# GENE NETWORKS IN HUMAN STEM CELLS

**KEVIN ANDREW UY GONZALES**

B.S. Biological Sciences (Hons.), NTU

**A THESIS SUBMITTED TO**

**FOR THE DEGREE OF DOCTOR OF PHILOSOPHY**

**NUS GRADUATE SCHOOL FOR INTEGRATIVE**

**SCIENCES AND ENGINEERING**

**NATIONAL UNIVERSITY OF SINGAPORE**

**2015**

**Declaration**

**I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.**

**This thesis has also not been submitted for any degree in any university previously.**

**Kevin Andrew Uy Gonzales**

**28 December 2014**

**Acknowledgments**

I would like to thank Dr Liou Yih Cherng and Dr Ng Huck Hui for their support and guidance throughout this project. I consider myself fortunate to be able to work under such motivating supervisors who drive a spirited environment in their laboratories.

I am deeply grateful to Dr Liang Hongqing, who I have closely worked with throughout the second half of my graduate studies. Her insight and proficiency on the cell cycle has been crucial in propelling this project to completion. I also thank Dr. Frederic Bard, who provided the facilities and expertise to perform high-throughput RNAi screens.

I express my appreciation to the people who have contributed in one way or another towards the completion of this project, including Lim Yee-Siang, Yeo Jia-Chi, Dr Gao Bin, Le Beilin, Tan Zi-Ying, Liu Kaiwen Ivy, Rachel Yuen and Lin Muhong. I also thank Chan Yun-Shen, Ng Jia-Hui and Friedrich Sachs, and all the people in the National University of Singapore and the Genome Institute of Singapore who have been around to help throughout my graduate studies. Their suggestions, encouragements and help have made my life as a Ph.D. student enjoyable and fulfilling.

Finally, to my family, who has been supporting me all the way in my journey through science and life.

**Table of Contents**

**Summary**

The pluripotent state of embryonic stem cells (ESCs) is shaped by complex interactions between multiple genes, transcripts and proteins that form the ESC regulatory network. During differentiation, ESCs exit from pluripotency by shutting down this network. As the potential of human ESCs (hESCs) lies in their ability to differentiate into any specific somatic lineage, it is imperative to understand the genetic regulations governing the exit from pluripotency. In this thesis, we report the first high-throughput RNA interference screen for factors crucial for the exit from pluripotency of hESCs in a comprehensive set of differentiation conditions. This systematic study enabled the identification of both centrally important and context-dependent processes for the exit from pluripotency. Strikingly, we found that the cell cycle is a critical process deterministically regulating the exit from pluripotency. Cell cycle factors that primarily operate during the S and G2 phases, such as the ATM/ATR pathway and Cyclin B1, actively promote the pluripotent state and inhibit differentiation. Altogether, this study uncovers a multitude of novel regulators of the exit from pluripotency, and establishes a new paradigm where the pluripotency network is hardwired into the specific cell cycle pathways.

## List of Tables

**List of Figures**

## Introduction

Proper growth and development of an organism is contingent on the precise balance of cells' decision between proliferation and specialization. The ability of cells to proliferate is regulated by the complex regulatory program known as the cell cycle. Specialization on the other hand is dictated by signalling cues that influence cell fate, and is often accompanied by a cessation of proliferation. Therefore, tight cross-regulation between the cell cycle and fate specification pathways underlies proper development (Budirahardja and Gonczy, 2009). This is exceptionally important for early embryonic development, which entails extensive growth and the earliest fate decisions. For example, pluripotent stem cells across animal species undergo rapid and successive cell divisions in order to give rise to the appropriate number of more specialized progeny (Edgar and Lehner, 1996; Lawson et al., 1991; Murray and Kirschner, 1989; Yarden and Geiger, 1996). This process of self-renewal necessitates a unique transcriptional network governing pluripotency coupled with a distinct cell cycle profile (Singh and Dalton, 2009). Therefore, thorough understanding of both pluripotency and cell cycle regulatory networks is necessary to grasp a complete comprehension of early development.

### I. Embryonic stem cells

Past studies of early development rely heavily on animal models in which it is ethical and convenient to extrinsically manipulate developmental processes at any embryonic stage. However, differences in the developmental progression between humans and even the closest animal models are evident. Therefore, an *in vitro* alternative to study human development would prove to be practical; a niche that was

fulfilled upon the derivation of embryonic stem cells (ESCs) from the inner cell mass of human blastocyst-stage embryos (Cowan et al., 2004; Thomson et al., 1998).

A. Potential applications of ESCs

ESCs have the ability to self-renew indefinitely in an undifferentiated state *in vitro*. They harbour the developmental potency to give rise to all cell types in the embryo proper, and are thus called pluripotent. These properties of ESCs make them ideal sources for generating substantial amounts of specific tissues for tissue replacement therapies (Figure 1), offering great promise in regenerative medicine. ESC-derived tissues can also be used for preliminary assessment of drug efficacy and toxicity, extending the clinical applications of ESCs (Figure 1).

Of course, ESCs also serve as an attractive cellular model to understand normal early developmental processes and development-related diseases (Figure 1). ESCs act as a convenient *in vitro* platform for dissecting the multitude of factors including transcriptional networks, signalling pathways, non-coding RNAs, epigenetic modifications and other processes that facilitate pluripotency maintenance and differentiation. As such, numerous studies have put effort in scrutinising the intricate machinery behind the ESC state.

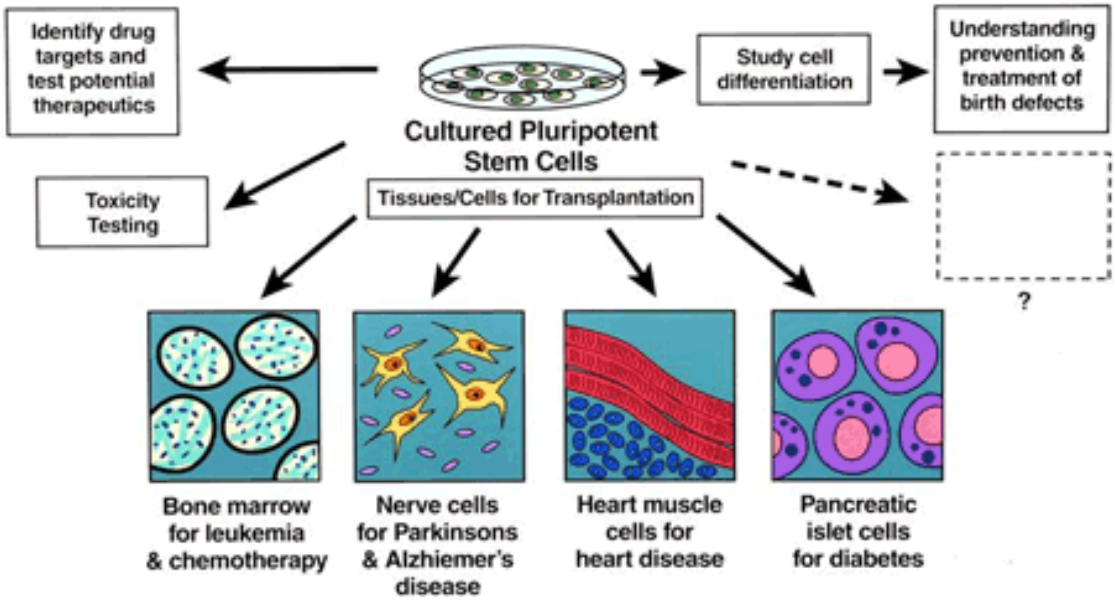



**Figure 1. Potential applications for human ESCs in clinic and research.** Embryonic stem cells can be used as sources of differentiated cells for scientific research, drug testing and transplantation therapy. They can also be used to study early developmental processes and genetic diseases. Additional applications for hESCs could still be developed in the future. Reprinted from the Virtual Genetics Education Centre, Department of Genetics, University of Leicester (http://www2.le.ac.uk/departments/genetics/genie/gs/law/lawembryonic) under Creative Commons License.

B. Challenges for the applications of human ESCs

However, use of human ESCs (hESCs) for clinical and scientific use is met with several obstacles. For example, hESC derivation methods require an exposure to animal-derived products (Vazin and Freed, 2010), which could potentially transfer pathogens that would endanger patients receiving ESC-derived transplantations. There is also the significant barrier of immune-mediated rejection of hESC-derived tissues due to genetic differences between the ESC donor and recipient (Drukker, 2008). Generation of hESC lines also obligates the destruction of the embryonic source (Vazin and Freed, 2010), which imposes ethical and socio-political issues. To address these, several alternative sources of human pluripotent cells have been

developed, including the very promising technique of factor-based reprogramming, which derives induced pluripotent stem cells from somatic cells (Takahashi et al., 2007; Takahashi and Yamanaka, 2006).

However, regardless of source, human pluripotent cells have to overcome the major hurdle of efficient directed differentiation. Despite the availability of a plethora of published differentiation protocols (Tabar and Studer, 2014), most do not provide a differentiation efficiency of 100%, which raises concerns on safety due to the risk of teratoma formation (Ben-David and Benvenisty, 2011). This also poses practical problems, as the application of hESCs in regenerative medicine often requires access to unlimited numbers of pure and functional populations of specific cell types. Thus, a prerequisite for the clinical application of hESCs is an improved understanding of their differentiation, which includes the initial exit from pluripotency (which we discuss in detail in this thesis) and all intermediate steps before the cell commits to its terminal differentiated fate.

C. Defining features of hESCs

To date, hundreds of hESC lines have been derived, necessitating the establishment of the integrity of hESC lines for further downstream use. First of all, hESCs must exhibit basic characteristics including a normal karyotype, indefinite proliferation *in vitro* and recovery from freeze-thaw cycles (Brivanlou et al., 2003).

In culture, hESCs should form flat, compact colonies, in contrast to the more disperse morphologies of their differentiated counterparts. hESCs must also exhibit the unique expression profile landmarked by the high expression of pluripotency-specific markers including *POU5F1*, *NANOG*, *SOX2*, *TDGF1* and *PRDM14* (Assou et al., 2007). hESCs are also characterized by the expression of cell-surface markers

such as TRA-1-60, TRA-1-81, SSEA3 and SSEA4, and the presence of alkaline phosphatase activity (Thomson et al., 1998) .

On top of possessing these molecular signatures, hESCs must have the capability to differentiate into cells from the three major embryonic germ layers *in vitro* and *in vivo*. This can be measured using directed differentiation protocols, embryoid body formation assays or teratoma formation assays. Fulfilment of all these characteristics ensures *bona fide* hESC identity.

## II. Maintenance of the human pluripotent state

The maintenance of ESCs *in vitro* is made possible by manipulating the extrinsic signalling environment, which integrates into the intrinsic transcriptional circuitry of cells. Given the correct conditions, proper signalling events can sustain the transcriptional program of pluripotency, allowing the undifferentiated self-renewal of ESCs.

### A. Regulatory network of hESCs

The sustenance of human pluripotency is mediated by an intricate regulatory network chiefly controlled by a few transcription factors, mainly the triumvirate of OCT4, SOX2 and NANOG (Boyer et al., 2005). These master regulators form multiple regulatory connections with one another (Figure 2A), with secondary transcription factors, and with other effectors such as epigenetic modifiers, transduction pathways and non-coding RNAs that together maintain self-renewal and pluripotency (Figure 2B) (Loh et al., 2011; Ng and Surani, 2011; Young, 2011; Yu and Thomson, 2008).

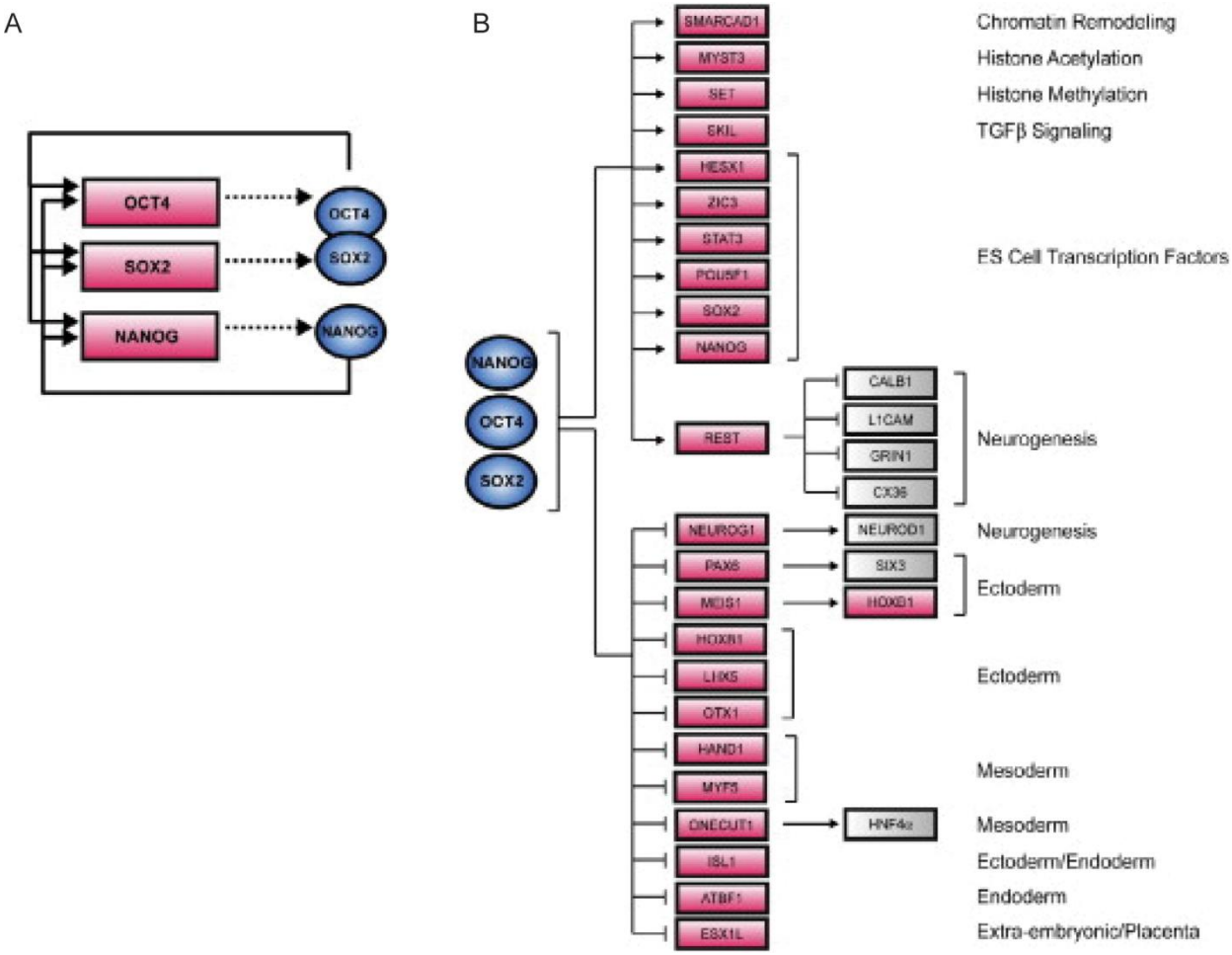



**Figure 2. Regulatory network of hESCs.** (**A**) OCT4, SOX2 and NANOG form multiple auto-regulatory loops that stabilize the core hESC transcriptional network. (**B**) OCT4, SOX2 and NANOG act on multiple downstream elements to promote pluripotency. Reprinted from Cell 122, Boyer, L.A. et al. Core transcriptional regulatory circuitry in human embryonic stem cells, 947-956, Copyright (2005), with permission from Elsevier.

B. Signalling control of hESC maintenance

Sustenance of the pluripotency network is dependent on the extrinsic environment, and is held together by a specific set of external signals. hESCs were initially cultured in the presence of mitotically-inactivated fibroblast feeders, which provided these signalling factors that were important for hESC maintenance (Thomson et al., 1998). Conditioning culture medium with feeders (widely known as 'conditioned medium') allowed an easy alternative to provide such external signalling

molecules in the absence of feeder cells. However, a requirement for defined hESC culture conditions called for efforts to dissect signalling control of hESC maintenance. This led to the identification of basic fibroblast growth factor (bFGF) and Activin/transforming growth factor beta (TGF-β) as the most critical signalling molecules for hESC cultivation (Beattie et al., 2005; Ludwig et al., 2006; Xu et al., 2005). Using this knowledge, chemically-defined media such as mTeSR1 (Ludwig and Thomson, 2007) and Essential 8 (Chen et al., 2011) have been designed for hESC maintenance.

The action of the bFGF pathway in pluripotency has largely been attributed to the downstream MEK-ERK cascade, which inhibits upregulation of neural specification genes (Greber et al., 2011; Li et al., 2007). In addition, the bFGF pathway has been attributed to promote pluripotency through pleiotropic effects including enhancing hESC proliferation, attachment and survival (Eiselleova et al., 2009; Wang et al., 2009). The TGFβ signalling pathway functions by activating ALK4 activity through its receptor, which in turn activates the downstream effector SMAD2/3. Upon activation and dimerization, SMAD2/3 binds the *NANOG* proximal promoter and upregulates *NANOG* expression, among other pluripotency genes, in hESCs (Vallier et al., 2009; Xu et al., 2008). Moreover, both the TGFβ and bFGF signalling pathways antagonize the BMP4 signalling pathway, which acts to repress pluripotency genes through its effectors SMAD1/5/8 (Xu et al., 2008).

It is also important to acknowledge that besides the major governance by the bFGF-MEK and TGFβ pathways, other signalling pathways also contribute to the human pluripotent state. bFGF and additional factors found in human pluripotency media such as insulin-like growth factor activate the phosphoinositide-3-kinase (PI3K) signalling pathway, which in turn cooperates with both the TGFβ pathway and

the bFGF-MEK axis to promote pluripotency (Bendall et al., 2007; McLean et al., 2007; Singh et al., 2012). In addition, the Wnt signalling pathway has been implicated to promote hESC maintenance (Sato et al., 2004), although its function appears to be highly context-dependent (Cai et al., 2007; Davidson et al., 2012; Dravid et al., 2005).

A precise understanding of how exactly these pathways interact to regulate pluripotency has not been achieved due to several confounding factors including the use of various and often disparate culture conditions and experimental designs, and the tendency to zoom in onto a specific pathway with a lack of consideration for its integration with other pathways. Nonetheless, it is clear that the cross-talk between these multiple pathways culminate to determine cell fate in hESCs (Figure 3).

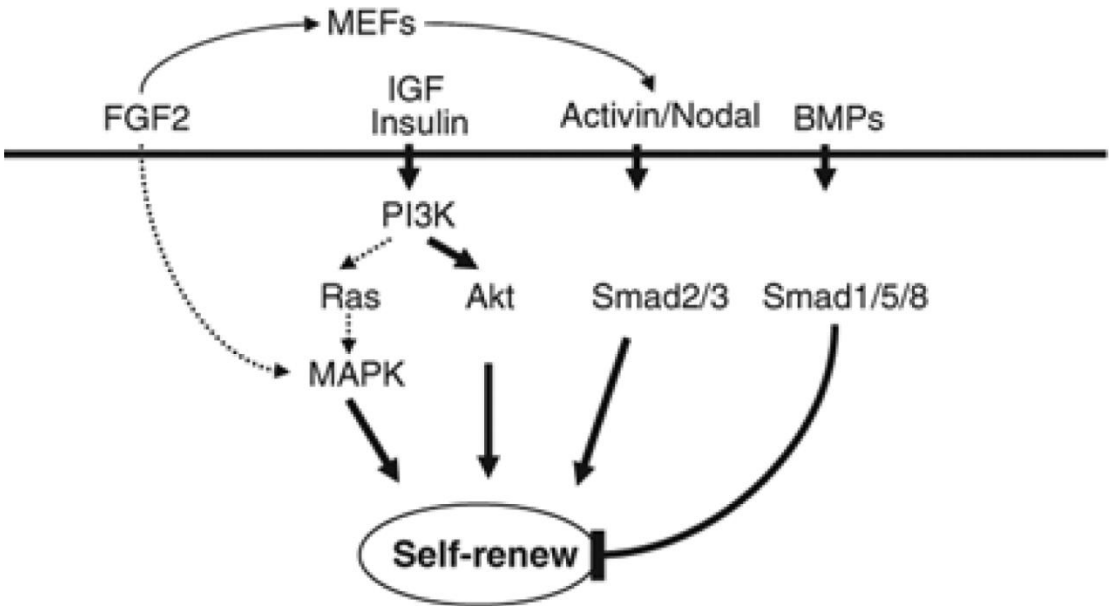



**Figure 3. Signalling pathways governing hESC self-renewal.** The bFGF (FGF2), PI3K and TGFβ (Activin/Nodal) pathways promote maintenance of hESCs, while the BMP4 pathway promotes differentiation. Reprinted by permission from Macmillan Publishers Ltd: Gene Therapy (Ohtsuka, S. and Dalton, S., Molecular and biological properties of pluripotent embryonic stem Cells), copyright (2007).

## III. High-throughput RNA interference screening in ESCs

Many studies on the maintenance of pluripotency have utilized a plethora of research technologies to prove the importance of certain factors in keeping the ESC state. Among these, the method of knocking down gene expression using RNA interference (RNAi) has immensely contributed to the search for crucial players in the regulation of the ESC state.

### A. Biology of RNAi technology

The method of delivering RNA to silence specific genes (Figure 4) was first demonstrated by Mello and colleagues in 1998 using the model system *Caenorhabditis elegans* (Fire et al., 1998). It was shown that injection of double-stranded RNA (dsRNA) into the nematode can potently and specifically suppress gene expression in the animal and its progeny. Subsequent studies discovered that a ribonuclease III enzyme, now known as Dicer, cleaves RNAi-inducing dsRNA into 21 to 22-nucleotide short interfering RNA (siRNA) (Bernstein et al., 2001; Elbashir et al., 2001b; Hammond et al., 2000; Zamore et al., 2000). The siRNA recruits the RNA-induced silencing complex (RISC), and together, they post-transcriptionally attenuate gene expression by binding to the target messenger RNA (mRNA) (Hammond et al., 2000; Zamore et al., 2000). While the siRNA confers sequence specificity to its target mRNA by perfect complementation, the RISC is the effector of gene silencing by recruiting nucleases like the Argonaute proteins for mRNA degradation. The utility of siRNAs in gene silencing extends from *C. elegans* to cells of higher animals including insects, lower mammals and humans (Caplen et al., 2000; Clemens et al., 2000; Elbashir et al., 2001a; Hammond et al., 2000; Ui-Tei et al., 2000). Since then, siRNAs

have become a robust tool in biological research to systematically shut down expression of genes of interest, as it conveniently bypasses the generation of genetically-modified cells or animals.

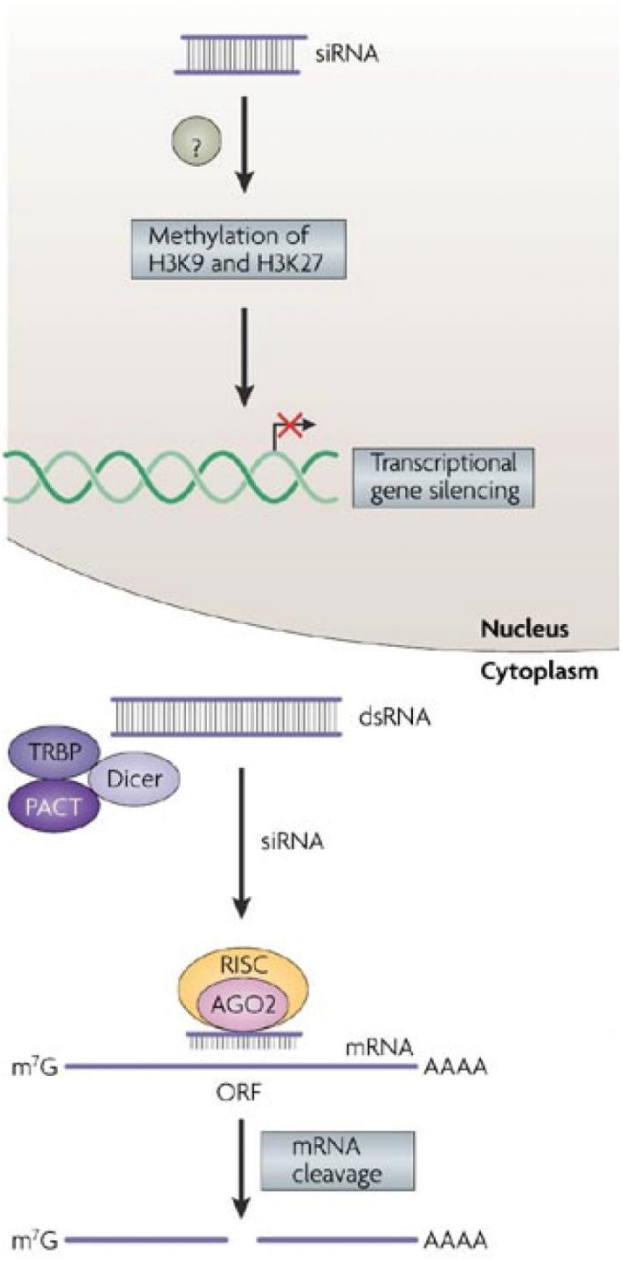



**Figure 4. Mechanisms of RNAi in mammalian cells.** dsRNA is cleaved by Dicer into siRNA, which associates with the RISC. AGO2 in the RISC catalyzes the cleavage of the mRNA complementary to the siRNA. siRNA are also postulated to silence transcription in the nucleus. Reprinted by permission from Macmillan Publishers Ltd: Nat Rev Genet (Kim, D. and Rossi J., Strategies for silencing human disease using RNA interference), copyright (2007).

B. High throughput RNAi screening

In the recent years, technological advances have facilitated the execution and analysis of biological assays in a high-throughput manner. The practicality of RNAi has propelled it to be at the forefront of high-throughput assays by enabling systematic loss-of-function screening. The application of reliable sequencing data to RNAi technology has made it possible to build libraries of RNAi reagents for high-throughput applications (Carpenter and Sabatini, 2004). This extended functional analyses that was only pragmatic for small sets of samples to a colossal number of samples observed in parallel, making it a powerful tool for cell biology. In addition, it can be used in combination with other assays or treatments (known as modifier screens), making it a method of choice for dissecting specific processes or pathways and for the development of therapeutic agents (Echeverri and Perrimon, 2006).

While RNAi screening has many advantages, this comes at the price of the equally numerous parameters that one needs to consider when conducting such experiments (Figure 5) (Sharma and Rao, 2009). These include: (1) choosing whether to do a genome-wide screen or focus on a subset of specific annotated genes, (2) selecting a cellular system to be used to answer a particular biological question, (3) deciding whether to pool all RNAi constructs and perform subsequent sequencing or to individually knockdown genes in an arrayed format, and (4) developing an appropriate and robust readout assay for hit identification. Furthermore, technical details such as the method of RNAi delivery, addition of reliable controls, inclusion of additional filters to eliminate false positives, and application of the appropriate statistical methods for normalization also have to be carefully deliberated.

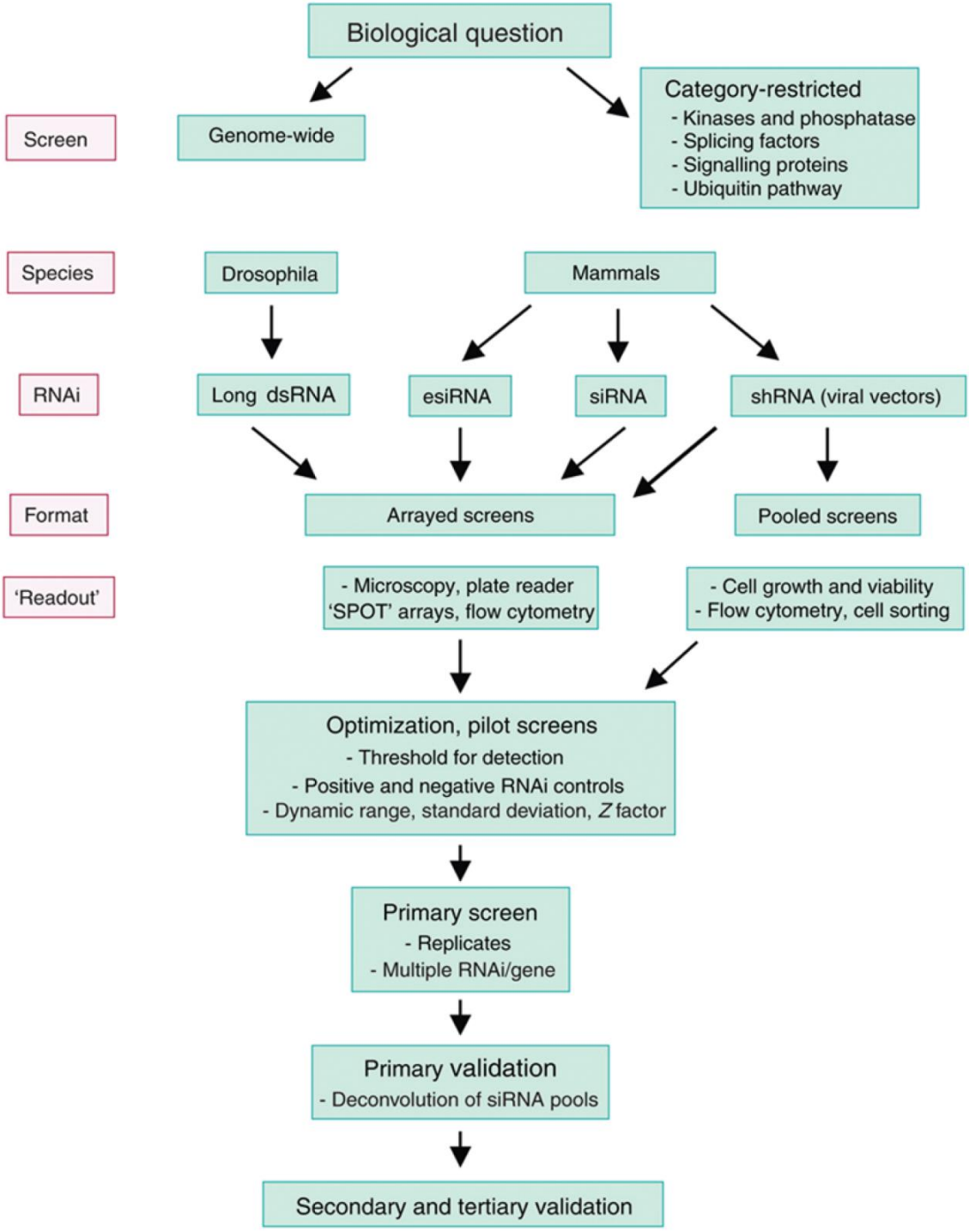



**Figure 5. Considerations for high-throughput RNAi screening.** The process of designing a high-throughput RNAi experiment encompasses multiple decision-making and technical steps, all of which need to be executed correctly for the success of the experiment. Reprinted by permission from Macmillan Publishers Ltd: Nat Immunol (Sharma, S. and Rao A., RNAi screening: tips and techniques), copyright (2009).

Designing a high-throughput assay is greatly dependent on the biological question of interest. Once the proper decisions have been made, high-throughput RNAi screening can be a very powerful discovery tool for both general biological processes and more specific pathways.

C. High-throughput RNAi screens in ESCs

Given the utility of high-throughput RNAi screening, it did not take long before the technique was applied in the context of pluripotency. The first research utilizing this technology to study the regulation of pluripotency was in mouse ESCs (mESCs) (Ivanova et al., 2006). This study revealed the importance for mESC self-renewal of Tbx3, Tcl1 and Esrrb, which acted together in a functional network distinct from the Oct4-Sox2-Nanog cluster (Ivanova et al., 2006). Thereafter, several RNAi screens have been performed, discovering various novel factors and complexes important for mESCs, such as the Paf1 complex (Ding et al., 2009), Cnot3 and Trim28 of the c-Myc cluster (Hu et al., 2009), and the mediator and cohesin complexes (Kagey et al., 2010).

A similar strategy of high-throughput RNAi screening has been adopted in hESCs to uncover regulators of maintenance of the pluripotent state in the human system (Chia et al., 2010). Using this approach, new factors such as PRDM14, NFRKB, the INO80 complex and the COP9 complex have been identified to play a role in hESCs. Moreover, dissimilarities in the pluripotency transcriptional circuitry between the mouse and human species were revealed in this study, highlighting the importance of species-specific differences in the network regulating pluripotency.

The various studies described above demonstrate the success of high-throughput RNAi screening in interrogating noteworthy biological questions, including that of pluripotency. The findings from these studies collectively reaffirm the concepts of multifaceted regulation of complex processes such as the maintenance of pluripotency, and the interdependence of regulatory functions between various pathways and complexes. More importantly, this platform enabled the identification of novel factors and pathways crucial for the regulation of the pluripotent state.

## IV. The exit from pluripotency

Through genome-wide RNAi screens and many other molecular tools that are available today, the regulation of pluripotency and the factors governing the self-renewing state of ESCs is relatively well-defined. However, the maintenance of pluripotency is only one side of the coin; ESC fate is determined by a balance between ESC maintenance and differentiation signals from the cell culture environment (Loh et al., 2011; Ng and Surani, 2011; Young, 2011; Yu and Thomson, 2008). A loss of ESC maintenance signals or dominance of differentiation signals can initiate the process of differentiation by driving stem cells out of the pluripotent ESC state (Figure 6). During this process, transcriptional networks conferring ESC identity need to be shut down and subsequently, lineages have to be determined. While many studies have identified factors for the maintenance of pluripotency and also for the determination of specific cell fates, the knowledge on the machinery behind the initiation of exit from the pluripotent state is scarce.

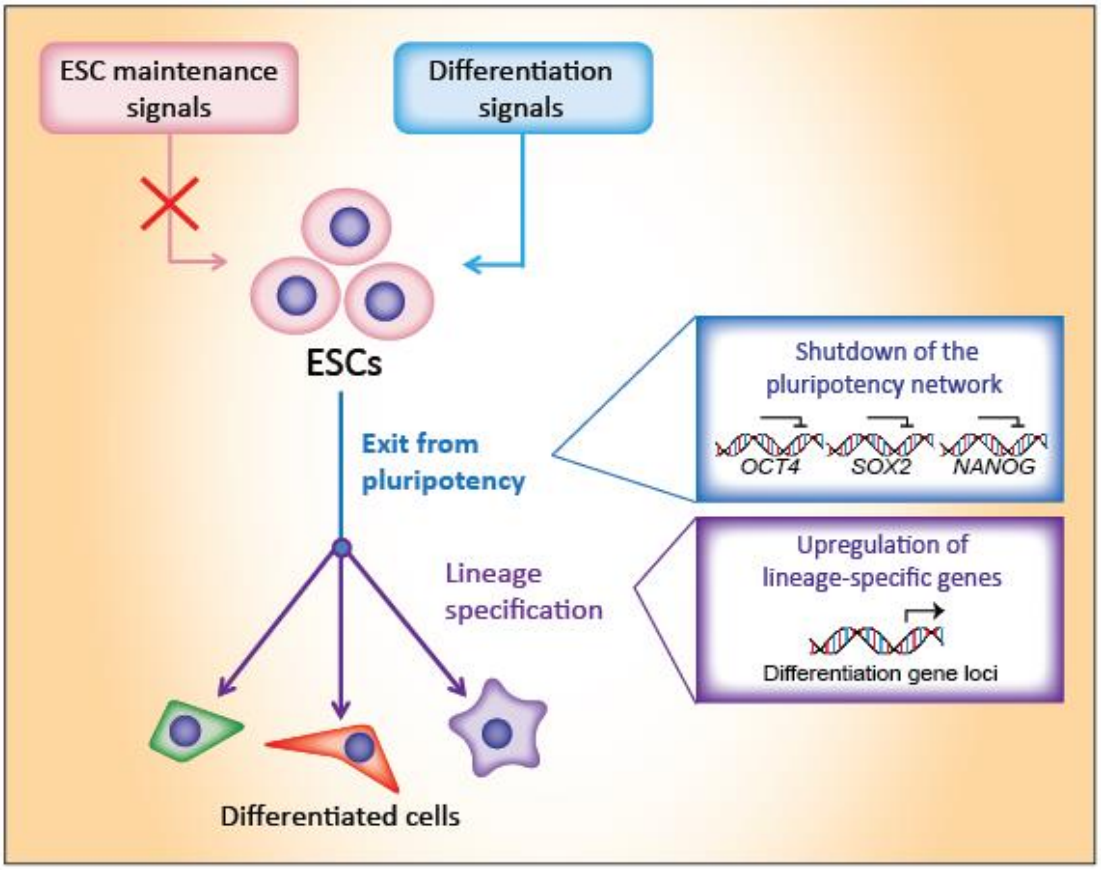



**Figure 6. The process of differentiation.** In the absence of self-renewal signals or in the presence of differentiation signals, ESCs undergo differentiation. This process starts with the exit from pluripotency, characterized by the shutdown of the pluripotency network. Cells then acquire a new transcriptional program in the process of lineage specification.

It is important to clarify that the exit from pluripotency is a distinct phenomenon from lineage specification (Figure 6). Whereas lineage specification marks the activation of the transcriptional program of a particular cell fate, the exit from pluripotency is defined here by the shutting down of the ESC program. It is important to understand the system that maintains the pluripotent state in order to fully harness the potential of ESCs in research and medicine, as cultivation of high-quality ESCs in large quantities is ideal for the mass production of specialized tissues. However, it is equally imperative to learn about the system that pushes ESCs out of

the pluripotent state and enables them to adopt a new program of their chosen lineage, as most applications of hESCs require precise control of their differentiation into specific somatic cell types.

The intrinsic and extrinsic signals regulating the maintenance of the ESC state *in vitro* are reasonably established. Using this knowledge, the machinery responsible for the exit from the pluripotent state can be studied by pushing ESCs out of self-renewal. This can be done by removal of essential growth factors, disruption of the transcriptional circuitry sustaining pluripotency or addition of differentiation inducers. These techniques have been applied in mESCs to identify currently known factors involved in the exit from pluripotency, including Tcf3 (Cole et al., 2008; Guo et al., 2011), Mp1 (Westerman et al., 2011), Dgcr8 (Wang et al., 2007), Dicer (Kanellopoulou et al., 2005), Dpy30 (Jiang et al., 2011), Jarid/Jumonji (Peng et al., 2009) and Spi1 (Abujarour et al., 2010).

RNAi screening has also been applied to study this question by Smith and colleagues on mESCs (Betschinger et al., 2013). In their study, they discovered that factors involved in the export of the transcription factor Tfe3 from the nucleus are essential for the exit from pluripotency. They similarly explored the regulation of exit from pluripotency in haploid mESCs through a piggyBac mutagenesis approach (Leeb et al., 2014), and identified a role for Zfp706 and Pum1 in the exit from pluripotency. These high-throughput studies not only reveal multiple levels of regulation in the exit from pluripotency, but also provide a great resource for future examination of mESC differentiation.

However, these studies have investigated the exit from pluripotency only in the context of mESCs, and no similar study has been conducted in hESCs to the best of our knowledge. Given that the regulation of pluripotency in mESCs is very distinct

from hESCs (Gonzales and Ng, 2011; Hanna et al., 2010; Nichols and Smith, 2009), studies conducted on hESCs are necessary to obtain accurate knowledge about how the hESC regulatory network is shut down for complete and efficient differentiation to occur; this gap in knowledge is addressed in this thesis.

## V. The cell cycle in pluripotency

Cell division requires the replication of the DNA blueprint, followed by the segregation of the replicated chromosomes into the two daughter cells. This requires elaborate coordination of many processes ultimately governed by the cell cycle. The cell cycle is divided into four phases: the S phase wherein DNA replication occurs, the M phase wherein cellular division happens, and the two gap phases wherein cells prepare all necessities before the S and M phases (Vermeulen et al., 2003).

### A. Regulation of cell cycle progression

The key regulatory proteins that govern cell cycle progression are cyclins and their interacting cyclin-dependent kinases (CDKs), with different CDKs active during each cell cycle phase (Figure 7). While CDK protein levels are stable across the cell cycle, their activity is controlled by their specific partner cyclins, whose protein levels rise and fall during the course of the cell cycle. Cyclin D is required for progression through the G1 phase by activating CDK4/6 activity (Sherr, 1994). Cyclin E peaks during the G1/S transition, activating CDK2 for entry to the S phase (Ohtsubo et al., 1995). Cyclin A expression increases at the S phase when it partners with CDK2, and persists until entry into mitosis when it partners with CDK1 (Girard et al., 1991; Walker and Maller, 1991). CDK1 activity is also enhanced by Cyclin B expression which peaks at the G2/M transition (King et al., 1994).

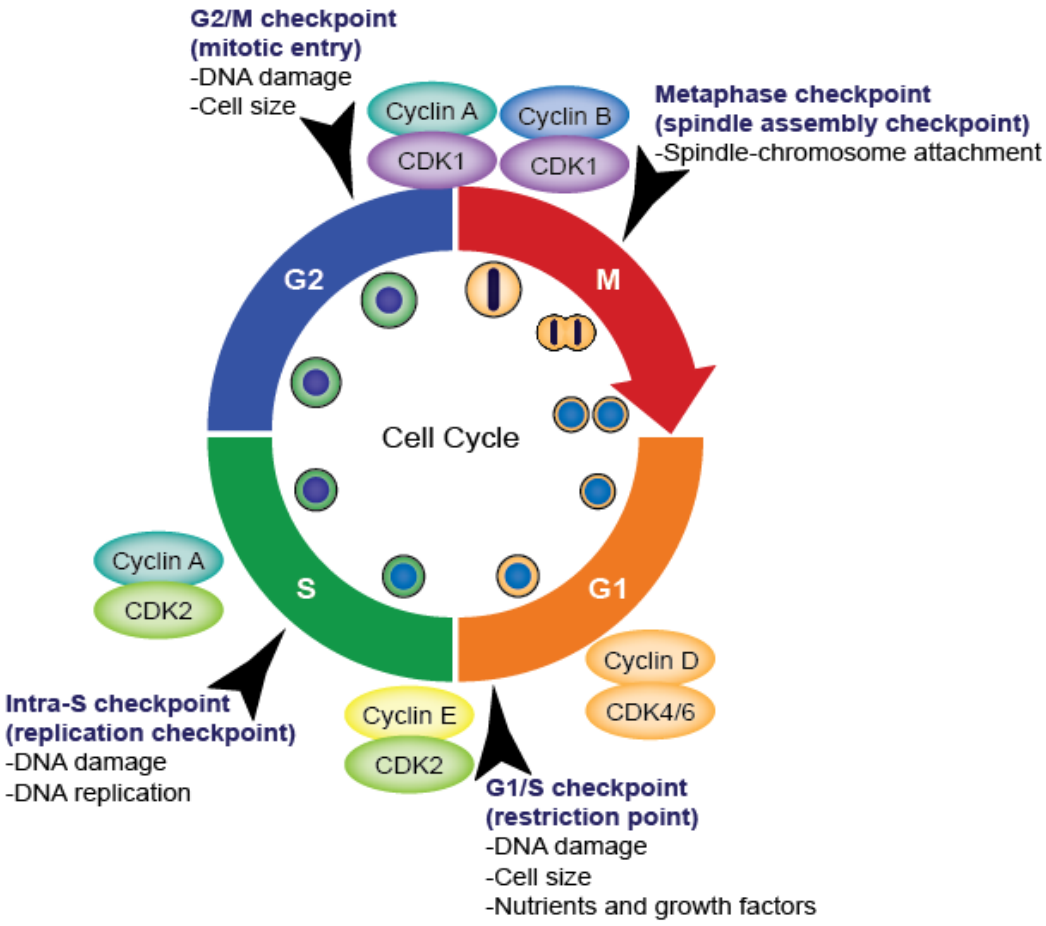



**Figure 7. The cell cycle and its checkpoints.** The cell cycle is divided into 4 phases. Progression through these phases is driven by various Cyclin-CDK complexes. Each phase has a checkpoint wherein cell cycle progression can be halted in case an error is detected.

In addition to the periodic activities of CDKs, progression through the cell cycle is gated by multiple checkpoints that ensure that all required processes are in place and the integrity of the cell and its DNA are intact before proceeding onto the next phase (Figure 7). These checkpoints are activated in the presence of stressors and act upstream of the CDKs to halt cell cycle progression. The G1/S checkpoint mainly halts cell cycle progression through the activity of tumor suppressors Rb (which

controls the restriction point), p21 and p53 (Kastan and Bartek, 2004). The intra-S checkpoint acts through the ATM/ATR signalling pathway to halt DNA replication (Bartek et al., 2004). The G2/M checkpoint targets multiple pathways leading to Cyclin B/CDK1 inhibition to prevent entry into mitosis (Kastan and Bartek, 2004). Finally, the metaphase checkpoint has unattached kinetochores signalling to the anaphase promoting complex to prevent initiation of anaphase (Lara-Gonzalez et al., 2012). Together, these checkpoints ensure that every round of the cell cycle is completed without error.

