent state of stem cells through the integration of mathematical and computational disciplines with high-throughput experimental methodologies. With systems biology, different layers of the stem cells biology are explored. The molecular determinants of stem cells are defined, protein–protein interaction map, transcriptional regulatory network that is constructed through the analysis of both the transcriptome and location binding and the epigenetic status of stem cells.

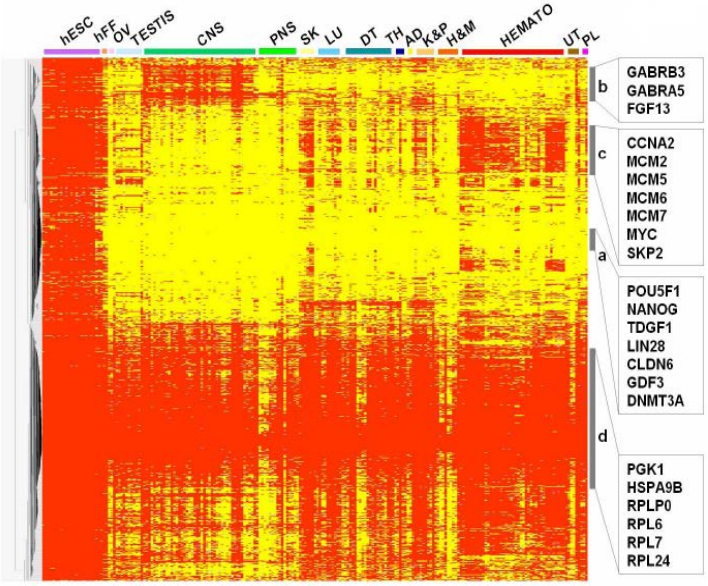
## 1.2 Identifying determinants of stem cell identity

High-throughput approach is popular for the parallel identification many determinants of ESCs in a considerably shorter amount of time. This can be applied in several of the approaches that aim to identify the regulators of ESCs such as in gene expression profiling of a cell, protein-protein interaction/mass spectrometry or RNAi experiments. It is likely that different approaches yield different findings as each method has its strengths that may not be complemented by the other.

### 1.2.1 Gene expression analysis

Expression profiling delineates the expression signature of a cell by quantifying the relative amount of mRNA that is expressed in at least two experimental conditions. This indicates whether a gene is actively required or inhibited in a particular condition, thereby revealing the significant molecular elements of a particular cell. This is a popular way to search for genes that are highly expressed and would thereby be essential for the maintenance in ESCs. With the advancement in technology, the scale in the number of genes that can be profiled is increased dramatically wherein DNA microarray and Serial Analysis of Gene expression (SAGE) are examples of high-throughput gene expression technologies that are capable of system-wide expression patterns. Thousands of genes among different cell types or conditions can be analyzed simultaneously, creating a global picture of cellular function. For instance, DNA microarray was adapted in the study of ESCs versus differentiated cell types in both mouse and human where the authors identified 88 genes which have consistent changes when mESCs differentiate[49] and in another study performed on 38 different transcriptome of hESCs, a gene expression atlas for pluripotent and differentiation markers was generated (Figure 6)[50]. Oct4 and Nanog are

among the highly expressed genes in both mESCs as well as hESCs but these genes are downregulated upon differentiation.




Stem Cells. 2007 Apr;25(4):961-73.

**Figure 6: Mico-array analysis of hESCs**
Heat map of gene expression detection for the hESC gene list across 24 hESC and 193 fetal and adult tissues samples analyzed with the U133A microarray (828 probesets). Red stands for a "Present" detection call (i.e. gene expression confidently detected according to the GCOS 1.2 software), grey for "Marginal" and yellow for "Absent".

RNA-sequencing (RNA-seq) is the next generation revolutionary tool to analyze the whole transcriptome. It uses high-throughput sequencing technologies to sequence RNA at base-level resolution[51; 52] resulting in a more comprehensive coverage than microarray as it is not limited by the probes on the chip. This technology further advances the transcriptional profiling by enabling the distinction of the different isoforms of a gene and the quantification of RNA transcripts.

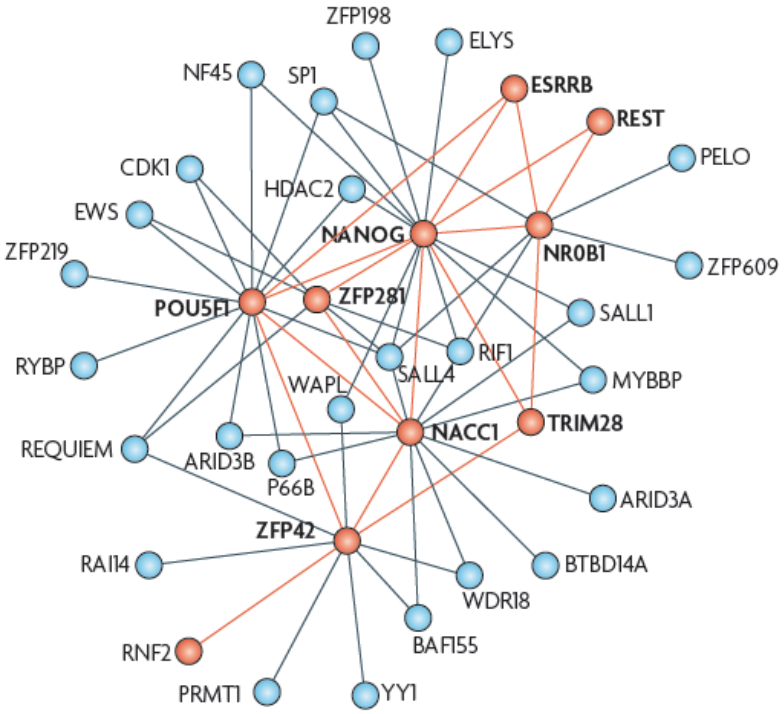
Hence, through expression profiling of the ESCs through various means, it is possible to identify factors that are essential for the maintenance of ESCs.

### 1.2.2 Protein-protein interaction

Besides differential gene expression which enables the identification of the molecular elements of stem cells, protein-protein interaction via affinity purification followed by mass spectrometry is another useful technique to identify the determinants of stem cells. Proteomics approach enables one to identify and expand the discovery of more factors through protein-protein interaction of a single critical factor. For instance, Wang *et al* applied affinity purification of Nanog followed by mass spectrometry and identified many interacting partners of Nanog like Dax1, Nac1, Zfp281 and Oct4. They confirmed the functional relevance of these new partners via RNAi and constructed a protein interaction network of Nanog[53] (Figure 7).

Along with Nanog interaction map, a protein interaction map centered on Oct4 was also constructed and more regulators were identified. Oct4 associates mainly with transcriptional regulators, but also with a variety of other chromatin binding proteins involved in DNA replication, recombination, and repair, proteins involved in nuclear assembly and/or organization, and diverse enzymes, some of which are responsible for addition of posttranslational modifications [54; 55].

Besides Nanog and Oct4, c-Myc (Myc) is an important transcriptional regulator in ESCs and recently, a c-Myc centered protein network was constructed with NuA4 HAT (or the Tip60-Ep400 complex) being discovered to be one of the interacting proteins[56]. The availability of protein-protein interactions may thereby reveal essential regulators of ESCs.




Nat Rev Mol Cell Biol. 2009 Oct;10(10):672-81. Epub 2009 Sep

**Figure 7: Stem cell regulatory networks.**
Schematic diagram showing high-confidence protein–protein interactions between NANOG and NANOG-associated proteins.

### 1.2.3 Genetic perturbation studies: Depletion/Overexpression

Oct4, Sox2 and Nanog are prototypic pluripotency factors well known to be important in defining the identity of both mESCs and hESCs. Evaluating the importance of candidate architects of pluripotency could be achieved through genetic perturbation of the expression of these regulators *in vitro* or *in vivo*. As transcriptional regulatory network is tightly regulated in stem cells, a disruption in the dosage level of any essential regulators beyond its critical threshold may have a negative impact on them. Oct4 is expressed during the early development of mouse embryos and is critical for in the ICM formation. Oct4-deficient mouse embryos display impaired preimplantation development [57; 58]. When Oct4 expression is genetically manipulated in mESCs *in vitro* using conditional

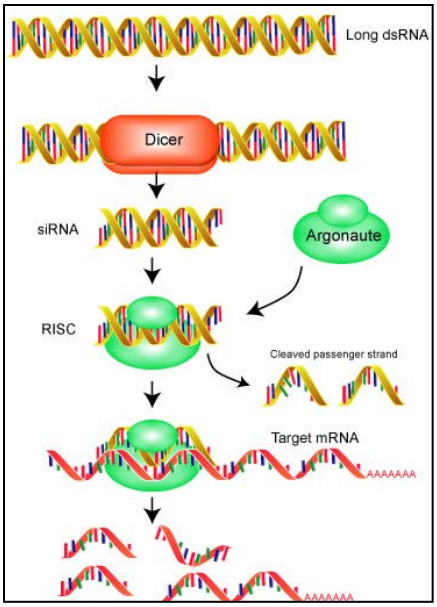
overexpression or repression by 50%, they succumb to differentiation [25]. Sox2 is required in the early mouse embryo and the knockout of *Sox2* displayed a defective epiblast[59] and ESCs derived from Sox2-null embryos differentiate into trophectoderm-like cells. Sox2 regulates the expression of Oct4 and interestingly, an overexpression of Oct4 is able to rescue these Sox2-null ESCs [60]. Given that mESCs are reliant on extrinsic instruction via the growth factor LIF to maintain their pluripotentiality; two groups premised that they could identify novel pluripotency factors by identifying genes whose overexpression could enable mESC self-renewal independent of LIF signaling. Through the screening of a cDNA library of more than 1000 clones, Chambers *et al* identified Nanog as a gene whose overexpression enables mESC self-renewal even in the absence of LIF[26]. Similarly, Mitsui *et al* showed that constitutive overexpression of Nanog renders the cells to grow normally without LIF. Additionally, Mitsui *et al* showed that genetic deletion of Nanog makes the ICM unable to generate the epiblast, and that depletion of Nanog in mESCs induces differentiation[24]. Thus, both gain-of-function and loss-of-function approaches constitute complementary methods to identify genes that underlie pluripotentiality in ESCs.

Recent techniques that have enabled the adaptation of RNA interference (RNAi) to high-throughput format for parallel, genome-wide targeted knockdown of specific gene products enabling the systematic query of gene function. HT-RNAi screen often involves the integration of instrumentation for automated imaging, sample preparation as well as computational aspects comprising of optimized automated algorithms that are designed to acquire, process and analyze images to extract specific cellular information and/or all possible cellular parameters. It is anticipated that large scale biology will be employed as a tool for mechanistic studies in to extend the value of genomics for stem cells.

## 1.3 High-throughput RNA interference technology

High-throughput RNA interference (HT-RNAi) is a powerful research tool that allows for the systematic identification of hitherto unappreciated ESC regulators. Such genome-scale RNAi platforms make it possible to assess the role of nearly every gene in the genome to produce a phenotype in an unbiased manner, due to its massively parallel nature. RNAi is advantageous for investigating the functional and causal role of genes in numerous cellular processes including signal transduction, cell survival, apoptosis, cell cycle and more recently, to study pluripotency in stem cells.

RNAi silences gene expression and is a natural biological mechanism, induced by long double stranded (ds)RNA that is homologous to the target gene. dsRNA is chopped into shorter fragments by an endoribonuclease of the RNase III family (Dicer). These fragments; small interfering RNA (siRNA) duplexes of approximately 21-23 base pairs in length are bound by the RNA-induced silencing complex (RISC) and separated into single strands, where one strand remains bound by RISC. This strand serves as the template for the recognition of the corresponding mRNA. Once a target mRNA is recognized, the protein Argonaut, which is a component of the RISC, now cleaves the mRNA and initiates its further degradation, resulting in the knockdown or silencing of the target gene and diminished protein expression (Figure 8).

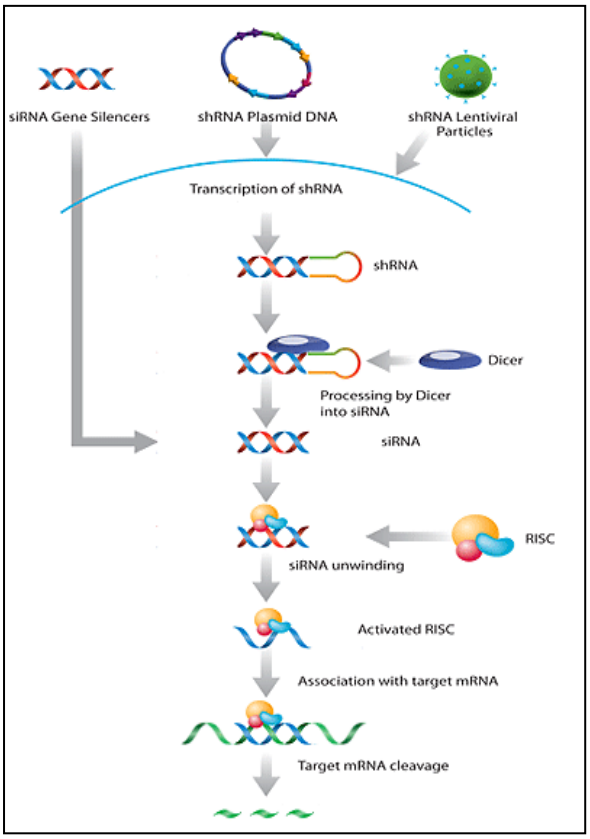



©RNAiWeb.com

**Figure 8: Mechanism of RNA interference**
The long dsRNAs enter a cellular pathway that is commonly referred to as the RNA interference (RNAi) pathway. First, the dsRNAs get processed into 20-25 nucleotide small interfering RNAs (siRNAs) by an RNase III-like enzyme called Dicer (initiation step). Then, the siRNAs assemble into endoribonuclease-containing complexes known as RNA-induced silencing complexes (RISCs), unwinding in the process. The siRNA strands subsequently guide the RISCs to complementary RNA molecules, where they cleave and destroy the cognate RNA. Cleavage of cognate RNA takes place near the middle of the region bound by the siRNA strand.

Shortly after its discovery RNAi became a widely used technology for loss-of-function studies in invertebrates like the nematode *Caenorhabditis elegans* or the fruit fly; *Drosophila melanogaster*. In most mammalian cells, however, long dsRNA induces a potent and often detrimental interferon response. Hence the approach of delivering long dsRNA is not useful to study the specific function of individual genes in most mammalian cells. Alternatively, synthetic versions of siRNA; that is similar to the Dicer products, can be introduced into the cell, thus triggering the remainder of the RNAi pathway. Importantly, these short interfering RNAs (siRNAs) do not induce an interferon response but mediate gene-silencing in mammalian cells. Besides siRNA, esiRNA, shRNA and

miRNA are several functional genomics technologies tool that can be introduced into the cells artificially for targeted gene knockdown (Figure 9).




http://www.sigmaaldrich.com/life-science/functional-genomics-and-rnai/shrna/library-information/feature-article.html

**Figure 9: shRNA and siRNA Mediated Gene Silencing.**
Different delivery strategies and the processing of shRNAs in the cell are shown.Cells may be directly transfected with siRNAs or shRNA plasmids or transduced with shRNA lentiviral particle for gene silencing. Following transcription of the shRNA in the nucleus, the hairpin enters the RNAi pathway when it is cleaved by Dicer to generate siRNA. The siRNA is recognized by RISC, which mediates cleavage of the target mRNA for gene silencing. Synthetic siRNA may be directly transfected and enters the RNAi pathway when it assembles with RISC. The activated RISC complex then associate with its target mRNA for degradation.

### 1.3.1 Different approaches to RNA interference mediated gene silencing

#### 1.3.1.1 siRNA

siRNA elicit transient gene repression. These synthetic siRNAs are available in RNA duplexes of 20-23 nucleotides in length and are complementary to the endogenous to mRNAs. They can be made accessible through chemical solid phase synthesis and several companies like Dharmacon and Sigma provide genome-scale siRNA libraries that are available for species such as human and mouse.

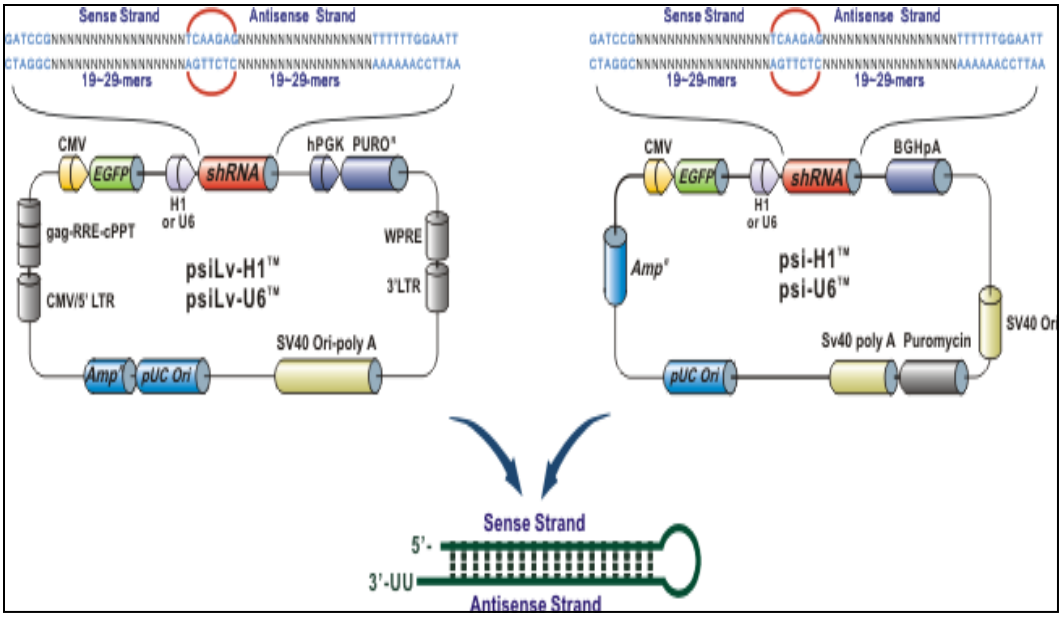
#### 1.3.1.2 esiRNA

esiRNA or endoribonuclease siRNAs are a mixture of siRNA oligos resulting from cleavage of long double-stranded RNA (dsRNA) with an endoribonuclease such as *Escherichia coli* RNase III or dicer. In this case, a cDNA template is amplified by PCR and tagged with two bacteriophage-promoter sequences. RNA polymerase is then used to generate long double stranded RNA that is homologous to the target-gene cDNA. This RNA is subsequently digested with Rnase III from *Escherichia coli* to generated short overlapping fragments of siRNAs with a length between 18-25 base pairs. This complex mixture of short double stranded RNAs is similar to the mixture generated by Dicer cleavage in vivo and is therefore called endoribonuclease-prepared siRNA or short esiRNA. esiRNA are a heterogeneous mixture of siRNAs that all target the same mRNA sequence. These multiple silencing triggers lead to highly specific and effective gene silencing.

#### 1.3.1.3 shRNA

shRNA refers to small hairpin or short hairpin RNA. It is generated from shRNA vector that utilizes the U6 or H1 promoter to ensure that the shRNA is always expressed. The

shRNA vectors express the sequence of RNA that makes a tight hairpin turn such that it can be used to silence gene expression through RNA intereference (Figure 10). shRNA vector can be introduced into the cells via lipid transfection or viruses. This vector is usually passed on to daughter cells, allowing the gene silencing to be inherited. The shRNA hairpin structure is cleaved by the cellular machinery into siRNA, which is then bound to the RNA-induced silencing complex (RISC). This complex binds to and cleaves mRNAs which match the siRNA that is bound to it. shRNA is transcribed by RNA polymerase III and shRNA production in a mammalian cell can sometimes cause the cell to mount an interferon response as the cell seeks to defend itself from what it perceives as viral attack. Studies have shown that dsRNA located and processed in the endosome may activate Toll-like receptor 3 or 7 leading to the induction of type 1 interferons[61; 62] In addition, cytoplasmic sensors including the RNA-dependent protein kinase can mediate dsRNA-triggered interferon response and in polymerase III driven shRNA expression systems, specific sequences around the transcription start site have been identified, which can also lead to interferon (IFN) induction[63; 64]. IFN then binds to cell surface receptors in an auto- or paracrine fashion and confers a more global antiviral state by inducing a complex array of IFN-stimulated genes (ISGs), including RNA-dependent protein kinase and a family of oligo adenylate synthetase (Oas) enzyme. This problem is not observed in miRNA, which is transcribed by RNA polymerase II (the same polymerase used to transcribe mRNA)[65].



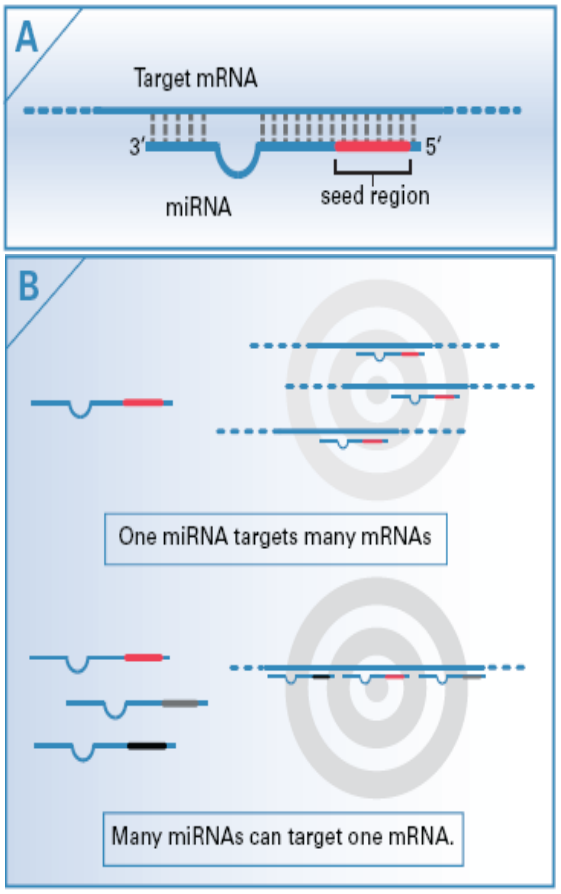
http://www.genecopoeia.com/product/shrna/?gclid=CNL4---ao6kCFUkb6wodFRBftA

**Figure 10. Lentiviral and non-viral expression vector-based shRNA clones with H1 or U6 promoter.**
Expression cassettes of all shRNA clones including the promoter, sense and antisense target sequences, hairpin, termination and other linker sequences.

#### 1.3.1.4 miRNA

MicroRNAs (miRNAs) are short ribonucleic acid (RNA) molecules, on average only 22 nucleotides long and are found in all eukaryotic cells, except fungi and marine plants[66]. miRNAs are post-transcriptional regulators that bind to complementary sequences on target messenger RNA transcripts (mRNAs), usually resulting in translational repression and gene silencing. miRNA-mediated gene regulation is achieved when an miRNA binds to an miRNA recognition element (MRE) in the 3'-untranslated region (UTR) of a target mRNA. Unlike siRNA-mediated gene silencing where, in most cases, complete sequence complementary between the siRNA and its mRNA target is necessary, miRNA mediated gene regulation can be achieved through partial complementary between the MRE and as few as six or seven nucleotides at the 5'-end of the miRNA, called the seed region (Figure 11A). The short length of the seed region predicts that each miRNA would have multiple

potential target genes in the genome. In addition, each target gene may be regulated by multiple miRNAs (Figure 11B). It is proposed that the synergistic action of multiple miRNAs is important for the regulation of a target gene.




http://www.dharmacon.com

**Figure 11: Mechanism of miRNA mediated gene suppression.**
**A)** MiRNA recognition region (MRE).
**B)** Targets of miRNAs.

### 1.3.2 Development of HT-RNAi screen procedure

siRNA, esiRNA, shRNA and miRNA could be adapted for HT-RNAi screen. HT-RNAi screen shares many similarities to HT-chemical compound screens. Both methods are carried out in a miniaturized plate format (384-well plate) and encompass the integration of automated and large-scale sample processing. One apparent difference between them is

that for RNA screens, it requires an additional transfection step which could be a major limiting step for cells that are difficult to transfect. A HT-RNAi screen involves four discreet experimental steps. Firstly, is the pre-experimental planning stage to define the experimental question/hypothesis and identify appropriate model systems and testing reagents. As such, it is crucial to select the appropriate cell lines and the particular RNAi library to be employed in a screen. For instance, targeted libraries like the kinome, druggable genome and even the whole-genome is available. However, given the experimental query to be addressed, the cost and the cumbersome handling of a whole-genome library, a smaller arrayed library is frequently employed. At this stage, high content screening (HCS) involving multi-parametric assays can be considered as they are used to better define the function of a specific gene target. RNAi-based gene function analysis is particularly powerful when combined with high content screening (HCS). Secondly, the assay development stage which involves the establishment of parameters to conduct robust and reproducible assays. This step requires the optimization of experimental variables such as cell number, the amount of transfection reagent, phenotype readout etc. This step is usually more time-consuming and requires a considerable effort to achieve an optimal condition for the screen and in some cases; this is the bottle-neck for the progression of the screening process. Thirdly, a small-scale assay validation screen using positive controls is performed to validate the experimental conditions that are established in the assay development. The goal of this is to serve as a preliminary run before the actual high-throughput RNAi screen to assess its performance with the statistical parameter known as Z' Factor and this determines the progression to the actual high-throughput RNAi screen. Other statistical calculations such as z-score are frequently

applied in high-throughput screening experiments. Z' Factor is a measure of the statistical effect size that assess the quality hence suitability of an assay for use in a full scale, high-throughput screening. z-score for each gene represents the distance each sample gene deviates from the population mean or negative controls in units of the standard deviation. With an appropriate z-score cutoff, "hit" genes of the screen could be determined and the application of bioinformatics analyses on the repertoire of "hit" genes helps to characterize their function and categorize them by grouping genes with similar functions together. Finally, with a good performance on the small-scale assay, a high-throughput RNAi screen that typically involves hundreds of plates can ensue. The screening process does not end at this high-throughput RNAi screen step as post activities are required to select and confirm the highly confidence positive hits. These require data normalization involving a number of statistical calculations such as z-score. Selected hits are advanced into hit confirmation on a smaller-scale and this process is also known as a secondary screen. This screen confirmation involves the re-arraying of prioritized hit in additional cell line under the same screen conditions. This validation is generally conducted with different siRNA from the actual screen targeting the same gene. The goal of the secondary screen is to filter off the false positives and increase the confidence of selecting the true positive hits from the high-throughput screen. With the completion of the HT-RNAi screen, a list of confirmed and validated genes is obtained. As with other genome-scale experiments like microarray or proteomics that generates a list of target genes, knowledge mining is required to extract biological information and this involves bioinformatics analysis to analyze and extract the biological meaning from the "hit" list data. Pathway analysis software like Ingenuity, Gene Ontology, Reactome etc or protein interaction

software such as STRING are several of the bioinformatics analyses that can provide comprehensive and informative conclusion from the pool of "hit" genes identified from the screen

### 1.3.3 HT-screening in mESCs

The application of the high-throughput screening technology has been applied in several ESCs studies. The first of such screen was carried out in mESCs by Ivanova *et al* in 2006 where they studied genes that govern self-renewal of mESCs[67]. They cloned shRNAs against 65 genes whose expression patterns suggest self-renewal regulatory functions. Self-renewal of the mESCs were measured with a fluorescence-based competition assay where the GFP-positive shRNA that were transduced into the cells using lentiviral vectors were mixed in a 4:1 ratio with non-transduced GFP-negative cells. A given shRNA that induces the cells to differentiate, changes the kinetics of cell cycle, compromised cell survival or changes in cell adhesion will result in the reduction of GFP+/GFP- ratio over time. The authors identified eight transcription (*Nanog, Oct4, Sox2, Tbx3, Esrrb, Tcl1, Dppa4, Mm 343880*) factors which affect the self-renewal of mESCs when these genes were depleted. Adaptation of RNAi studies to a genome-wide scale have massively enhanced the number of genes that may be screened. In 2009, two groups managed to improve the throughput of the number of genes screen to the genome-wide level. Hu *et al* employed the use of Dharmacon genome-wide siRNA library to screen against 16,683 mouse genes using *Oct4-GFP* mESCs reporter lines. They identified many genes that function in gene regulation and/or development, and are expressed at a high level in ESCs and embryonic tissue upon which two novel transcriptional regulators in mESCs Cnot3 and Trim28 were studied in greater detail[68]. The other group, carried out by Ding *et al*

employed the use of esiRNAs for the genome-wide RNAi screen. Similarly, they also use *Oct4-GFP* mESCs reporter lines in their screen and they discovered Paf1 complex to be essential for the maintenance of mESCs identity[69]. In addition, Mediator complex was also discovered in another high-throughput shRNA screen and in conjunction with the protein complex known as Cohesin, the authors showed that they regulate gene expression and chromatin structure in mESCs[70]. Therefore, these studies emphasize the value of a screen to identify novel regulators of pluripotentiality in mESCs.

As hESCs and mESCs are dissimilar in various regards such as difference in the signalling pathways that governs them as well as other characteristics, it will be interesting to identify regulators pertaining to hESCs. Although it is known to be technically challenging to implement and adapt high-throughput assays in hESCs, Desbordes *et al* have successfully carried out a high-throughput screening of chemical libraries in hESCs[71]. They identified several marketed drugs and natural compounds promoting short-term hESC maintenance and compounds directing early lineage choice during differentiation. Their study demonstrated the feasibility of performing a primary HTS assay in hESCs and paves the way for establishing high-content screening platforms. A HT-RNAi screen for hESCs like mESCs could be employed to pinpoint the determinants of hESCs. The discovery of essential regulators of hESCs will be crucial to improve our understanding in hESCs and therefore, the goal of my experiment is to carry out a genome-wide RNAi screen to identify the determinants of hESCs.

# 2 MATERIALS AND METHODS

## 2.1 Cell culture

### 2.1.1 hESCs

The hESC lines H1 (WA-01, passage 28), H9 (WA-09, passage 26), HES2 (ES-02, passage 79), HES3 (ES-03, passage 97) and H1 *OCT4-GFP* reporter cells (passage 56) were used for this study[1; 72]. They were cultured feeder free on matri-gel (BD)[73]. Condition medium used for culturing hESCs contained 20% knockout serum replacement, 1 mM L-glutamine, 1% non-essential amino acids and 0.1 mM β-mercaptoethanol and an additional 8ng/ml of basic fibroblast growth factor (Invitrogen) supplemented to the hESC unconditioned medium. Medium was changed daily. The hESCs were subcultured with 1 mg/ml collagenase IV (GIBCO-BRL) every 5-7 days.

### 2.1.2 mESCs

E14 mouse ES cells, cultured under feeder-free conditions, were maintained in Dulbecco's modified Eagle's medium (DMEM, GIBCO-BRL), with 15% de-activated ES-qualified fetal bovine serum (FBS, GIBCO-BRL), 0.055 mM β-mercaptoethano, 2mM L-glutamine, 0.1 mM MEM nonessential amino acid, 5000 U/ml penicillin/streptomycin, and 1000 U/ml LIF (Chemicon).

### 2.1.3 293-T cells

293T cells were cultured in DMEM with 10% FBS and maintained at 37 °C with 5% CO2.

## 2.2 Generation of reporter lines

A 3,064 bp upstream region of human *OCT4* gene was cloned upstream of a *GFP* reporter gene in a N1-EGFP plasmid with Geneticin (Gibco) drug selection marker. Two µg of the *OCT4-GFP* construct was transfected into 10,000 H1 hESCs using 6 µl of Fugene (Roche). Drug resistant colonies appeared after 2 weeks of drug selection. Notably, the efficiency of transfection is very low where only a few green colonies were observed and fewer colonies retained the GFP expression after several passages, indicating that the latter colonies may carry the integrated transgene permanently. Several individual colonies were picked and expanded for further characterization to ensure that the *OCT4* promoter and the reporter gene fragment remained intact. Genomic DNA was extracted and four different pairs of primers priming the *OCT4*-reporter gene junction were used to amplify the region using PCR. Control hESCs that were not subjected to plasmid transfection were negative for the *OCT4*-reporter gene region while the *OCT4*-reporter plasmids that were stably integrated into the genome of hESCs were shown to be positive for the PCR product for the *OCT4*-reporter gene region

## 2.3 Transfection

### 2.3.1 siRNA transfection in 384-well plate

384-well plates (Grenier) were coated with 10 µl of matrigel for 30 mins at 37 °C before removing the excess matrigel. 5 µl of 500 nM pooled siRNAs (siGenome, Dharmacon) or 5 µl of 500 nM individual siRNAs were printed onto the plates and frozen at -20 °C before use. During reverse transfection, a master mix of 0.05 µl of DharmaFect1 (Dharmacon) transfection reagent and 4.95 µl of OptiMEM (Invitrogen) mix was added to siRNA plates and incubated for 20 mins. Subsequently, 3,000 cells in 40 µl of conditioned medium with

10 μM ROCK inhibitor (Calbiochem) were seeded in each well. Reagents and cells were dispensed onto the plate using a multidrop (Thermoscientific) and the above-mentioned volume refers to the amount added to each well.

### 2.3.2 shRNA transfection in 6-well plate

WI siRNA selection program http://jura.wi.mit.edu/bioc/siRNAext/ was used for the design of the shRNAs. hESCs were trypsinized for 30s at 37 °C and passaged in a ratio of 1:9. hESCs were transfected at around 70% confluency with 1.5 μg of shRNA construct (pSuper, Oligoengine) and 4.5 μl of Fugene HD (Roche). 0.8 μg/ml of puromycin was added to the condition medium 24 hrs after transfection. All the knockdown experiments were carried out in biological triplicates

## 2.4 Informatics analysis

### 2.4.1 Data analysis

Z' factor was calculated for the primary screen based on the formula $Z'=1-3(\sigma p+\sigma n)/(\mu p-\mu n)$ where σp= standard deviation of the positive control, σn= standard deviation of the negative control, μp= mean of the positive control and μn=mean of the negative controls. z-score was calculated using the formula $z= (X-\mu)/SD$ where μ is the mean of the negative controls and SD is the standard deviation of the whole population. X is the sample value calculated based on the integrated fluorescent intensity/ number of cells. The Z' factor for the entire screen was 0.76.

### 2.4.2 Gene ontology (GO) analysis

GO analysis was performed with Panther classification (www.pantherdb.org) for the molecular functions and biological processes.

### 2.4.3 Reactome analysis

A web-resource Reactome (www.reactome.org) was used for the analysis of reactions and/or pathways that were statistically over-represented from the 566 genes with z score>2 submitted. The Entrez gene ID of the hits were input as gene-identifiers using Skypainter tool which calculates a one-tailed Fisher's exact test for the probability of observing at least N genes from an event if the event is not over-represented among the 566 genes. Events with p-value<0.05 were statistically significant and over-represented. These events are highlighted in the map with the accompanying genes listed.

### 2.4.4 STRING network analysis

Protein-protein interaction network was generated using STRING database which comprises of known and predicted protein interactions (http://string.embl.de/). 566 genes were input into STRING and 263 genes formed interactions among themselves. A medium confidence score criterion was set for the building of the protein network. Active prediction methods used are experiments, databases and text mining. The resulting network was imported into Cytoscape (www.cytoscape.org). Stem cells and transcription related genes based on Gene Ontology prediction are indicated in green in the Cytoscape. The rest of the genes were indicated in pink. A high confidence score criterion was set for the individual protein complexes; INO80 complex, mediator complex, TAF complex, COP9 signalosome, eukaryotic initiation factor complex and spliceosome complex. Active prediction methods used for these smaller protein-protein network networks were the same as that for the 263 genes protein-protein network. The protein complexes were imported into pathway studio for further text-mining and additional interactions. Pathway studio highlighted the entities that have a z-score of greater than 2 in red.

## 2.5 Staining

### 2.5.1 Alkaline phosphatase staining

Alkaline phosphatase detection was performed using a commercial ESC characterization kit (Chemicon) according to the manufacturer's protocol.

### 2.5.2 Immunostaining

Human ESCs, iPSCs or differentiated cultures were fixed with 4% paraformaldehyde in PBS. After permeablization in 1% triton X-100/PBS for 30 min, immunostaining was performed using the following primary antibodies: NANOG (AF1997, R&D system), OCT4 (ab19857, Abcam), TRA-1-60 (sc-21705, Santa Cruz), TRA-1-81 (sc-21706, Santa Cruz), SSEA-4 (sc-21704, Santa Cruz), NESTIN (ab5968, Abcam ), cardiac actin (10R-C116a, Fitzgerald), SOX17 (sc-17355, Santa Cruz), $p57^{kip2}$ (RB-1637-P, Neomarkers), anti-α-Smooth Muscle Actin (ab18460, Abcam), RUNX1 (ab61753,Abcam), MAFB (sc-22830,Santa Cruz) and IGFBP5 (sc-6006, Santa Cruz). Secondary antibodies used were Alexa Fluor 488/546 anti-mouse IgM, and Alexa Fluor 488/546 anti-mouse or anti-rabbit IgG (Invitrogen). DAPI or Hoechst (Invitrogen) was used for staining the nuclei.

## 2.6 Imaging

Cells were imaged with IXU ultra confocal microscope (Research Instruments) at 20x magnification and 4 frames per well were taken. Integrated fluorescent intensity and nuclei number were quantitated using MetaXpress Image Acquisition and Analysis software V1.7.

## 2.7 Microarray analysis

mRNAs derived from hESCs were reverse transcribed, labelled and analyzed on Illumina microarray platform (HumanRef-8 v3 .0 Expression BeadChips). Arrays were processed according to manufacturer's instructions. For each cell type or cell line, biological replicate microarray data were generated. Rank invariant normalization was used to normalize the microarrays. For *PRDM14* knockdown, mRNAs derived from *PRDM14* shRNA and *luciferase* shRNA-treated H1 hESCs were reverse transcribed, labelled and similarly analyzed on Illumina microarray platform (HumanRef-8 v3 .0 Expression BeadChips). Biological triplicates were included in the profiling of *PRDM14*-depleted H1 cells. Cluster 3.0 was used for hierarchical clustering and Java TreeView for visualization.

## 2.8 Western blot analysis

After 48 hrs transfection, 293-T cells were lysed with RIPA buffer (Pierce) supplemented with protease inhibitor cocktail (Roche). Protein concentration was measured with a Bradford assay kit (Bio-Rad). 50 μg of cell lysate was resolved on a 10% SDS-polyacrylamide gel and transferred to a polyvinylidine difluoride membrane (Millipore). The membrane was blocked with 5% skim milk. After blocking, the blot was incubated with either anti-PRDM14 (1:2000, custom-made), anti-OCT4 (1:5000 Abcam), anti-NANOG (1:800 R&D) or anti-GAPDH (1:5000 Santa-Cruz) primary antibodies for 1 hr, washed with PBST and incubated with either horse-radish peroxidase (HRP)-conjugated anti-rabbit IgG (1:5000, Santa Cruz), HRP-conjugated anti-goat IgG (1:5000, Santa Cruz) or HRP-conjugated anti-mouse IgG (1:5000, Santa Cruz), respectively. After washing with PBST, signals were detected using the Western Blotting Luminol Reagents (Santa Cruz).

## 2.9 RNA extraction, reverse transcription and quantitative real-time RT-PCR

Total RNA was extracted using TRIzol Reagent (Invitrogen) and purified with the RNAeasy Mini Kit (Qiagen). Reverse transcription was performed using SuperScript II Kit (Invitrogen). DNA contamination was removed by DNase (Ambion) treatment, and the RNA was further purified by an RNAeasy column (Qiagen). Quantitative PCR analyses were performed in real time using ABI PRISM 7900 Sequence Detection System and SYBR Green Master MixFor all the primers used, each gave a single product of the right size. In all our controls lacking reverse transcriptase, no signal was detected (Threshold cycle (Ct) 435). The real-time RT-PCR conditions was at 95 °C for 10s followed by 40 cycles at 95 °C for 10s and 60 °C for 1 min. Each RNAi experiment was performed in triplicates. The forward and reverse primers used in my thesis are indicated in the table below.

| | |
|---|---|
| JMJD2B | CGCGGCAGACGTATGATGACATCGACGACGTG<br>CACGTCGTCGATGTCATCATACGTCTGCCGCG |
| PRDM14 | CTGTTACCTGAGGGGCTGAG<br>CTCAGCCCCTCAGGTAACAG |
| WDR82 | GGGATCTCCGGTCTCCTAAC<br>GTTAGGAGACCGGAGATCCC |
| NFRKB | GCAGGAAGTGTTAAGTGATTCTCAACGTGAAC<br>GTTCACGTTGAGAATCACTTAACACTTCCTGC |
| SERTAD2 | GAGCTCTGTCCCACATCTACCTC<br>GAGGTAGATGTGGGACAGAGCTC |
| RUNX1 | CCCTAGGGGATGTTCCAGAT<br>ATCTGGAACATCCCCTAGGG |
| MAFB | GCCATGGAGTATGTCAACGA<br>TCGTTGACATACTCCATGGC |
| IGFBP5 | GACCGCAGAAAGAAGCTGAC<br>GTCAGCTTCTTTCTGCGGTC |
| OCT4 | TTGTGCCAGGGTTTTTGGGA |

| | |
|---|---|
| | TCCCAAAAACCCTGGCACAA |
| DPPA4 | TGTGGTGACAACTTCTGCCCCAG<br>CTGGGGCAGAAGTTGTCACCACA |
| HESX1 | TAGAGGCCGAAGACCAAGAA<br>TTCTTGGTCTTCGGCCTCTA |
| TDGF1 | ACAGAACCTGCTGCCTGAAT<br>ATTCAGGCAGCAGGTTCTGT |
| SEMA6A | CCTGGACACCAGTTCCTGAT<br>ATCAGGAACTGGTGTCCAGG |
| NANOG | TGCAGTTCCAGCCAAATTCTC<br>GAGAATTTGGCTGGAACTGCA |
| SOX2 | AACCCCAAGATGCACAACTC<br>GAGTTGTGCATCTTGGGGTT |
| HELLS | GCTTGATGGGTCCATGTCTT<br>AAGACATGGACCCATCAAGC |
| N-MYC | CTAGAGCGCGCAGTGAACGA<br>TCGTTCACTGCGCGCTCTAG |
| LIN28 | CCCCCAGTGGATGTCTTTGTGCACCAGAGTA<br>TACTCTGGTGCACAAAGACATCCACTGGGGG |
| HEY1 | CGAGGTGGAGAAGGAGAGTG<br>CACTCTCCTTCTCCACCTCG |
| SALL2 | AGAGAGCAGCAGCAGAAAGG<br>CCTTTCTGCTGCTGCTCTCT |
| ETV4 | CGCCTACGACTCAGATGTCA<br>TGACATCTGAGTCGTAGGCG |
| HESX1 | TAGAGGCCGAAGACCAAGAA<br>TTCTTGGTCTTCGGCCTCTA |
| ACTA2 | TTCAATGTCCCAGCCATGTA<br>TACATGGCTGGGACATTGAA |
| NASP | TCTCCTTGCAGAGACCCACT<br>AGTGGGTCTCTGCAAGGAGA |
| PHF17 | GCAGCGATGCTACGACAATA<br>TATTGTCGTAGCATCGCTGC |
| ZSCAN2 | GACTTGACCCAGACCCTTCA<br>TGAAGGGTCTGGGTCAAGTC |
| MAP7 | GGAAGAGCGGAAGAAGAGGT<br>ACCTCTTCTTCCGCTCTTCC |

| | |
|---|---|
| AK3L1 | GGTCTCCAGCATCTCTCCAG<br>CTGGAGAGATGCTGGAGACC |
| TGIF2 | GAAGCAACGGGACCAATCGC<br>GCGATTGGTCCCGTTGCTTC |
| BCL11A | GATAAGCCACCTTCCCCTTC<br>GAAGGGGAAGGTGGCTTATC |
| ZNF649 | GAATCACTGACCCTGGAGGA<br>TCCTCCAGGGTCAGTGATTC |
| FOXO4 | GCCTGGGGAAATCAGTCATA<br>TATGACTGATTTCCCCAGGC |
| SEMG1 | GAAAGCACAGGGCAAGTCTC<br>GAGACTTGCCCTGTGCTTTC |
| TCF7L1 | AGCCCATGTACTCCCTTCCT<br>AGGAAGGGAGTACATGGGCT |
| CDA | TGAAGCCTGAGTGTGTCCAG<br>CTGGACACACTCAGGCTTCA |
| HOOK2 | GGCTTCAGCTGGAGAACAAG<br>CTTGTTCTCCAGCTGAAGCC |
| FOXO1 | GCTGCCAAGAAGAAAGCATC<br>GATGCTTTCTTCTTGGCAGC |
| FOXD1 | CTCGTATATCGCGCTCATCA<br>TGATGAGCGCGATATACGAG |
| PAPPA | TGGCCTCCATCCTACATCTC<br>GAGATGTAGGATGGAGGCCA |
| PTHLH | CCCTCTCCCAACACAAAGAA<br>TTCTTTGTGTTGGGAGAGGG |
| PRICKLE2 | TGACAACGAGGGCTATTTCC<br>GGAAATAGCCCTCGTTGTCA |
| ZFHX3 | ACGTGAAGAAGGAGCCACTG<br>CAGTGGCTCCTTCTTCACGT |
| OAF | GGAGCAAGGTGTGGACAGTT<br>AACTGTCCACACCTTGCTCC |
| COL1A2 | CTGCAAGAACAGCATTGCAT<br>ATGCAATGCTGTTCTTGCAG |
| PHC | CATTGTGAAACCCCAAATCC<br>GGATTTGGGGTTTCACAATG |
| RUNX3 | TACGGTGGTGACTGTGATGG |

| | |
|---|---|
| | CCATCACAGTCACCACCGTA |
| BTG2 | AAGATGGACCCCATCATCAG<br>CTGATGATGGGGTCCATCTT |
| TNFRSF19 | CTATGGGGAGGATGCACAGT<br>ACTGTGCATCCTCCCCATAG |
| CALCA | TTCCTGGCTCTCAGCATCTT<br>AAGATGCTGAGAGCCAGGAA |
| GPRC5C | CGTAAGCATGGGGTCTTTGT<br>ACAAAGACCCCATGCTTACG |
| PHC2 | AATGTGGGTCTCAACCTTCG<br>CGAAGGTTGAGACCCACATT |
| PAG1 | CTGATGAACGTGCCTTCAGA<br>TCTGAAGGCACGTTCATCAG |
| SEMA3C | ATATGCGAACCACCAAGGAG<br>CTCCTTGGTGGTTCGCATAT |
| LHX1 | ATCCTGGACCGCTTTCTCTT<br>AAGAGAAAGCGGTCCAGGAT |
| MSX1 | TCCTCAAGCTGCCAGAAGAT<br>ATCTTCTGGCAGCTTGAGGA |

# 3 Results

Majority of the following results section has been published in the Journal Nature and is entitled:

## A genome-wide RNAi screen reveals determinants of human embryonic stem cells

Chia NY, Chan YS, Feng B, Lu X, Orlov YL, Moreau D, Kumar P, Yang L, Jiang J, Lau MS, Huss M, Soh BS, Kraus P, Li P, Lufkin T, Lim B, Clarke ND, Bard F, Ng HH.
Nature. 2010 Nov 11; 468(7321):316-20. Epub 2010 Oct 17. PMID: 20953172

### 3.1 Generation of reporter lines

Aiming to ascertain genes that are novel *bona fide* regulators of pluripotency in hESCs, I will perform genome-wide siRNA screen to identify genes whose knockdown prompted the differentiation of hESCs. As aforementioned, *OCT4* is a cardinal marker of undifferentiated hESCs and mESCs that also marks various pluripotent cell types in the embryo; it shows largely restricted expression to ESCs but not most differentiated cell types. Moreover, *OCT4* is generally downregulated within several days of ESCs embarking on differentiation[74; 75] and hence, its expression in hESCs may be considered a proxy to mark undifferentiated and self-renewing hESCs; siRNA targeting important pluripotency factor should presumably lead secondarily to downregulation of *OCT4* and differentiation. To remotely monitor *OCT4* expression in hESCs, I constructed an *OCT4-GFP* reporter wherein the *GFP* fluorescent reporter was placed under the control of a human *OCT4* cis-regulatory (CR) module carrying the proximal promoter (CR1), the proximal enhancer (CR2 and CR3), and the distal enhancer (CR4) (Figure 12). In the mouse, the proximal and distal enhancers are specifically and selectively activated in

pluripotent cells in the embryo and it is often used in the assessment of pluripotency in mESCs[67; 68; 69]. Likewise to mouse, human *OCT4* cis-regulatory module reporter is much more active in undifferentiated pluripotent hESCs than differentiated cell types and can be used to monitor the differentiation status of hESCs[76].