B. Influence of the cell cycle on biological processes

During the cell cycle, cells undergo dramatic physical and biochemical changes, engendering different cell states. These cell states are found to influence cellular decisions in perspectives other than proliferation such as metabolism and inflammatory responses. For example, studies in yeast have discovered that enzymes involved in glycolysis, oxidative phosphorylation, fatty acid biosynthesis were all induced during the G1 phase of the cell cycle (Cho et al., 1998). Research on human fibroblasts demonstrated that the replication checkpoint can activate inflammatory responses by inducing interleukin-6 secretion (Rodier et al., 2009). In mESCs, the physical state of DNA during replication was found to promote cell fate changes after nuclear fusion with somatic cells (Tsubouchi et al., 2013), while shortened telomeres were found to block fate transitions (Pucci et al., 2013). These findings provide evidence that cells have evolved to utilize cell cycle states in order to prime and regulate other biological events, which may include pluripotency.

C. Cell cycle of ESCs

Given the importance of the cell cycle, its regulation and progression is widely conserved across animal species and cell types. However, unlike most cells, *in vivo* pluripotent stem cells participate in rapid and uninterrupted symmetric cell divisions prior to gastrulation (Edgar and Lehner, 1996; Lawson et al., 1991; Murray and Kirschner, 1989; Yarden and Geiger, 1996), in order to quickly amplify up to cell numbers required for early embryonic development. Their extraordinary speed in cell division is enabled by a unique cell cycle profile featuring alternating rounds of DNA synthesis (S phase) and mitosis (M phase) with very short or no interjecting gap phases.

Interestingly, this unique cell cycle profile is preserved in ESCs *in vitro* (Figure 8A). Both mouse and human ESCs spend the majority of their time in the S and G2 phases, while the length of the G1 phase is drastically shortened compared to their somatic counterparts (Coronado et al., 2013; Hindley and Philpott, 2013; Singh and Dalton, 2009; Stead et al., 2002). Accordingly, the differentiation of hESCs reverts their cell cycle profile by the lengthening of the G1 phase (Calder et al., 2013; Filipczyk et al., 2007), potentially facilitated by the increase in negative regulators of proliferation, such as p21 and p27, during the differentiation of hESCs (Egozi et al., 2007). This abbreviation of the G1 phase is very specific to the pluripotent state, and this is because the hESC pluripotency network is responsible for enforcing it (Figure 8B). For example, OCT4, SOX2 and MYC all promote expression of miRNAs such as miR-302 (Card et al., 2008; Judson et al., 2009; Singh and Dalton, 2009), which belong to the class known as ESC cycle (ESCC) miRNAs, which specifically promote the G1-to-S transition in ESCs. NANOG on the other hand promotes expression of

CDK6 and CDC25A to facilitate S phase entry of hESCs (Zhang et al., 2009). Hence, cell cycle regulation by the transcriptional program intrinsic to pluripotency couples the ESC state with its unique cell cycle profile.

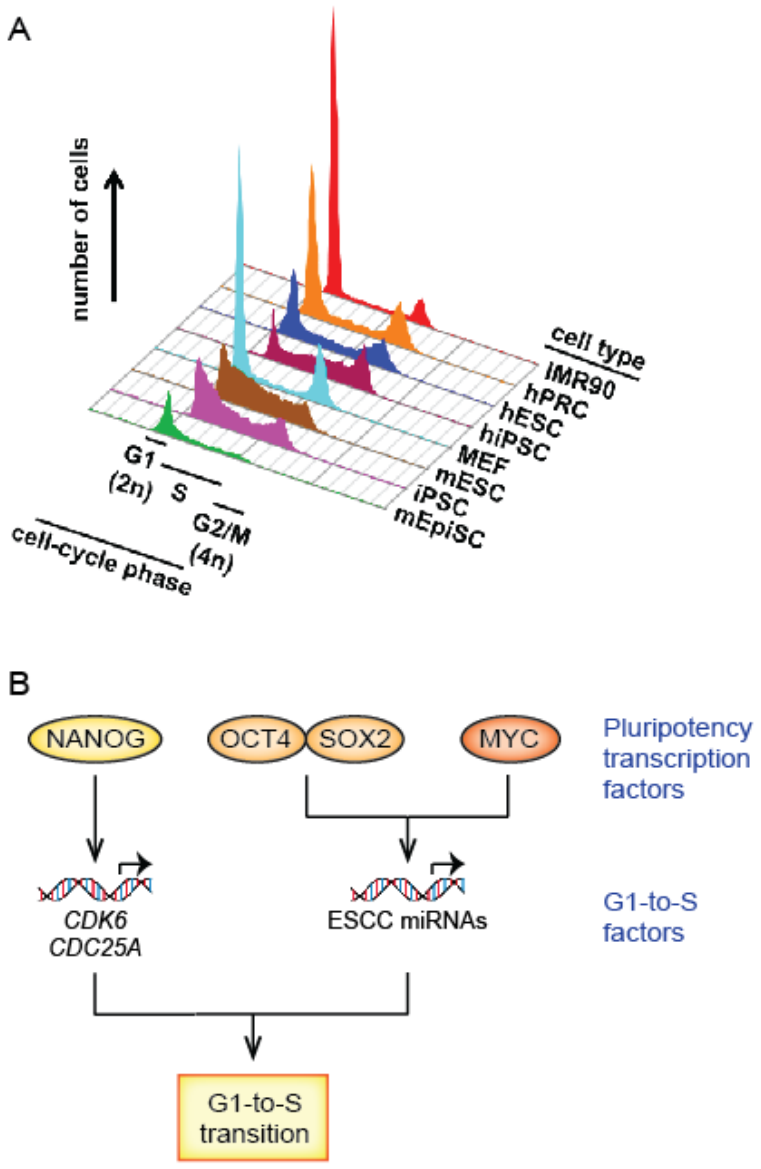



**Figure 8. The cell cycle profile of ESCs.** (**A**) A smaller proportion of ESCs spend their time in the G1 phase compared to somatic cells. A larger proportion of ESCs are instead in the S phase. Reprinted from Cell Stem Cell 5, Singh, A.M. and Dalton, S. The cell cycle and Myc intersect with mechanisms that regulate pluripotency and reprogramming, 141-149, Copyright (2009), with permission from Elsevier. (**B**) Pluripotency-associated transcription factors promote transcription of genes that facilitate the G1-to-S transition.

D. Regulation of the exit from pluripotency by the cell cycle machinery

While the pluripotency network acts to facilitate the unusual cell cycle profile of ESCs, this in turn is believed to contribute to the maintenance of pluripotency. First, ESCs are characterized by a globally euchromatic genome (Gaspar-Maia et al., 2011; Meshorer et al., 2006), which is important for their ability to give rise to the various cell types of the embryo, as it facilitates rapid epigenetic changes in response to differentiation cues. The accelerated G1-to-S transition displayed by ESCs could help maintain this open chromatin state, as experimentally observed in Rb-deficient fibroblasts (Herrera et al., 1996). Second, and perhaps more importantly, the current paradigm dictates that G1 is the only phase permissive to differentiation (Mummery et al., 1987; Pauklin and Vallier, 2013; Sela et al., 2012), and the shortening of this phase in ESCs serves to minimize aberrant exit from pluripotency.

The window of opportunity for differentiation provided by the G1 phase in hESCs can potentially be explained by three varying but non-exclusive hypotheses. First, the G1 phase boasts an enrichment of cell cycle factors like CDK inhibitors and Cyclin D that contribute to lineage specification (Li et al., 2012a; Pauklin and Vallier, 2013). Second, the state of chromatin modifications in the G1 phase, such as elevated 5-hydroxymethylcytosine levels, leads to higher expression of differentiation markers (Singh et al., 2013) that lowers the differentiation threshold. Third, the displacement of transcription factors in the M phase (Martinez-Balbas et al., 1995) creates a blank slate for establishing new transcriptional programs in the G1 phase (Egli et al., 2008). However, none of these hypotheses have been compellingly proven as studies on the regulation of pluripotency by the cell cycle in hESCs are mostly limited to correlation. Hence, there is no direct evidence hitherto that factors within the core cell cycle machinery deterministically control the maintenance of or the exit from pluripotency.

How ESCs make their fate choices in response to cell cycle events have direct consequences on genome stability, tissue development and stem cell maintenance. Therefore, it is imperative to interrogate the functional relationships between the cell cycle and pluripotency, which we tackle in the second half of this thesis.

## Objectives

In this thesis project, we aimed to search for factors crucial for the exit from the human pluripotent state. We did this by developing a systematic approach that capitalizes on the investigative strength of high-throughput RNAi screening and applied it to hESCs to study the exit from pluripotency. We found multiple processes that regulate the exit from pluripotency, which encompass those that function in a context-dependent manner and those which are universally important for the exit from pluripotency. These pathways range from histone acetylation and chromatin remodelling to development-related signalling, showcasing that control of the exit from pluripotency occurs in a diverse fashion at multiple levels.

Importantly, we found that the cell cycle has an assertive effect on the exit from pluripotency, as the exit from pluripotency is blocked upon cell cycle arrest at the S and G2 phases. Therefore, we additionally aimed to find mechanisms by which the exit from pluripotency becomes hard-wired to the cell cycle machinery. We found that cell cycle pathways stimulated during S and G2 phase perturbation play an active role in promoting the pluripotent state and preventing its exit. Thus, this thesis addresses the unanswered question of whether factors in the core cell cycle machinery enforce a deterministic regulation on pluripotency and differentiation in ESCs.

## Materials and Methods

### Cell Culture

The hESC lines H1 (WA-01, passage 30), *NANOG-GFP* H1 reporter cells (passage 50) and *ACTIN-GFP* H1 reporter cells (passage 45) were used for this study. They were cultured feeder-free on Matrigel (BD) with mTeSR1 medium (STEMCELL Technologies). Medium was changed daily. The hESCs were routinely subcultured with 1 U/mL Dispase in DMEM:F12 (STEMCELL Technologies) every 4–5 days. For experiments, hESCs were passaged using TrypLE™ Express (Life Technologies) before treatment.

For treatment with small molecules, chemicals diluted in DMSO were added to cell culture medium to achieve 0.2% DMSO by volume. The compounds used are as follows: PD0332991 (6 μM, Santa Cruz), Aphidicolin (75 ng/mL, Sigma), 5-hydroxyurea (0.2 mM, Sigma), Alisertib (65 nM, Selleckchem), RO3306 (10 μM, Enzo Life Science), Nocodazole (18 ng/mL, SciMed), Barasertib (50 nM, Selleckchem), BI2539 (5.5 nM, Selleckchem), Taxol (1.1 nM, Santa Cruz), Caffeine (1 mM, Sigma), AZD7762 (100 nM, Selleckchem), Nutlin (4uM, Cayman Chemical).

### Reporter Line Generation

EGFP cassette with kanamycin selection was inserted into BACs for NANOG (CTD-2317D19, BacPac) and beta-Actin (CTD-3223H20, BacPac) immediately before the initiating Methionine (ATG) of respective genes using recombineering (Quick & Easy BAC Modification Kit, KD-001, Gene Bridges GmbH). The Tol2 transposon cassette with Ampicillin selection mark was inserted into the loxP site of the BAC in the backbone using recombineering. 10 million H1 cells were cultured in

CF1 conditioned medium (20% KO serum replacement, 1 mM l-glutamine, 1% non-essential amino acids, 0.1 mM 2-mercaptoethanol and 8 ng $ml^{-1}$ of basic fibroblast growth factor in DMEM:F12) for 6 days and dissociated into single cells with TrypLE™ Express (Life Technologies) and electroporated with 20 micrograms of Tol2 transposes and 100 micrograms of Tol2/EGFP modified Transposon-BACs. After electroporation, the cells were resuspended in conditioned medium with 10 μM ROCK inhibitor Y27632 (Calbiochem). Rock inhibitor was added for the first 48 hours after electroporation. 50 micrograms/ml geneticin (Gibco) was added for selection of positive clones 72 hours post-electroporation. 14 days later after drug selection, single colonies were picked into 24 well plates for expansion. Fluorescent *in situ* hybridyzation (FISH) using non-modified BACs as probes was carried out to validate the incorporation of BAC construct into genome of ES cells (Cytogenetics Services, Genome Institute of Singapore).

High-throughput RNAi screening for the exit from pluripotency

384-well plates (Greiner) were coated with 10 uL of 30x-diluted Matrigel (BD) overnight before removing excess Matrigel. Pooled siRNAs (2.5 uL of 1 μM; Dharmacon, Ambion) were printed on the plates and frozen at -80°C before use. Reverse transfection was performed by adding a master mix of Lipofectamine RNAi Max (Invitrogen) diluted 200x in 5 uL OptiMEM (Invitrogen) per well, and incubated at room temperature for 20 mins. 45 uL mTeSR1 (STEMCELL Technologies) containing approximately 3,000 *NANOG-GFP* H1 hESCs were seeded as single cells with 500 nM Thiazovivin (Tocris). mTeSR was replaced with differentiation media 24hrs after seeding. The differentiation media used are as follows: 1) - bFGF, -TGFβ condition: mTeSR1 without select growth factors (STEMCELL Technologies),

namely bFGF, TGFβ, GABA and pipecolic acid, 2) TGFβ pathway inhibition: mTeSR + 1 μM A8301 (Stemolecule), 3) bFGF pathway inhibition: mTeSR1 + 2.5 μM PD0325901 (Sigma), 4) + Retinoic acid: mTeSR1 + 20 μM retinoic acid (Sigma). Cells were incubated in differentiation media for 120hrs for condition 1, and 48 hrs for conditions 2, 3 and 4. Media were then replaced with mTeSR1 and incubated for another 48 hrs to enrich for pluripotent cells. The cells were then fixed with 4% paraformaldehyde (Sigma) for 30 mins, and then stained with Hoechst 3342 (1:4000, Invitrogen) in 1% Triton X-100 for 30 mins. GFP fluorescence microscope photos were acquired using ImageXpress Ultra (Research Instruments) at 20x magnification with 9 pictures taken per well, and quantified using MetaXpress Image Acquisition and Analysis software. *z*-score was calculated using the formula $z = (X - \mu)/\text{s.d.}$ where $\mu$ is the mean and s.d. is the standard deviation of the whole population. $X$ is the sample value calculated based on the integrated fluorescent intensity divided by total number of cells. Cell viability scores were calculated using the formula $v = C_s/C_{nt}$ where $C_s$ is the total number of cells in the sample well and $C_{nt}$ is the mean total number of cells in the negative control wells in the same plate. The $Z'$ factors for all screens are > 0.5. Screen analyses was done using Screensifter software(Kumar et al., 2013).

Informatics analysis for screen results

For individual analyses, genes considered hits have z > 1.25 or z > 1.5 (above noise) in at least 2 out of 3 replicates for each condition. For the combined analysis, genes considered hits have z > 1.25 in at least 4 out of 9 sets in - bFGF, - TGFβ, TGFβ pathway inhibition and bFGF pathway inhibition conditions. Hits which also have z > 1.5 in at least 2 out of 3 replicates in + Retinoic acid condition are excluded

in the combined analysis. For the ACTIN-GFP H1 RNAi screens, genes considered hits have z > 1.25 in both replicates. For the analysis of hits that enhance the exit from pluripotency, genes considered hits have z < -1, -1.25 or -1.5 in at least 2 out of 3 replicates for each condition. List of hits for the - bFGF, - TGFβ, TGFβ pathway inhibition and bFGF pathway inhibition conditions were pooled for combined analysis.

Hierarchical clustering of screen conditions was performed by calculating Euclidian distances between the set of z-scores from each condition using the distance matrix algorithm in R. Gene ontology analysis was performed with DAVID functional annotation tool (david.abcc.ncifcrf.gov) for biological processes and functional clusters. Reactome pathway analysis was performed at www.reactome.org was used for the analysis of reactions and/or pathways that were statistically over-represented from the hits. Protein-protein interaction network was generated using STRING database (string.embl.de), and filtered only for experimentally-validated interactions. The resulting network was imported into Cytoscape (www.cytoscape.org), and then HPRD (www.hprd.org) validated interactions were added to the network. Enriched clusters were derived using Markov clustering algorithm (Van Dongen, 2008).

RNA quantification and RT-qPCR

RNA was extracted from cells using TRIzol® (Invitrogen). 500 ng RNA was reverse-transcribed using Superscript II (Invitrogen) utilizing the oligo(dT) primer. mRNA expression changes were quantified from qPCR using SYBR Green (KAPA). Measured transcripts were normalized to GAPDH and all samples were run in triplicate. qPCR primer sequences used are available in Appendix 1.

Knockdown of gene expression with RNAi constructs

Knockdowns were performed using shRNAs designed using BLOCK-iT RNAi designer (Invitrogen) cloned into the pLKO.1 vector. pLKO.1 lentiviruses were packaged using HEK293T cells. 2 uL of concentrated viruses was added to culture medium with 4 ug/mL polybrene (Sigma). Knockdowns were also performed using siRNAs (Bioneer) at a final concentration of 50 nM transfected using Lipofectamine RNAi Max (Invitrogen). A full list of sequences of RNAi constructs used in this thesis can be found in Appendix 2.

Immunoblotting

Immunoblotting was performed as conventional procedures using cells lysed with RIPA buffer with proteinase inhibitor (Merck). Antibodies against Cyclin D (AF4196, 1:200), Cyclin E (AF6810, 1:200), Cyclin A2 (AF5999, R&D system 1:1000), Cyclin B1 (sc-245, Santa Cruz, 1:500), p53 (sc-126, Santa Cruz, 1:500), p21 (sc-397, Santa Cruz, 1:500), pChk1 (2348, Cell Signalling, 1:1000), pChk2 (2661, Cell Signalling, 1:1000), pSmad2 (3108, Cell Signalling, 1:500) and Glyceraldehyde 3-phosphate dehydrogenase (sc-25778, 1:1000) were used.

Immunostaining, microscopy and FACS

To perform OCT4, NANOG, TRA-1-60 and TRA-1-81 immunostaining for microscopy, cells were fixed with 4% formaldehyde (Sigma) on tissue culture plates (Falcon) and permeabilized using 1% Triton X-100. Primary antibodies used are as follows: OCT4 (ab19857, Abcam, 1:5000), NANOG (AF1997, R&D system, 500

ng/mL), TRA-1-60 (sc-21705, Santa Cruz, 1:500), TRA-1-81 (sc-21706, Santa Cruz, 1:500), p53 (sc-126, Santa Cruz, 1:500). To perform γH2AX immunostaining for microscopy, cells were harvested by cytospinning, fixed with 4% formaldehyde (Sigma) and stained with antibody against γH2AX (#2577, Cell Signalling, 1:800). Secondary antibodies used were Alexa Fluor 488 anti-mouse IgG, Alexa Fluor 488 anti-rabbit IgG, and Alexa Fluor 546 anti-goat IgG (Invitrogen). Hoechst 3342 (1:4000, Invitrogen) was used for nuclear staining. All microscope images were taken using Observer Z.1 (Zeiss).

To perform immunostaining for FACS, cells were harvested, fixed with 4% formaldehyde (Sigma) and 80% ethanol. To obtain cell cycle profiles, cells were stained with antibody against phospho-histone H3 (06-570, Merck, 1:1000) and Hoechst 33342. FACS analyses were done using LSRII flow cytometer (BD).

Microarray analysis

mRNAs derived from hESCs were reverse transcribed, labelled and analyzed on Illumina microarray platform (HumanHT-12 v4 Expression BeadChips). Arrays were processed according to manufacturer's instructions. Biological triplicate microarray data were generated. Rank invariant normalization was used to normalize the microarrays. SAM (statweb.stanford.edu/~tibs/SAM) was used for analysis and matrix2png (www.chibi.ubc.ca/matrix2png) for visualization.

Teratoma assay

hESCs were dissociated with TrypLE™ Express (Life Technologies) and resuspended in Matrigel (BD) diluted 3x in DMEM:F12 (Nacalai Tesque) at a concentration of 1,000,000 cells/mL. 100uL of the cell suspension was injected into

the dorsal flanks of BALB/c nude mice anesthetized with isoflurane. After 6-8 weeks, teratomas were harvested, surgically dissected, fixed in Bouin's solution, embedded in paraffin blocks and sectioned. Tissue sections were analysed with Mallory's tetrachrome staining.

Karyotyping

hESCs were treated with colcemid to induce mitotic arrest. Cells were then harvested by standard hypotonic treatment and methanol:acetic acid (3:1 v/v) fixation. Slides were prepared by standard air drying and G-band karyotyping was executed.

## Results and Discussion

### I. The high-throughput RNAi screening conditions are optimized to assay for the exit from pluripotency

With the aim of uncovering factors and processes that are required for the exit from pluripotency, we designed a high-throughput RNAi screen in hESCs under multiple differentiation conditions. To ensure the robustness of the experimental design, initial optimization experiments were performed as follows.

#### A. Search for a reliable marker of the pluripotent state

To perform a high-throughput assay for the exit from pluripotency, we needed a reliable marker with an expression that changes rapidly and drastically upon induction of differentiation. We triggered the exit from pluripotency of hESCs through the withdrawal of bFGF and TGFβ (- bFGF, - TGFβ condition), as these two factors are necessary for hESC self-renewal (Beattie et al., 2005; Ludwig et al., 2006; Xu et al., 2005). The expression profiles of an array of genes which are known to have differential expression in hESCs and somatic cells (Assou et al., 2007) were then compared at 0, 48 and 96 hours of incubation in - bFGF, - TGFβ (Figure 9). We found that some genes like *GPC4* and *INDO* undergo a transient upregulation upon exit from pluripotency; these are not suitable markers for the pluripotent state. Conversely, majority of the genes showed a gradual decrease in expression in - bFGF, - TGFβ. Among these, *NANOG* showed the greatest decrease in expression at both the earlier and later time points. Thus, we chose *NANOG* as the ideal marker for the human pluripotent state, as it provides the most rapid and drastic decrease in expression upon induction of the exit from pluripotency.

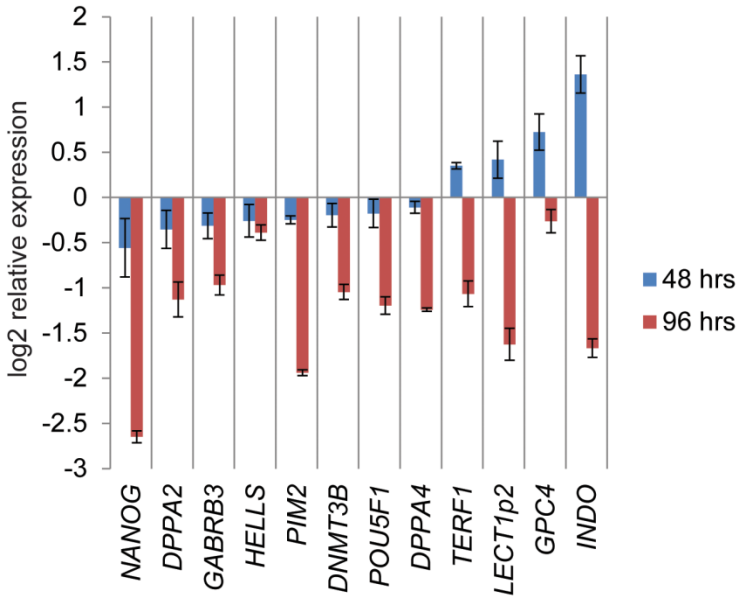



**Figure 9. Time-course expression profile of pluripotency markers upon the exit from pluripotency.** Transcript levels were measured using microarray analysis. NANOG expression level experiences the most drastic decrease at both earlier and later time points. Error bars denote standard deviation of triplicate data.

B. Generation and validation of the *NANOG-GFP* hESC line

To easily monitor *NANOG* expression of hESCs, we created a transgenic hESC line using a Transposon-BAC system (see Materials and Methods), wherein GFP expression is under the control of the *NANOG* promoter. In order to validate that green fluorescence reflects the pluripotent state of *NANOG-GFP* hESCs, we transfected these cells with siRNAs against *POU5F1* and *GFP* (Figure 10A-B). As expected, knockdown of *POU5F1* has induced differentiation, as seen with the decrease in both green fluorescence and *NANOG* transcript levels. This demonstrates that the pluripotent state, *NANOG* expression levels, and the intensity of green fluorescence are all coupled together.

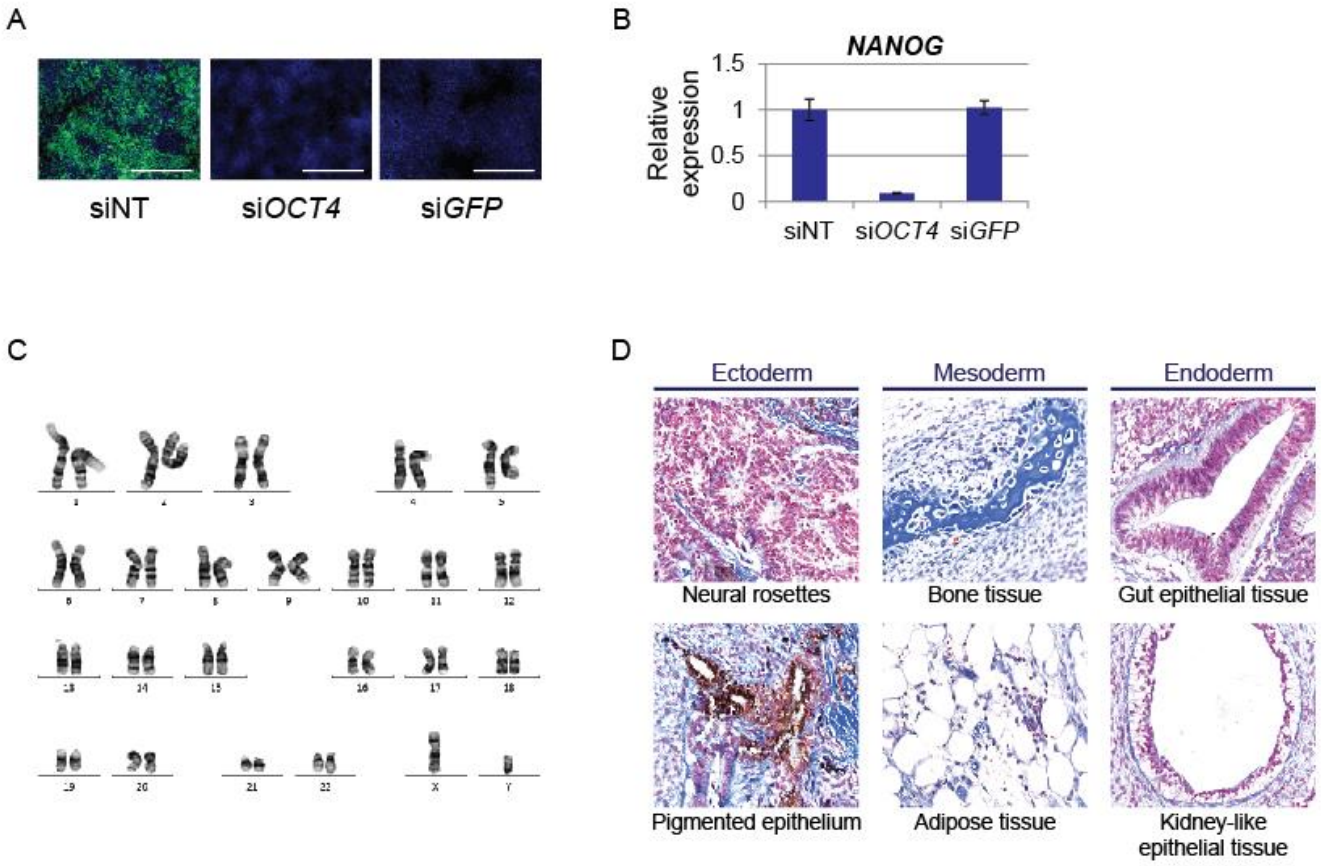



**Figure 10. Validations for the *NANOG-GFP* hESC line.** (**A**) Images for *NANOG*-GFP fluorescence (green) and nuclei (blue) after transfection with siRNA controls. Scale bar = 300μm. (**B**) Expression level of *NANOG* as measured by qPCR after transfection with siRNA controls. Error bars denote standard deviation for triplicate data. (**C**) Cytogenetic analysis of NANOG-GFP cells confirms a normal karyotype of 46XY chromosomes. (**D**) *NANOG-GFP* hESCs gave rise to teratomas consisting of tissues from ectoderm, mesoderm and endoderm lineages. *NANOG-GFP* hESC line was generated with the help of Dr Gao Bin and Le Beilin.

To confirm that the modified hESC line is suitable for downstream purposes, we performed cytogenetic analysis to find a normal karyotype with 46XY chromosomes (Figure 10C). Teratoma formation assay for *NANOG-GFP* hESCs also ensured that their pluripotency is uncompromised; *NANOG-GFP* hESC-derived teratomas comprised tissues from ectoderm, mesoderm and endoderm lineages 6

weeks after injection into SCID mice (Figure 10D). Altogether, these results confirm that the hESC line generated is perfectly normal, and that GFP fluorescence levels in these cells accurately reflect their state of pluripotency.

C. Optimization of differentiation conditions

To achieve robust and unbiased identification of universal and specific factors governing the exit from pluripotency, we performed the siRNA screen under multiple differentiation conditions. To dissect the signalling control of hESC maintenance, we wanted to individually perturb the bFGF and TGFβ pathways, in addition to the - bFGF, - TGFβ condition. Hence, we used the ALK5 kinase inhibitor A8301 to inhibit the TGFβ pathway (TGFβ pathway inhibition condition), and the MEK inhibitor PD0325901 to inhibit the bFGF-MEK axis (bFGF pathway inhibition condition). In addition, we included a fourth condition wherein we introduce a differentiation signal in the form of retinoic acid without withdrawing self-renewal signals (+ Retinoic acid condition). We ensured that all four conditions induce efficient exit from pluripotency by checking *NANOG*-GFP fluorescence (Figure 11A) and pluripotency marker expression at both the transcript (Figure 11B) and protein (Figure 11C) levels. Importantly, we observed a varied response in the upregulation of lineage-specific markers in the four conditions (Figure 11B), indicating that the different conditions we are using for the experiment lead to distinct differentiated fates.

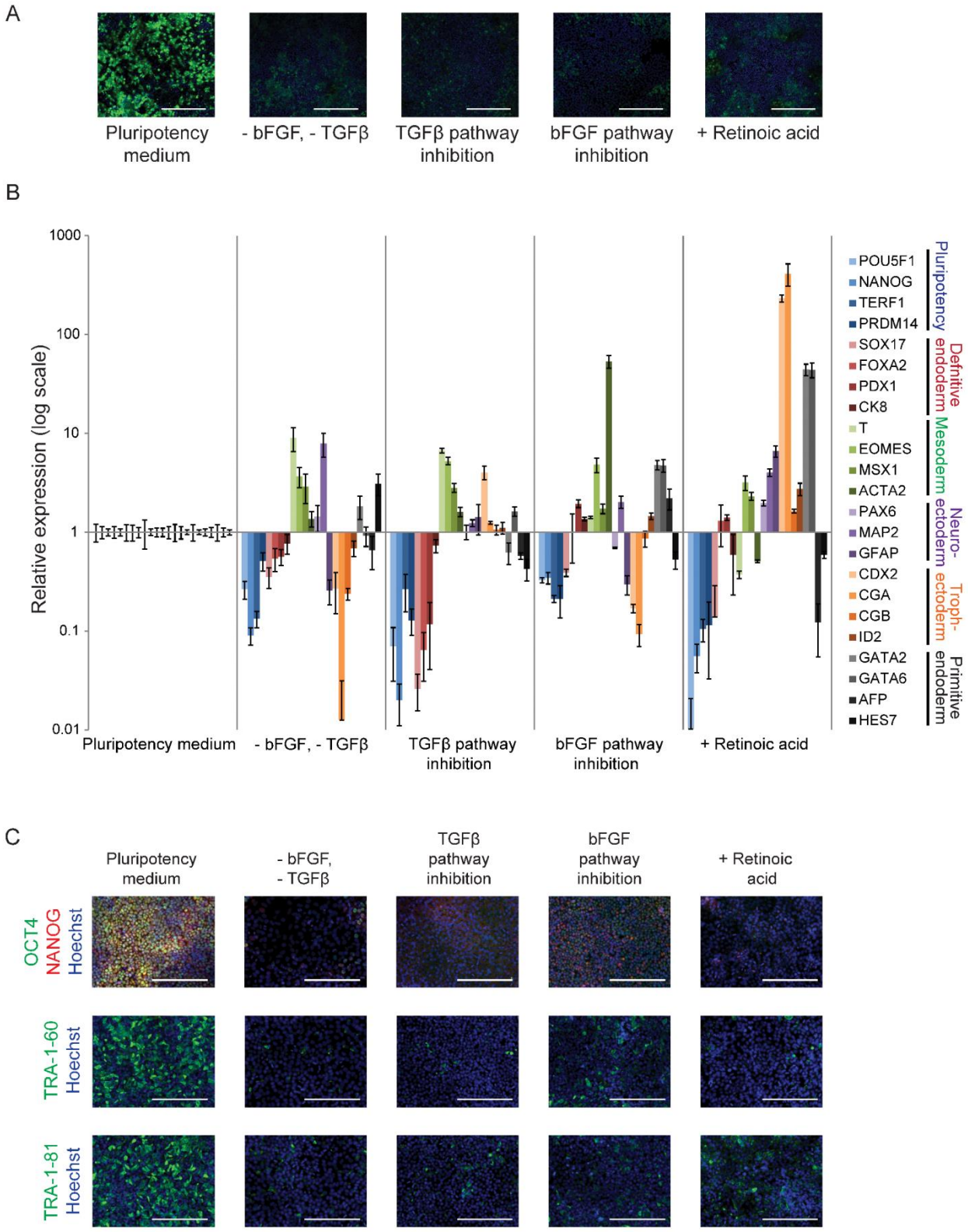



**Figure 11. Characterization of the four differentiation conditions.** (**A**) Images for *NANOG*-GFP fluorescence (green) and nuclei (blue) after incubation in the respective conditions. Scale bar = 300um. (**B**) Expression levels of pluripotency and lineage-specific markers as measured by qPCR. Error bars denote standard deviation for triplicate data. (**C**) Images for immunofluorescence staining against pluripotency markers after incubation in the differentiation media. Scale bar = 200μm.

D. Experimental design of the high-throughput RNAi screen

Assembly of all these primary results guided the final design of our high-throughput RNAi screen (Figure 12). First, siRNA pools against genes of interest were introduced into the *NANOG-GFP* hESC line by reverse transfection in a 384-well format. Given the large scale of the experiment, we limited our siRNA libraries to target a subset of the genome encompassing factors that can easily be targeted with small molecules (kinases, phosphatases and G-protein coupled receptors) and factors directly involved in gene expression regulation (chromatin modifiers and transcription factors). 24 hours after transfection, we initiate the exit from pluripotency by changing the culture medium to one of the four aforementioned conditions. After the differentiation period, hESCs are returned to the pluripotency medium to enrich for hESCs that have resisted the exit from pluripotency. The degree of preservation of hESC identity was subsequently measured through the average *NANOG*-GFP fluorescence intensity per cell, as hESCs depleted of genes crucial for the exit from pluripotency are expected to retain higher GFP signals. All in all, we screened a total of 4,558 genes in triplicate for all four conditions using pools of 3-4 independent siRNAs, summing up to 54,696 data points.

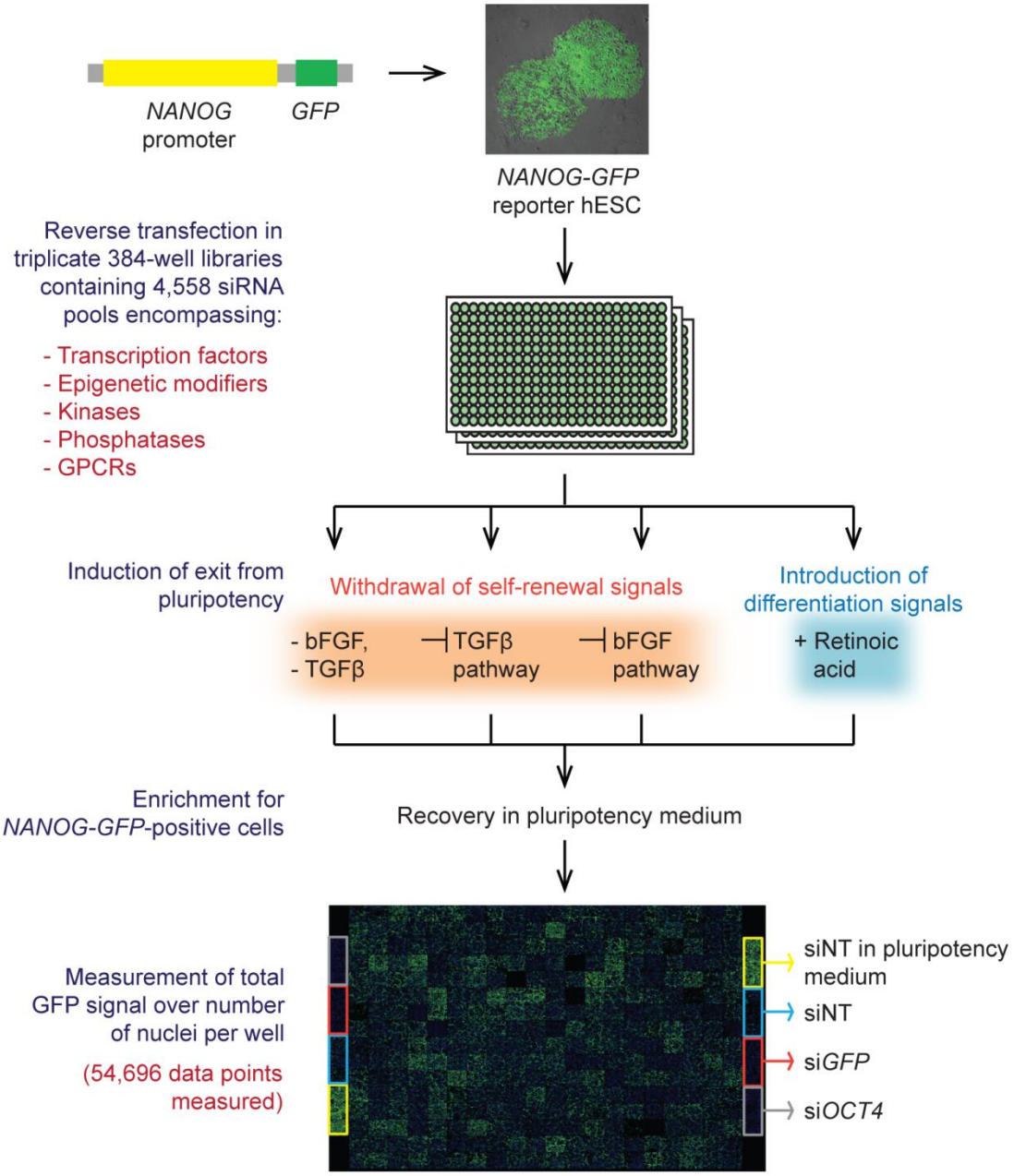



**Figure 12. Experimental design of the high-throughput RNAi screen.**
*NANOG-GFP* hESCs were transfected with siRNA pools in a high-throughput platform. Exit from pluripotency was induced using one of four differentiation conditions, and hESCs that resisted the exit from pluripotency were enriched by recovery in pluripotency medium. GFP fluorescence and cell number were finally measured from each well up to a total of 54,696 data points.

II. Quality control checks certify the robustness of the high-throughput RNAi screen

Before proceeding with the analysis of the screen results, it is imperative to check that the results obtained are of high quality. Quality control checks were put in place by assessing parameters such as plate alignment, plate layout effects, correlation, accuracy of fluorescence quantification, and cell number bias.

A. Plate alignment

High-throughput assays require samples to be distributed into experimental groups. In this large-scale RNAi screen, samples are divided into groups of up to 352, which is the maximum number that can be fitted into each culture plate. As environmental variations between plates are inevitable, plate-to-plate variation has to be accounted for to enable accurate analysis of screen results. We did this by aligning every plate with each other by calculating z-scores using the formula $z = (X - \mu)/\text{s.d.}$ where $\mu$ is the mean and s.d. is the standard deviation of of all samples within a plate (the full dataset of z-scores can be found in Supp. Tables 1-4). After z-score calculation, z-score distributions between each plate are perfectly aligned (Figure 13), negating any systematic bias that could have arisen between plates.

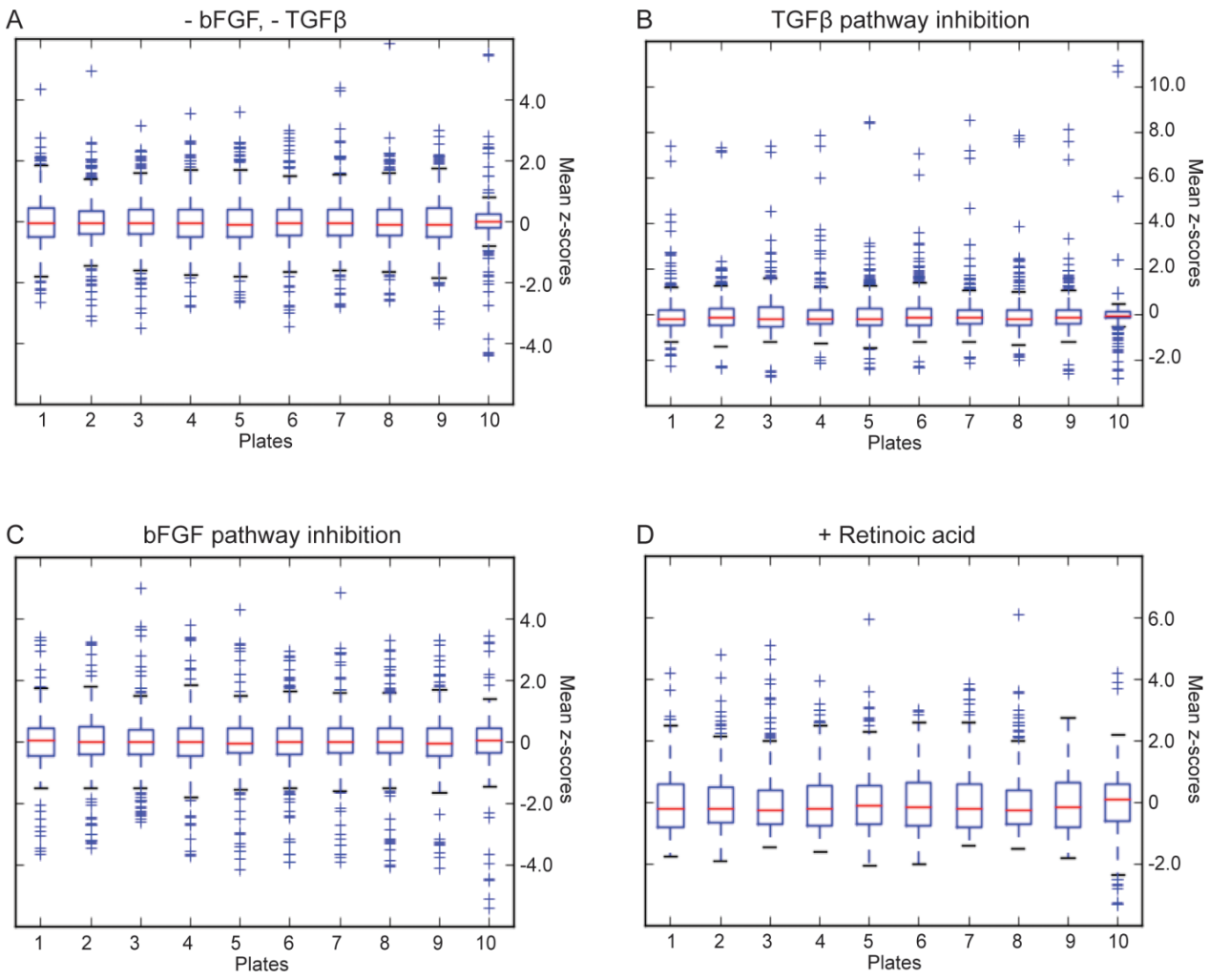



**Figure 13. Plate alignment of z-scores.** Boxplots showing that data from each plate is properly aligned for the (**A**) - bFGF, - TGFβ, (**B**) TGFβ pathway inhibition, (**C**) bFGF pathway inhibition and (**D**) + Retinoic acid conditions after z-score normalization. Data shown are from the transcription factor and epigenetic modifier subsets of the library.

B. Plate layout effects

Ideally, all conditions within each group are kept identical. High-throughput assays are however vulnerable to plate layout effects, which arise due to irregularities in the conditions between the numerous samples in each plate (Maddox et al., 2008). Most common is the edge effect, wherein samples near the edges of the plate layout experience slightly different conditions (mainly variations in humidity levels) from

those around the centre of the plate. Heatmaps of normalized fluorescence values for representative plates from each of the four screening conditions show no observable pattern, indicating no plate layout biases (Figure 14).