Previous studies have defined regulatory regions that are important for driving *Oct4* expression in different cell types of the early mouse embryo through the analysis of *lac*Z reporter genes under the control of different mouse *Oct4* genomic fragments[77]. The mouse *Oct4* genomic fragments include the core promoter which is located within the first 250 bp of the transcription initiation site and 4 enhancer regions named conserved region (CR) 1, 2, 3 and 4 that were conserved among species like human, bovine, and mouse (Figure 12). CR1 is located at the proximal promoter, CR2 and CR3 is within the proximal enhancer and CR4 is within the distal enhancer of the *Oct4* promoter. A proximal enhancer with CR2 is located about 1.2 kb upstream and is responsible for *Oct4* expression in the epiblast and a distal enhancer region with CR4 is located about 2 kb upstream and is responsible for driving *Oct4* expression in the morula, ICM, PGCs as well as in mESCs. Several pluripotency factors have been reported to regulate the different conserved regions of the *Oct4* regulatory region. For instance in mESCs, Oct4, Sox2 and Zfx bind to CR4 while Lrh1 binds to CR2 and Nanog binds to both CR2 and CR4.

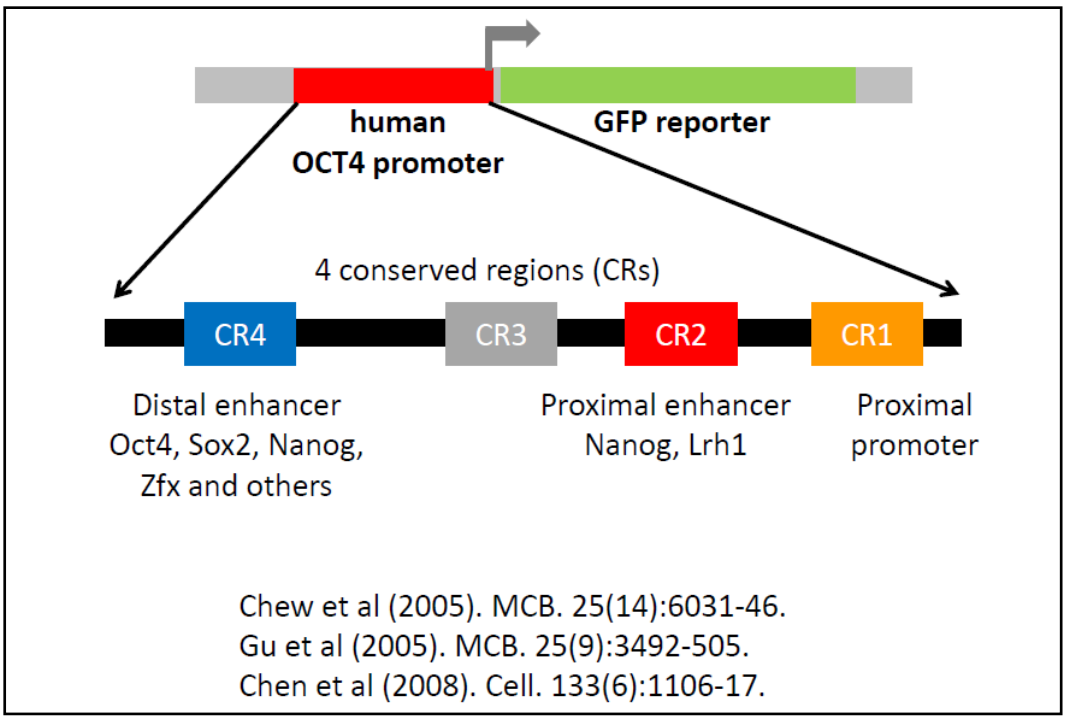



**Figure 12: schematic diagram of the human *OCT4* promoter that was used in the *OCT4-GFP* construct.**
The promoter contains the 4 conserved regions CR1, CR2, CR3 and CR4 that are bound by the different factors.

In our previous study[78] to determine the functional importance of the conserved regions (CR1 to CR4), Chew *et al* cloned a human *OCT4* 3kb regulatory fragment that contains all conserved regions (CR1 to CR4) from a bacterial artificial chromosome (BAC) and fused it to an upstream Renilla luciferase (*Luc*) gene plasmid[78]. This *OCT4* promoter is appropriate for use in my screen as this promoter comprises of important regulatory regions that will be required to screen for *OCT4* regulators. I cloned this *OCT4* regulatory region into an EGFP-reporter plasmid and transfected it into hESCs to create stable cell lines (Figure 13). The use of a stable reporter cell line in a high throughput screening negates the need of staining with antibody. This could effectively reduce the complication and cost in the screen. Furthermore, with the use of a stable reporter line, the

homogeneity of the reporter gene expression from experiments to experiments could be improved than with immunostaining.

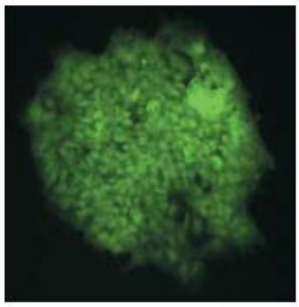


**Figure 13: *OCT4-GFP* reporter line**
Fluorescent image of H1 hESC line with a GFP reporter gene driven by the POU5F1 promoter was used for the screen.

### 3.1.1 *OCT4-GFP* reporter lines

Several *OCT4-GFP* colonies were picked and subjected to GFP fluorescent activated cell sorting (FACS) to enhance the purity of GFP positive cells. These colonies were further expanded and the brightest colony was selected for further characterization. In order to ensure that the GFP expression of the *OCT4-GFP* reporter cells correlates faithfully with the state of ESCs and the quantification of GFP could be used as a proxy to monitor the differentiation status; these cells were induced to differentiate. I seeded these *OCT4-GFP* stable hESCs in a 384-well plate and half of the plate was subjected to retinoic acid (RA) treatment. From the montage, the stable lines lost its GFP expression under the RA treatment but not the untreated cells (Figure 14). In addition, *OCT4* siRNA knockdown resulted in a decreased in GFP expression as indicated by both image analysis as well as FACS analysis. On the other hand, the negative control siRNA; non-targeting (NT) siRNA which does not target any known human genes does not affect the GFP expression (Figure 15). It is imperative that the reporter cells' genomic integrity is not comprised with the

introduction of this *OCT4-GFP* transgene and from the karyotype analysis that is performed by Cacheux-Rataboul V, the *OCT4-GFP* reporter cells do not lead to chromosomal abnormalities as they have 46XY chromosomes (Figure 16, top). Additionally, these cells remained pluripotent as it is able to form teratoma of the different lineages; ectoderm, mesoderm and endoderm upon injection into a SCID mouse (Figure 16, bottom).

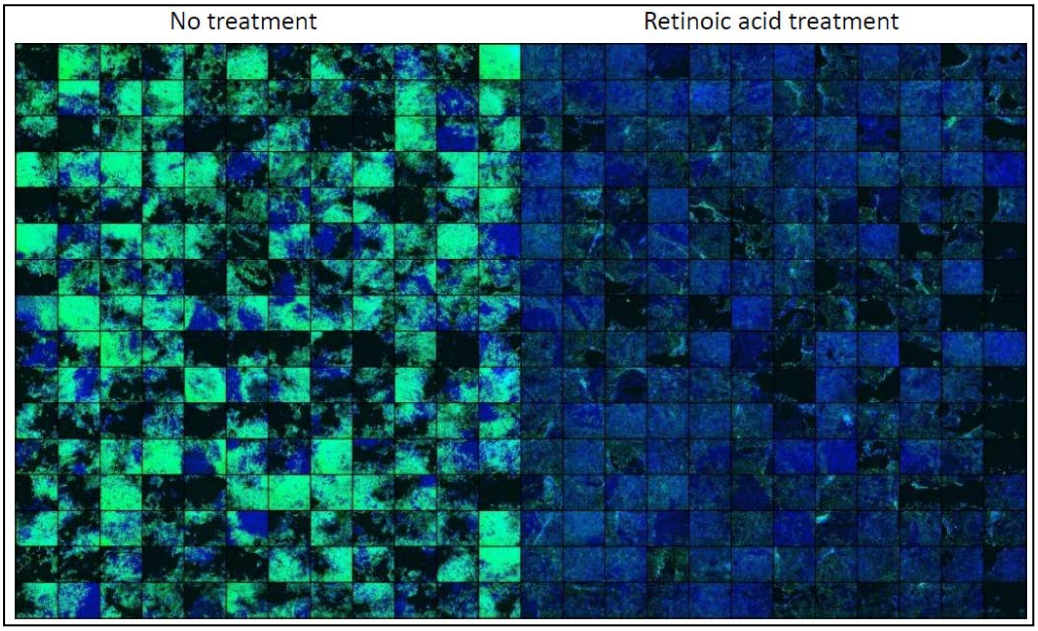



**Figure 14: Montage of *OCT4-GFP* cells in a 384 well.**
RA treatment turns off the GFP expression of *OCT4-GFP* reporter cells (right) while these cells express GFP under non-differentiating condition (left). Some wells in the non-differentiation condition appear to be non-green due to self-differentiation when the wells are over-crowded.

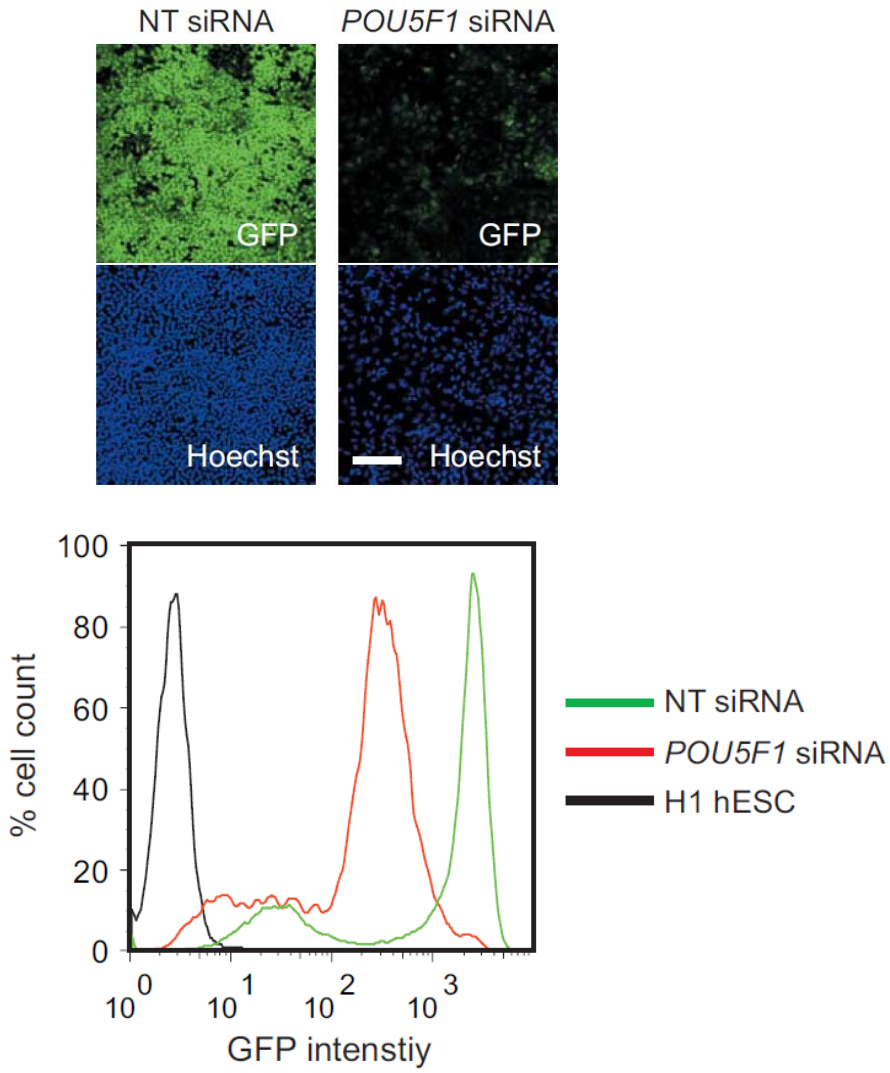
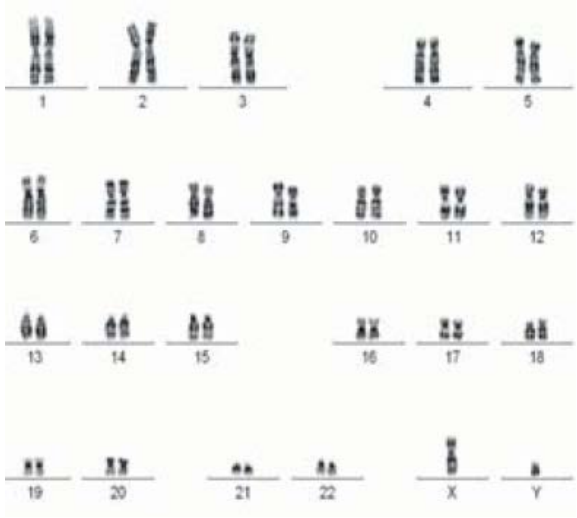
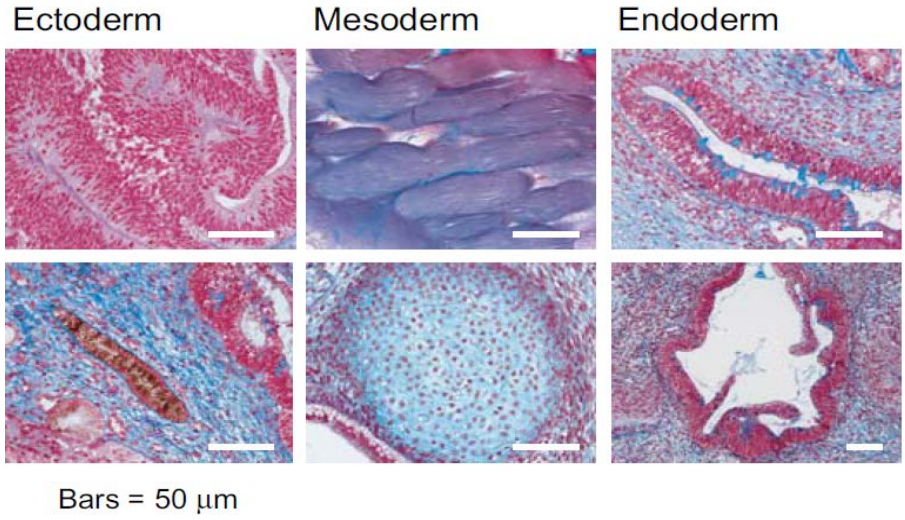



**Figure 15: Reduction of GPF is observed when *OCT4-GFP* hESCs differentiates**
Image analysis (top) and fluorescent activated sorting quantification (bottom) of GFP expression of cells subjected to *NT* siRNA and *OCT4* siRNA knockdown. Knockdown of *OCT4* resulted in the loss of GFP expression from both image analysis and FACs quantification



by Cacheux-Rataboul V




by Petra Kraus and Thomas Lufkin

**Figure 16: Karyotype and teratoma images**
Karyotypic analysis of *OCT4-GFP* cells (top). Teratoma images of *OCT4-GFP* cells that were injected into SCID mice. Tissues from the three different lineages, ectoderm, mesoderm and endoderm were obtained (bottom).

## 3.2 Optimization procedures

Subsequent to the generation of *OCT4-GFP* reporter cell lines will be to establish parameters for robust and reproducible measurements that can be used for high-throughput siRNA screen. This will include the optimization of a number of important experimental variables such as reverse transfection, number of cells and amount of transfection reagent to be used.

### 3.2.1 Reverse transfection

I adopted a reverse transfection method for the optimization procedure (Figure 17). In reverse transfection, the order of addition of DNA and adherent cells is reversed to that of conventional transfection[79]. This was necessary for my screen as the plates were first printed with siRNAs and frozen before use. This method has higher transfection efficiency albeit higher cell toxicity. This was compensated with the supplementation of ROCK inhibitor which had been reported to protect hESCs from apoptosis[80]. ROCK inhibitor was tested on different cell number ranging from 3000, 6000 and 10 000 cells. In the presence of ROCK inhibitor, there was an increase in survival rate for hESCs by 8-fold from the different initial seeding cell number (Figure 18).

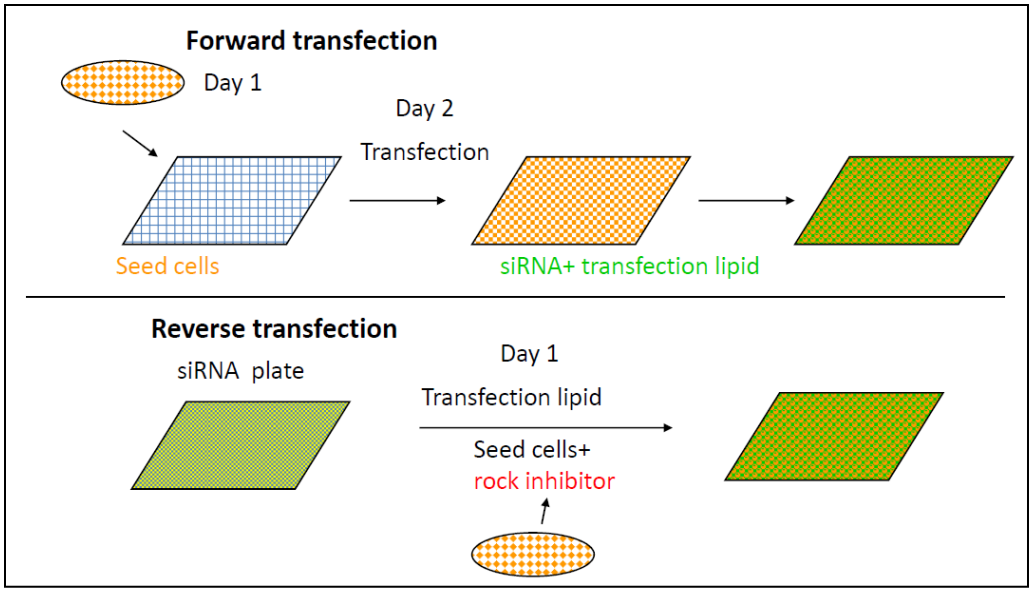
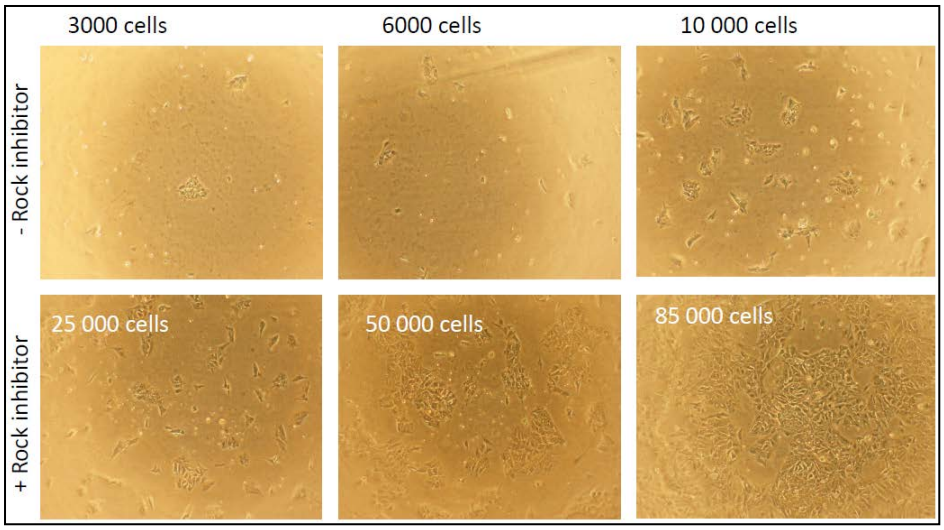



**Figure 17: Schematic diagram of forward and reverse transfection.**
In a forward transfection, the cells are seeded and allowed to attach to the plate prior to transfection. In a reverse transfection, the transfection lipid is added to the plate and immediately after, the cells in the presence of ROCK inhibitor are added.



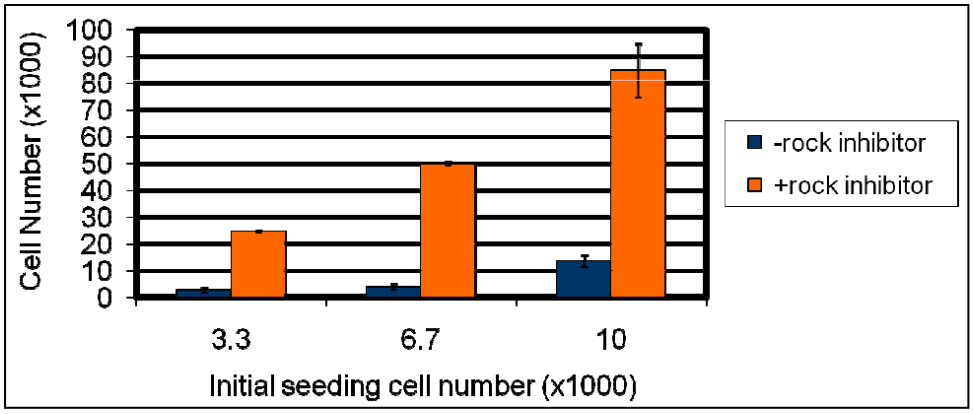
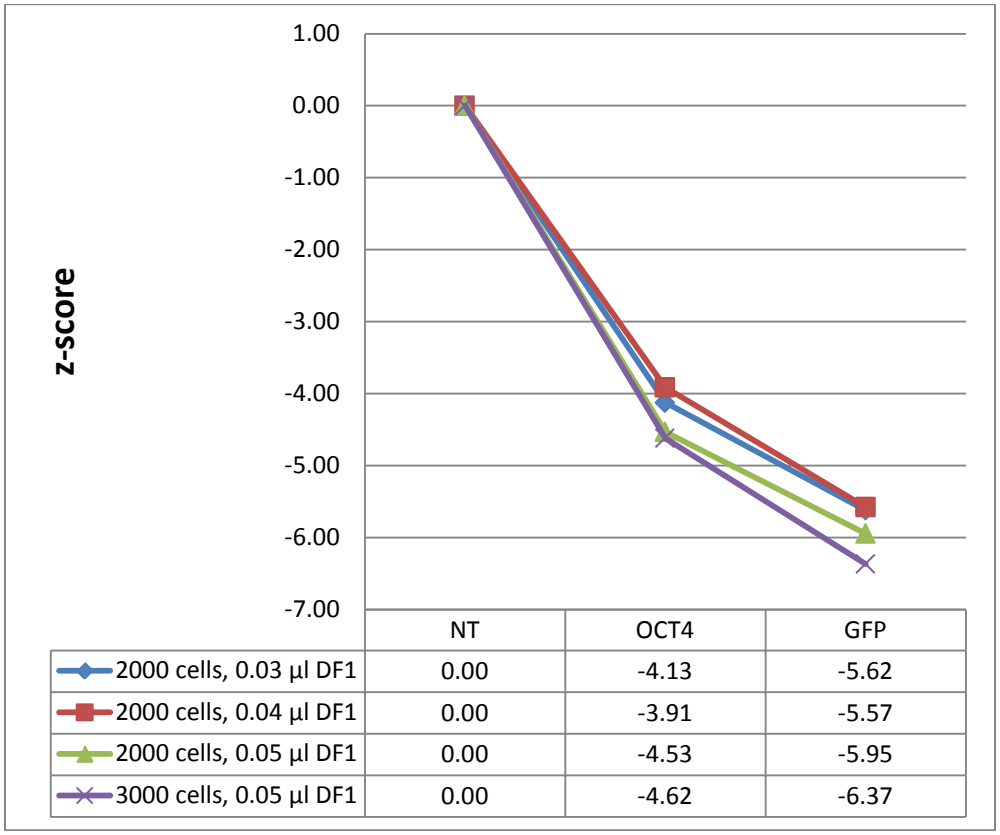



**Figure 18: ROCK inhibitor protects the cells from apoptosis**
3300, 6700 and 10 000 cells were seeded and the survivability was examined in the presence and absence of ROCK inhibitor. Bright field images of the cells were taken (top) and the cell number under different seeding number was quantified (bottom).

### 3.2.2 Number of cells and volume of transfection lipid

DharmaFect transfection lipid was the recommended lipid for the siRNAs that will be used in the subsequent screening assays. DharmaFect is available in four formulation and all four of them were tested for its efficiency in siRNA transfection in hESCs. Only DharmaFect-1 resulted in greatest transfection efficiency as determined by the strongest lost in GFP expression when OCT4 siRNAs were transfected into *OCT4-GFP* hESCs (results not shown). Next I went on to test the number of *OCT4-GFP* hESCs that were required for an efficient transfection in the 384-well format. Negative control: Non-targeting (NT) siRNA and positive controls: GFP siRNA and OCT4 siRNA were used. Each control was seeded in triplicate in the designated wells and the transfection efficiency was calculated based on the average z-score for each control. A starting number of 2000 hESCs were used for the titration of DF1 volume of 0.03 µl, 0.04 µl and 0.05 µl. Under the condition of 2000 cells with 0.05 µl DF1, the loss of GFP in the presence of OCT4 siRNA and GFP siRNA knockdown is the greatest as compared to

0.03ul and 0.04 µl DF1 used (Figure 19). As some toxicity is observed under this condition, I increased the number of cells to 3000 and I obtained an improvement in transfection efficiency. The z-score of OCT4 siRNA and GFP siRNA under the condition of 2000 cells with 0.03, 0.04 and 0.05 µl and 3000 cells with 0.05 µl was shown in Figure 19. A DF1 volume greater than 0.05 µl resulted in cell toxicity and therefore, 3000 cells and 0.05 µl of DF1 is the optimal condition for transfection.




| | NT | OCT4 | GFP |
|---|---|---|---|
| 2000 cells, 0.03 µl DF1 | 0.00 | -4.13 | -5.62 |
| 2000 cells, 0.04 µl DF1 | 0.00 | -3.91 | -5.57 |
| 2000 cells, 0.05 µl DF1 | 0.00 | -4.53 | -5.95 |
| 3000 cells, 0.05 µl DF1 | 0.00 | -4.62 | -6.37 |

**Figure 19: Optimization of cell number and amount of DF1**
Average z-score of each of the positive controls, in the presence of different amount of DF1 is represented in the graph and table.

### 3.3 Differentiation assay

In order to assess the differentiation status of the hESCs, it is desirable to examine more parameters in addition to the loss of GFP expression in the *OCT4-GFP* reporter lines. As such, scrutinized for morphological changes through Phalloidin staining (Figure 20). Phalloidin is a bicyclic peptide that belongs to a family of toxins isolated from the deadly *Amanita phalloides* “death cap" mushroom. It is commonly applied in imaging applications to selectively label F-actin. I differentiated the cells with OCT4 siRNA knockdown and compared the actin distribution between differentiated and undifferentiated cells. A less expanded actin distribution was observed in differentiated cells where actin distribution was quantified at around 70% with respect to undifferentiated cells via NT siRNA and GFP siRNA (Figure 21). Based on actin node distribution, the change in hESCs transfected with NT siRNA and OCT4 siRNA was 30% as compared to the GFP reporter activity where the change in expression in hESCs transfected with NT siRNA and OCT4 siRNA was 70%. Notably, the decrease in actin distribution was not as apparent as compared to the change in GFP expression in differentiated cells. This parameter for actin distribution was comparatively less sensitive in this differentiation assay assessment as it has a smaller dynamic range between differentiated and undifferentiated cells. As such, it may not be an efficient indicator for hESC differentiation and hence we focused on GFP expression as an assessment for differentiation.

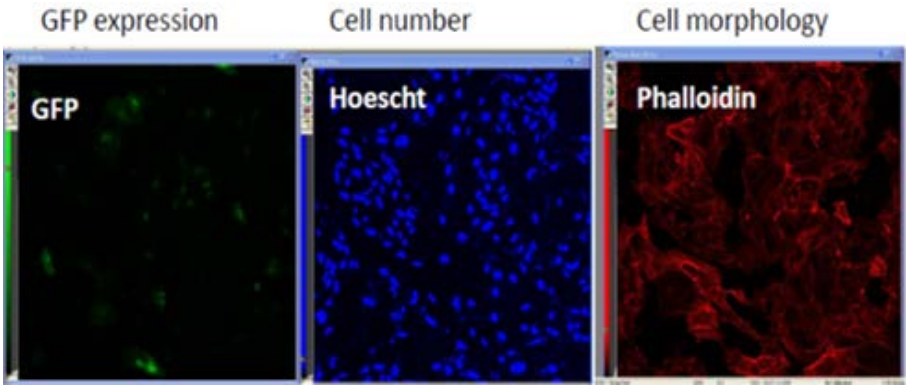



**Figure 20: Different transflour module was tested.** *OCT4-GFP* hESCs were subjected to OCT4 knockdown as well as GFP knock down and measured for GFP, Hoescht staining and phalloidin staining. GFP measures the OCT4 reporter activity, Hoescht staining measures the cell number and phalloidin staining determines the cell morphology.

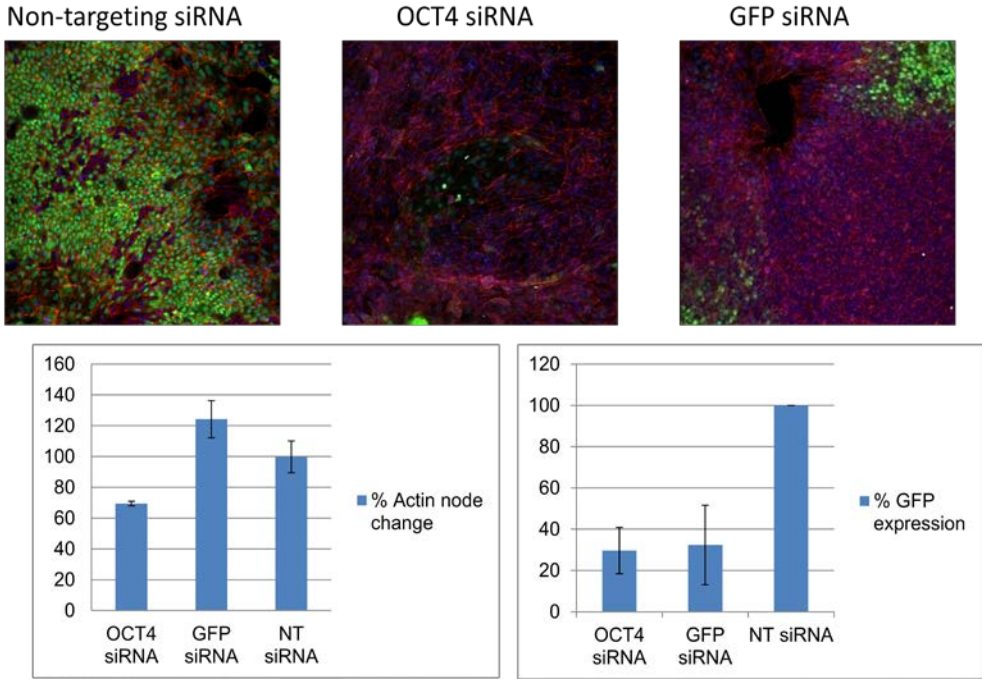



**Figure 21: Phalloidin analysis of *OCT4-GFP* hESCs**
Phalloidin staining (red) of *OCT4-GFP* hESCs (top). The graph measures the percentage of GFP expression of both the actin node and the GFP expression in the presence of non-targeting (NT), OCT4 and GFP siRNAs (bottom). GFP expression is quantified from the confocal image analysis. The GFP expression from the well NT siRNA knockdown is at 100% and the % GFP expression of OCT4 siRNA and GFP siRNA is compared with respect to NT siRNA.

### 3.4 Survival of hESCs' assay

In addition to the pluripotency parameter, self-renewal could be examined by quantifying cell number via Hoescht staining. The Hoechst stains are part of a family of fluorescent stains that labels DNA in fluorescence microscopy and fluorescence-activated cell sorting (FACS) and are commonly used to visualize nuclei and mitochondria. From the quantification of cell number, it is possible to identify genes that affect survival of hESCs when depleted. On the cautionary note, this parameter does not distinguish genes that control cell survival due to their role in the general housekeeping function or their role in the self-renewal of ESCs. The data has to be scrutinized carefully to filter off the housekeeping genes. As genome-wide siRNA screen has been performed on other human cell line such as Hela cells, siRNA that causes cell death in both hESCs and Hela cells may suggest that this gene may not be specific to hESCs can may be excluded for further studies. Alternatively, a counter screen could be performed and this will be elaborated in detail in the section 3.5.4.

### 3.5 siRNA screens

#### 3.5.1 Kinome Pilot screen

The kinome library is a smaller library than the whole-genome library, it comprises of siRNAs targeting 799 genes and they are distributed across three 384-well plates with control siRNAs that are manually added in the designated wells. A pilot screen replicates the high-throughput screening conditions that will be used in the actual screen and this will allow us to presage how well the whole-genome screen will fare through the statistical calculation of Z'-factor (Box 1). From this kinome screen, the Z' Factor is calculated to be 0.56. A Z' Factor of greater than 0.5 indicates that the assay is robust for a

high throughput screen as we will be able to distinguish the "hits" from the background. It is also indicative of an optimal condition for the actual screen.

| Z-factor | Interpretation |
|---|---|
| 1.0 | Ideal. Z-factors can never exceed 1. |
| between 0.5 and 1.0 | An excellent assay. Note that if $\sigma_p = \sigma_n$, 0.5 is equivalent to a separation of 12 standard deviations between $\mu_p$ and $\mu_n$. |
| between 0 and 0.5 | A marginal assay. |
| less than 0 | There is too much overlap between the positive and negative controls for the assay to be useful. |

**Box 1: Interpretations for the Z-factor, adapted from Zhang, *et al*. 1999[81]**

From the kinome screen, the positive controls (OCT4 siRNA and GFP siRNA) showed a reduction in GFP but not the negative control (NT siRNA) from the image analysis. This indicates that the optimal screening condition was attained and with that, one of the kinases known as CSNK1A1 was identified as a potential candidate (Figure 22).

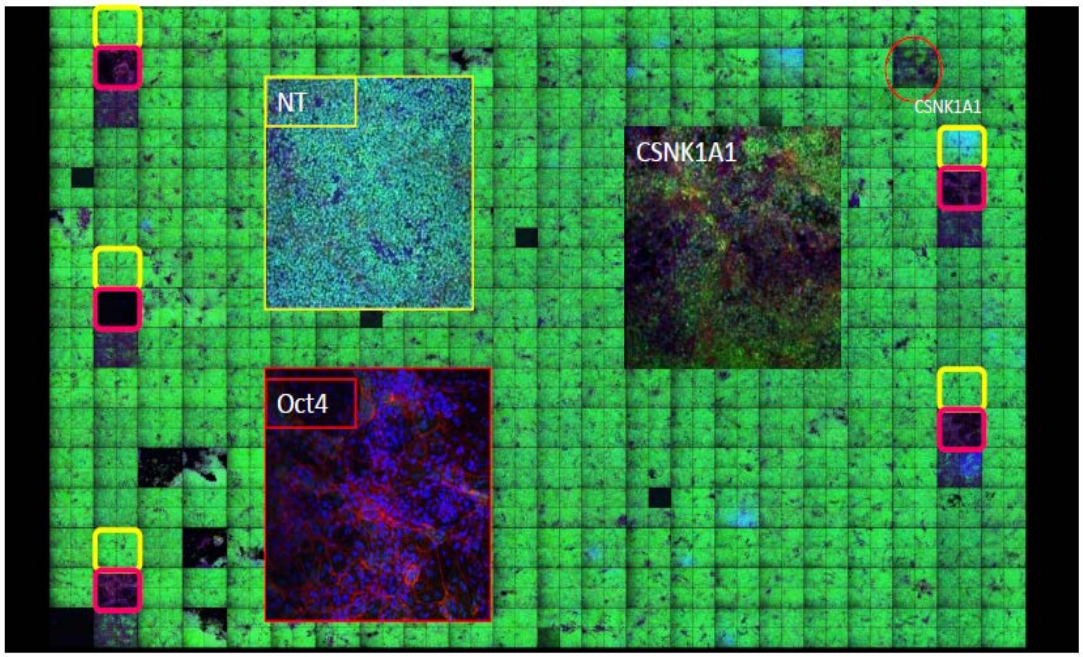



**Figure 22: Montage of the kinome screen**
The NT siRNA and OCT4 siRNA were added in the yellow box and red box respectively. GFP siRNA was added in the well below OCT4 siRNA. CSNK1A1 was one of the potential hits from the screen.

### 3.5.2 Primary screen: Genome wide siRNA screen

With the success of the preliminary run on the kinome library, I commenced on the actual screen with the whole genome siRNA library. This library comprises of siRNA against 21,121 genes involving the kinases, phosphatases, G-protein coupled receptors (GPCR) druggable genes and all annotated genes. Each gene is targeted by a mixture of four siRNAs that recognizes the different parts of the same gene and these siRNAs were pooled together in each well to maximize siRNA knockdown effect thereby reducing the chances of getting false negatives in the initial phase of the actual screen. The siRNAs were distributed in 67 individual 384-well plates and the screen was carried out in duplicate. The plates were assayed for GFP intensity and cell numbers and GFP expression per cell was quantified for each well as represented in the dot plot (Figure 23). We calculated the Z' Factor from all plates and obtained a good value of 0.76 which indicates that the assay is working well and with a good GFP difference between the postitive and negative controls. We represented the RNAi data for each targeted gene in terms of positive z-score which indicates the number of standard deviation in the reduction of GFP fluorescent from the negative control. A cutoff with an average z-score of +2 was used to determine the genes that reduce GFP expression (Fav) from the negative control. We obtained 566 potential "hits" that downregulate GFP expression and it is reassuring that *OCT4* which is essential for the maintenance of ESCs was ranked first on the list, with a z-score of 5.14. In addition to that, we identified several regulators that were implicated in mESCs, including *HCFC1*, *TCL1A*, *ZSCAN10*, *ZIC3*, *NANOG* and *ZNF143* to be among the top 5% of the gene list (Table 1).

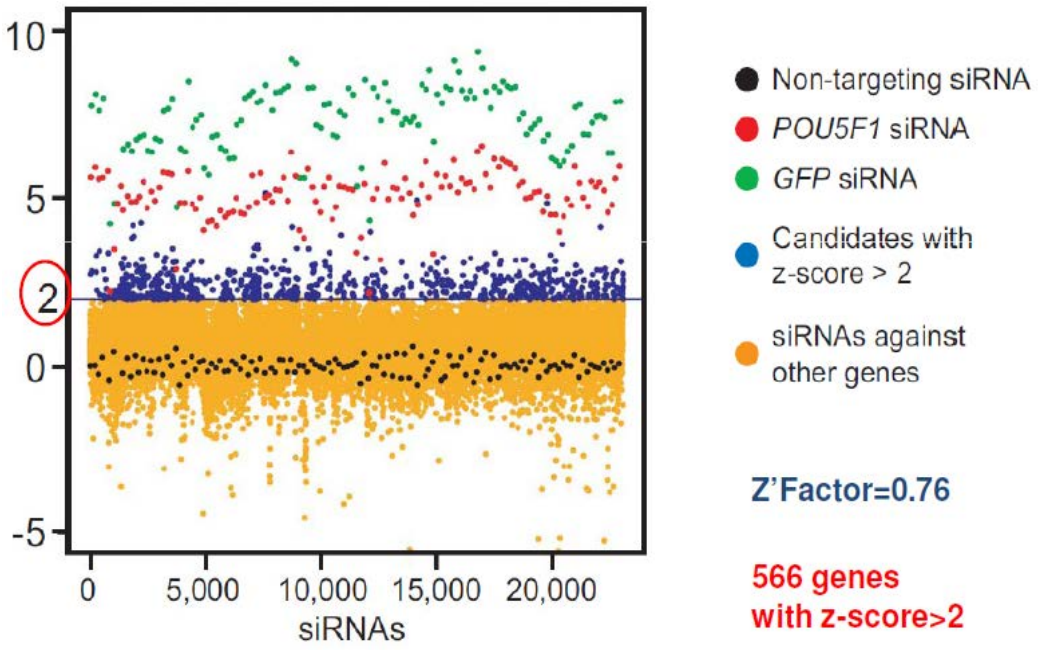
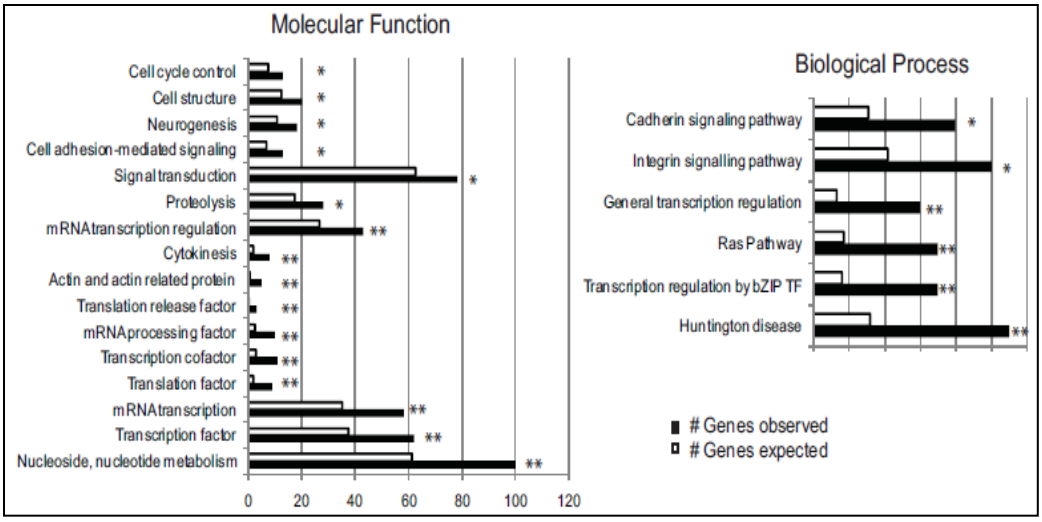



**Figure 23: Dot plot of the genome-wide RNAi screen.**
The y-axis represents the average z-scores for the GFP reduction for each targeted gene. Controls are represented by the black (non-targeting siRNA), red (OCT4 siRNA) and green (GFP siRNA) dots. Genes with z-score >2, highlighted in blue, are potential candidates required for the maintenance of hESC identity. The rest of the genes are indicated in orange.

In order to understand the characteristics and roles of these 566 factors that are essential for hESCs, several pathway analysis tools were utilized. Yuriy L. Orlov utilised Gene Ontology (GO) to describe these genes in terms of their associated biological processes and molecular functions. A great proportion of these 566 genes were enriched in the categories of transcription factors or transcription related factors and many of them were involved in nucleoside or nucleotide metabolism, signal transduction and signalling like cadherin and integrin pathways (Figure 24). In addition to gene ontology, Reactome skypainter was utilized. It is a tool for determining the reactions or pathways that are statistically overrepresented in a set of genes. Similarly, we detected events involved in transcription and translation and other events that were involved in cell cycle, cell

migration and metabolism (Figure 25, Table 3). In addition to the above knowledge acquired, we proceeded with further bioinformatics analysis to study protein-protein interactions. Pankaj Kumar carried out analysis on the 566 genes and 263 were found to have protein-protein interactions between/among themselves and these interactions were statistically more significant with a value of 0.192 compared to a random interaction of 0.034 (Figure 26a). Interestingly, I identified genes coding for proteins in known biochemical complexes which have not been previously implicated for its importance in ESCs maintenance. For instance, the INO80 chromatin remodelling complex, the mediator complex[3], the COP9 signalosome[4] and the TAF complex[5] were discovered (Figure 26b). In the context of mESCs, INO80 chromatic remodelling complex has been reported to be associated with OCT4 from the OCT4 protein interactome and Mediator has been shown to physically and functionally connect the enhancers and core promoters of active genes coupled with Cohesin[82]. COP9 signalosome can act as a transcriptional regulator and TAFs belong to the classes of coactivators that regulate transcription.




by Yuriy L. Orlov

**Figure 24: Gene Ontology analysis of the 566 genes with z-score>2.**
Graphs represent the functional categorization of the biological process and molecular function categories that are over-represented. Under the molecular function, a great proportion of these 566 genes were enriched in the categories of signal transduction, mRNA transcription regulation, transcription factors and transcription related factors and many of them were involved in nucleoside or nucleotide metabolism. The biological process of these genes are mainly in signalling pathways such as cadherin and integrin. Categories with p-value<0.05 are indicated with * and categories with p-value<0.01 are indicated with **.

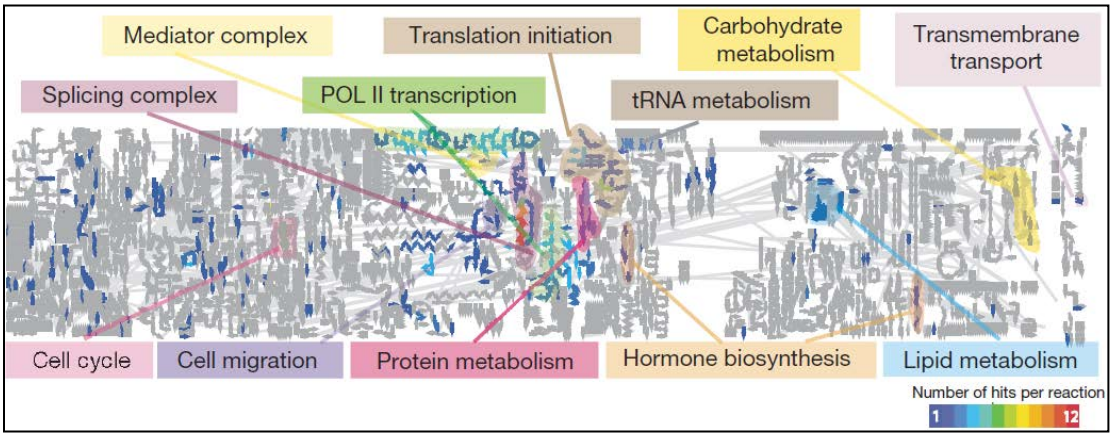
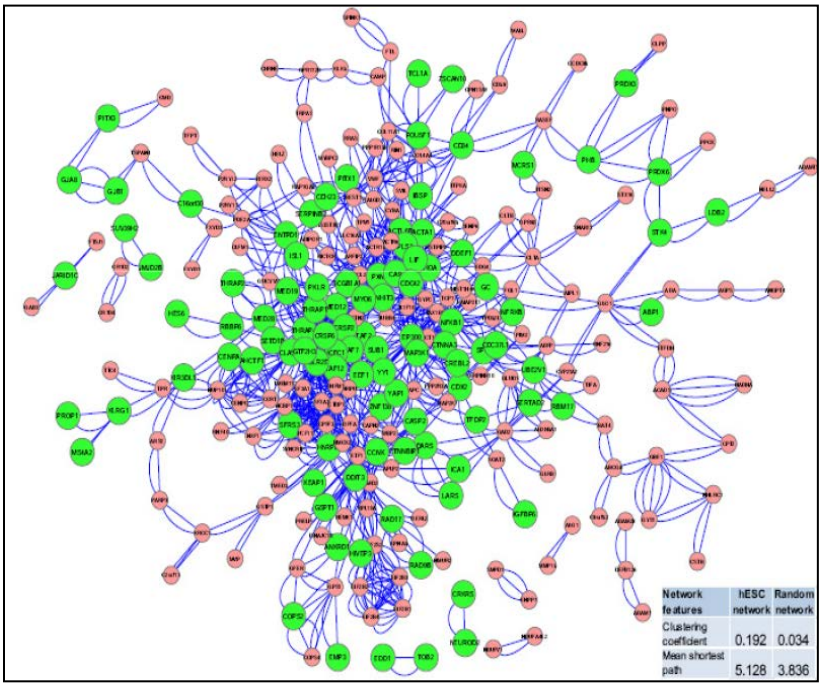



**Figure 25: Reactome analysis.**
REACTOME is an open-source, open access, manually curated and peer-reviewed pathway database. The reactions/pathways of the 566 genes (identifiers) were analysed using this web resource to determine those that were statistically over-represented. Twelve categories of reactions/pathways with P-value,0.05 were over-represented. The colour of each reaction arrow on the reaction map indicates the the number of genes in the submitted list that participates in the reaction. For instance, blue means 1 matching identifier and red means 12 matching identifiers).

A)




by Pankaj Kumar

B)

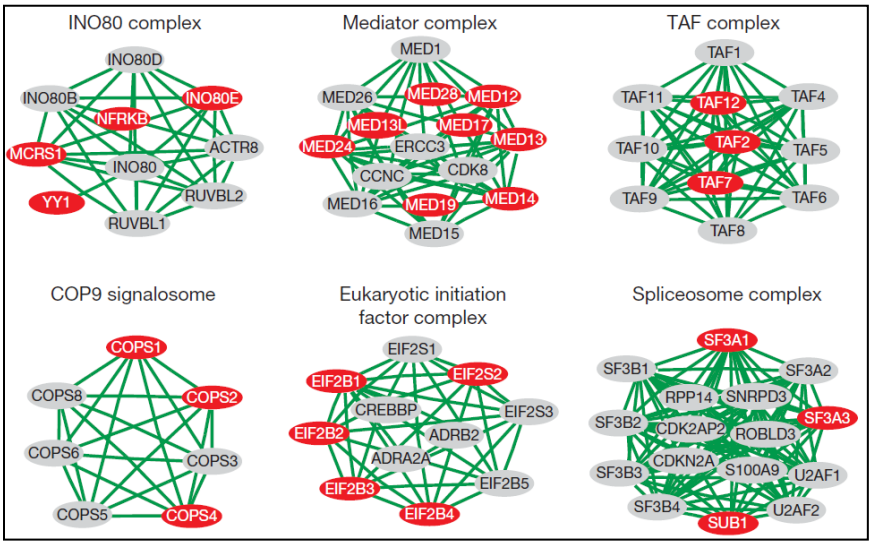
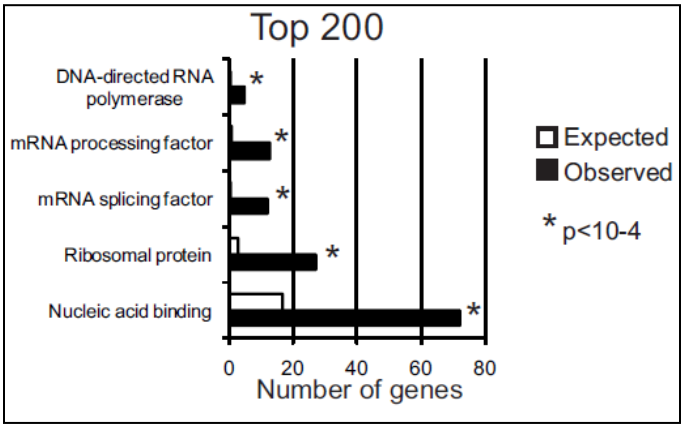



**Figure 26: Protein-protein interactions of the 566 genes**

**A)** 263 out of the 566 genes have protein-protein interactions between/among themselves and these interactions were statistically more significant with the value of 0.192 compared to a random interaction of 0.034. Transcription factors/ ES related genes are indicated in green while the rest of the genes are indicated in pink.

**B)** Components of the INO80 chromatin remodelling complex, mediator complex, TAF complex, COP9 signalosome, eukaryotic initiation complex and spliceosome complex with z-score >2 are indicated in red.

Next, I went on to identify genes that could affect hESCs' survival when depleted by siRNAs. This was analysed by quantifying the nuclei number (Table 2). A cutoff with an average z-score of +2 was used to determine the genes that reduce cell number (Nav) from the negative control (NT siRNA). GO analysis of the top 200 genes ranked by $N_{av}$ score revealed that gene categories such as nucleic acid binding protein, ribosomal protein and DNA-directed RNA polymerase were significantly enriched (Figure 27) and several members of ribosomal protein and DNA-directed RNA polymerase were among the top 25% of the hits. This is expected given the essential roles of ribosomal proteins and RNA polymerase II complex in eukaryotes. Hence, this proves that our RNAi assay was effective in identifying genes that affect hESC viability as well as genes that regulate the *OCT4* promoter.



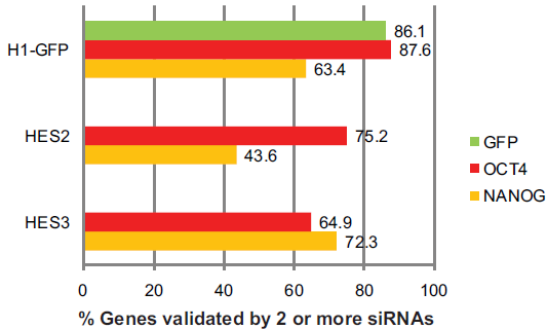

by Yuriy L. Orlov

**Figure 27: Gene ontology of the top 200 genes that affects cell survival**
Gene ontology analysis on the molecular function of the top 200 genes ranked by Nav scores (Panther classification). Candidate genes involved in cell survivability were enriched in GO categories for nucleic acid binding and ribosomal proteins, mRNA splicing and processing factor and DNA-directed RNA polymerase.

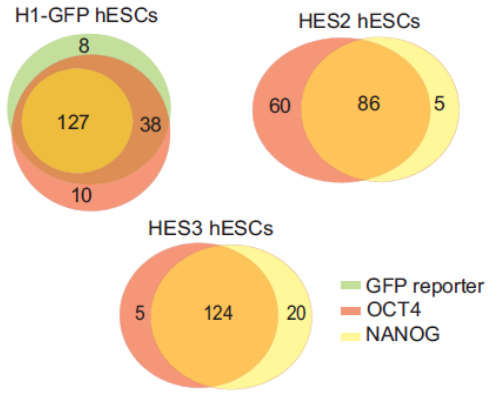
### 3.5.3 Secondary screen validation

Next, I performed secondary validation screen for 200 of the 566 candidates from the primary screen. The pooled siRNAs for each gene were deconvoluted into 4 individual siRNAs. Candidates were considered positive if they were scored by at least 2 siRNAs. To further enhance the confidence of the hit genes, a multiparametric approach was adopted where the importance of each gene in the maintenance of hESCs was assessed by different stemness markers (OCT4 and NANOG) for different hESCs cell lines. For H1-GFP hESCs, the validation rate based on the reduction of GFP reporter, OCT4 and NANOG expression were 86.1%, 87.6% and 63.4% respectively and 127 genes were validated by all three markers. For HES2 hESCs, 86 common genes were obtained based on OCT4 and NANOG expression and the validation rate was 75.2% and 43.6% respectively. Repeating with HES3 hESCs, 124 common genes with a validation rate of 64.9% and 72.3% for OCT4 and NANOG was observed respectively (Figure 28a, b, Table 4,5). The higher validation rate for H1-GFP hESCs as compared to other hESC lines corroborated the fact that this cell line was used for the primary screen. I identified 93 genes that downregulated OCT4 expression (Figure 28c, Table 6) and 54 genes that downregulated NANOG expression in the 3 different hESC lines (Figure 28c, Table 7). Of the 93 genes that regulate OCT4 expression, several interesting factors like PRDM14, NFRKB, JMJD2B, TAF7 etc appeared positive. Other categories of genes such as zinc finger/related proteins, cytoskeletal proteins and lipid/hormone related synthesis proteins were also enriched. In addition, I also observed a positive correlation between the stemness markers (Figure 28d).

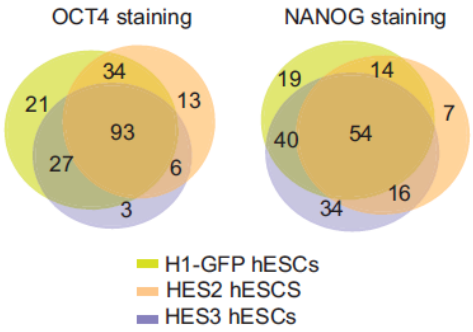
A)


H1-GFP
HES2
HES3
86.1
87.6
63.4
75.2
43.6
64.9
72.3
GFP
OCT4
NANOG
0
20
40
60
80
100
% Genes validated by 2 or more siRNAs

B)


H1-GFP hESCs
8
127
38
10
HES2 hESCs
60
86
5
HES3 hESCs
5
124
20
GFP reporter
OCT4
NANOG

C)


OCT4 staining
NANOG staining
34
21
13
93
27
6
3
19
14
7
40
54
16
34
H1-GFP hESCs
HES2 hESCS
HES3 hESCs

D)

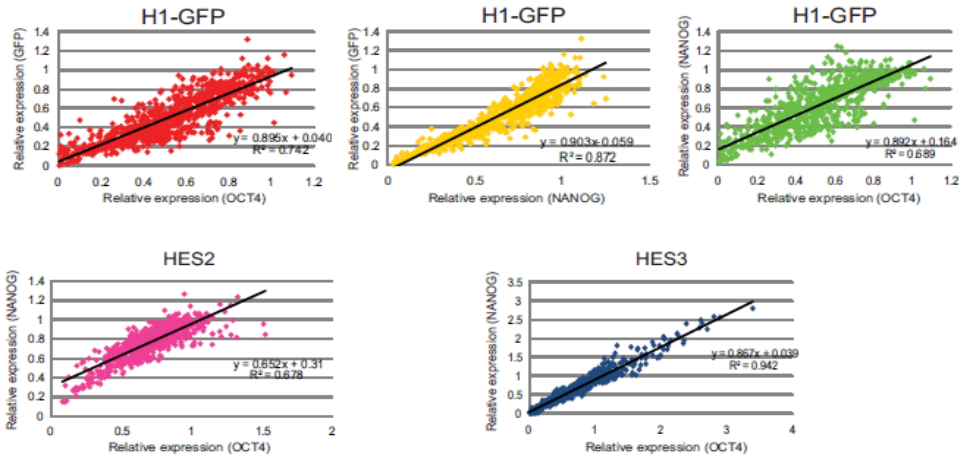


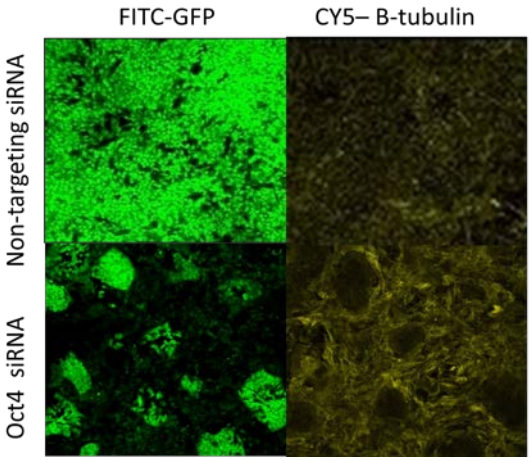

**Figure 28: Secondary screen**
**A)** Deconvoluted siRNA screen using H1-GFP, HES2 and HES3 hESCs lines. 200 genes from the 566 genes with z score>2 were subjected to further validation by deconvoluting the pooled mixture of 4 siRNAs. The screen was performed on 3 different hESCs lines and different stemness markers were used for analysis. H1-GFP hESC line was analyzed for GFP, OCT4 and NANOG expression, HES2 and HES3 hESCs were analyzed for OCT4 and NANOG expression. Genes were considered positive hits if 2 or more siRNA downregulated GFP/OCT4/NANOG expression. The percentage of genes that were validated/cell line/stemness marker are indicated beside the respective bars.
**B)** Venn diagram showing the overlapping hits for the different marker of analysis in each of the different cell lines. 127 genes are validated by GFP, OCT4 and NANOG downregulation in H1-GFP hESCs. 86 genes in HES2 and 124 genes in HES3 were validated based on OCT4 and NANOG downregulation.
**C)** Venn diagram showing the common overlapping genes among the 3 different hESC lines based on OCT4 or NANOG stemness marker for analysis. 93 common genes are involved in the downregulation of OCT4 and 54 common genes are involved in the down regulation of NANOG in all the 3 hESCs lines.
**D)** Graphs depicting the GFP VS OCT4, GFP VS NANOG correlation for H1-GFP hESCs and NANOG versus OCT4 correlation for each of H1-GFP, HES2 and HES3 hESCs.

After ascertaining that knockdown of these select genes led to loss of pluripotency factor expression in hESCs, I wanted to confirm that differentiation was indeed induced—and more specifically, I was interested in descrying what specific lineages hESCs differentiated to after the depletion of specific pluripotency genes, because pluripotency factors often selectively blockade differentiation to some but not all lineages. In addition to the identification of stemness genes, I sought for genes that induce differentiation of

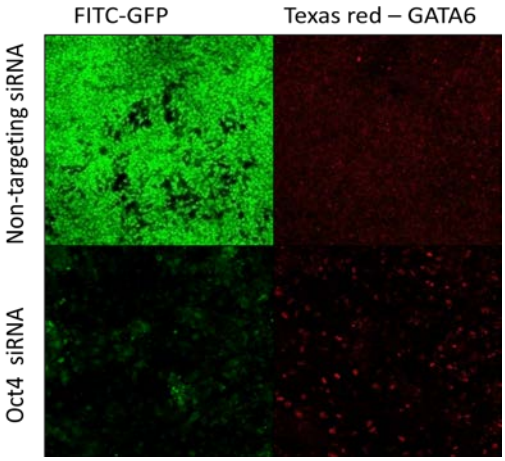
hESCs into the different lineages. β-tubulin, Brachrury and GATA6 are representative of the early differentiation markers of neuroectoderm, mesoderm and endoderm, respectively. Prior to the secondary screen, these markers were tested on hESCs that were induced to differentiate via OCT4 siRNA knockdown. The immunostaining of these lineage markers showed a marked difference between differentiated and undifferentiated cells for β-tubulin and GATA6 but slightly increased for brachrury (Figure 29a). It is possible that knockdown of OCT4 induce hESCs to differentiate less to the mesoderm lineage than the other lineages. Subsequently, the secondary screen proceeded with the staining of the cells with these different lineage antibodies. With a cut-off of greater than 2 fold increase in these lineage markers expressions, we discovered that depletion of *APC (*adenomatous polyposis coli) drives hESCs to the mesodermal as well as endodermal lineage. This gene encodes a tumour suppressor protein that acts as an antagonist of the Wnt signalling pathway. It is also involved in other processes including cell migration and adhesion, transcriptional activation, and apoptosis. Other genes associated with GATA6 induction are *OCT4, HES6, IGFBP6, TADA2B, RHOA* and *FLJ45684* while *TAF2* induced β-tublin upregulation (Figure 29b). These genes may be applicable in the differentiation procedure to the respective lineages for future studies.

A)

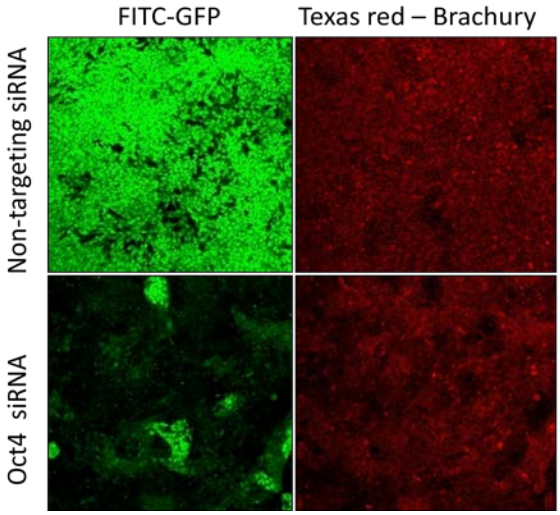
i)


FITC-GFP
CY5– B-tubulin
Non-targeting siRNA
Oct4 siRNA

ii)


FITC-GFP
Texas red – GATA6
Non-targeting siRNA
Oct4 siRNA

iii)


FITC-GFP
Texas red – Brachury
Non-targeting siRNA
Oct4 siRNA

B)

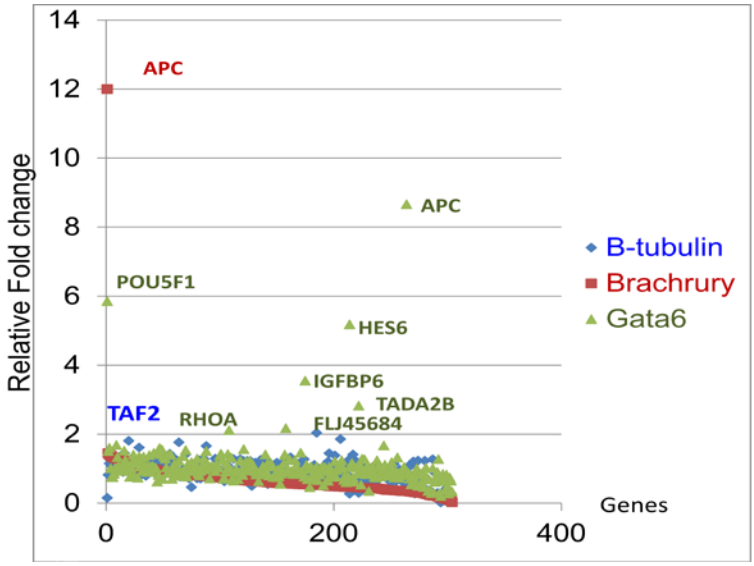



**Figure 29A) Immunostaining of the different lineage antibodies.** i) β-tubulin, ii) GATA6 and iii) Brachury that stains ectoderm, endoderm and mesoderm respectively were used. The cells were stained on cells with non-targeting siRNA knockdown and OCT4 siRNA knockdown.
**B) Lineage marker expression.** The relative fold change of the 3 lineages; ectoderm, mesoderm and endoderm is represented by β-tubulin, brachrury and GATA6 respectively.

### 3.5.4 Counter screen

In order to reduce the chances of identifying global regulators that could affect the general transcription or translation which are not specific for OCT4 regulation, we further conducted counter-screens against these 200 genes that were screened in the secondary screen. 3 control promoters, EF1α, β-ACTIN and GAPDH were selected for use in the screen (Figure 30).

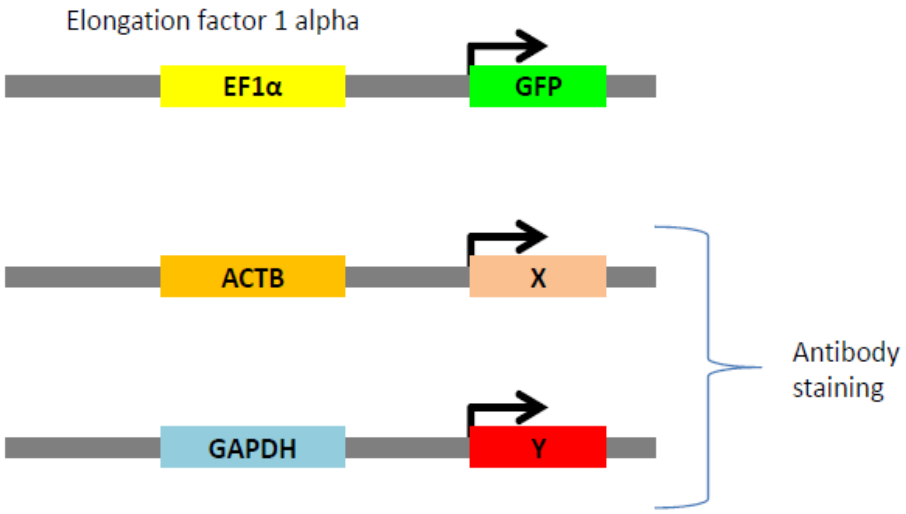



**Figure 30: Control promoters for counter screen.**
EF1a, ACTB and GAPDH are the different control promoters that are chosen for the counter screen.

We utilize a stable hESC line harboring an *EF1α* promoter (a gift from Lim Bing's lab) driving *GFP* reporter gene for the screen. EF1α promoters have been used for long term constitutive transgene expression in hESCs and we hypothesize that genes that could affect both the OCT4 and EF1α promoter are basic regulators for hESCs' survival and might not be specific for maintaining pluripotentiality. Additionally we assessed the expression level of housekeeping proteins like β-ACTIN and GAPDH via immuostaining (Table 7). Only a small fraction of the hits from the secondary assay could down-regulate EF1α promoter or the housekeeping genes (Figure 31). For instance *NXF1, PHB, PSMD2* and *SF3A1* which are generally involved in the basic transcription machinery were identified in the 3 different assays of the counter screen.

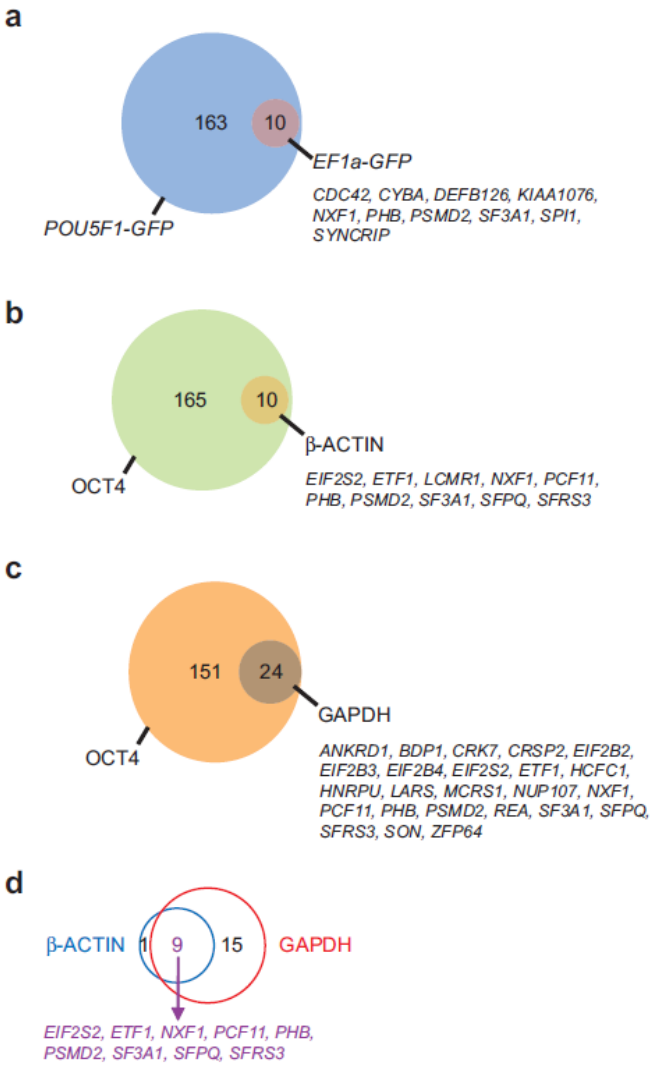



**Figure 31: Counter screen for 200 genes.**
**A)** Deconvoluted siRNA counter-screen using *EF1a-GFP* cell-line was carried out to assess the genes that non-specifically down-regulate an irrelevant house-keeping promoter after RNAi depletion. Genes were considered positive hits if 2 or more siRNA downregulated *EF1a-GFP* expression. In our secondary screen using the same set of siRNAs, 173 genes were confirmed by 2 or more siRNAs .
**B)** Venn diagram showing the overlapping hits between OCT4 and β-ACTIN. In our secondary screen using *OCT4-GFP* reporter, depletion of 175 genes using siRNAs led to down-regulation of OCT4 protein.
**C)** Venn diagram showing the overlapping hits between OCT4 and GAPDH. In our secondary screen using *OCT4-GFP* reporter, depletion of 175 genes using siRNAs led to down-regulation of OCT4 protein.
**D)** Common genes identified in both β-ACTIN and GAPDH counter-screens. Four of the common genes (*NXF1, PHB, PSMD2, SF3A1*) were also identified in the *EF1a-GFP* counter-screen.

### 3.6 Validation of short listed targets

A completed HT-RNAi screen comprises of the primary, secondary and counter screens and the final outcome resulted in a list of confirmed and high confidence genes that are modulators of the pluripotency of hESCs. It is of interest to study the biological roles of these targets in order to understand how they regulate hESCs. Among these high-confidence 93 candidates, we selected 4 genes that may be expected to play functional roles such as transcriptional factors or chromatin modifiers for further validation using loss of hESC morphology and reduction in AP staining as additional criteria. PRDM14 and NFRKB are transcription factors while JMJD2B and WDR82 (PRO2730) are chromatin modifiers. For each of these genes, we designed four shRNAs; independent of the siRNA sequence used in the primary screen and targeting different regions of the transcript. These shRNA constructs were transfected into the parental H1 line of the *OCT4-GFP* reporter hESCs. Evidence for the loss of hESC identity from the screens by independent siRNAs and shRNAs will corroborate the importance of the gene in the maintenance of hESCs, with minimal possibility of off-target effects. shRNA-mediated knockdown of all four genes resulted in apparent changes in morphology from the undifferentiated state and a reduction in AP staining. Like mESCs, hESCs express high levels of AP. Thus, in addition to stem cell markers such as OCT4 and NANOG expression, AP activity by immunocytochemistry staining is an alternative widely used stem cell marker for monitoring ESC differentiation status[1]. Besides these 4 genes, I included a negative control; *SERTAD2* which was scored negative in the secondary screen and expectedly depletion of *SERTAD2* did not induce differentiation (Figure 32, marked

by **). Therefore, we further affirmed the importance of these genes to be essential for the maintenance of hESCs.

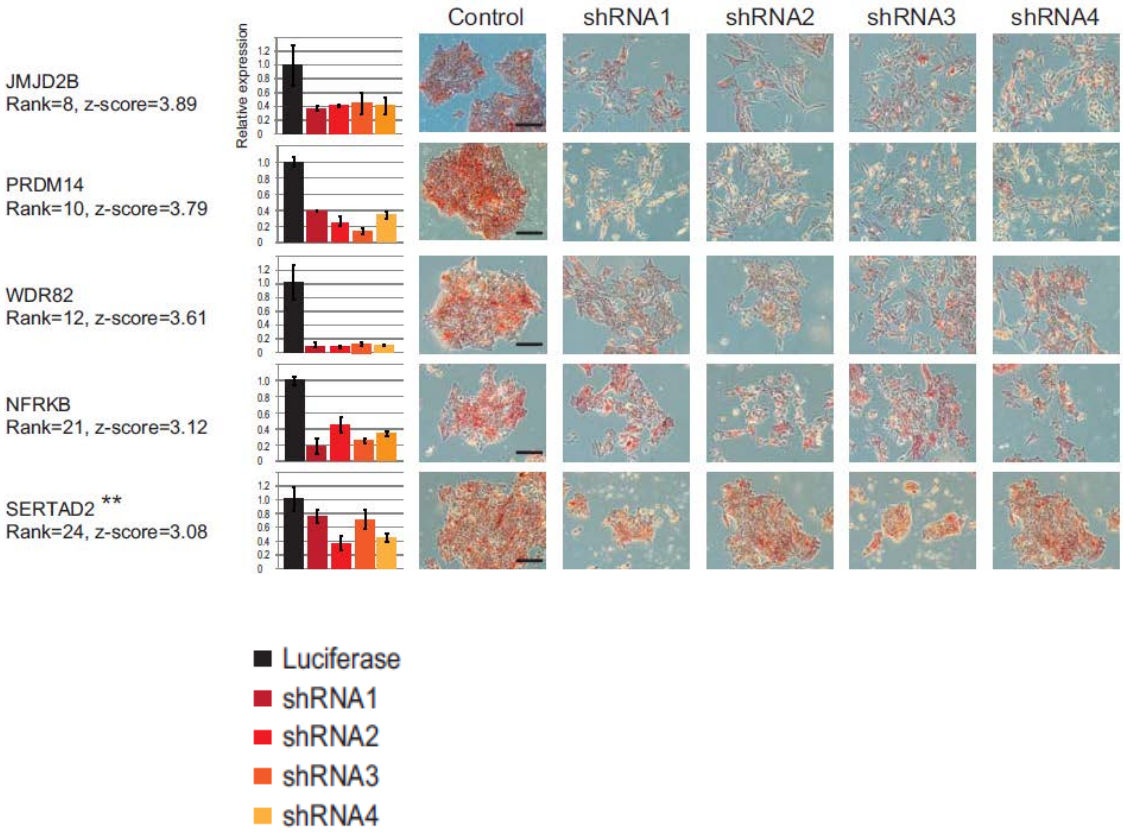
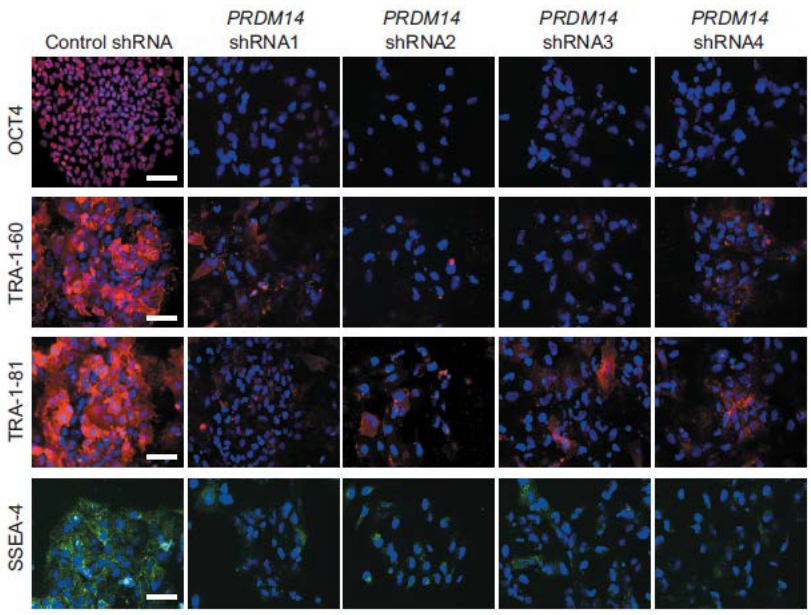



**Figure 32: shRNA validation of the 5 selected genes.**
The graphs show the relative mRNA expression of each of the individual genes with the four shRNA constructs. The cells are stained with AP staining. Bars represent 50 µM. ** represents the negative control.

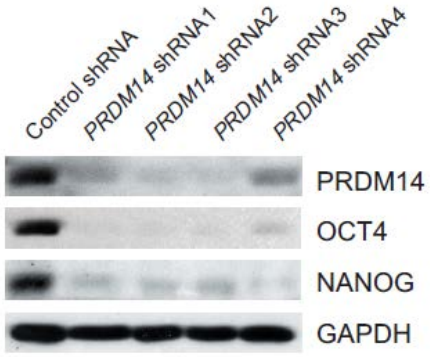
### 3.6.1 PRDM14 validation

The role of *PRDM14* in regulating the pluripotency of hESCs was studied in greater detail as it was ranked highly in the primary screen and its importance in the maintenance of hESCs was recapitulated in the secondary screen as well as validation using the four independent shRNAs. Furthermore, PRDM14 is highly expressed in a variety of independently derived hESC lines[50]. It is interesting to study genes that are specifically expressed in undifferentiated hESCs as they may potentially play important roles in hESCs. In addition PRDM14 is a target of the core transcription factors in hESCs[15]. *PRDM14* is bound by the key pluripotent transcription factor OCT4, indicating that *PRDM14* might be an essential downstream target of OCT4 and this also corroborate PRDM14's importance in the maintenance of hESCs when depleted. PRDM14 belongs to the PR (PRDI-BF1 and RIZ) domain proteins (PRDM) family which is a subclass of the SET domain proteins, a common domain found in histone modifying enzymes. PRDM14 has also been previously implicated to regulate self-renewal of hESCs since its depletion induced expression of differentiation marker genes and altered the cellular morphology[83]. We repeated the knockdown of PRDM14 using four shRNA constructs in non-reporter hESCs for more assays of pluripotency. Depletion of PRDM14 resulted in a reduction in the expression of hESC-associated genes such as OCT4, TRA-1-60, TRA-1-81 and SSEA-4 in three different hESC lines (H1, H9 and HES3) (Figure 33 a–c). Furthermore, we observed an upregulation in the differentiation markers expression of RUNX1, MAFB and IGFBP5 in the mRNA expression as well as protein expression when PRDM14 was depleted, strongly indicating differentiation of hESCs (Figure 33d).

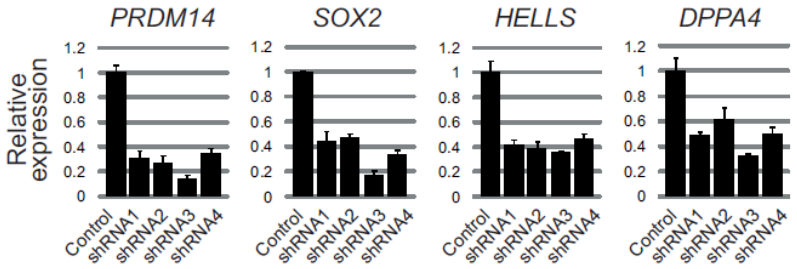
A)


Control shRNA
PRDM14 shRNA1
PRDM14 shRNA2
PRDM14 shRNA3
PRDM14 shRNA4
OCT4
TRA-1-60
TRA-1-81
SSEA-4

B)


Control shRNA
PRDM14 shRNA1
PRDM14 shRNA2
PRDM14 shRNA3
PRDM14 shRNA4
PRDM14
OCT4
NANOG
GAPDH

C)


PRDM14
SOX2
HELLS
DPPA4
Relative expression
1.2
1
0.8
0.6
0.4
0.2
0
Control
shRNA1
shRNA2
shRNA3
shRNA4

D)

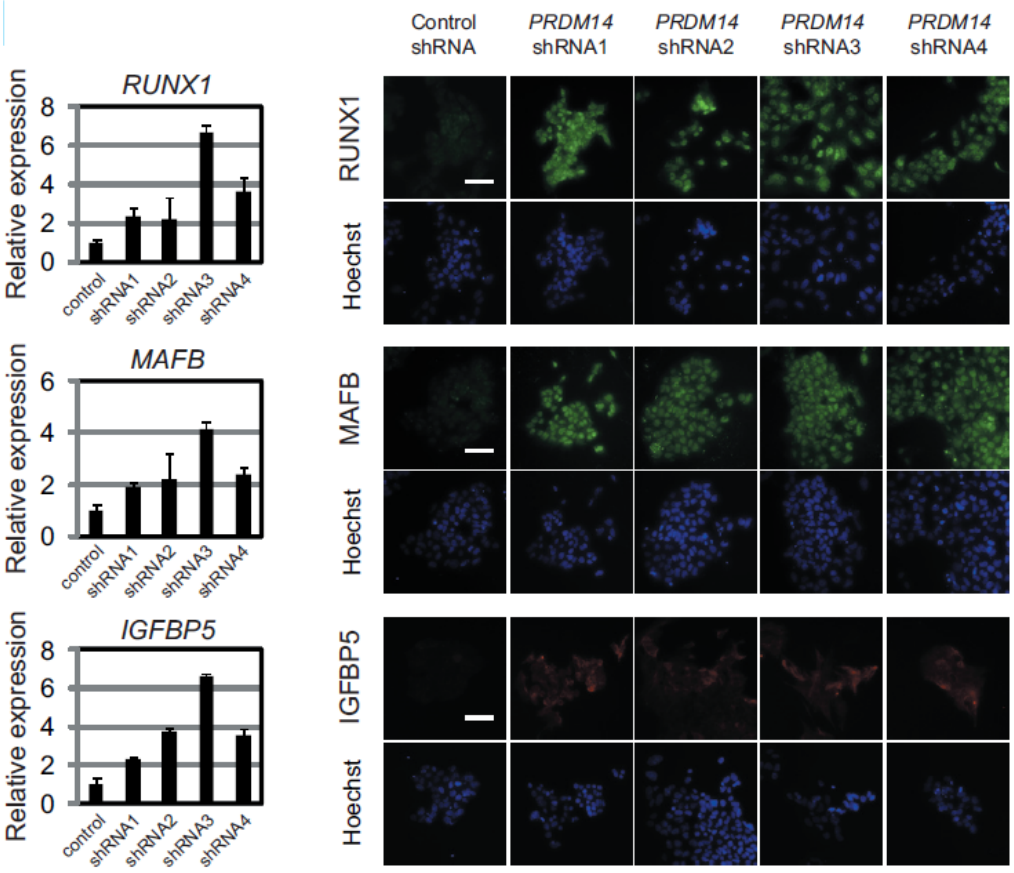



**Figure 33: PRDM14 is essential for the maintenance of hESCs.**
**A)** Immunofluorescence staining for stemness markers. *PRDM14* and control knockdown H1 hESCs were stained for OCT4, TRA-1-60, TRA-1-81 and SSEA-4 after 4 days of transfection. The nuclei of the cells were counterstained with Hoechst. Scale bars represent 50 µm.
**B)** Analysis of PRDM14, OCT4 and NANOG protein levels by Western blot.
**C)** Quantification of pluripotency-related transcript levels. Pluripotency-associated genes *SOX2*, *HELLS* and *DPPA4* were quantified for mRNA expression changes by qPCR. All values are means ± s.e.m from 3 independent experiments (n=3) and fold changes were normalized against control *luciferase* RNAi samples.
**D)** Quantification of differentiation-related transcript and protein levels. *RUNX1*, *MAFB* and *IGFBP5* were quantified for mRNA expression changes by qPCR. All values are means ± s.e.m from 3 independent experiments (n=3) and fold changes were normalized to control RNAi samples. Immunofluorescence assays were used to detect protein expression upon PRDM14 depletion. Scale bars represent 100 µm.

### 3.6.2 Rescue of PRDM14 knockdown phenotype

In order to ensure that the knockdown effect is specific, PRDM14 rescue experiment was performed where both RNAi-immune PRDM14 cDNA and PRDM14 shRNA were co-expressed. In the parallel control experiment, hESCs were transfected with the PRDM14 shRNA construct only. RNAi immune construct was generated through the introduction of synonymous substitution (also called a *silent* substitution) into the coding region of PRDM14 that was targeted by its respective shRNA sequence. The third base codon of the PRDM14 shRNA targeted region was mutated such that the amino acid sequence was not modified (Box 2). In the rescue experiment, endogenous PRDM14 will be downregulated but PRDM14 continues to be expressed from the RNAi-immune PRDM14 cDNA as the PRDM14 shRNA does not recognize the PRDM14 mutated sequence. If the knockdown of PRDM14 is specific, the exogenous expression of the PRDM14 will rescue the knockdown effect of PRDM14 (Figure 34).

PRDM14 WT: ggagactgctatgagaaat
PRDM14 silent mutation: ggGgaTtgTtaCgaAaaGt

**Box 2: Example of silent mutation of PRDM14**

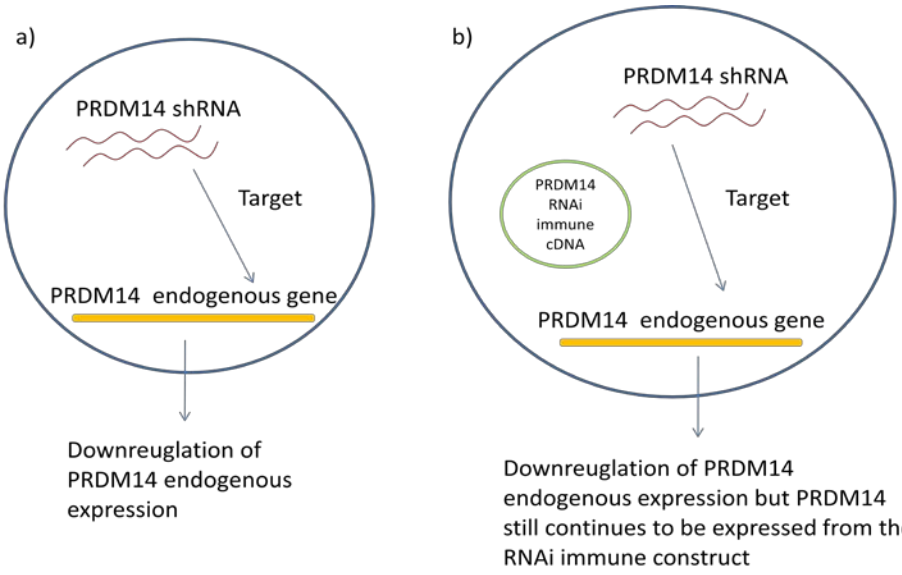
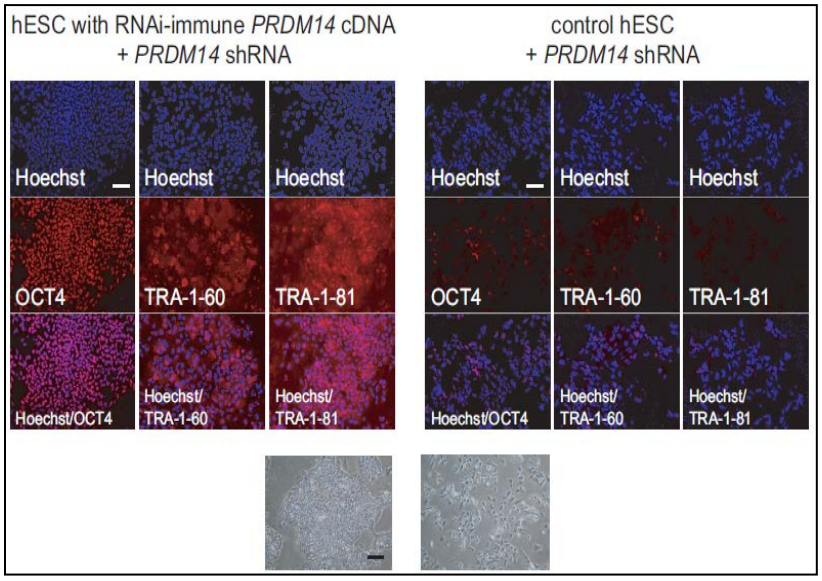



**Figure 34: Schematic diagram of PRDM14 rescue experiment.**
**A)** Downregulation of endogenous PRDM14 by shRNA. The endogenous PRDM14 mRNA can be downregulated by the PRDM14 shRNA.
**B)** Rescue of PRDM14 knockdown. In the presence of the PRDM14 RNAi immune cDNA, the PRDM14 mRNA will not not recognized by PRDM14 shRNA and hence, the continued expression of PRDM14 protein will be prevent the phenotype that is observed when PRDM14 is downregulated.

In the presence of RNAi-immune PRDM14 cDNA, PRDM14 knockdown effect was not observed from the immunostaining (Figure 35A). Additionally, there was no downregulation in the pluripotency marker gene expression although its endogenous PRDM14 level was reduced to the same level the control experiment as evident from the 3'UTR mRNA expression of PRDM14 (Figure 35B). Since the phenotypes induced by PRDM14 shRNA could be rescued by co-expression of RNAi-immune PRDM14 cDNA, this indicates that the knockdown effect was specific and was not due to the off-target effects of RNAi. Since the depletion of PRDM14 affects OCT4 level, OCT4 might be the downstream target gene of *PRDM14*, I asked whether the overexpression of OCT4 could rescue the knockdown effect of PRDM14. OCT4 cDNA and PRDM14 shRNA were co-

transfected into hESCs. PRDM14 knockdown resulted in the downregulation of the pluripotent markers and a upregulation of differentiation markers such as MAFB, BMP4 and ACTA2. However, the overexpression of OCT4 did not bring the pluripotent and differentiation markers to the same level as the negative control. Hence, OCT4 over-expression is unable to rescue the PRDM14 knockdown phenotype (Figure 36A,B).

A)




B)

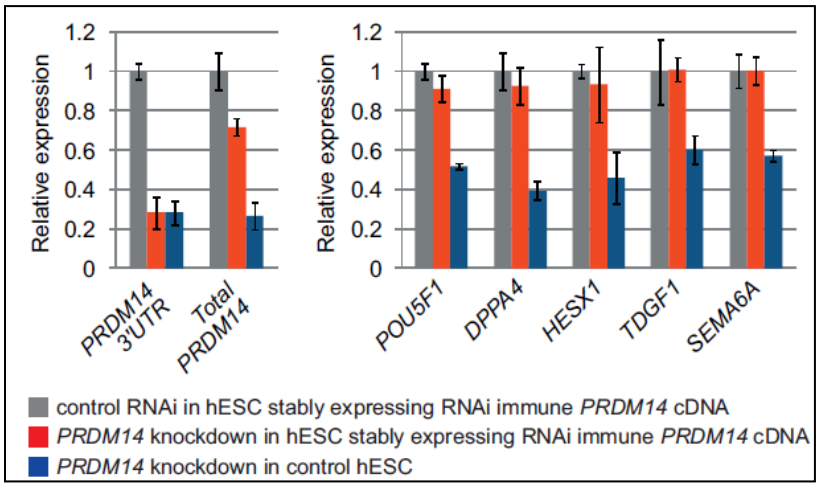
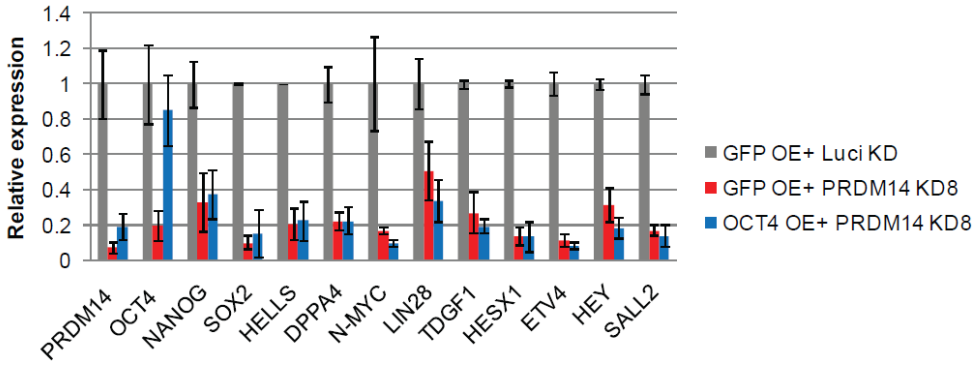


**Figure 35: Co-expression of RNAi-immune PRDM14 cDNA can rescue the phenoptye induced by PRDM14 shRNA.**
**A)** Knockdown of *PRDM14* in hESCs resulted in differentiation with a reduction in OCT4, TRA-1-60 and TRA-1-81while hESCs stably expressing *PRDM14* RNAi immune cDNA in the presence of *PRDM14* shRNA remained undifferentiated and maintain expression of OCT4, TRA-1-60 and TRA-1-81. The brighfield images show morphological change induced by *PRDM14* depletion for control hESCs but not for hESCs with RNAi-immune *PRDM14*. The scale bars represent 100 µm for the fluorescence images, and 50 µM for the bright field images.
**B)** *PRDM14* shRNA induced knockdown of endogenous *PRDM14* in both control hESCs and hESCs with stably expressing *PRDM14* RNAi immune cDNA (left graph). Endogenous *PRDM14* was detected by primer pairs targeting the 3'UTR of *PRDM14*. Primer pairs against *PRDM14* coding region detect total *PRDM14* (wild-type and RNAi-immune *PRDM14*). Expression of pluripotentcy markers are sustained in the rescue cells as compared to control cells (right graph).