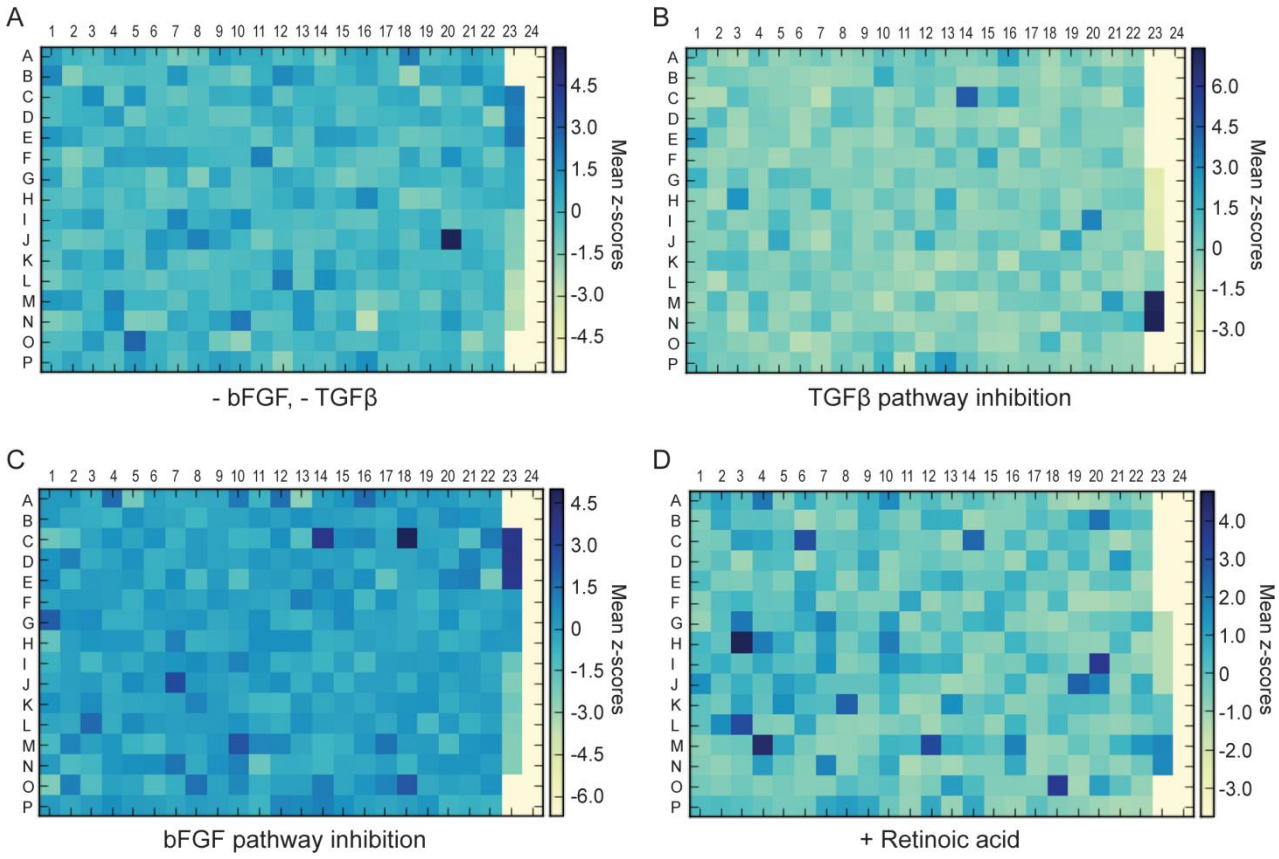



**Figure 14. Plate layout effects.** Heatmaps of z-scores for representative plates from the transcription factor subset for the (**A**) - bFGF, - TGFβ, (**B**) TGFβ pathway inhibition, (**C**) bFGF pathway inhibition and (**D**) + Retinoic acid conditions. No visible plate layout biases were observed for all conditions.

C. Correlation between replicates

We next checked for the correlation between replicates of each screening condition. There is a high linear correlation ($R^2$>0.6) between all replicates of the TGFβ inhibition, bFGF inhibition and + Retinoic acid conditions, demonstrating

excellent reproducibility (Figure 15A-C). The - bFGF, - TGFβ condition also showed good linear correlation ($R^2$>0.4), albeit lower than the other three conditions (Figure 15D). This can be attributed to the longer incubation time needed for exit from pluripotency to take place in - bFGF, - TGFβ, which provides more time for random variables to introduce noise into the results. Nevertheless, all conditions exhibit good reproducibility, certifying the robustness of the results.

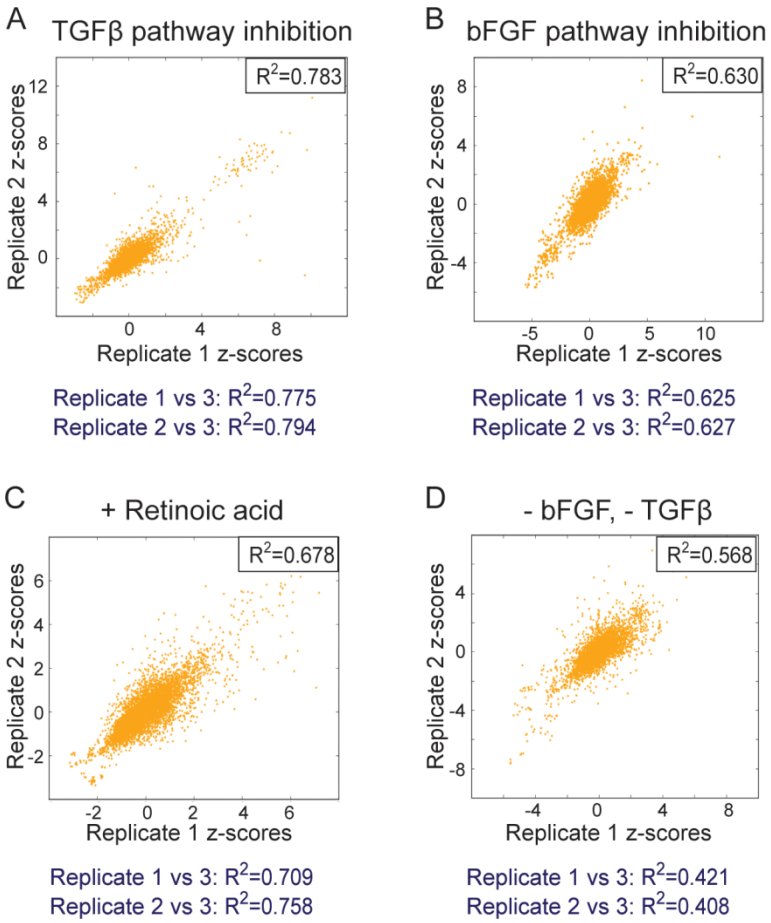



**Figure 15. Correlation between replicates.** Scatter plots depicting linear correlation between two replicates from screens in the (**A**) TGFβ pathway inhibition, (**B**) bFGF pathway inhibition, (**C**) + Retinoic acid conditions and (**D**) - bFGF, - TGFβ are shown. Pearson correlation values ($R^2$) for each combination of replicates are also indicated.

D. Accuracy of fluorescence quantification

To certify that the z-scores accurately reflect GFP fluorescence, several samples from the wells with relatively high z-scores (z-score>1.25) were examined. We ascertained that all these samples have visibly higher fluorescence levels (Figure 16), guaranteeing that hits have truly maintained a high *NANOG*-GFP signal.

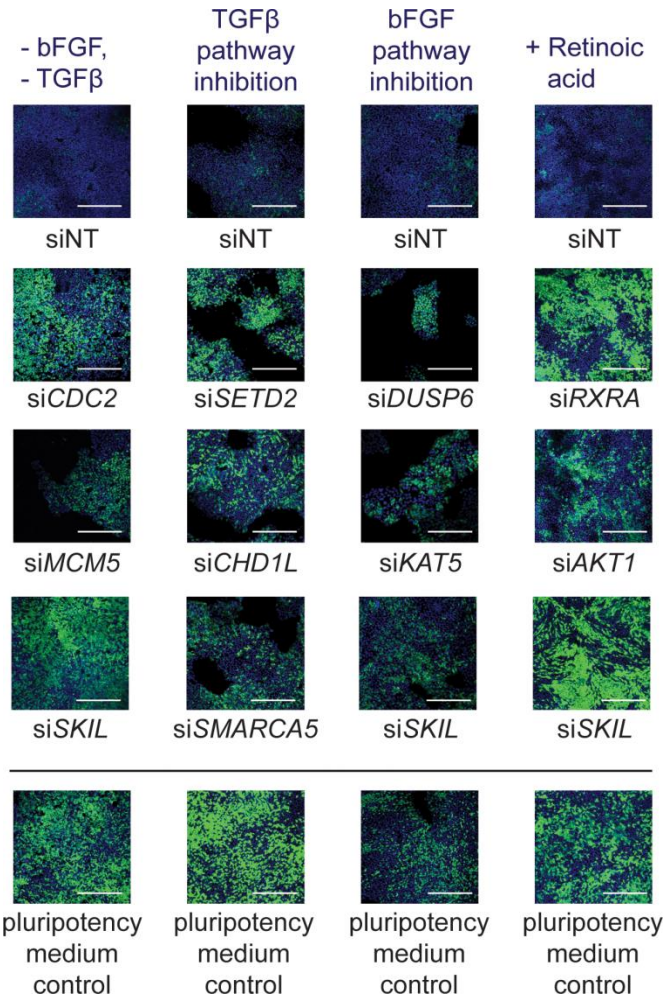



**Figure 16. GFP fluorescence of wells with high z-scores.** Representative images for *NANOG*-GFP fluorescence (green) and Hoescht staining (blue) for hits in the four differentiation conditions are shown. All samples show visibly higher GFP signals compared to the siNT negative controls. Scale bar = 300μm.

E. Cell number bias

Furthermore, correlations between the z-scores and cell viability scores were calculated to detect if there is any cell number bias. All four conditions demonstrate that there is no substantial correlation ($R^2$<0.1) between cell number and GFP fluorescence (Figure 17). Therefore, potential hits and non-hits from the screen both

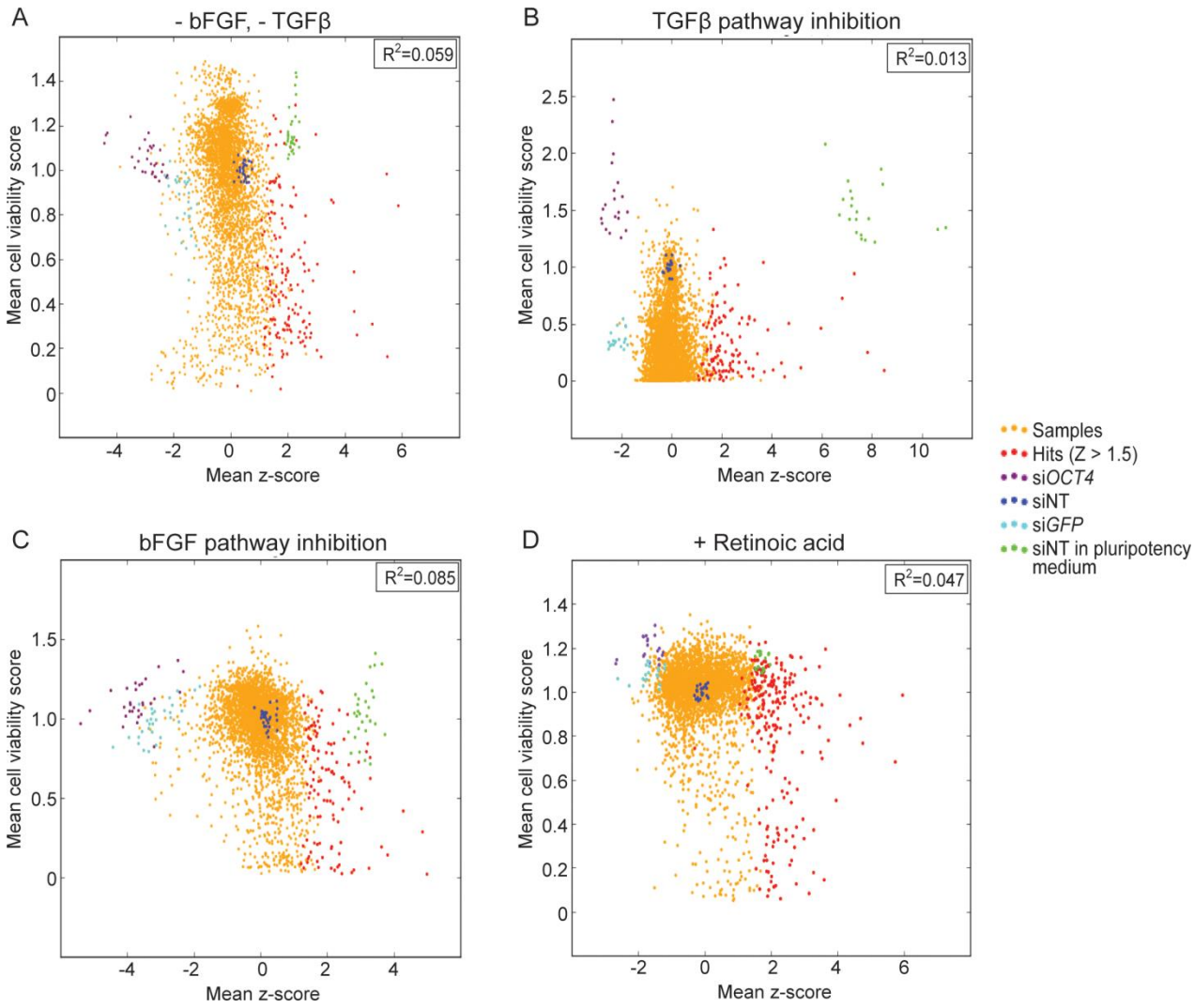



**Figure 17. Cell number bias.** Scatter plots depicting linear correlation between the mean z-scores and the mean cell viability scores from screens in the (**A**) - bFGF, - TGFβ, (**B**) TGFβ pathway inhibition, (**C**) bFGF pathway inhibition and (**D**) + Retinoic acid conditions are shown. Pearson correlation values ($R^2$) for each combination of replicates are also indicated. All conditions show very poor correlation, indicating the absence of cell number bias in all of the conditions.

span a wide range of cell numbers, eliminating the possibility that housekeeping genes would be wrongfully detected as false positives or false negatives.

F. Counter-screening using an *ACTIN-GFP* reporter

Finally, counter-screens using a hESC line harbouring an *ACTIN-GFP* reporter were conducted on a subset of the conditions. Samples that consistently give out a higher GFP signal versus the negative control (z-score > 1.25 in at least 2 replicates) were tallied, and no overlap between the list of genes for the *NANOG-GFP* and *ACTIN-GFP* reporters were found (Figure 18). This is conceivable as *ACTIN-GFP* should not decrease with the exit from pluripotency. Hence, the only hits that could be detected are those that those directly upregulate *ACTIN* expression, which should not overlap with siRNAs that confer maintenance of *NANOG* expression.

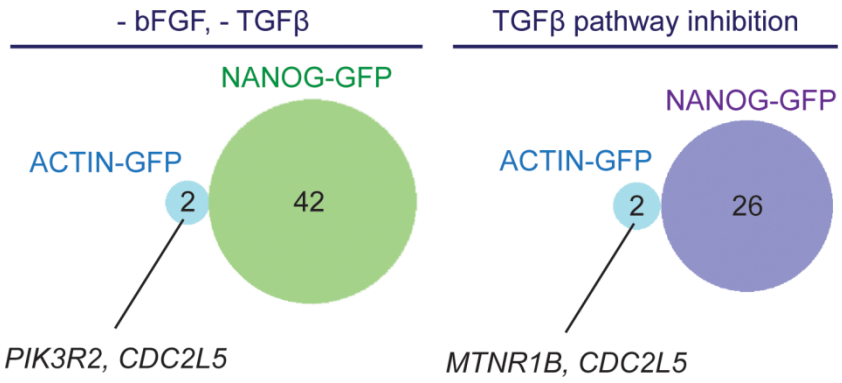



**Figure 18. Overlap with *ACTIN-GFP* hESC counter-screens.** Venn diagrams showing overlap of samples with z > 1.25 for at least 2 replicates in the *ACTIN-GFP* screens and *NANOG-GFP* screens for the kinase, phosphatase and GPCR subsets of the library. No overlap is observed.

G. Hit selection from the high-throughput RNAi screen results

The goal of this high-throughput assay is to identify genes that are critical for the exit from pluripotency of hESCs. Filtering these from the rest of the samples requires a sensible cut-off. With this in mind, samples that reproducibly gave a z-score > 1.25 or 1.5 in at least 2 replicates were considered as hits. These cut-offs are stringently above noise levels (Figure 19) while allowing a reasonable number of samples to be considered as hits. We obtained 206 hits for - bFGF, - TGFβ, 146 hits

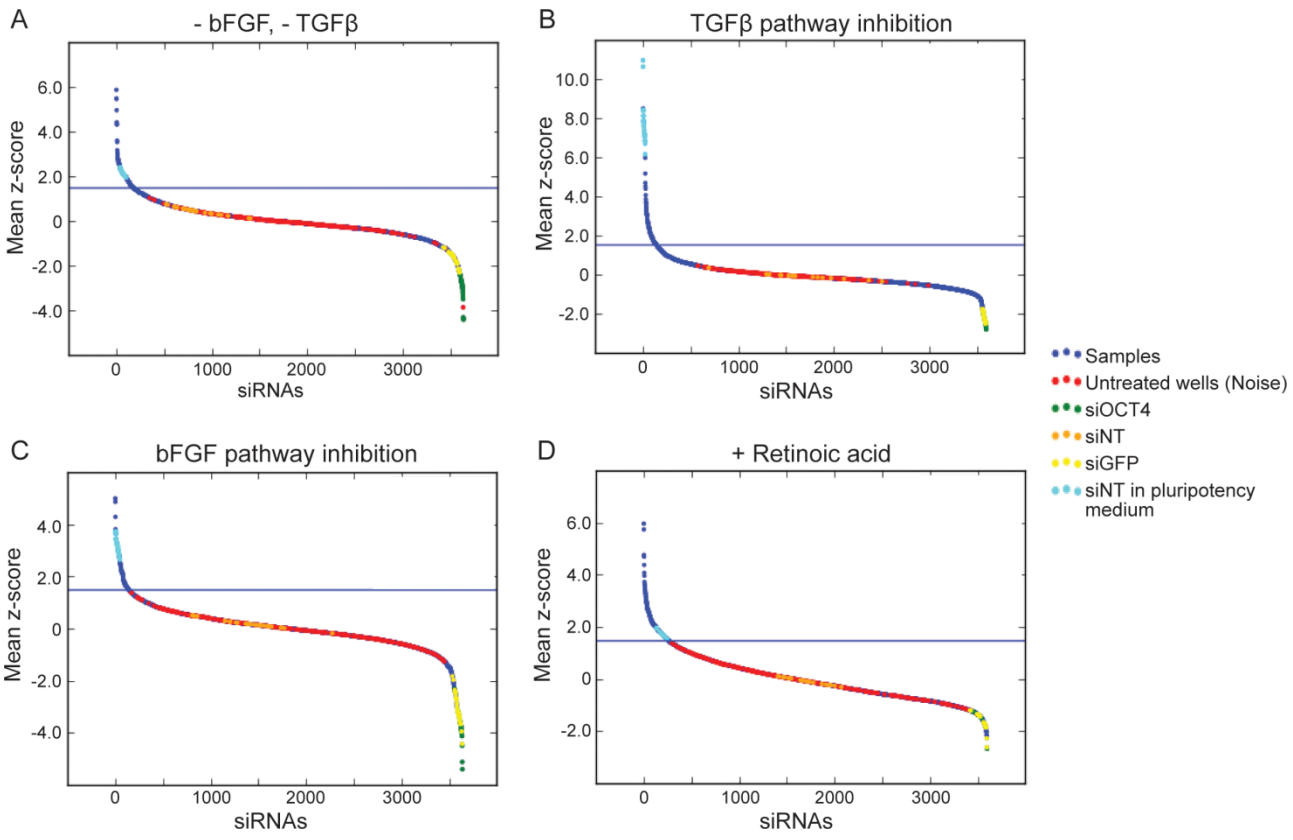
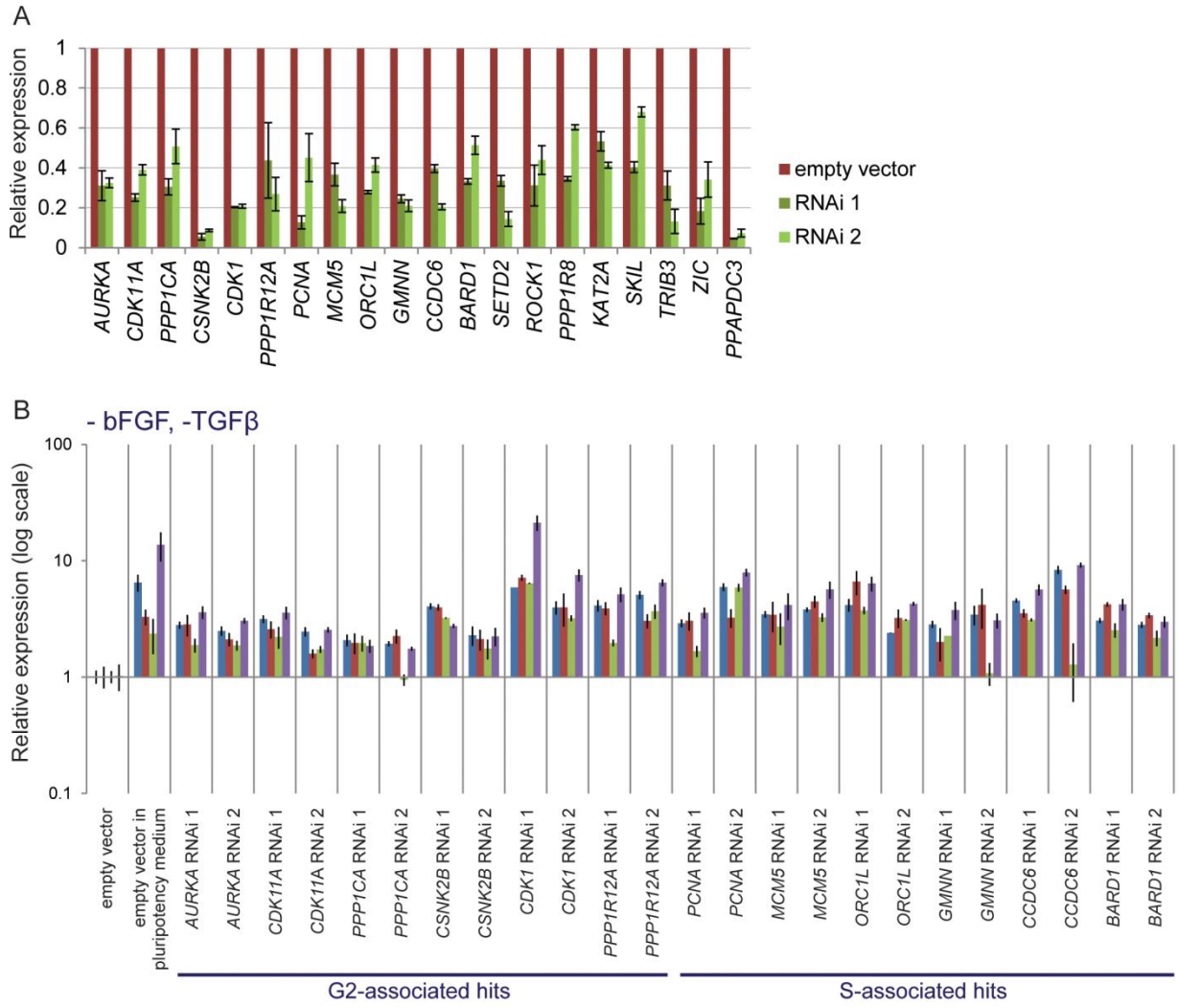



**Figure 19. Z-score cut-off for hit selection.** Descending plots of mean z-scores of screens in the (**A**) - bFGF, - TGFβ, (**B**) TGFβ pathway inhibition, (**C**) bFGF pathway inhibition and (**D**) + Retinoic acid conditions are shown. Blue horizontal line demarcates the selected cut-off, which are all above noise levels (as measured using untreated wells and the inflection point).

for TGFβ pathway inhibition, 138 hits for the bFGF pathway inhibition and 325 hits for + Retinoic acid. The full list of hits for each condition can be viewed in Appendices 3-6.

To validate the hits filtered out by this experiment, we selected a panel of genes and designed 2 independent RNAi constructs for each gene (Figure 20A). This includes many cell cycle-associated genes (Figure 20B), which we will focus on later in this thesis. We considered hits to be true positives if both constructs prevented the decrease of majority of the pluripotency markers tested upon exit from pluripotency. Altogether, we managed to validate 80.6% of the hits across the first three conditions (Figure 20B-F), confirming that these hits are true positives.

Altogether, these quality control checks establish that the screen is able to rigorously identify genes important for the exit from pluripotency of hESCs.


A
Relative expression
1
0.8
0.6
0.4
0.2
0
AURKA
CDK11A
PPP1CA
CSNK2B
CDK1
PPP1R12A
PCNA
MCM5
ORC1L
GMNN
CCDC6
BARD1
SETD2
ROCK1
PPP1R8
KAT2A
SKIL
TRIB3
ZIC
PPAPDC3
empty vector
RNAi 1
RNAi 2
B
- bFGF, -TGFβ
Relative expression (log scale)
100
10
1
0.1
empty vector
empty vector in pluripotency medium
AURKA RNAi 1
AURKA RNAi 2
CDK11A RNAi 1
CDK11A RNAi 2
PPP1CA RNAi 1
PPP1CA RNAi 2
CSNK2B RNAi 1
CSNK2B RNAi 2
CDK1 RNAi 1
CDK1 RNAi 2
PPP1R12A RNAi 1
PPP1R12A RNAi 2
PCNA RNAi 1
PCNA RNAi 2
MCM5 RNAi 1
MCM5 RNAi 2
ORC1L RNAi 1
ORC1L RNAi 2
GMNN RNAi 1
GMNN RNAi 2
CCDC6 RNAi 1
CCDC6 RNAi 2
BARD1 RNAi 1
BARD1 RNAi 2
G2-associated hits
S-associated hits
C
- bFGF, -TGFβ
Relative expression (log scale)
100
10
1
0.1
empty vector
empty vector in pluripotency medium
SETD2 RNAi 1
SETD2 RNAi 2
ROCK1 RNAi 1
ROCK1 RNAi 1
PPP1R8 RNAi 2
PPP1R8 RNAi 1
KAT2A RNAi 2
KAT2A RNAi 1
SKIL RNAi 2
SKIL RNAi 1
TRIB3 RNAi 2
TRIB3 RNAi 1
ZIC3 RNAi 2
ZIC3 RNAi 1
Legend (B-E):
POU5F1
NANOG
PRDM14
TDGF1

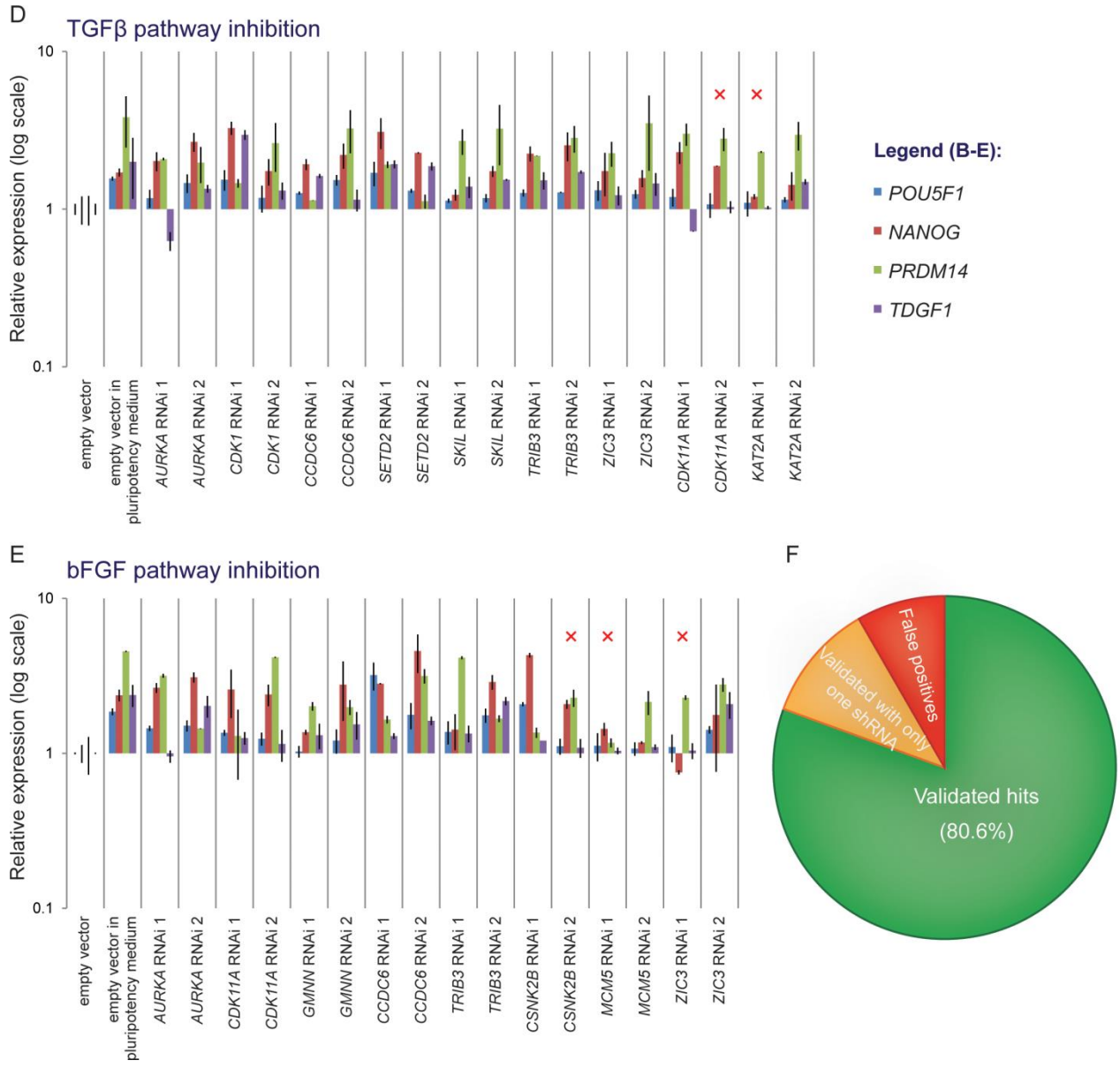



**Figure 20. Independent validations of hits.** (**A**) Transcript levels of respective genes upon RNAi knockdown as measured by qPCR. All RNAi constructs effectively decreased expression of their target gene. (**B-E**) Expression levels of pluripotency markers upon knockdown of hits in the (**B,C**) - bFGF, - TGFβ, (**D**) TGFβ pathway inhibition and (**E**) bFGF pathway inhibition conditions as measured by qPCR. Red crosses indicate false positives. All error bars denote standard deviations of triplicate data. (**F**) Pie chart depicting the percentage of hits validated. qPCR experiments were done with the help of Yee Siang Lim.

III. Independent analyses of screen results from each condition identify context-specific mechanisms regulating the exit from pluripotency

We first ventured to independently analyse the results from each screening condition by looking for processes, pathways and associations enriched within the list of hits from each screening condition. We employed protein interaction analysis using STRING and HPRD (Figure 21), pathway enrichment analysis using Reactome (Figure 22) and gene ontology analysis using DAVID (Figure 23) (see Materials and Methods). These platforms revealed the prevalence of context-dependent processes that are critical for the exit from pluripotency.

A. Procurement of hits expected in each screening condition

We expect that the list of hits should include factors that are associated with the primary pathways transducing the distinct initial differentiation cues in each condition. This is because perturbation of these primary pathways can nullify the differentiation-inducing effect stemming from the same pathway. Indeed, we find repressors of TGFβ-responsive genes (such as *SKI* and *SKIL*) to be enriched when upstream TGFβ signalling is inactive (Figure 21, 22B). It was reassuring to find similar results in other conditions, wherein the retinoic acid receptor and inactivators of ERK (*DUSP4*, *DUSP6*) are enriched in the hits for the + Retinoic acid and bFGF pathway inhibition conditions, respectively (Appendices 5,6).

Furthermore, we observe that retinoic acid-mediated exit from pluripotency relies heavily on transcription as members of the RNA polymerase II and TFIID/Mediator complexes are enriched in the hits (Figure 21, 22D, 23D). This is anticipated as the retinoic acid receptor is a transcription factor, requiring high activity of the general transcriptional machinery to elicit its effects on the exit from

pluripotency (Lefebvre et al., 2005). Altogether, these results demonstrate the robustness of the high-throughput screening assay.

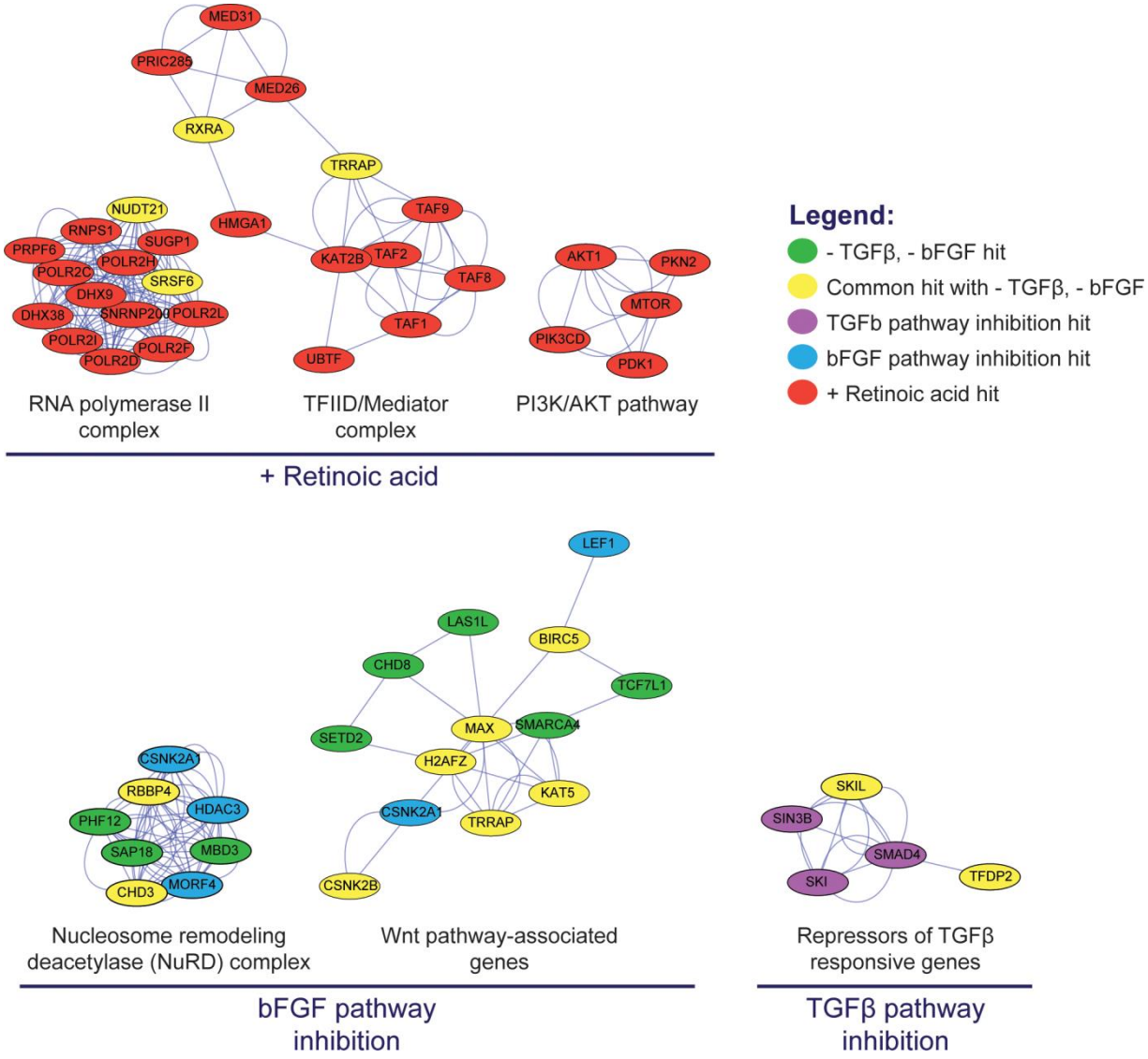



**Figure 21. Protein interaction networks of hits.** Protein-protein interaction networks of genes that are uniquely enriched for the different screening conditions. Node color indicates the screening conditions wherein the gene was identified as a hit.

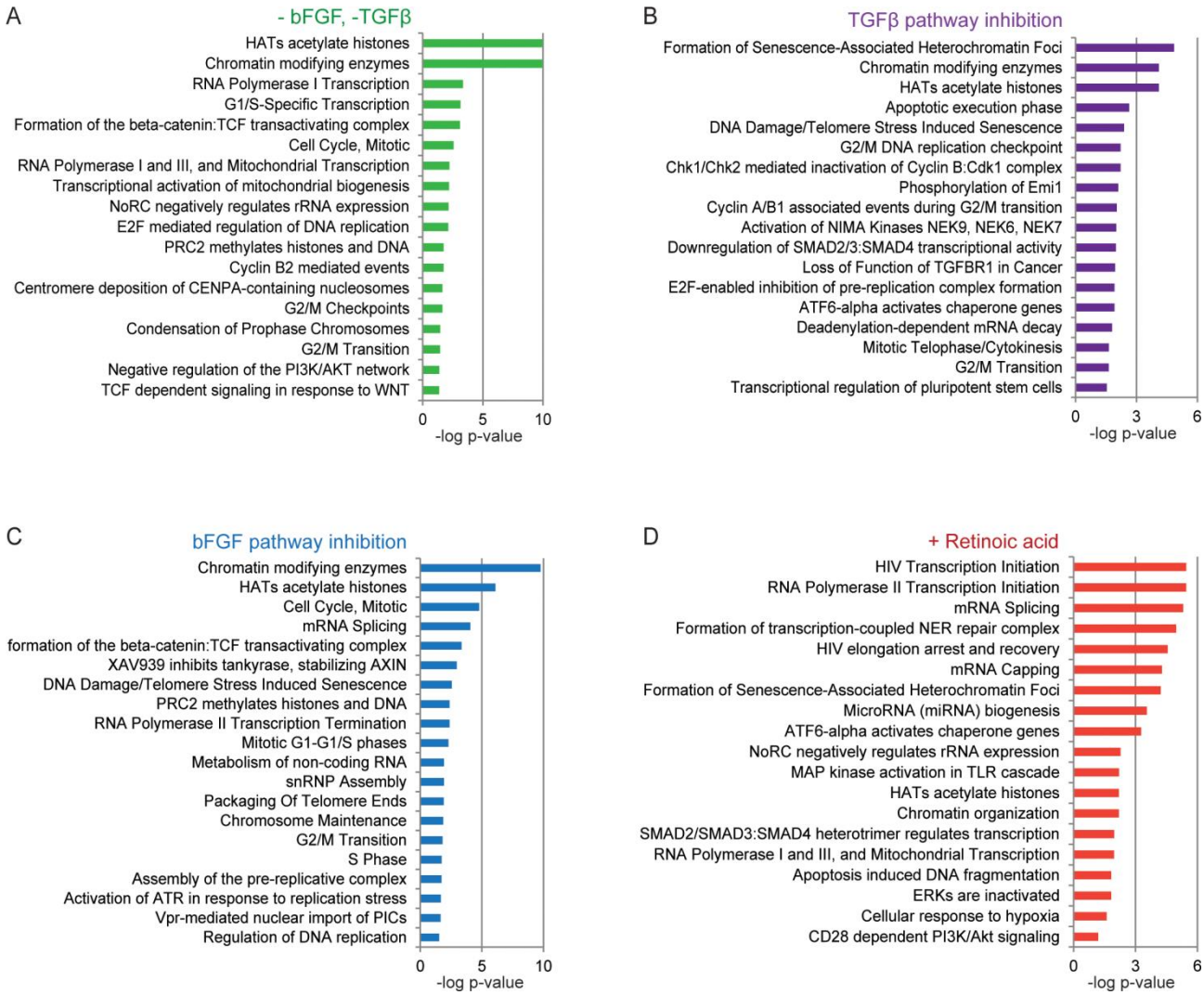



**Figure 22. Reactome analysis.** Top pathways that are overrepresented ($p<0.05$) among the hits as determined using the web resource Reactome in the (**A**) - bFGF, - TGFβ, (**B**) TGFβ pathway inhibition, (**C**) bFGF pathway inhibition and (**D**) + Retinoic acid conditions conditions are shown.

B. The NuRD complex in the exit from pluripotency

More interestingly, some processes seem to be more important in the exit from pluripotency given specific differentiation triggers. For example, members of the NuRD were identified as hits for regulating the exit from pluripotency particularly during bFGF-MEK inhibition (Figure 21, 23C). The NuRD complex, especially its fundamental member Mbd3, has an established role in promoting PSD in mESCs (Kaji et al., 2006; Reynolds et al., 2012), and our results demonstrate the conservation

of this function in hESCs. Interestingly, while the LIF-Stat3 pathway opposes the action of the NuRD complex in mESCs (Hu and Wade, 2012), this role seems to be assumed by the bFGF-MEK pathway in hESCs.

In seeming contrast to our results, there are also specific NuRD complex members that function to promote the pluripotent state instead (Liang et al., 2008). These discrepancies potentially arise from the presence of multiple NuRD complex isoforms that play opposite functions in ESCs. Our screen provides a starting point to classify the function of different NuRD members, with our hits specifically being the ones important for the exit from pluripotency in hESCs.

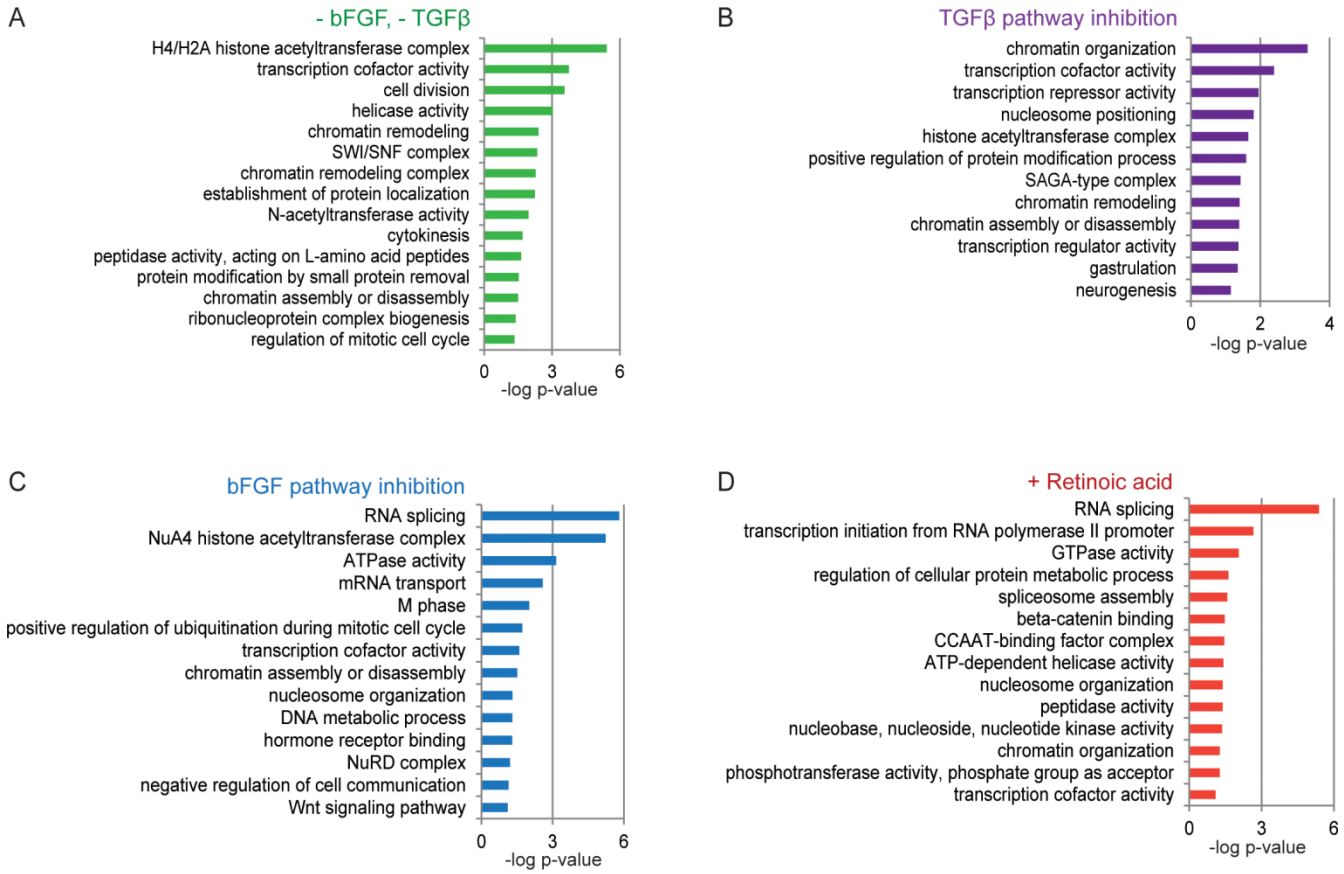



**Figure 23. Gene ontology analysis.** Top gene ontology functional clusters that are overrepresented ($p<0.05$) among the hits as determined using the web resource DAVID in the (**A**) - bFGF, - TGFβ, (**B**) TGFβ pathway inhibition, (**C**) bFGF pathway inhibition and (**D**) + Retinoic acid conditions conditions are shown.