A)




B)

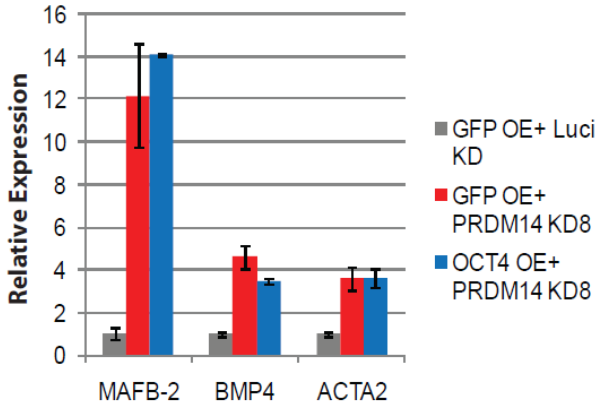
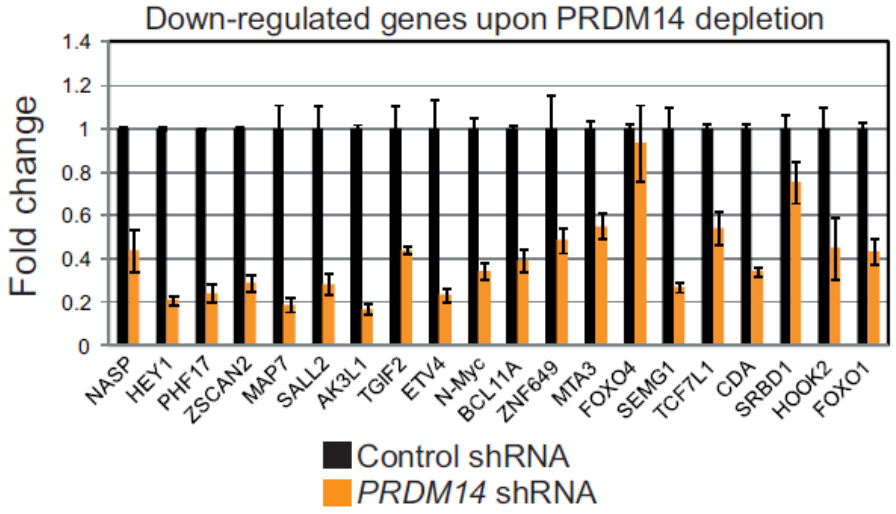



**Figure 36: Overexpression of OCT4 does not rescue PRDM14 knockdown phenotype.**
**A)** Effect of OCT4 over-expression on pluripotency markers.
**B)** Differentiation markers upon the knockdown of PRDM14.

Next, we performed microarray analysis to profile the expression changes on the *PRDM14*–RNAi-depleted hESCs. From the microarray data, we observed downregulation of pluripotency and upregulation of differentiation markers. I selected 20 genes that are downregulated and 20 genes that are upregulated and the same expression pattern can be validated using qPCR (Figure 37). The downregulation of *PRDM14* results in the downregulation of *OCT4* expression. This positive regulation of *OCT4* expression by PRDM14 is unexpected as previous studies implicate PRDM14 as a transcriptional repressor[83; 84]. Based on our genome-wide PRDM14 binding site profiling data and expression analysis, we unveiled that PRDM14's target genes are involved in a diversity of cellular processes. Genes coding for transcription factors (*OCT4*, *N-MYC*, *ETV4*, *TCF7L1*), chromatin modifiers (*TET2*), growth factors (*TDGF1*, *GDF3*), microRNA biogenesis factor (*LIN28*) and cell cycle regulator (*CDC25A*) are positively regulated by PRDM14 (Figure 38). On the other hand, genes coding for tissue-specific transcription factors and certain growth factor (BMP4) are negatively regulated by PRDM14.This finding suggests that PRDM14 can play both positive and negative roles on transcription.

A)




B)

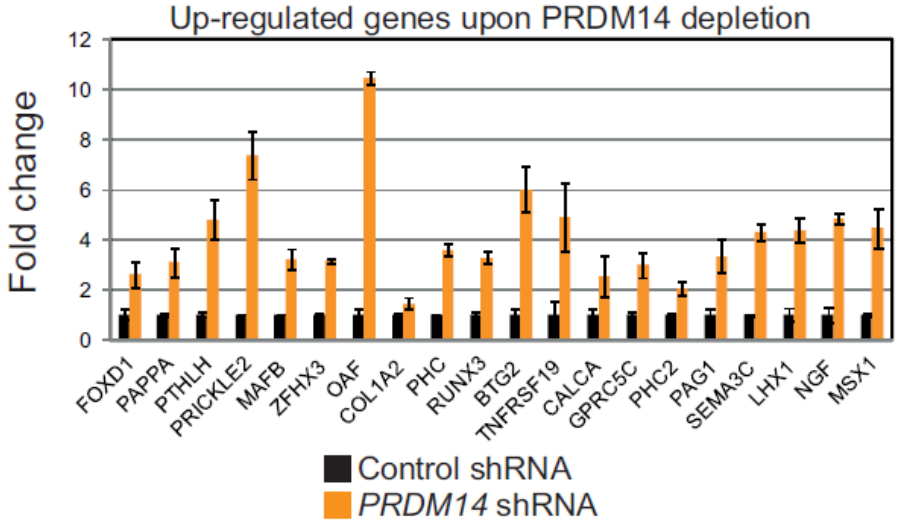



**Figure 37: Validation of gene expression upon *PRDM14* depletion**
**A)** 20 genes from the different Gene Ontology groups were selected for qPCR validation and 19 genes were downregulated upon *PRDM14* depletion. All values are means ± s.e.m from 3 independent experiments (n=3).
**B)** 20 genes from the different Gene Ontology groups were selected for qPCR validation and all the 20 genes were upregulated upon *PRDM14* depletion. All values are means ± s.e.m from 3 independent experiments (n=3).

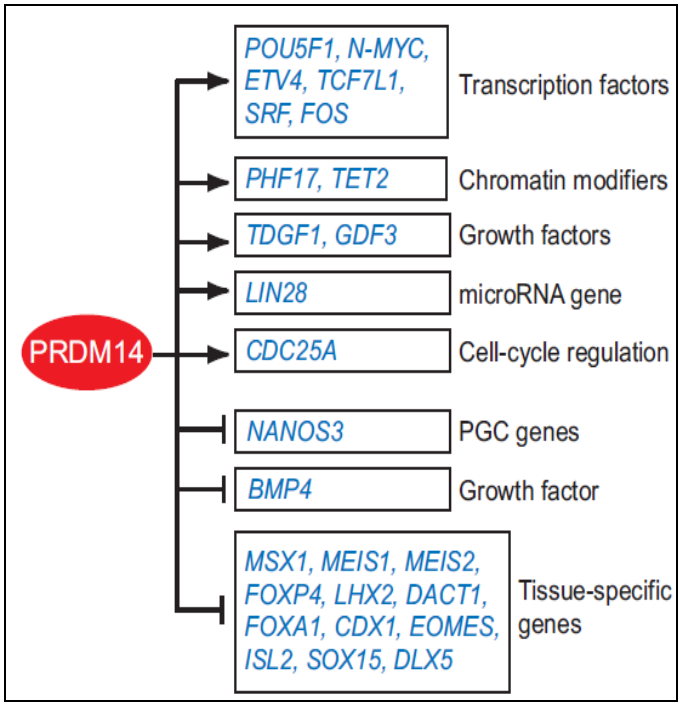


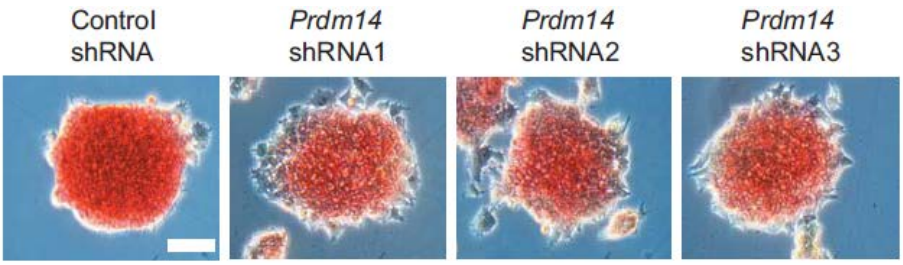

**Figure 38. Regulation of target genes by PRDM14**
Schematic representation of a model for the transcriptional regulatory network governed by PRDM14 in hESC

### 3.6.3 Human PRDM14 vs mouse Prdm14

Since hESCs and mESCs share similar ESCs characteristics where they are both pluripotent and can self-renew indefinitely, I sought to understand the importance of Prdm14 in mESCs. Previous work on a Prdm14 knockout mouse model showed that Prdm14 is critical for the establishment of the germ cell lineage[84]. In addition, Prdm14 is essential for the derivation of embryonic germ cells from primordial germ cells (PGCs). However, the knockout animals do not show early embryonic lethal phenotype, unlike that of other key regulators essential for the maintenance of pluripotency in mESCs[24; 26; 57; 85]. Hence, Prdm14 is not required to maintain mESCs and pluripotent stem cells in the blastocysts[68; 69; 84]. In PGCs, the expression of *OCT4* is maintained in the absence of

Prdm14. However, our study indicates that PRDM14 is critical for activating *OCT4* in hESCs. Therefore, PRDM14 maintains pluripotency and promotes the acquisition of pluripotency of the germ cell lineage and hESCs through distinctive mechanisms. These differences may arise through cell-type specific or species-specific differences in *OCT4* regulation. Given that Prdm14 knockout mice is not embryonic lethal, its role in hESCs is unexpected and it could potentially be hESC-specific. Nevertheless, we cannot exclude the possibility that functional redundancy by other molecules or pathways has masked the role of *Prdm14 in vivo* in the mouse system. To ascertain the functional redundancy of *Prdm14 in vitro*, we proceeded with the knock down of *Prdm14* in mESCs. *Prdm14* was depleted using three independent shRNAs that target different regions of *Prdm14*. There was no phenotypic change in the morphology of mESCs and in addition there was no reduction in *OCT4* and *Sox2* mRNA expression levels (Figure 39A, B). Hence, this result showed that Prdm14 is not essential *in vivo* and *in vitro* for mouse.

A)




B)

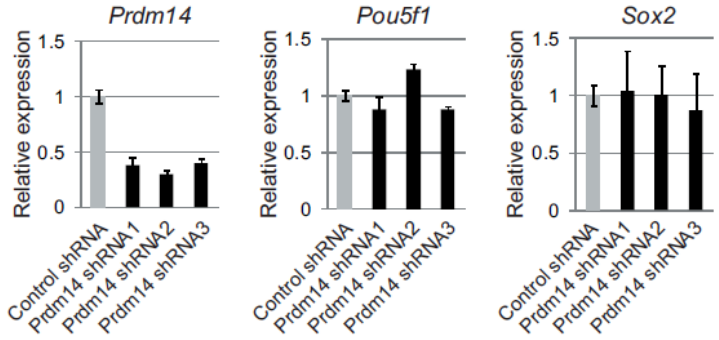
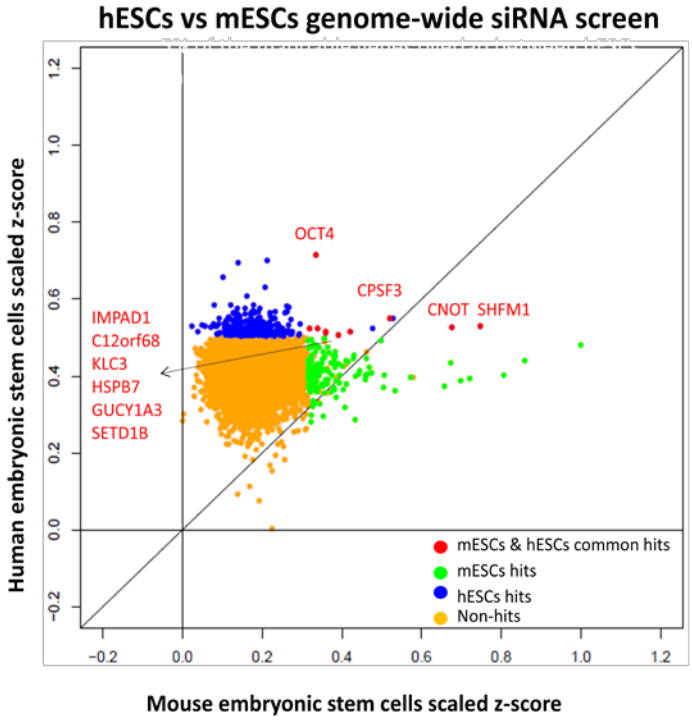



**Figure 39: Prdm14 is not required for the maintenance of mouse ESCs**
**A)** Knockdown of *Prdm14* in mESCs with 3 different shRNA constructs did not induce differentiation as indicated by the AP staining and morphology. Scale bar represents 50 µm.
**B)** Knockdown of *Prdm14* in mESCs does not reduce *OCT4*, *Nanog* and *Sox2* expression. All qPCR values are means ± s.e.m from 3 independent experiments (n=3) and fold changes were normalized to control RNAi samples.

Besides PRDM14, it is likely that hESCs have its specific regulators that are not shared by mESCs. To test this hypothesis, Pankaj Kumar compared the "hits" from our study[86] with the genome-wide siRNA screen study in mESCs by Hu et al[68] to examine the extent of the number of shared regulators such as OCT4 between the two species. Surprisingly, there is only a 5% overlap in the candidate genes between hESCs and mESCs (Figure 40). There are 9 genes, namely, OCT4, CNOT, CPSF3(cleavage and polyadenylation specific factor 3), IMPAD1(inositol monophosphatase domain containing 1), C12orf68 (chromosome 12 open reading frame 68), KLC3 (kinesin light chain 3), HSPB7 (heat shock protein family,

member 7), GUCY1A3 (guanylate cyclase 1), SETD1B (SET domain containing 1B) that are shared between hESCs and mESCs. SETD1B is a component of a histone methyltransferase complex that produces trimethylated histone H3 at Lys4. SETD1B may be an interesting gene which may reveal how it regulates both hESCs and mESCs in a similar way and this may be worth investigating. The low overlapping percentage is also due to our stringent cutoff  in the screen. As a result, well known genes such as NANOG and SOX2 were not within the top candidates.

From this genome-wide siRNA screen comparison between hESCs and mESCs, the small overlap in the regulators between them revealed that human and mouse have distinctive regulators and this could be a possible explanation for the differences between them.




Analysed by Pankaj Kumar

**Figure 40: Overlap between mESCs and hESCs siRNA screen results**
There is a 5% overlap of genes that are essential for the maintenance of the pluripotency of both hESCs and mESCs. The 9 common genes are indicated in red.

# 4 DISCUSSION

## 4.1 Comparison of "hits" between hESCs and mESCs

Both small-scale and genome-wide RNAi screens in murine ESCs [67; 68; 69; 87] have uncovered many important regulators. Despite these efforts, little is known about the key players in hESCs. Given the substantial differences between the two, it is of interest to investigate the key genetic components that are specific to hESCs. I conducted the first ever large scale whole genome wide screening in hESCs identified many regulators that are essential for the maintenance of hESCs. In comparison to the work carried out by Hu *et al* in mESCs using the Dharmacon siRNA library[68], there is only a 5% overlap in the hit genes with z-score of above 2 (Figure 40). This indicates that hESCs are governed by a set of regulators that are different from mESCs thereby accounting for the differences between them. The identification of complexes like the Mediator complex, INO80 complex, TAF complexes and COP9 signalosome is interesting and has yet to be implicated for its importance in hESCs. In mESCs, it has been reported that mediator in conjunction with Cohesin are involved in gene activation by physically and functionally connecting the enhancers and core promoters of active genes[82]. It would be interesting to study if mediator in hESCs function in the same way as mESCs. The 93 high confidence OCT4 regulators that were identified in the secondary screen might reveal insights into the possible partners of OCT4 hence illuminating how the transcription network of hESCs is established and regulated. In mESCs, INO80 complex has been shown to be associated with Oct4 via FLAG-affinity-based protein purification procedure[88]. It is possible that some of the components of the INO80 complex may be interacting with OCT4 in hESCs. Studying these OCT4 regulators has implications in the identification of disease associated

genes in human as a majority of Oct4-associated proteins resulted in early lethal phenotype when mutated in mouse models[88]. As there has been no reports of association of TAF complex and COP9 signalosome with ESCs, their importance could be specific to hESCs and their specific role remains to be determined.

## 4.2 Application of my screening data in induced pluripotent cells (iPSCs) studies

A detailed molecular understanding on how genetic factors affect the balance between pluripotency and differentiation of mammalian cells is essential to develop ESCs for their potential therapeutic use. This understanding is becoming increasingly important as it is possible to generate ESC-like cells (also known as induced pluripotent cells; iPSCs) from somatic cells via direct reprogramming of somatic cells[89], avoiding the hurdles of ethical issues. In the near future, it is possible to bank personalized iPSCs that may be used to generate differentiated cells to replace damaged tissues in the body when needed. Pluripotency can be reinstated in somatic cells through the introduction of defined transcription factors such as OCT4(O), SOX2(S), KLF4(K) and c-Myc(C)[90; 91]. This process is also known as reprogramming which induces a differentiated cell to regress back to its initial stage of a pluripotent cell. It has become apparent since the initial finding of these four factors that this specific combination does not have to be strictly adhered to as some of these factors can be replaced. For instance, OCT4 can be replaced with Nr5a2 in the OSKC cocktail and it can enhance reprogramming efficiency[92], also, KLF4 of the OSKC cocktail can be replaced with Esrrb[93]. These factors usually play an important role in defining mESCs characteristics. With the identification of essential determinants of hESCs, it is likely that they may serve as potential reprogramming factors in the conversion of human somatic cells to iPSCs. This gain-of-function assay is successfully

exemplified with the introduction of transcription factors such as PRDM14 or NFRKB to the OSKC cocktail where we found that PRDM14 and NFRKB can enhance the efficiency of human iPSC genesis. This indicated the value in our identification of these essential determinants of hESCs and their role in hESCs could be further explored.

As the depletion of these factors via RNAi leads to the differentiation of hESCs, these factors could be responsible for the derivation of a particular lineage. Understanding the lineages that these factors are driving ESCs to will enhance our knowledge or improve on the differentiation protocol of hESCs, and this is useful for the application in tissue transplantation.

# 5 FUTURE STUDY

## 5.1 To understand the mechanism of the essential genes in hESCs

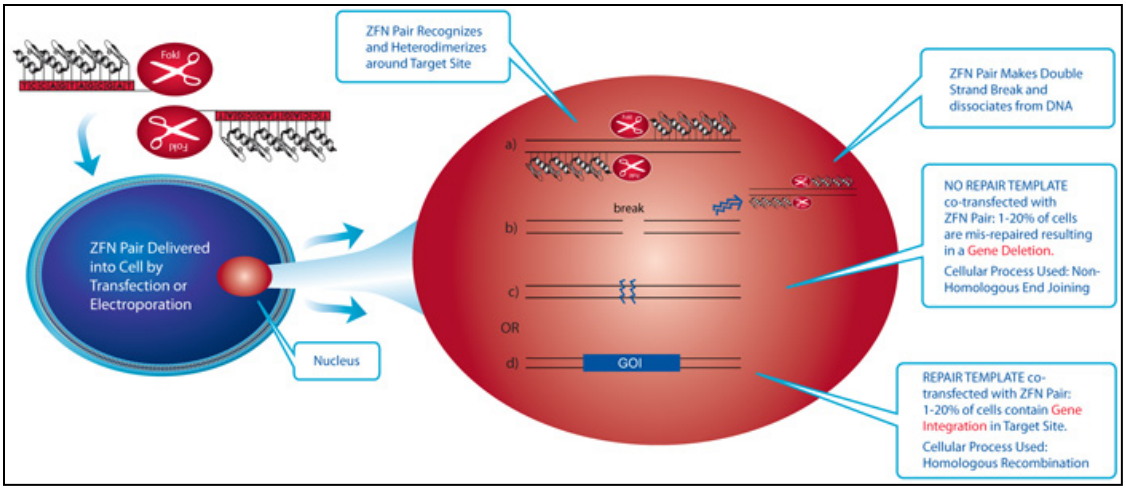
I have identified many regulators as well as complexes that are essential for the pluripotency and survival of hESCs and these genes could be further investigated to understand the mechanism in which these genes contributes to the maintenance of hESCs. Of particular interest are INO80 chromatin remodelling complex, mediator complex, TAF complex and COP9 signalosome. Further studies in characterizing the role of these complexes would reveal how these complexes work together to contribute to the stemness of hESCs as some of these complexes have proven to be important in mESCs[70; 88]. In addition, it is also worthwhile to explore genes that are involved in the identified signalling pathways. It will be interesting to decipher how the signalling pathways are wired into the transcriptional regulatory network of hESCs.

## 5.2 Site specific transgene integrated reporter lines using Zing finger nucleases

An improvement to my current screen is to generate a reporter line that has a reporter transgene integrated at a specific site instead into random sites. Random integration may disrupt the function of essential genes or it may be integrated into a constitutively active region, making the promoter immune to changes in OCT4 levels. This may result in a collection of clones with greatly varied expression levels and expression stability. This is made less tedious with the recent launched of CompoZr® Zinc Finger Nuclease Technology by the company Sigma which enables the editing of hESC genome using well-established and robust protocols (Figure 41). This allows for targeted gene deletions (knockouts), integrations, or modifications at a greater ease and efficiency than the

traditional tedious way via homologous recombination method. In addition the mutations are permanent and heritable and this is opportune for making stable cell lines. This technology employs the use of Zinc finger nucleases (ZFNs) which are a class of engineered DNA-binding proteins that facilitate the editing of the genome a designated target site via the creation of double-strand breaks in DNA. Double-strand breaks serves to stimulate the cell's inherent DNA-repair processes; homologous recombination and non-homologous end joining (NHEJ). *OCT4-GFP* transgene can then be directed to replace the natural adeno-associated virus (AAVS1) locus on human chromosome 19. This is the preferred site for integration because this region has an open chromatin conformation as indicated by the presence of a DNase I-hypersensitive site (DHS-S1)[94].

Thus, it is possible to hESCs reporter lines easily at any of the designated targeted site, circumventing the disadvantages and problems created by random integration of transgene. In addition, this screen can be evolved into a high-content (many parameters) screening by incorporating transgenes of reporter constructs driven by promoter of both pluripotent maker as well as differentiation markers. For example, a rainbow cell line can be generated by incorporating *OCT4-GFP*, ectodermal marker-RFP, mesodermal-YFP and endodermal marker-CFP transgenes together in the same cell line. This allows both its pluripotent and differentiation status at the same time.




Adapted from Sigma's website at http://www.sigmaaldrich.com/life-science/zinc-finger-nuclease-technology/learning-center/what-is-zfn.html

**Figure 41: CompoZr® Zinc Finger Nuclease Technology**
Zinc Finger Nuclease (ZFN) technology, allows easy creation of novel cell lines and model organisms with precise and heritable gene additions, deletions or modifications.

## 5.3 Other screening assays for hESCs

The realization of this genome-wide RNAi screen in hESCs opens the opportunities in the scientific field to carry out different RNAi screening assays to understand hESCs further in the near future. For instance, we could screen hESCs under differentiating conditions in search for genes that prevents the progression to a specific lineage; thereby identifying pluripotency genes. Additionally, we can investigate on the miRNA regulation in hESCs as miRNA library as it is not included in the genome-wide screen.

In this genome-wide siRNA screen, we have identified many false positives that have been eliminated from the subsequent screening assays. These false positives comprises of genes that are not expressed or expressed at a very low level in hESCs and are therefore not suitable for the RNAi screen. Hence, for any subsequent screening assays, it is necessary to filter away these genes from the very beginning to reduce the number of false positives obtained.

# 6 CONCLUSION

The adaptation of high-throughput siRNA screen to hESCs has been hampered by multiple technical challenges as hESCs have been known to be genetically less amenable than mESCs. Here, I have presented a strategy that has enabled the adaption of hESCs to HTS conditions. This is made possible under précised optimized conditions during the screening process. Use of this new assay has led to the identification of several novel regulators for the maintenance of hESCs. We revealed the involvement of transcription factors as well as several components of the different functionally distinct complexes in the maintenance of hESCs. It is particularly intriguing to find that chromatin remodelling complex (INO80 complex), transcriptional regulatory complexes (mediator complex and TAF complex) and signalling complex (COP9 signalosome) are implicated in hESC biology. Additionally, the function of PRDM14 and NFRKB are also crucial for hESC maintenance as well as enhancing the process of reprogramming. Future work will enable us to define more specifically their individual function and mechanisms through which these regulators operate collectively in hESCs. Our study shows that these regulators may hold the potential in advancing the understanding of hESCs biology and the mechanisms of human somatic cell reprogramming, thereby accelerating the progress in basic and translational hESC biology.

# 7 Biblography


1. Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S. & Jones, J. M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145-7.
2. Conaway, R. C. & Conaway, J. W. (2009). The INO80 chromatin remodeling complex in transcription, replication and repair. *Trends Biochem Sci* **34**, 71-7.
3. Casamassimi, A. & Napoli, C. (2007). Mediator complexes and eukaryotic transcription regulation: an overview. *Biochimie* **89**, 1439-46.
4. Chamovitz, D. A. (2009). Revisiting the COP9 signalosome as a transcriptional regulator. *EMBO Rep* **10**, 352-8.
5. Albright, S. R. & Tjian, R. (2000). TAFs revisited: more data reveal new twists and confirm old ideas. *Gene* **242**, 1-13.
6. Jackson, R. J., Hellen, C. U. & Pestova, T. V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* **11**, 113-27.
7. Rino, J. & Carmo-Fonseca, M. (2009). The spliceosome: a self-organized macromolecular machine in the nucleus? *Trends Cell Biol* **19**, 375-84.
8. Obeso, J. A., Rodriguez-Oroz, M. C., Goetz, C. G., Marin, C., Kordower, J. H., Rodriguez, M., Hirsch, E. C., Farrer, M., Schapira, A. H. & Halliday, G. (2010). Missing pieces in the Parkinson's disease puzzle. *Nat Med* **16**, 653-61.
9. Gilbert, S. F. (2000). *Developmental biology*, Sinauer Associates, Sunderland (MA).
10. Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A* **78**, 7634-8.
11. Evans, M. J. & Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-6.
12. Brons, I. G., Smithers, L. E., Trotter, M. W., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S. M., Howlett, S. K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R. A. & Vallier, L. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191-5.
13. Tesar, P. J., Chenoweth, J. G., Brook, F. A., Davies, T. J., Evans, E. P., Mack, D. L., Gardner, R. L. & McKay, R. D. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, 196-9.
14. Nichols, J. & Smith, A. (2009). Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487-92.
15. Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G., Gifford, D. K., Melton, D. A., Jaenisch, R. & Young, R. A. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947-56.
16. Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L. & Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. *Nature* **450**, 1230-4.
17. Yu, J. & Thomson, J. A. (2008). Pluripotent stem cell lines. *Genes Dev* **22**, 1987-97.
18. Wei, C. L., Miura, T., Robson, P., Lim, S. K., Xu, X. Q., Lee, M. Y., Gupta, S., Stanton, L., Luo, Y., Schmitt, J., Thies, S., Wang, W., Khrebtukova, I., Zhou, D., Liu, E. T., Ruan, Y. J., Rao, M. & Lim, B. (2005). Transcriptome profiling of human and murine ESCs identifies divergent paths required to maintain the stem cell state. *Stem Cells* **23**, 166-85.

19. Sun, Y., Li, H., Liu, Y., Shin, S., Mattson, M. P., Rao, M. S. & Zhan, M. (2007). Cross-species transcriptional profiles establish a functional portrait of embryonic stem cells. *Genomics* **89**, 22-35.
20. Kim, J., Chu, J., Shen, X., Wang, J. & Orkin, S. H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* **132**, 1049-61.
21. Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V. B., Wong, E., Orlov, Y. L., Zhang, W., Jiang, J., Loh, Y. H., Yeo, H. C., Yeo, Z. X., Narang, V., Govindarajan, K. R., Leong, B., Shahab, A., Ruan, Y., Bourque, G., Sung, W. K., Clarke, N. D., Wei, C. L. & Ng, H. H. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133**, 1106-17.
22. Scholer, H. R., Ruppert, S., Suzuki, N., Chowdhury, K. & Gruss, P. (1990). New type of POU domain in germ line-specific protein Oct-4. *Nature* **344**, 435-9.
23. Scholer, H. R., Dressler, G. R., Balling, R., Rohdewohld, H. & Gruss, P. (1990). Oct-4: a germline-specific transcription factor mapping to the mouse t-complex. *EMBO J* **9**, 2185-95.
24. Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M. & Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631-42.
25. Niwa, H., Miyazaki, J. & Smith, A. G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* **24**, 372-6.
26. Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S. & Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643-55.
27. Smith, A. G., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M. & Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* **336**, 688-90.
28. Niwa, H., Burdon, T., Chambers, I. & Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev* **12**, 2048-60.
29. Vallier, L., Touboul, T., Brown, S., Cho, C., Bilican, B., Alexander, M., Cedervall, J., Chandran, S., Ahrlund-Richter, L., Weber, A. & Pedersen, R. A. (2009). Signaling pathways controlling pluripotency and early cell fate decisions of human induced pluripotent stem cells. *Stem cells* **27**, 2655-66.
30. Bendall, S. C., Stewart, M. H., Menendez, P., George, D., Vijayaragavan, K., Werbowetski-Ogilvie, T., Ramos-Mejia, V., Rouleau, A., Yang, J., Bosse, M., Lajoie, G. & Bhatia, M. (2007). IGF and FGF cooperatively establish the regulatory stem cell niche of pluripotent human cells in vitro. *Nature* **448**, 1015-21.
31. Hombria, J. C. & Lovegrove, B. (2003). Beyond homeosis--HOX function in morphogenesis and organogenesis. *Differentiation; research in biological diversity* **71**, 461-76.
32. Hay, D. C., Sutherland, L., Clark, J. & Burdon, T. (2004). Oct-4 knockdown induces similar patterns of endoderm and trophoblast differentiation markers in human and mouse embryonic stem cells. *Stem cells* **22**, 225-35.
33. Matin, M. M., Walsh, J. R., Gokhale, P. J., Draper, J. S., Bahrami, A. R., Morton, I., Moore, H. D. & Andrews, P. W. (2004). Specific knockdown of Oct4 and beta2-microglobulin expression by RNA interference in human embryonic stem cells and embryonic carcinoma cells. *Stem cells* **22**, 659-68.
34. Botquin, V., Hess, H., Fuhrmann, G., Anastassiadis, C., Gross, M. K., Vriend, G. & Scholer, H. R. (1998). New POU dimer configuration mediates antagonistic control of an

osteopontin preimplantation enhancer by Oct-4 and Sox-2. *Genes & development* **12**, 2073-90.

35. Chew, J. L., Loh, Y. H., Zhang, W., Chen, X., Tam, W. L., Yeap, L. S., Li, P., Ang, Y. S., Lim, B., Robson, P. & Ng, H. H. (2005). Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Molecular and cellular biology* **25**, 6031-46.
36. Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N. & Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes & development* **17**, 126-40.
37. Vallier, L., Mendjan, S., Brown, S., Chng, Z., Teo, A., Smithers, L. E., Trotter, M. W., Cho, C. H., Martinez, A., Rugg-Gunn, P., Brons, G. & Pedersen, R. A. (2009). Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. *Development* **136**, 1339-49.
38. Loh, Y. H., Wu, Q., Chew, J. L., Vega, V. B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., Wong, K. Y., Sung, K. W., Lee, C. W., Zhao, X. D., Chiu, K. P., Lipovich, L., Kuznetsov, V. A., Robson, P., Stanton, L. W., Wei, C. L., Ruan, Y., Lim, B. & Ng, H. H. (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet* **38**, 431-40.
39. Marson, A., Levine, S. S., Cole, M. F., Frampton, G. M., Brambrink, T., Johnstone, S., Guenther, M. G., Johnston, W. K., Wernig, M., Newman, J., Calabrese, J. M., Dennis, L. M., Volkert, T. L., Gupta, S., Love, J., Hannett, N., Sharp, P. A., Bartel, D. P., Jaenisch, R. & Young, R. A. (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* **134**, 521-33.
40. Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281-97.
41. Houbaviy, H. B., Murray, M. F. & Sharp, P. A. (2003). Embryonic stem cell-specific MicroRNAs. *Dev Cell* **5**, 351-8.
42. Houbaviy, H. B., Dennis, L., Jaenisch, R. & Sharp, P. A. (2005). Characterization of a highly variable eutherian microRNA gene. *RNA* **11**, 1245-57.
43. Suh, M. R., Lee, Y., Kim, J. Y., Kim, S. K., Moon, S. H., Lee, J. Y., Cha, K. Y., Chung, H. M., Yoon, H. S., Moon, S. Y., Kim, V. N. & Kim, K. S. (2004). Human embryonic stem cells express a unique set of microRNAs. *Dev Biol* **270**, 488-98.
44. Kanellopoulou, C., Muljo, S. A., Kung, A. L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D. M. & Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev* **19**, 489-501.
45. Murchison, E. P., Partridge, J. F., Tam, O. H., Cheloufi, S. & Hannon, G. J. (2005). Characterization of Dicer-deficient murine embryonic stem cells. *Proc Natl Acad Sci U S A* **102**, 12135-40.
46. Wang, Y., Medvid, R., Melton, C., Jaenisch, R. & Blelloch, R. (2007). DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nat Genet* **39**, 380-5.
47. Hawkins, R. D., Hon, G. C., Lee, L. K., Ngo, Q., Lister, R., Pelizzola, M., Edsall, L. E., Kuan, S., Luu, Y., Klugman, S., Antosiewicz-Bourget, J., Ye, Z., Espinoza, C., Agarwahl, S., Shen, L., Ruotti, V., Wang, W., Stewart, R., Thomson, J. A., Ecker, J. R. & Ren, B. (2010). Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* **6**, 479-91.
48. Chia, N. Y. & Ng, H. H. (2011). Stem cell genome-to-systems biology. *Wiley interdisciplinary reviews. Systems biology and medicine*.

49. Glover, C. H., Marin, M., Eaves, C. J., Helgason, C. D., Piret, J. M. & Bryan, J. (2006). Meta-analysis of differentiating mouse embryonic stem cell gene expression kinetics reveals early change of a small gene set. *PLoS Comput Biol* **2**, e158.
50. Assou, S., Le Carrour, T., Tondeur, S., Strom, S., Gabelle, A., Marty, S., Nadal, L., Pantesco, V., Reme, T., Hugnot, J. P., Gasca, S., Hovatta, O., Hamamah, S., Klein, B. & De Vos, J. (2007). A meta-analysis of human embryonic stem cells transcriptome integrated into a web-based expression atlas. *Stem Cells* **25**, 961-73.
51. Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M. & Snyder, M. (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320**, 1344-9.
52. Wilhelm, B. T., Marguerat, S., Watt, S., Schubert, F., Wood, V., Goodhead, I., Penkett, C. J., Rogers, J. & Bahler, J. (2008). Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* **453**, 1239-43.
53. Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D. N., Theunissen, T. W. & Orkin, S. H. (2006). A protein interaction network for pluripotency of embryonic stem cells. *Nature* **444**, 364-8.
54. van den Berg, D. L., Snoek, T., Mullin, N. P., Yates, A., Bezstarosti, K., Demmers, J., Chambers, I. & Poot, R. A. (2010). An Oct4-centered protein interaction network in embryonic stem cells. *Cell Stem Cell* **6**, 369-81.
55. Pardo, M., Lang, B., Yu, L., Prosser, H., Bradley, A., Babu, M. M. & Choudhary, J. (2010). An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. *Cell Stem Cell* **6**, 382-95.
56. Kim, J., Woo, A. J., Chu, J., Snow, J. W., Fujiwara, Y., Kim, C. G., Cantor, A. B. & Orkin, S. H. (2010). A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs. *Cell* **143**, 313-24.
57. Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H. & Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379-91.
58. Rosner, M. H., Vigano, M. A., Ozato, K., Timmons, P. M., Poirier, F., Rigby, P. W. & Staudt, L. M. (1990). A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. *Nature* **345**, 686-92.
59. Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N. & Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* **17**, 126-40.
60. Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A. A., Ko, M. S. & Niwa, H. (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol* **9**, 625-35.
61. Kariko, K., Bhuyan, P., Capodici, J. & Weissman, D. (2004). Small interfering RNAs mediate sequence-independent gene suppression and induce immune activation by signaling through toll-like receptor 3. *Journal of immunology* **172**, 6545-9.
62. Hornung, V., Guenthner-Biller, M., Bourquin, C., Ablasser, A., Schlee, M., Uematsu, S., Noronha, A., Manoharan, M., Akira, S., de Fougerolles, A., Endres, S. & Hartmann, G. (2005). Sequence-specific potent induction of IFN-alpha by short interfering RNA in plasmacytoid dendritic cells through TLR7. *Nature medicine* **11**, 263-70.
63. Sledz, C. A., Holko, M., de Veer, M. J., Silverman, R. H. & Williams, B. R. (2003). Activation of the interferon system by short-interfering RNAs. *Nature cell biology* **5**, 834-9.

64. Pebernard, S. & Iggo, R. D. (2004). Determinants of interferon-stimulated gene induction by RNAi vectors. *Differentiation; research in biological diversity* **72**, 103-11.
65. Bauer, M., Kinkl, N., Meixner, A., Kremmer, E., Riemenschneider, M., Forstl, H., Gasser, T. & Ueffing, M. (2009). Prevention of interferon-stimulated gene expression using microRNA-designed hairpins. *Gene therapy* **16**, 142-7.
66. Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215-33.
67. Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCoste, C., Schafer, X., Lun, Y. & Lemischka, I. R. (2006). Dissecting self-renewal in stem cells with RNA interference. *Nature* **442**, 533-8.
68. Hu, G., Kim, J., Xu, Q., Leng, Y., Orkin, S. H. & Elledge, S. J. (2009). A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. *Genes Dev* **23**, 837-48.
69. Ding, L., Paszkowski-Rogacz, M., Nitzsche, A., Slabicki, M. M., Heninger, A. K., de Vries, I., Kittler, R., Junqueira, M., Shevchenko, A., Schulz, H., Hubner, N., Doss, M. X., Sachinidis, A., Hescheler, J., Iacone, R., Anastassiadis, K., Stewart, A. F., Pisabarro, M. T., Caldarelli, A., Poser, I., Theis, M. & Buchholz, F. (2009). A genome-scale RNAi screen for Oct4 modulators defines a role of the Paf1 complex for embryonic stem cell identity. *Cell Stem Cell* **4**, 403-15.
70. Kagey, M. H., Newman, J. J., Bilodeau, S., Zhan, Y., Orlando, D. A., van Berkum, N. L., Ebmeier, C. C., Goossens, J., Rahl, P. B., Levine, S. S., Taatjes, D. J., Dekker, J. & Young, R. A. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430-5.
71. Desbordes, S. C., Placantonakis, D. G., Ciro, A., Socci, N. D., Lee, G., Djaballah, H. & Studer, L. (2008). High-throughput screening assay for the identification of compounds regulating self-renewal and differentiation in human embryonic stem cells. *Cell Stem Cell* **2**, 602-12.
72. Reubinoff, B. E., Pera, M. F., Fong, C. Y., Trounson, A. & Bongso, A. (2000). Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. *Nature biotechnology* **18**, 399-404.
73. Xu, C., Inokuma, M. S., Denham, J., Golds, K., Kundu, P., Gold, J. D. & Carpenter, M. K. (2001). Feeder-free growth of undifferentiated human embryonic stem cells. *Nature biotechnology* **19**, 971-4.
74. Zafarana, G., Avery, S. R., Avery, K., Moore, H. D. & Andrews, P. W. (2009). Specific knockdown of OCT4 in human embryonic stem cells by inducible short hairpin RNA interference. *Stem cells* **27**, 776-82.
75. Ma, Y., Jin, J., Dong, C., Cheng, E. C., Lin, H., Huang, Y. & Qiu, C. (2010). High-efficiency siRNA-based gene knockdown in human embryonic stem cells. *RNA* **16**, 2564-9.
76. Gerrard, L., Zhao, D., Clark, A. J. & Cui, W. (2005). Stably transfected human embryonic stem cell clones express OCT4-specific green fluorescent protein and maintain self-renewal and pluripotency. *Stem cells* **23**, 124-33.
77. Yeom, Y. I., Fuhrmann, G., Ovitt, C. E., Brehm, A., Ohbo, K., Gross, M., Hubner, K. & Scholer, H. R. (1996). Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development* **122**, 881-94.
78. Chew, J. L., Loh, Y. H., Zhang, W., Chen, X., Tam, W. L., Yeap, L. S., Li, P., Ang, Y. S., Lim, B., Robson, P. & Ng, H. H. (2005). Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol* **25**, 6031-46.

79. Ziauddin, J. & Sabatini, D. M. (2001). Microarrays of cells expressing defined cDNAs. *Nature* **411**, 107-10.
80. Watanabe, K., Ueno, M., Kamiya, D., Nishiyama, A., Matsumura, M., Wataya, T., Takahashi, J. B., Nishikawa, S., Muguruma, K. & Sasai, Y. (2007). A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nat Biotechnol* **25**, 681-6.
81. Zhang, J. H., Chung, T. D. & Oldenburg, K. R. (1999). A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *Journal of biomolecular screening : the official journal of the Society for Biomolecular Screening* **4**, 67-73.
82. Kagey, M. H., Newman, J. J., Bilodeau, S., Zhan, Y., Orlando, D. A., van Berkum, N. L., Ebmeier, C. C., Goossens, J., Rahl, P. B., Levine, S. S., Taatjes, D. J., Dekker, J. & Young, R. A. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**, 430-5.
83. Tsuneyoshi, N., Sumi, T., Onda, H., Nojima, H., Nakatsuji, N. & Suemori, H. (2008). PRDM14 suppresses expression of differentiation marker genes in human embryonic stem cells. *Biochem Biophys Res Commun* **367**, 899-905.
84. Yamaji, M., Seki, Y., Kurimoto, K., Yabuta, Y., Yuasa, M., Shigeta, M., Yamanaka, K., Ohinata, Y. & Saitou, M. (2008). Critical function of Prdm14 for the establishment of the germ cell lineage in mice. *Nat Genet* **40**, 1016-22.
85. Silva, J., Nichols, J., Theunissen, T. W., Guo, G., van Oosten, A. L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I. & Smith, A. (2009). Nanog is the gateway to the pluripotent ground state. *Cell* **138**, 722-37.
86. Chia, N. Y., Chan, Y. S., Feng, B., Lu, X., Orlov, Y. L., Moreau, D., Kumar, P., Yang, L., Jiang, J., Lau, M. S., Huss, M., Soh, B. S., Kraus, P., Li, P., Lufkin, T., Lim, B., Clarke, N. D., Bard, F. & Ng, H. H. (2010). A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. *Nature* **468**, 316-20.
87. Fazzio, T. G., Huff, J. T. & Panning, B. (2008). An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. *Cell* **134**, 162-74.
88. Pardo, M., Lang, B., Yu, L., Prosser, H., Bradley, A., Babu, M. M. & Choudhary, J. An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. *Cell Stem Cell* **6**, 382-95.
89. Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., Nie, J., Jonsdottir, G. A., Ruotti, V., Stewart, R., Slukvin, II & Thomson, J. A. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917-20.
90. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K. & Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861-72.
91. Park, I. H., Lerou, P. H., Zhao, R., Huo, H. & Daley, G. Q. (2008). Generation of human-induced pluripotent stem cells. *Nat Protoc* **3**, 1180-6.
92. Heng, J. C., Feng, B., Han, J., Jiang, J., Kraus, P., Ng, J. H., Orlov, Y. L., Huss, M., Yang, L., Lufkin, T., Lim, B. & Ng, H. H. (2010). The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. *Cell Stem Cell* **6**, 167-74.
93. Feng, B., Jiang, J., Kraus, P., Ng, J. H., Heng, J. C., Chan, Y. S., Yaw, L. P., Zhang, W., Loh, Y. H., Han, J., Vega, V. B., Cacheux-Rataboul, V., Lim, B., Lufkin, T. & Ng, H. H. (2009). Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nature cell biology* **11**, 197-203.

94. Ogata, T., Kozuka, T. & Kanda, T. (2003). Identification of an insulator in AAVS1, a preferred region for integration of adeno-associated virus DNA. *Journal of virology* **77**, 9000-7.