C. Context-dependent function of development-related signalling pathways

Wnt-associated factors, particularly those involved with β-catenin, were also enriched in the bFGF pathway inhibition condition (Figure 21, 22C, 23C), indicating that the exit from pluripotency upon bFGF-MEK inhibition crosstalks with the canonical Wnt signalling pathway. While the role of Wnt signalling in hESC pluripotency has been controversial, most reports now agree that Wnt-β-catenin signaling works towards differentiation (Cai et al., 2007; Davidson et al., 2012; Dravid et al., 2005; Sumi et al., 2008; Ullmann et al., 2008). Our results support this current consensus, as the knockdown of canonical Wnt pathway members was able to promote the pluripotent state. The specificity of this observation to the context of inactive bFGF-MEK signalling is intriguing, as it suggests the existence of an additional layer of crosstalk between the two involved pathways. It is known that ERK reinforces the canonical Wnt pathway by phosphorylating and consequently inhibiting GSK3 (Ding et al., 2005; Singh et al., 2012). Disruption of ERK and canonical Wnt signalling might have a synergistic effect that converges on β-catenin regulation and perhaps other downstream connecting nodes.

Another notable pathway that highlights the divergent roles that pathways can play in different contexts is the PI3K pathway. The PI3K pathway is known to support pluripotency in hESCs by coordinating with both bFGF and TGFβ pathways (Li et al., 2007; Singh et al., 2012). Accordingly, we find negative regulators of this pathway to be enriched in the - bFGF, - TGFβ condition (Figure 22A). Yet, in the context of RA-induced exit from pluripotency, the PI3K seems to switch function to one antagonizing the pluripotent state (Figure 21, 22D), like its role downstream of RA in promoting differentiation in other cell types (Bastien et al., 2006; Qiao et al., 2012). This observation can be independently validated with LY294002, a small

molecule inhibitor of PI3K. Supplementation of LY294002 accelerated the decrease in pluripotency markers in the pluripotency medium and - bFGF, - TGFβ conditions, but prevented the exit from pluripotency in the + Retinoic acid condition (Figure 24A,B).

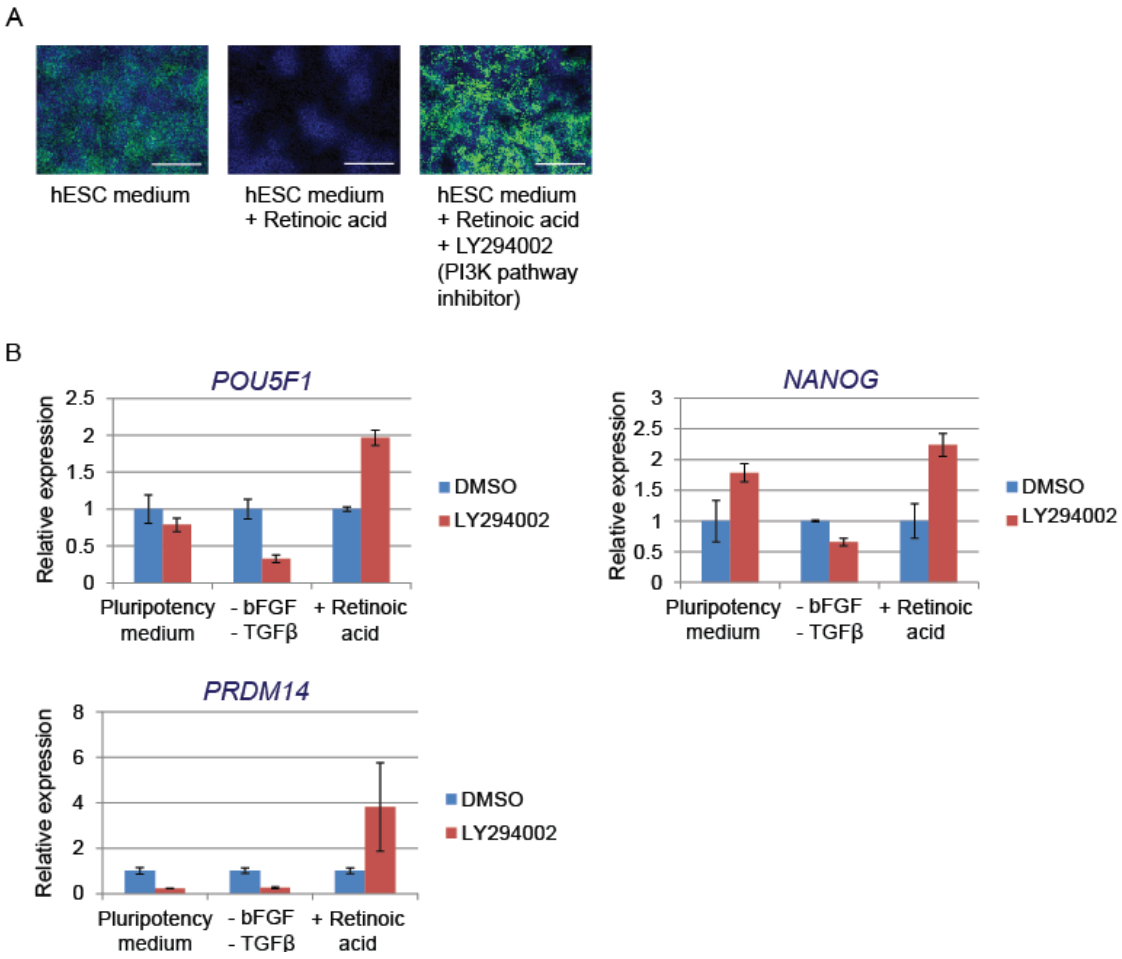



**Figure 24. Validation of context-dependent roles of the PI3K pathway with small molecules.** (**A**) Representative images for *NANOG*-GFP fluorescence (green) and nuclear staining (blue) after treatment with small molecules. Scale bar = 500μm. (**B**) Expression levels of pluripotency markers upon treatment with LY294002 (PI3K inhibitor) as measured by qPCR. All error bars denote standard deviation of triplicate data.

These results emphasize the importance of studying the role of developmental pathways in the proper context, as it is commonplace for these pathways to play opposite functions depending on environmental cues and crosstalk with other pathways (Kim et al., 2013; Singh et al., 2012).

D. The role of splicing in the exit from pluripotency

Interestingly, the RNA splicing machinery also seems to play a role in the exit from pluripotency, especially during retinoic acid addition and bFGF-MEK inhibition (Figure 22C-D, 23C-D). This is notable because although multiple RNA splicing factors have been reported to control hESC pluripotency (Gabut et al., 2011; Lu et al., 2013; Lu et al., 2014), it was unknown hitherto whether they also regulate its exit.

RNA splicing has been shown in other species to play a major role in regulating the Ras-MAPK-ERK cascade (Ashton-Beaucage et al., 2014; Shilo et al., 2014), which is pivotal to the role of bFGF in maintaining the human pluripotent state (Li et al., 2007; Xu et al., 2005). Similarly, retinoic acid has been reported to influence RNA splicing in human embryonic kidney cells through its interaction with Acinus (Wang et al., 2014). Similar mechanisms might be in place in hESCs, explaining why RNA splicing could affect the exit from pluripotency triggered specifically by bFGF pathway inhibition or retinoic acid introduction. Our study therefore opens the door for studying the crosstalk between the splicing machinery, the exit from pluripotency, and these developmental pathways.

In conclusion, independent analyses of the various screening conditions identify context-dependent processes that are crucial for the exit from pluripotency, in addition to demonstrating the robustness of our screen results.

## IV. Combined analysis of screen results reveals universal pathways governing the exit from pluripotency upon the withdrawal of self-renewal signals

### A. Hierarchical clustering of the four screening conditions

To get a general overview of the relationship between the four conditions, we performed hierarchical clustering analysis of the overall results from the four conditions. Each condition clustered within its own replicates, but away from other conditions (Figure 25). This reinforces the presence of diverse responses during the exit from pluripotency of hESCs given distinct differentiation triggers, which we have discussed in the previous section.

Interestingly, the - bFGF, - TGFβ, TGFβ pathway inhibition and bFGF pathway inhibition conditions weakly cluster together, but not with the + Retinoic acid condition (Figure 25). This was conceivable given that the + Retinoic acid condition introduces a differentiation signal in contrast to the withdrawal or inhibition of self-renewal signals in the other 3 conditions. This implies that the effectors of PSD vary when self-renewal signals are withdrawn versus when a differentiation signal is introduced, and that the mode of differentiation induction greatly influences PSD regulation.

Given the similarity between the aforementioned clustering conditions, we next performed a combined analysis of these three conditions to find central pathways that are important for the exit from pluripotency in the absence of self-renewal signals (Figure 26). The significance of epigenetic modification and the cell cycle in the exit from pluripotency is highlighted in this combined analysis, with members of multiple chromatin modifying complexes and cell cycle pathways being enriched.

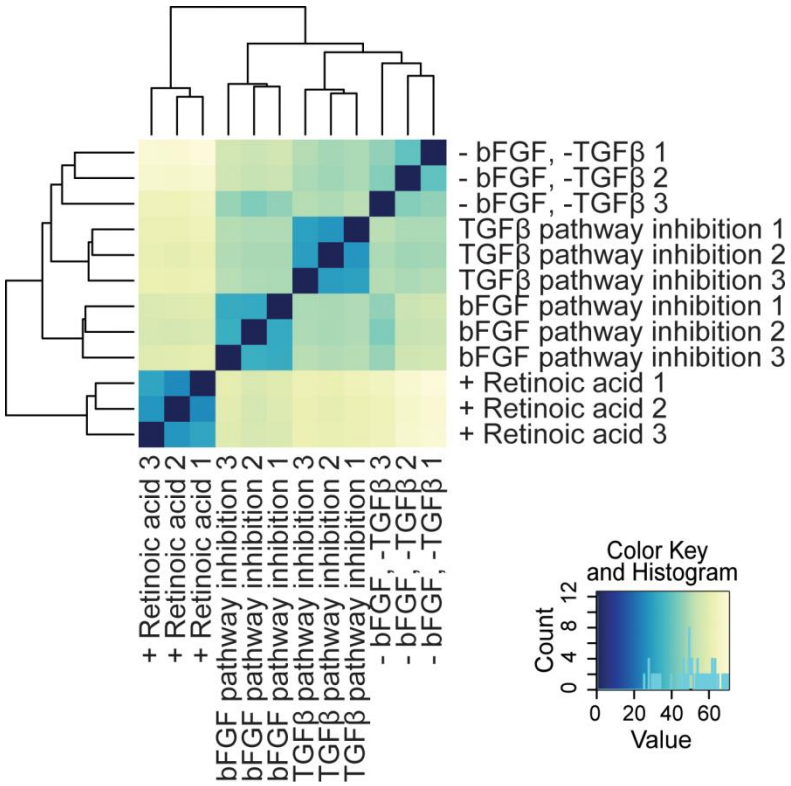



**Figure 25. Hierarchical clustering of screening conditions.** Heatmap depicting hierarchical clustering by Euclidean distance between different screening conditions. - bFGF, - TGFβ, TGFβ pathway inhibition and bFGF pathway inhibition results weakly cluster together and away from + Retinoic acid results.

B. Histone acetylation and the exit from pluripotency

Differentiation is often accompanied by changes in histone acetylation (Golob et al., 2008; Legartova et al., 2014), and accordingly, histone acetyltransferase (HAT) complex proteins appear to be top hits (Figure 26A-C), most prominently those that belong to SAGA-type and NuA4 (also known as Tip60-p400) HAT complexes. Markedly, the catalytic subunit of the NuA4 complex *KAT5* is among the top 5 hits of all three conditions (Appendices 3-5).

The NuA4 HAT complex has been previously implicated in the regulation of pluripotency, wherein its downregulation compromises the ability of mESCs to robustly give rise to differentiated cells (Chen et al., 2013; Fazzio et al., 2008). This feature of the NuA4 HAT complex seems to be conserved in hESCs. In contrast, the

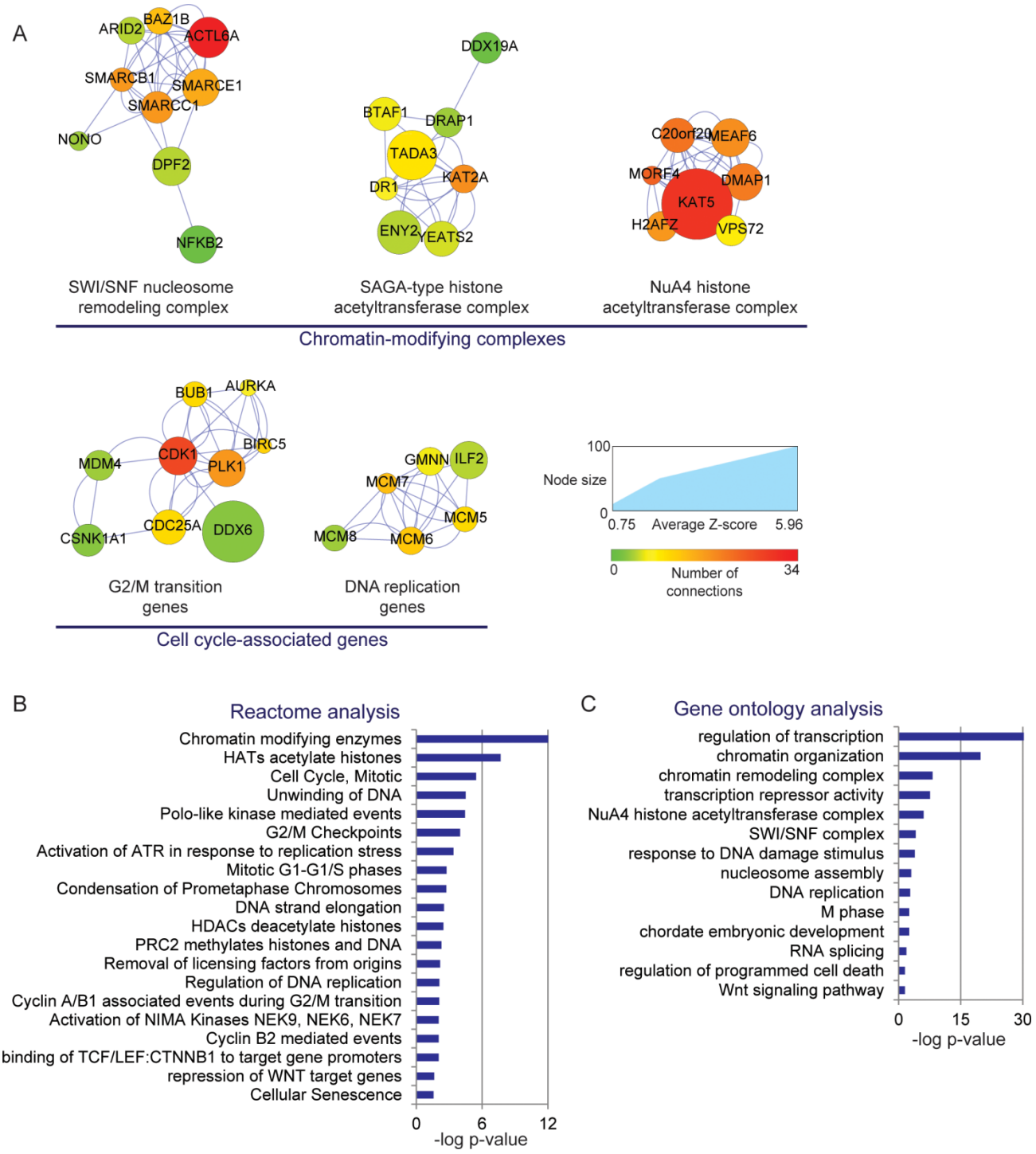



**Figure 26. Combined analysis of conditions wherein self-renewal signalling is withdrawn.** (**A**) Enriched gene clusters from the protein-protein interaction network analysis of the combined hits. Node size indicates the average z-score of the hits, while node color indicates the degree of integration (number of edges) of the gene with the entire network of hits. (**B**) Reactome analysis. Overrepresented pathways ($p<0.05$) from the combined hits as determined using the web resource Reactome are shown. (**C**) Gene ontology analysis. Representative terms ($p<0.05$) for enriched functional clusters from the combined hits are shown.

effect of SAGA-type HAT complexes on the exit from pluripotency seems to be unique for hESCs. While there is little literature on this topic, a study has demonstrated that knockout of the major SAGA catalytic subunit Gcn5 (or KAT2A in humans) in mESCs diminished their viability but not their differentiation capacity (Lin et al., 2007). In fact, the Gcn5-null mESCs even exhibited accelerated downregulation of a few pluripotency markers. Therefore, in contrast to the conserved function of the NuA4 HAT complex, species-specific differences might exist for the role of SAGA-type complexes in the exit from pluripotency. Interestingly, it has also been reported that Gcn5 depletion provokes cell cycle arrest at the G2/M checkpoint upon induction of differentiation (Lin et al., 2007), in line with our findings on the regulation of the exit from pluripotency by the cell cycle, which will be discussed later in this thesis.

Histone acetylation levels are higher in ESCs compared to somatic cells (Meshorer et al., 2006); we would therefore expect that differentiation is accompanied by histone deacetylation. Contrary to this expectation, two major HAT complexes were universally implicated in the exit from pluripotency upon withdrawal of self-renewal signals, while HDACs were not highly enriched in the combined screen hits. It thus seems that during the initiation of the exit from pluripotency, histone acetylation rather than deacetylation is required.

This could be attributed to the global but transient upsurge in histone acetylation that occurs at the onset of the exit from pluripotency (Golob et al., 2008; McCool et al., 2007). We hypothesize that this global increase in histone acetylation levels is an essential part of the exit from pluripotency that occurs specifically at differentiation gene loci. This enables ESCs to expansively acquire expression of genes from across multiple lineages, thus conferring pluripotency. Blockade of this

histone acetylation wave could thus inhibit exit from pluripotency due to the failure to upregulate differentiation-associated gene expression. Alternatively, these complexes might target specific loci, likely other effectors of the exit from pluripotency, necessitating their activity to complete the process of the exit from pluripotency. While these complexes target a broad range of histone marks and loci, there is a definite degree of specificity by these complexes that make them uniquely important for the exit from pluripotency, as other histone acetylation complexes such as HBO1 and PCAF complexes were not detected as hits in the screen. However, there is no study to date that has identified the loci targeted by either of these complexes in hESCs; this will be the next important step to understand how these HAT complexes regulate the exit from pluripotency by determining the identity of the genes they regulate.

C. The SWI/SNF nucleosome remodelling complex in the exit from pluripotency

Another chromatin modifying complex implicated in our study is the SWI/SNF nucleosome remodelling complex (Figure 26A,C). Multiple members of the SWI/SNF complex that were detected as hits in our screen have been demonstrated to function in the differentiation of multiple cell types (de la Serna et al., 2006). Specifically, the SWI/SNF factor BRG1 (encoded by *SMARCA4*) has been reported to antagonize Myc activity in cancer cells (Romero et al., 2012). The SWI/SNF complex might promote the exit from pluripotency by shutting down the Myc module, since Myc is well-known to be important for self-renewal of ESCs (Smith et al., 2010; Varlakhanova et al., 2010). However, there also exists seemingly contrasting evidence, where the SWI/SNF complex was regarded to be important for pluripotency maintenance, not exit (Zhang et al., 2014). Specifically, they report that BRG1

negatively modulates H3K27ac levels in multiple lineage-specific gene loci, thus contributing to pluripotency maintenance. However, the same study also showed that a unique SWI/SNF composition containing BAF170 (encoded by *SMARCC2*) but excluding BAF155 (encoded by *SMARCC1*) was responsible for this effect (Zhang et al., 2014). Interestingly, our screen identified *SMARCC1* but not *SMARCC2* to be enriched in the hit list. This means that different compositions of the same complex perform divergent functions in hESCs, and the screen results enable us to sift out which members of the SWI/SNF complex are specifically important for the exit from pluripotency in hESCs.

Notably, the role of these various complexes in the exit from pluripotency seems to be highly conserved, as inhibition of these complexes have been shown to impair differentiation in mouse embryonic stem cells (mESCs) (Chen et al., 2013; Lin et al., 2007) and cancer cells (Romero and Sanchez-Cespedes, 2013). Certain members of these complexes, such as *Trrap* of the NuA4 HAT complex and *Arid1a* of the SWI/SNF complex, have also been found to be crucial for exit from pluripotency of mESCs (Betschinger et al., 2013). Thus, our study systematically identified several chromatin-modifying complexes to be universally important for the exit from hESC pluripotency upon withdrawal of self-renewal signals.

D. <u>The cell cycle in the exit from pluripotency</u>

Genes involved in cell cycle regulation are among the most enriched in conditions where self-renewal signals were withdrawn (Figure 26A-C). Strikingly, the cell cycle-associated hits are mostly involved in the G2-to-M transition or in DNA replication during the S phase (Figure 26A). On the contrary, we found no strong

enrichment of processes specific to the G0/G1 phase and mitotic progression. This suggests that the exit from pluripotency is gated by specific mechanisms hardwired to cell cycle progression. We therefore decided to focus on examining how the exit from pluripotency is regulated by the cell cycle machinery, which will be discussed in detail later in this thesis.

V. Combined analysis of screen results additionally identify genes that enhance the exit from pluripotency

While the focus of this project is to look for genes that are important for the exit from pluripotency, the high-throughput RNAi screen also enabled the identification of genes that function in the opposite direction. Hence, we likewise performed a combined analysis of genes whose knockdown enhanced the exit from pluripotency (negative z-scores) in the three clustering conditions (Figure 27).

A. Known guardians of the pluripotent state

As expected, direct knockdown of members of the core pluripotency network itself accelerated the exit from pluripotency (Figure 27A-B). This analysis also encouragingly sieved out positive regulators of the chief pathways of the maintenance of hESC pluripotency, namely the TGFβ and bFGF signalling pathways (Figure 27A-C). This is in parallel to the enrichment of negative regulators of these same pathways in the hits that prevent the exit from pluripotency. Hence, we see that genes with opposite functions yielded opposite scores in the RNAi screen, underlining the robustness of the RNAi screen results.

B. Chromatin modification in the maintenance of pluripotency

It is well-known that cell fate maintenance and transitions are highly dependent on the epigenetic state of the cell. This is strikingly exemplified by the importance of chromatin modifying complexes in promoting the exit from pluripotency as discussed in the previous section. Likewise, a different set of chromatin modifiers are essential for upholding the pluripotent state.

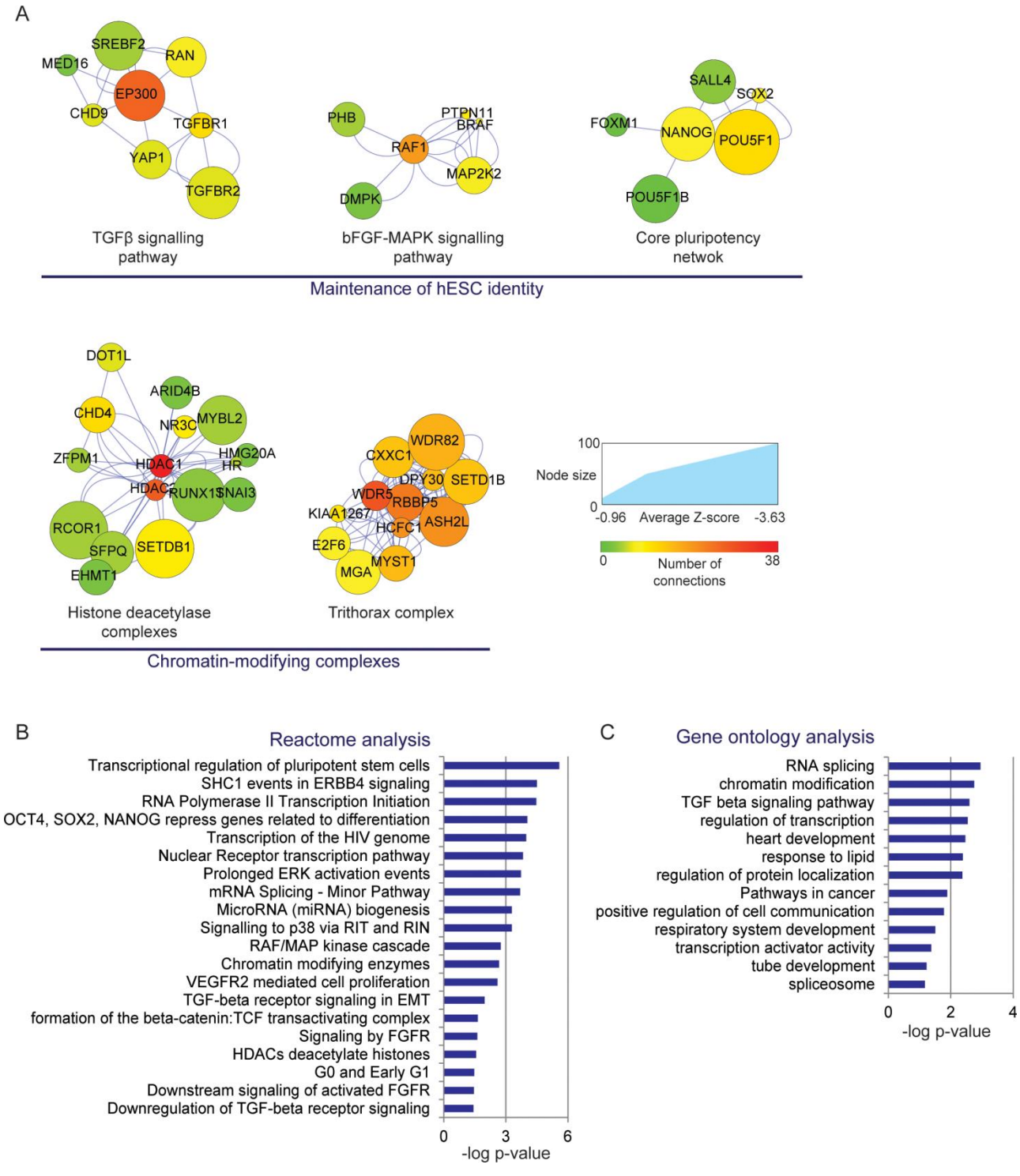



**Figure 27. Combined analysis of hits that enhance the exit from pluripotency (negative z-scores).** (**A**) Enriched gene clusters from the protein-protein interaction network analysis of the combined hits. Node size indicates the average z-score of the hits, while node color indicates the degree of integration (number of edges) of the gene with the entire network of hits. (**B**) Reactome analysis. Overrepresented pathways ($p<0.05$) from the combined hits as determined using the web resource Reactome are shown. (**C**) Gene ontology analysis. Representative terms ($p<0.05$) for enriched functional clusters from the combined hits are shown.

Particularly, members of histone deacetylase (HDAC) complexes were enriched in the hits (Figure 27A). Whilst the knockdown of HATs attenuated the exit from pluripotency as discussed earlier, knockdown of HDACs accelerated the exit from pluripotency. This reiterates that histone acetylation is a central process necessary for the exit from pluripotency to manifest.

A distinct epigenetic feature of the pluripotent state is the presence of bivalent domains, characterized by the presence of both the repressive histone 3 lysine 27 trimethylation mark (H3K27me3) and the activating histone 3 lysine 4 trimethylation mark (H3K4me3) (Azuara et al., 2006; Bernstein et al., 2002; Pan et al., 2007; Zhao et al., 2007). Bivalent domains are hence believed to mark developmental genes in a poised state, which are primed for rapid and immediate induction upon differentiation (Sha and Boyer, 2008). These modifications are deposited by the Polycomb repressive complex 2 (PRC2) and the Trithorax complex, respectively. Although both complexes are conceivably crucial for the integrity of the pluripotent state, only the PRC2 complex has been extensively studied to be essential for the maintenance of pluripotency (Boyer et al., 2006; Lee et al., 2006). Nevertheless, the Trithorax core member Ash2l has recently been implicated in the maintenance of mESCs (Wan et al., 2013). Our screen results highlight the role of the Trithorax complex in specifically in promoting hESC pluripotency, as knockdown of multiple Trithorax group proteins all resulted in the accelerated exit from pluripotency (Figure 27A).

In conclusion, our high-throughput RNAi screen has enabled the identification of both positive and negative regulators of the exit from pluripotency. The results from this study will henceforth serve as a unique resource for the stem cell field to explore in detail the mechanisms underlying the exit from pluripotency.

VI. The cell cycle exerts profound control over the exit from pluripotency

We have found that many genes involved in cell cycle regulation affect the exit from pluripotency, specifically those functioning during the S and G2 phases of the cell cycle (Figure 26A-C). Specifically, we have found 14 genes that are involved in DNA replication, and 14 genes involved in the G2/M transition.

During proliferation, cells experience dramatic biochemical and physical changes, which cells have evolved to utilize in order to prime and regulate other events that are not immediately related to proliferation. Biochemical differences within various cell cycle states were shown to extend their function to regulate immune response, metabolism and lineage specification (Handschick et al., 2014; Lee et al., 2014; Pauklin and Vallier, 2013; Rodier et al., 2009). Despite increasing evidence showing the cell cycle regulating other cellular process, there is no direct and functional evidence that cell cycle states can control the pluripotency network and its dissolution. As the screen results seem to suggest that the core cell cycle machinery may enforce a deterministic regulation on pluripotency and differentiation in hESCs, we focused on examining how the exit from pluripotency is regulated by the cell cycle machinery.

A. Validation of S- and G2-associated hits

We first sought to validate the prevention of the exit from pluripotency by knockdown of S- and G2-associated hits. In the - bFGF, -TGFβ condition, knockdown of all hits examined can remarkably be validated with 2 independent shRNAs (Figure 20B). Knockdown of majority of these genes similarly caused retention of *NANOG*-GFP fluorescence in the TGFβ pathway inhibition and bFGF pathway inhibition

conditions (Figure 28). Therefore, we are confident that these S- and G2-associated genes truly play a role in the exit from pluripotency.

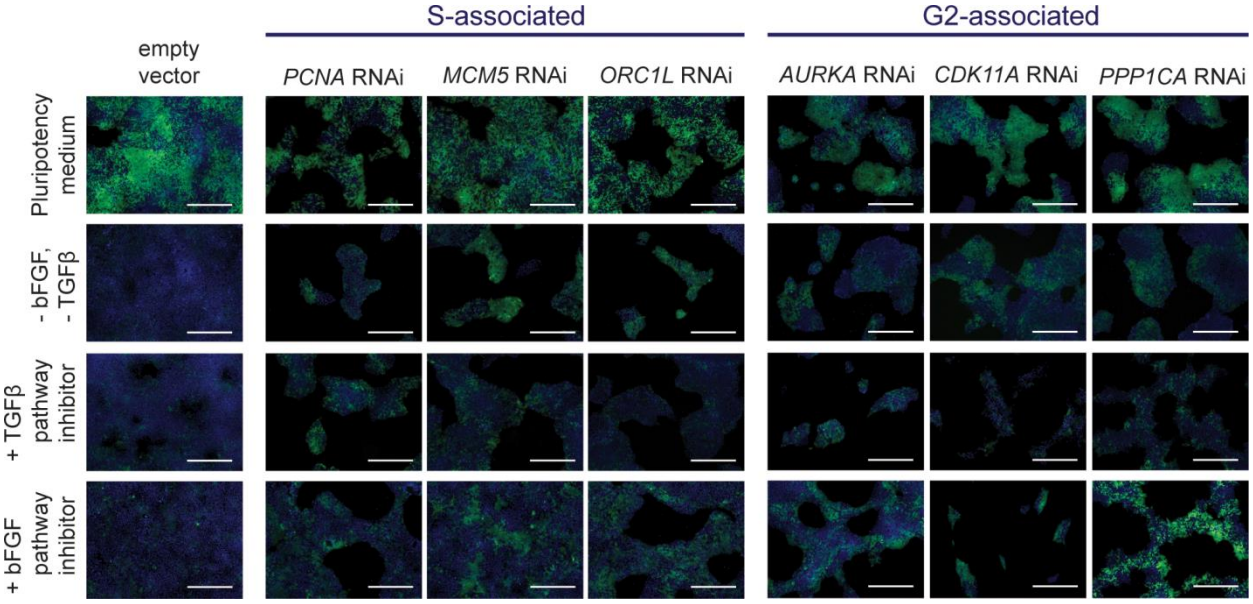



**Figure 28. Cell cycle-associated hit knockdown in conditions wherein self-renewal signalling is withdrawn.** Representative images for *NANOG*-GFP fluorescence (green) and nuclear staining (blue) after RNAi knockdown of select cell cycle-associated hits. Scale bar = 500μm.

B. Perturbation of cell cycle progression by knockdown of hits

We next checked if their knockdown affects cell cycle progression in hESCs. Knockdown of the S-associated hits led to the elevation of γH2AX foci (Figure 29), which results from the activation of the replication checkpoint upon perturbation of DNA replication. This can be mimicked by treatment of hESCs with the DNA polymerase inhibitor Aphidicolin (Figure 29), which is proven to block DNA replication during the S phase (Pedrali-Noy et al., 1980). On the other hand, knockdown of G2-associated hits increased the relative time spent by hESCs in the

G2 phase (Figure 30). Thus, we confirm that knockdown of the S- and G2-associated hits impedes progression at the S and G2 phases, respectively.

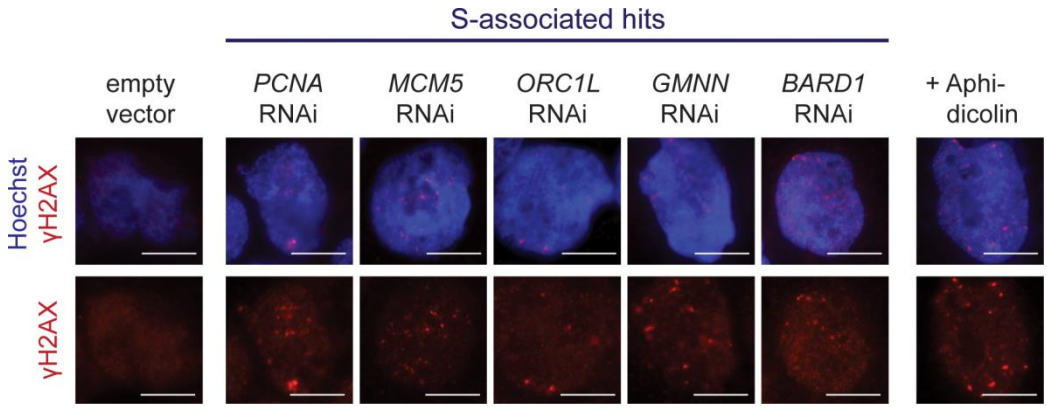



**Figure 29. Presence of γH2A.X foci upon knockdown of S-associated hits.** Representative images of nuclei (blue) and anti-γ-histone 2A.X immunofluorescence staining (red) of hESCs after RNAi knockdown of replication-associated hits or treatment with Aphidicolin in the pluripotency medium. Scale bar = 10μm. Staining was done with the help of Dr Liang Hongqing.

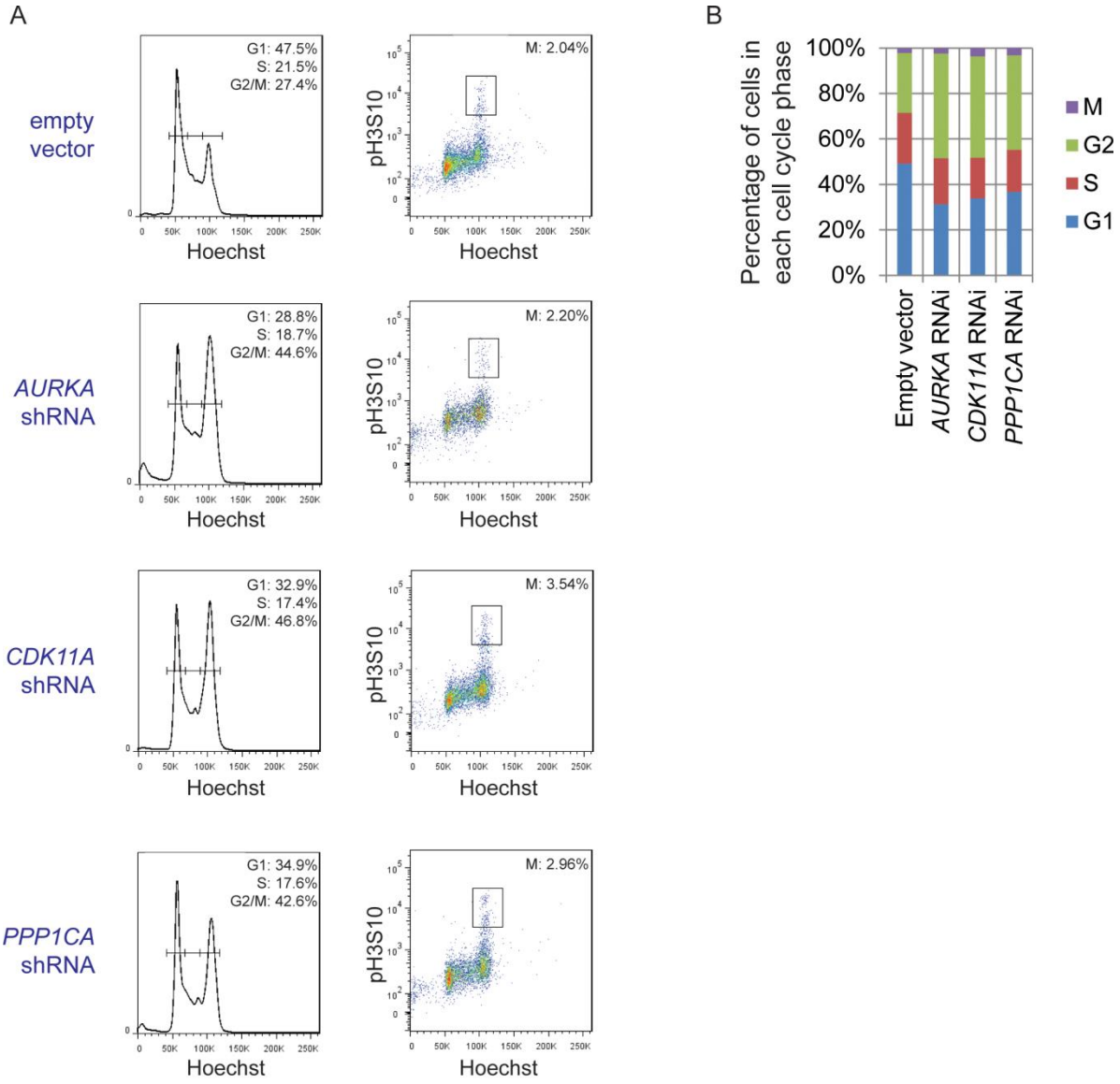



**Figure 30. Cell cycle profile of hESCs upon knockdown of G2-associated hits.** (**A**) FACS profiles and (**B**) Flow cytometry quantification indicating the cell cycle status of hESCs after RNAi knockdown of G2-associated hits in the pluripotency medium.

C. Perturbation of cell cycle progression by small molecules

To validate that resistance to the exit from pluripotency can indeed be conferred by specific cell cycle changes, we employed a collection of chemical inhibitors known to perturb various cell cycle pathways. With extensive dosage titration, we managed to enrich but not completely arrest hESCs in the various cell cycle phases (Figure 31) without inducing extensive apoptosis. In concordance with

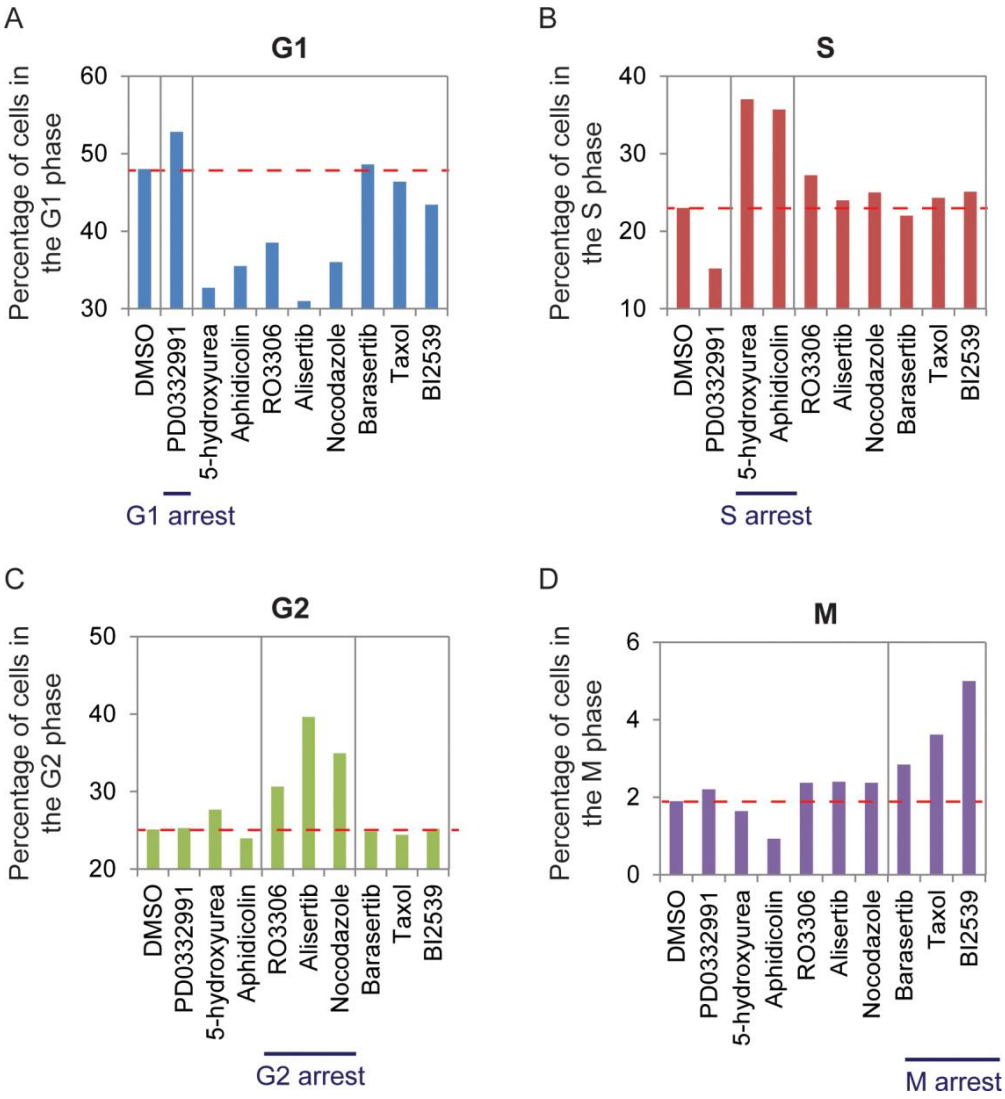



**Figure 31. Cell cycle profile of hESCs upon treatment with cell cycle inhibitors.** Flow cytometry quantification indicating the percentage of hESCs in the (**A**) G1, (**B**) S, (**C**) G2 and (**D**) M phases of the cell cycle after treatment with various chemicals known to perturb the cell cycle in the pluripotency medium.

our screen results, chemicals that perturb DNA replication in the S phase (5-hydroxyurea, Aphidicolin) or delay progression through the G2 phase by inhibiting mitotic entry (RO3306, Alisertib, Nocodazole) can significantly preserve the expression of pluripotency markers (Figure 32A-B) and N*ANOG*-GFP fluorescence (Figure 32C) when self-renewal signals were withdrawn. In contrast, inhibitors that led to enrichment in other cell cycle phases did not exhibit such an effect (Figure 30). Importantly, these observations are not exclusive to the hESC line used, as the same trend is also observed in other hESC lines (Figure 32D).

Therefore, we provide evidence through both genetic and chemical means that hESCs inevitably preserve their pluripotent state despite withdrawal of self-renewal signals when their progression through the S and G2 phases is hindered (Figure 33). In contrast, the exit from pluripotency ensues despite perturbations in the G1 and M phase (Figure 33). These suggest that the exit from pluripotency is strongly governed by specific factors or pathways within the cell cycle machinery.

Some studies have previously shown that hESCs in the G1 phase are more sensitive to the initiation of differentiation compared to S and G2 phase cells (Pauklin and Vallier, 2013; Sela et al., 2012). Yet, this is the first demonstration that perturbation of the cell cycle at the S and G2 phases is sufficient to attenuate the exit from pluripotency in hESCs, providing proof-of-concept that manipulation of the hES cell cycle can deterministically influence its cell fate decisions.