# 8. TABLE

**Table 1: Gene list sorted by $F_{av}$ score.** F1: z-score of GFP fluorescence change for replicate 1, F2: z-score of GFP fluorescence change for replicate 2, Fav: average z-score of the GFP fluorescence change of the duplicates.

| No. | Entrez Gene Symbol | Accession no. | F1 | F2 | Fav |
|---|---|---|---|---|---|
| 1 | OCT4 | NM_002701 | -6.141 | -4.155 | -5.148 |
| 2 | SLC25A23 | NM_024103 | -6.426 | -3.441 | -4.934 |
| 3 | LRRC33 | NM_198565 | -4.710 | -4.974 | -4.842 |
| 4 | YAP1 | NM_006106 | -3.563 | -4.950 | -4.256 |
| 5 | PSMD2 | NM_002808 | -4.082 | -4.261 | -4.172 |
| 6 | LOC400134 | XM_378416 | -4.626 | -3.644 | -4.135 |
| 7 | HEMK1 | NM_016173 | -4.107 | -3.858 | -3.982 |
| 8 | JMJD2B | NM_015015 | -3.702 | -4.082 | -3.892 |
| 9 | SON | NM_003103 | -4.044 | -3.644 | -3.844 |
| 10 | PRDM14 | NM_024504 | -4.049 | -3.545 | -3.797 |
| 11 | GPS1 | NM_004127 | -2.377 | -4.865 | -3.621 |
| 12 | WDR82 | XM_293514 | -4.104 | -3.123 | -3.614 |
| 13 | CROP | NM_016424 | -1.996 | -5.050 | -3.523 |
| 14 | ZNF136 | NM_003437 | -1.669 | -4.950 | -3.309 |
| 15 | LOC344165 | XM_292957 | -4.126 | -2.441 | -3.283 |
| 16 | SLC7A5P1 | NM_031211 | -3.695 | -2.740 | -3.217 |
| 17 | GUSB | NM_000181 | -2.766 | -3.656 | -3.211 |
| 18 | COPS4 | NM_016129 | -4.233 | -2.162 | -3.198 |
| 19 | ZFP64 | NM_018197 | -3.416 | -2.933 | -3.174 |
| 20 | RUFY4 | NM_198483 | -2.927 | -3.353 | -3.140 |
| 21 | NFRKB | NM_006165 | -3.277 | -2.981 | -3.129 |
| 22 | RASEF | NM_152573 | -3.735 | -2.493 | -3.114 |
| 23 | CAPN2 | NM_001748 | -4.057 | -2.108 | -3.083 |
| 24 | SERTAD2 | XM_376059 | -2.882 | -3.277 | -3.080 |
| 25 | TADA2B | XM_291105 | -2.295 | -3.827 | -3.061 |
| 26 | ZIC4 | NM_032153 | -3.446 | -2.670 | -3.058 |
| 27 | EIF2S2 | NM_003908 | -2.573 | -3.425 | -2.999 |
| 28 | SFRS3 | NM_003017 | -3.592 | -2.392 | -2.992 |
| 29 | NPEPL1 | NM_024663 | -2.944 | -3.028 | -2.986 |
| 30 | ZFP36 | NM_003407 | -2.615 | -3.322 | -2.968 |
| 31 | SLC25A42 | XM_209204 | -2.372 | -3.518 | -2.945 |
| 32 | HCFC1 | NM_005334 | -2.969 | -2.881 | -2.925 |
| 33 | PPP2R3A | NM_002718 | -2.556 | -3.292 | -2.924 |
| 34 | HELZ | XM_375485 | -3.218 | -2.596 | -2.907 |

| | | | | | |
|---|---|---|---|---|---|
| 35 | CYBA | NM_000101 | -3.293 | -2.510 | -2.901 |
| 36 | MOCS1 | NM_005943 | -2.107 | -3.669 | -2.888 |
| 37 | SFXN3 | NM_030971 | -3.472 | -2.303 | -2.887 |
| 38 | TMEM14B | NM_030969 | -3.474 | -2.283 | -2.879 |
| 39 | HIVEP3 | NM_024503 | -2.389 | -3.365 | -2.877 |
| 40 | HNRNPU | NM_004501 | -2.241 | -3.498 | -2.869 |
| 41 | TRPA1 | NM_007332 | -2.845 | -2.876 | -2.861 |
| 42 | PPAPDC2 | NM_203453 | -3.317 | -2.393 | -2.855 |
| 43 | MVP | NM_005115 | -2.696 | -2.996 | -2.846 |
| 44 | ABP1 | NM_001091 | -2.278 | -3.403 | -2.840 |
| 45 | ZNF434 | NM_017810 | -3.304 | -2.375 | -2.840 |
| 46 | ETF1 | NM_004730 | -2.589 | -3.069 | -2.829 |
| 47 | DDIT3 | NM_004083 | -2.725 | -2.926 | -2.825 |
| 48 | SOX14 | NM_004189 | -2.581 | -3.047 | -2.814 |
| 49 | NRSN1 | NM_080723 | -2.946 | -2.630 | -2.788 |
| 50 | SF3A1 | NM_005877 | -2.227 | -3.347 | -2.787 |
| 51 | C19orf24 | NM_017914 | -2.978 | -2.572 | -2.775 |
| 52 | ANGPT4 | NM_015985 | -2.170 | -3.366 | -2.768 |
| 53 | MGC35434 | NM_198543 | -2.815 | -2.713 | -2.764 |
| 54 | CCDC130 | NM_030818 | -1.578 | -3.944 | -2.761 |
| 55 | ZNF138 | NM_006524 | -2.911 | -2.602 | -2.757 |
| 56 | PHB | NM_002634 | -1.656 | -3.831 | -2.743 |
| 57 | LOC391012 | XM_372768 | -2.787 | -2.699 | -2.743 |
| 58 | LOC400942 | XM_379075 | -2.275 | -3.197 | -2.736 |
| 59 | ARL5C | XM_372668 | -2.883 | -2.581 | -2.732 |
| 60 | APLP2 | NM_001642 | -2.627 | -2.837 | -2.732 |
| 61 | CREBL2 | NM_001310 | -2.448 | -3.008 | -2.728 |
| 62 | LOC387898 | XM_373556 | -3.054 | -2.398 | -2.726 |
| 63 | CHCHD10 | NM_213720 | -2.824 | -2.628 | -2.726 |
| 64 | BCL6B | NM_181844 | -2.397 | -3.048 | -2.723 |
| 65 | MIRHG1 | NM_213723 | -3.043 | -2.399 | -2.721 |
| 66 | CRKRS | NM_016507 | -2.555 | -2.873 | -2.714 |
| 67 | FLJ44881 | NM_207461 | -2.820 | -2.600 | -2.710 |
| 68 | SOAT2 | NM_003578 | -2.520 | -2.880 | -2.700 |
| 69 | CRLF1 | NM_004750 | -3.063 | -2.320 | -2.691 |
| 70 | TPM1 | NM_000366 | -2.810 | -2.562 | -2.686 |
| 71 | CPSF3 | NM_016207 | -2.755 | -2.617 | -2.686 |
| 72 | PCF11 | NM_015885 | -2.792 | -2.572 | -2.682 |
| 73 | TTC23L | NM_144725 | -2.927 | -2.434 | -2.681 |
| 74 | MMP24 | NM_006690 | -3.364 | -1.995 | -2.679 |

| | | | | | |
|---|---|---|---|---|---|
| 75 | ZDHHC20 | NM_153251 | -2.710 | -2.636 | -2.673 |
| 76 | INO80E | NM_173618 | -3.166 | -2.159 | -2.663 |
| 77 | NUDT8 | NM_181843 | -2.991 | -2.333 | -2.662 |
| 78 | ZNF35 | NM_003420 | -2.764 | -2.557 | -2.660 |
| 79 | MCRS1 | NM_006337 | -2.706 | -2.614 | -2.660 |
| 80 | MED19 | NM_153450 | -2.284 | -3.030 | -2.657 |
| 81 | PRDM9 | NM_020227 | -2.682 | -2.624 | -2.653 |
| 82 | LOC151658 | XM_379205 | -2.723 | -2.570 | -2.647 |
| 83 | TPR | NM_003292 | -3.039 | -2.238 | -2.638 |
| 84 | MAGT1 | NM_032121 | -2.877 | -2.398 | -2.638 |
| 85 | PROP1 | NM_006261 | -2.493 | -2.780 | -2.636 |
| 86 | CLRN3 | NM_152311 | -2.831 | -2.438 | -2.634 |
| 87 | IGFBP6 | NM_002178 | -3.035 | -2.215 | -2.625 |
| 88 | LOC401725 | XM_377278 | -3.158 | -2.091 | -2.624 |
| 89 | C19orf68 | NM_199341 | -2.652 | -2.589 | -2.620 |
| 90 | NEUROD2 | NM_006160 | -1.827 | -3.387 | -2.607 |
| 91 | GLTSCR1 | NM_015711 | -2.229 | -2.983 | -2.606 |
| 92 | CGGBP1 | NM_003663 | -2.407 | -2.796 | -2.601 |
| 93 | LOC254571 | XM_170783 | -2.659 | -2.538 | -2.599 |
| 94 | RPESP | NM_153225 | -2.663 | -2.530 | -2.597 |
| 95 | POLH | NM_006502 | -2.558 | -2.614 | -2.586 |
| 96 | ADAMTS1 | NM_006988 | -2.808 | -2.351 | -2.579 |
| 97 | LARS | NM_020117 | -3.013 | -2.136 | -2.574 |
| 98 | YY1 | NM_003403 | -2.113 | -3.031 | -2.572 |
| 99 | ABTB1 | NM_172027 | -2.285 | -2.854 | -2.570 |
| 100 | ENPP7 | NM_178543 | -2.824 | -2.315 | -2.569 |
| 101 | ODF2 | NM_002540 | -2.305 | -2.832 | -2.568 |
| 102 | C20orf95 | XM_293123 | -2.728 | -2.393 | -2.561 |
| 103 | HES6 | NM_018645 | -2.074 | -3.041 | -2.557 |
| 104 | ZNF43 | NM_003423 | -2.312 | -2.798 | -2.555 |
| 105 | PXN | NM_002859 | -2.948 | -2.161 | -2.554 |
| 106 | INCA1 | NM_213726 | -2.719 | -2.384 | -2.551 |
| 107 | MED13L | NM_015335 | -2.380 | -2.719 | -2.550 |
| 108 | SLC16A3 | NM_004207 | -2.461 | -2.638 | -2.549 |
| 109 | MR1 | NM_001531 | -4.217 | -0.877 | -2.547 |
| 110 | EIF2B3 | NM_020365 | -1.563 | -3.530 | -2.546 |
| 111 | TRAIP | NM_005879 | -1.121 | -3.971 | -2.546 |
| 112 | TBC1D10A | NM_031937 | -2.777 | -2.311 | -2.544 |
| 113 | MGC10981 | XM_378193 | -3.032 | -2.050 | -2.541 |
| 114 | SERPINB2 | NM_002575 | -2.438 | -2.635 | -2.536 |

| | | | | | |
|---|---|---|---|---|---|
| 115 | COPS2 | NM_004236 | -1.977 | -3.085 | -2.531 |
| 116 | GLRB | NM_000824 | -3.924 | -1.137 | -2.530 |
| 117 | DRG2 | NM_001388 | -2.501 | -2.552 | -2.526 |
| 118 | CCDC60 | NM_178499 | -2.494 | -2.558 | -2.526 |
| 119 | COL11A1 | NM_001854 | -2.735 | -2.306 | -2.520 |
| 120 | TPD52L1 | NM_003287 | -2.499 | -2.540 | -2.520 |
| 121 | SAMD7 | NM_182610 | -2.886 | -2.138 | -2.512 |
| 122 | SORDL | XM_007651 | -2.692 | -2.332 | -2.512 |
| 123 | NCBP1 | NM_002486 | -2.649 | -2.370 | -2.510 |
| 124 | LIF | NM_002309 | -2.508 | -1.712 | -2.508 |
| 125 | FAM19A1 | NM_213609 | -3.587 | -1.419 | -2.503 |
| 126 | ADA | NM_000022 | -1.706 | -3.292 | -2.499 |
| 127 | SLC46A3 | NM_181785 | -2.784 | -2.186 | -2.485 |
| 128 | GJA8 | NM_005267 | -2.186 | -2.781 | -2.483 |
| 129 | MED24 | NM_014815 | -2.579 | -2.386 | -2.482 |
| 130 | LOC400726 | XM_375670 | -2.276 | -2.686 | -2.481 |
| 131 | ANXA4 | NM_001153 | -2.177 | -2.780 | -2.478 |
| 132 | NDUFA4L2 | NM_020142 | -3.385 | -1.569 | -2.477 |
| 133 | EIF2B4 | NM_015636 | -3.596 | -1.347 | -2.471 |
| 134 | TMEM204 | NM_024600 | -2.177 | -2.758 | -2.468 |
| 135 | MBTD1 | NM_017643 | -1.655 | -3.279 | -2.467 |
| 136 | FAM82A1 | NM_144713 | -2.649 | -2.284 | -2.467 |
| 137 | LOC387783 | XM_373505 | -2.656 | -2.277 | -2.467 |
| 138 | KIR3DL1 | NM_013289 | -3.365 | -1.566 | -2.465 |
| 139 | MED12 | NM_005120 | -2.094 | -2.835 | -2.465 |
| 140 | FTSJ1 | NM_012280 | -1.561 | -3.365 | -2.463 |
| 141 | EIF2B2 | NM_014239 | -2.874 | -2.050 | -2.462 |
| 142 | KIRREL2 | NM_032123 | -2.905 | -2.018 | -2.461 |
| 143 | PODNL1 | NM_024825 | -2.538 | -2.383 | -2.461 |
| 144 | ACPL2 | NM_152282 | -2.218 | -2.698 | -2.458 |
| 145 | TCL1A | NM_021966 | -3.312 | -1.603 | -2.457 |
| 146 | LOC401440 | XM_379539 | -2.426 | -2.488 | -2.457 |
| 147 | PDZD11 | NM_016484 | -2.619 | -2.295 | -2.457 |
| 148 | WDR82 | NM_025222 | -1.585 | -3.324 | -2.455 |
| 149 | C20orf59 | NM_022082 | -2.739 | -2.167 | -2.453 |
| 150 | NALCN | NM_052867 | -2.615 | -2.290 | -2.453 |
| 151 | AIPL1 | NM_014336 | -2.590 | -2.312 | -2.451 |
| 152 | SYF2 | NM_015484 | -2.307 | -2.591 | -2.449 |
| 153 | ANKRD1 | NM_014391 | -2.460 | -2.435 | -2.448 |
| 154 | MALL | NM_005434 | -2.709 | -2.179 | -2.444 |

| | | | | | |
|---|---|---|---|---|---|
| 155 | KREMEN1 | NM_032045 | -2.090 | -2.793 | -2.442 |
| 156 | ANKRD31 | XM_293911 | -2.283 | -2.600 | -2.441 |
| 157 | LOC389257 | XM_371722 | -2.755 | -2.127 | -2.441 |
| 158 | FLJ45244 | NM_207443 | -2.769 | -2.110 | -2.440 |
| 159 | RBM17 | NM_032905 | -2.339 | -2.540 | -2.439 |
| 160 | LOC400604 | XM_378684 | -2.881 | -1.998 | -2.439 |
| 161 | FAM169B | NM_182562 | -2.601 | -2.272 | -2.437 |
| 162 | CORT | NM_001302 | -2.132 | -2.733 | -2.433 |
| 163 | LCE1E | NM_178353 | -2.599 | -2.257 | -2.428 |
| 164 | CORT | NM_001302 | -2.043 | -2.810 | -2.427 |
| 165 | ZC3H18 | NM_144604 | -2.782 | -2.071 | -2.426 |
| 166 | NXF1 | NM_006362 | -1.686 | -3.166 | -2.426 |
| 167 | AGPS | NM_003659 | -2.808 | -2.039 | -2.424 |
| 168 | ALDH1L2 | XM_090294 | -2.503 | -2.341 | -2.422 |
| 169 | H1FX | NM_006026 | -2.087 | -2.756 | -2.421 |
| 170 | MED14 | NM_004229 | -1.870 | -2.970 | -2.420 |
| 171 | LOC390565 | XM_374355 | -2.609 | -2.226 | -2.418 |
| 172 | FLJ41047 | XM_374945 | -2.609 | -2.224 | -2.416 |
| 173 | FUBP1 | NM_003902 | -2.888 | -1.939 | -2.414 |
| 174 | DEFB126 | NM_030931 | -2.163 | -2.653 | -2.408 |
| 175 | RHOA | NM_001664 | -2.020 | -2.795 | -2.408 |
| 176 | FOXJ3 | NM_014947 | -2.125 | -2.690 | -2.407 |
| 177 | NBPF15 | NM_173638 | -2.069 | -2.742 | -2.405 |
| 178 | MED28 | NM_025205 | -3.356 | -1.452 | -2.404 |
| 179 | LPPR2 | NM_022737 | -2.785 | -2.022 | -2.404 |
| 180 | SYNCRIP | NM_006372 | -2.724 | -2.079 | -2.401 |
| 181 | SYTL4 | NM_080737 | -1.118 | -3.678 | -2.398 |
| 182 | DNAH8 | NM_001371; XM_353628 | -1.990 | -2.806 | -2.398 |
| 183 | MBOAT1 | XM_351649 | -2.847 | -1.948 | -2.398 |
| 184 | FTL | NM_000146 | -3.132 | -1.660 | -2.396 |
| 185 | FAM127B | NM_015582 | -2.287 | -2.495 | -2.391 |
| 186 | ZSCAN10 | NM_032805 | -2.969 | -1.810 | -2.390 |
| 187 | FIG4 | NM_014845 | -2.314 | -2.464 | -2.389 |
| 188 | FLJ45684 | NM_207462 | -2.306 | -2.472 | -2.389 |
| 189 | LOC346910 | XM_294456 | -3.026 | -1.746 | -2.386 |
| 190 | TAF2 | NM_003184 | -2.786 | -1.982 | -2.384 |
| 191 | TCP1 | NM_030752 | -1.834 | -2.933 | -2.383 |
| 192 | VWF | NM_000552 | -1.732 | -3.031 | -2.381 |
| 193 | LOC401513 | XM_379637 | -2.256 | -2.505 | -2.381 |
| 194 | LDB2 | NM_001290 | -1.662 | -3.099 | -2.380 |

| | | | | | |
|---|---|---|---|---|---|
| 195 | FXYD7 | NM_022006 | -2.918 | -1.840 | -2.379 |
| 196 | LOC338731 | XM_294688 | -2.925 | -1.831 | -2.378 |
| 197 | EDF1 | NM_003792 | -2.115 | -2.639 | -2.377 |
| 198 | SLC2A12 | NM_145176 | -2.710 | -2.033 | -2.372 |
| 199 | TAAR5 | NM_003967 | -1.990 | -2.752 | -2.371 |
| 200 | MMP15 | NM_002428 | -2.753 | -1.987 | -2.370 |
| 201 | E4F1 | NM_004424 | -1.537 | -3.203 | -2.370 |
| 202 | EMX1 | NM_004097 | -2.316 | -2.422 | -2.369 |
| 203 | CYP11B2 | NM_000498 | -3.065 | -1.672 | -2.368 |
| 204 | GSTT2 | NM_000854 | -3.263 | -1.471 | -2.367 |
| 205 | SHFM1 | NM_006304 | -2.074 | -2.658 | -2.366 |
| 206 | VARS | NM_006295 | -1.988 | -2.743 | -2.366 |
| 207 | ODF3L1 | NM_175881 | -2.262 | -2.469 | -2.365 |
| 208 | CLASP1 | NM_015282 | -1.953 | -2.777 | -2.365 |
| 209 | CCDC33 | NM_182791 | -2.952 | -1.777 | -2.365 |
| 210 | PRDX6 | NM_004905 | -2.440 | -2.289 | -2.364 |
| 211 | PSTPIP2 | NM_024430 | -2.573 | -2.155 | -2.364 |
| 212 | OR1D2 | NM_002548 | -2.736 | -1.991 | -2.363 |
| 213 | C6orf195 | NM_152554 | -2.591 | -2.125 | -2.358 |
| 214 | CCNK | NM_003858 | -2.856 | -1.853 | -2.354 |
| 215 | SH3BP4 | NM_014521 | -1.508 | -3.200 | -2.354 |
| 216 | SPINK7 | NM_032566 | -2.465 | -2.238 | -2.351 |
| 217 | CAMP | NM_004345 | -1.054 | -3.649 | -2.351 |
| 218 | LOC400733 | XM_378835 | -2.461 | -2.238 | -2.349 |
| 219 | C16orf81 | NM_173617 | -2.157 | -2.539 | -2.348 |
| 220 | CNOT1 | NM_016284 | -2.538 | -2.156 | -2.347 |
| 221 | KIAA1999 | XM_114447 | -2.431 | -2.260 | -2.346 |
| 222 | SMU1 | NM_018225 | -1.677 | -3.014 | -2.345 |
| 223 | KRT39 | NM_213656 | -2.466 | -2.218 | -2.342 |
| 224 | KIAA1853 | XM_375000; NM_194286 | -2.729 | -1.951 | -2.340 |
| 225 | LOC402679 | XM_380022 | -3.185 | -1.493 | -2.339 |
| 226 | CCT7 | NM_006429 | -1.771 | -2.898 | -2.335 |
| 227 | LOC284214 | XM_378734 | -2.325 | -2.342 | -2.334 |
| 228 | RBM10 | NM_005676 | -2.444 | -2.214 | -2.329 |
| 229 | CHRDL1 | NM_145234 | -1.631 | -3.022 | -2.327 |
| 230 | ACSS1 | NM_032501 | -2.110 | -2.542 | -2.326 |
| 231 | HIST1H4K | NM_003541 | -2.194 | -2.454 | -2.324 |
| 232 | ADIPOR1 | NM_015999 | -2.244 | -2.404 | -2.324 |
| 233 | ACTR1B | NM_005735 | -2.011 | -2.630 | -2.320 |
| 234 | WDR53 | NM_182627 | -2.548 | -2.093 | -2.320 |

| | | | | | |
|---|---|---|---|---|---|
| 235 | TTC4 | NM_004623 | -2.394 | -2.246 | -2.320 |
| 236 | LOC392425 | XM_373338 | -2.434 | -2.192 | -2.313 |
| 237 | GSPT1 | NM_002094 | -2.408 | -2.218 | -2.313 |
| 238 | LOC389159 | XM_374055 | -2.706 | -1.919 | -2.312 |
| 239 | LOC389834 | XM_374318 | -2.232 | -2.392 | -2.312 |
| 240 | KIF19 | NM_153209 | -2.046 | -2.577 | -2.312 |
| 241 | TUBB4 | NM_006087 | -2.340 | -2.283 | -2.312 |
| 242 | CDH18 | NM_004934 | -2.283 | -2.339 | -2.311 |
| 243 | ATOH8 | NM_032827 | -1.209 | -3.413 | -2.311 |
| 244 | LOC387706 | XM_373471 | -2.221 | -2.398 | -2.310 |
| 245 | MGC9913 | XM_378178 | -2.361 | -2.258 | -2.310 |
| 246 | ARL4C | NM_005737 | -1.637 | -2.978 | -2.307 |
| 247 | MAGEH1 | NM_014061 | -1.686 | -2.928 | -2.307 |
| 248 | SUV39H2 | NM_024670 | -2.890 | -1.724 | -2.307 |
| 249 | LOC283951 | XM_370925 | -2.437 | -2.171 | -2.304 |
| 250 | TRAM2 | NM_012288 | -1.774 | -2.833 | -2.304 |
| 251 | REXO1 | NM_020695 | -2.982 | -1.623 | -2.303 |
| 252 | LOC138729 | XM_072554 | -2.942 | -1.660 | -2.301 |
| 253 | NHLRC1 | NM_198586 | -2.767 | -1.825 | -2.296 |
| 254 | GLUD1 | NM_005271 | -1.936 | -2.654 | -2.295 |
| 255 | CTNNA3 | NM_013266 | -2.647 | -1.940 | -2.293 |
| 256 | APC | NM_000038 | -2.344 | -2.242 | -2.293 |
| 257 | PPY2 | NM_021092 | -2.755 | -1.831 | -2.293 |
| 258 | ULK2 | NM_014683 | -2.175 | -2.409 | -2.292 |
| 259 | FBXW2 | NM_012164 | -3.293 | -1.291 | -2.292 |
| 260 | KLRG2 | NM_198508 | -2.567 | -2.016 | -2.292 |
| 261 | ISL1 | NM_002202 | -2.627 | -1.956 | -2.291 |
| 262 | DERL2 | NM_016041 | -1.962 | -2.620 | -2.291 |
| 263 | IMPAD1 | NM_017813 | -1.890 | -2.676 | -2.283 |
| 264 | HS3ST3B1 | NM_006041 | -2.327 | -2.236 | -2.281 |
| 265 | UFSP1 | XM_380026 | -2.544 | -2.017 | -2.280 |
| 266 | C12orf68 | XM_370691 | -2.627 | -1.930 | -2.279 |
| 267 | ARS2 | NM_015908 | -3.128 | -1.428 | -2.278 |
| 268 | C22orf15 | NM_182520 | -2.919 | -1.637 | -2.278 |
| 269 | ETFDH | NM_004453 | -2.659 | -1.894 | -2.276 |
| 270 | EP300 | NM_001429 | -1.511 | -3.038 | -2.274 |
| 271 | EMP3 | NM_001425 | -3.273 | -1.273 | -2.273 |
| 272 | C1orf157 | NM_182579 | -2.187 | -2.357 | -2.272 |
| 273 | LOC392584 | XM_373397 | -2.404 | -2.137 | -2.270 |
| 274 | WDR5B | NM_019069 | -1.994 | -2.546 | -2.270 |

| | | | | | |
|---|---|---|---|---|---|
| 275 | NUP107 | NM_020401 | -3.452 | -1.081 | -2.267 |
| 276 | LOC341333 | XM_296117 | -2.464 | -2.066 | -2.265 |
| 277 | PMS2L1 | XM_377962; XM_380025 | -2.838 | -1.691 | -2.264 |
| 278 | CLTB | NM_001834 | -2.250 | -2.276 | -2.263 |
| 279 | ZNF513 | NM_144631 | -2.114 | -2.398 | -2.256 |
| 280 | OR5K4 | NM_001005517 | -2.132 | -2.380 | -2.256 |
| 281 | LOC402677 | XM_380021 | -2.918 | -1.593 | -2.255 |
| 282 | GK5 | NM_152776 | -2.558 | -1.949 | -2.254 |
| 283 | FBXL6 | NM_012162 | -2.250 | -2.257 | -2.253 |
| 284 | CSTB | NM_000100 | -2.750 | -1.754 | -2.252 |
| 285 | ZDHHC15 | NM_144969 | -2.796 | -1.708 | -2.252 |
| 286 | GSTP1 | NM_000852 | -2.402 | -2.099 | -2.251 |
| 287 | BEST1 | NM_004183 | -1.903 | -2.598 | -2.250 |
| 288 | ADH5P4 | XM_208352 | -2.121 | -2.374 | -2.247 |
| 289 | ASTN2 | NM_014010 | -1.733 | -2.760 | -2.247 |
| 290 | MS4A6E | NM_139249 | -2.585 | -1.901 | -2.243 |
| 291 | LOC284402 | XM_378794 | -2.508 | -1.975 | -2.242 |
| 292 | CERCAM | NM_016174 | -4.335 | -0.139 | -2.237 |
| 293 | PHB2 | NM_007273 | -1.604 | -2.868 | -2.236 |
| 294 | LPAR2 | NM_004720 | -1.883 | -2.589 | -2.236 |
| 295 | GBE1 | NM_000158 | -2.527 | -1.944 | -2.236 |
| 296 | MRPL51 | NM_016497 | -2.626 | -1.843 | -2.235 |
| 297 | LOC154092 | XM_379434 | -2.684 | -1.783 | -2.234 |
| 298 | MASP1 | NM_001879 | -2.357 | -2.105 | -2.231 |
| 299 | ZNF546 | NM_178544 | -2.379 | -2.083 | -2.231 |
| 300 | TAS2R45 | NM_176886 | -2.600 | -1.857 | -2.229 |
| 301 | CPT2 | NM_000098 | -2.468 | -1.970 | -2.219 |
| 302 | NMUR2 | NM_020167 | -1.707 | -2.723 | -2.215 |
| 303 | PARP3 | NM_005485 | -2.653 | -1.777 | -2.215 |
| 304 | LOC389667 | XM_372046 | -2.163 | -2.260 | -2.211 |
| 305 | ICT1 | NM_001545 | -3.251 | -1.169 | -2.210 |
| 306 | MYBPHL | XM_291485 | -2.790 | -1.628 | -2.209 |
| 307 | BAIAP2L1 | NM_018842 | -2.836 | -1.578 | -2.207 |
| 308 | GPR172B | NM_017986 | -1.454 | -2.958 | -2.206 |
| 309 | APLF | NM_173545 | -2.967 | -1.444 | -2.206 |
| 310 | CASC1 | NM_018272 | -2.472 | -1.937 | -2.204 |
| 311 | ADAM7 | NM_003817 | -1.894 | -2.514 | -2.204 |
| 312 | ADAM28 | NM_014265 | -2.007 | -2.399 | -2.203 |
| 313 | DNAJC10 | NM_018981 | -2.375 | -2.029 | -2.202 |
| 314 | CTSE | NM_001910 | -1.201 | -3.203 | -2.202 |

| | | | | | |
|---|---|---|---|---|---|
| 315 | IBSP | NM_004967 | -2.764 | -1.636 | -2.200 |
| 316 | CLPP | NM_006012 | -2.222 | -2.178 | -2.200 |
| 317 | MYBPC2 | NM_004533 | -1.907 | -2.491 | -2.199 |
| 318 | CDC42 | NM_001791 | -2.290 | -2.107 | -2.199 |
| 319 | LOC257054 | NM_152783 | -2.760 | -1.636 | -2.198 |
| 320 | ANGPTL6 | NM_031917 | -2.076 | -2.319 | -2.198 |
| 321 | ZNF197 | NM_006991 | -1.458 | -2.933 | -2.195 |
| 322 | ATP1A3 | NM_152296 | -2.325 | -2.064 | -2.194 |
| 323 | CLTA | NM_001833 | -1.980 | -2.403 | -2.192 |
| 324 | ABCG8 | NM_022437 | -1.955 | -2.422 | -2.189 |
| 325 | HIST1H3B | NM_003537 | -2.773 | -1.599 | -2.186 |
| 326 | OMD | NM_005014 | -1.606 | -2.766 | -2.186 |
| 327 | CENPA | NM_001809 | -2.062 | -2.309 | -2.185 |
| 328 | FAM120B | NM_032448 | -2.813 | -1.555 | -2.184 |
| 329 | LOC388340 | XM_373715 | -2.532 | -1.833 | -2.183 |
| 330 | LOC340515 | XM_379650 | -2.598 | -1.766 | -2.182 |
| 331 | MS4A2 | NM_000139 | -2.333 | -2.031 | -2.182 |
| 332 | LOC388934 | XM_373977 | -3.105 | -1.255 | -2.180 |
| 333 | LOC391405 | XM_372941 | -2.606 | -1.752 | -2.179 |
| 334 | TAF7 | NM_005642 | -1.300 | -3.052 | -2.176 |
| 335 | SETD1B | XM_037523 | -2.063 | -2.286 | -2.174 |
| 336 | GTF2H3 | NM_001516 | -3.742 | -0.605 | -2.173 |
| 337 | P2RY11 | NM_002566 | -2.214 | -2.130 | -2.172 |
| 338 | LOC345537 | XM_293868 | -2.537 | -1.802 | -2.170 |
| 339 | LOC284551 | XM_375713 | -1.813 | -2.526 | -2.170 |
| 340 | KEAP1 | NM_012289 | -2.997 | -1.340 | -2.169 |
| 341 | TTC39B | NM_152574 | -2.612 | -1.722 | -2.167 |
| 342 | RALGDS | NM_006266 | -2.995 | -1.330 | -2.163 |
| 343 | CTDP1 | NM_004715 | -1.827 | -2.496 | -2.161 |
| 344 | MED17 | NM_004268 | -2.412 | -1.910 | -2.161 |
| 345 | ACTR6 | NM_022496 | -2.031 | -2.292 | -2.161 |
| 346 | GAB3 | NM_080612 | -2.548 | -1.770 | -2.159 |
| 347 | CA9 | NM_001216 | -2.731 | -1.582 | -2.157 |
| 348 | HINT1 | NM_005340 | -1.444 | -2.868 | -2.156 |
| 349 | KRTAP22-1 | NM_181620 | -2.580 | -1.732 | -2.156 |
| 350 | ECHDC1 | NM_018479 | -3.401 | -0.911 | -2.156 |
| 351 | RNF216 | NM_207111 | -1.753 | -2.556 | -2.155 |
| 352 | LOC386597 | XM_379073 | -2.200 | -2.109 | -2.154 |
| 353 | HNRNPD | NM_002138 | -2.688 | -1.618 | -2.153 |
| 354 | PITX1 | NM_002653 | -1.952 | -2.353 | -2.153 |

| | | | | | |
|---|---|---|---|---|---|
| 355 | LOC389183 | XM_374065 | -2.221 | -2.081 | -2.151 |
| 356 | RAD9B | NM_152442 | -2.082 | -2.220 | -2.151 |
| 357 | KLK5 | NM_012427 | -2.137 | -2.163 | -2.150 |
| 358 | TFDP2 | NM_006286 | -1.550 | -2.749 | -2.150 |
| 359 | LOC389089 | XM_374029 | -3.496 | -0.802 | -2.149 |
| 360 | C13orf35 | NM_207440 | -2.860 | -1.438 | -2.149 |
| 361 | LOC402414 | XM_378124 | -2.286 | -2.010 | -2.148 |
| 362 | LOC442401 | XM_378044 | -1.910 | -2.386 | -2.148 |
| 363 | GUCY1A3 | NM_000856 | -1.756 | -2.540 | -2.148 |
| 364 | ALKBH8 | NM_138775 | -1.913 | -2.382 | -2.147 |
| 365 | LTA4H | NM_000895 | -1.658 | -2.636 | -2.147 |
| 366 | WBP2 | NM_012478 | -1.601 | -2.692 | -2.147 |
| 367 | KPNA5 | NM_002269 | -1.663 | -2.631 | -2.147 |
| 368 | FAM114A1 | NM_138389 | -2.124 | -2.166 | -2.145 |
| 369 | SVIL | NM_003174 | -2.334 | -1.954 | -2.144 |
| 370 | GFER | NM_005262 | -2.789 | -1.495 | -2.142 |
| 371 | COX6A2 | NM_005205 | -2.058 | -2.226 | -2.142 |
| 372 | LOC402671 | XM_380013 | -3.663 | -0.619 | -2.141 |
| 373 | WAC | NM_016628 | -2.512 | -1.766 | -2.139 |
| 374 | LOC400751 | XM_378859 | -2.885 | -1.393 | -2.139 |
| 375 | CENPQ | NM_018132 | -2.634 | -1.640 | -2.137 |
| 376 | CD59 | NM_000611 | -1.978 | -2.295 | -2.136 |
| 377 | UPP2 | NM_173355 | -2.778 | -1.494 | -2.136 |
| 378 | HMCN2 | XM_175125 | -2.551 | -1.720 | -2.136 |
| 379 | GYPC | NM_002101 | -2.720 | -1.550 | -2.135 |
| 380 | ZNF711 | NM_021998 | -3.184 | -1.085 | -2.134 |
| 381 | ITPKA | NM_002220 | -2.083 | -2.186 | -2.134 |
| 382 | TAF12 | NM_005644 | -2.034 | -2.234 | -2.134 |
| 383 | PLCL1 | NM_006226 | -1.776 | -2.488 | -2.132 |
| 384 | KLC3 | NM_177417 | -2.244 | -2.016 | -2.130 |
| 385 | PPP1R15B | NM_032833 | -1.806 | -2.452 | -2.129 |
| 386 | C15orf29 | NM_024713 | -1.853 | -2.405 | -2.129 |
| 387 | ACTA1 | NM_001100 | -3.241 | -1.016 | -2.128 |
| 388 | ACBD3 | NM_022735 | -2.107 | -2.149 | -2.128 |
| 389 | OR2H1 | NM_030883 | -1.413 | -2.842 | -2.127 |
| 390 | LOC150759 | NM_175853 | -2.213 | -2.038 | -2.126 |
| 391 | LOC389628 | XM_374249 | -1.854 | -2.395 | -2.125 |
| 392 | RAP1GAP | NM_002885 | -2.700 | -1.549 | -2.125 |
| 393 | TAX1BP1 | NM_006024 | -1.957 | -2.289 | -2.123 |
| 394 | DBX2 | NM_001004329 | -2.009 | -2.236 | -2.123 |

| | | | | | |
|---|---|---|---|---|---|
| 395 | NFKB1 | NM_003998 | -1.441 | -2.804 | -2.123 |
| 396 | TSPAN4 | NM_003271 | -1.444 | -2.799 | -2.122 |
| 397 | LOC402282 | XM_377949 | -2.179 | -2.063 | -2.121 |
| 398 | BLM | NM_000057 | -2.082 | -2.159 | -2.121 |
| 399 | C2orf69 | NM_153689 | -1.654 | -2.587 | -2.121 |
| 400 | LOC285733 | XM_379432 | -2.001 | -2.240 | -2.121 |
| 401 | HIST1H4G | NM_003547 | -1.970 | -2.270 | -2.120 |
| 402 | C6orf163 | XM_116497 | -2.127 | -2.109 | -2.118 |
| 403 | KIAA0195 | NM_014738 | -2.939 | -1.295 | -2.117 |
| 404 | GALNTL4 | NM_198516 | -1.950 | -2.282 | -2.116 |
| 405 | HMOX2 | NM_002134 | -1.348 | -2.884 | -2.116 |
| 406 | C9orf94 | NM_152702 | -2.353 | -1.875 | -2.114 |
| 407 | ECE2 | NM_014693 | -1.691 | -2.537 | -2.114 |
| 408 | RAD17 | NM_002873 | -2.278 | -1.950 | -2.114 |
| 409 | CCDC74B | NM_207310 | -1.687 | -2.540 | -2.113 |
| 410 | OR7D4 | XM_064879 | -2.827 | -1.400 | -2.113 |
| 411 | NDUFV1 | NM_007103 | -1.748 | -2.476 | -2.112 |
| 412 | AADACL2 | NM_207365 | -2.451 | -1.772 | -2.112 |
| 413 | CXorf59 | NM_173695 | -1.566 | -2.651 | -2.109 |
| 414 | MGC39821 | NM_182576 | -1.369 | -2.845 | -2.107 |
| 415 | PRDM13 | NM_021620 | -1.941 | -2.272 | -2.107 |
| 416 | BDP1 | NM_018429 | -2.227 | -1.986 | -2.107 |
| 417 | CLCN1 | NM_000083 | -1.941 | -2.267 | -2.104 |
| 418 | GPR119 | NM_178471 | -1.582 | -2.625 | -2.103 |
| 419 | MAGEB4 | NM_002367 | -3.239 | -0.966 | -2.102 |
| 420 | DFFA | NM_004401 | -2.060 | -2.141 | -2.101 |
| 421 | GLO1 | NM_006708 | -2.062 | -2.139 | -2.101 |
| 422 | LY6D | NM_003695 | -1.981 | -2.219 | -2.100 |
| 423 | HADHA | NM_000182 | -1.900 | -2.301 | -2.100 |
| 424 | CD34 | NM_001773 | -2.063 | -2.136 | -2.100 |
| 425 | CYP21A2 | NM_000500 | -1.771 | -2.422 | -2.097 |
| 426 | PPOX | NM_000309 | -1.672 | -2.519 | -2.096 |
| 427 | TIFA | NM_052864 | -2.171 | -2.018 | -2.095 |
| 428 | TRYX3 | NM_001001317 | -1.852 | -2.335 | -2.093 |
| 429 | TOB2 | NM_016272 | -2.320 | -1.866 | -2.093 |
| 430 | MAGEB5 | XM_293407 | -2.617 | -1.568 | -2.093 |
| 431 | XRCC1 | NM_006297 | -1.839 | -2.346 | -2.092 |
| 432 | ALS2CR16 | NM_205543 | -2.505 | -1.679 | -2.092 |
| 433 | EFCAB4A | NM_173584 | -2.502 | -1.681 | -2.092 |
| 434 | FLJ44450 | XM_373981 | -2.132 | -2.051 | -2.092 |

| | | | | | |
|---|---|---|---|---|---|
| 435 | C16orf13 | NM_032366 | -2.059 | -2.121 | -2.090 |
| 436 | PRELP | NM_002725 | -2.411 | -1.769 | -2.090 |
| 437 | TMEM14C | NM_016462 | -2.131 | -2.048 | -2.089 |
| 438 | FGL1 | NM_004467 | -2.400 | -1.778 | -2.089 |
| 439 | LOC400053 | XM_378367 | -2.666 | -1.510 | -2.088 |
| 440 | ICA1 | NM_004968 | -1.567 | -2.609 | -2.088 |
| 441 | ZNF786 | NM_152411 | -2.340 | -1.835 | -2.088 |
| 442 | MAP3K1 | XM_042066 | -1.736 | -2.439 | -2.088 |
| 443 | DARS | NM_001349 | -2.005 | -2.168 | -2.086 |
| 444 | ZNHIT3 | NM_004773 | -1.621 | -2.549 | -2.085 |
| 445 | TFPT | NM_013342 | -1.647 | -2.522 | -2.084 |
| 446 | GC | NM_000583 | -1.515 | -2.652 | -2.083 |
| 447 | SRP14P1 | XM_372448 | -2.439 | -1.728 | -2.083 |
| 448 | GAD2 | NM_000818 | -2.287 | -1.878 | -2.082 |
| 449 | RNF40 | NM_014771 | -1.772 | -2.392 | -2.082 |
| 450 | SENP6 | NM_015571 | -2.309 | -1.855 | -2.082 |
| 451 | CCDC74A | NM_138770 | -1.733 | -2.430 | -2.081 |
| 452 | C17orf74 | NM_175734 | -2.149 | -2.013 | -2.081 |
| 453 | C10orf40 | XM_378230 | -2.110 | -2.052 | -2.081 |
| 454 | PCDHA1 | NM_018900 | -1.962 | -2.200 | -2.081 |
| 455 | ADFP | NM_001122 | -1.984 | -2.173 | -2.079 |
| 456 | CDH19 | NM_021153 | -2.273 | -1.880 | -2.077 |
| 457 | PDE2A | NM_002599 | -1.493 | -2.660 | -2.077 |
| 458 | FXYD3 | NM_005971 | -1.563 | -2.588 | -2.076 |
| 459 | RRAS | NM_006270 | -2.277 | -1.869 | -2.073 |
| 460 | SNAP23 | NM_003825 | -1.564 | -2.582 | -2.073 |
| 461 | MED13 | NM_005121 | -1.425 | -2.721 | -2.073 |
| 462 | LOC220906 | XM_374781 | -2.167 | -1.978 | -2.072 |
| 463 | ASAP1 | NM_018482 | -2.451 | -1.692 | -2.072 |
| 464 | PITX3 | NM_005029 | -1.424 | -2.719 | -2.071 |
| 465 | SF3A3 | NM_006802 | -1.663 | -2.479 | -2.071 |
| 466 | FAM174B | NM_207446 | -2.192 | -1.950 | -2.071 |
| 467 | TIGD3 | NM_145719 | -2.455 | -1.687 | -2.071 |
| 468 | TALDO1 | NM_006755 | -1.727 | -2.414 | -2.070 |
| 469 | EYA3 | NM_001990 | -1.895 | -2.245 | -2.070 |
| 470 | AHCTF1 | NM_015446 | -1.458 | -2.680 | -2.069 |
| 471 | PNPO | NM_018129 | -2.353 | -1.786 | -2.069 |
| 472 | SCGB1A1 | NM_003357 | -1.991 | -2.141 | -2.066 |
| 473 | LOC391749 | XM_373061 | -3.083 | -1.048 | -2.066 |
| 474 | BAT4 | NM_033177 | -2.253 | -1.877 | -2.065 |

| | | | | | |
|---|---|---|---|---|---|
| 475 | LOC283440 | XM_211040 | -2.857 | -1.269 | -2.063 |
| 476 | LOC392217 | XM_373249 | -1.938 | -2.188 | -2.063 |
| 477 | PSPH | NM_004577 | -3.035 | -1.090 | -2.063 |
| 478 | ZNF137 | NM_003438 | -2.507 | -1.617 | -2.062 |
| 479 | LOC285407 | XM_209597 | -2.580 | -1.543 | -2.062 |
| 480 | OPN1SW | NM_001708 | -1.744 | -2.375 | -2.059 |
| 481 | EIF2B1 | NM_001414 | -2.224 | -1.894 | -2.059 |
| 482 | MGC33407 | NM_178525 | -1.951 | -2.167 | -2.059 |
| 483 | LOC391763 | XM_373075 | -2.617 | -1.501 | -2.059 |
| 484 | LOC388174 | XM_370905 | -2.371 | -1.746 | -2.058 |
| 485 | IMPG2 | NM_016247 | -1.414 | -2.702 | -2.058 |
| 486 | GABRP | NM_014211 | -1.879 | -2.236 | -2.057 |
| 487 | RIN1 | NM_004292 | -1.983 | -2.129 | -2.056 |
| 488 | P2RY12 | NM_022788 | -2.134 | -1.975 | -2.055 |
| 489 | OXA1L | NM_005015 | -1.818 | -2.292 | -2.055 |
| 490 | CDC37L1 | NM_017913 | -2.386 | -1.716 | -2.051 |
| 491 | ZNF354A | NM_005649 | -2.214 | -1.888 | -2.051 |
| 492 | FAM45A | NM_207009 | -2.437 | -1.664 | -2.051 |
| 493 | SFPQ | NM_005066 | -2.526 | -1.575 | -2.051 |
| 494 | GJB1 | NM_000166 | -2.016 | -2.082 | -2.049 |
| 495 | PIM2 | NM_006875 | -2.550 | -1.547 | -2.048 |
| 496 | STK4 | NM_006282 | -2.861 | -1.236 | -2.048 |
| 497 | ENTPD1 | NM_001776 | -1.976 | -2.120 | -2.048 |
| 498 | GBP2 | NM_004120 | -2.449 | -1.644 | -2.046 |
| 499 | P2RX2 | NM_012226 | -1.439 | -2.653 | -2.046 |
| 500 | PLS3 | NM_005032 | -1.846 | -2.244 | -2.045 |
| 501 | RBBP6 | NM_006910 | -0.305 | -3.786 | -2.045 |
| 502 | LOC137107 | XM_070233 | -2.311 | -1.777 | -2.044 |
| 503 | STK10 | NM_005990 | -1.955 | -2.132 | -2.044 |
| 504 | LOC389695 | XM_374286 | -1.621 | -2.466 | -2.043 |
| 505 | C21orf54 | XM_295017 | -2.376 | -1.710 | -2.043 |
| 506 | ACTL6B | NM_016188 | -1.905 | -2.180 | -2.042 |
| 507 | SUB1 | NM_006713 | -1.938 | -2.146 | -2.042 |
| 508 | UBXN8 | NM_005671 | -1.656 | -2.426 | -2.041 |
| 509 | IL1F10 | NM_032556 | -2.365 | -1.716 | -2.040 |
| 510 | TMED3 | NM_007364 | -2.076 | -2.004 | -2.040 |
| 511 | JARID1C | NM_004187 | -3.015 | -1.064 | -2.040 |
| 512 | CCL2 | NM_002982 | -1.364 | -2.714 | -2.039 |
| 513 | UBE2V1 | NM_021988 | -2.466 | -1.611 | -2.039 |
| 514 | CDX2 | NM_001265 | -1.854 | -2.220 | -2.037 |

| | | | | | |
|---|---|---|---|---|---|
| 515 | MYO6 | NM_004999 | -2.292 | -1.782 | -2.037 |
| 516 | POLR2E | NM_002695 | -1.800 | -2.273 | -2.037 |
| 517 | LOC388922 | XM_371476 | -3.063 | -1.007 | -2.035 |
| 518 | CTNNBIP1 | NM_020248 | -2.143 | -1.925 | -2.034 |
| 519 | KLRG1 | NM_005810 | -1.619 | -2.446 | -2.033 |
| 520 | ALDH5A1 | NM_001080 | -2.025 | -2.039 | -2.032 |
| 521 | PKP1 | NM_000299 | -1.660 | -2.404 | -2.032 |
| 522 | LOC388780 | XM_373904 | -2.203 | -1.856 | -2.029 |
| 523 | SMCR5 | NM_144774 | -3.608 | -0.450 | -2.029 |
| 524 | HSPB7 | NM_014424 | -1.581 | -2.473 | -2.027 |
| 525 | OLFM1 | NM_014279 | -1.357 | -2.694 | -2.026 |
| 526 | PKLR | NM_000298 | -2.308 | -1.744 | -2.026 |
| 527 | DOC2A | NM_003586 | -2.136 | -1.915 | -2.025 |
| 528 | SERPINB10 | NM_005024 | -1.643 | -2.405 | -2.024 |
| 529 | C10orf26 | NM_017787 | -2.113 | -1.935 | -2.024 |
| 530 | HDX | NM_144657 | -1.351 | -2.695 | -2.023 |
| 531 | ARID4B | NM_016374 | -2.293 | -1.752 | -2.022 |
| 532 | CCDC103 | NM_213607 | -2.383 | -1.662 | -2.022 |
| 533 | SLC2A9 | NM_020041 | -1.786 | -2.257 | -2.021 |
| 534 | C21orf128 | NM_152507 | -2.227 | -1.814 | -2.021 |
| 535 | ARFIP2 | NM_012402 | -1.910 | -2.129 | -2.020 |
| 536 | LPAR5 | NM_020400 | -1.564 | -2.475 | -2.019 |
| 537 | COL4A4 | NM_000092 | -0.985 | -3.053 | -2.019 |
| 538 | FIBIN | NM_203371 | -2.267 | -1.770 | -2.019 |
| 539 | ACADS | NM_000017 | -2.308 | -1.728 | -2.018 |
| 540 | FLJ10404 | NM_019057 | -2.600 | -1.433 | -2.016 |
| 541 | LMO1 | NM_002315 | -2.586 | -1.445 | -2.015 |
| 542 | SMPD1 | NM_000543 | -1.077 | -2.952 | -2.014 |
| 543 | PRDX3 | NM_006793 | -2.970 | -1.059 | -2.014 |
| 544 | FLJ41841 | NM_207499 | -2.664 | -1.362 | -2.013 |
| 545 | GYS1 | NM_002103 | -1.417 | -2.608 | -2.013 |
| 546 | CHST1 | NM_003654 | -2.512 | -1.514 | -2.013 |
| 547 | FAM118B | NM_024556 | -2.023 | -2.001 | -2.012 |
| 548 | LVRN | NM_173800 | -1.921 | -2.103 | -2.012 |
| 549 | ITSN2 | NM_006277 | -1.308 | -2.715 | -2.012 |
| 550 | NELL2 | NM_006159 | -2.143 | -1.879 | -2.011 |
| 551 | PRG3 | NM_006093 | -1.977 | -2.045 | -2.011 |
| 552 | CDH23 | NM_022124 | -2.137 | -1.883 | -2.010 |
| 553 | CHRNE | NM_000080 | -1.807 | -2.209 | -2.008 |
| 554 | CASP2 | NM_001224 | -2.312 | -1.701 | -2.007 |

| | | | | | |
|---|---|---|---|---|---|
| 555 | UBR5 | NM_015902 | -1.648 | -2.364 | -2.006 |
| 556 | ADRBK1 | NM_001619 | -1.492 | -2.519 | -2.006 |
| 557 | YIPF7 | NM_182592 | -1.560 | -2.451 | -2.005 |
| 558 | NT5DC1 | NM_152729 | -1.731 | -2.279 | -2.005 |
| 559 | DCTN2 | NM_006400 | -1.844 | -2.166 | -2.005 |
| 560 | LOC400743 | XM_378843 | -2.699 | -1.310 | -2.005 |
| 561 | SPI1 | NM_003120 | -2.093 | -1.913 | -2.003 |
| 562 | TCEAL1 | NM_004780 | -1.207 | -2.798 | -2.002 |
| 563 | KCTD18 | NM_152387 | -2.040 | -1.963 | -2.002 |
| 564 | MAP2K7 | NM_145185 | -1.961 | -2.040 | -2.001 |
| 565 | LOC401606 | XM_377028 | -1.704 | -2.297 | -2.001 |