Moreover, cells in the M phase were excluded or pooled into the other phases in the previous studies, but we show here that perturbation at the M phase does not prevent exit from pluripotency. Multiple characteristics of the M phase prevent it from enforcing cell fate transitions: its very short duration compared to other cell

cycle phases (Becker et al., 2006; Gordon and Lane, 1980), the lack of transcriptional activity throughout the phase (Gottesfeld and Forbes, 1997; Martinez-Balbas et al., 1995; Prescott and Bender, 1962), and its inherent intolerance to perturbations. Thus, it is unlikely that there are mechanisms that take place for promotion of either pluripotency or differentiation in the M phase. Instead, the observations for cells in the M phase are probably an extension from the following G1 phase. Mitotic arrest results in transcriptional inhibition and DNA damage, that ends up either in apoptosis or mitotic slippage into a senescent G1 phase (Blagosklonny, 2007; Lanni and Jacks, 1998). As our experiments for the M phase only involved chemical perturbations, most of the surviving cells are probably the ones which have slipped into the G1 phase, which we demonstrate here to be similarly amenable to the exit from pluripotency.

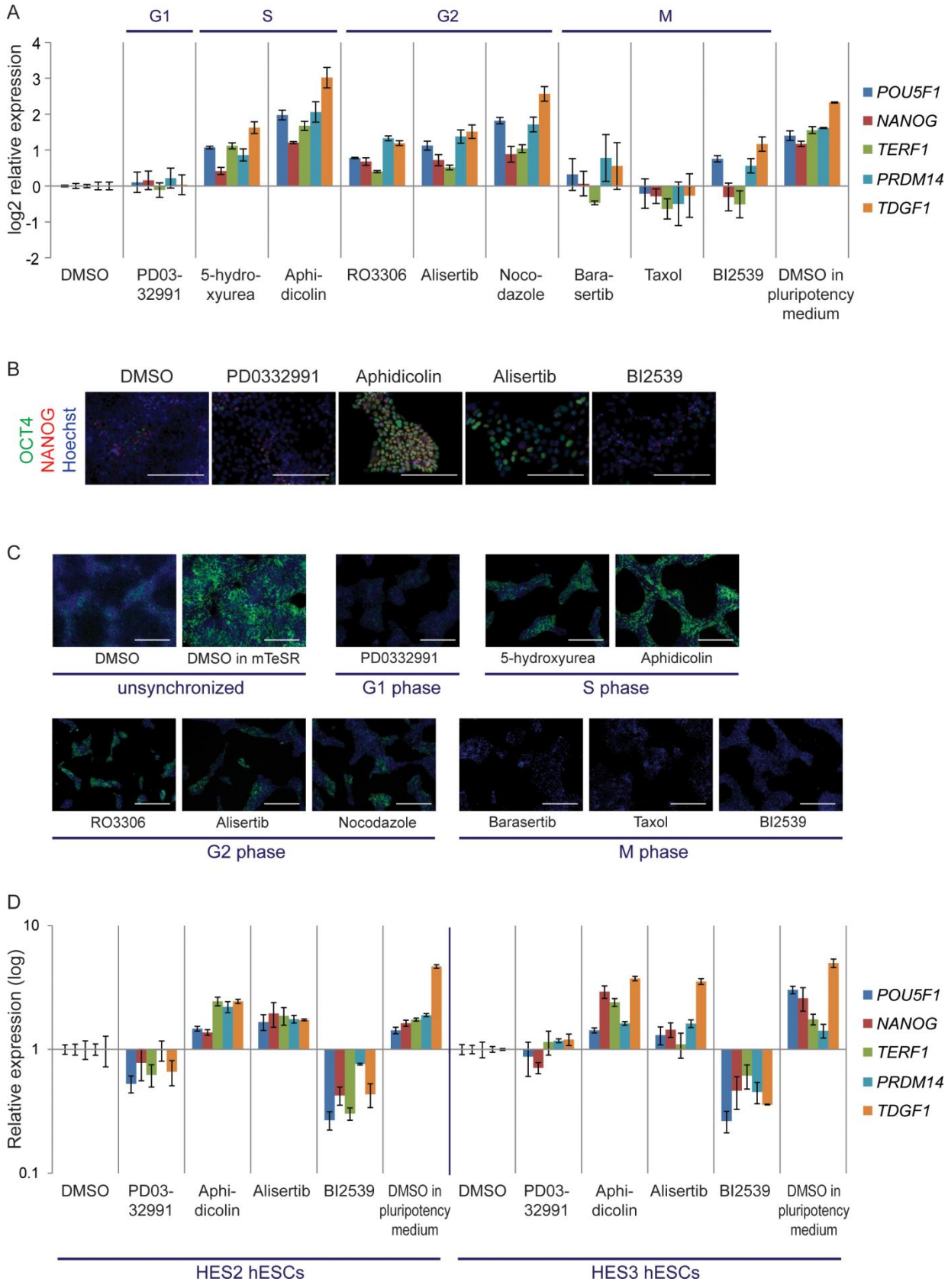



**Figure 32. Effects of various cell cycle inhibitors on the exit from pluripotency.** (**A-C**) Expression levels of pluripotency markers upon treatment with cell cycle inhibitors in the - bFGF, - TGFβ condition as measured by (**A**) qPCR, (**B**) immunofluorescence staining (Scale bar = 200um) and (**C**) *NANOG*-GFP fluorescence (Scale bar = 500μm). (**D**) Expression levels of pluripotency markers upon treatment of additional hESC lines with cell cycle inhibitors in the - bFGF, - TGFβ condition as measured by qPCR. All error bars denote standard deviation of triplicate data.

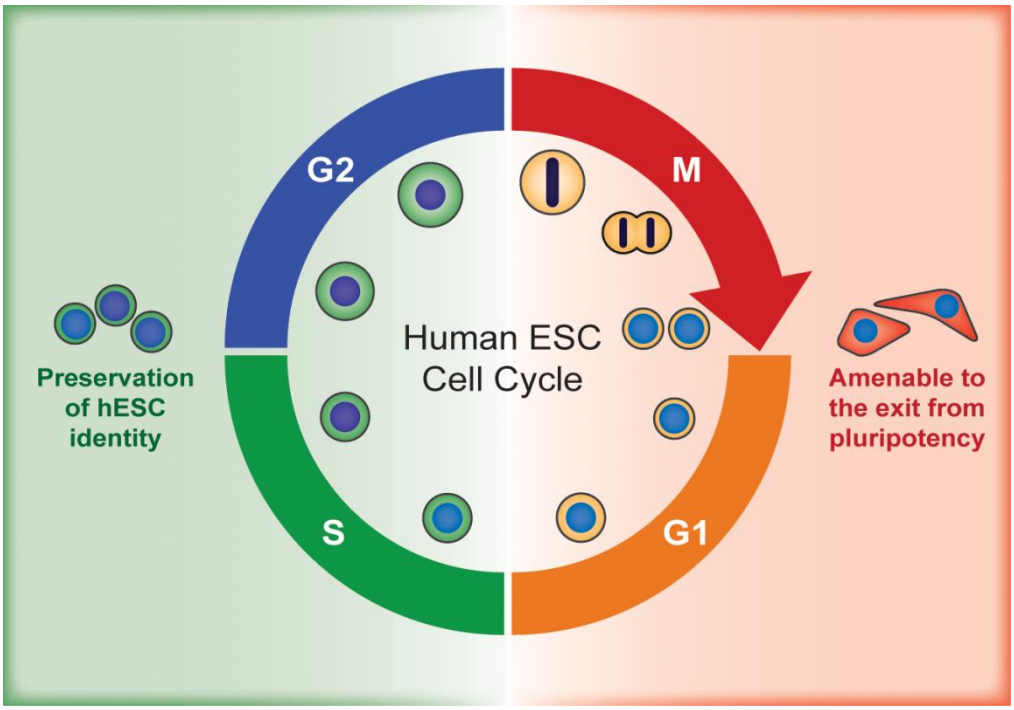



**Figure 33. Model for cell cycle dependency of the exit from pluripotency.** hESCs are amenable to the exit from pluripotency during the G1 and M phases, while they tend to preserve their identity and prevent the exit from pluripotency during the S and G2 phases.

## VII. G1-specific factors do not deterministically regulate the exit from pluripotency

Embryonic stem cells spend the majority of their time in the S and G2 phases of the cell cycle, in contrast to differentiated cells which spend more time in the G1 phase (Coronado et al., 2013; Hindley and Philpott, 2013; Singh and Dalton, 2009). The current paradigm is that G1 is the only phase permissive to differentiation, attributed to a higher expression of differentiation markers (Singh et al., 2013) and the enrichment of cell cycle factors like CDK inhibitors and Cyclin D that contribute to lineage specification (Li et al., 2012a; Pauklin and Vallier, 2013). In the other phases of the cell cycle, hESCs are believed to be passively retained in the pluripotent state due to the "lack" of a differentiation response pathway. Hence, one could argue that prolongation of the S and G2 phases is simply a result of blocked access to the G1 phase.

### A. Expression levels of G1-associated factors upon perturbation of S and G2 phases

To test if the delayed exit from pluripotency resulting from cell cycle perturbation at the S and G2 phases was due to an inaccessibility to G1-associated factors, we firstly examined the expression levels of major G1-specific factors that are implicated in differentiation (Li et al., 2012a; Pauklin and Vallier, 2013). Neither Cyclin D1 nor p21 protein was downregulated upon knockdown of S and G2 phase-related hits (Figure 34), indicating that the consequent prevention of the exit from pluripotency is not due to an inaccessibility to these G1-associated factors.

### B. Knockdown of G1-specific factors

We next knocked down factors required for G1 phase progression such as Cyclin D1/2 or CDK4/6. If these factors mediate the exit from pluripotency, their

depletion is expected to attenuate the exit from pluripotency. If other G1-specific factors are instead required, their depletion should enhance the exit from pluripotency due to the resulting elongation of the G1 phase. However, neither of these were observed as knockdown of any of these factors had no consistent effect on pluripotency gene expression (Figure 35), suggesting that G1-associated factors do not deterministically govern the shutdown of the pluripotency network.

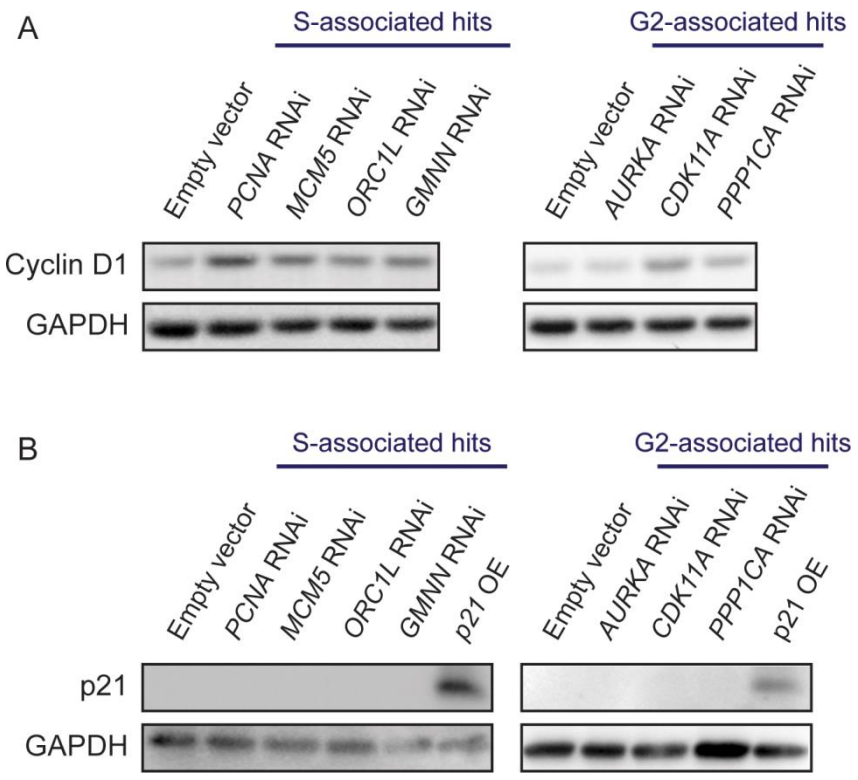



**Figure 34. Expression levels of G1-specific factors upon S and G2 phase perturbation.** Western blot of (**A**) Cyclin D1 and (**B**) p21 levels upon knockdown of S- and G2-associated hits. No downregulation of either protein was observed.

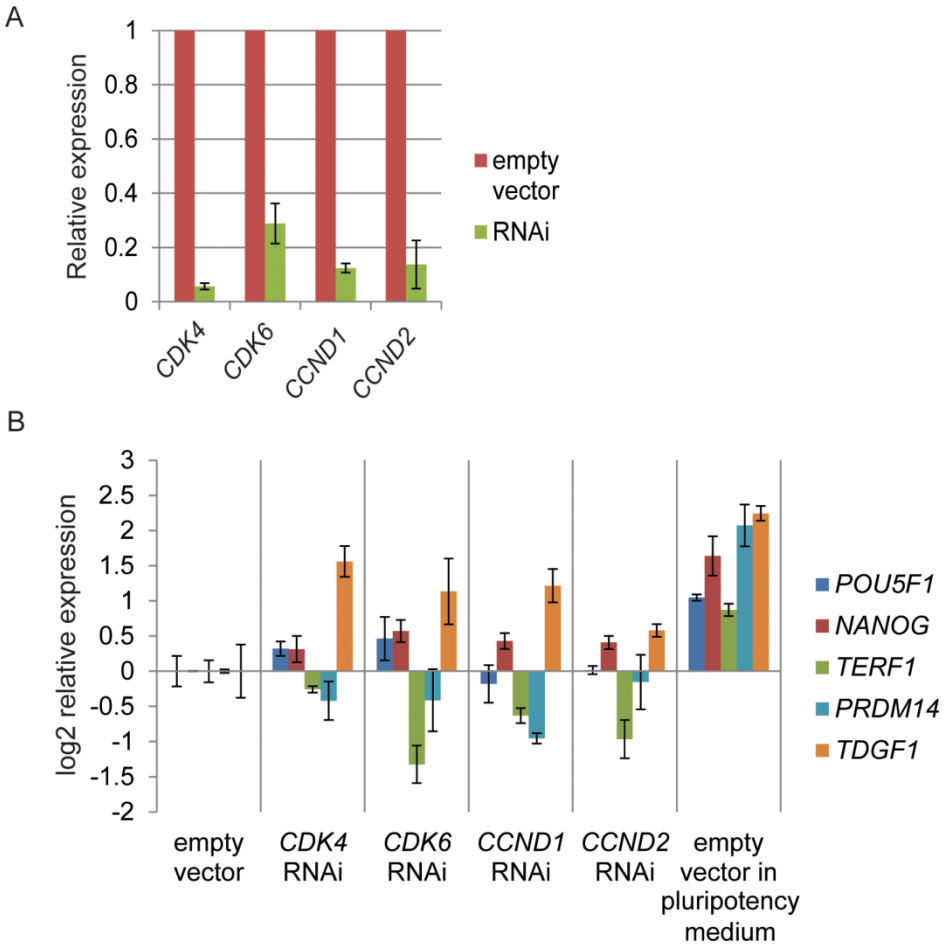



**Figure 35. Effect of G1-associated factor knockdown on the exit from pluripotency.** (**A**) Transcript levels of respective genes upon RNAi knockdown as measured by qPCR. All RNAi constructs effectively decreased expression of their target gene. (**B**) Expression levels of pluripotency markers upon knockdown of G1-associated genes in the - bFGF, - TGFβ condition as measured by qPCR. All error bars denote standard deviation of triplicate data. Experiments were done with the help of Dr Liang Hongqing.

C. Overexpression of G1-specific factors

To finalize that G1-specific factors do not directly regulate the exit from pluripotency, we examined whether the shutdown of pluripotency marker expression can be enhanced by ectopic overexpression of certain G1-specific factors that are implicated in differentiation. Direct overexpression of either Cyclin D1 or p21 did not significantly affect the exit from pluripotency (Figure 36), in line with the prior observations.

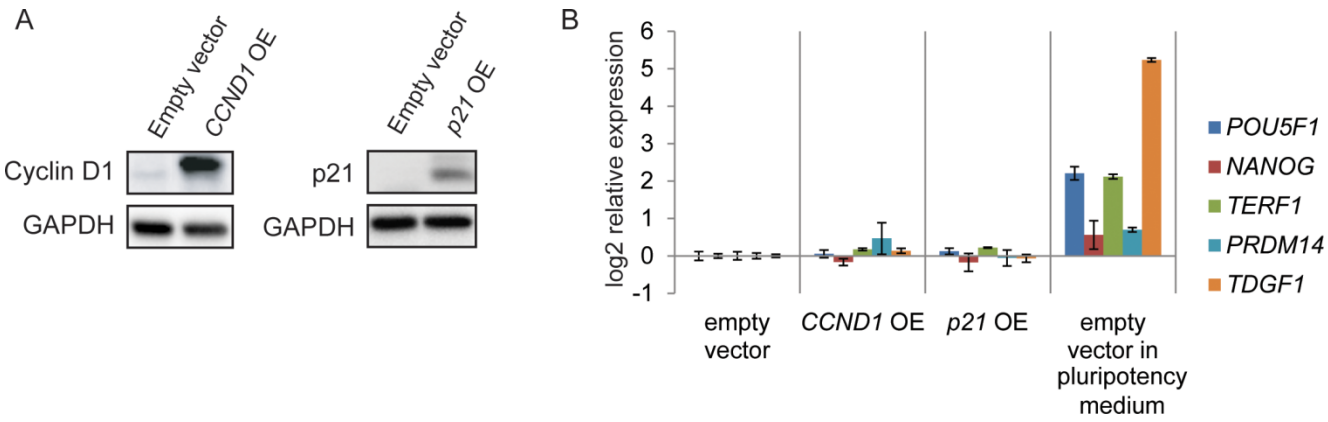



**Figure 36. Effect of G1-associated factor overexpression on the exit from pluripotency.** (**A**) Protein levels of respective genes upon overexpression as measured by Western blot. Both overexpression constructs effectively increased expression of their target gene. (**B**) Expression levels of pluripotency markers upon overexpression of G1-associated genes in the - bFGF, - TGFβ condition as measured by qPCR. All error bars denote standard deviation of triplicate data. Experiments were done with the help of Dr Liang Hongqing.

Altogether, these suggest that although G1-associated factors were reported to contribute to lineage specification, they are not determinants in shutting down the pluripotency network in this context. These findings in hESCs highly resemble previous findings in mESCs, in which a longer G1 phase is coupled to differentiation but targeting of G1-specific factors does not affect the exit from pluripotency (Li et al., 2012c). Therefore, the attenuation of the exit from pluripotency by S and G2 phase perturbation cannot be completely explained to be driven by G1-specific factors, breaking the G1-centric paradigm of pluripotency control by the cell cycle. Instead, we hypothesize that pathways which are independent of G1 but are overstimulated by either DNA replication error or G2 phase perturbation may play a more deterministic role in the exit from pluripotency.

VIII. Activation of the DNA replication checkpoint in the S phase attenuates the exit from pluripotency

To test the hypothesis that the cell cycle-dependent regulation of the exit from pluripotency might stem from a direct effect of S- and G2-specific pathways, we looked for specific cell cycle machineries that could crosstalk with the pluripotency network. One major outcome of DNA replication perturbation during the S phase is the activation of the DNA replication checkpoint (Bakkenist and Kastan, 2003; Zou and Elledge, 2003), which halts the cell cycle until the replication fork stalling and the consequent DNA damage are resolved (Recolin et al., 2014). In brief, replication checkpoint signalling is mainly mediated by the sensor kinases ATM and ATR, which phosphorylate the downstream kinases CHEK1 and CHEK2 (Figure 37). Substrates of

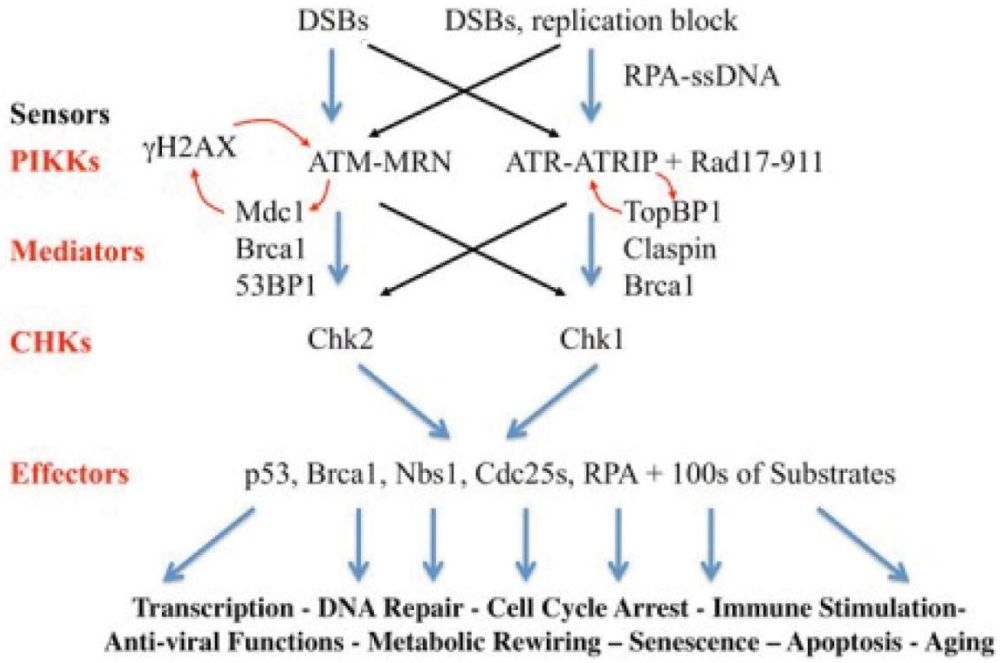



**Figure 37. The DNA replication checkpoint pathway.** Activation of the DNA replication checkpoint activates the sensor kinases ATM and ATR, which in turn, activates CHEK1 and CHEK2 directly or through several mediators. The CHEKs activate a multitude of effectors to regulate various processes including transcription, cell cycle progression and DNA repair. Adapted from Elledge Lab 2014 (http://elledgelab.med.harvard.edu) with permission from Dr. Steve Elledge.

both ATM/ATR and the CHEKs are then activated, coordinating processes such as DNA repair, cell cycle progression, apoptosis and gene expression to rectify errors before resumption of DNA replication.

A. Activation of the DNA replication checkpoint

To ensure that the DNA replication checkpoint is activated in our experiments, we looked at phosphorylation levels of the effector kinases CHEK1 and CHEK2 upon exposure to DNA polymerase inhibitor Aphidicolin. Aphidicolin treatment increased levels of both pCHEK1 and pCHEK2 (Figure 38). Treatment with the pan-CHEK inhibitor AZD7762 further enhanced phosphorylation CHEK1 and CHEK2 (Figure 36), resulting from negative feedback. Furthermore, activation of the checkpoint is substantiated by the appearance of γH2AX foci upon Aphidicolin treatment or S-associated hit knockdown (Figure 29). These results certify that the replication checkpoint is indeed activated upon stalling of DNA replication.

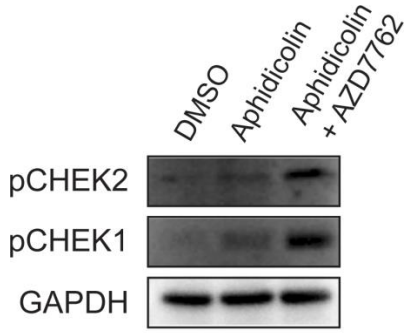



**Figure 38. Checkpoint kinase activation.** Protein levels of phosphorylated CHEK1 and CHEK2 upon treatment with replication inhibitor Aphidicolin and checkpoint inhibitor AZD7762. Experiments were done with the help of Dr Liang Hongqing.

B. Preservation of the pluripotency state by the ATM/ATR-CHEK2 axis

We next inspected if checkpoint activation plays a dominant effect on suspending the exit from pluripotency. To accomplish this, checkpoint signalling was abolished in hESCs through the inhibition of ATM and ATR using Caffeine or RNAi. Under these conditions, pluripotency markers and *NANOG*-GFP fluorescence were downregulated upon withdrawal of self-renewal signals despite replication perturbation by Aphidicolin (Figure 39). Similarly, pan-inhibition of the effector kinases CHEKs rescued the exit from pluripotency from Aphidicolin (Figure 39C-D). Interestingly, knockdown of CHEK2 but not CHEK1 released the block on the exit from pluripotency (Figure 39A,B,D), suggesting that CHEK2 is the major effector of sustaining the pluripotent state.

Since abolishment of the DNA replication checkpoint restored the ability of hESCs to exit the pluripotent state, we examined whether this also resulted in the restoration of cell cycle progression. In all these cases, inhibition of the DNA replication checkpoint pathway did not alter cell cycle profiles significantly (Figure 40). With Aphidicolin, DNA replication remained largely delayed despite checkpoint abolishment (Figure 40), indicating that exit from pluripotency can be restored even though cell cycle progression remains halted. This excludes the possibility that cells are unable to differentiate as due to a simple S phase lock or the physical state of DNA during replication. These evidences from both genetic and chemical approaches independently validate our hypothesis that signalling through the ATM/ATR-CHEK2 axis, directly contributes to pluripotency.

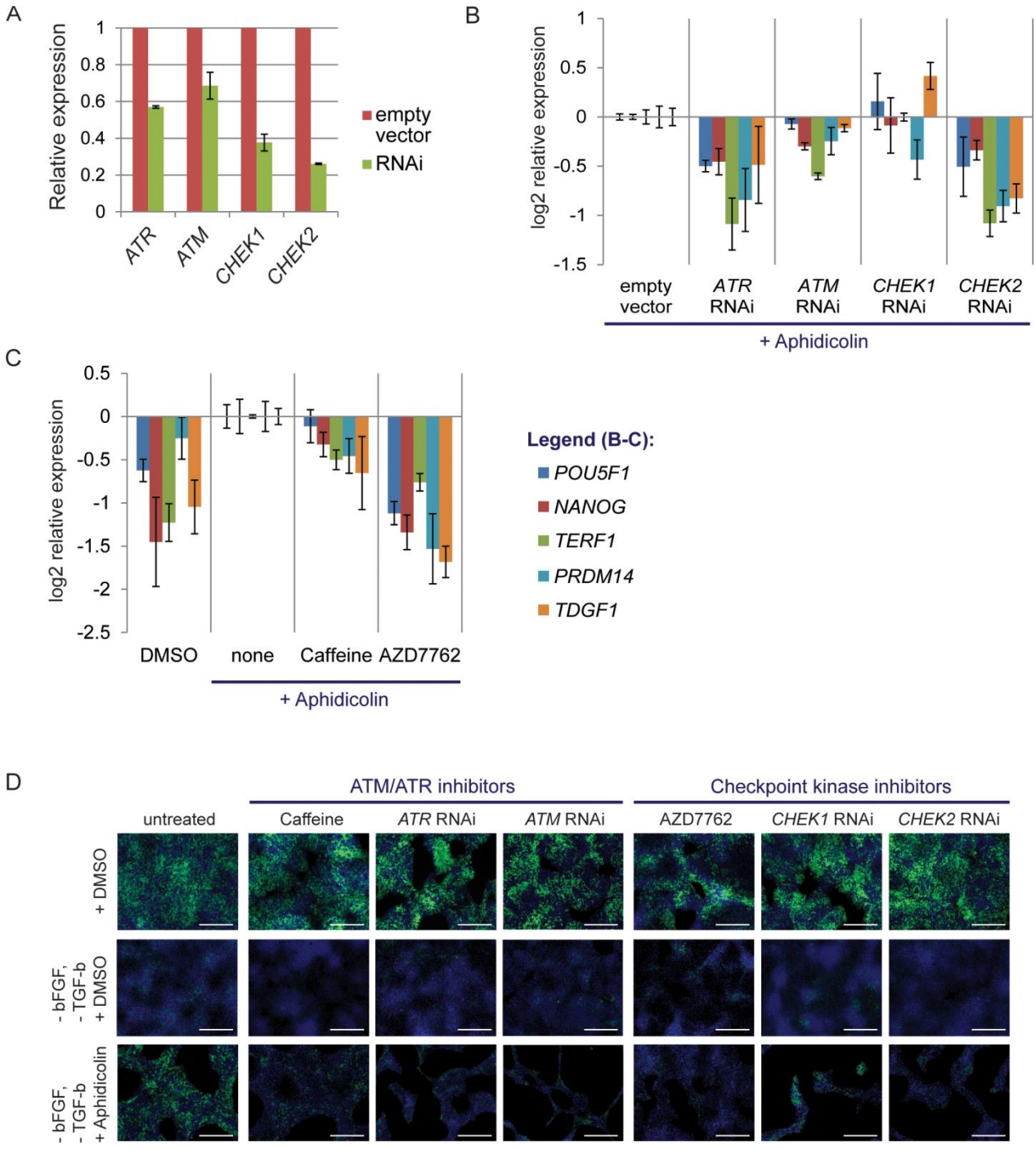



**Figure 39. Effect of replication checkpoint inhibition on the exit from pluripotency.** (**A**) Transcript levels of respective genes upon RNAi knockdown as measured by qPCR. All RNAi constructs effectively decreased expression of their target gene. (**B-C**) Expression levels of pluripotency markers upon (**B**) knockdown of replication checkpoint pathway members and (**C**) chemical inhibition of replication checkpoint signalling in the - bFGF, - TGFβ condition as measured by qPCR. All error bars denote standard deviation of triplicate data. (**D**) Representative images for *NANOG*-GFP fluorescence (green) and nuclear staining (blue) after inhibition of the replication checkpoint. Scale bar = 500μm.

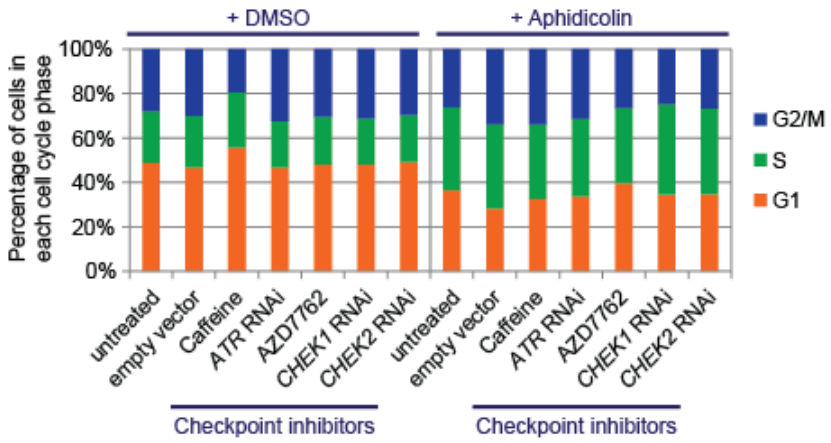



**Figure 40. Cell cycle profile of hESCs upon abolishment of the replication checkpoint.** Inhibition of the replication checkpoint pathway does not alter cell cycle progression as seen from flow cytometry quantification indicating the cell cycle profiles of hESCs.

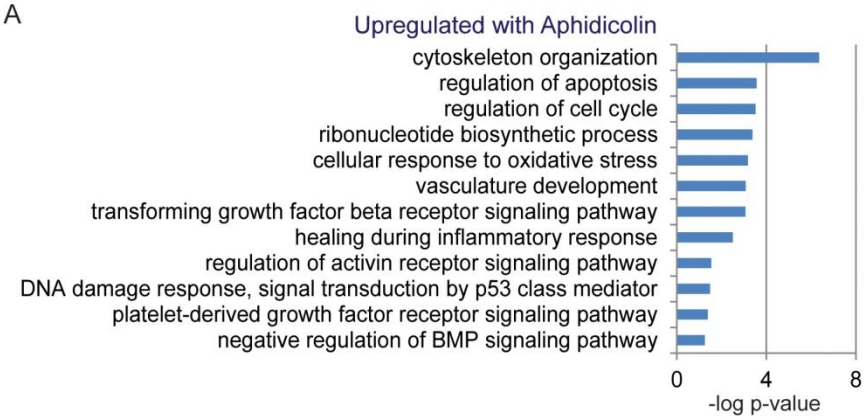
During proliferation, stem cells can encounter various situations like replication stress, DNA damage and checkpoint activation. How stem cells make their fate choices in response to these special but prevalent cell cycle events have direct consequences on genome stability, tissue development and stem cell maintenance. From our experiments probing the exit from pluripotency upon DNA damage, we demonstrate that the pluripotent state is reinforced in the presence of replication stress via the ATM/ATR-CHEK2 axis. This is astonishing given that adult stem cells commit to differentiation upon checkpoint activation (Wang et al., 2012). We theorize that this could be a beneficial adaptation in hESCs because: (1) ESCs are known to express higher levels of homologous recombination and damage repair proteins (Roos et al., 2007; Tichy et al., 2010), which promotes resolution of DNA damage caused by replication stress, and (2) ESCs are able to enter apoptosis more efficiently in case the damage cannot be rectified, due to a lower apoptotic threshold in hESCs compared to differentiated cells (Dumitru et al., 2012; Liu et al., 2013). Hence, preservation of the pluripotent state in the presence of DNA damage decreases the chances of giving rise to a mutant progeny of differentiated cells.

C. Enhancement of TGFβ signalling by the DNA replication checkpoint

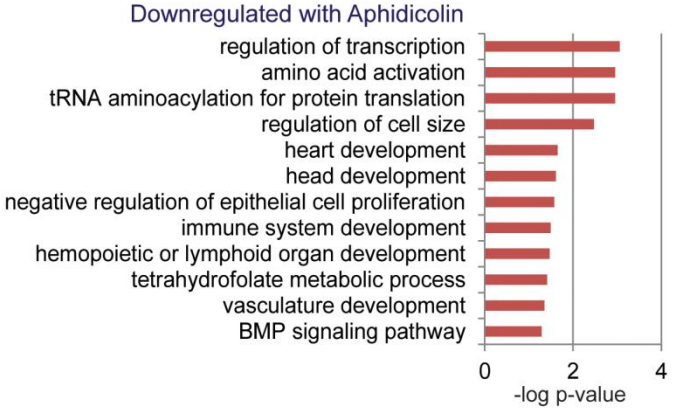
To obtain a deeper mechanistic connection between ATM/ATR-CHEK2 activation and pluripotency, we performed a time-course microarray analysis in hESCs treated with Aphidicolin and AZD7762. After 48 hours of Aphidicolin treatment, we observed upregulation of genes involved in cell cycle regulation, apoptosis and other housekeeping processes, as would be expected from induction of cell cycle arrest (Figure 41A,B). Notably, we also saw an upregulation of genes positively regulating TGFβ signalling and a simultaneous downregulation of BMP4 pathway genes and TGFβ pathway antagonists (Figure 41A-C), which we validated by qPCR (Figure 41D). The TGFβ pathway has a well-known role in promoting the human pluripotent state (Beattie et al., 2005; Xu et al., 2008). Specifically, TGFβ signalling directly regulates *NANOG* transcription (Vallier et al., 2009; Xu et al., 2008), which we observed to be upregulated following the changes in TGFβ-related gene expression (Figure 41C-E). Hence, this offers a potential explanation of how the ATM/ATR-CHEK2 axis could uphold pluripotency.

Comparison of time-course expression levels demonstrated that TGFβ-related gene expression changes occur earlier and to a greater degree compared to the pluripotency marker expression changes that follow the withdrawal of self-renewal signals (Figure 41C-D). This confirms that the changes in TGFβ-related genes occur upstream of the pluripotency network. We also observed the same changes in TGFβ-related gene expression upon stalling of DNA replication even when hESCs are in the pluripotency medium (Figure 41E), further validating the causative effect of DNA replication perturbation on TGFβ-related gene expression.

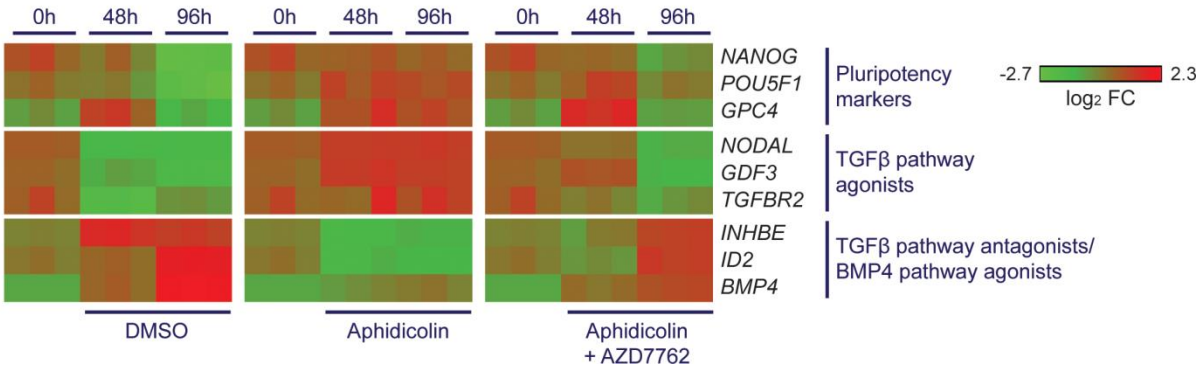
A


Upregulated with Aphidicolin
cytoskeleton organization
regulation of apoptosis
regulation of cell cycle
ribonucleotide biosynthetic process
cellular response to oxidative stress
vasculature development
transforming growth factor beta receptor signaling pathway
healing during inflammatory response
regulation of activin receptor signaling pathway
DNA damage response, signal transduction by p53 class mediator
platelet-derived growth factor receptor signaling pathway
negative regulation of BMP signaling pathway
0
4
8
-log p-value

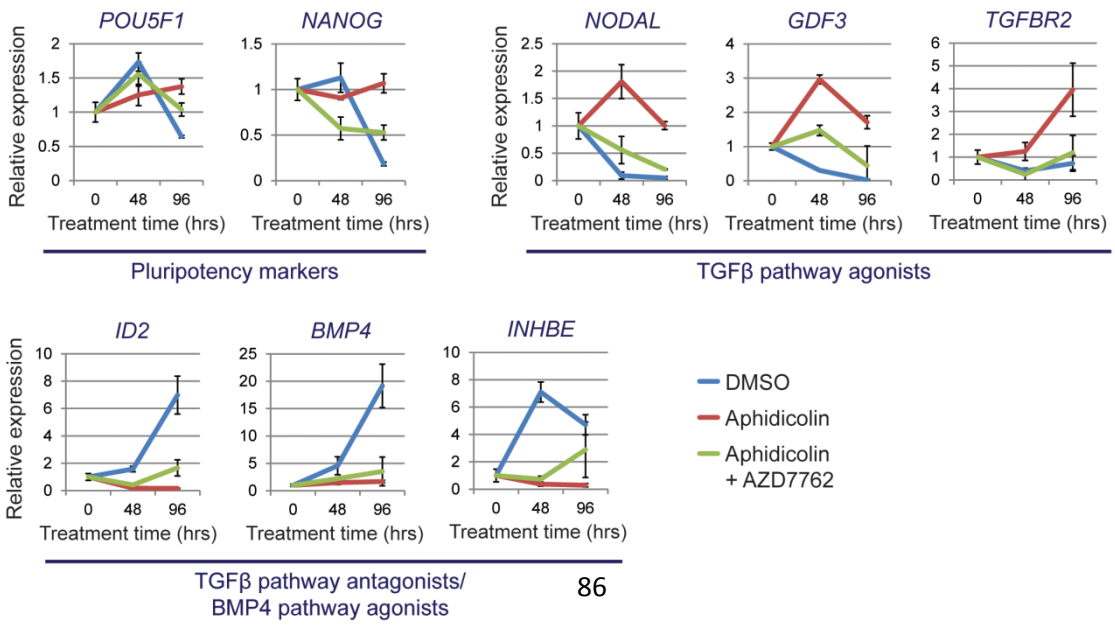
B


Downregulated with Aphidicolin
regulation of transcription
amino acid activation
tRNA aminoacylation for protein translation
regulation of cell size
heart development
head development
negative regulation of epithelial cell proliferation
immune system development
hemopoietic or lymphoid organ development
tetrahydrofolate metabolic process
vasculature development
BMP signaling pathway
0
2
4
-log p-value

C


0h
48h
96h
NANOG
POU5F1
GPC4
Pluripotency markers
-2.7
2.3
log2 FC
NODAL
GDF3
TGFBR2
TGFβ pathway agonists
INHBE
ID2
BMP4
TGFβ pathway antagonists/ BMP4 pathway agonists
DMSO
Aphidicolin
Aphidicolin + AZD7762

D


POU5F1
NANOG
NODAL
GDF3
TGFBR2
ID2
BMP4
INHBE
Relative expression
Treatment time (hrs)
Pluripotency markers
TGFβ pathway agonists
TGFβ pathway antagonists/ BMP4 pathway agonists
DMSO
Aphidicolin
Aphidicolin + AZD7762

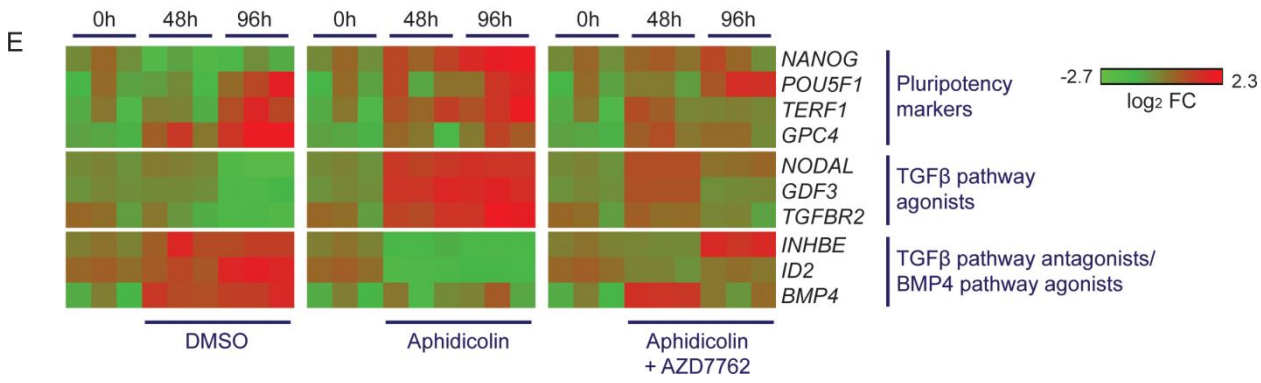



**Figure 41. Time-course microarray analysis of replication checkpoint-activated hESCs.** (**A-B**) Gene ontology analysis. Enriched biological process terms ($p<0.05$) for (**A**) upregulated and (**B**) downregulated genes identified in the microarray analysis upon Aphidicolin (DNA polymerase inhibitor) treatment for 48 hours in the - bFGF, - TGFβ condition. (**C**) Microarray heatmap for differentially expressed genes upon treatment with Aphidicolin and AZD7762 (CHEK1/2 inhibitor) in the - bFGF, -TGFβ condition. (**D**) Time-course quantitative PCR for pluripotency and TGFβ-related genes upon treatment with Aphidicolin and AZD7762 in the - bFGF, - TGFβ condition. Error bars denote standard deviations of triplicate data. (**E**) Microarray heatmap for differentially expressed genes upon treatment with Aphidicolin and AZD7762 in pluripotency medium. Microarray experiments were done with the help of Yeo Jia-Chi.

Importantly, these changes in TGFβ pathway gene expression lead to enhanced TGFβ signalling activation as indicated by increased SMAD2 phosphorylation (Figure 42). Heightened TGFβ signalling can be reversed by treatment with AZD7762 (Figure 41, 42) without releasing stalled replication (Figure 40), indicating that the ATM/ATR-CHEK2 axis directly augmented TGFβ signalling.

Moreover, although TGFβ gene expression patterns are expected to change with cellular state, they do not simply correlate with pluripotency status (Figure 43). This again confirms that the changes in TGFβ pathway gene expression are a consequence of checkpoint activation rather than of delayed exit from pluripotency.

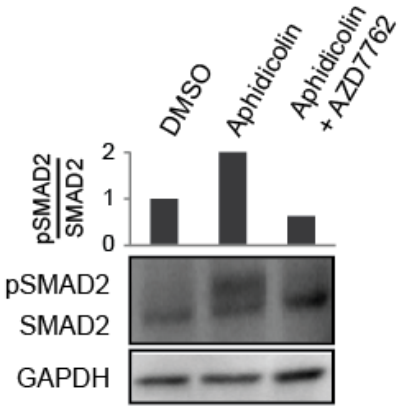



**Figure 42. Phospho-SMAD2 levels during replication perturbation.** Western blot of SMAD2 and phospho-SMAD2 from hESCs in the - bFGF, - TGFβ condition. SMAD2 phosphorylation is increased upon Aphidicolin (DNA polymerase inhibitor) treatment. Treatment with AZD7762 (CHEK1/2 inhibitor) prevents the increase in SMAD2 phosphorylation. Experiment was done with the help of Dr Liang Hongqing.

Collectively, these data strongly argue that in the presence of DNA replication perturbation, activation of the ATM/ATR-CHEK2 cascade alters expression of TGFβ-related genes. This ultimately enhances TGFβ pathway signalling, which consequently sustains *NANOG* expression, preserves the hESC state and prevents the exit from pluripotency. Alteration of cytokine signalling pathways by the ATM/ATR-CHEK2 axis is not farfetched, as it also regulates the interleukin-6 pathway in human fibroblasts (Rodier et al., 2009). Here we provide novel evidence that in hESCs, the ATM/ATR-CHEK2 axis uniquely affects the TGFβ pathway to regulate pluripotency.