**Table 2: Gene list sorted by $N_{av}$ score.**
N1: z-score of nucleic number change for replicate 1, N2: z-score of nuclei number change for replicate 2, Fav: average z-score of the nuclei number change of the duplicates.

| No. | Entrez Gene Symbol | Accession no. | N1 | N2 | Nav |
|---|---|---|---|---|---|
| 1 | HNRNPK | NM_002140 | -5.307 | -8.224 | -6.765 |
| 2 | PSMD8 | NM_002812 | -5.195 | -8.064 | -6.629 |
| 3 | RAN | NM_006325 | -2.372 | -9.362 | -5.867 |
| 4 | RBM8A | NM_005105 | -5.694 | -5.748 | -5.721 |
| 5 | RPL36 | NM_015414 | -7.748 | -3.606 | -5.677 |
| 6 | NAPA | NM_003827 | -5.825 | -5.238 | -5.532 |
| 7 | RPL14 | NM_003973 | -3.706 | -6.885 | -5.296 |
| 8 | SC4MOL | NM_006745 | -5.103 | -5.404 | -5.254 |
| 9 | AMPH | NM_001635 | -5.237 | -5.237 | -5.237 |
| 10 | RPL23 | NM_000978 | -4.356 | -5.884 | -5.120 |
| 11 | PSMB6 | NM_002798 | -4.214 | -5.972 | -5.093 |
| 12 | dJ612B15.1 | XM_370872 | -5.262 | -4.908 | -5.085 |
| 13 | SF3B2 | XM_290506 | -6.891 | -3.234 | -5.063 |
| 14 | SNW1 | NM_012245 | -5.876 | -4.215 | -5.045 |
| 15 | SON | NM_003103 | -5.294 | -4.781 | -5.037 |
| 16 | EIF4A3 | NM_014740 | -4.844 | -5.213 | -5.028 |
| 17 | COPZ1 | NM_016057 | -6.872 | -3.138 | -5.005 |
| 18 | LOC387907 | XM_370713 | -5.168 | -4.750 | -4.959 |
| 19 | XAB2 | NM_020196 | -7.225 | -2.677 | -4.951 |
| 20 | KPNB1 | NM_002265 | -3.811 | -6.061 | -4.936 |
| 21 | PSMD2 | NM_002808 | -5.117 | -4.693 | -4.905 |
| 22 | BCL2L1 | NM_138578 | -3.775 | -5.829 | -4.802 |

| | | | | | |
|---|---|---|---|---|---|
| 23 | ISY1 | NM_020701 | -6.828 | -2.659 | -4.744 |
| 24 | NXF1 | NM_006362 | -5.303 | -4.115 | -4.709 |
| 25 | IK | NM_006083 | -2.883 | -6.349 | -4.616 |
| 26 | SF3B14 | NM_016047 | -6.591 | -2.613 | -4.602 |
| 27 | TUBA1C | NM_032704 | -5.217 | -3.931 | -4.574 |
| 28 | CDC5L | NM_001253 | -5.275 | -3.854 | -4.565 |
| 29 | RPL37 | NM_000997 | -4.932 | -4.120 | -4.526 |
| 30 | hCG_1783417 | XM_376154 | -5.960 | -3.015 | -4.487 |
| 31 | RPL37A | NM_000998 | -3.893 | -4.942 | -4.417 |
| 32 | AQR | NM_014691 | -3.724 | -5.094 | -4.409 |
| 33 | hCG_21078 | XM_371853 | -4.500 | -4.191 | -4.345 |
| 34 | RPL32 | NM_000994 | -4.452 | -4.176 | -4.314 |
| 35 | PRPF8 | NM_006445 | -4.693 | -3.913 | -4.303 |
| 36 | RPL10A | NM_007104 | -4.045 | -4.553 | -4.299 |
| 37 | CCL17 | NM_002987 | -3.339 | -5.237 | -4.288 |
| 38 | RBM22 | NM_018047 | -7.394 | -1.107 | -4.250 |
| 39 | LOC388122 | XM_370865 | -5.164 | -3.311 | -4.238 |
| 40 | SF3A1 | NM_005877 | -4.357 | -4.105 | -4.231 |
| 41 | RPL10 | NM_006013 | -4.306 | -4.145 | -4.226 |
| 42 | RPL7P6 | XM_029805 | -4.200 | -4.225 | -4.213 |
| 43 | RPL14L | XM_056681 | -4.029 | -4.384 | -4.206 |
| 44 | LOC389305 | XM_371757 | -3.825 | -4.583 | -4.204 |
| 45 | RPS16 | NM_001020 | -3.528 | -4.835 | -4.181 |
| 46 | LOC389425 | XM_371843 | -3.260 | -5.083 | -4.172 |
| 47 | POLR2C | NM_002694 | -4.173 | -4.155 | -4.164 |
| 48 | SETD8 | NM_020382 | -4.482 | -3.844 | -4.163 |

| | | | | | |
|---|---|---|---|---|---|
| 49 | TRHR | NM_003301 | -5.320 | -2.998 | -4.159 |
| 50 | LOC388460 | XM_371107 | -4.337 | -3.944 | -4.141 |
| 51 | RPS4X | NM_001007 | -3.738 | -4.516 | -4.127 |
| 52 | CCNA2 | NM_001237 | -3.202 | -5.047 | -4.124 |
| 53 | RPL37P6 | XM_294473 | -4.438 | -3.791 | -4.114 |
| 54 | NUP205 | XM_058073 | -5.285 | -2.835 | -4.060 |
| 55 | C19orf29 | XM_375557 | -5.120 | -2.994 | -4.057 |
| 56 | FMC1 | NM_197964 | -3.587 | -4.525 | -4.056 |
| 57 | FBL | NM_001436 | -2.785 | -5.301 | -4.043 |
| 58 | SF3B3 | NM_012426 | -5.083 | -2.988 | -4.036 |
| 59 | POLR1A | NM_015425 | -5.416 | -2.636 | -4.026 |
| 60 | POLR2D | NM_004805 | -3.577 | -4.418 | -3.998 |
| 61 | SF3B1 | NM_012433 | -5.091 | -2.896 | -3.994 |
| 62 | RBBP9 | NM_006606 | -4.681 | -3.306 | -3.993 |
| 63 | FATE1 | NM_033085 | -5.295 | -2.661 | -3.978 |
| 64 | hCG_1992539 | XM_039218 | -4.011 | -3.836 | -3.924 |
| 65 | GRIP2 | XM_042936 | -5.278 | -2.558 | -3.918 |
| 66 | EEF1A1 | NM_001402 | -2.972 | -4.860 | -3.916 |
| 67 | SFRS3 | NM_003017 | -4.351 | -3.479 | -3.915 |
| 68 | WDR43 | XM_087089 | -5.568 | -2.249 | -3.908 |
| 69 | SETD8 | NM_020382 | -3.377 | -4.423 | -3.900 |
| 70 | RPS3AP6 | XM_039702 | -4.263 | -3.503 | -3.883 |
| 71 | ZNF574 | NM_022752 | -5.249 | -2.451 | -3.850 |
| 72 | CKAP5 | NM_014756 | -3.139 | -4.547 | -3.843 |
| 73 | C14orf177 | NM_182560 | -4.414 | -3.241 | -3.827 |
| 74 | LOC285053 | XM_208281 | -4.146 | -3.503 | -3.824 |

| | | | | | |
|---|---|---|---|---|---|
| 75 | COPB2 | NM_004766 | -3.655 | -3.953 | -3.804 |
| 76 | LOC146053 | XM_016713 | -4.164 | -3.426 | -3.795 |
| 77 | LOC284393 | XM_209178 | -5.349 | -2.233 | -3.791 |
| 78 | LOC285658 | XM_209704 | -5.000 | -2.580 | -3.790 |
| 79 | PRPF19 | NM_014502 | -4.677 | -2.900 | -3.788 |
| 80 | RPS29 | NM_001032 | -3.214 | -4.347 | -3.780 |
| 81 | RRM1 | NM_001033 | -3.425 | -4.132 | -3.778 |
| 82 | SF3A2 | NM_007165 | -4.063 | -3.487 | -3.775 |
| 83 | LOC389342 | XM_371781 | -3.033 | -4.515 | -3.774 |
| 84 | RRM2 | NM_001034 | -2.923 | -4.615 | -3.769 |
| 85 | POLR2A | NM_000937 | -2.705 | -4.823 | -3.764 |
| 86 | LOC340228 | XM_291204 | -4.165 | -3.355 | -3.760 |
| 87 | EIF2B2 | NM_014239 | -5.317 | -2.196 | -3.756 |
| 88 | BDP1 | NM_018429 | -4.232 | -3.252 | -3.742 |
| 89 | PLK1 | NM_005030; NM_005030 | -5.069 | -2.398 | -3.733 |
| 90 | AATF | NM_012138 | -4.721 | -2.725 | -3.723 |
| 91 | RPSA | NM_002295 | -2.796 | -4.623 | -3.709 |
| 92 | LOC388519 | XM_371151 | -3.384 | -4.031 | -3.708 |
| 93 | LOC149329 | XM_086494 | -3.976 | -3.398 | -3.687 |
| 94 | COL20A1 | NM_020882 | -4.250 | -3.074 | -3.662 |
| 95 | TMEM61 | NM_182532 | -3.972 | -3.348 | -3.660 |
| 96 | LOC342994 | XM_292836 | -5.094 | -2.175 | -3.634 |
| 97 | ZNF643 | NM_023070 | -5.241 | -2.027 | -3.634 |
| 98 | SF3B4 | NM_005850 | -3.557 | -3.683 | -3.620 |
| 99 | ERAF | NM_016633 | -4.028 | -3.202 | -3.615 |
| 100 | POLR2F | NM_021974 | -2.818 | -4.409 | -3.613 |

| | | | | | |
|---|---|---|---|---|---|
| 101 | KNCN | NM_182516 | -4.198 | -3.017 | -3.608 |
| 102 | CHD4 | NM_001273 | -4.372 | -2.811 | -3.592 |
| 103 | P2RX6 | NM_005446 | -3.475 | -3.687 | -3.581 |
| 104 | MGC50273 | NM_214461 | -3.824 | -3.332 | -3.578 |
| 105 | RPS19 | NM_001022 | -3.416 | -3.730 | -3.573 |
| 106 | U2AF1 | NM_006758 | -3.385 | -3.732 | -3.558 |
| 107 | RPS5 | NM_001009 | -2.266 | -4.825 | -3.546 |
| 108 | U2AF2 | NM_007279 | -4.928 | -2.148 | -3.538 |
| 109 | DHRS13 | NM_144683 | -4.828 | -2.248 | -3.538 |
| 110 | CCT6A | NM_001762 | -4.719 | -2.349 | -3.534 |
| 111 | NHP2L1 | NM_005008 | -2.218 | -4.847 | -3.533 |
| 112 | ABCE1 | NM_002940 | -3.031 | -4.022 | -3.527 |
| 113 | DHX8 | NM_004941 | -2.499 | -4.553 | -3.526 |
| 114 | hCG_1984468 | XM_372048 | -3.137 | -3.913 | -3.525 |
| 115 | EIF2S2 | NM_003908 | -2.806 | -4.226 | -3.516 |
| 116 | POLR2B | NM_000938 | -2.710 | -4.322 | -3.516 |
| 117 | CNGA2 | NM_005140 | -3.514 | -3.514 | -3.514 |
| 118 | ISCA2 | NM_194279 | -4.785 | -2.188 | -3.486 |
| 119 | PSMD6 | NM_014814 | -3.628 | -3.321 | -3.475 |
| 120 | CASP8AP2 | NM_012115 | -3.330 | -3.600 | -3.465 |
| 121 | RP11-365K22.1 | XM_370727 | -4.239 | -2.677 | -3.458 |
| 122 | CCT7 | NM_006429 | -4.625 | -2.274 | -3.449 |
| 123 | LOC130773 | XM_065899 | -4.825 | -2.007 | -3.416 |
| 124 | LOC389748 | XM_372108 | -3.045 | -3.785 | -3.415 |
| 125 | LOC400652 | XM_375543 | -3.593 | -3.222 | -3.407 |
| 126 | GABPA | NM_002040 | -3.387 | -3.427 | -3.407 |

| | | | | | |
|---|---|---|---|---|---|
| 127 | C8orf31 | NM_173687 | -3.490 | -3.314 | -3.402 |
| 128 | MAGOH | NM_002370 | -2.706 | -4.095 | -3.401 |
| 129 | RPL13A | NM_012423 | -4.021 | -2.773 | -3.397 |
| 130 | UBC | NM_021009 | -3.619 | -3.171 | -3.395 |
| 131 | CCT2 | NM_006431 | -4.224 | -2.552 | -3.388 |
| 132 | CDC2 | NM_001786 | -3.603 | -3.161 | -3.382 |
| 133 | LOC388720 | XM_371330 | -3.159 | -3.604 | -3.382 |
| 134 | HNRNPC | NM_004500 | -4.033 | -2.709 | -3.371 |
| 135 | CDC27 | NM_001256 | -3.270 | -3.468 | -3.369 |
| 136 | NDE1 | NM_017668 | -3.500 | -3.193 | -3.347 |
| 137 | LOC388474 | XM_371115 | -3.325 | -3.356 | -3.340 |
| 138 | GPS1 | NM_004127 | -3.428 | -3.192 | -3.310 |
| 139 | HIVEP3 | NM_024503 | -2.625 | -3.986 | -3.306 |
| 140 | LOC343153 | XM_291428 | -3.787 | -2.817 | -3.302 |
| 141 | FLJ21075 | NM_025031 | -3.448 | -3.135 | -3.292 |
| 142 | LOC388532 | XM_371160 | -3.360 | -3.212 | -3.286 |
| 143 | WDR46 | NM_005452 | -2.771 | -3.794 | -3.283 |
| 144 | POLR2J2 | NM_032958 | -4.404 | -2.148 | -3.276 |
| 145 | RPL35 | NM_007209 | -3.721 | -2.820 | -3.270 |
| 146 | RPS3 | NM_001005 | -1.907 | -4.622 | -3.264 |
| 147 | ESPL1 | NM_012291 | -3.473 | -3.044 | -3.258 |
| 148 | CSE1L | NM_001316 | -2.840 | -3.665 | -3.252 |
| 149 | NBPF4 | NM_152488 | -3.579 | -2.920 | -3.249 |
| 150 | EFTUD2 | NM_004247 | -3.195 | -3.303 | -3.249 |
| 151 | SLC22A16 | NM_033125 | -3.551 | -2.945 | -3.248 |
| 152 | TBL3 | NM_006453 | -4.087 | -2.399 | -3.243 |

| | | | | | |
|---|---|---|---|---|---|
| 153 | KIAA1604 | XM_034594 | -3.639 | -2.847 | -3.243 |
| 154 | WDR78 | NM_024763 | -3.956 | -2.529 | -3.242 |
| 155 | LACTB | NM_032857 | -3.901 | -2.571 | -3.236 |
| 156 | BOP1 | NM_015201 | -3.970 | -2.480 | -3.225 |
| 157 | SF3A3 | NM_006802 | -3.239 | -3.198 | -3.218 |
| 158 | MYBL2 | NM_002466 | -4.100 | -2.322 | -3.211 |
| 159 | LOC158345 | XM_034640 | -3.324 | -3.093 | -3.208 |
| 160 | HCFC1 | NM_005334 | -4.374 | -2.042 | -3.208 |
| 161 | CDC42 | NM_001791 | -3.267 | -3.116 | -3.191 |
| 162 | RPS21 | NM_001024 | -1.894 | -4.473 | -3.183 |
| 163 | C12orf41 | NM_017822 | -5.237 | -1.097 | -3.167 |
| 164 | LOC392208 | XM_373246 | -3.349 | -2.952 | -3.150 |
| 165 | UBA52 | NM_003333 | -2.169 | -4.124 | -3.147 |
| 166 | hCG_26523 | XM_374987 | -3.732 | -2.556 | -3.144 |
| 167 | RPS2 | NM_002952 | -2.025 | -4.260 | -3.142 |
| 168 | RPL11 | NM_000975 | -2.838 | -3.436 | -3.137 |
| 169 | INCENP | NM_020238 | -2.611 | -3.660 | -3.136 |
| 170 | KIF23 | NM_004856 | -4.243 | -2.025 | -3.134 |
| 171 | POLA2 | NM_002689 | -4.554 | -1.714 | -3.134 |
| 172 | C16orf54 | NM_175900 | -3.189 | -3.078 | -3.134 |
| 173 | LOC149224 | XM_015717 | -3.233 | -3.031 | -3.132 |
| 174 | PES1 | NM_014303 | -3.170 | -3.085 | -3.127 |
| 175 | RPL23AP9 | XM_063202 | -3.435 | -2.788 | -3.111 |
| 176 | OR1A1 | NM_014565 | -2.919 | -3.296 | -3.107 |
| 177 | LCE3A | NM_178431 | -3.150 | -3.037 | -3.094 |
| 178 | hCG_1644323 | XM_373343 | -2.923 | -3.201 | -3.062 |

| | | | | | |
|---|---|---|---|---|---|
| 179 | CAPRIN1 | NM_005898 | -3.586 | -2.533 | -3.060 |
| 180 | NDUFB7 | NM_004146 | -3.095 | -3.020 | -3.057 |
| 181 | CCDC103 | NM_213607 | -3.203 | -2.896 | -3.050 |
| 182 | WAC | NM_016628 | -4.357 | -1.737 | -3.047 |
| 183 | MFAP1 | NM_005926 | -3.948 | -2.144 | -3.046 |
| 184 | DNAH9 | NM_001372 | -3.037 | -3.037 | -3.037 |
| 185 | C12orf57 | NM_138425 | -2.373 | -3.698 | -3.036 |
| 186 | DMAP1 | NM_019100 | -3.972 | -2.092 | -3.032 |
| 187 | COPS5 | NM_006837 | -1.933 | -4.107 | -3.020 |
| 188 | SPSB3 | NM_080861 | -3.532 | -2.504 | -3.018 |
| 189 | RAD51 | NM_002875 | -4.266 | -1.767 | -3.016 |
| 190 | RPS9 | NM_001013 | -3.376 | -2.651 | -3.014 |
| 191 | OR1D2 | NM_002548 | -3.227 | -2.794 | -3.010 |
| 192 | GLB1L2 | NM_138342 | -3.867 | -2.140 | -3.003 |
| 193 | PSMD1 | NM_002807 | -3.401 | -2.574 | -2.988 |
| 194 | LOC389814 | XM_372160 | -2.093 | -3.881 | -2.987 |
| 195 | FAM87A | XM_379540 | -2.871 | -3.095 | -2.983 |
| 196 | LOC388344 | XM_371023 | -2.591 | -3.333 | -2.962 |
| 197 | SNX30 | XM_376902 | -3.439 | -2.463 | -2.951 |
| 198 | SFPQ | NM_005066 | -2.861 | -3.041 | -2.951 |
| 199 | NOP2 | NM_006170 | -1.949 | -3.929 | -2.939 |

**Table 3: Reactome Analysis**

| Categories | Matching identifiers |
|---|---|
| Cell attachement and migration | CDC42, RRAS, COL4A4, RHOA, CLASP1 |
| Cell cycle | AHCTF1, NUP107, TUBB4, PSMD2, ODF2, CENPA, DCTN2, APITD1, CLASP1, CENPQ, RAD9B, RAD17 |
| Hormone synthesis | GC, CYP21A2, LTA4H ,CYP11B2 |
| Mediator complex | MED14, MED12, MED24, MED17, MED13 |
| Metabolism of carbohydrates | NUP107, PKLR, GBE1, TPR, TALDO1, GYS1 |
| Metabolism of lipids | CPT2, ACADS, AGPS, ABCG8, CYP21A2, HADHA, CYP11B2, PLIN2 |
| Metabolism of proteins | EIF2B4,TUBB4, ETF1, CCT7, TCP1, EIF2B2, EIF2S2, EIF2B1, EIF2B3 |
| Polymerase II transcription | TAF12, SFRS3, GTF2H3, CTDP1, POLR2E, CPSF3, PCF11, NCBP1 |
| Splicing complex | HNRNPU, SFRS3, SF3A3, POLR2E, CPSF3, PCF11, NCBP1, SF3A1, NXF1, HNRNPD, GTF2H3 |
| Translation initiation | EIF2S2, EIF2B2, EIF2B1, EIF2B4, EIF2B3 |
| Transmembrane transport | NUP107, SLC2A9, SLC16A3, TPR, SLC2A12 |
| tRNA metabolism | LARS, DARS, VARS |

**Table 4a: H1 hESCs secondary screen data (relative GFP expression)**
Deconvoluted siRNA screen data for the 200 genes

| No. | Gene symbol | GFP_S1 | GFP_S2 | GFP_S3 | GFP_S4 |
|---|---|---|---|---|---|
| 1 | ABP1 | 0.226 | 0.357 | 0.360 | 0.266 |
| 2 | ABTB1 | 0.431 | 0.653 | 0.525 | 0.539 |
| 3 | ADA | 0.415 | 0.339 | 0.386 | 0.431 |
| 4 | ADAMTS1 | 0.548 | 0.362 | 0.365 | 0.578 |
| 5 | AGPS | 0.820 | 0.641 | 0.782 | 0.931 |
| 6 | AIPL1 | 0.700 | 0.559 | 0.713 | 0.735 |
| 7 | ANGPT4 | 0.772 | 0.439 | 0.304 | 0.247 |
| 8 | ANKRD1 | 0.767 | 0.612 | 0.704 | 0.672 |
| 9 | ANKRD31 | 0.765 | 0.804 | 0.487 | 0.671 |
| 10 | ANXA4 | 0.332 | 0.283 | 0.573 | 0.223 |
| 11 | APLP2 | 0.317 | 0.407 | 0.424 | 0.516 |
| 12 | ATOH8 | 0.438 | 0.709 | 0.834 | 0.639 |
| 13 | BCL6B | 1.017 | 0.765 | 1.128 | 1.326 |
| 14 | BDP1 | 0.306 | 0.470 | 0.386 | 0.608 |
| 15 | BENE | 0.612 | 0.917 | 0.845 | 0.554 |
| 16 | C22ORF16 | 0.563 | 0.308 | 0.355 | 0.429 |
| 17 | CAMP | 0.694 | 0.729 | 0.857 | 0.971 |
| 18 | CAPN2 | 0.285 | 0.329 | 0.344 | 0.478 |
| 19 | CCL2 | 0.700 | 1.024 | 0.931 | 0.860 |
| 20 | CDC42 | 0.639 | 0.506 | 0.593 | 0.559 |
| 21 | CDX2 | 0.510 | 0.755 | 0.744 | 0.741 |
| 22 | CGGBP1 | 0.558 | 0.794 | 0.639 | 0.416 |
| 23 | COL11A1 | 0.530 | 0.463 | 0.197 | 0.266 |

| | | | | | |
|---|---|---|---|---|---|
| 24 | COPS4 | 0.512 | 0.531 | 0.440 | 0.495 |
| 25 | CORT | 0.761 | 0.847 | 0.934 | 0.810 |
| 26 | CPSF3 | 0.500 | 0.341 | 0.318 | 0.698 |
| 27 | CREBL2 | 0.331 | 0.169 | 0.864 | 0.837 |
| 28 | CRK7 | 0.417 | 0.534 | 0.447 | 0.614 |
| 29 | CRLF1 | 0.382 | 0.429 | 0.264 | 0.471 |
| 30 | CRSP2 | 0.379 | 0.315 | 0.336 | 0.491 |
| 31 | CTNNA3 | 0.584 | 0.776 | 0.831 | 0.843 |
| 32 | CYBA | 0.465 | 0.575 | 0.360 | 0.472 |
| 33 | DDEF1 | 0.914 | 0.546 | 0.843 | 0.681 |
| 34 | DDIT3 | 0.790 | 0.466 | 1.055 | 0.799 |
| 35 | DEFB126 | 0.390 | 0.336 | 0.774 | 0.557 |
| 36 | DKFZP564B147 | 0.529 | 0.660 | 0.694 | 0.649 |
| 37 | DKFZP564K142 | 0.309 | 0.407 | 0.368 | 0.330 |
| 38 | DRG2 | 0.713 | 0.532 | 0.796 | 0.711 |
| 39 | E4F1 | 0.599 | 0.438 | 0.589 | 1.015 |
| 40 | EDF1 | 0.531 | 0.388 | 0.800 | 0.288 |
| 41 | EIF2B1 | 0.421 | 0.440 | 0.851 | 0.712 |
| 42 | EIF2B2 | 0.237 | 0.203 | 0.210 | 0.230 |
| 43 | EIF2B3 | 0.160 | 0.160 | 0.112 | 0.201 |
| 44 | EIF2B4 | 0.239 | 0.101 | 0.441 | 0.174 |
| 45 | EIF2S2 | 0.177 | 0.152 | 0.150 | 0.105 |
| 46 | ELYS | 0.525 | 0.553 | 0.738 | 0.648 |
| 47 | ENPP7 | 0.294 | 0.186 | 0.253 | 0.246 |
| 48 | EP300 | 0.712 | 0.854 | 0.649 | 0.711 |
| 49 | ETF1 | 0.124 | 0.265 | 0.168 | 0.086 |

| | | | | | |
|---|---|---|---|---|---|
| 50 | FAM19A1 | 0.308 | 0.419 | 0.155 | 0.432 |
| 51 | FLJ20898 | 0.392 | 0.464 | 0.507 | 0.397 |
| 52 | FLJ23447 | 0.481 | 0.245 | 0.391 | 0.397 |
| 53 | FLJ23751 | 0.585 | 0.741 | 0.876 | 0.797 |
| 54 | FLJ25439 | 0.378 | 0.453 | 0.167 | 0.292 |
| 55 | FLJ25952 | 0.692 | 0.827 | 0.604 | 0.684 |
| 56 | FLJ32954 | 0.189 | 0.258 | 0.616 | 0.345 |
| 57 | FLJ38508 | 0.861 | 0.571 | 0.898 | 0.712 |
| 58 | FLJ39743 | 0.734 | 0.850 | 0.779 | 0.661 |
| 59 | FLJ46536 | 0.305 | 0.173 | 0.465 | 0.170 |
| 60 | FLJ90652 | 0.856 | 0.697 | 0.516 | 0.873 |
| 61 | FOXJ3 | 0.792 | 0.831 | 0.649 | 0.871 |
| 62 | FTL | 0.807 | 0.846 | 0.784 | 0.872 |
| 63 | FTSJ1 | 0.297 | 0.425 | 0.209 | 0.540 |
| 64 | FUBP1 | 0.608 | 0.672 | 0.629 | 0.571 |
| 65 | GJA8 | 0.273 | 0.382 | 0.248 | 0.348 |
| 66 | GLRB | 0.335 | 0.428 | 0.291 | 0.613 |
| 67 | GLTSCR1 | 0.411 | 0.440 | 0.355 | 0.352 |
| 68 | GPS1 | 0.509 | 0.558 | 0.484 | 0.380 |
| 69 | GSPT1 | 0.640 | 0.609 | 0.555 | 0.660 |
| 70 | GSTP1 | 0.763 | 0.731 | 0.722 | 1.003 |
| 71 | GUSB | 0.363 | 0.586 | 0.357 | 0.244 |
| 72 | H1FX | 0.735 | 0.473 | 0.257 | 0.622 |
| 73 | HCFC1 | 0.153 | 0.230 | 0.122 | 0.219 |
| 74 | HELZ | 0.335 | 0.343 | 0.315 | 0.329 |
| 75 | HEMK1 | 0.319 | 0.476 | 0.292 | 0.200 |

| | | | | | |
|---|---|---|---|---|---|
| 76 | HES6 | 0.488 | 0.519 | 0.588 | 0.375 |
| 77 | HIVEP3 | 0.894 | 0.643 | 0.605 | 0.749 |
| 78 | HNRPD | 0.879 | 0.934 | 0.817 | 0.664 |
| 79 | HNRPU | 0.118 | 0.323 | 0.225 | 0.290 |
| 80 | IBSP | 0.664 | 0.947 | 0.722 | 0.750 |
| 81 | IGFBP6 | 0.457 | 0.575 | 0.291 | 0.699 |
| 82 | INCA1 | 0.164 | 0.348 | 0.423 | 0.249 |
| 83 | ITSN2 | 0.521 | 0.779 | 0.932 | 0.684 |
| 84 | JMJD2B | 0.966 | 1.126 | 0.292 | 0.562 |
| 85 | KIAA0274 | 0.679 | 0.737 | 0.763 | 0.814 |
| 86 | KIAA1076 | 0.673 | 0.715 | 0.964 | 0.885 |
| 87 | KIR3DL1 | 0.318 | 0.521 | 0.419 | 0.335 |
| 88 | KIRREL2 | 0.409 | 0.302 | 0.535 | 0.284 |
| 89 | KLK5 | 0.723 | 0.862 | 0.810 | 0.645 |
| 90 | KREMEN1 | 0.805 | 0.840 | 0.965 | 0.816 |
| 91 | LARS | 0.352 | 0.440 | 0.576 | 0.173 |
| 92 | LCE1E | 0.608 | 0.770 | 0.761 | 0.303 |
| 93 | LCMR1 | 0.178 | 0.375 | 0.287 | 0.251 |
| 94 | LIF | 0.411 | 0.230 | 0.201 | 0.035 |
| 95 | LOC124245 | 0.370 | 0.604 | 0.283 | 0.307 |
| 96 | LOC374654 | 0.679 | 0.362 | 0.422 | 0.444 |
| 97 | LOC390790 | 0.217 | 0.076 | 0.322 | 0.307 |
| 98 | LOC400221 | 0.458 | 0.329 | 0.504 | 0.313 |
| 99 | LOC56901 | 0.340 | 0.405 | 0.328 | 0.494 |
| 100 | LPPR2 | 0.847 | 0.633 | 0.560 | 0.697 |
| 101 | LRRC33 | 0.258 | 0.117 | 0.251 | 0.301 |

| | | | | | |
|---|---|---|---|---|---|
| 102 | LUC7A | 0.289 | 0.349 | 0.303 | 0.357 |
| 103 | MAP2K7 | 0.819 | 0.755 | 0.693 | 1.029 |
| 104 | MAP3K1 | 0.975 | 0.789 | 1.021 | 0.606 |
| 105 | MBTD1 | 0.445 | 0.462 | 0.341 | 0.418 |
| 106 | MCRS1 | 0.155 | 0.372 | 0.375 | 0.658 |
| 107 | MED28 | 0.634 | 0.678 | 0.750 | 0.687 |
| 108 | MGC10471 | 0.287 | 0.340 | 0.295 | 0.343 |
| 109 | MGC21874 | 0.343 | 0.616 | 0.952 | 0.508 |
| 110 | MGC32871 | 0.270 | 0.287 | 0.269 | 0.351 |
| 111 | MGC39827 | 0.227 | 0.132 | 0.227 | 0.313 |
| 112 | MGC8902 | 1.051 | 0.762 | 0.849 | 0.782 |
| 113 | MMP15 | 0.721 | 0.537 | 0.611 | 0.481 |
| 114 | MMP24 | 0.554 | 0.527 | 0.836 | 0.915 |
| 115 | MOCS1 | 0.507 | 0.354 | 0.406 | 0.256 |
| 116 | MR1 | 0.387 | 0.197 | 0.168 | 0.346 |
| 117 | MVP | 0.387 | 0.143 | 0.380 | 0.643 |
| 118 | NANOG | 0.281 | 0.360 | 0.482 | 0.928 |
| 119 | NCBP1 | 0.151 | 0.252 | 0.261 | 0.180 |
| 120 | NEUROD2 | 0.333 | 0.458 | 0.369 | 0.333 |
| 121 | NFKB1 | 0.659 | 0.844 | 0.933 | 1.020 |
| 122 | NFRKB | 0.559 | 0.657 | 0.374 | 0.361 |
| 123 | NPEPL1 | 0.325 | 0.352 | 0.431 | 0.334 |
| 124 | NUDT8 | 0.250 | 0.273 | 0.218 | 0.461 |
| 125 | NUP107 | 0.516 | 0.491 | 0.844 | 0.183 |
| 126 | NXF1 | 0.127 | 0.068 | 0.586 | 0.145 |
| 127 | OACT1 | 0.854 | 0.759 | 0.815 | 0.849 |

| | | | | | |
|---|---|---|---|---|---|
| 128 | ODF2 | 0.135 | 0.316 | 0.365 | 0.247 |
| 129 | P29 | 0.630 | 0.889 | 0.764 | 0.826 |
| 130 | PCF11 | 0.049 | 0.153 | 0.222 | 0.114 |
| 131 | PDZK11 | 0.518 | 0.777 | 0.717 | 0.581 |
| 132 | PHB | 0.129 | 0.079 | 0.202 | 0.209 |
| 133 | PITX1 | 0.689 | 0.970 | 0.688 | 0.949 |
| 134 | PKP1 | 0.715 | 0.761 | 0.809 | 0.764 |
| 135 | POLH | 0.295 | 0.253 | 0.535 | 0.351 |
| 136 | OCT4 | 0.128 | 0.245 | 0.086 | 0.116 |
| 137 | PPAPDC2 | 0.256 | 0.192 | 0.204 | 0.398 |
| 138 | PPP2R3A | 0.583 | 0.370 | 0.551 | 0.666 |
| 139 | PRDM14 | 0.728 | 0.341 | 0.832 | 0.191 |
| 140 | PRDM9 | 0.536 | 0.131 | 0.697 | 0.968 |
| 141 | PRDX6 | 0.841 | 0.982 | 0.639 | 0.988 |
| 142 | PRO2730 | 0.405 | 0.263 | 0.682 | 0.857 |
| 143 | PROP1 | 0.390 | 0.529 | 0.416 | 0.543 |
| 144 | PSMD2 | 0.042 | 0.207 | 0.114 | 0.027 |
| 145 | PSTPIP2 | 0.671 | 0.800 | 0.709 | 0.341 |
| 146 | PXN | 0.309 | 0.420 | 0.338 | 0.162 |
| 147 | RALGDS | 0.900 | 0.796 | 0.734 | 0.691 |
| 148 | RASEF | 0.449 | 0.584 | 0.554 | 0.698 |
| 149 | RBM17 | 0.752 | 0.677 | 0.745 | 0.580 |
| 150 | REA | 0.259 | 0.210 | 0.388 | 0.127 |
| 151 | RHOA | 0.753 | 0.510 | 0.289 | 0.534 |
| 152 | RICTOR | 0.683 | 0.783 | 0.709 | 0.755 |
| 153 | RPESP | 0.292 | 0.312 | 0.253 | 0.224 |

| | | | | | |
|---|---|---|---|---|---|
| 154 | RRAS | 0.749 | 0.692 | 0.888 | 0.707 |
| 155 | SAMD7 | 0.428 | 0.346 | 0.203 | 0.476 |
| 156 | SERPINB2 | 0.263 | 0.157 | 0.117 | 0.123 |
| 157 | SERTAD2 | 0.758 | 0.754 | 0.837 | 0.903 |
| 158 | SF3A1 | 0.160 | 0.055 | 0.325 | 0.124 |
| 159 | SF3A3 | 0.967 | 0.064 | 0.245 | 0.755 |
| 160 | SFPQ | 0.502 | 0.288 | 0.345 | 0.177 |
| 161 | SFRS3 | 0.479 | 0.123 | 0.320 | 0.371 |
| 162 | SFXN3 | 0.370 | 0.425 | 0.385 | 0.400 |
| 163 | SOAT2 | 0.292 | 0.497 | 0.337 | 0.466 |
| 164 | SON | 0.017 | 0.021 | 0.039 | 0.004 |
| 165 | SOX14 | 0.792 | 0.608 | 0.382 | 0.644 |
| 166 | SPI1 | 0.726 | 0.742 | 0.553 | 0.752 |
| 167 | STK4 | 0.942 | 0.955 | 0.948 | 0.762 |
| 168 | SUV39H2 | 0.845 | 0.869 | 0.382 | 0.271 |
| 169 | SYNCRIP | 0.713 | 0.843 | 0.413 | 0.808 |
| 170 | SYTL4 | 0.761 | 0.842 | 0.468 | 0.786 |
| 171 | TAF2 | 0.529 | 0.540 | 0.843 | 0.869 |
| 172 | TAF7 | 0.741 | 0.980 | 0.640 | 0.950 |
| 173 | TBC1D10 | 0.463 | 0.247 | 0.398 | 0.378 |
| 174 | TCL1A | 0.543 | 0.763 | 0.918 | 0.884 |
| 175 | THRAP2 | 0.674 | 0.642 | 0.563 | 0.879 |
| 176 | THRAP4 | 0.902 | 0.729 | 0.981 | 0.856 |
| 177 | TMEM14B | 0.351 | 0.156 | 0.419 | 0.392 |
| 178 | TNRC11 | 0.634 | 0.919 | 0.580 | 0.504 |
| 179 | TPD52L1 | 0.350 | 0.240 | 0.684 | 0.312 |

| | | | | | |
|---|---|---|---|---|---|
| 180 | TPM1 | 0.215 | 0.271 | 0.361 | 0.382 |
| 181 | TPR | 0.284 | 0.259 | 0.315 | 0.145 |
| 182 | TRIP | 0.424 | 0.377 | 0.762 | 0.746 |
| 183 | TRIP15 | 0.699 | 0.773 | 0.436 | 0.567 |
| 184 | TRPA1 | 0.272 | 0.291 | 0.206 | 0.248 |
| 185 | ULK2 | 0.773 | 0.643 | 0.917 | 0.765 |
| 186 | VGCNL1 | 0.649 | 0.842 | 0.759 | 0.839 |
| 187 | VMP | 0.368 | 0.400 | 0.282 | 0.354 |
| 188 | VWF | 0.716 | 0.731 | 0.828 | 0.762 |
| 189 | XRCC1 | 0.904 | 0.744 | 0.751 | 1.009 |
| 190 | YAP1 | 0.647 | 0.692 | 0.368 | 0.663 |
| 191 | YY1 | 0.594 | 0.521 | 0.698 | 0.623 |
| 192 | ZFP36 | 0.199 | 0.769 | 0.518 | 0.389 |
| 193 | ZFP64 | 0.701 | 0.656 | 0.225 | 0.853 |
| 194 | ZIC4 | 0.640 | 0.789 | 0.278 | 0.273 |
| 195 | ZNF136 | 0.966 | 0.642 | 0.689 | 0.624 |
| 196 | ZNF138 | 0.801 | 0.947 | 0.651 | 0.521 |
| 197 | ZNF206 | 0.383 | 0.250 | 0.528 | 0.186 |
| 198 | ZNF35 | 0.638 | 0.856 | 0.608 | 0.740 |
| 199 | ZNF43 | 0.858 | 0.527 | 0.771 | 1.165 |
| 200 | ZNF434 | 0.863 | 0.163 | 0.414 | 0.814 |