D. <u>The role of p53 in the ATM/ATR-CHEK2-TGFβ cascade</u>

Given that the ATM/ATR-CHEK2 axis regulates the TGFβ pathway by regulating transcript levels of associated genes, we ought to determine the mediating factor by looking for transcription factors that satisfy two criteria: 1) it must act downstream of the ATM/ATR-CHEK2 axis and 2) it has binding sites near TGFβ-related genes. We first examined p53, a key transcription factor activated downstream

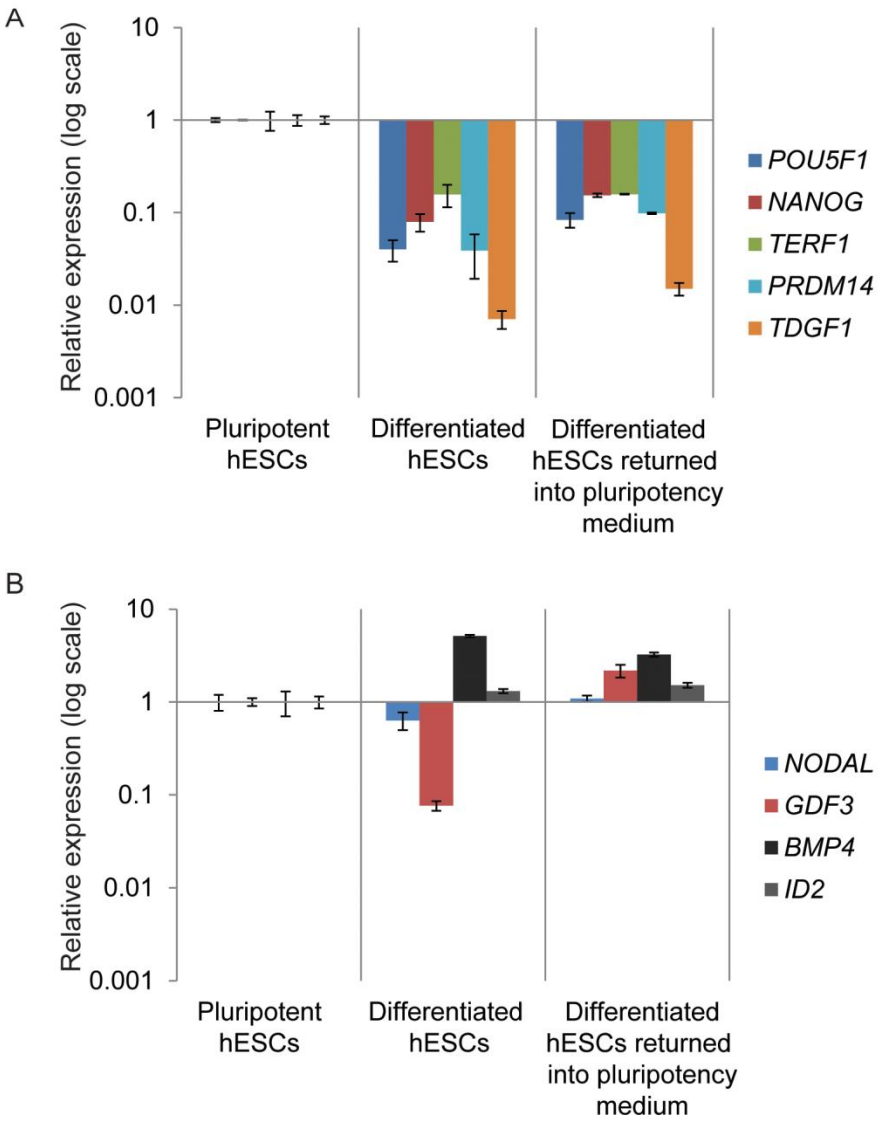



**Figure 43. Comparison between changes in pluripotency and TGFβ-associated gene expression.** Expression levels of (**A**) pluripotency markers and (**B**) TGFβ agonists (color) and antagonists (grayscale) upon incubation either in pluripotency medium, in the - bFGF, - TGFβ condition, or in the - bFGF, - TGFβ followed by recovery in pluripotency medium. TGFβ-associated gene expression does not correlate with pluripotency marker expression.

of ATM/ATR-CHEK2 during DNA damage response (Banin et al., 1998; Hirao et al., 2000; Tibbetts et al., 1999). A recent study performing chromatin immunoprecipitation with massively parallel tag sequencing (ChIP-Seq) for p53 revealed binding sites near TGFβ-related genes in hESCs, albeit varied binding

depending on environmental conditions (Akdemir et al., 2014). Therefore, p53 could be the perfect candidate mediating ATM/ATR-CHEK2 with TGFβ signalling.

We first confirmed that p53 is activated in this context by looking at p53 protein levels. Both immunostaining and Western blot ascertain the upregulation of p53 levels upon DNA replication stalling of differentiating hESCs (Figure 44). Importantly, inhibition of the CHEK2 activity using AZD7762 reverted p53 levels, validating that the heightened p53 protein level is due to p53 stabilization by CHEK2 (Hirao et al., 2000).

We sought to determine if p53 activity is responsible for the effects of the ATM/ATR-CHEK2 axis on the exit from pluripotency by altering p53 levels in hESCs during DNA replication perturbation in the absence of self-renewal signals.

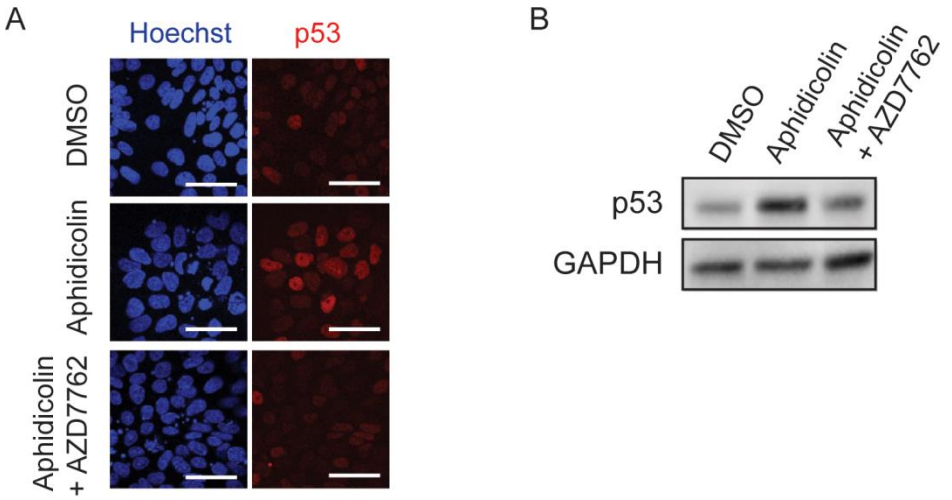



**Figure 44. Expression levels of p53 in replication-perturbed hESCs.** Protein levels of p53 as measured by (**A**) immunofluorescence staining (Scale bar = 50μm) and (**B**) Western blot upon treatment with Aphidicolin (DNA replication inhibitor) and AZD7762 (CHEK1/2 inhibitor). Experiments were done with the help of Dr Liang Hongqing.

In this condition, p53 knockdown reversed the altered expression of TGFβ pathway genes, enabling downregulation of pluripotency markers (Figure 45). Thus, p53 knockdown can overturn the effects of replication checkpoint activation, substantiating that p53 acts downstream of the ATM/ATR-CHEK2 axis in regulating the exit from pluripotency.

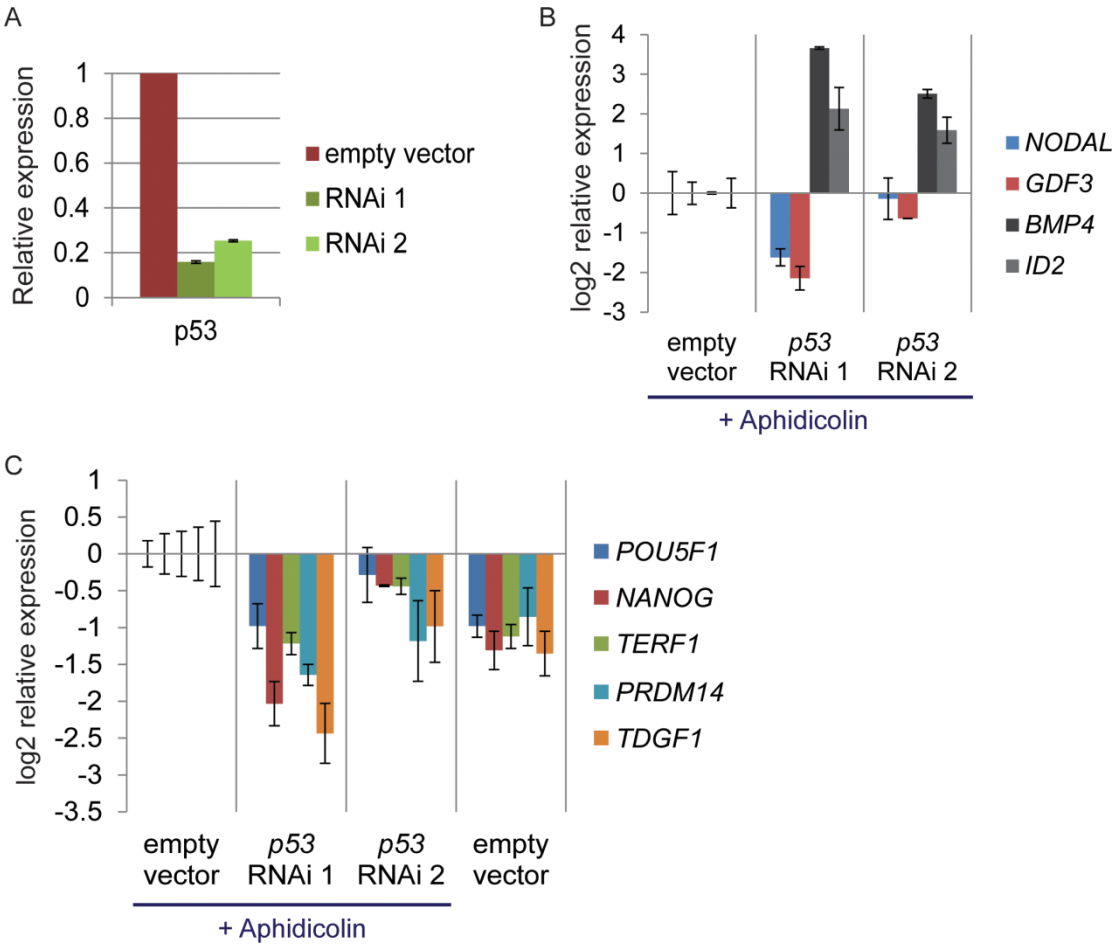



**Figure 45. Effect of *p53* knockdown on the exit of pluripotency.** (**A**) Transcript levels of *p53* upon RNAi knockdown as measured by qPCR. Both RNAi constructs effectively decreased expression of *p53*. (**B-C**) Expression levels of (**B**) TGFβ agonists (color) and antagonists (grayscale) and (**C**) pluripotency markers upon *p53* knockdown in the - bFGF, - TGFβ condition with Aphidicolin (DNA polymerase inhibitor). Experiments were done with the help of Dr Liang Hongqing.

To test whether p53 alone is sufficient to enhance the TGFβ signalling pathway in hESCs, we activated p53 in the absence of DNA replication stalling using the small molecule Nutlin-3, which prevents p53 degradation by inhibiting the interaction of p53 with MDM2 (Vassilev et al., 2004), similar to CHEK2's mechanism (Hirao et al., 2000). Nutlin-3 effectively increased p53 levels in hESCs (Figure 46A-B), which led to the alteration of TGFβ signalling pathway gene expression similar to that induced by DNA replication checkpoint activation (Figure

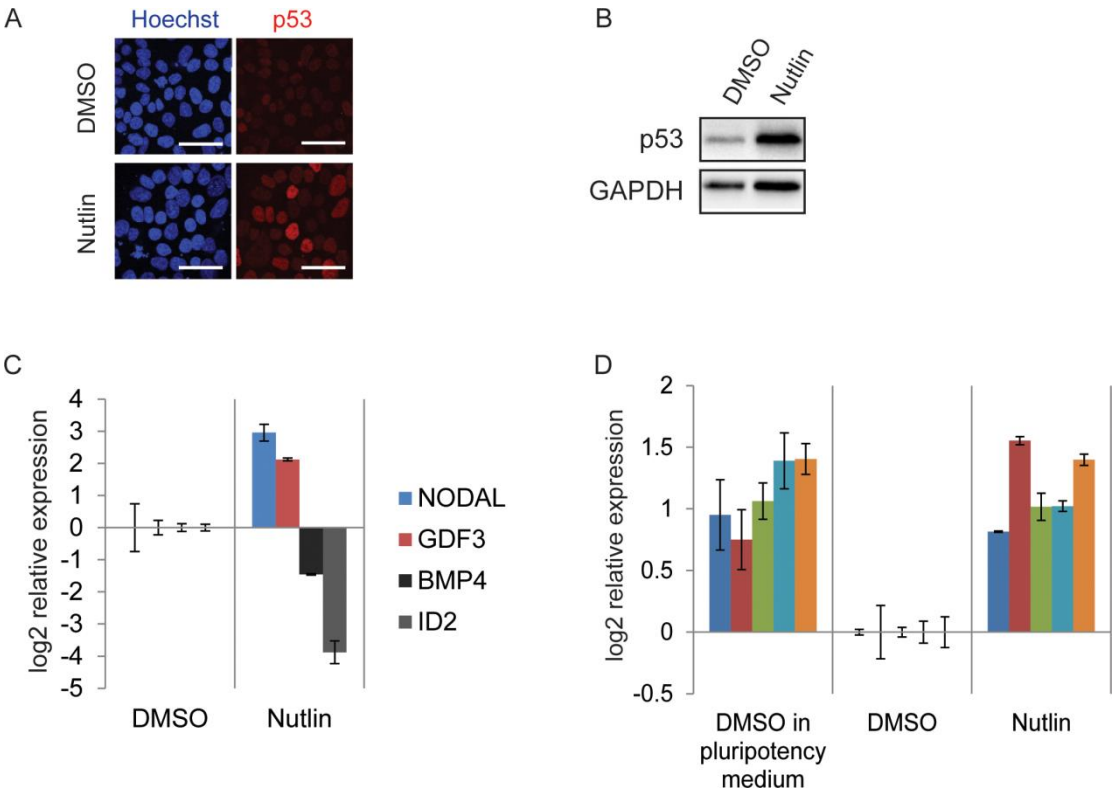



**Figure 46. Effect of Nutlin-3 treatment on the exit of pluripotency.** (**A-B**) Protein levels of p53 as measured by (**A**) immunofluorescence staining (Scale bar = 50μm) and (**B**) Western blot upon treatment of hESCs with Nutlin-3 (p53 stabilizer). (**C-D**) Expression levels of (**C**) TGFβ agonists (color) and antagonists (grayscale) and (**D**) pluripotency markers upon Nutlin-3 treatment. Experiments were done with the help of Dr Liang Hongqing.

46C). Accordingly, this resulted in delayed exit from pluripotency (Figure 46D), confirming that p53 alone can attenuate the exit from pluripotency by modulating the expression of TGFβ-related genes.

Altogether, our results prove that p53 enforces the prevention of the exit from pluripotency following DNA replication perturbation. This is surprising given that current literature assigns p53 as a factor opposing both the maintenance and acquisition of the pluripotent state (Hong et al., 2009; Jain et al., 2012; Kawamura et al., 2009; Li et al., 2012b; Lin et al., 2005; Qin et al., 2007), which seemingly contradicts our findings. However, it is important to note that as a hub transcription factor, p53 is known to regulate different gene sets (Akdemir et al., 2014) and mediate diverse cell fate transitions under different cellular contexts (Chang et al., 2011; Ubil et al., 2014). On top of that, different p53 activation dynamics under similar contexts can even trigger distinct outcomes (Aylon and Oren, 2007; Lee et al., 2010). Hitherto, p53 activity has not been studied in the context where both differentiation cues and DNA damage cues are present. Now we show that in this context, p53 switches to a role that upholds the human pluripotent state by enhancing TGFβ signalling, again emphasizing the importance of studying regulatory mechanisms within specific contexts.

We therefore propose a model of how perturbation of DNA replication during the S phase prevents the exit from pluripotency (Figure 47). Stalled DNA replication leads to the activation of the DNA replication checkpoint, initiated by ATM and ATR. These kinases phosphorylate CHEK2, which in turn, stabilizes p53 protein. p53 then binds to the chromatin and upregulates the transcription of TGFβ agonists while downregulating TGFβ antagonist expression. These changes in TGFβ-related gene

expression synergize to enhance pathway transduction. Finally, enhanced TGFβ signalling promotes NANOG expression and preserves the pluripotent state.

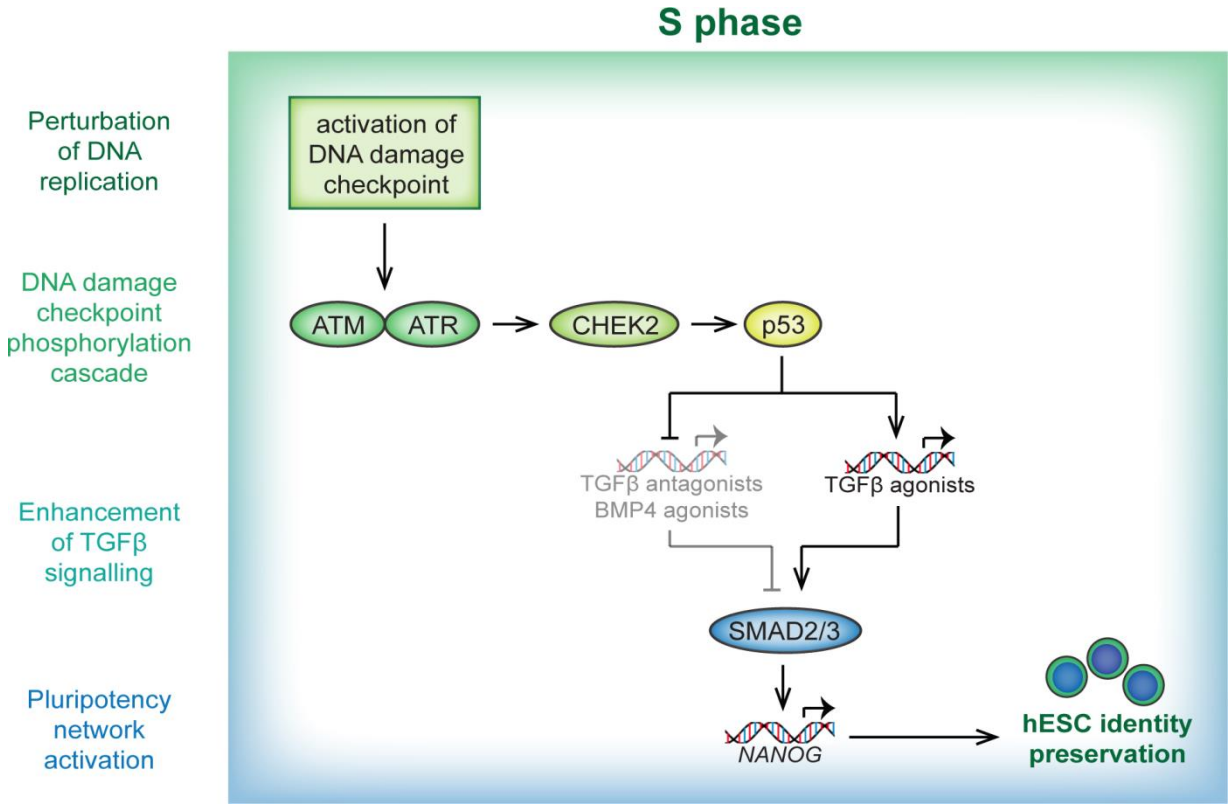



**Figure 47. The role of the DNA replication checkpoint in the exit from pluripotency.** Perturbation of DNA replication during the S phase of hESCs lead to the activation of the DNA replication checkpoint, initiated by the ATR and ATM sensor kinases. These phosphorylate and activate CHEK2, which in turn stabilizes p53. p53 then activates transcription of TGFβ agonists and repress TGFβ antagonists. These transcriptional changes work to enhance TGFβ signalling, SMAD2 phosphorylation and NANOG transcription, ultimately supporting the pluripotent state.

IX. Cyclin B1 upholds the pluripotent state during the G2 phase

While the mechanism above explains how cell cycle perturbations in the S phase prevents the exit from pluripotency, it cannot explain how arrest at the G2 phase provides the same effect. To find the mechanism underlying the G2 phase's role in the exit from pluripotency, we started by looking for factors or pathways that get upregulated during perturbation at the G2 phase of hESCs, starting with the cyclins. We observed that knockdown of G2-associated hits resulted in a prolonged G2 phase (Figure 30) concomitant with elevated expression of Cyclin B1, but not other cyclins (Figure 48). We therefore focused on whether Cyclin B1 plays a role in preserving the pluripotent state.

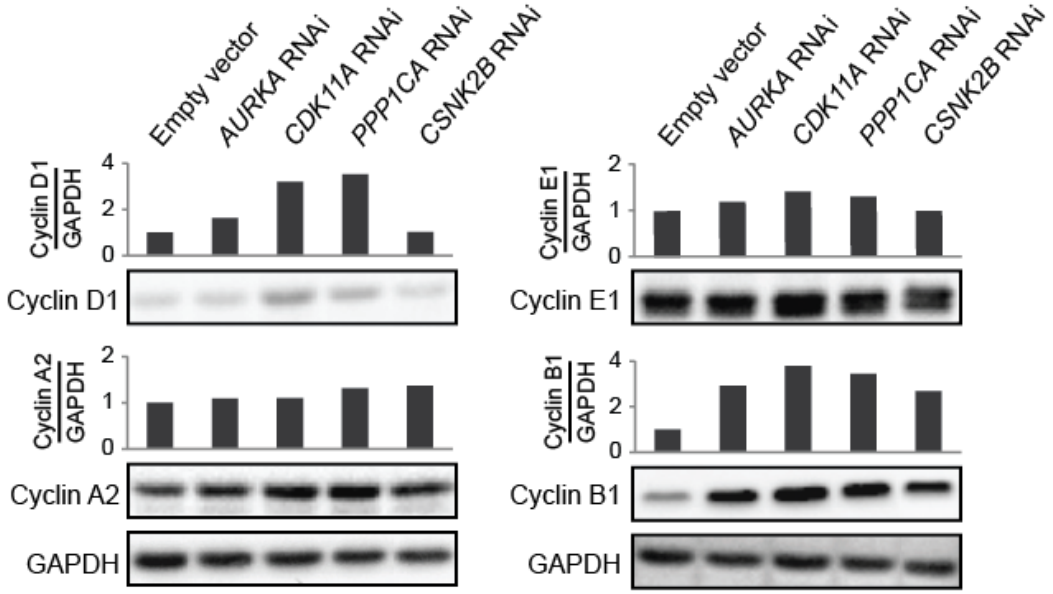



**Figure 48. Cyclin levels upon knockdown of G2-associated hits.** Western blot of Cyclins D1, A2, E1 and B1 upon knockdown of G2-associated hits in hESCs. Only Cyclin B1 shows a consistent upregulation amongst samples.

A. Exit from pluripotency by Cyclin B1 knockdown

To test if Cyclin B1 plays a role in promoting the pluripotent state, we ectopically modified expression levels of Cyclin B1 in hESCs. Knockdown of Cyclin B1 (Figure 49A) caused a dramatic downregulation of pluripotency marker expression prior to cell death (Figure 49B-C). Cyclin B1 knockdown also induced a prolongation of the G2 phase and a shortening of the G1 phase (Figure 49D-E), similar to other G2-specific factors (Figure 30). This is intriguing as knockdown of other G2-specific factors led to the preservation of pluripotency (Figure 20B), but Cyclin B1 knockdown caused the opposite. In addition, the concomitant differentiation and G1 shortening caused by Cyclin B1 knockdown demonstrates a decoupling of the expected cell cycle profile associated with differentiation. These deviations imply that the effects of Cyclin B1 on pluripotency are downstream of G2 phase arrest and indicate a tight linkage between Cyclin B1 and pluripotency.

B. Preservation of the pluripotent state by Cyclin B1 overexpression

To further validate whether Cyclin B1 is the underlying factor for the causative effect of a prolonged G2 phase in blocking exit from pluripotency, we overexpressed Cyclin B1 in hESCs (Figure 50A) and demonstrated for the first time that overexpression of a single cell cycle factor can attenuate the exit from pluripotency (Figure 50B-C). Moreover, this overexpression of Cyclin B1 did not cause significant enrichment of cells in the G2 phase (Figure 50D-E), indicating that a prolongation of the G2 phase is not necessary for Cyclin B1 to preserve the pluripotent state. Altogether, these results reveal that Cyclin B1 is the cardinal connecting node between the G2 phase and its ability to prevent the exit from pluripotency.

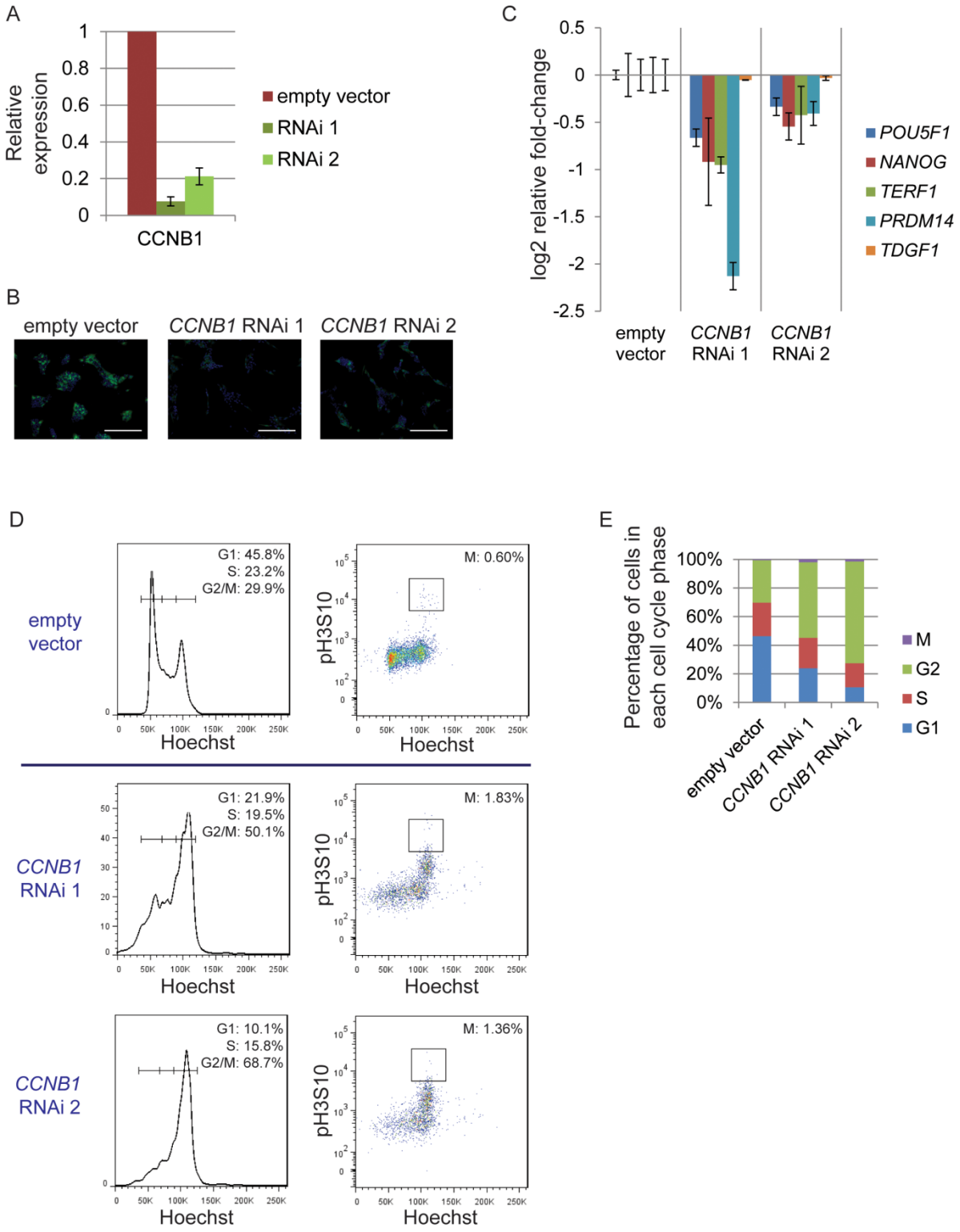



**Figure 49. Effect of Cyclin B1 knockdown on pluripotency.** (**A**) Transcript levels of *CCNB1* upon RNAi knockdown as measured by qPCR. Both RNAi constructs effectively decreased expression of p53. (**B**) Representative images for *NANOG*-GFP fluorescence (green) and nuclear staining (blue) after *CCNB1* knockdown in the pluripotency medium. Scale bar = 500μm. (**C**) Expression levels of pluripotency markers upon *CCNB1* knockdown in the pluripotency medium. (**D**) Flow cytometry profiles and (**E**) flow cytometry quantification indicating the cell cycle status of hESCs after *CCNB1* knockdown in the pluripotency medium.

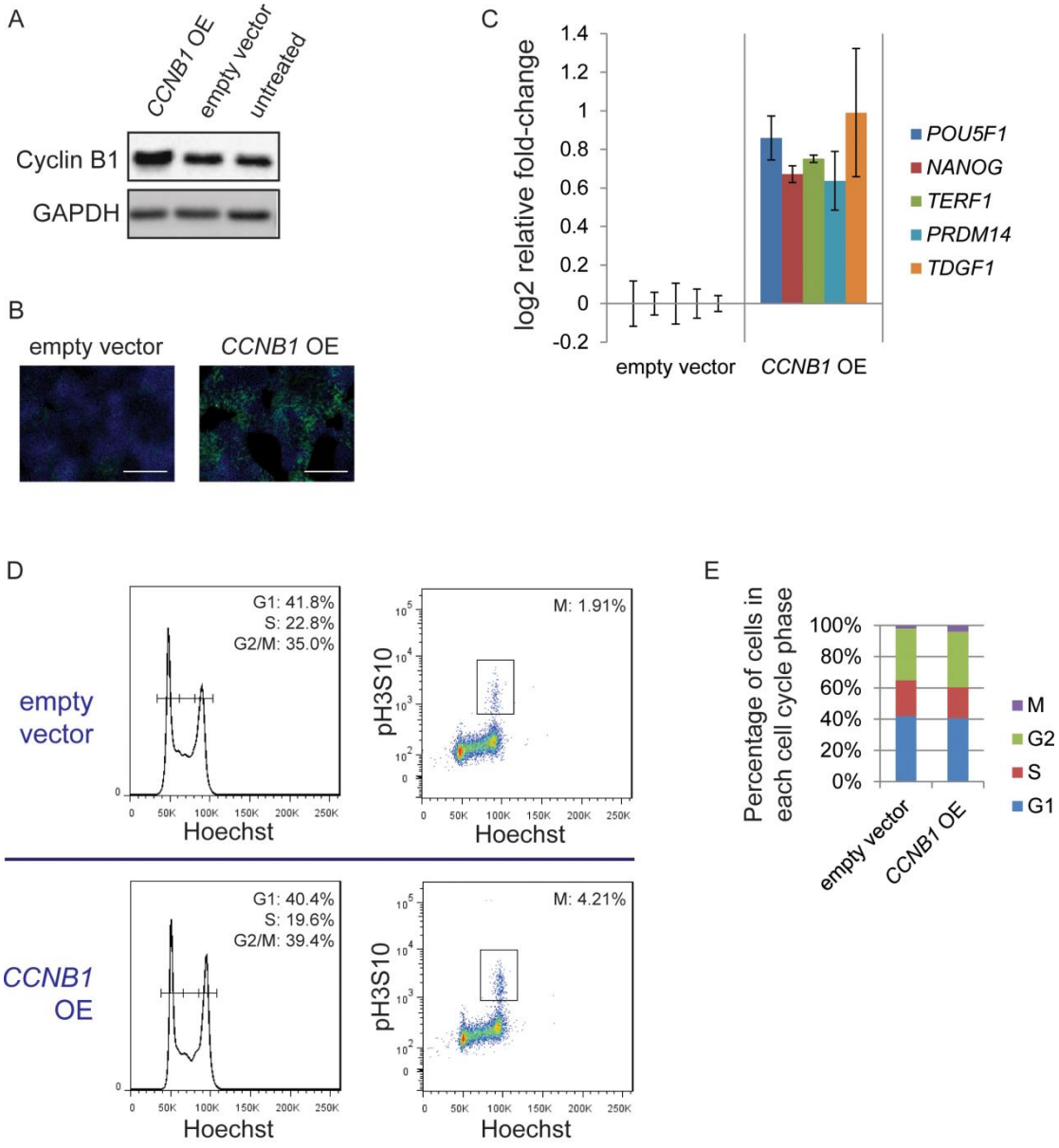



**Figure 50. Effect of Cyclin B1 overexpression on the exit from pluripotency.** (**A**) Protein levels of Cyclin B1 upon overexpression as measured by Western blot. (**B**) Representative images for *NANOG*-GFP fluorescence (green) and nuclear staining (blue) after CCNB1 overexpression in the - bFGF, - TGFβ condition. Scale bar = 500µm. (**C**) Expression levels of pluripotency markers upon CCNB1 overexpression in the - bFGF, - TGFβ condition. (**D**) Flow cytometry profiles and (**E**) flow cytometry quantification indicating the cell cycle status of hESCs after CCNB1 overexpression in the pluripotency medium. Experiments were done with the help of Dr Liang Hongqing.

C. Possible mechanisms behind the role of Cyclin B1 in the exit from pluripotency

To investigate how Cyclin B1 promotes the pluripotent state, we performed a time-course microarray analysis of hESCs overexpressing Cyclin B1. Interestingly, we also found an upregulated expression of TGFβ agonists, concomitant with the preservation of pluripotency marker expression (Figure 51A-B). Therefore, Cyclin B1 might also work through the TGFβ pathway to prevent the exit from pluripotency. Notably, Cyclin B1 overexpression also resulted in decreased expression of genes involved in specification of neuroectodermic lineages and an opposite trend for mesodermic lineages (Figure 51C), suggesting that Cyclin B1 might supplement its role in the exit from pluripotency by keeping the balance in expression of lineage-specifying genes (Loh and Lim, 2011).

Finally, we checked if Cyclin B1 activity is dependent on its partner kinase CDK1. Surprisingly, when CDK1 is knocked down while self-renewal signals were withdrawn, the exit from pluripotency is prevented, similar to other G2-associated factors (Figure 52A-C). Inhibition of CDK1 with RO3306 had the same effect (Figure 52A,D). This indicates that while CDK1's effect on the exit from pluripotency could be dependent on Cyclin B1, Cyclin B1's functions on pluripotency are downstream of and autonomous from CDK1. This is very intriguing as Cyclin B1 is not yet known to have CDK-independent functions. Interestingly, the G1 phase-associated Cyclin D1 has been demonstrated in mice to have a role in regulating transcription independent of its CDK partners (Bienvenu et al., 2010). It is possible that Cyclin B1 has adopted similar CDK-independent functions to influence the exit from pluripotency in hESCs.

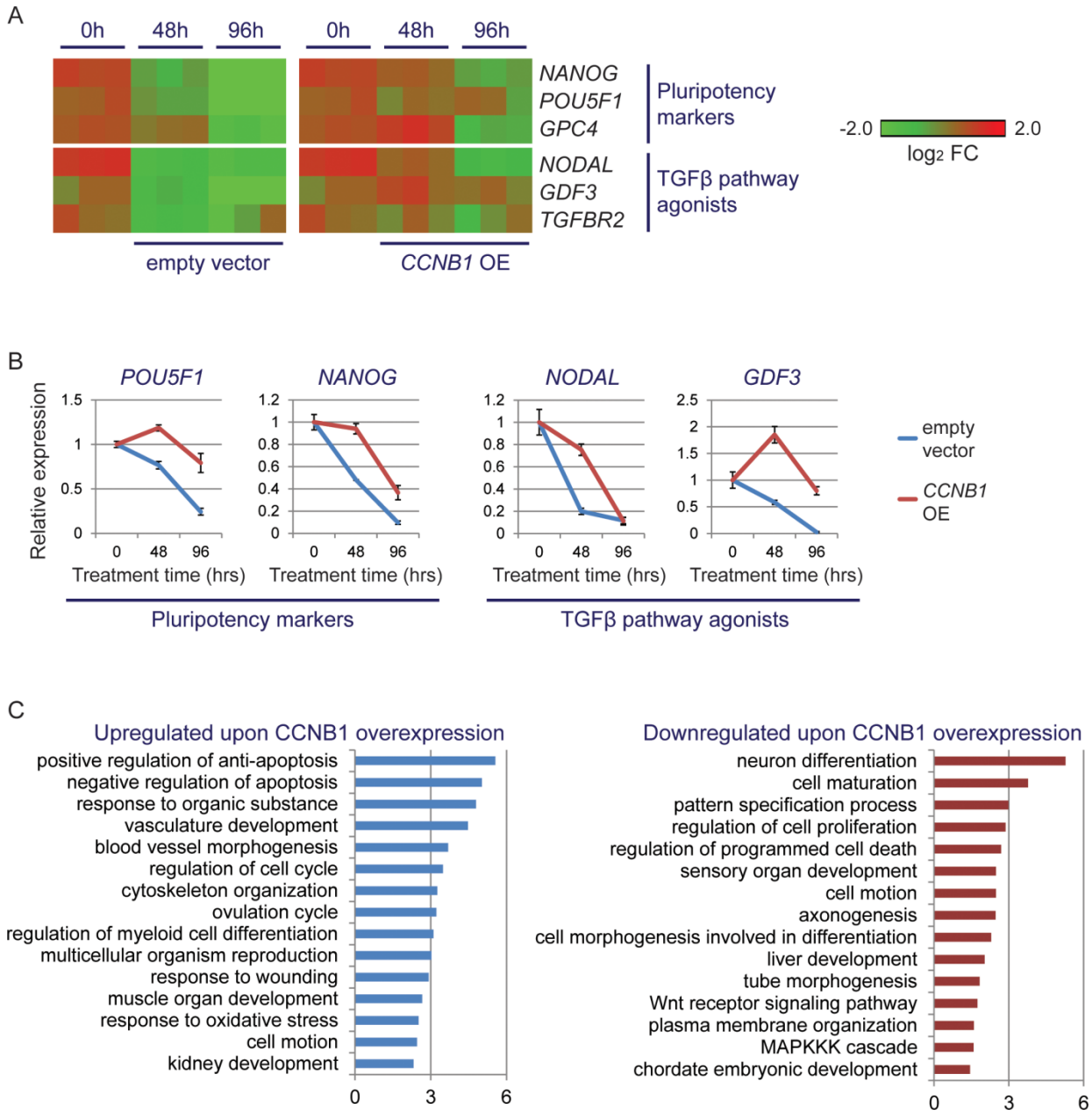



**Figure 51. Time-course microarray analysis upon Cyclin B1 overexpression in hESCs.** (**A**) Microarray heatmap for differentially expressed genes upon overexpression of Cyclin B1 in the - bFGF, -TGFβ condition. (**B**) Time-course quantitative PCR for pluripotency and TGFβ-related genes upon overexpression of Cyclin B1 in the - bFGF, - TGFβ condition. Error bars denote standard deviations of triplicate data. (**C**) Gene ontology analysis. Enriched biological process terms ($p<0.05$) for upregulated and downregulated genes identified in the microarray analysis upon Cyclin B1 overexpression in the - bFGF, - TGFβ condition. Microarray experiments were done with the help of Tan Zi-Ying.

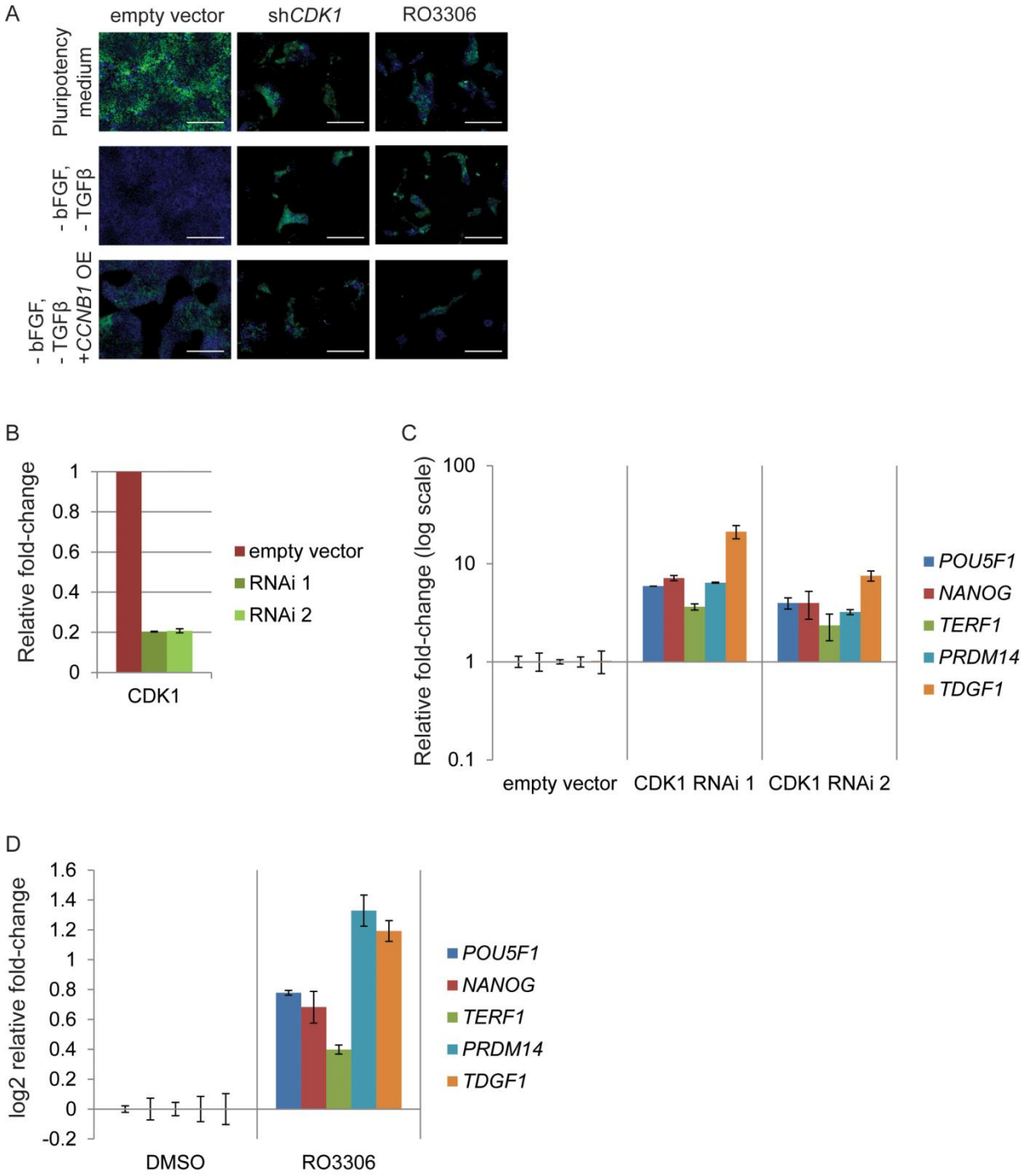



**Figure 52. Effect of CDK1 inhibition on the exit from pluripotency.** (**A**) Representative images for *NANOG*-GFP fluorescence (green) and nuclear staining (blue) after CDK1 inhibition. Scale bar = 500μm. (**B**) Transcript levels of *CDK1* upon RNAi knockdown as measured by qPCR. Both RNAi constructs effectively decreased expression of *CDK1*. (**C**-**D**) Expression levels of pluripotency markers of hESCs in the - bFGF, - TGFβ condition upon (**C**) *CDK1* knockdown or (**D**) treatment with RO3306 (CDK1 inhibitor).