**H1 hESCs secondary screen data (relative NANOG and OCT4 expression)**
Deconvoluted siRNA screen data for the 200 genes

| No. | Gene symbol | NANOG_S1 | NANOG_S2 | NANOG_S3 | NANOG_S4 | OCT4_S1 | OCT4_S2 | OCT4_S3 | OCT4_S4 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | ABP1 | 0.296 | 0.488 | 0.502 | 0.395 | 0.308 | 0.561 | 0.478 | 0.396 |
| 2 | ABTB1 | 0.608 | 0.846 | 0.697 | 0.690 | 0.566 | 0.822 | 0.586 | 0.623 |
| 3 | ADA | 0.523 | 0.424 | 0.592 | 0.559 | 0.740 | 0.310 | 0.492 | 0.511 |
| 4 | ADAMTS1 | 0.712 | 0.519 | 0.449 | 0.672 | 0.539 | 0.266 | 0.307 | 0.671 |
| 5 | AGPS | 0.870 | 0.766 | 0.851 | 0.963 | 0.768 | 0.674 | 0.746 | 0.836 |
| 6 | AIPL1 | 0.824 | 0.718 | 0.868 | 0.857 | 0.679 | 0.587 | 0.754 | 0.681 |
| 7 | ANGPT4 | 0.805 | 0.576 | 0.451 | 0.368 | 1.072 | 0.461 | 0.245 | 0.217 |
| 8 | ANKRD1 | 0.912 | 0.807 | 0.780 | 0.832 | 0.879 | 0.722 | 0.620 | 0.672 |
| 9 | ANKRD31 | 0.903 | 0.845 | 0.775 | 0.783 | 0.733 | 0.764 | 0.587 | 0.612 |
| 10 | ANXA4 | 0.387 | 0.397 | 0.708 | 0.273 | 0.442 | 0.349 | 0.726 | 0.123 |
| 11 | APLP2 | 0.481 | 0.495 | 0.486 | 0.692 | 0.369 | 0.422 | 0.368 | 0.577 |
| 12 | ATOH8 | 0.490 | 0.800 | 1.080 | 0.849 | 0.506 | 0.629 | 0.587 | 0.612 |
| 13 | BCL6B | 1.086 | 1.071 | 1.077 | 1.114 | 0.766 | 0.998 | 0.986 | 0.890 |
| 14 | BDP1 | 0.344 | 0.593 | 0.526 | 0.648 | 0.227 | 0.427 | 0.359 | 0.489 |
| 15 | BENE | 0.716 | 0.969 | 0.970 | 0.623 | 0.554 | 0.978 | 0.904 | 0.501 |
| 16 | C22ORF16 | 0.701 | 0.414 | 0.430 | 0.623 | 0.428 | 0.189 | 0.343 | 0.398 |
| 17 | CAMP | 0.896 | 0.858 | 0.986 | 1.039 | 0.691 | 0.686 | 0.847 | 0.878 |
| 18 | CAPN2 | 0.462 | 0.577 | 0.432 | 0.551 | 0.410 | 0.195 | 0.278 | 0.450 |
| 19 | CCL2 | 0.889 | 1.032 | 0.935 | 0.955 | 0.714 | 0.984 | 0.975 | 0.874 |
| 20 | CDC42 | 0.804 | 0.669 | 0.731 | 0.724 | 0.666 | 0.506 | 0.738 | 0.618 |
| 21 | CDX2 | 0.703 | 0.850 | 0.838 | 0.800 | 0.511 | 0.756 | 0.656 | 0.684 |
| 22 | CGGBP1 | 0.836 | 0.841 | 0.677 | 0.657 | 0.738 | 0.784 | 0.686 | 0.382 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 23 | COL11A1 | 0.763 | 0.580 | 0.291 | 0.364 | 0.492 | 0.482 | 0.215 | 0.331 |
| 24 | COPS4 | 0.812 | 0.878 | 0.694 | 0.740 | 0.413 | 0.402 | 0.462 | 0.484 |
| 25 | CORT | 0.923 | 0.923 | 0.949 | 0.899 | 0.838 | 0.775 | 0.935 | 0.859 |
| 26 | CPSF3 | 0.657 | 0.417 | 0.431 | 0.970 | 0.553 | 0.575 | 0.514 | 0.669 |
| 27 | CREBL2 | 0.495 | 0.426 | 0.909 | 1.049 | 0.265 | 0.044 | 0.734 | 0.843 |
| 28 | CRK7 | 0.725 | 0.940 | 0.787 | 0.925 | 0.419 | 0.631 | 0.551 | 0.622 |
| 29 | CRLF1 | 0.555 | 0.565 | 0.368 | 0.595 | 0.315 | 0.363 | 0.289 | 0.530 |
| 30 | CRSP2 | 0.613 | 0.402 | 0.533 | 0.549 | 0.348 | 0.391 | 0.325 | 0.401 |
| 31 | CTNNA3 | 0.771 | 0.923 | 1.004 | 0.882 | 0.694 | 0.858 | 0.916 | 0.820 |
| 32 | CYBA | 0.592 | 0.718 | 0.486 | 0.614 | 0.409 | 0.618 | 0.325 | 0.388 |
| 33 | DDEF1 | 0.867 | 0.761 | 0.771 | 0.803 | 0.829 | 0.678 | 0.756 | 0.729 |
| 34 | DDIT3 | 1.029 | 0.714 | 0.991 | 0.928 | 0.487 | 0.429 | 0.895 | 0.746 |
| 35 | DEFB126 | 0.546 | 0.494 | 0.849 | 0.683 | 0.385 | 0.306 | 0.728 | 0.541 |
| 36 | DKFZP564 B147 | 0.737 | 0.956 | 0.834 | 0.738 | 0.588 | 0.874 | 0.660 | 0.596 |
| 37 | DKFZP564 K142 | 0.455 | 0.526 | 0.638 | 0.458 | 0.492 | 0.607 | 0.856 | 0.452 |
| 38 | DRG2 | 0.783 | 0.727 | 0.838 | 0.967 | 0.717 | 0.661 | 0.788 | 0.588 |
| 39 | E4F1 | 0.627 | 0.769 | 0.583 | 0.936 | 0.602 | 0.405 | 0.703 | 0.913 |
| 40 | EDF1 | 0.768 | 0.715 | 1.031 | 0.474 | 0.438 | 0.368 | 0.630 | 0.231 |
| 41 | EIF2B1 | 0.616 | 0.737 | 0.877 | 0.880 | 0.462 | 0.637 | 0.716 | 0.799 |
| 42 | EIF2B2 | 0.279 | 0.285 | 0.289 | 0.343 | 0.058 | 0.203 | 0.159 | 0.188 |
| 43 | EIF2B3 | 0.230 | 0.253 | 0.215 | 0.314 | 0.362 | 0.047 | 0.069 | 0.124 |
| 44 | EIF2B4 | 0.311 | 0.153 | 0.578 | 0.254 | 0.254 | 0.065 | 0.256 | 0.093 |
| 45 | EIF2S2 | 0.241 | 0.211 | 0.192 | 0.135 | 0.031 | 0.122 | 0.143 | 0.062 |
| 46 | ELYS | 0.643 | 0.593 | 0.801 | 0.695 | 0.577 | 0.472 | 0.684 | 0.608 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 47 | ENPP7 | 0.390 | 0.273 | 0.301 | 0.316 | 0.422 | 0.374 | 0.449 | 0.339 |
| 48 | EP300 | 0.840 | 0.895 | 0.789 | 0.884 | 0.873 | 0.902 | 0.731 | 0.783 |
| 49 | ETF1 | 0.182 | 0.324 | 0.262 | 0.134 | 0.040 | 0.466 | 0.092 | 0.054 |
| 50 | FAM19A1 | 0.548 | 0.660 | 0.278 | 0.606 | 0.384 | 0.519 | 0.176 | 0.501 |
| 51 | FLJ20898 | 0.498 | 0.549 | 0.619 | 0.625 | 0.454 | 0.716 | 0.560 | 0.416 |
| 52 | FLJ23447 | 0.612 | 0.411 | 0.583 | 0.545 | 0.615 | 0.299 | 0.350 | 0.278 |
| 53 | FLJ23751 | 0.809 | 0.980 | 0.962 | 1.005 | 0.718 | 0.833 | 0.904 | 0.916 |
| 54 | FLJ25439 | 0.472 | 0.473 | 0.248 | 0.409 | 0.705 | 1.016 | 0.253 | 0.551 |
| 55 | FLJ25952 | 0.921 | 0.890 | 0.626 | 0.848 | 0.457 | 0.706 | 0.669 | 0.607 |
| 56 | FLJ32954 | 0.251 | 0.345 | 0.820 | 0.428 | 0.226 | 0.501 | 0.762 | 0.675 |
| 57 | FLJ38508 | 0.892 | 0.881 | 0.817 | 0.860 | 0.796 | 0.723 | 0.724 | 0.753 |
| 58 | FLJ39743 | 0.894 | 0.858 | 0.854 | 0.703 | 0.802 | 0.850 | 0.710 | 0.565 |
| 59 | FLJ46536 | 0.401 | 0.269 | 0.631 | 0.236 | 0.583 | 0.180 | 0.325 | 0.206 |
| 60 | FLJ90652 | 0.975 | 0.678 | 0.638 | 0.954 | 0.748 | 0.694 | 0.400 | 0.755 |
| 61 | FOXJ3 | 0.894 | 1.073 | 0.887 | 1.100 | 0.761 | 0.553 | 0.462 | 0.740 |
| 62 | FTL | 0.918 | 0.945 | 0.881 | 0.971 | 0.711 | 0.800 | 0.820 | 0.915 |
| 63 | FTSJ1 | 0.420 | 0.583 | 0.357 | 0.660 | 0.489 | 0.656 | 0.300 | 0.524 |
| 64 | FUBP1 | 0.944 | 0.992 | 0.908 | 0.901 | 0.668 | 0.594 | 0.608 | 0.574 |
| 65 | GJA8 | 0.386 | 0.517 | 0.339 | 0.508 | 0.224 | 0.419 | 0.285 | 0.535 |
| 66 | GLRB | 0.439 | 0.538 | 0.365 | 0.771 | 0.385 | 0.495 | 0.481 | 0.530 |
| 67 | GLTSCR1 | 0.503 | 0.632 | 0.397 | 0.517 | 0.433 | 0.389 | 0.719 | 0.299 |
| 68 | GPS1 | 0.698 | 0.795 | 0.759 | 0.648 | 0.535 | 0.580 | 0.535 | 0.418 |
| 69 | GSPT1 | 0.899 | 0.892 | 0.806 | 0.866 | 0.745 | 0.656 | 0.676 | 0.735 |
| 70 | GSTP1 | 0.885 | 0.868 | 0.827 | 0.996 | 0.789 | 0.692 | 0.781 | 0.997 |
| 71 | GUSB | 0.491 | 0.658 | 0.507 | 0.337 | 0.428 | 0.483 | 0.308 | 0.240 |
| 72 | H1FX | 0.906 | 0.630 | 0.380 | 0.831 | 0.774 | 0.461 | 0.255 | 0.602 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 73 | HCFC1 | 0.373 | 0.475 | 0.296 | 0.568 | 0.156 | 0.352 | 0.159 | 0.170 |
| 74 | HELZ | 0.521 | 0.488 | 0.389 | 0.441 | 0.497 | 0.261 | 0.517 | 0.396 |
| 75 | HEMK1 | 0.466 | 0.560 | 0.399 | 0.204 | 0.338 | 0.578 | 0.466 | 0.297 |
| 76 | HES6 | 0.556 | 0.628 | 0.664 | 0.410 | 0.541 | 0.374 | 0.555 | 0.149 |
| 77 | HIVEP3 | 0.852 | 0.693 | 0.757 | 0.801 | 0.952 | 0.567 | 0.364 | 0.800 |
| 78 | HNRPD | 0.949 | 0.925 | 0.888 | 0.839 | 0.870 | 0.893 | 0.737 | 0.728 |
| 79 | HNRPU | 0.152 | 0.461 | 0.294 | 0.375 | 0.072 | 0.162 | 0.259 | 0.090 |
| 80 | IBSP | 0.742 | 0.962 | 0.830 | 0.910 | 0.651 | 0.991 | 0.771 | 0.760 |
| 81 | IGFBP6 | 0.593 | 0.622 | 0.505 | 0.674 | 0.505 | 0.465 | 0.236 | 0.604 |
| 82 | INCA1 | 0.268 | 0.525 | 0.492 | 0.364 | 0.224 | 0.329 | 0.450 | 0.200 |
| 83 | ITSN2 | 0.687 | 0.950 | 0.966 | 0.793 | 0.526 | 0.893 | 0.919 | 0.656 |
| 84 | JMJD2B | 1.030 | 1.104 | 0.792 | 0.780 | 0.728 | 0.866 | 0.519 | 0.434 |
| 85 | KIAA0274 | 0.812 | 0.744 | 0.934 | 0.965 | 0.687 | 0.653 | 0.784 | 0.793 |
| 86 | KIAA1076 | 0.904 | 0.995 | 0.958 | 0.968 | 0.793 | 0.809 | 1.016 | 0.996 |
| 87 | KIR3DL1 | 0.507 | 0.742 | 0.567 | 0.476 | 0.463 | 0.657 | 0.507 | 0.424 |
| 88 | KIRREL2 | 0.536 | 0.415 | 0.694 | 0.425 | 0.501 | 0.453 | 0.478 | 0.485 |
| 89 | KLK5 | 0.874 | 0.899 | 0.848 | 0.844 | 0.782 | 0.708 | 0.705 | 0.730 |
| 90 | KREMEN1 | 0.891 | 0.907 | 0.926 | 0.918 | 0.809 | 0.914 | 0.837 | 0.853 |
| 91 | LARS | 0.441 | 0.549 | 0.709 | 0.252 | 0.307 | 0.421 | 0.583 | 0.176 |
| 92 | LCE1E | 0.811 | 0.913 | 0.826 | 0.420 | 0.752 | 0.788 | 0.619 | 0.248 |
| 93 | LCMR1 | 0.263 | 0.560 | 0.333 | 0.340 | 0.194 | 0.582 | 0.339 | 0.450 |
| 94 | LIF | 0.536 | 0.345 | 0.283 | 0.056 | 0.344 | 0.322 | 0.329 | 0.044 |
| 95 | LOC124245 | 0.559 | 0.855 | 0.384 | 0.317 | 0.576 | 0.688 | 0.407 | 0.478 |
| 96 | LOC374654 | 0.754 | 0.541 | 0.528 | 0.556 | 0.748 | 0.444 | 0.533 | 0.711 |
| 97 | LOC390790 | 0.224 | 0.105 | 0.355 | 0.488 | 0.274 | 0.100 | 0.238 | 0.337 |
| 98 | LOC400221 | 0.560 | 0.402 | 0.587 | 0.416 | 0.742 | 0.360 | 0.460 | 0.389 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 99 | LOC56901 | 0.383 | 0.524 | 0.417 | 0.555 | 0.298 | 0.510 | 0.555 | 0.487 |
| 100 | LPPR2 | 0.965 | 0.795 | 0.702 | 0.751 | 0.893 | 0.607 | 0.561 | 0.549 |
| 101 | LRRC33 | 0.336 | 0.190 | 0.382 | 0.393 | 0.427 | 0.092 | 0.328 | 0.307 |
| 102 | LUC7A | 0.401 | 0.647 | 0.417 | 0.570 | 0.466 | 0.508 | 0.197 | 0.477 |
| 103 | MAP2K7 | 0.884 | 0.984 | 0.864 | 1.007 | 0.881 | 0.903 | 0.797 | 0.978 |
| 104 | MAP3K1 | 0.904 | 0.872 | 0.896 | 0.720 | 0.858 | 0.783 | 0.875 | 0.656 |
| 105 | MBTD1 | 0.702 | 0.556 | 0.449 | 0.529 | 0.578 | 0.462 | 0.405 | 0.439 |
| 106 | MCRS1 | 0.335 | 0.637 | 0.639 | 0.844 | 0.089 | 0.477 | 0.364 | 0.736 |
| 107 | MED28 | 0.787 | 0.815 | 0.974 | 0.791 | 0.674 | 0.711 | 0.914 | 0.569 |
| 108 | MGC10471 | 0.380 | 0.467 | 0.347 | 0.404 | 0.511 | 0.548 | 0.450 | 0.359 |
| 109 | MGC21874 | 0.484 | 0.621 | 0.909 | 0.748 | 0.260 | 0.667 | 1.098 | 0.471 |
| 110 | MGC32871 | 0.457 | 0.374 | 0.369 | 0.435 | 0.711 | 0.524 | 0.571 | 0.641 |
| 111 | MGC39827 | 0.317 | 0.185 | 0.418 | 0.453 | 0.235 | 0.116 | 0.207 | 0.460 |
| 112 | MGC8902 | 0.936 | 0.849 | 0.880 | 0.906 | 0.938 | 0.783 | 0.739 | 0.723 |
| 113 | MMP15 | 0.898 | 0.833 | 0.730 | 0.635 | 0.693 | 0.710 | 0.597 | 0.545 |
| 114 | MMP24 | 0.820 | 1.052 | 0.846 | 1.005 | 0.541 | 0.625 | 0.798 | 0.693 |
| 115 | MOCS1 | 0.628 | 0.472 | 0.529 | 0.574 | 0.348 | 0.406 | 0.333 | 0.502 |
| 116 | MR1 | 0.555 | 0.271 | 0.215 | 0.477 | 0.397 | 0.241 | 0.169 | 0.610 |
| 117 | MVP | 0.514 | 0.212 | 0.520 | 0.702 | 0.618 | 0.123 | 0.514 | 0.379 |
| 118 | NANOG | 0.202 | 0.440 | 0.642 | 1.240 | 0.153 | 0.260 | 0.338 | 0.637 |
| 119 | NCBP1 | 0.215 | 0.335 | 0.386 | 0.285 | 0.263 | 0.411 | 0.508 | 0.303 |
| 120 | NEUROD2 | 0.461 | 0.602 | 0.467 | 0.467 | 0.724 | 0.566 | 0.362 | 0.333 |
| 121 | NFKB1 | 0.826 | 0.920 | 0.879 | 0.948 | 0.815 | 0.847 | 0.852 | 0.934 |
| 122 | NFRKB | 0.826 | 0.848 | 0.540 | 0.676 | 0.378 | 0.432 | 0.270 | 0.264 |
| 123 | NPEPL1 | 0.427 | 0.384 | 0.564 | 0.467 | 0.385 | 0.299 | 0.480 | 0.371 |
| 124 | NUDT8 | 0.383 | 0.345 | 0.273 | 0.568 | 0.417 | 0.402 | 0.428 | 0.571 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 125 | NUP107 | 0.670 | 0.781 | 0.982 | 0.321 | 0.472 | 0.643 | 0.986 | 0.252 |
| 126 | NXF1 | 0.213 | 0.122 | 0.824 | 0.231 | 0.142 | 0.090 | 0.673 | 0.159 |
| 127 | OACT1 | 0.893 | 0.889 | 0.907 | 0.908 | 0.894 | 0.791 | 0.831 | 0.800 |
| 128 | ODF2 | 0.161 | 0.435 | 0.467 | 0.279 | 0.173 | 0.805 | 0.447 | 0.287 |
| 129 | P29 | 0.886 | 0.957 | 0.908 | 1.012 | 0.732 | 0.909 | 0.800 | 0.958 |
| 130 | PCF11 | 0.093 | 0.263 | 0.397 | 0.195 | 0.093 | 0.164 | 0.231 | 0.098 |
| 131 | PDZK11 | 0.716 | 0.962 | 0.808 | 0.712 | 0.605 | 0.890 | 0.665 | 0.560 |
| 132 | PHB | 0.157 | 0.113 | 0.224 | 0.328 | 0.071 | 0.023 | 0.074 | 0.149 |
| 133 | PITX1 | 0.836 | 0.932 | 0.808 | 0.918 | 0.753 | 0.990 | 0.655 | 0.909 |
| 134 | PKP1 | 0.844 | 0.947 | 0.833 | 0.871 | 0.658 | 0.897 | 0.846 | 0.701 |
| 135 | POLH | 0.348 | 0.319 | 0.610 | 0.412 | 0.596 | 0.363 | 0.858 | 0.630 |
| 136 | OCT4 | 0.083 | 0.177 | 0.050 | 0.080 | 0.029 | 0.076 | 0.023 | 0.041 |
| 137 | PPAPDC2 | 0.334 | 0.254 | 0.291 | 0.532 | 0.299 | 0.177 | 0.147 | 0.415 |
| 138 | PPP2R3A | 0.766 | 0.453 | 0.858 | 0.782 | 0.533 | 0.498 | 0.500 | 0.810 |
| 139 | PRDM14 | 0.972 | 0.675 | 0.924 | 0.202 | 0.556 | 0.249 | 0.625 | 0.352 |
| 140 | PRDM9 | 0.783 | 0.111 | 0.764 | 1.088 | 0.443 | 0.275 | 0.652 | 0.800 |
| 141 | PRDX6 | 1.030 | 0.986 | 0.854 | 1.007 | 0.843 | 1.011 | 0.767 | 0.943 |
| 142 | PRO2730 | 0.774 | 0.557 | 0.815 | 0.930 | 0.440 | 0.351 | 0.641 | 0.699 |
| 143 | PROP1 | 0.598 | 0.542 | 0.580 | 0.685 | 0.392 | 0.619 | 0.626 | 0.559 |
| 144 | PSMD2 | 0.057 | 0.245 | 0.164 | 0.036 | 0.052 | 0.013 | 0.162 | 0.018 |
| 145 | PSTPIP2 | 0.809 | 0.988 | 0.989 | 0.687 | 0.675 | 0.937 | 0.851 | 0.525 |
| 146 | PXN | 0.433 | 0.525 | 0.421 | 0.314 | 0.258 | 0.430 | 0.195 | 0.374 |
| 147 | RALGDS | 0.819 | 0.890 | 0.784 | 0.748 | 0.632 | 0.712 | 0.563 | 0.618 |
| 148 | RASEF | 0.544 | 0.720 | 0.662 | 0.814 | 0.759 | 0.623 | 0.510 | 0.654 |
| 149 | RBM17 | 0.887 | 0.908 | 0.926 | 0.867 | 0.769 | 0.700 | 0.833 | 0.698 |
| 150 | REA | 0.330 | 0.330 | 0.549 | 0.183 | 0.231 | 0.243 | 0.413 | 0.130 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 151 | RHOA | 0.866 | 0.729 | 0.531 | 0.872 | 0.781 | 0.638 | 0.434 | 0.795 |
| 152 | RICTOR | 0.666 | 0.825 | 0.757 | 0.948 | 0.724 | 0.760 | 0.640 | 0.690 |
| 153 | RPESP | 0.412 | 0.445 | 0.362 | 0.326 | 0.376 | 0.493 | 0.478 | 0.490 |
| 154 | RRAS | 0.902 | 0.764 | 0.864 | 0.796 | 0.795 | 0.571 | 0.767 | 0.729 |
| 155 | SAMD7 | 0.571 | 0.462 | 0.409 | 0.637 | 0.399 | 0.499 | 0.399 | 0.507 |
| 156 | SERPINB2 | 0.339 | 0.376 | 0.161 | 0.164 | 0.352 | 0.531 | 0.227 | 0.250 |
| 157 | SERTAD2 | 1.025 | 0.882 | 1.100 | 0.980 | 0.663 | 0.621 | 0.679 | 0.678 |
| 158 | SF3A1 | 0.182 | 0.047 | 0.328 | 0.156 | 0.013 | 0.014 | 0.012 | 0.006 |
| 159 | SF3A3 | 0.911 | 0.087 | 0.396 | 0.849 | 0.839 | 0.071 | 0.299 | 0.766 |
| 160 | SFPQ | 0.565 | 0.370 | 0.440 | 0.219 | 0.542 | 0.313 | 0.384 | 0.198 |
| 161 | SFRS3 | 0.522 | 0.154 | 0.387 | 0.479 | 0.103 | 0.026 | 0.352 | 0.259 |
| 162 | SFXN3 | 0.469 | 0.520 | 0.492 | 0.473 | 0.201 | 0.705 | 0.459 | 0.597 |
| 163 | SOAT2 | 0.439 | 0.614 | 0.397 | 0.572 | 0.368 | 0.591 | 0.051 | 0.359 |
| 164 | SON | 0.073 | 0.048 | 0.070 | 0.036 | 0.016 | 0.044 | 0.073 | 0.006 |
| 165 | SOX14 | 0.744 | 0.607 | 0.836 | 0.747 | 0.573 | 0.616 | 0.749 | 0.566 |
| 166 | SPI1 | 0.830 | 0.838 | 0.732 | 0.969 | 0.691 | 0.711 | 0.651 | 0.826 |
| 167 | STK4 | 0.994 | 0.953 | 0.919 | 0.905 | 0.845 | 0.858 | 0.876 | 0.802 |
| 168 | SUV39H2 | 0.900 | 0.872 | 0.454 | 0.293 | 0.764 | 0.852 | 0.616 | 0.406 |
| 169 | SYNCRIP | 0.860 | 0.896 | 0.554 | 0.949 | 0.857 | 0.787 | 0.378 | 0.758 |
| 170 | SYTL4 | 0.909 | 0.938 | 0.631 | 0.989 | 0.811 | 0.918 | 0.468 | 0.990 |
| 171 | TAF2 | 0.676 | 0.588 | 0.885 | 0.886 | 0.573 | 0.511 | 0.718 | 0.829 |
| 172 | TAF7 | 0.811 | 0.981 | 0.664 | 1.023 | 0.691 | 0.964 | 0.573 | 0.955 |
| 173 | TBC1D10 | 0.537 | 0.351 | 0.510 | 0.514 | 0.413 | 0.496 | 0.592 | 0.320 |
| 174 | TCL1A | 0.762 | 0.928 | 0.925 | 0.949 | 0.605 | 0.883 | 0.933 | 0.808 |
| 175 | THRAP2 | 0.762 | 0.609 | 0.748 | 1.043 | 0.620 | 0.553 | 0.533 | 0.789 |
| 176 | THRAP4 | 0.894 | 0.799 | 1.071 | 0.962 | 0.665 | 0.546 | 0.734 | 0.784 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 177 | TMEM14B | 0.562 | 0.248 | 0.564 | 0.524 | 0.389 | 0.122 | 0.525 | 0.406 |
| 178 | TNRC11 | 0.694 | 1.014 | 0.886 | 0.678 | 0.744 | 0.683 | 0.450 | 0.415 |
| 179 | TPD52L1 | 0.463 | 0.361 | 0.802 | 0.404 | 0.703 | 0.217 | 0.299 | 0.425 |
| 180 | TPM1 | 0.250 | 0.311 | 0.456 | 0.511 | 0.269 | 0.306 | 0.381 | 0.425 |
| 181 | TPR | 0.444 | 0.420 | 0.510 | 0.262 | 0.566 | 0.503 | 0.352 | 0.420 |
| 182 | TRIP | 0.540 | 0.252 | 0.940 | 0.959 | 0.546 | 0.587 | 0.788 | 0.758 |
| 183 | TRIP15 | 0.790 | 0.844 | 0.808 | 0.785 | 0.522 | 0.810 | 0.455 | 0.692 |
| 184 | TRPA1 | 0.357 | 0.485 | 0.306 | 0.290 | 0.263 | 0.724 | 0.571 | 0.422 |
| 185 | ULK2 | 1.027 | 0.928 | 1.087 | 0.958 | 0.705 | 0.397 | 0.692 | 0.676 |
| 186 | VGCNL1 | 0.930 | 0.904 | 0.929 | 0.931 | 0.815 | 0.835 | 0.828 | 0.885 |
| 187 | VMP | 0.511 | 0.539 | 0.415 | 0.469 | 0.462 | 0.393 | 0.422 | 0.324 |
| 188 | VWF | 0.960 | 0.847 | 0.964 | 0.951 | 0.933 | 0.776 | 0.914 | 0.897 |
| 189 | XRCC1 | 0.892 | 0.866 | 0.808 | 0.891 | 0.824 | 0.788 | 0.697 | 0.878 |
| 190 | YAP1 | 1.187 | 1.254 | 0.850 | 0.946 | 0.661 | 0.616 | 0.413 | 0.786 |
| 191 | YY1 | 0.798 | 0.684 | 0.879 | 0.614 | 0.406 | 0.462 | 0.615 | 0.463 |
| 192 | ZFP36 | 0.233 | 0.867 | 0.658 | 0.492 | 0.201 | 0.481 | 0.322 | 0.499 |
| 193 | ZFP64 | 0.906 | 0.832 | 0.478 | 0.992 | 0.264 | 0.577 | 0.162 | 0.655 |
| 194 | ZIC4 | 0.978 | 0.964 | 0.703 | 0.246 | 0.490 | 0.540 | 0.474 | 0.308 |
| 195 | ZNF136 | 1.006 | 0.839 | 0.801 | 0.681 | 0.764 | 0.617 | 0.665 | 0.713 |
| 196 | ZNF138 | 0.942 | 1.137 | 0.824 | 0.725 | 0.736 | 0.567 | 0.602 | 0.494 |
| 197 | ZNF206 | 0.718 | 0.343 | 0.537 | 0.370 | 0.299 | 0.303 | 0.677 | 0.114 |
| 198 | ZNF35 | 0.649 | 0.881 | 0.706 | 0.898 | 0.721 | 0.819 | 0.633 | 0.657 |
| 199 | ZNF43 | 0.978 | 0.754 | 0.920 | 1.013 | 0.712 | 0.452 | 0.580 | 1.064 |
| 200 | ZNF434 | 0.893 | 0.277 | 0.607 | 0.728 | 0.779 | 0.190 | 0.297 | 0.881 |

**Table 4b: HES2 hESCs secondary screen data (relative NANOG and OCT4 expression expression)**

Deconvoluted siRNA screen data for the 200 genes

| No. | Gene symbol | NANOG _S1 | NANOG _S2 | NANOG _S3 | NANOG _S4 | OCT4 _S1 | OCT4 _S2 | OCT4 _S3 | OCT4 _S4 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | ABP1 | 0.717 | 1.021 | 0.933 | 0.844 | 0.671 | 1.099 | 0.912 | 0.949 |
| 2 | ABTB1 | 0.876 | 0.767 | 0.976 | 0.953 | 0.971 | 1.021 | 0.893 | 0.938 |
| 3 | ADA | 0.795 | 0.772 | 1.007 | 0.868 | 0.871 | 0.709 | 0.982 | 0.719 |
| 4 | ADAMTS1 | 0.795 | 0.714 | 0.818 | 0.917 | 0.843 | 0.626 | 0.762 | 0.872 |
| 5 | AGPS | 0.767 | 0.809 | 0.791 | 0.907 | 0.665 | 0.585 | 0.654 | 0.761 |
| 6 | AIPL1 | 0.828 | 0.682 | 0.808 | 0.781 | 0.591 | 0.549 | 0.626 | 0.541 |
| 7 | ANGPT4 | 1.022 | 0.937 | 0.932 | 1.058 | 1.057 | 0.845 | 0.794 | 0.888 |
| 8 | ANKRD1 | 0.730 | 0.777 | 0.731 | 0.827 | 0.650 | 0.722 | 0.530 | 0.634 |
| 9 | ANKRD31 | 0.751 | 0.848 | 0.823 | 0.712 | 0.537 | 0.612 | 0.571 | 0.473 |
| 10 | ANXA4 | 0.827 | 0.659 | 0.826 | 0.466 | 0.968 | 0.727 | 0.776 | 0.415 |
| 11 | APLP2 | 0.948 | 0.945 | 0.894 | 0.974 | 0.836 | 0.913 | 0.862 | 0.932 |
| 12 | ATOH8 | 0.517 | 0.802 | 0.581 | 0.933 | 0.452 | 0.745 | 0.583 | 1.246 |
| 13 | BCL6B | 0.981 | 0.626 | 0.619 | 0.887 | 0.942 | 0.498 | 0.630 | 1.082 |
| 14 | BDP1 | 0.644 | 0.828 | 0.701 | 0.966 | 0.420 | 0.664 | 0.544 | 0.752 |
| 15 | BENE | 0.687 | 0.811 | 0.658 | 0.630 | 0.440 | 0.873 | 0.666 | 0.461 |
| 16 | C22ORF16 | 0.709 | 0.909 | 0.855 | 0.720 | 0.725 | 0.839 | 0.759 | 0.723 |
| 17 | CAMP | 0.687 | 0.639 | 0.855 | 0.770 | 0.479 | 0.440 | 0.720 | 0.602 |
| 18 | CAPN2 | 0.686 | 0.695 | 0.856 | 0.781 | 0.631 | 0.509 | 0.750 | 0.870 |
| 19 | CCL2 | 0.835 | 0.894 | 0.906 | 0.867 | 0.758 | 0.950 | 1.044 | 0.894 |
| 20 | CDC42 | 0.826 | 0.827 | 0.597 | 0.693 | 0.659 | 0.646 | 0.571 | 0.577 |
| 21 | CDX2 | 0.660 | 0.703 | 0.798 | 0.801 | 0.486 | 0.578 | 0.564 | 0.657 |
| 22 | CGGBP1 | 0.744 | 0.813 | 1.013 | 0.568 | 1.147 | 0.755 | 0.945 | 0.529 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 23 | COL11A1 | 0.804 | 0.600 | 0.761 | 0.752 | 0.720 | 0.593 | 0.583 | 0.593 |
| 24 | COPS4 | 0.767 | 0.700 | 0.806 | 0.836 | 0.600 | 0.439 | 0.644 | 0.564 |
| 25 | CORT | 0.946 | 0.869 | 0.837 | 0.833 | 0.800 | 0.840 | 0.745 | 0.689 |
| 26 | CPSF3 | 0.415 | 0.431 | 0.536 | 0.539 | 0.411 | 0.362 | 0.510 | 0.619 |
| 27 | CREBL2 | 0.980 | 0.600 | 0.763 | 0.611 | 1.094 | 0.499 | 0.667 | 0.565 |
| 28 | CRK7 | 0.826 | 1.100 | 1.164 | 0.741 | 0.584 | 0.809 | 1.203 | 0.569 |
| 29 | CRLF1 | 0.932 | 1.059 | 0.691 | 0.937 | 0.797 | 1.036 | 0.560 | 0.984 |
| 30 | CRSP2 | 0.916 | 0.753 | 0.691 | 0.570 | 0.884 | 0.542 | 0.469 | 0.541 |
| 31 | CTNNA3 | 0.707 | 0.914 | 0.879 | 0.842 | 0.620 | 0.843 | 0.887 | 0.713 |
| 32 | CYBA | 0.821 | 0.893 | 0.648 | 0.641 | 0.753 | 0.852 | 0.587 | 0.520 |
| 33 | DDEF1 | 0.867 | 0.701 | 0.903 | 0.869 | 0.853 | 0.624 | 0.784 | 0.761 |
| 34 | DDIT3 | 0.608 | 1.037 | 0.807 | 0.965 | 0.612 | 0.840 | 0.727 | 1.068 |
| 35 | DEFB126 | 0.620 | 0.779 | 0.754 | 0.810 | 0.452 | 0.667 | 0.605 | 0.648 |
| 36 | DKFZP564B147 | 0.796 | 0.818 | 0.955 | 0.916 | 0.740 | 0.751 | 0.752 | 0.808 |
| 37 | DKFZP564K142 | 0.764 | 0.919 | 0.746 | 0.796 | 0.665 | 0.937 | 0.774 | 0.855 |
| 38 | DRG2 | 0.880 | 0.627 | 1.086 | 0.808 | 0.652 | 0.692 | 1.237 | 0.589 |
| 39 | E4F1 | 0.809 | 0.661 | 0.804 | 0.744 | 0.724 | 0.672 | 0.650 | 0.929 |
| 40 | EDF1 | 0.762 | 0.836 | 0.711 | 0.754 | 0.712 | 0.770 | 0.673 | 1.062 |
| 41 | EIF2B1 | 0.826 | 0.722 | 0.941 | 0.932 | 0.597 | 0.665 | 0.745 | 0.777 |
| 42 | EIF2B2 | 0.608 | 0.540 | 0.646 | 0.855 | 0.455 | 0.522 | 0.575 | 0.729 |
| 43 | EIF2B3 | 0.810 | 0.287 | 0.298 | 0.468 | 0.688 | 0.227 | 0.237 | 0.423 |
| 44 | EIF2B4 | 0.417 | 0.408 | 0.748 | 0.355 | 0.398 | 0.351 | 0.691 | 0.290 |
| 45 | EIF2S2 | 0.450 | 0.345 | 0.351 | 0.303 | 0.363 | 0.330 | 0.297 | 0.276 |
| 46 | ELYS | 0.535 | 0.592 | 0.789 | 0.746 | 0.400 | 0.412 | 0.613 | 0.582 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 47 | ENPP7 | 0.970 | 0.837 | 0.812 | 0.762 | 0.807 | 0.700 | 0.703 | 0.677 |
| 48 | EP300 | 0.769 | 0.809 | 0.735 | 0.723 | 0.877 | 0.818 | 0.700 | 0.558 |
| 49 | ETF1 | 0.463 | 0.712 | 0.722 | 0.458 | 0.458 | 0.621 | 0.538 | 0.376 |
| 50 | FAM19A1 | 0.834 | 0.774 | 0.653 | 0.601 | 0.798 | 0.847 | 0.584 | 0.592 |
| 51 | FLJ20898 | 0.651 | 0.821 | 0.914 | 0.894 | 0.592 | 0.823 | 0.759 | 0.850 |
| 52 | FLJ23447 | 0.961 | 0.646 | 0.799 | 0.806 | 1.029 | 0.650 | 0.847 | 0.795 |
| 53 | FLJ23751 | 0.585 | 0.810 | 0.769 | 0.850 | 0.496 | 0.760 | 0.626 | 0.800 |
| 54 | FLJ25439 | 0.869 | 0.942 | 1.056 | 0.971 | 0.837 | 1.074 | 1.038 | 0.986 |
| 55 | FLJ25952 | 0.980 | 0.896 | 1.068 | 0.882 | 0.843 | 0.899 | 1.105 | 0.840 |
| 56 | FLJ32954 | 0.863 | 0.792 | 0.862 | 0.847 | 0.964 | 0.883 | 0.901 | 0.907 |
| 57 | FLJ38508 | 0.776 | 0.968 | 0.760 | 0.800 | 0.670 | 0.820 | 0.793 | 0.825 |
| 58 | FLJ39743 | 0.816 | 0.901 | 0.888 | 0.823 | 0.681 | 0.836 | 0.660 | 0.595 |
| 59 | FLJ46536 | 0.931 | 0.699 | 0.901 | 0.826 | 0.985 | 0.686 | 0.806 | 1.213 |
| 60 | FLJ90652 | 0.791 | 0.856 | 0.712 | 0.606 | 0.731 | 0.969 | 0.613 | 0.655 |
| 61 | FOXJ3 | 0.756 | 0.871 | 0.750 | 0.557 | 0.554 | 0.730 | 0.626 | 0.645 |
| 62 | FTL | 0.715 | 0.768 | 0.717 | 0.639 | 0.588 | 0.648 | 0.742 | 0.662 |
| 63 | FTSJ1 | 0.672 | 0.831 | 0.734 | 0.696 | 0.570 | 0.866 | 0.728 | 0.610 |
| 64 | FUBP1 | 0.801 | 0.749 | 0.912 | 0.912 | 0.859 | 0.659 | 0.911 | 0.673 |
| 65 | GJA8 | 0.841 | 0.956 | 0.911 | 0.890 | 0.710 | 0.960 | 0.885 | 0.794 |
| 66 | GLRB | 0.784 | 0.814 | 0.940 | 0.891 | 0.593 | 0.680 | 0.795 | 0.820 |
| 67 | GLTSCR1 | 0.809 | 0.764 | 1.066 | 0.910 | 0.704 | 0.631 | 0.908 | 0.713 |
| 68 | GPS1 | 0.768 | 0.903 | 0.817 | 1.043 | 0.603 | 0.675 | 0.611 | 1.127 |
| 69 | GSPT1 | 0.800 | 0.667 | 0.724 | 0.654 | 0.725 | 0.537 | 0.628 | 0.575 |
| 70 | GSTP1 | 0.916 | 0.845 | 0.723 | 0.995 | 0.898 | 0.641 | 0.601 | 0.948 |
| 71 | GUSB | 0.879 | 0.745 | 0.742 | 0.842 | 0.791 | 0.555 | 0.644 | 0.703 |
| 72 | H1FX | 0.822 | 0.617 | 0.683 | 0.692 | 0.624 | 0.445 | 0.535 | 0.708 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 73 | HCFC1 | 0.552 | 0.471 | 0.564 | 0.670 | 0.219 | 0.329 | 0.232 | 0.341 |
| 74 | HELZ | 0.908 | 0.621 | 0.686 | 0.843 | 0.776 | 0.447 | 0.688 | 0.905 |
| 75 | HEMK1 | 0.652 | 1.059 | 0.707 | 0.600 | 0.530 | 0.888 | 0.694 | 0.566 |
| 76 | HES6 | 0.749 | 0.674 | 0.873 | 0.660 | 0.667 | 0.593 | 0.831 | 0.664 |
| 77 | HIVEP3 | 0.739 | 0.683 | 0.698 | 0.783 | 0.734 | 0.599 | 0.643 | 0.683 |
| 78 | HNRPD | 0.895 | 0.885 | 0.964 | 0.836 | 0.818 | 0.779 | 0.827 | 0.848 |
| 79 | HNRPU | 0.565 | 0.401 | 0.530 | 0.333 | 0.484 | 0.381 | 0.569 | 0.353 |
| 80 | IBSP | 0.718 | 0.848 | 0.833 | 0.819 | 0.701 | 0.918 | 0.744 | 0.681 |
| 81 | IGFBP6 | 0.737 | 0.821 | 0.487 | 0.569 | 0.778 | 0.663 | 0.402 | 0.623 |
| 82 | INCA1 | 0.731 | 0.913 | 0.934 | 0.790 | 0.664 | 0.738 | 0.817 | 0.571 |
| 83 | ITSN2 | 0.822 | 0.959 | 0.887 | 0.601 | 0.642 | 0.896 | 0.807 | 0.408 |
| 84 | JMJD2B | 0.544 | 0.713 | 0.514 | 0.753 | 0.476 | 0.802 | 0.425 | 0.600 |
| 85 | KIAA0274 | 0.666 | 0.703 | 0.787 | 0.854 | 0.525 | 0.569 | 0.672 | 0.687 |
| 86 | KIAA1076 | 0.919 | 1.036 | 0.961 | 0.753 | 0.751 | 0.865 | 0.864 | 0.576 |
| 87 | KIR3DL1 | 0.914 | 0.823 | 1.068 | 0.768 | 0.725 | 0.720 | 1.055 | 0.684 |
| 88 | KIRREL2 | 0.817 | 1.018 | 0.959 | 0.958 | 0.667 | 0.885 | 0.931 | 0.887 |
| 89 | KLK5 | 0.702 | 0.647 | 0.715 | 0.905 | 0.608 | 0.438 | 0.511 | 0.930 |
| 90 | KREMEN1 | 0.754 | 0.865 | 0.819 | 0.802 | 0.607 | 0.775 | 0.695 | 0.684 |
| 91 | LARS | 0.773 | 0.854 | 0.844 | 0.693 | 0.621 | 0.809 | 0.705 | 0.491 |
| 92 | LCE1E | 0.827 | 0.834 | 0.865 | 0.616 | 0.846 | 0.846 | 0.575 | 0.390 |
| 93 | LCMR1 | 0.399 | 0.631 | 0.684 | 0.934 | 0.296 | 0.510 | 0.612 | 0.788 |
| 94 | LIF | 0.712 | 0.828 | 0.905 | 0.579 | 0.561 | 0.699 | 0.792 | 0.453 |
| 95 | LOC124245 | 0.923 | 0.818 | 0.846 | 1.270 | 0.588 | 1.321 | 0.585 | 0.949 |
| 96 | LOC374654 | 0.834 | 0.718 | 0.793 | 0.855 | 0.812 | 0.656 | 0.840 | 0.994 |
| 97 | LOC390790 | 0.623 | 0.551 | 0.691 | 0.667 | 0.534 | 0.484 | 0.699 | 0.669 |
| 98 | LOC400221 | 0.813 | 0.809 | 0.942 | 0.942 | 0.747 | 0.676 | 0.939 | 0.942 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 99 | LOC56901 | 0.794 | 0.857 | 0.891 | 0.824 | 0.597 | 0.747 | 0.773 | 0.771 |
| 100 | LPPR2 | 0.896 | 0.746 | 0.803 | 0.969 | 0.898 | 0.542 | 0.769 | 0.730 |
| 101 | LRRC33 | 0.825 | 0.776 | 0.929 | 0.896 | 0.757 | 0.705 | 0.897 | 0.810 |
| 102 | LUC7A | 0.787 | 0.777 | 0.499 | 0.838 | 0.739 | 0.810 | 0.421 | 0.744 |
| 103 | MAP2K7 | 0.872 | 0.894 | 0.667 | 1.055 | 0.896 | 0.806 | 0.626 | 0.973 |
| 104 | MAP3K1 | 0.790 | 0.936 | 0.954 | 0.701 | 0.650 | 0.787 | 0.815 | 0.586 |
| 105 | MBTD1 | 0.773 | 0.808 | 0.954 | 0.965 | 0.683 | 0.770 | 0.892 | 0.898 |
| 106 | MCRS1 | 0.560 | 0.724 | 0.761 | 0.713 | 0.506 | 0.837 | 0.803 | 0.765 |
| 107 | MED28 | 0.728 | 0.804 | 0.836 | 0.824 | 0.734 | 0.615 | 0.827 | 0.687 |
| 108 | MGC10471 | 0.945 | 0.996 | 0.545 | 0.722 | 0.871 | 0.916 | 0.558 | 0.718 |
| 109 | MGC21874 | 0.628 | 0.732 | 0.592 | 0.868 | 0.662 | 0.664 | 0.712 | 0.869 |
| 110 | MGC32871 | 0.748 | 0.839 | 0.836 | 1.008 | 0.726 | 0.754 | 0.738 | 0.868 |
| 111 | MGC39827 | 0.871 | 0.685 | 0.780 | 0.726 | 0.721 | 0.640 | 0.623 | 0.630 |
| 112 | MGC8902 | 0.847 | 0.899 | 0.899 | 0.976 | 0.687 | 0.715 | 0.676 | 0.729 |
| 113 | MMP15 | 0.912 | 0.657 | 0.737 | 0.797 | 0.741 | 0.509 | 0.598 | 0.561 |
| 114 | MMP24 | 0.920 | 0.697 | 0.772 | 1.071 | 1.055 | 0.677 | 0.695 | 1.026 |
| 115 | MOCS1 | 0.617 | 0.959 | 0.824 | 0.945 | 0.386 | 0.858 | 0.690 | 0.770 |
| 116 | MR1 | 0.774 | 0.587 | 0.744 | 0.752 | 0.749 | 0.565 | 0.708 | 0.801 |
| 117 | MVP | 0.885 | 0.555 | 0.959 | 1.035 | 0.868 | 0.506 | 0.881 | 0.878 |
| 118 | NANOG | 0.684 | 0.859 | 0.615 | 0.805 | 0.415 | 0.835 | 0.606 | 0.959 |
| 119 | NCBP1 | 0.686 | 0.843 | 0.901 | 0.572 | 0.621 | 0.867 | 1.017 | 0.543 |
| 120 | NEUROD2 | 0.962 | 0.792 | 0.933 | 0.834 | 0.877 | 0.807 | 0.873 | 0.729 |
| 121 | NFKB1 | 0.843 | 0.963 | 0.973 | 0.953 | 0.708 | 0.789 | 0.959 | 0.868 |
| 122 | NFRKB | 0.754 | 0.694 | 0.739 | 0.680 | 0.724 | 0.774 | 0.830 | 0.633 |
| 123 | NPEPL1 | 0.502 | 0.699 | 0.982 | 0.760 | 0.499 | 0.779 | 1.009 | 0.730 |
| 124 | NUDT8 | 0.785 | 0.751 | 0.860 | 0.921 | 0.769 | 0.516 | 0.945 | 0.977 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 125 | NUP107 | 0.637 | 0.878 | 0.934 | 0.408 | 0.410 | 0.760 | 0.930 | 0.264 |
| 126 | NXF1 | 0.439 | 0.287 | 0.707 | 0.453 | 0.298 | 0.200 | 0.541 | 0.281 |
| 127 | OACT1 | 0.851 | 0.707 | 0.797 | 0.842 | 0.759 | 0.580 | 0.601 | 0.717 |
| 128 | ODF2 | 0.585 | 0.751 | 0.884 | 0.854 | 0.615 | 0.775 | 0.767 | 0.642 |
| 129 | P29 | 0.853 | 0.764 | 0.833 | 0.790 | 0.754 | 0.678 | 0.829 | 0.780 |
| 130 | PCF11 | 0.542 | 0.886 | 0.858 | 0.635 | 0.430 | 0.774 | 0.701 | 0.455 |
| 131 | PDZK11 | 0.747 | 0.728 | 0.407 | 0.676 | 0.794 | 0.777 | 0.263 | 0.470 |
| 132 | PHB | 0.463 | 0.391 | 0.463 | 0.524 | 0.349 | 0.356 | 0.402 | 0.468 |
| 133 | PITX1 | 0.767 | 0.929 | 0.584 | 0.876 | 0.588 | 0.841 | 0.419 | 0.723 |
| 134 | PKP1 | 0.856 | 0.925 | 0.756 | 0.895 | 0.744 | 0.962 | 0.786 | 0.736 |
| 135 | POLH | 0.872 | 0.980 | 0.856 | 0.850 | 0.768 | 0.715 | 0.977 | 0.922 |
| 136 | OCT4 | 0.389 | 0.663 | 0.269 | 0.490 | 0.254 | 0.459 | 0.157 | 0.246 |
| 137 | PPAPDC2 | 0.870 | 0.857 | 0.635 | 0.847 | 0.811 | 0.692 | 0.510 | 0.754 |
| 138 | PPP2R3A | 0.955 | 0.608 | 0.825 | 0.746 | 0.965 | 0.488 | 0.986 | 0.767 |
| 139 | PRDM14 | 0.875 | 0.606 | 0.958 | 0.596 | 0.775 | 0.405 | 1.512 | 0.664 |
| 140 | PRDM9 | 0.580 | 0.440 | 0.631 | 0.707 | 0.707 | 0.510 | 0.775 | 0.782 |
| 141 | PRDX6 | 0.848 | 0.680 | 0.756 | 0.884 | 0.666 | 0.696 | 0.733 | 0.763 |
| 142 | PRO2730 | 0.708 | 0.704 | 0.852 | 0.976 | 0.659 | 0.396 | 0.906 | 1.117 |
| 143 | PROP1 | 0.735 | 0.727 | 0.592 | 0.831 | 1.004 | 0.456 | 0.701 | 0.848 |
| 144 | PSMD2 | 0.267 | 0.322 | 0.730 | 0.212 | 0.188 | 0.226 | 0.690 | 0.179 |
| 145 | PSTPIP2 | 0.568 | 0.768 | 0.732 | 0.523 | 0.382 | 0.749 | 0.677 | 0.351 |
| 146 | PXN | 0.829 | 0.736 | 0.597 | 0.756 | 0.781 | 0.659 | 0.515 | 0.724 |
| 147 | RALGDS | 0.889 | 0.835 | 0.968 | 0.739 | 0.608 | 0.630 | 0.786 | 0.596 |
| 148 | RASEF | 0.675 | 1.036 | 0.940 | 0.735 | 0.658 | 0.953 | 0.852 | 0.683 |
| 149 | RBM17 | 0.816 | 0.914 | 0.803 | 0.829 | 0.726 | 0.785 | 0.756 | 0.758 |
| 150 | REA | 0.298 | 0.422 | 0.506 | 0.366 | 0.179 | 0.340 | 0.408 | 0.272 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 151 | RHOA | 0.761 | 0.793 | 0.775 | 0.765 | 0.802 | 0.800 | 0.728 | 0.747 |
| 152 | RICTOR | 0.746 | 0.971 | 0.960 | 0.927 | 0.723 | 0.773 | 1.005 | 0.797 |
| 153 | RPESP | 0.864 | 0.841 | 0.777 | 0.716 | 0.839 | 0.895 | 0.707 | 0.511 |
| 154 | RRAS | 0.877 | 0.751 | 0.798 | 0.776 | 0.798 | 0.477 | 0.620 | 0.702 |
| 155 | SAMD7 | 0.784 | 0.823 | 0.849 | 0.909 | 0.672 | 0.750 | 0.761 | 0.926 |
| 156 | SERPINB2 | 0.852 | 0.854 | 0.777 | 0.780 | 0.831 | 0.884 | 0.854 | 0.748 |
| 157 | SERTAD2 | 0.767 | 0.849 | 0.833 | 0.910 | 0.926 | 0.936 | 0.777 | 0.950 |
| 158 | SF3A1 | 0.158 | 0.332 | 0.284 | 0.146 | 0.112 | 0.267 | 0.195 | 0.099 |
| 159 | SF3A3 | 0.798 | 0.262 | 0.269 | 0.682 | 0.610 | 0.164 | 0.171 | 0.543 |
| 160 | SFPQ | 0.379 | 0.455 | 0.675 | 0.339 | 0.260 | 0.362 | 0.493 | 0.266 |
| 161 | SFRS3 | 0.640 | 0.151 | 0.811 | 0.456 | 0.478 | 0.082 | 0.587 | 0.339 |
| 162 | SFXN3 | 0.763 | 0.792 | 0.700 | 0.924 | 0.785 | 0.822 | 0.733 | 0.929 |
| 163 | SOAT2 | 0.809 | 0.835 | 0.529 | 0.898 | 0.697 | 0.865 | 0.482 | 0.772 |
| 164 | SON | 0.683 | 0.524 | 0.317 | 0.405 | 0.312 | 0.184 | 0.103 | 0.126 |
| 165 | SOX14 | 0.665 | 0.714 | 0.738 | 0.574 | 0.623 | 0.699 | 0.886 | 0.500 |
| 166 | SPI1 | 0.794 | 0.939 | 0.754 | 0.883 | 0.593 | 0.801 | 0.645 | 0.712 |
| 167 | STK4 | 0.965 | 1.015 | 0.864 | 0.869 | 0.833 | 0.887 | 0.708 | 0.708 |
| 168 | SUV39H2 | 1.025 | 0.986 | 0.600 | 0.573 | 1.152 | 1.027 | 0.270 | 0.687 |
| 169 | SYNCRIP | 0.756 | 0.851 | 0.710 | 0.861 | 0.645 | 0.738 | 0.564 | 0.757 |
| 170 | SYTL4 | 0.761 | 0.782 | 0.653 | 0.843 | 0.721 | 0.851 | 0.572 | 0.832 |
| 171 | TAF2 | 0.769 | 0.753 | 0.766 | 0.924 | 0.716 | 0.816 | 0.874 | 0.834 |
| 172 | TAF7 | 0.661 | 0.903 | 0.690 | 0.853 | 0.417 | 0.873 | 0.483 | 0.901 |
| 173 | TBC1D10 | 0.689 | 0.733 | 0.886 | 0.946 | 0.610 | 0.735 | 0.839 | 0.832 |
| 174 | TCL1A | 0.613 | 0.726 | 0.743 | 0.813 | 0.474 | 0.610 | 0.642 | 0.595 |
| 175 | THRAP2 | 0.878 | 0.601 | 0.766 | 0.857 | 0.803 | 0.581 | 0.831 | 1.149 |
| 176 | THRAP4 | 1.024 | 0.947 | 0.556 | 0.779 | 0.986 | 0.884 | 0.597 | 0.826 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 177 | TMEM14B | 0.958 | 0.775 | 0.943 | 0.965 | 0.914 | 0.627 | 0.913 | 0.805 |
| 178 | TNRC11 | 0.603 | 0.811 | 0.605 | 0.721 | 0.626 | 0.870 | 0.665 | 0.932 |
| 179 | TPD52L1 | 0.849 | 0.913 | 0.926 | 0.671 | 0.694 | 0.776 | 0.866 | 0.594 |
| 180 | TPM1 | 0.801 | 0.827 | 0.867 | 0.768 | 0.855 | 0.842 | 1.021 | 0.692 |
| 181 | TPR | 0.728 | 0.800 | 0.925 | 0.687 | 0.602 | 0.742 | 0.886 | 0.571 |
| 182 | TRIP | 0.735 | 0.820 | 0.920 | 0.716 | 1.016 | 0.786 | 1.088 | 0.703 |
| 183 | TRIP15 | 0.978 | 0.940 | 0.997 | 1.154 | 0.701 | 0.805 | 1.294 | 1.277 |
| 184 | TRPA1 | 0.636 | 0.791 | 0.765 | 0.799 | 0.582 | 0.835 | 0.864 | 0.764 |
| 185 | ULK2 | 0.937 | 1.145 | 0.941 | 0.960 | 0.994 | 0.992 | 1.040 | 1.009 |
| 186 | VGCNL1 | 0.923 | 0.928 | 0.796 | 0.623 | 0.866 | 0.870 | 0.661 | 0.466 |
| 187 | VMP | 0.848 | 1.001 | 0.654 | 0.939 | 0.778 | 0.960 | 0.613 | 0.849 |
| 188 | VWF | 0.681 | 0.676 | 0.697 | 0.800 | 0.674 | 0.604 | 0.726 | 0.808 |
| 189 | XRCC1 | 0.962 | 0.931 | 0.868 | 0.965 | 0.903 | 0.958 | 0.740 | 0.982 |
| 190 | YAP1 | 0.853 | 0.579 | 0.507 | 0.764 | 1.151 | 0.765 | 0.572 | 0.782 |
| 191 | YY1 | 0.877 | 0.767 | 0.846 | 0.906 | 0.889 | 0.925 | 1.281 | 0.550 |
| 192 | ZFP36 | 0.677 | 0.920 | 0.880 | 0.879 | 0.538 | 0.851 | 0.746 | 0.844 |
| 193 | ZFP64 | 0.843 | 1.009 | 0.645 | 0.742 | 0.659 | 1.065 | 0.451 | 0.580 |
| 194 | ZIC4 | 0.459 | 0.649 | 0.789 | 0.617 | 0.517 | 0.585 | 0.652 | 0.727 |
| 195 | ZNF136 | 0.718 | 0.822 | 0.694 | 0.849 | 0.571 | 0.689 | 0.834 | 1.525 |
| 196 | ZNF138 | 0.880 | 0.844 | 0.816 | 0.621 | 1.201 | 0.800 | 0.771 | 0.785 |
| 197 | ZNF206 | 0.677 | 0.700 | 0.860 | 0.580 | 0.694 | 0.505 | 0.672 | 0.323 |
| 198 | ZNF35 | 0.818 | 0.860 | 0.660 | 0.857 | 0.913 | 0.895 | 0.649 | 0.875 |
| 199 | ZNF43 | 0.998 | 0.683 | 0.815 | 1.035 | 1.090 | 0.647 | 0.749 | 1.071 |
| 200 | ZNF434 | 1.027 | 0.454 | 0.644 | 1.240 | 1.080 | 0.286 | 0.515 | 1.328 |