D. Role of other cyclins in the exit from pluripotency

We finally examined whether these effects are exclusive to the G2/M-specific Cyclin B1 or if it can be conferred by other cyclins as well. Modulation of expression levels of Cyclins A2 and D1 neither induced nor prevented the exit from pluripotency (Figure 53,54). Surprisingly, knockdown of Cyclin E1 caused downregulation of pluripotency markers, and its overexpression also prevented the exit from

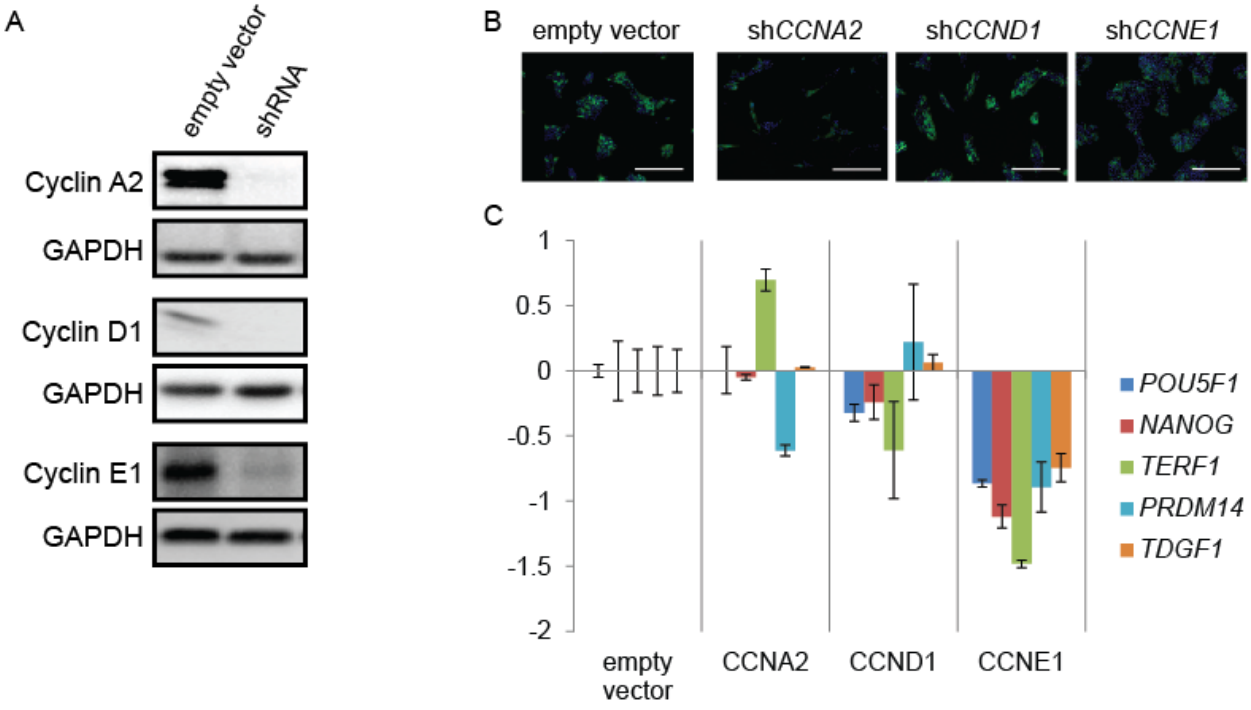



**Figure 53. Effect of knockdown of other cyclins on pluripotency.** (**A**) Protein levels of Cyclins A2, D1 and E1 upon RNAi knockdown as measured by Western blot. (**B**) Representative images for *NANOG*-GFP fluorescence (green) and nuclear staining (blue) after cyclin knockdown in the pluripotency medium. Scale bar = 500μm. (**C**) Expression levels of pluripotency markers upon cyclin knockdown in the pluripotency medium.

pluripotency (Figure 53,54). Thus, Cyclin E1 has the same effects on the exit from pluripotency to that of Cyclin B1. As Cyclin E1 is expressed at the G1/S interface, this factor could provide an additional route from cell cycle stalling at the S phase could act to preserve the pluripotent state.

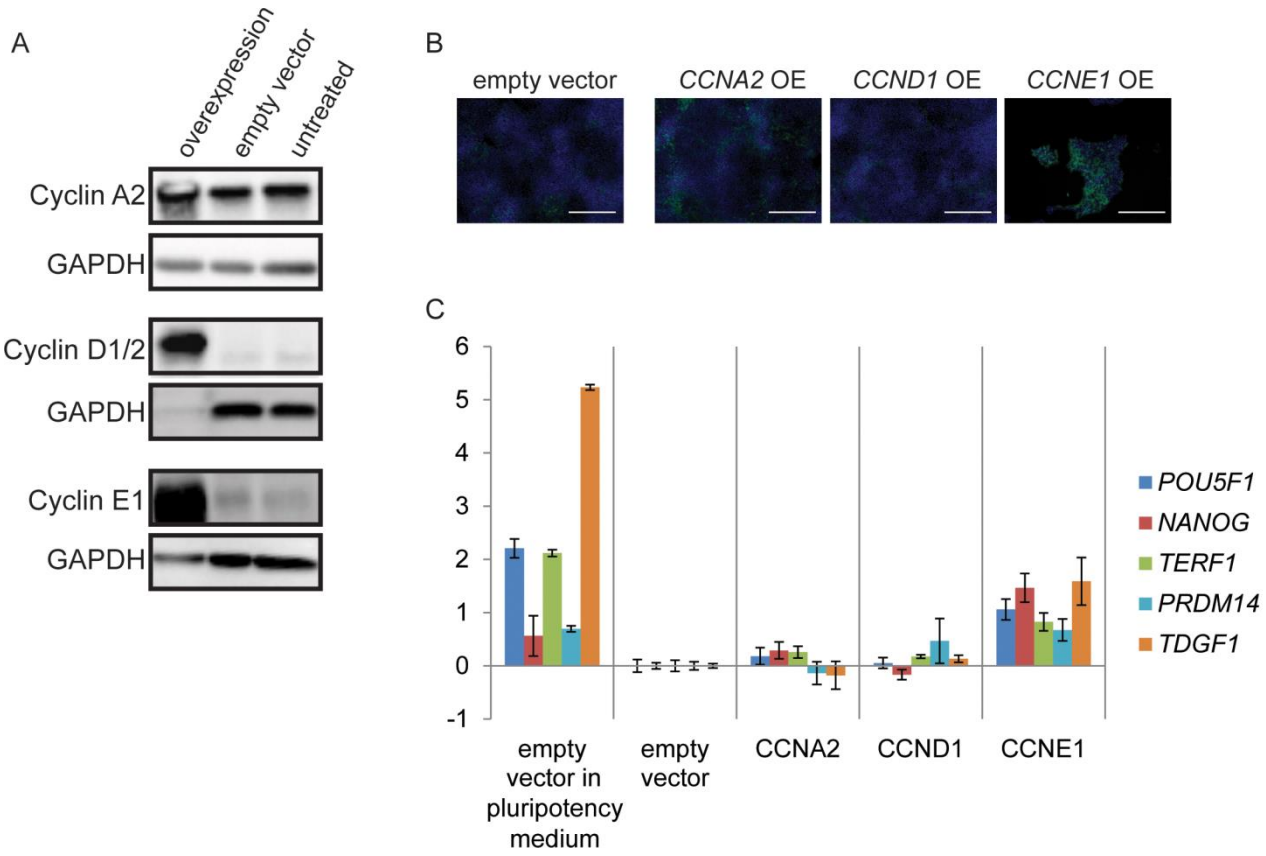



**Figure 54. Effect of overexpression of other cyclins on the exit from pluripotency.** (**A**) Protein levels of Cyclins A2, D1 and E1 upon overexpression as measured by Western blot. (**B**) Representative images for *NANOG*-GFP fluorescence (green) and nuclear staining (blue) after cyclin overexpression in the - bFGF, - TGFβ condition. Scale bar = 500μm. (**C**) Expression levels of pluripotency markers upon cyclin overexpression in the - bFGF, - TGFβ condition. Experiments were done with the help of Dr Liang Hongqing.

In a broader perspective, these results demonstrate that cyclins not only act as the major regulators of cell cycle progression (Evans et al., 1983; Pines, 1991), but also influence other biological processes unrelated to the cell cycle. In fact, cyclins have been implicated in various processes including transcription, DNA damage response, epigenetic regulation and lipogenesis (Lim and Kaldis, 2013). A role for cyclins in other cellular functions inevitably couples these processes to the cell cycle, contributing to the distinct cell states that are represented by the different cell cycle phases. While the results here only refer to the regulation of pluripotency exit, it is likely that similar mechanisms are in place for other biological processes, calling for future studies to extensively characterize the non-classical roles of cyclins, which may even be CDK-independent, as we have demonstrated for Cyclin B1.

Altogether, these results indicate that Cyclin B1 is the key mediator for upholding the pluripotent state during the G2 phase. Although further experiments need to be done to conclusively understand its mechanism, we performed preliminary analyses hinting that Cyclin B1 could partially work through the TGFβ pathway, and by managing the balance in lineage-specifying gene expression (Figure 55).

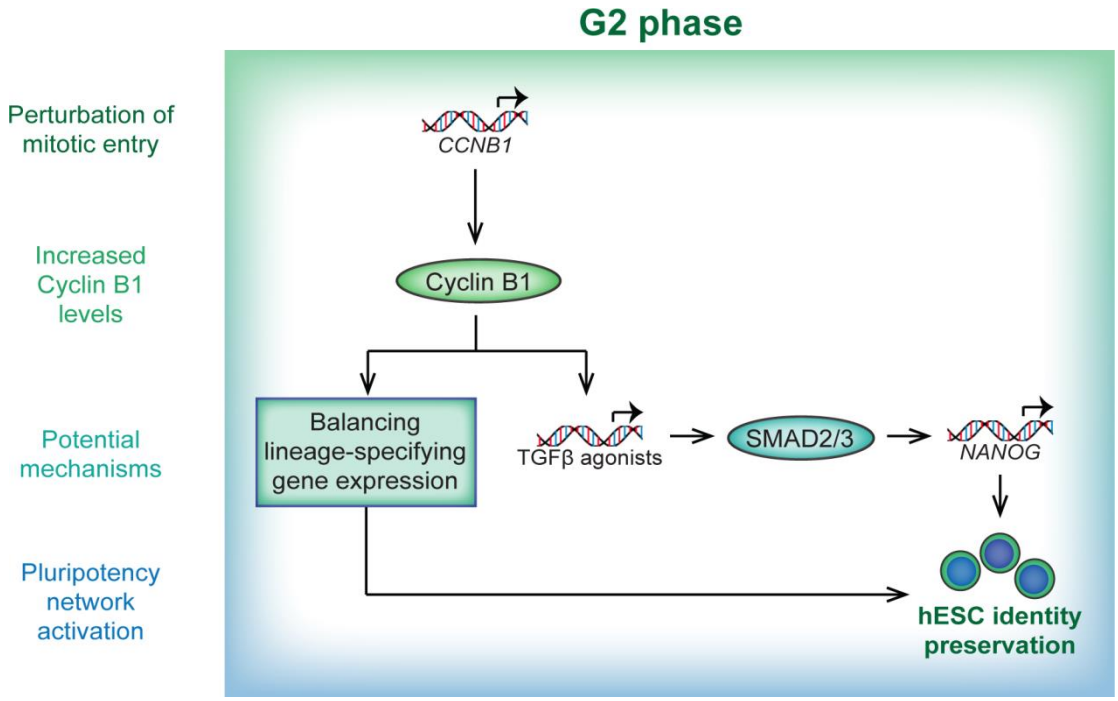



**Figure 55. The role of Cyclin B1 in the exit from pluripotency.** Arrest of cell cycle progression at the G2 phase leads to the accumulation of Cyclin B1. Heightened Cyclin B1 levels induce transcriptional changes that enhance TGFβ signalling and maintain the balance of lineage-specifying genes, both of which ultimately support the pluripotent state.

## Conclusions

The exit from pluripotency necessitates the collapse of the pluripotency network. Yet, how differentiation cues lead to the breakdown of this network is ill-defined, especially in hESCs. We performed a systematic large-scale RNAi screen under multiple differentiation conditions to unveil genes required for this process. Our systematic functional screens of hESCs under multiple differentiation conditions enabled the identification of both universal and specific regulatory axes for the exit from pluripotency (Figure 56). Our results thus provide insight on how some regulatory modules have conserved functions regardless of other cues, while others perform in a very context-specific manner. Members of chromatin-modifying complexes such as HAT complexes and the SWI/SNF complex were highly enriched, emphasizing a universal need to restructure the chromatin to enable complete cell fate transition during PSD. Conversely, other PSD regulators like development-related signalling pathways and RNA splicing tend to function in a more context-dependent manner. Altogether, the results of the large-scale RNAi screen presented in this thesis delivers a unique resource for dissecting the mechanisms behind PSD by providing new insight on the roles of various factors and pathways on PSD and opens the avenue for future work on unravelling the complete regulatory network of this process.

Moreover, the data presented in this thesis address the unanswered question of whether factors in the core cell cycle machinery enforce a deterministic regulation on pluripotency and differentiation in ESCs. We found that cell cycle pathways

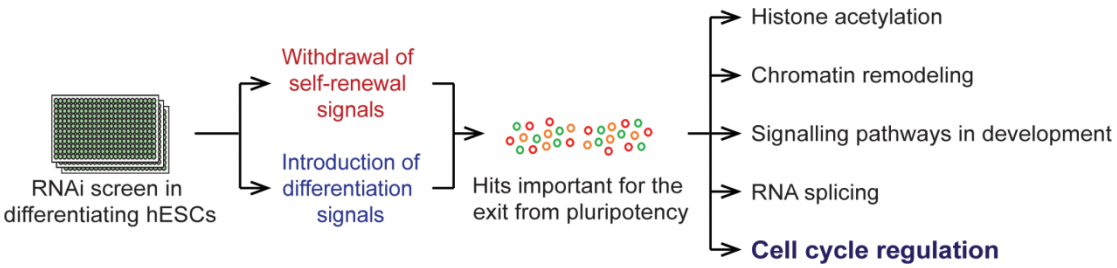



**Figure 56. Summary of high-throughput RNAi screen.** High-throughput RNAi screening in multiple differentiation conditions enabled the identification of factors and processes that are important for the exit from pluripotency.

stimulated during S and G2 phase perturbation play an active role in promoting the pluripotent state and preventing its exit (Figure 57). Specifically, we find the ATM/ATR-CHEK2-p53 axis in the S phase and Cyclin B1 in the G2 phase functioning to uphold the hESC state. Such a direct linkage from S- and G2-specific pathways to the exit from pluripotency has not been reported before, thus introducing a new paradigm for the coupling of cell cycle and pluripotency. In contrast, we did not find G1-specific factors implicated in differentiation to have an effect on the exit from pluripotency in this context. Instead, our results suggest that the G1 phase is amenable to the exit from pluripotency because of the absence of such pathways that actively preserve hESC identity. Finally, as the ATR/ATM-mediated checkpoint can be activated by endogenous damage generated during DNA replication and Cyclin B1 is more abundant during the G2 phase of the unperturbed cell cycle, these pathways and their effects in upholding pluripotency can be applied to normal progression through the S and G2 phases.

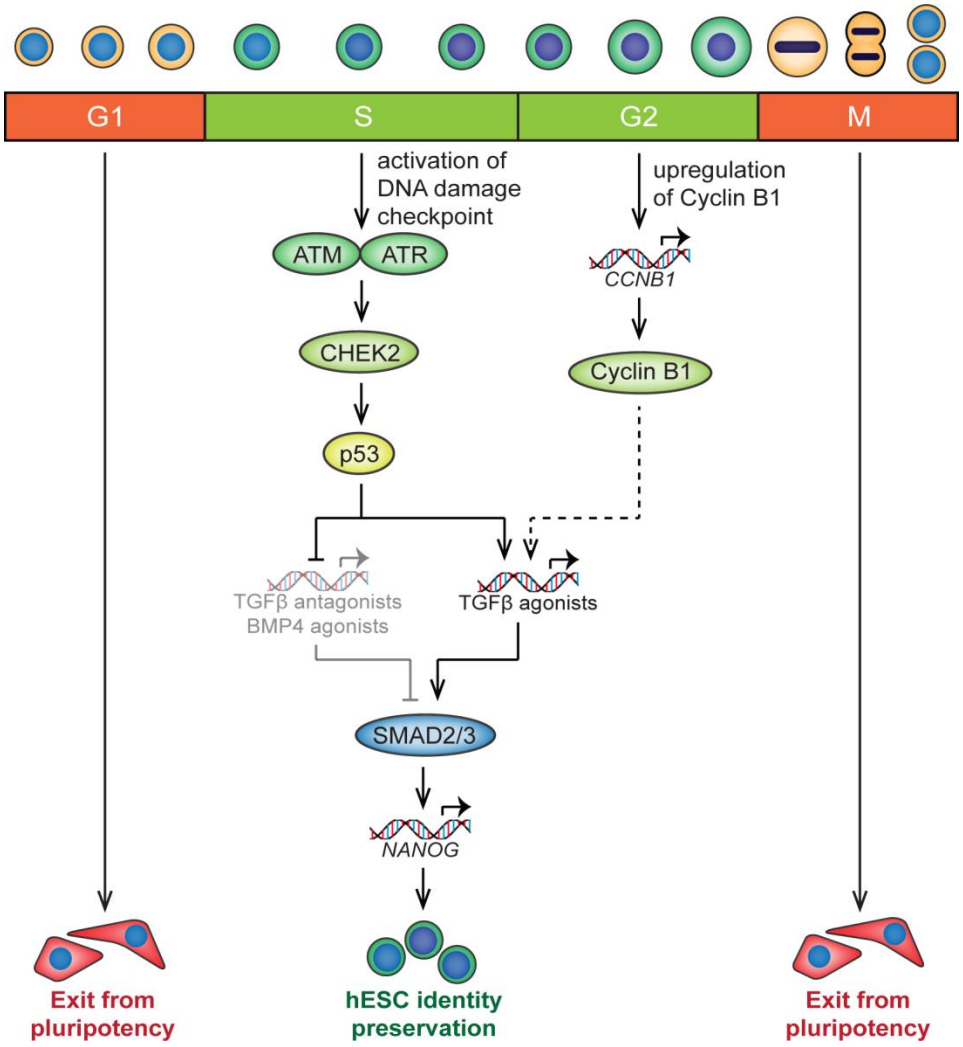



**Figure 57. Outline of mechanisms behind the prevention of the exit from pluripotency in the S and G2 phases.** hESCs resist the exit from pluripotency during the S and G2 phases due to the presence of active mechanisms that uphold the pluripotent state, namely the DNA replication checkpoint and Cyclin B1, respectively.

ESCs display a unique cell cycle profile with a shortened G1 phase (Coronado et al., 2013; Hindley and Philpott, 2013; Singh and Dalton, 2009; Stead et al., 2002). This was thought to maintain pluripotency by limiting time spent in the G1 phase, which is the only phase permissive to differentiation (Mummery et al., 1987; Pauklin and Vallier, 2013; Sela et al., 2012). However, there is little evidence to unambiguously declare that the cell cycle phases can dominantly and deterministically

regulate the pluripotency network. In this thesis, we show that the S and G2 phases employ specific pathways to actively boost the pluripotent state, rather than just passively retain the pluripotent state in the absence of the G1 phase's effects. In fact, our results imply that it is the absence of such pathways in the G1 phase that underlies its responsiveness to differentiation cues. The need to actively boost the pluripotent state may be critical given the fact that the inclination of the G1 phase towards differentiation may lead to a transient tendency for pluripotency decay every time cells pass through it. Therefore, active resumption of the pluripotency network in the following S and G2 phases may be critical to tilt the cell fate balance back to pluripotent state maintenance. Thus, the results in this thesis can be culminated in a proposed model wherein the S/G2 and M/G1 phases shift the weights between pluripotency maintenance and exit across the hES cell cycle (Figure 58). This model suggests that a balance between cell cycle phases is critical for ESC fate determination, in contrast to the previously recognized G1-centric model.

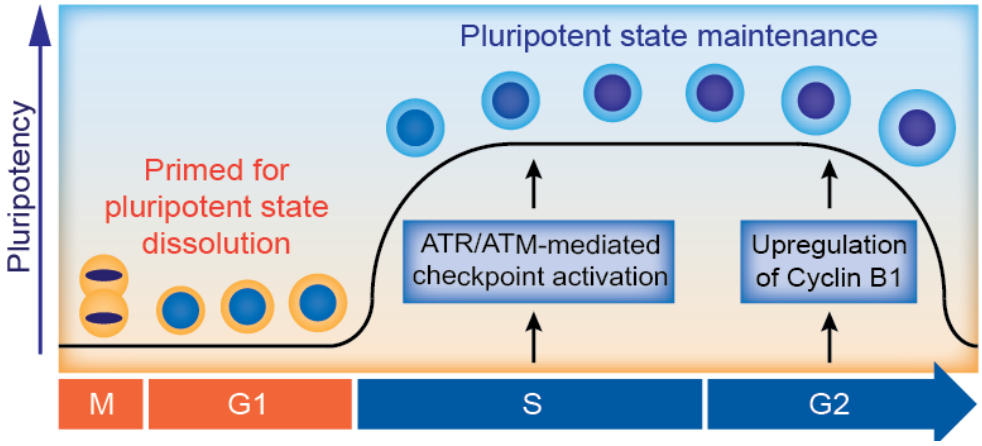



**Figure 58. Model of how pluripotency maintenance and exit is regulated by a balance between the different cell cycle phases.** Active pathways in the S and G2 phases promote the pluripotent state, enforcing maintenance of hESC identity. Conversely, hESCs are primed for the exit from pluripotency at the G1 and M phases due to the absence of such pathways. Cell cycle progression alternates between these two states, enabling both self-renewal and ability to differentiate.

While these findings provide answers to the fundamental question of whether the cell cycle machinery directly regulates the exit from pluripotency in hESCs, they also open up new questions for future research. First, it is imperative to find out whether the observed cell cycle-dependent restriction of cell fate transition in hESCs is also applicable to other undifferentiated cell types such as adult stem cells and cancer cells, and if so, whether the underlying mechanisms are conserved or not. Second, as our results prove that cell cycle manipulations in hESCs can influence cell fate transitions, this knowledge can be applied to improve the efficiency of directed differentiation protocols. Finally, given the complexity of cell cycle regulation, it is almost certain that there are additional mechanisms in place for the effects of the cell cycle on the exit from pluripotency, besides the ones shown in this thesis. Therefore, this pilot study jumpstarts the search for more connective nodes between the cell cycle machinery and pluripotency.

In conclusion, this thesis summarizes the results of our studies, which contribute to the understanding of the exit from pluripotency in hESCs, and how this process is hard-wired to the cell cycle machinery.

## Bibliography


Abujarour, R., Efe, J., and Ding, S. (2010). Genome-wide gain-of-function screen identifies novel regulators of pluripotency. Stem Cells *28*, 1487-1497.

Akdemir, K.C., Jain, A.K., Allton, K., Aronow, B., Xu, X., Cooney, A.J., Li, W., and Barton, M.C. (2014). Genome-wide profiling reveals stimulus-specific functions of p53 during differentiation and DNA damage of human embryonic stem cells. Nucleic Acids Res *42*, 205-223.

Ashton-Beaucage, D., Udell, C.M., Gendron, P., Sahmi, M., Lefrancois, M., Baril, C., Guenier, A.S., Duchaine, J., Lamarre, D., Lemieux, S., *et al.* (2014). A functional screen reveals an extensive layer of transcriptional and splicing control underlying RAS/MAPK signaling in Drosophila. PLoS Biol *12*, e1001809.

Assou, S., Le Carrour, T., Tondeur, S., Strom, S., Gabelle, A., Marty, S., Nadal, L., Pantesco, V., Reme, T., Hugnot, J.P., *et al.* (2007). A meta-analysis of human embryonic stem cells transcriptome integrated into a web-based expression atlas. Stem Cells *25*, 961-973.

Aylon, Y., and Oren, M. (2007). Living with p53, dying of p53. Cell *130*, 597-600.

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M., *et al.* (2006). Chromatin signatures of pluripotent cell lines. Nat Cell Biol *8*, 532-538.

Bakkenist, C.J., and Kastan, M.B. (2003). DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. Nature *421*, 499-506.

Banin, S., Moyal, L., Shieh, S., Taya, Y., Anderson, C.W., Chessa, L., Smorodinsky, N.I., Prives, C., Reiss, Y., Shiloh, Y., *et al.* (1998). Enhanced phosphorylation of p53 by ATM in response to DNA damage. Science *281*, 1674-1677.

Bartek, J., Lukas, C., and Lukas, J. (2004). Checking on DNA damage in S phase. Nat Rev Mol Cell Biol *5*, 792-804.

Bastien, J., Plassat, J.L., Payrastre, B., and Rochette-Egly, C. (2006). The phosphoinositide 3-kinase/Akt pathway is essential for the retinoic acid-induced differentiation of F9 cells. Oncogene *25*, 2040-2047.

Beattie, G.M., Lopez, A.D., Bucay, N., Hinton, A., Firpo, M.T., King, C.C., and Hayek, A. (2005). Activin A maintains pluripotency of human embryonic stem cells in the absence of feeder layers. Stem Cells *23*, 489-495.

Becker, K.A., Ghule, P.N., Therrien, J.A., Lian, J.B., Stein, J.L., van Wijnen, A.J., and Stein, G.S. (2006). Self-renewal of human embryonic stem cells is supported by a shortened G1 cell cycle phase. J Cell Physiol *209*, 883-893.

Ben-David, U., and Benvenisty, N. (2011). The tumorigenicity of human embryonic and induced pluripotent stem cells. Nat Rev Cancer *11*, 268-277.

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

Bernstein, B.E., Humphrey, E.L., Erlich, R.L., Schneider, R., Bouman, P., Liu, J.S., Kouzarides, T., and Schreiber, S.L. (2002). Methylation of histone H3 Lys 4 in coding regions of active genes. Proc Natl Acad Sci U S A *99*, 8695-8700.

Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature *409*, 363-366.

Betschinger, J., Nichols, J., Dietmann, S., Corrin, P.D., Paddison, P.J., and Smith, A. (2013). Exit from Pluripotency Is Gated by Intracellular Redistribution of the bHLH Transcription Factor Tfe3. Cell *153*, 335-347.

Bienvenu, F., Jirawatnotai, S., Elias, J.E., Meyer, C.A., Mizeracka, K., Marson, A., Frampton, G.M., Cole, M.F., Odom, D.T., Odajima, J., *et al.* (2010). Transcriptional role of cyclin D1 in development revealed by a genetic-proteomic screen. Nature *463*, 374-378.

Blagosklonny, M.V. (2007). Mitotic arrest and cell fate: why and how mitotic inhibition of transcription drives mutually exclusive events. Cell Cycle *6*, 70-74.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., *et al.* (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947-956.

Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K., *et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature *441*, 349-353.

Brivanlou, A.H., Gage, F.H., Jaenisch, R., Jessell, T., Melton, D., and Rossant, J. (2003). Stem cells. Setting standards for human embryonic stem cells. Science *300*, 913-916.

Budirahardja, Y., and Gonczy, P. (2009). Coupling the cell cycle to development. Development *136*, 2861-2872.

Cai, L., Ye, Z., Zhou, B.Y., Mali, P., Zhou, C., and Cheng, L. (2007). Promoting human embryonic stem cell renewal or differentiation by modulating Wnt signal and culture conditions. Cell Res *17*, 62-72.

Calder, A., Roth-Albin, I., Bhatia, S., Pilquil, C., Lee, J.H., Bhatia, M., Levadoux-Martin, M., McNicol, J., Russell, J., Collins, T., *et al.* (2013). Lengthened g1 phase indicates differentiation status in human embryonic stem cells. Stem Cells Dev *22*, 279-295.

Caplen, N.J., Fleenor, J., Fire, A., and Morgan, R.A. (2000). dsRNA-mediated gene silencing in cultured Drosophila cells: a tissue culture model for the analysis of RNA interference. Gene *252*, 95-105.

Card, D.A., Hebbar, P.B., Li, L., Trotter, K.W., Komatsu, Y., Mishina, Y., and Archer, T.K. (2008). Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells. Mol Cell Biol *28*, 6426-6438.

Carpenter, A.E., and Sabatini, D.M. (2004). Systematic genome-wide screens of gene function. Nat Rev Genet *5*, 11-22.

Chang, C.J., Chao, C.H., Xia, W., Yang, J.Y., Xiong, Y., Li, C.W., Yu, W.H., Rehman, S.K., Hsu, J.L., Lee, H.H., *et al.* (2011). p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs. Nat Cell Biol *13*, 317-323.

Chen, G., Gulbranson, D.R., Hou, Z., Bolin, J.M., Ruotti, V., Probasco, M.D., Smuga-Otto, K., Howden, S.E., Diol, N.R., Propson, N.E., *et al.* (2011). Chemically defined conditions for human iPSC derivation and culture. Nat Methods *8*, 424-429.

Chen, P.B., Hung, J.H., Hickman, T.L., Coles, A.H., Carey, J.F., Weng, Z., Chu, F., and Fazzio, T.G. (2013). Hdac6 regulates Tip60-p400 function in stem cells. Elife (Cambridge) *2*, e01557.

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

Cho, R.J., Campbell, M.J., Winzeler, E.A., Steinmetz, L., Conway, A., Wodicka, L., Wolfsberg, T.G., Gabrielian, A.E., Landsman, D., Lockhart, D.J., *et al.* (1998). A genome-wide transcriptional analysis of the mitotic cell cycle. Mol Cell *2*, 65-73.

Clemens, J.C., Worby, C.A., Simonson-Leff, N., Muda, M., Maehama, T., Hemmings, B.A., and Dixon, J.E. (2000). Use of double-stranded RNA interference in Drosophila cell lines to dissect signal transduction pathways. Proc Natl Acad Sci U S A *97*, 6499-6503.

Cole, M.F., Johnstone, S.E., Newman, J.J., Kagey, M.H., and Young, R.A. (2008). Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. Genes Dev *22*, 746-755.

Coronado, D., Godet, M., Bourillot, P.Y., Tapponnier, Y., Bernat, A., Petit, M., Afanassieff, M., Markossian, S., Malashicheva, A., Iacone, R., *et al.* (2013). A short G1 phase is an intrinsic determinant of naive embryonic stem cell pluripotency. Stem Cell Res *10*, 118-131.

Cowan, C.A., Klimanskaya, I., McMahon, J., Atienza, J., Witmyer, J., Zucker, J.P., Wang, S., Morton, C.C., McMahon, A.P., Powers, D., *et al.* (2004). Derivation of embryonic stem-cell lines from human blastocysts. N Engl J Med *350*, 1353-1356.

Davidson, K.C., Adams, A.M., Goodson, J.M., McDonald, C.E., Potter, J.C., Berndt, J.D., Biechele, T.L., Taylor, R.J., and Moon, R.T. (2012). Wnt/beta-catenin signaling

promotes differentiation, not self-renewal, of human embryonic stem cells and is repressed by Oct4. Proc Natl Acad Sci U S A *109*, 4485-4490.

de la Serna, I.L., Ohkawa, Y., and Imbalzano, A.N. (2006). Chromatin remodelling in mammalian differentiation: lessons from ATP-dependent remodellers. Nat Rev Genet *7*, 461-473.

Ding, L., Paszkowski-Rogacz, M., Nitzsche, A., Slabicki, M.M., Heninger, A.K., de Vries, I., Kittler, R., Junqueira, M., Shevchenko, A., Schulz, H., *et al.* (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*, 403-415.

Ding, Q., Xia, W., Liu, J.C., Yang, J.Y., Lee, D.F., Xia, J., Bartholomeusz, G., Li, Y., Pan, Y., Li, Z., *et al.* (2005). Erk associates with and primes GSK-3beta for its inactivation resulting in upregulation of beta-catenin. Mol Cell *19*, 159-170.

Dravid, G., Ye, Z., Hammond, H., Chen, G., Pyle, A., Donovan, P., Yu, X., and Cheng, L. (2005). Defining the role of Wnt/beta-catenin signaling in the survival, proliferation, and self-renewal of human embryonic stem cells. Stem Cells *23*, 1489-1501.

Drukker, M. (2008). Immunological considerations for cell therapy using human embryonic stem cell derivatives, 2010 edn (Cambridge (MA), Harvard Stem Cell Institute).

Dumitru, R., Gama, V., Fagan, B.M., Bower, J.J., Swahari, V., Pevny, L.H., and Deshmukh, M. (2012). Human embryonic stem cells have constitutively active Bax at the Golgi and are primed to undergo rapid apoptosis. Mol Cell *46*, 573-583.

Echeverri, C.J., and Perrimon, N. (2006). High-throughput RNAi screening in cultured cells: a user's guide. Nat Rev Genet *7*, 373-384.

Edgar, B.A., and Lehner, C.F. (1996). Developmental control of cell cycle regulators: a fly's perspective. Science *274*, 1646-1652.

Egli, D., Birkhoff, G., and Eggan, K. (2008). Mediators of reprogramming: transcription factors and transitions through mitosis. Nat Rev Mol Cell Biol *9*, 505-516.

Egozi, D., Shapira, M., Paor, G., Ben-Izhak, O., Skorecki, K., and Hershko, D.D. (2007). Regulation of the cell cycle inhibitor p27 and its ubiquitin ligase Skp2 in differentiation of human embryonic stem cells. FASEB J *21*, 2807-2817.

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

Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001a). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature *411*, 494-498.

Elbashir, S.M., Lendeckel, W., and Tuschl, T. (2001b). RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes Dev *15*, 188-200.

Evans, T., Rosenthal, E.T., Youngblom, J., Distel, D., and Hunt, T. (1983). Cyclin: a protein specified by maternal mRNA in sea urchin eggs that is destroyed at each cleavage division. Cell *33*, 389-396.

Fazzio, T.G., Huff, J.T., and Panning, B. (2008). An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. Cell *134*, 162-174.

Filipczyk, A.A., Laslett, A.L., Mummery, C., and Pera, M.F. (2007). Differentiation is coupled to changes in the cell cycle regulatory apparatus of human embryonic stem cells. Stem Cell Res *1*, 45-60.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806-811.

Gabut, M., Samavarchi-Tehrani, P., Wang, X., Slobodeniuc, V., O'Hanlon, D., Sung, H.K., Alvarez, M., Talukder, S., Pan, Q., Mazzoni, E.O., *et al.* (2011). An Alternative Splicing Switch Regulates Embryonic Stem Cell Pluripotency and Reprogramming. Cell *147*, 132-146.

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

Girard, F., Strausfeld, U., Fernandez, A., and Lamb, N.J. (1991). Cyclin A is required for the onset of DNA replication in mammalian fibroblasts. Cell *67*, 1169-1179.

Golob, J.L., Paige, S.L., Muskheli, V., Pabon, L., and Murry, C.E. (2008). Chromatin remodeling during mouse and human embryonic stem cell differentiation. Dev Dyn *237*, 1389-1398.

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

Gordon, R.E., and Lane, B.P. (1980). Duration of cell cycle and its phases measured in synchronized cells of squamous cell carcinoma of rat trachea. Cancer Res *40*, 4467-4472.

Gottesfeld, J.M., and Forbes, D.J. (1997). Mitotic repression of the transcriptional machinery. Trends Biochem Sci *22*, 197-202.

Greber, B., Coulon, P., Zhang, M., Moritz, S., Frank, S., Muller-Molina, A.J., Arauzo-Bravo, M.J., Han, D.W., Pape, H.C., and Scholer, H.R. (2011). FGF signalling inhibits neural induction in human embryonic stem cells. EMBO J *30*, 4874-4884.

Guo, G., Huang, Y., Humphreys, P., Wang, X., and Smith, A. (2011). A PiggyBac-based recessive screening method to identify pluripotency regulators. PLoS One *6*, e18189.

Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G.J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature *404*, 293-296.

Handschick, K., Beuerlein, K., Jurida, L., Bartkuhn, M., Muller, H., Soelch, J., Weber, A., Dittrich-Breiholz, O., Schneider, H., Scharfe, M., *et al.* (2014). Cyclin-dependent kinase 6 is a chromatin-bound cofactor for NF-kappaB-dependent gene expression. Mol Cell *53*, 193-208.

Hanna, J.H., Saha, K., and Jaenisch, R. (2010). Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. Cell *143*, 508-525.

Herrera, R.E., Chen, F., and Weinberg, R.A. (1996). Increased histone H1 phosphorylation and relaxed chromatin structure in Rb-deficient fibroblasts. Proc Natl Acad Sci U S A *93*, 11510-11515.

Hindley, C., and Philpott, A. (2013). The cell cycle and pluripotency. Biochem J *451*, 135-143.

Hirao, A., Kong, Y.Y., Matsuoka, S., Wakeham, A., Ruland, J., Yoshida, H., Liu, D., Elledge, S.J., and Mak, T.W. (2000). DNA damage-induced activation of p53 by the checkpoint kinase Chk2. Science *287*, 1824-1827.

Hong, H., Takahashi, K., Ichisaka, T., Aoi, T., Kanagawa, O., Nakagawa, M., Okita, K., and Yamanaka, S. (2009). Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. Nature *460*, 1132-1135.

Hu, G., Kim, J., Xu, Q., Leng, Y., Orkin, S.H., and Elledge, S.J. (2009). A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. Genes Dev *23*, 837-848.

Hu, G., and Wade, P.A. (2012). NuRD and pluripotency: a complex balancing act. Cell Stem Cell *10*, 497-503.

Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCoste, C., Schafer, X., Lun, Y., and Lemischka, I.R. (2006). Dissecting self-renewal in stem cells with RNA interference. Nature *442*, 533-538.

Jain, A.K., Allton, K., Iacovino, M., Mahen, E., Milczarek, R.J., Zwaka, T.P., Kyba, M., and Barton, M.C. (2012). p53 regulates cell cycle and microRNAs to promote differentiation of human embryonic stem cells. PLoS Biol *10*, e1001268.

Jiang, H., Shukla, A., Wang, X., Chen, W.Y., Bernstein, B.E., and Roeder, R.G. (2011). Role for Dpy-30 in ES cell-fate specification by regulation of H3K4 methylation within bivalent domains. Cell *144*, 513-525.

Judson, R.L., Babiarz, J.E., Venere, M., and Blelloch, R. (2009). Embryonic stem cell-specific microRNAs promote induced pluripotency. Nat Biotechnol *27*, 459-461.

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., *et al.* (2010). Mediator

and cohesin connect gene expression and chromatin architecture. Nature *467*, 430-435.

Kaji, K., Caballero, I.M., MacLeod, R., Nichols, J., Wilson, V.A., and Hendrich, B. (2006). The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. Nat Cell Biol *8*, 285-292.

Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., and Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev *19*, 489-501.

Kastan, M.B., and Bartek, J. (2004). Cell-cycle checkpoints and cancer. Nature *432*, 316-323.

Kawamura, T., Suzuki, J., Wang, Y.V., Menendez, S., Morera, L.B., Raya, A., Wahl, G.M., and Belmonte, J.C. (2009). Linking the p53 tumour suppressor pathway to somatic cell reprogramming. Nature *460*, 1140-1144.

Kim, H., Wu, J., Ye, S., Tai, C.I., Zhou, X., Yan, H., Li, P., Pera, M., and Ying, Q.L. (2013). Modulation of beta-catenin function maintains mouse epiblast stem cell and human embryonic stem cell self-renewal. Nat Commun *4*, 2403.

King, R.W., Jackson, P.K., and Kirschner, M.W. (1994). Mitosis in transition. Cell *79*, 563-571.

Kumar, P., Goh, G., Wongphayak, S., Moreau, D., and Bard, F. (2013). ScreenSifter: analysis and visualization of RNAi screening data. BMC Bioinformatics *14*, 290.

Lanni, J.S., and Jacks, T. (1998). Characterization of the p53-dependent postmitotic checkpoint following spindle disruption. Mol Cell Biol *18*, 1055-1064.

Lara-Gonzalez, P., Westhorpe, F.G., and Taylor, S.S. (2012). The spindle assembly checkpoint. Curr Biol *22*, R966-980.

Lawson, K.A., Meneses, J.J., and Pedersen, R.A. (1991). Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. Development *113*, 891-911.

Lee, K.H., Li, M., Michalowski, A.M., Zhang, X., Liao, H., Chen, L., Xu, Y., Wu, X., and Huang, J. (2010). A genomewide study identifies the Wnt signaling pathway as a major target of p53 in murine embryonic stem cells. Proc Natl Acad Sci U S A *107*, 69-74.

Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K., *et al.* (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. Cell *125*, 301-313.

Lee, Y., Dominy, J.E., Choi, Y.J., Jurczak, M., Tolliday, N., Camporez, J.P., Chim, H., Lim, J.H., Ruan, H.B., Yang, X., *et al.* (2014). Cyclin D1-Cdk4 controls glucose metabolism independently of cell cycle progression. Nature *510*, 547-551.

Leeb, M., Dietmann, S., Paramor, M., Niwa, H., and Smith, A. (2014). Genetic exploration of the exit from self-renewal using haploid embryonic stem cells. Cell Stem Cell *14*, 385-393.

Lefebvre, P., Martin, P.J., Flajollet, S., Dedieu, S., Billaut, X., and Lefebvre, B. (2005). Transcriptional activities of retinoic acid receptors. Vitam Horm *70*, 199-264.

Legartova, S., Kozubek, S., Franek, M., Zdrahal, Z., Lochmanova, G., Martinet, N., and Bartova, E. (2014). Cell differentiation along multiple pathways accompanied by changes in histone acetylation status. Biochem Cell Biol *92*, 85-93.

Li, H., Collado, M., Villasante, A., Matheu, A., Lynch, C.J., Canamero, M., Rizzoti, K., Carneiro, C., Martinez, G., Vidal, A., *et al.* (2012a). p27(Kip1) directly represses Sox2 during embryonic stem cell differentiation. Cell Stem Cell *11*, 845-852.

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

Li, M., He, Y., Dubois, W., Wu, X., Shi, J., and Huang, J. (2012b). Distinct regulatory mechanisms and functions for p53-activated and p53-repressed DNA damage response genes in embryonic stem cells. Mol Cell *46*, 30-42.

Li, V.C., Ballabeni, A., and Kirschner, M.W. (2012c). Gap 1 phase length and mouse embryonic stem cell self-renewal. Proc Natl Acad Sci U S A *109*, 12550-12555.

Liang, J., Wan, M., Zhang, Y., Gu, P., Xin, H., Jung, S.Y., Qin, J., Wong, J., Cooney, A.J., Liu, D., *et al.* (2008). Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. Nat Cell Biol *10*, 731-739.

Lim, S., and Kaldis, P. (2013). Cdks, cyclins and CKIs: roles beyond cell cycle regulation. Development *140*, 3079-3093.

Lin, T., Chao, C., Saito, S., Mazur, S.J., Murphy, M.E., Appella, E., and Xu, Y. (2005). p53 induces differentiation of mouse embryonic stem cells by suppressing Nanog expression. Nat Cell Biol *7*, 165-171.

Lin, W., Srajer, G., Evrard, Y.A., Phan, H.M., Furuta, Y., and Dent, S.Y. (2007). Developmental potential of Gcn5(-/-) embryonic stem cells in vivo and in vitro. Dev Dyn *236*, 1547-1557.

Liu, J.C., Guan, X., Ryan, J.A., Rivera, A.G., Mock, C., Agrawal, V., Letai, A., Lerou, P.H., and Lahav, G. (2013). High mitochondrial priming sensitizes hESCs to DNA-damage-induced apoptosis. Cell Stem Cell *13*, 483-491.

Loh, K.M., and Lim, B. (2011). A precarious balance: pluripotency factors as lineage specifiers. Cell Stem Cell *8*, 363-369.

Loh, Y.H., Yang, L., Yang, J.C., Li, H., Collins, J.J., and Daley, G.Q. (2011). Genomic approaches to deconstruct pluripotency. Annu Rev Genomics Hum Genet *12*, 165-185.

Lu, X., Goke, J., Sachs, F., Jacques, P.E., Liang, H., Feng, B., Bourque, G., Bubulya, P.A., and Ng, H.H. (2013). SON connects the splicing-regulatory network with pluripotency in human embryonic stem cells. Nat Cell Biol *15*, 1141-1152.

Lu, Y., Loh, Y.H., Li, H., Cesana, M., Ficarro, S.B., Parikh, J.R., Salomonis, N., Toh, C.X., Andreadis, S.T., Luckey, C.J., *et al.* (2014). Alternative splicing of MBD2 supports self-renewal in human pluripotent stem cells. Cell Stem Cell *15*, 92-101.

Ludwig, T., and Thomson, J.A. (2007). Defined, feeder-independent medium for human embryonic stem cell culture. Curr Protoc Stem Cell Biol *Chapter 1*, Unit 1C 2.

Ludwig, T.E., Levenstein, M.E., Jones, J.M., Berggren, W.T., Mitchen, E.R., Frane, J.L., Crandall, L.J., Daigh, C.A., Conard, K.R., Piekarczyk, M.S., *et al.* (2006). Derivation of human embryonic stem cells in defined conditions. Nat Biotechnol *24*, 185-187.

Maddox, C.B., Rasmussen, L., and White, E.L. (2008). Adapting Cell-Based Assays to the High Throughput Screening Platform: Problems Encountered and Lessons Learned. JALA Charlottesv Va *13*, 168-173.

Martinez-Balbas, M.A., Dey, A., Rabindran, S.K., Ozato, K., and Wu, C. (1995). Displacement of sequence-specific transcription factors from mitotic chromatin. Cell *83*, 29-38.

McCool, K.W., Xu, X., Singer, D.B., Murdoch, F.E., and Fritsch, M.K. (2007). The role of histone acetylation in regulating early gene expression patterns during early embryonic stem cell differentiation. J Biol Chem *282*, 6696-6706.

McLean, A.B., D'Amour, K.A., Jones, K.L., Krishnamoorthy, M., Kulik, M.J., Reynolds, D.M., Sheppard, A.M., Liu, H., Xu, Y., Baetge, E.E., *et al.* (2007). Activin a efficiently specifies definitive endoderm from human embryonic stem cells only when phosphatidylinositol 3-kinase signaling is suppressed. Stem Cells *25*, 29-38.