**Table 4c: HES3 hESCs secondary screen data (relative NANOG and OCT4 expression)**

Deconvoluted siRNA screen data for the 200 genes

| No. | Gene symbol | NANOG_ S1 | NANOG_ S2 | NANOG_ S3 | NANOG_ S4 | OCT4_ S1 | OCT4_ S2 | OCT4_ S3 | OCT4_ S4 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | ABP1 | 0.630 | 1.015 | 0.759 | 0.749 | 0.734 | 1.216 | 0.805 | 0.906 |
| 2 | ABTB1 | 1.036 | 0.926 | 0.849 | 0.718 | 1.167 | 1.276 | 0.831 | 0.785 |
| 3 | ADA | 1.056 | 0.728 | 0.773 | 0.809 | 1.234 | 0.766 | 0.788 | 0.767 |
| 4 | ADAMTS1 | 1.028 | 0.351 | 0.740 | 1.304 | 1.625 | 0.299 | 0.726 | 1.593 |
| 5 | AGPS | 0.592 | 0.249 | 0.598 | 0.569 | 0.740 | 0.341 | 0.703 | 0.725 |
| 6 | AIPL1 | 0.775 | 0.620 | 0.899 | 0.835 | 0.938 | 0.705 | 1.036 | 1.078 |
| 7 | ANGPT4 | 1.014 | 0.834 | 0.455 | 0.240 | 1.151 | 0.977 | 0.412 | 0.204 |
| 8 | ANKRD1 | 0.784 | 0.659 | 0.736 | 0.591 | 0.849 | 0.754 | 0.901 | 0.709 |
| 9 | ANKRD31 | 0.830 | 0.746 | 0.523 | 0.699 | 0.970 | 0.922 | 0.802 | 0.931 |
| 10 | ANXA4 | 1.008 | 0.576 | 1.116 | 0.378 | 1.268 | 0.676 | 1.360 | 0.326 |
| 11 | APLP2 | 0.631 | 0.871 | 1.006 | 0.951 | 0.634 | 1.101 | 1.314 | 1.138 |
| 12 | ATOH8 | 0.992 | 0.169 | 2.290 | 0.319 | 1.196 | 0.152 | 2.592 | 0.295 |
| 13 | BCL6B | 0.239 | 1.340 | 0.667 | 0.869 | 0.262 | 1.379 | 0.722 | 1.008 |
| 14 | BDP1 | 0.091 | 1.146 | 0.250 | 0.248 | 0.185 | 1.152 | 0.345 | 0.345 |
| 15 | BENE | 0.427 | 1.030 | 0.814 | 1.006 | 0.587 | 0.970 | 0.722 | 1.096 |
| 16 | C22ORF16 | 1.030 | 0.813 | 0.755 | 0.735 | 1.410 | 0.936 | 0.768 | 0.912 |
| 17 | CAMP | 0.824 | 0.702 | 0.620 | 0.917 | 0.959 | 0.853 | 0.746 | 1.001 |
| 18 | CAPN2 | 0.315 | 0.241 | 0.659 | 1.085 | 0.331 | 0.179 | 0.644 | 1.397 |
| 19 | CCL2 | 0.202 | 1.116 | 0.848 | 0.764 | 0.268 | 1.078 | 0.800 | 0.762 |
| 20 | CDC42 | 0.507 | 0.323 | 0.602 | 0.695 | 0.619 | 0.422 | 0.594 | 0.688 |
| 21 | CDX2 | 0.284 | 0.669 | 0.792 | 1.243 | 0.354 | 0.744 | 1.013 | 1.234 |
| 22 | CGGBP1 | 0.491 | 0.239 | 0.628 | 0.253 | 0.601 | 0.312 | 0.855 | 0.265 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 23 | COL11A1 | 0.714 | 0.877 | 0.498 | 0.661 | 0.850 | 0.967 | 0.456 | 0.657 |
| 24 | COPS4 | 1.737 | 0.923 | 0.635 | 1.140 | 1.654 | 0.797 | 0.514 | 1.057 |
| 25 | CORT | 0.579 | 1.246 | 0.831 | 0.778 | 0.776 | 1.251 | 0.994 | 0.842 |
| 26 | CPSF3 | 1.823 | 0.277 | 2.249 | 1.990 | 1.896 | 0.322 | 2.709 | 2.062 |
| 27 | CREBL2 | 0.746 | 0.290 | 2.050 | 0.289 | 0.820 | 0.274 | 2.044 | 0.285 |
| 28 | CRK7 | 0.810 | 0.493 | 0.378 | 0.404 | 0.854 | 0.442 | 0.414 | 0.376 |
| 29 | CRLF1 | 0.716 | 0.835 | 0.693 | 1.324 | 0.692 | 0.782 | 0.702 | 1.775 |
| 30 | CRSP2 | 1.908 | 0.106 | 1.075 | 0.814 | 1.875 | 0.104 | 0.925 | 0.843 |
| 31 | CTNNA3 | 0.161 | 0.694 | 0.654 | 0.877 | 0.217 | 0.802 | 0.759 | 0.897 |
| 32 | CYBA | 0.792 | 0.734 | 0.653 | 0.956 | 0.876 | 0.807 | 0.703 | 1.057 |
| 33 | DDEF1 | 0.977 | 0.166 | 1.116 | 0.895 | 0.941 | 0.215 | 1.068 | 0.940 |
| 34 | DDIT3 | 0.153 | 0.093 | 0.822 | 1.382 | 0.155 | 0.062 | 0.919 | 1.420 |
| 35 | DEFB126 | 0.261 | 0.166 | 0.819 | 0.581 | 0.346 | 0.215 | 0.908 | 0.743 |
| 36 | DKFZP564B147 | 0.251 | 1.002 | 0.557 | 1.167 | 0.276 | 0.970 | 0.657 | 1.149 |
| 37 | DKFZP564K142 | 0.829 | 0.826 | 1.088 | 0.880 | 0.884 | 0.896 | 1.169 | 1.020 |
| 38 | DRG2 | 0.672 | 2.094 | 1.132 | 1.703 | 0.804 | 2.264 | 1.178 | 1.646 |
| 39 | E4F1 | 0.489 | 0.111 | 1.807 | 0.543 | 0.505 | 0.070 | 2.113 | 0.607 |
| 40 | EDF1 | 0.638 | 0.846 | 0.927 | 0.398 | 0.709 | 0.899 | 0.958 | 0.457 |
| 41 | EIF2B1 | 0.263 | 0.150 | 0.914 | 0.304 | 0.354 | 0.168 | 1.025 | 0.368 |
| 42 | EIF2B2 | 0.150 | 0.214 | 0.383 | 0.124 | 0.118 | 0.208 | 0.386 | 0.083 |
| 43 | EIF2B3 | 0.511 | 0.155 | 0.059 | 0.112 | 0.586 | 0.146 | 0.067 | 0.119 |
| 44 | EIF2B4 | 0.253 | 0.221 | 0.339 | 0.176 | 0.274 | 0.179 | 0.303 | 0.151 |
| 45 | EIF2S2 | 0.109 | 0.181 | 0.274 | 0.071 | 0.090 | 0.164 | 0.260 | 0.053 |
| 46 | ELYS | 0.087 | 0.202 | 0.593 | 0.074 | 0.133 | 0.260 | 0.612 | 0.095 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 47 | ENPP7 | 0.649 | 0.535 | 0.667 | 0.512 | 0.613 | 0.497 | 0.791 | 0.469 |
| 48 | EP300 | 1.609 | 1.045 | 1.165 | 0.965 | 1.308 | 1.017 | 0.996 | 1.074 |
| 49 | ETF1 | 0.217 | 0.523 | 0.171 | 0.078 | 0.201 | 0.471 | 0.135 | 0.069 |
| 50 | FAM19A1 | 0.822 | 0.742 | 0.475 | 0.688 | 0.915 | 0.782 | 0.375 | 0.799 |
| 51 | FLJ20898 | 0.977 | 1.170 | 0.837 | 1.044 | 1.080 | 1.673 | 0.810 | 1.108 |
| 52 | FLJ23447 | 1.069 | 0.582 | 0.764 | 0.582 | 1.224 | 0.546 | 0.772 | 0.570 |
| 53 | FLJ23751 | 0.173 | 0.936 | 0.851 | 0.780 | 0.212 | 0.917 | 0.936 | 0.792 |
| 54 | FLJ25439 | 0.935 | 1.044 | 0.350 | 0.602 | 0.984 | 1.412 | 0.342 | 0.665 |
| 55 | FLJ25952 | 0.338 | 0.285 | 0.754 | 0.423 | 0.313 | 0.344 | 0.848 | 0.441 |
| 56 | FLJ32954 | 0.725 | 0.859 | 1.045 | 1.012 | 0.836 | 1.076 | 1.423 | 1.229 |
| 57 | FLJ38508 | 0.692 | 0.644 | 1.217 | 0.670 | 0.848 | 0.708 | 1.055 | 0.706 |
| 58 | FLJ39743 | 0.630 | 1.150 | 0.634 | 0.856 | 0.709 | 1.107 | 0.838 | 1.029 |
| 59 | FLJ46536 | 0.823 | 0.385 | 0.546 | 0.878 | 0.969 | 0.406 | 0.531 | 1.192 |
| 60 | FLJ90652 | 0.409 | 0.921 | 0.249 | 0.387 | 0.556 | 1.303 | 0.188 | 0.501 |
| 61 | FOXJ3 | 1.201 | 0.473 | 1.576 | 1.274 | 1.419 | 0.599 | 1.711 | 1.293 |
| 62 | FTL | 0.647 | 0.739 | 0.938 | 0.798 | 0.876 | 0.855 | 0.930 | 0.737 |
| 63 | FTSJ1 | 0.717 | 0.742 | 0.512 | 0.665 | 0.624 | 0.796 | 0.574 | 0.659 |
| 64 | FUBP1 | 0.805 | 0.348 | 0.473 | 0.692 | 0.796 | 0.254 | 0.435 | 0.763 |
| 65 | GJA8 | 0.782 | 0.817 | 0.660 | 0.900 | 0.888 | 0.894 | 0.707 | 0.984 |
| 66 | GLRB | 0.779 | 0.826 | 0.778 | 0.906 | 0.782 | 0.951 | 0.856 | 1.067 |
| 67 | GLTSCR1 | 0.661 | 0.853 | 0.645 | 0.641 | 0.723 | 0.945 | 0.620 | 0.592 |
| 68 | GPS1 | 0.484 | 0.504 | 0.531 | 0.511 | 0.383 | 0.459 | 0.495 | 0.517 |
| 69 | GSPT1 | 0.772 | 0.802 | 0.573 | 0.329 | 0.851 | 0.887 | 0.635 | 0.386 |
| 70 | GSTP1 | 0.680 | 0.209 | 1.112 | 1.057 | 0.856 | 0.300 | 1.090 | 1.130 |
| 71 | GUSB | 0.742 | 0.487 | 0.467 | 0.581 | 0.822 | 0.410 | 0.454 | 0.602 |
| 72 | H1FX | 0.697 | 0.383 | 0.162 | 0.584 | 0.774 | 0.491 | 0.204 | 0.621 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 73 | HCFC1 | 0.515 | 0.247 | 0.219 | 0.630 | 0.705 | 0.271 | 0.198 | 0.851 |
| 74 | HELZ | 0.702 | 0.463 | 0.847 | 0.806 | 0.665 | 0.353 | 1.023 | 0.870 |
| 75 | HEMK1 | 0.755 | 0.706 | 0.779 | 0.667 | 0.691 | 0.693 | 0.861 | 0.780 |
| 76 | HES6 | 1.075 | 0.920 | 1.812 | 0.242 | 1.319 | 1.092 | 1.984 | 0.271 |
| 77 | HIVEP3 | 1.586 | 0.609 | 0.763 | 1.172 | 1.780 | 0.646 | 0.935 | 1.346 |
| 78 | HNRPD | 0.688 | 0.946 | 0.685 | 0.629 | 0.756 | 1.078 | 0.855 | 0.623 |
| 79 | HNRPU | 0.279 | 0.280 | 0.892 | 0.451 | 0.221 | 0.189 | 1.211 | 0.529 |
| 80 | IBSP | 0.507 | 0.980 | 0.647 | 0.586 | 0.536 | 0.961 | 0.726 | 0.682 |
| 81 | IGFBP6 | 0.507 | 0.512 | 0.378 | 1.490 | 0.584 | 0.596 | 0.475 | 1.949 |
| 82 | INCA1 | 0.691 | 0.722 | 0.704 | 0.513 | 0.758 | 0.782 | 0.629 | 0.429 |
| 83 | ITSN2 | 0.708 | 1.237 | 0.767 | 0.428 | 0.829 | 1.122 | 0.822 | 0.551 |
| 84 | JMJD2B | 0.988 | 0.660 | 0.528 | 0.431 | 1.151 | 0.649 | 0.830 | 0.474 |
| 85 | KIAA0274 | 1.020 | 0.426 | 0.753 | 0.801 | 1.051 | 0.572 | 0.909 | 0.987 |
| 86 | KIAA1076 | 0.582 | 0.403 | 0.767 | 0.712 | 0.646 | 0.658 | 0.795 | 0.872 |
| 87 | KIR3DL1 | 0.906 | 0.956 | 1.064 | 0.577 | 1.000 | 0.928 | 1.154 | 0.538 |
| 88 | KIRREL2 | 0.631 | 0.675 | 0.784 | 0.635 | 0.658 | 0.674 | 0.856 | 0.656 |
| 89 | KLK5 | 0.479 | 0.323 | 0.611 | 1.029 | 0.555 | 0.513 | 0.787 | 0.981 |
| 90 | KREMEN1 | 0.928 | 0.746 | 1.347 | 0.525 | 0.963 | 0.824 | 1.267 | 0.657 |
| 91 | LARS | 0.389 | 0.808 | 0.636 | 0.325 | 0.316 | 0.833 | 0.553 | 0.247 |
| 92 | LCE1E | 0.721 | 0.724 | 0.469 | 0.183 | 0.724 | 0.854 | 0.802 | 0.318 |
| 93 | LCMR1 | 0.234 | 0.898 | 0.523 | 0.579 | 0.232 | 0.975 | 0.443 | 0.553 |
| 94 | LIF | 0.756 | 1.005 | 0.766 | 0.109 | 0.670 | 1.079 | 0.762 | 0.067 |
| 95 | LOC124245 | 0.481 | 0.678 | 0.327 | 2.563 | 0.386 | 0.534 | 0.413 | 2.904 |
| 96 | LOC374654 | 0.921 | 0.863 | 0.836 | 0.813 | 1.108 | 1.001 | 1.168 | 1.004 |
| 97 | LOC390790 | 0.783 | 0.394 | 0.726 | 0.528 | 0.843 | 0.294 | 0.877 | 0.628 |
| 98 | LOC400221 | 1.033 | 0.658 | 0.862 | 0.954 | 1.569 | 0.616 | 0.961 | 1.083 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 99 | LOC56901 | 0.699 | 1.194 | 0.601 | 0.842 | 0.659 | 1.333 | 0.633 | 0.930 |
| 100 | LPPR2 | 0.767 | 0.553 | 0.666 | 0.256 | 0.813 | 0.723 | 0.649 | 0.391 |
| 101 | LRRC33 | 0.849 | 0.515 | 0.807 | 0.651 | 0.914 | 0.474 | 0.913 | 0.746 |
| 102 | LUC7A | 0.532 | 0.449 | 0.268 | 0.468 | 0.518 | 0.435 | 0.222 | 0.424 |
| 103 | MAP2K7 | 0.811 | 0.619 | 0.339 | 1.385 | 0.728 | 0.693 | 0.448 | 1.204 |
| 104 | MAP3K1 | 0.976 | 0.955 | 0.935 | 0.400 | 0.981 | 1.053 | 0.963 | 0.377 |
| 105 | MBTD1 | 0.685 | 0.734 | 0.895 | 1.014 | 0.677 | 0.779 | 1.077 | 1.189 |
| 106 | MCRS1 | 0.149 | 0.523 | 0.371 | 0.108 | 0.163 | 0.648 | 0.383 | 0.114 |
| 107 | MED28 | 0.928 | 0.218 | 1.093 | 0.626 | 0.826 | 0.320 | 1.012 | 0.795 |
| 108 | MGC10471 | 1.070 | 1.074 | 0.883 | 0.600 | 1.188 | 1.283 | 1.204 | 0.728 |
| 109 | MGC21874 | 0.559 | 0.381 | 0.517 | 0.294 | 0.481 | 0.372 | 0.586 | 0.301 |
| 110 | MGC32871 | 0.701 | 0.789 | 0.580 | 0.925 | 0.800 | 0.827 | 0.559 | 0.954 |
| 111 | MGC39827 | 0.413 | 0.711 | 0.486 | 0.557 | 0.340 | 0.758 | 0.464 | 0.510 |
| 112 | MGC8902 | 1.102 | 1.046 | 0.620 | 0.411 | 1.127 | 0.935 | 0.774 | 0.552 |
| 113 | MMP15 | 0.285 | 0.374 | 0.775 | 0.836 | 0.424 | 0.497 | 0.887 | 0.985 |
| 114 | MMP24 | 0.114 | 0.511 | 0.678 | 0.344 | 0.083 | 0.549 | 0.740 | 0.320 |
| 115 | MOCS1 | 0.569 | 0.823 | 0.737 | 0.807 | 0.381 | 0.838 | 0.659 | 0.781 |
| 116 | MR1 | 0.730 | 0.640 | 0.366 | 0.883 | 0.704 | 0.746 | 0.369 | 1.110 |
| 117 | MVP | 0.798 | 0.256 | 0.876 | 0.660 | 0.928 | 0.262 | 0.916 | 0.724 |
| 118 | NANOG | 0.176 | 0.361 | 0.134 | 0.407 | 0.136 | 0.403 | 0.101 | 0.551 |
| 119 | NCBP1 | 0.422 | 0.667 | 0.353 | 0.178 | 0.411 | 0.680 | 0.345 | 0.120 |
| 120 | NEUROD2 | 0.939 | 0.652 | 0.869 | 0.746 | 1.085 | 0.666 | 0.956 | 0.902 |
| 121 | NFKB1 | 1.320 | 1.053 | 1.008 | 1.248 | 1.114 | 1.060 | 1.067 | 1.208 |
| 122 | NFRKB | 0.464 | 0.304 | 1.401 | 0.290 | 0.465 | 0.405 | 1.597 | 0.356 |
| 123 | NPEPL1 | 0.439 | 0.646 | 1.078 | 0.841 | 0.514 | 0.746 | 1.350 | 0.861 |
| 124 | NUDT8 | 0.543 | 0.763 | 1.010 | 0.796 | 0.571 | 0.728 | 1.127 | 0.954 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 125 | NUP107 | 0.275 | 0.025 | 0.590 | 0.019 | 0.365 | 0.034 | 0.647 | 0.042 |
| 126 | NXF1 | 0.062 | 0.042 | 0.551 | 0.009 | 0.097 | 0.061 | 0.673 | 0.015 |
| 127 | OACT1 | 0.630 | 1.119 | 0.656 | 0.528 | 0.676 | 1.217 | 0.885 | 0.657 |
| 128 | ODF2 | 0.725 | 1.089 | 0.724 | 0.680 | 0.889 | 1.572 | 0.670 | 0.507 |
| 129 | P29 | 0.972 | 0.919 | 0.666 | 0.815 | 1.036 | 1.001 | 0.743 | 0.862 |
| 130 | PCF11 | 0.248 | 0.329 | 0.716 | 0.267 | 0.179 | 0.299 | 0.728 | 0.258 |
| 131 | PDZK11 | 0.471 | 0.946 | 0.336 | 0.607 | 0.517 | 0.874 | 0.414 | 0.736 |
| 132 | PHB | 0.119 | 0.105 | 0.154 | 0.134 | 0.081 | 0.086 | 0.128 | 0.102 |
| 133 | PITX1 | 0.707 | 0.990 | 0.365 | 0.939 | 0.750 | 1.014 | 0.543 | 1.026 |
| 134 | PKP1 | 0.641 | 1.216 | 1.814 | 0.825 | 0.773 | 1.201 | 1.337 | 0.997 |
| 135 | POLH | 0.688 | 0.835 | 1.026 | 0.959 | 0.749 | 0.804 | 1.232 | 1.116 |
| 136 | OCT4 | 0.239 | 0.645 | 0.357 | 0.156 | 0.182 | 0.421 | 0.205 | 0.045 |
| 137 | PPAPDC2 | 0.452 | 0.535 | 0.487 | 0.539 | 0.461 | 0.421 | 0.489 | 0.559 |
| 138 | PPP2R3A | 1.711 | 0.557 | 1.350 | 0.363 | 2.043 | 0.521 | 1.497 | 0.502 |
| 139 | PRDM14 | 0.779 | 0.423 | 1.433 | 1.097 | 0.683 | 0.383 | 1.624 | 1.064 |
| 140 | PRDM9 | 0.326 | 0.475 | 0.643 | 0.555 | 0.349 | 0.603 | 0.749 | 0.627 |
| 141 | PRDX6 | 0.890 | 1.339 | 0.478 | 0.930 | 1.031 | 1.153 | 0.546 | 1.013 |
| 142 | PRO2730 | 0.717 | 0.096 | 0.526 | 0.349 | 0.785 | 0.131 | 0.654 | 0.371 |
| 143 | PROP1 | 0.484 | 1.273 | 0.212 | 1.872 | 0.571 | 1.229 | 0.202 | 2.345 |
| 144 | PSMD2 | 0.067 | 0.038 | 0.063 | 0.054 | 0.021 | 0.020 | 0.060 | 0.026 |
| 145 | PSTPIP2 | 0.440 | 0.850 | 0.775 | 0.296 | 0.496 | 0.874 | 0.816 | 0.414 |
| 146 | PXN | 0.458 | 0.727 | 0.750 | 0.572 | 0.458 | 0.688 | 0.721 | 0.589 |
| 147 | RALGDS | 0.526 | 0.627 | 0.832 | 0.741 | 0.789 | 0.759 | 1.022 | 0.945 |
| 148 | RASEF | 1.154 | 1.011 | 0.850 | 0.927 | 1.475 | 1.000 | 0.880 | 1.023 |
| 149 | RBM17 | 0.442 | 0.667 | 0.746 | 0.334 | 0.640 | 0.911 | 0.791 | 0.437 |
| 150 | REA | 0.084 | 0.084 | 0.091 | 0.083 | 0.160 | 0.101 | 0.099 | 0.136 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 151 | RHOA | 1.467 | 1.032 | 0.890 | 0.548 | 1.145 | 0.950 | 0.757 | 0.597 |
| 152 | RICTOR | 0.886 | 1.294 | 0.568 | 1.480 | 0.977 | 1.698 | 0.643 | 1.705 |
| 153 | RPESP | 0.808 | 1.009 | 0.665 | 0.664 | 0.944 | 1.237 | 0.737 | 0.673 |
| 154 | RRAS | 0.405 | 0.425 | 0.725 | 0.632 | 0.468 | 0.606 | 0.848 | 0.680 |
| 155 | SAMD7 | 0.641 | 0.537 | 0.739 | 0.777 | 0.655 | 0.571 | 0.714 | 0.903 |
| 156 | SERPINB2 | 0.884 | 0.703 | 0.394 | 0.671 | 1.070 | 0.773 | 0.445 | 0.794 |
| 157 | SERTAD2 | 0.818 | 1.006 | 0.467 | 1.442 | 1.121 | 1.242 | 0.509 | 1.564 |
| 158 | SF3A1 | 0.069 | 0.094 | 0.134 | 0.032 | 0.078 | 0.083 | 0.133 | 0.022 |
| 159 | SF3A3 | 0.790 | 0.013 | 0.017 | 0.460 | 0.912 | 0.031 | 0.043 | 0.575 |
| 160 | SFPQ | 0.011 | 0.016 | 0.039 | 0.007 | 0.094 | 0.126 | 0.145 | 0.078 |
| 161 | SFRS3 | 0.191 | 0.061 | 0.633 | 0.340 | 0.109 | 0.062 | 0.470 | 0.289 |
| 162 | SFXN3 | 0.987 | 0.965 | 0.683 | 0.818 | 1.270 | 1.314 | 0.906 | 1.007 |
| 163 | SOAT2 | 0.624 | 0.844 | 0.440 | 0.903 | 0.662 | 0.989 | 0.395 | 0.987 |
| 164 | SON | 0.069 | 0.028 | 0.062 | 0.025 | 0.036 | 0.007 | 0.037 | 0.006 |
| 165 | SOX14 | 1.195 | 0.402 | 1.474 | 0.538 | 1.460 | 0.467 | 1.770 | 0.729 |
| 166 | SPI1 | 1.061 | 0.494 | 0.357 | 0.464 | 1.149 | 0.657 | 0.412 | 0.602 |
| 167 | STK4 | 1.299 | 0.797 | 1.027 | 0.805 | 1.302 | 0.827 | 1.055 | 0.827 |
| 168 | SUV39H2 | 2.494 | 2.420 | 0.390 | 0.445 | 2.612 | 2.628 | 0.415 | 0.507 |
| 169 | SYNCRIP | 0.661 | 0.782 | 0.228 | 0.735 | 0.695 | 0.888 | 0.305 | 0.856 |
| 170 | SYTL4 | 0.776 | 1.122 | 0.659 | 0.952 | 0.798 | 1.026 | 0.643 | 0.901 |
| 171 | TAF2 | 0.524 | 0.359 | 0.904 | 0.950 | 0.688 | 0.415 | 0.984 | 1.115 |
| 172 | TAF7 | 0.423 | 0.705 | 0.794 | 0.481 | 0.611 | 0.814 | 0.895 | 0.547 |
| 173 | TBC1D10 | 0.930 | 1.017 | 1.031 | 0.643 | 0.865 | 1.197 | 1.346 | 0.626 |
| 174 | TCL1A | 0.453 | 0.824 | 0.951 | 0.868 | 0.548 | 0.896 | 0.961 | 0.984 |
| 175 | THRAP2 | 1.338 | 1.501 | 1.466 | 1.314 | 1.273 | 1.277 | 1.432 | 1.604 |
| 176 | THRAP4 | 0.459 | 1.045 | 0.611 | 0.171 | 0.403 | 1.063 | 0.594 | 0.165 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 177 | TMEM14B | 0.583 | 0.303 | 0.635 | 0.518 | 0.628 | 0.252 | 0.685 | 0.473 |
| 178 | TNRC11 | 1.711 | 0.885 | 0.954 | 1.103 | 2.100 | 0.929 | 0.826 | 0.967 |
| 179 | TPD52L1 | 0.903 | 0.286 | 0.613 | 0.582 | 0.959 | 0.261 | 0.617 | 0.676 |
| 180 | TPM1 | 0.739 | 0.675 | 0.508 | 0.356 | 0.850 | 0.710 | 0.590 | 0.386 |
| 181 | TPR | 0.820 | 0.595 | 0.461 | 0.407 | 0.805 | 0.572 | 0.441 | 0.413 |
| 182 | TRIP | 1.951 | 1.997 | 1.526 | 0.212 | 2.153 | 2.048 | 1.784 | 0.202 |
| 183 | TRIP15 | 1.418 | 1.156 | 0.458 | 0.308 | 1.456 | 1.164 | 0.371 | 0.212 |
| 184 | TRPA1 | 0.666 | 0.898 | 0.661 | 0.477 | 0.647 | 1.041 | 0.757 | 0.521 |
| 185 | ULK2 | 0.359 | 0.482 | 0.728 | 2.584 | 0.399 | 0.568 | 0.762 | 2.811 |
| 186 | VGCNL1 | 0.895 | 0.954 | 0.993 | 0.820 | 1.006 | 1.026 | 1.126 | 0.888 |
| 187 | VMP | 0.567 | 0.670 | 0.255 | 0.937 | 0.583 | 0.726 | 0.219 | 1.011 |
| 188 | VWF | 0.770 | 0.618 | 1.152 | 0.877 | 0.746 | 0.651 | 1.131 | 0.812 |
| 189 | XRCC1 | 0.689 | 0.618 | 0.571 | 0.802 | 0.742 | 0.651 | 0.592 | 0.784 |
| 190 | YAP1 | 1.509 | 1.266 | 1.557 | 1.946 | 1.985 | 1.580 | 2.000 | 2.355 |
| 191 | YY1 | 1.298 | 0.336 | 1.153 | 1.778 | 1.279 | 0.315 | 1.060 | 1.550 |
| 192 | ZFP36 | 0.548 | 0.780 | 0.833 | 0.784 | 0.460 | 0.800 | 0.896 | 0.820 |
| 193 | ZFP64 | 2.397 | 1.324 | 0.181 | 0.531 | 2.392 | 1.524 | 0.132 | 0.491 |
| 194 | ZIC4 | 1.189 | 2.135 | 0.951 | 2.805 | 1.328 | 2.251 | 1.127 | 3.401 |
| 195 | ZNF136 | 0.541 | 0.094 | 0.086 | 0.142 | 0.493 | 0.076 | 0.056 | 0.163 |
| 196 | ZNF138 | 0.373 | 0.429 | 0.642 | 0.357 | 0.494 | 0.551 | 0.701 | 0.287 |
| 197 | ZNF206 | 0.861 | 0.209 | 1.856 | 0.398 | 0.929 | 0.231 | 2.116 | 0.468 |
| 198 | ZNF35 | 0.669 | 0.275 | 0.190 | 0.734 | 0.789 | 0.249 | 0.160 | 0.743 |
| 199 | ZNF43 | 0.261 | 0.493 | 0.598 | 1.748 | 0.285 | 0.621 | 0.602 | 1.761 |
| 200 | ZNF434 | 1.015 | 0.353 | 1.556 | 1.090 | 1.234 | 0.360 | 1.879 | 1.248 |

**Table 5: Gene list of positive hits scored by all the different stemness markers of assessment for each of the three hESCs lines**

| H1 hESCs (GFP, OCT4 and NANOG hits) | HES2 hESCs (OCT4 and NANOG hits) | HES3 hESCs (OCT4 and NANOG hits) |
|---|---|---|
| ABP1 | ANKRD1 | ADAMTS1 |
| ABTB1 | ANXA4 | AGPS |
| ADA | ATOH8 | ANGPT4 |
| ADAMTS1 | BCL6B | ANXA4 |
| ANGPT4 | BDP1 | ATOH8 |
| ANXA4 | BENE | BCL6B |
| APLP2 | C22ORF16 | BDP1 |
| BDP1 | CAMP | CAPN2 |
| BENE | CAPN2 | CDC42 |
| C22ORF16 | CDC42 | CDX2 |
| CAPN2 | CDX2 | CGGBP1 |
| CDC42 | CPSF3 | COL11A1 |
| CGGBP1 | CREBL2 | CREBL2 |
| COL11A1 | CRSP2 | CRK7 |
| COPS4 | CYBA | CRLF1 |
| CPSF3 | E4F1 | DDIT3 |
| CREBL2 | EIF2B2 | DEFB126 |
| CRLF1 | EIF2B3 | DKFZP564B147 |
| CRSP2 | EIF2B4 | E4F1 |
| CYBA | EIF2S2 | EDF1 |
| DEFB126 | ELYS | EIF2B1 |
| DKFZP564B147 | EP300 | EIF2B2 |

| | | |
|---|---|---|
| DKFZP564K142 | ETF1 | EIF2B3 |
| E4F1 | FAM19A1 | EIF2B4 |
| EDF1 | FLJ90652 | EIF2S2 |
| EIF2B1 | FOXJ3 | ELYS |
| EIF2B2 | FTL | ENPP7 |
| EIF2B3 | FTSJ1 | ETF1 |
| EIF2B4 | GSPT1 | FLJ23447 |
| EIF2S2 | GUSB | FLJ25439 |
| ELYS | H1FX | FLJ25952 |
| ENPP7 | HCFC1 | FLJ38508 |
| ETF1 | HELZ | FLJ46536 |
| FAM19A1 | HEMK1 | FLJ90652 |
| FLJ20898 | HES6 | FTSJ1 |
| FLJ23447 | HIVEP3 | FUBP1 |
| FLJ25439 | HNRPU | GLTSCR1 |
| FLJ32954 | IGFBP6 | GPS1 |
| FLJ46536 | JMJD2B | GSPT1 |
| FLJ90652 | KIAA0274 | GUSB |
| FTSJ1 | KLK5 | H1FX |
| GJA8 | LCMR1 | HCFC1 |
| GLRB | LIF | HELZ |
| GLTSCR1 | LOC390790 | HEMK1 |
| GPS1 | MGC10471 | HNRPU |
| GUSB | MGC21874 | IBSP |
| H1FX | MGC39827 | IGFBP6 |
| HCFC1 | MMP15 | INCA1 |

| | | |
|---|---|---|
| HELZ | MR1 | JMJD2B |
| HEMK1 | NANOG | KIAA1076 |
| HES6 | NCBP1 | KIRREL2 |
| HNRPU | NFRKB | KLK5 |
| IGFBP6 | NPEPL1 | LARS |
| INCA1 | NUP107 | LCE1E |
| KIR3DL1 | NXF1 | LCMR1 |
| KIRREL2 | PCF11 | LOC124245 |
| LARS | PDZK11 | LOC390790 |
| LCMR1 | PHB | LOC56901 |
| LIF | OCT4 | LPPR2 |
| LOC124245 | PRDM14 | LRRC33 |
| LOC374654 | PRDM9 | LUC7A |
| LOC390790 | PRO2730 | MAP2K7 |
| LOC400221 | PROP1 | MCRS1 |
| LOC56901 | PSMD2 | MGC21874 |
| LRRC33 | PSTPIP2 | MGC39827 |
| LUC7A | PXN | MMP15 |
| MBTD1 | RASEF | MMP24 |
| MCRS1 | REA | MOCS1 |
| MGC10471 | SF3A1 | MR1 |
| MGC21874 | SF3A3 | MVP |
| MGC32871 | SFPQ | NANOG |
| MGC39827 | SFRS3 | NCBP1 |
| MMP15 | SON | NFRKB |
| MOCS1 | SOX14 | NPEPL1 |

| | | |
|---|---|---|
| MR1 | SUV39H2 | NUP107 |
| MVP | TAF7 | NXF1 |
| NANOG | TBC1D10 | OACT1 |
| NCBP1 | TCL1A | ODF2 |
| NEUROD2 | TNRC11 | PCF11 |
| NFRKB | TPR | PDZK11 |
| NPEPL1 | VWF | PHB |
| NUDT8 | ZFP64 | PITX1 |
| NUP107 | ZIC4 | OCT4 |
| NXF1 | ZNF136 | PPAPDC2 |
| ODF2 | ZNF206 | PPP2R3A |
| PCF11 | ZNF434 | PRDM9 |
| PDZK11 | | PRO2730 |
| PHB | | PROP1 |
| POLH | | PSMD2 |
| OCT4 | | PSTPIP2 |
| PPAPDC2 | | PXN |
| PRDM14 | | RBM17 |
| PROP1 | | REA |
| PSMD2 | | RPESP |
| PXN | | RRAS |
| RASEF | | SAMD7 |
| REA | | SF3A1 |
| RHOA | | SF3A3 |
| RPESP | | SFPQ |
| SAMD7 | | SFRS3 |

| | | |
|---|---|---|
| SERPINB2 | | SOAT2 |
| SF3A1 | | SON |
| SF3A3 | | SOX14 |
| SFPQ | | SPI1 |
| SFRS3 | | SUV39H2 |
| SFXN3 | | SYNCRIP |
| SOAT2 | | TAF2 |
| SON | | TAF7 |
| SOX14 | | THRAP4 |
| SUV39H2 | | TMEM14B |
| TAF2 | | TPD52L1 |
| TBC1D10 | | TPM1 |
| THRAP2 | | TPR |
| TMEM14B | | TRIP15 |
| TNRC11 | | TRPA1 |
| TPD52L1 | | ULK2 |
| TPM1 | | VMP |
| TPR | | XRCC1 |
| TRIP | | ZFP64 |
| TRPA1 | | ZNF136 |
| VMP | | ZNF138 |
| YY1 | | ZNF206 |
| ZFP36 | | ZNF35 |
| ZIC4 | | ZNF43 |
| ZNF206 | | |
| ZNF35 | | |

| | | |
|---|---|---|
| ZNF434 | | |

**Table 6: Gene list of consolidated positive hits identified by OCT4 reduction in all three hESC lines**

| No. | Gene |
|---|---|
| 1 | AGPS |
| 2 | ANXA4 |
| 3 | ATOH8 |
| 4 | BDP1 |
| 5 | BENE |
| 6 | CAPN2 |
| 7 | CDC42 |
| 8 | CDX2 |
| 9 | COL11A1 |
| 10 | CREBL2 |
| 11 | CRK7 |
| 12 | DDIT3 |
| 13 | DEFB126 |
| 14 | E4F1 |
| 15 | EDF1 |
| 16 | EIF2B1 |
| 17 | EIF2B2 |
| 18 | EIF2B3 |
| 19 | EIF2B4 |
| 20 | EIF2S2 |
| 21 | ELYS |
| 22 | ENPP7 |

| | |
|---|---|
| 23 | ETF1 |
| 24 | FLJ90652 |
| 25 | FTSJ1 |
| 26 | FUBP1 |
| 27 | GLTSCR1 |
| 28 | GPS1 |
| 29 | GSPT1 |
| 30 | GUSB |
| 31 | H1FX |
| 32 | HCFC1 |
| 33 | HELZ |
| 34 | HEMK1 |
| 35 | HNRPU |
| 36 | IGFBP6 |
| 37 | INCA1 |
| 38 | JMJD2B |
| 39 | KLK5 |
| 40 | LARS |
| 41 | LCE1E |
| 42 | LCMR1 |
| 43 | LIF |
| 44 | LOC124245 |
| 45 | LOC390790 |
| 46 | LOC56901 |
| 47 | LPPR2 |
| 48 | LUC7A |

| | |
|---|---|
| 49 | MGC21874 |
| 50 | MGC39827 |
| 51 | MMP15 |
| 52 | MMP24 |
| 53 | MOCS1 |
| 54 | MR1 |
| 55 | NANOG |
| 56 | NCBP1 |
| 57 | NFRKB |
| 58 | NPEPL1 |
| 59 | NUP107 |
| 60 | NXF1 |
| 61 | ODF2 |
| 62 | PCF11 |
| 63 | PDZK11 |
| 64 | PHB |
| 65 | OCT4 |
| 66 | PPAPDC2 |
| 67 | PRDM14 |
| 68 | PRDM9 |
| 69 | PRO2730 (WDR82) |
| 70 | PROP1 |
| 71 | PSMD2 |
| 72 | PSTPIP2 |
| 73 | PXN |
| 74 | REA |

| | |
|---|---|
| 75 | RPESP |
| 76 | RRAS |
| 77 | SAMD7 |
| 78 | SF3A1 |
| 79 | SF3A3 |
| 80 | SFPQ |
| 81 | SFRS3 |
| 82 | SOAT2 |
| 83 | SON |
| 84 | SOX14 |
| 85 | SPI1 |
| 86 | SUV39H2 |
| 87 | TAF7 |
| 88 | TPD52L1 |
| 89 | TPR |
| 90 | ZFP64 |
| 91 | ZNF136 |
| 92 | ZNF206 |
| 93 | ZNF43 |

**Table 7: Gene list of consolidated positive hits identified by NANOG reduction in all three hESC lines**

| No. | Gene |
|---|---|
| 1 | ANXA4 |
| 2 | BDP1 |
| 2 | CAPN2 |
| 3 | CDC42 |
| 3 | CGGBP1 |
| 4 | CREBL2 |
| 4 | CYBA |
| 5 | E4F1 |
| 5 | EIF2B2 |
| 6 | EIF2B3 |
| 6 | EIF2B4 |
| 7 | EIF2S2 |
| 7 | ELYS |
| 8 | ETF1 |
| 8 | FAM19A1 |
| 9 | FLJ90652 |
| 9 | FTSJ1 |
| 10 | GUSB |
| 10 | H1FX |
| 11 | HCFC1 |
| 11 | HELZ |
| 12 | HEMK1 |
| 12 | HNRPU |
| 13 | IGFBP6 |

| | |
|---|---|
| 13 | LCMR1 |
| 14 | LOC390790 |
| 14 | MCRS1 |
| 15 | MGC21874 |
| 15 | MGC39827 |
| 16 | MMP15 |
| 16 | MR1 |
| 17 | NANOG |
| 17 | NCBP1 |
| 18 | NFRKB |
| 18 | NPEPL1 |
| 19 | NUP107 |
| 19 | NXF1 |
| 20 | PCF11 |
| 20 | PDZK11 |
| 21 | PHB |
| 21 | OCT4 |
| 22 | PROP1 |
| 22 | PSMD2 |
| 23 | PXN |
| 23 | REA |
| 24 | SF3A1 |
| 24 | SF3A3 |
| 25 | SFPQ |
| 25 | SFRS3 |
| 26 | SON |

| | |
|---|---|
| 26 | SOX14 |
| 27 | SUV39H2 |
| 27 | TPR |
| 28 | ZNF206 |

# 9. LIST OF MY PUBLICATIONS

**Stem cell genome-to-systems biology**
Chia NY, Ng HH
Wiley Interdiscip Rev Syst Biol Med. 2011 Apr 11. doi: 10.1002/wsbm.151.
PMID: 21485016

**A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity.**
Chia NY, Chan YS, Feng B, Lu X, Orlov YL, Moreau D, Kumar P, Yang L, Jiang J, Lau MS, Huss M, Soh BS, Kraus P, Li P, Lufkin T, Lim B, Clarke ND, Bard F, Ng HH.
Nature. 2010 Nov 11;468(7321):316-20. Epub 2010 Oct 17.PMID: 20953172

**Transposable elements have rewired the core regulatory network of human embryonic stem cells.**
Kunarso G, Chia NY, Jeyakani J, Hwang C, Lu X, Chan YS, Ng HH, Bourque G.
Nat Genet. 2010 Jul;42(7):631-4. Epub 2010 Jun 6.PMID: 20526341

**KLF4 and PBX1 directly regulate NANOG expression in human embryonic stem cells.**
Chan KK, Zhang J, Chia NY, Chan YS, Sim HS, Tan KS, Oh SK, Ng HH, Choo AB.
Stem Cells. 2009 Sep;27(9):2114-25.PMID: 19522013

**Cross-species de novo identification of cis-regulatory modules with GibbsModule: application to gene regulation in embryonic stem cells.**
Xie D, Cai J, Chia NY, Ng HH, Zhong S.
Genome Res. 2008 Aug;18(8):1325-35. Epub 2008 May 15.PMID: 18490265