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

Mummery, C.L., van Rooijen, M.A., van den Brink, S.E., and de Laat, S.W. (1987). Cell cycle analysis during retinoic acid induced differentiation of a human embryonal carcinoma-derived cell line. Cell Differ *20*, 153-160.

Murray, A.W., and Kirschner, M.W. (1989). Cyclin synthesis drives the early embryonic cell cycle. Nature *339*, 275-280.

Ng, H.H., and Surani, M.A. (2011). The transcriptional and signalling networks of pluripotency. Nat Cell Biol *13*, 490-496.

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

Ohtsubo, M., Theodoras, A.M., Schumacher, J., Roberts, J.M., and Pagano, M. (1995). Human cyclin E, a nuclear protein essential for the G1-to-S phase transition. Mol Cell Biol *15*, 2612-2624.

Pan, G., Tian, S., Nie, J., Yang, C., Ruotti, V., Wei, H., Jonsdottir, G.A., Stewart, R., and Thomson, J.A. (2007). Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. Cell Stem Cell *1*, 299-312.

Pauklin, S., and Vallier, L. (2013). The cell-cycle state of stem cells determines cell fate propensity. Cell *155*, 135-147.

Pedrali-Noy, G., Spadari, S., Miller-Faures, A., Miller, A.O., Kruppa, J., and Koch, G. (1980). Synchronization of HeLa cell cultures by inhibition of DNA polymerase alpha with aphidicolin. Nucleic Acids Res *8*, 377-387.

Peng, J.C., Valouev, A., Swigut, T., Zhang, J., Zhao, Y., Sidow, A., and Wysocka, J. (2009). Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells. Cell *139*, 1290-1302.

Pines, J. (1991). Cyclins: wheels within wheels. Cell Growth Differ *2*, 305-310.

Prescott, D.M., and Bender, M.A. (1962). Synthesis of RNA and protein during mitosis in mammalian tissue culture cells. Exp Cell Res *26*, 260-268.

Pucci, F., Gardano, L., and Harrington, L. (2013). Short Telomeres in ESCs Lead to Unstable Differentiation. Cell Stem Cell *12*, 479-486.

Qiao, J., Paul, P., Lee, S., Qiao, L., Josifi, E., Tiao, J.R., and Chung, D.H. (2012). PI3K/AKT and ERK regulate retinoic acid-induced neuroblastoma cellular differentiation. Biochem Biophys Res Commun *424*, 421-426.

Qin, H., Yu, T., Qing, T., Liu, Y., Zhao, Y., Cai, J., Li, J., Song, Z., Qu, X., Zhou, P., *et al.* (2007). Regulation of apoptosis and differentiation by p53 in human embryonic stem cells. J Biol Chem *282*, 5842-5852.

Recolin, B., van der Laan, S., Tsanov, N., and Maiorano, D. (2014). Molecular mechanisms of DNA replication checkpoint activation. Genes (Basel) *5*, 147-175.

Reynolds, N., Latos, P., Hynes-Allen, A., Loos, R., Leaford, D., O'Shaughnessy, A., Mosaku, O., Signolet, J., Brennecke, P., Kalkan, T., *et al.* (2012). NuRD suppresses pluripotency gene expression to promote transcriptional heterogeneity and lineage commitment. Cell Stem Cell *10*, 583-594.

Rodier, F., Coppe, J.P., Patil, C.K., Hoeijmakers, W.A., Munoz, D.P., Raza, S.R., Freund, A., Campeau, E., Davalos, A.R., and Campisi, J. (2009). Persistent DNA damage signalling triggers senescence-associated inflammatory cytokine secretion. Nat Cell Biol *11*, 973-979.

Romero, O.A., and Sanchez-Cespedes, M. (2013). The SWI/SNF genetic blockade: effects in cell differentiation, cancer and developmental diseases. Oncogene *33*, 2681-2689.

Romero, O.A., Setien, F., John, S., Gimenez-Xavier, P., Gomez-Lopez, G., Pisano, D., Condom, E., Villanueva, A., Hager, G.L., and Sanchez-Cespedes, M. (2012). The tumour suppressor and chromatin-remodelling factor BRG1 antagonizes Myc activity and promotes cell differentiation in human cancer. EMBO Mol Med *4*, 603-616.

Roos, W.P., Christmann, M., Fraser, S.T., and Kaina, B. (2007). Mouse embryonic stem cells are hypersensitive to apoptosis triggered by the DNA damage O(6)-methylguanine due to high E2F1 regulated mismatch repair. Cell death and differentiation *14*, 1422-1432.

Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and Brivanlou, A.H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. Nat Med *10*, 55-63.

Sela, Y., Molotski, N., Golan, S., Itskovitz-Eldor, J., and Soen, Y. (2012). Human embryonic stem cells exhibit increased propensity to differentiate during the G1 phase prior to phosphorylation of retinoblastoma protein. Stem Cells *30*, 1097-1108.

Sha, K., and Boyer, L.A. (2008). The chromatin signature of pluripotent cells, 2010 edn (Cambridge (MA), Harvard Stem Cell Institute).

Sharma, S., and Rao, A. (2009). RNAi screening: tips and techniques. Nat Immunol *10*, 799-804.

Sherr, C.J. (1994). G1 phase progression: cycling on cue. Cell *79*, 551-555.

Shilo, A., Ben Hur, V., Denichenko, P., Stein, I., Pikarsky, E., Rauch, J., Kolch, W., Zender, L., and Karni, R. (2014). Splicing factor hnRNP A2 activates the Ras-MAPK-ERK pathway by controlling A-Raf splicing in hepatocellular carcinoma development. RNA *20*, 505-515.

Singh, A.M., Chappell, J., Trost, R., Lin, L., Wang, T., Tang, J., Wu, H., Zhao, S., Jin, P., and Dalton, S. (2013). Cell-cycle control of developmentally regulated transcription factors accounts for heterogeneity in human pluripotent cells. Stem Cell Reports *1*, 532-544.

Singh, A.M., and Dalton, S. (2009). The cell cycle and Myc intersect with mechanisms that regulate pluripotency and reprogramming. Cell Stem Cell *5*, 141-149.

Singh, A.M., Reynolds, D., Cliff, T., Ohtsuka, S., Mattheyses, A.L., Sun, Y., Menendez, L., Kulik, M., and Dalton, S. (2012). Signaling network crosstalk in human pluripotent cells: a Smad2/3-regulated switch that controls the balance between self-renewal and differentiation. Cell Stem Cell *10*, 312-326.

Smith, K.N., Singh, A.M., and Dalton, S. (2010). Myc represses primitive endoderm differentiation in pluripotent stem cells. Cell Stem Cell *7*, 343-354.

Stead, E., White, J., Faast, R., Conn, S., Goldstone, S., Rathjen, J., Dhingra, U., Rathjen, P., Walker, D., and Dalton, S. (2002). Pluripotent cell division cycles are driven by ectopic Cdk2, cyclin A/E and E2F activities. Oncogene *21*, 8320-8333.

Sumi, T., Tsuneyoshi, N., Nakatsuji, N., and Suemori, H. (2008). Defining early lineage specification of human embryonic stem cells by the orchestrated balance of canonical Wnt/beta-catenin, Activin/Nodal and BMP signaling. Development *135*, 2969-2979.

Tabar, V., and Studer, L. (2014). Pluripotent stem cells in regenerative medicine: challenges and recent progress. Nat Rev Genet *15*, 82-92.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861-872.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

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

Tibbetts, R.S., Brumbaugh, K.M., Williams, J.M., Sarkaria, J.N., Cliby, W.A., Shieh, S.Y., Taya, Y., Prives, C., and Abraham, R.T. (1999). A role for ATR in the DNA damage-induced phosphorylation of p53. Genes Dev *13*, 152-157.

Tichy, E.D., Pillai, R., Deng, L., Liang, L., Tischfield, J., Schwemberger, S.J., Babcock, G.F., and Stambrook, P.J. (2010). Mouse embryonic stem cells, but not somatic cells, predominantly use homologous recombination to repair double-strand DNA breaks. Stem Cells Dev *19*, 1699-1711.

Tsubouchi, T., Soza-Ried, J., Brown, K., Piccolo, F.M., Cantone, I., Landeira, D., Bagci, H., Hochegger, H., Merkenschlager, M., and Fisher, A.G. (2013). DNA synthesis is required for reprogramming mediated by stem cell fusion. Cell *152*, 873-883.

Ubil, E., Duan, J., Pillai, I.C., Rosa-Garrido, M., Wu, Y., Bargiacchi, F., Lu, Y., Stanbouly, S., Huang, J., Rojas, M., *et al.* (2014). Mesenchymal-endothelial transition contributes to cardiac neovascularization. Nature *514*, 585-590.

Ui-Tei, K., Zenno, S., Miyata, Y., and Saigo, K. (2000). Sensitive assay of RNA interference in Drosophila and Chinese hamster cultured cells using firefly luciferase gene as target. FEBS Lett *479*, 79-82.

Ullmann, U., Gilles, C., De Rycke, M., Van de Velde, H., Sermon, K., and Liebaers, I. (2008). GSK-3-specific inhibitor-supplemented hESC medium prevents the epithelial-mesenchymal transition process and the up-regulation of matrix metalloproteinases in hESCs cultured in feeder-free conditions. Mol Hum Reprod *14*, 169-179.

Vallier, L., Mendjan, S., Brown, S., Chng, Z., Teo, A., Smithers, L.E., Trotter, M.W., Cho, C.H., Martinez, A., Rugg-Gunn, P., *et al.* (2009). Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. Development *136*, 1339-1349.

Van Dongen, S. (2008). Graph Clustering Via a Discrete Uncoupling Process. SIAM Journal on Matrix Analysis and Applications *30*, 121-141.

Varlakhanova, N.V., Cotterman, R.F., deVries, W.N., Morgan, J., Donahue, L.R., Murray, S., Knowles, B.B., and Knoepfler, P.S. (2010). myc maintains embryonic stem cell pluripotency and self-renewal. Differentiation *80*, 9-19.

Vassilev, L.T., Vu, B.T., Graves, B., Carvajal, D., Podlaski, F., Filipovic, Z., Kong, N., Kammlott, U., Lukacs, C., Klein, C., *et al.* (2004). In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. Science *303*, 844-848.

Vazin, T., and Freed, W.J. (2010). Human embryonic stem cells: derivation, culture, and differentiation: a review. Restor Neurol Neurosci *28*, 589-603.

Vermeulen, K., Van Bockstaele, D.R., and Berneman, Z.N. (2003). The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. Cell Prolif *36*, 131-149.

Walker, D.H., and Maller, J.L. (1991). Role for cyclin A in the dependence of mitosis on completion of DNA replication. Nature *354*, 314-317.

Wan, M., Liang, J., Xiong, Y., Shi, F., Zhang, Y., Lu, W., He, Q., Yang, D., Chen, R., Liu, D., *et al.* (2013). The trithorax group protein Ash2l is essential for pluripotency and maintaining open chromatin in embryonic stem cells. J Biol Chem *288*, 5039-5048.

Wang, F., Soprano, K.J., and Soprano, D.R. (2014). Role of Acinus in Regulating Retinoic Acid-responsive Gene Pre-mRNA Splicing. J Cell Physiol *230*, 791-801.

Wang, J., Sun, Q., Morita, Y., Jiang, H., Gross, A., Lechel, A., Hildner, K., Guachalla, L.M., Gompf, A., Hartmann, D., *et al.* (2012). A differentiation checkpoint limits hematopoietic stem cell self-renewal in response to DNA damage. Cell *148*, 1001-1014.

Wang, X., Lin, G., Martins-Taylor, K., Zeng, H., and Xu, R.H. (2009). Inhibition of caspase-mediated anoikis is critical for basic fibroblast growth factor-sustained culture of human pluripotent stem cells. J Biol Chem *284*, 34054-34064.

Wang, Y., Medvid, R., Melton, C., Jaenisch, R., and Blelloch, R. (2007). DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat Genet *39*, 380-385.

Westerman, B.A., Braat, A.K., Taub, N., Potman, M., Vissers, J.H., Blom, M., Verhoeven, E., Stoop, H., Gillis, A., Velds, A., *et al.* (2011). A genome-wide RNAi screen in mouse embryonic stem cells identifies Mp1 as a key mediator of differentiation. J Exp Med *208*, 2675-2689.

Xu, C., Rosler, E., Jiang, J., Lebkowski, J.S., Gold, J.D., O'Sullivan, C., Delavan-Boorsma, K., Mok, M., Bronstein, A., and Carpenter, M.K. (2005). Basic fibroblast growth factor supports undifferentiated human embryonic stem cell growth without conditioned medium. Stem Cells *23*, 315-323.

Xu, R.H., Sampsell-Barron, T.L., Gu, F., Root, S., Peck, R.M., Pan, G., Yu, J., Antosiewicz-Bourget, J., Tian, S., Stewart, R., *et al.* (2008). NANOG is a direct target of TGFbeta/activin-mediated SMAD signaling in human ESCs. Cell Stem Cell *3*, 196-206.

Yarden, A., and Geiger, B. (1996). Zebrafish cyclin E regulation during early embryogenesis. Dev Dyn *206*, 1-11.

Young, R.A. (2011). Control of the embryonic stem cell state. Cell *144*, 940-954.

Yu, J., and Thomson, J.A. (2008). Pluripotent stem cell lines. Genes Dev *22*, 1987-1997.

Zamore, P.D., Tuschl, T., Sharp, P.A., and Bartel, D.P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell *101*, 25-33.

Zhang, X., Li, B., Li, W., Ma, L., Zheng, D., Li, L., Yang, W., Chu, M., Chen, W., Mailman, R.B., *et al.* (2014). Transcriptional repression by the BRG1-SWI/SNF complex affects the pluripotency of human embryonic stem cells. Stem Cell Reports *3*, 460-474.

Zhang, X., Neganova, I., Przyborski, S., Yang, C., Cooke, M., Atkinson, S.P., Anyfantis, G., Fenyk, S., Keith, W.N., Hoare, S.F., *et al.* (2009). A role for NANOG in G1 to S transition in human embryonic stem cells through direct binding of CDK6 and CDC25A. J Cell Biol *184*, 67-82.

Zhao, X.D., Han, X., Chew, J.L., Liu, J., Chiu, K.P., Choo, A., Orlov, Y.L., Sung, W.K., Shahab, A., Kuznetsov, V.A., *et al.* (2007). Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. Cell Stem Cell *1*, 286-298.

Zou, L., and Elledge, S.J. (2003). Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. Science *300*, 1542-1548.

## Appendices

Appendix 1. List of qPCR primers

| Gene | Forward primer | Reverse primer |
|---|---|---|
| GAPDH | GGCTGTGGGCAAGGTCATCCCTGAG | GTCGCTGTTGAAGTCAGAGGAGACCACCTG |
| POU5F1 | TCTCCGTCACCACTCTGGGCTCTCC | GCATCATTGAACTTCACCTTCCCTCCAACC |
| NANOG | ACCAGAACTGTGTTCTCTTCCACC | CCATTGCTATTCTTCGGCCAGTTG |
| TERF1 | TGTTTCAAAGAGTCAGCCGGTAA | GTCTTCGCTGTCTGAGGAAATCA |
| PRDM14 | GCGGATGCTTCTCTGTTACC | ACCCCGTACAGAACGAAGTG |
| TDGF1 | TGCTGGGGTCCTTTTGTGCCT | ATGCCTGAGGAAAGCAGCGGA |
| AURKA | TCCAGTCACAAGCCGGTTCA | ACCCAGAGGGCGACCAATTT |
| CDK11A | CAAGCAGGAGCTGCCCAAGT | TGGGCCTTGAGGATGGTGTT |
| PPP1CA | CGAGTGTGCCAGCATCAACC | CCGCCGAATCTGCTCCATAG |
| CSNK2B | CGCCCGCTACATCCTTACCA | GTAGGCGCCATCCGTGTGAT |
| CDK1 | TGAAGTGTGGCCAGAAGTGGAA | CAGAAATTCGTTTGGCTGGATCA |
| PPP1R12A | CCTCTGGCCCGCACTCATAG | CTGCTCGTTCCGCTTCTGCT |
| PCNA | CCGGAGGCTCTAGCCTGACA | AGTGGCAACAACGCCGCTAC |
| MCM5 | TGCCTTGTCCGGTACCCTGT | GAGCTGCAGCACCTTGTGGA |
| ORC1L | TTACCCGGCCTGTGACAGCAA | TGGCACTTGGCTGCACTCTCT |
| GMNN | GGACAATGAAATTGCCCGCCTGA | TTCCTTCAGCACACGTGCCAAT |
| CCDC6 | CCAGCCCGATCCCTTACACA | CGTGGGCCGTTTGAATTTGT |
| BARD1 | GCTTGCAATCATGGGCACCT | TAATCGACAGGCCGCAGACC |
| SETD2 | CAGCCCTTGGTGGGACATTC | GCACTGGCAAGACAGCAACG |
| ROCK1 | GGCGGTCTCCGTTTGTTTGA | TCGATGCCCGATGGAGACTT |
| PPP1R8 | CTTGCTCTCGGGTCCATGCT | TTGTGGATGCGCCAAATGAG |

| | | |
|---|---|---|
| KAT2A | AGCTTCAGTCAGTGCAGCGGT | ATGTCACCCATCACACGGAGCC |
| SKIL | TGCGGCCACGAACTTTTCCTCA | ACACGGAGCATCAGGCTGAACA |
| TRIB3 | AAGCCCTGTTCTCGGTGCTG | TGGAGGCCGACACTGGTACA |
| ZIC3 | AGGCCAGAGCTCGGCTTTCA | TGCGCTGGCGGAACAGAAACT |
| PPAPDC3 | CATGCAGCTGAACCCCTCCTT | AGTGTGCTGCTCTTCACCAGG |
| CDK4 | GGAGGAGGCCTTCCCATCAG | GGATGTGGCACAGACGTCCA |
| CDK6 | TGGAGTGCCCACTGAAACCA | CAAGGCCGAAGTCAGCGAGT |
| CCND1 | AGCAGATCGAAGCCCTGCTG | CCGGAGAGGAGGGACTGTCA |
| CCND2 | GCCATGTACCCACCGTCGAT | CACCGCCTCAATCTGCTCCT |
| ATR | TGGACCTGGAGGCAACCATT | CGTGATTGTGCCTGTGGTGAA |
| ATM | GCATGGGCATTACGGGTGTT | CATCTTCCGGCCTCTGCTGT |
| CHEK1 | CCTGGCAGCGGTTGGTCAAAA | GCCAAATCTTCTGGCTGCTCACA |
| CHEK2 | AAGACACCCGTGGCTTCAGGA | ACAGCTGGGCGCTTTGTGG |
| P53 | GTGCGTGTTTGTGCCTGTCC | GCTGGTGTTGTTGGGCAGTG |
| NODAL | CTTCGGCAGACACAGCCTGA | CTGAGGTCGCACAGCTTCCA |
| GDF3 | TCTGCCACCGTCACCAGCTA | ATGCATCAGGGCTTGCATGA |
| BMP4 | TCCACTGGCTGACCACCTCA | CGGCACCCACATCCCTCTAC |
| ID2 | ACGACCCGATGAGCCTGCTA | GTCCTGGACGCCTGGTTCTG |
| CCNB1 | GATGCTGCAGCTGGTTGGTG | AGGCCGACCCAGACCAAAGT |

Appendix 2. List of RNAi sequences

| **Gene** | **RNAi 1** | **RNAi 2** |
|---|---|---|
| Non-targeting | siGENOME D-001210-01-20 (Thermo Scientific) | |
| POU5F1 | siGENOME D-019591-05 (Thermo Scientific) | |
| AURKA | GGTCTTGTGTCCTTCAAATTC | GCACCACTTGGAACAGTTTAT |
| CDK11A | GCCTGATGGAGACCATGAAAC | GCTGCTTGGTGCCAAGGAATA |
| PPP1CA | GCAGTCTATGGAGCAGATTCG | GCAATTCCGCCAAAGCCAAGA |
| CSNK2B | GCAAGGAGACTTTGGTTACTG | GGCTCCGTGGCAATGAATTCT |
| CDK1 | GGTCAGTACATGGATTCTTCA | GGGTCAGCTCGTTACTCAACT |
| PPP1R12A | GCTCTGGATCATACAGTTACT | GCTAATGCAGAGTAGGCTACT |
| PCNA | CACGTATATGCCGAGATCT | GGAGGCTCTAGCCTGACAAAT |
| MCM5 | CTGAAGTCGGACATGATGT | ACGTGAGCGCACTGACACA |
| ORC1L | GCTTCACCTGAACATCGCA | CAGAGATTTCAGACTCTAG |
| GMNN | GAGCTAATAGAGAGACTGA | GTATATGGCAGAGCTAATA |
| CCDC6 | CAGTGTTGTGGGTCTACAA | GCAACAGTCCTGACAAATTCA |
| BARD1 | GTGGTTTAGCCCTCGAAGT | CTCCTTCGAGCTGGTTTAT |
| SETD2 | GCAGGACACTATATCTAATAG | CTCAACACACCTGATCCTT |
| ROCK1 | GCATTTGGAGAAGTTCAATTG | GCTTGAAGCTGAGCAATATTT |
| PPP1R8 | GCACTTTCTTGGGTCACATTC | GCCCACAACAAGCGGATTTCT |
| KAT2A | GCGCATGCCTAAGGAGTATAT | GCAGCTCTTTCTGGACCTTCA |
| SKIL | GGGCATACTTCCATTCAATGC | GGAATGGAATTACAGTCATGG |
| TRIB3 | GCCCTACAGGCACTGAGTATA | GGTTGGAGTTGGATGACAACT |
| ZIC3 | GGGAAGATCTTTGCCCGTTCT | GCCAGTTCAGGCTATGAATCT |
| PPAPDC3 | GCAACTCTCCACACTATATTT | GCCATCGATATCTGTATGTCC |
| CDK4 | GCATGTAGACCAGGACCTAAG | |
| CDK6 | GGATATGATGTTTCAGCTTCT | |
| CCND1 | GCATGTTCGTGGCCTCTAAGA | GCTTCCTCTCCAGAGTGATCA |

| | | |
|---|---|---|
| CCND2 | GCATGCTCAGACCTTCATTGC | |
| ATR | GGGAGAAACCTTTGAAGTTCC | |
| ATM | GCATTCAGATTCCAAACAAGG | |
| CHEK1 | GCAGACAAATCTTATCAATGC | |
| CHEK2 | GGAGAGGTAAAGCTGGCTTTC | |
| P53 | GCGCACAGAGGAAGAGAATCT | GGAAGACTCCAGTGGTAATCT |
| CCNB1 | GGTTGTTGCAGGAGACCATGT | |
| CCNA2 | GGATCTTCCTGTAAATGATGA | |
| CCNE1 | GCAGCCAAACTTGAGGAAATC | |

Appendix 3. List of hits in the - bFGF, - TGFβ condition

| Gene Symbol | Entrez Gene ID | Mean Z-score | Mean Cell Viability Score |
|---|---|---|---|
| SMARCA5 | 8467 | 5.869282 | 0.839829 |
| ACTL6A | 86 | 5.483016 | 0.161976 |
| CCDC6 | 8030 | 5.460845 | 0.982402 |
| DDX6 | 1656 | 4.9581 | 0.308243 |
| KAT5 | 10524 | 4.420716 | 0.259281 |
| CNOT2 | 4848 | 4.326923 | 0.364522 |
| CNOT3 | 4849 | 4.318167 | 0.542582 |
| ENY2 | 56943 | 3.597846 | 0.854299 |
| HSF1 | 3297 | 3.528116 | 0.865967 |
| POLR1A | 25885 | 3.166428 | 0.160843 |
| SETD2 | 29072 | 3.035043 | 0.577263 |
| NOC2L | 26155 | 3.022793 | 0.190277 |
| TADA3 | 10474 | 2.989501 | 1.159938 |
| MEAF6 | 64769 | 2.884428 | 0.502794 |
| PRPF4B | 8899 | 2.87604 | 0.730647 |
| ILF3 | 3609 | 2.82386 | 0.267254 |
| DHX16 | 8449 | 2.785978 | 0.247583 |
| BARD1 | 580 | 2.75891 | 0.357742 |
| SMARCC1 | 6599 | 2.738959 | 0.549892 |
| TSG101 | 7251 | 2.651138 | 0.292815 |
| ZNF140 | 7699 | 2.620046 | 0.794513 |
| CSNK1A1 | 1452 | 2.612363 | 0.82705 |
| MAX | 4149 | 2.594695 | 0.674808 |
| NAT10 | 55226 | 2.570848 | 0.322444 |
| GABPB1 | 2553 | 2.529705 | 0.317756 |
| RXRA | 6256 | 2.524585 | 0.455905 |
| CHD3 | 1107 | 2.447744 | 0.225399 |
| BPTF | 2186 | 2.282481 | 0.887256 |
| RARB | 5915 | 2.276966 | 0.396075 |
| ZNRF1 | 84937 | 2.270899 | 1.291893 |
| BUB1 | 699 | 2.270332 | 0.704695 |
| FOXC2 | 2303 | 2.242375 | 0.357211 |
| TRIB3 | 57761 | 2.234426 | 0.595736 |
| ZNF860 | 344787 | 2.222835 | 0.621504 |
| PHF12 | 57649 | 2.201726 | 0.922017 |
| CDC2 | 983 | 2.18586 | 0.880317 |
| HIST1H2BE | 8344 | 2.170879 | 0.372116 |
| UNCX | 340260 | 2.143311 | 0.611737 |
| RBM7 | 10179 | 2.136006 | 0.462185 |
| PATZ1 | 23598 | 2.031076 | 0.329792 |
| SMARCA4 | 6597 | 1.968374 | 0.302959 |
| GPR4 | 2828 | 1.89507 | 0.913843 |

| | | | |
|---|---|---|---|
| SKIL | 6498 | 1.880279 | 1.119721 |
| HIST1H1C | 3006 | 1.776697 | 0.641745 |
| LOC91461 | 91461 | 1.666532 | 0.925055 |
| PPP1R12A | 4659 | 1.595832 | 0.895704 |
| GPR22 | 2845 | 1.465623 | 0.844729 |
| YEATS2 | 55689 | 2.762648 | 0.307423 |
| DMAP1 | 55929 | 2.641341 | 0.248067 |
| ILF2 | 3608 | 2.598684 | 0.208361 |
| CHD8 | 57680 | 2.595751 | 0.544948 |
| DHX37 | 57647 | 2.589043 | 0.265184 |
| MAZ | 4150 | 2.573902 | 0.577391 |
| NFKB2 | 4791 | 2.566391 | 0.580524 |
| ESPL1 | 9700 | 2.531161 | 0.231478 |
| NR2C2 | 7182 | 2.454664 | 0.382861 |
| C20orf20 | 55257 | 2.441838 | 0.606945 |
| LAS1L | 81887 | 2.426053 | 0.273103 |
| BRD8 | 10902 | 2.383031 | 0.641544 |
| SNIP1 | 79753 | 2.37501 | 0.248388 |
| TRRAP | 8295 | 2.358355 | 0.352415 |
| ZIC3 | 7547 | 2.346794 | 0.484757 |
| C11orf30 | 56946 | 2.300292 | 1.133289 |
| CDC2L2 | 728642 | 2.281274 | 0.200104 |
| CRK7 | 51755 | 2.275951 | 0.877251 |
| CSNK2B | 1460 | 2.187604 | 0.858716 |
| FUS | 2521 | 2.149051 | 0.83369 |
| EIF5B | 9669 | 2.132064 | 0.223154 |
| SAP18 | 10284 | 2.118492 | 0.274155 |
| CIC | 23152 | 2.09946 | 0.482677 |
| KAT2A | 2648 | 2.094975 | 0.8386 |
| ZBTB12 | 221527 | 2.092462 | 1.102796 |
| PYGO1 | 26108 | 2.082673 | 0.282389 |
| DPF2 | 5977 | 2.048047 | 0.342108 |
| COPS5 | 10987 | 2.047223 | 0.30076 |
| TCF7L1 | 83439 | 2.035807 | 0.338124 |
| ETV5 | 2119 | 2.028036 | 0.379003 |
| EXOSC10 | 5394 | 2.015847 | 0.546531 |
| ZNHIT6 | 54680 | 1.992057 | 0.54084 |
| RBMX | 27316 | 1.989799 | 0.304618 |
| BRCA2 | 675 | 1.987551 | 0.282202 |
| MTF1 | 4520 | 1.982405 | 0.558391 |
| EYA1 | 2138 | 1.975624 | 0.174405 |
| CDC25A | 993 | 1.958147 | 0.800849 |
| TPR | 7175 | 1.943308 | 0.52714 |
| SMARCE1 | 6605 | 1.932042 | 0.508457 |

| | | | |
|---|---|---|---|
| ORC1L | 4998 | 1.930002 | 0.428054 |
| OPN3 | 23596 | 1.920588 | 0.834964 |
| GMNN | 51053 | 1.91049 | 0.312537 |
| SIN3B | 23309 | 1.908024 | 0.740513 |
| AURKA | 6790 | 1.893505 | 0.963257 |
| BARX1 | 56033 | 1.890243 | 0.28115 |
| AHRR | 57491 | 1.84469 | 0.172904 |
| CCT4 | 10575 | 1.842673 | 0.282939 |
| ARID1A | 8289 | 1.833016 | 0.437272 |
| MYEOV2 | 150678 | 1.814735 | 0.759619 |
| MBD3 | 53615 | 1.814098 | 0.539505 |
| WDR77 | 79084 | 1.807729 | 0.553592 |
| YY1 | 7528 | 1.80516 | 0.160275 |
| SETD8 | 387893 | 1.801512 | 0.187206 |
| SFRS6 | 6431 | 1.799853 | 0.842518 |
| ZSCAN12 | 9753 | 1.799274 | 0.493184 |
| H2AFZ | 3015 | 1.793877 | 0.578258 |
| KLHL41 | 10324 | 1.793181 | 0.78237 |
| GMEB1 | 10691 | 1.774919 | 0.638417 |
| OGT | 8473 | 1.772843 | 0.636149 |
| MDM4 | 4194 | 1.762111 | 0.452708 |
| KIAA1310 | 55683 | 1.757319 | 0.01651 |
| TSPYL3 | 128854 | 1.749925 | 1.172408 |
| CETN2 | 1069 | 1.727207 | 0.922015 |
| KLHL14 | 57565 | 1.721915 | 0.697751 |
| RBBP4 | 5928 | 1.709623 | 0.246354 |
| ATXN2 | 6311 | 1.706602 | 0.926146 |
| BTAF1 | 9044 | 1.701417 | 0.934329 |
| ZMYND8 | 23613 | 1.698474 | 0.971272 |
| DIDO1 | 11083 | 1.691884 | 0.405602 |
| PROP1 | 5626 | 1.660303 | 0.267848 |
| TCF23 | 150921 | 1.630527 | 0.610948 |
| GPR147 | 64106 | 1.624486 | 0.630393 |
| NRF1 | 4899 | 1.621229 | 0.248604 |
| MC3R | 4159 | 1.621174 | 0.501381 |
| IL13 | 3596 | 1.603541 | 0.475624 |
| TBX1 | 6899 | 1.60115 | 0.963401 |
| DDX27 | 55661 | 1.597007 | 0.258032 |
| HTT | 3064 | 1.594034 | 1.113343 |
| C12orf41 | 54934 | 1.582244 | 0.43316 |
| ZNF48 | 197407 | 1.578245 | 0.691743 |
| GON4L | 54856 | 1.578158 | 1.230509 |
| TFDP2 | 7029 | 1.566201 | 0.949364 |
| ROCK1 | 6093 | 1.561698 | 0.839824 |

| | | | |
|---|---|---|---|
| CASP2 | 835 | 1.555531 | 0.562889 |
| TACR2 | 6865 | 1.553305 | 0.744828 |
| CBFA2T3 | 863 | 1.551773 | 0.327733 |
| ZDHHC1 | 29800 | 1.543371 | 0.804975 |
| TCEB2 | 6923 | 1.541395 | 0.305174 |
| THAP11 | 57215 | 1.540047 | 0.94406 |
| JARID2 | 3720 | 1.536811 | 0.7062 |
| ARID3B | 10620 | 1.534729 | 0.439162 |
| NAGK | 55577 | 1.532449 | 0.969585 |
| TIGD4 | 201798 | 1.506258 | 1.245067 |
| AIP | 9049 | 1.506237 | 0.124029 |
| SVEP1 | 79987 | 1.501175 | 0.38907 |
| CDK8 | 1024 | 1.479287 | 0.94925 |
| IRF2 | 3660 | 1.475936 | 0.857523 |
| MMP9 | 4318 | 1.475932 | 0.329346 |
| SETMAR | 6419 | 1.457933 | 0.071716 |
| NTRK2 | 4915 | 1.457137 | 0.984119 |
| YBX1 | 4904 | 1.44806 | 0.418119 |
| NFKBIL2 | 4796 | 1.447428 | 0.531313 |
| EMX1 | 2016 | 1.437284 | 0.449538 |
| TFAP2B | 7021 | 1.430711 | 0.810662 |
| ATG4B | 23192 | 1.417206 | 0.860999 |
| WBP11 | 51729 | 1.416754 | 0.391203 |
| DDX19A | 55308 | 1.407622 | 0.947153 |
| EP400 | 57634 | 1.400543 | 0.329437 |
| TIMM50 | 92609 | 1.397515 | 0.888602 |
| BPNT1 | 10380 | 1.3933 | 0.929577 |
| ZNF324 | 25799 | 1.388203 | 0.614092 |
| FIGLA | 344018 | 1.387577 | 0.485383 |
| TICRR | 90381 | 1.382642 | 0.500659 |
| NUDT21 | 11051 | 1.378719 | 0.196149 |
| STON1 | 11037 | 1.377385 | 0.78346 |
| BIRC5 | 332 | 1.373331 | 0.266309 |
| PCNA | 5111 | 1.372453 | 0.265116 |
| DRAP1 | 10589 | 1.371026 | 1.161495 |
| ZNF70 | 7621 | 1.369033 | 0.941417 |
| PPP1CA | 5499 | 1.368785 | 0.348042 |
| SRCAP | 10847 | 1.366842 | 0.172093 |
| DUSP6 | 1848 | 1.348194 | 0.885348 |
| DUSP10 | 11221 | 1.336001 | 0.774118 |
| PDX1 | 3651 | 1.326863 | 0.803294 |
| DPF1 | 8193 | 1.31987 | 0.466464 |
| DR1 | 1810 | 1.308609 | 1.109319 |
| DHX58 | 79132 | 1.307035 | 0.556235 |

| | | | |
|---|---|---|---|
| ZNF574 | 64763 | 1.300668 | 0.109876 |
| GPR153 | 387509 | 1.299487 | 0.860915 |
| MCM5 | 4174 | 1.278632 | 0.416658 |
| ASB12 | 142689 | 1.276394 | 0.377423 |
| RAP1GAP | 5909 | 1.274286 | 0.916226 |
| FFAR1 | 2864 | 1.265985 | 0.951375 |
| OTX2 | 5015 | 1.244985 | 0.561147 |
| UMPK | 7371 | 1.244717 | 1.04597 |
| DHX9 | 1660 | 1.233013 | 0.294378 |
| HES2 | 54626 | 1.229457 | 0.562685 |
| NUP62 | 23636 | 1.215847 | 0.766179 |
| GPR156 | 165829 | 1.206866 | 0.872205 |
| CDKL3 | 51265 | 1.205719 | 0.997066 |
| DMRTB1 | 63948 | 1.201494 | 0.249506 |
| PTGFR | 5737 | 1.199582 | 0.895978 |
| NEK1 | 4750 | 1.193928 | 0.911078 |
| IRAK2 | 3656 | 1.189881 | 0.97351 |
| NKAP | 79576 | 1.185625 | 0.289486 |
| HELZ | 9931 | 1.185075 | 0.723814 |
| ZFP3 | 124961 | 1.177928 | 0.216375 |
| EDG5 | 9294 | 1.167201 | 0.880713 |
| PTEN | 5728 | 1.141416 | 0.928846 |
| CHMP6 | 79643 | 1.106687 | 0.707924 |
| FUK | 197258 | 1.100659 | 0.995415 |
| ZNF467 | 168544 | 1.100314 | 0.417385 |
| PTRF | 284119 | 1.082397 | 0.775244 |
| GPR12 | 2835 | 1.079001 | 0.947832 |
| NT5C | 30833 | 0.988326 | 0.809417 |
| ZIC4 | 84107 | 0.943337 | 0.570531 |
| IKBKAP | 8518 | 0.935737 | 0.796106 |
| SPEN | 23013 | 0.752955 | 0.191599 |
| ZNF407 | 55628 | 0.242782 | 0.029625 |

Appendix 4. List of hits for the TGFβ pathway inhibition condition

| Gene Symbol | Entrez Gene ID | Mean Z-score | Mean Cell Viability Score |
|---|---|---|---|
| KAT5 | 10524 | 8.500555 | 0.093213 |
| SMARCA5 | 8467 | 7.834006 | 0.251136 |
| DDX6 | 1656 | 7.306216 | 0.940609 |
| TADA3 | 10474 | 6.822607 | 0.724203 |
| MAX | 4149 | 5.972157 | 0.461944 |
| CCDC6 | 8030 | 5.168458 | 0.116037 |
| SETD2 | 29072 | 4.687285 | 0.504832 |
| TRRAP | 8295 | 4.522646 | 0.037156 |
| CIC | 23152 | 4.394846 | 0.158667 |
| CNOT2 | 4848 | 4.061287 | 0.096676 |
| SKIL | 6498 | 3.850636 | 0.448934 |
| HSF1 | 3297 | 3.706731 | 0.080219 |
| AIP1 | 9863 | 3.702754 | 0.001365 |
| SOX11 | 6664 | 3.674094 | 1.040297 |
| AOF2 | 23028 | 3.43429 | 0.239363 |
| ZMYND8 | 23613 | 3.317655 | 0.533291 |
| BRD4 | 23476 | 3.263565 | 0.107031 |
| EP400 | 57634 | 3.255099 | 0.042811 |
| PIK3R2 | 5296 | 3.199677 | 0.001781 |
| HELT | 391723 | 3.10906 | 0.563268 |
| ENY2 | 56943 | 3.108157 | 0.546719 |
| KDM5B | 10765 | 3.097936 | 0.110372 |
| CNOT3 | 4849 | 3.088913 | 0.109915 |
| ZBTB12 | 221527 | 3.048454 | 0.499277 |
| DNMT1 | 1786 | 2.969742 | 0.167631 |
| SKI | 6497 | 2.819455 | 0.552766 |
| STON1 | 11037 | 2.80089 | 0.038367 |
| RSF1 | 51773 | 2.736912 | 0.387407 |
| LSM12 | 124801 | 2.730703 | 0.239472 |
| NFYC | 4802 | 2.710035 | 0.504298 |
| XRCC3 | 7517 | 2.700963 | 0.163053 |
| ZBTB39 | 9880 | 2.66349 | 0.843745 |
| ZNF287 | 57336 | 2.657818 | 0.222059 |
| MAF1 | 84232 | 2.657211 | 0.074596 |
| PTGS2 | 5743 | 2.593612 | 0.271994 |
| HIST1H2AK | 8330 | 2.479352 | 0.225366 |
| ZNF846 | 162993 | 2.460631 | 0.088816 |
| ZBTB46 | 140685 | 2.455202 | 0.172015 |
| ZNF140 | 7699 | 2.443684 | 0.093163 |
| BAZ1B | 9031 | 2.434823 | 0.632911 |
| ZNRF1 | 84937 | 2.42661 | 0.232081 |
| TIGD4 | 201798 | 2.419531 | 0.337323 |

| | | | |
|---|---|---|---|
| ZNF689 | 115509 | 2.400125 | 0.35854 |
| TFDP2 | 7029 | 2.367838 | 0.107992 |
| PPP1CC | 5501 | 2.363951 | 0.00525 |
| EED | 8726 | 2.354056 | 0.61863 |
| UBN1 | 29855 | 2.343187 | 0.154717 |
| SMARCE1 | 6605 | 2.334844 | 0.135732 |
| SETD5 | 55209 | 2.302616 | 0.082448 |
| CDC2 | 983 | 2.249203 | 0.08773 |
| RAI14 | 26064 | 2.23928 | 0.402841 |
| NIPBL | 25836 | 2.193568 | 0.03373 |
| DHX58 | 79132 | 2.188633 | 0.089407 |
| SMAD4 | 4089 | 2.164124 | 1.018242 |
| ZNF69 | 7620 | 2.156333 | 0.255653 |
| HTT | 3064 | 2.14767 | 0.560795 |
| JMJD7-PLA2G4B | 8681 | 2.133011 | 0.688662 |
| TOX4 | 9878 | 2.127174 | 0.832769 |
| NUP62 | 23636 | 2.113548 | 0.002104 |
| NFYB | 4801 | 2.106238 | 1.07298 |
| DFFB | 1677 | 2.086489 | 0.572511 |
| ZIC3 | 7547 | 2.085392 | 0.088481 |
| MZF1 | 7593 | 2.07303 | 0.004678 |
| IL13 | 3596 | 2.043556 | 0.145339 |
| CDC2L2 | 728642 | 2.040397 | 0.002263 |
| MEIS1 | 4211 | 2.037857 | 0.418983 |
| L3MBTL2 | 83746 | 2.035535 | 0.084081 |
| GMEB1 | 10691 | 2.003581 | 0.169052 |
| DPF2 | 5977 | 1.994861 | 0.036312 |
| BPTF | 2186 | 1.988137 | 0.176727 |
| UBP1 | 7342 | 1.985456 | 0.51958 |
| CHD1L | 9557 | 1.952948 | 0.993246 |
| HIST1H1A | 3024 | 1.952727 | 0.179962 |
| HIST1H1C | 3006 | 1.943209 | 0.051822 |
| ZMYM2 | 7750 | 1.929547 | 0.306278 |
| MCM8 | 84515 | 1.927351 | 0.017134 |
| ZNF793 | 390927 | 1.925971 | 0.51161 |
| SUPT20H | 55578 | 1.915236 | 0.266075 |
| SRRM1 | 10250 | 1.913023 | 0.087678 |
| ZNF354C | 30832 | 1.903431 | 0.553552 |
| VEGFA | 7422 | 1.858479 | 0.289521 |
| ZNF423 | 23090 | 1.855819 | 0.152005 |
| ZNF695 | 57116 | 1.849036 | 0.221648 |
| ADORA1 | 134 | 1.847487 | 0.295649 |
| ZNF92 | 168374 | 1.838197 | 0.658696 |
| ACIN1 | 22985 | 1.831692 | 0.773224 |

| | | | |
|---|---|---|---|
| ATG7 | 10533 | 1.799631 | 0.261506 |
| HKR1 | 284459 | 1.783316 | 0.760659 |
| CDC2L1 | 984 | 1.77972 | 0.00845 |
| SIN3B | 23309 | 1.773454 | 0.530537 |
| UBTF | 7343 | 1.771513 | 0.072984 |
| ZIC5 | 85416 | 1.757049 | 0.208026 |
| ZNF432 | 9668 | 1.741552 | 0.489139 |
| MLNR | 2862 | 1.733637 | 0.121295 |
| FLJ34389 | 197259 | 1.727372 | 0.161107 |
| CREB1 | 1385 | 1.720396 | 0.391923 |
| ZNF671 | 79891 | 1.720267 | 0.504045 |
| NDN | 4692 | 1.710066 | 0.096457 |
| TRIB3 | 57761 | 1.708707 | 0.010218 |
| PLK1 | 5347 | 1.697705 | 0.002944 |
| RBKS | 64080 | 1.695117 | 0.402948 |
| GPRC5B | 51704 | 1.685528 | 0.036406 |
| WEE1 | 7465 | 1.681142 | 0.004052 |
| BTAF1 | 9044 | 1.681116 | 0.180609 |
| LSM14A | 26065 | 1.679036 | 0.114269 |
| ARID2 | 196528 | 1.676843 | 1.328193 |
| ZIC1 | 7545 | 1.675335 | 0.193986 |
| RABGAP1L | 9910 | 1.651894 | 0.655652 |
| ZNF27 | 0 | 1.629934 | 0.611814 |
| CSNK2A1 | 1457 | 1.622431 | 0.594907 |
| GON4L | 54856 | 1.619349 | 0.056148 |
| EGR3 | 1960 | 1.617023 | 0.458048 |
| ASCL3 | 56676 | 1.61609 | 0.172599 |
| KLHDC3 | 116138 | 1.615708 | 0.662741 |
| PPARGC1B | 133522 | 1.595375 | 0.187351 |
| AURKA | 6790 | 1.584807 | 0.056916 |
| KLF4 | 9314 | 1.576385 | 0.287166 |
| ZNF668 | 79759 | 1.574289 | 0.312597 |
| KAT2A | 2648 | 1.574077 | 0.563696 |
| PTPN12 | 5782 | 1.565333 | 0.401493 |
| MAPT | 4137 | 1.559559 | 0.054902 |
| UBE2K | 3093 | 1.545062 | 0.945313 |
| TCF7L2 | 6934 | 1.54139 | 0.89919 |
| CNOT1 | 23019 | 1.518457 | 0.00741 |
| SCML4 | 256380 | 1.512091 | 0.14198 |
| IL24 | 11009 | 1.487028 | 0.512251 |
| LOC91461 | 91461 | 1.481985 | 0.004375 |
| KIAA2002 | 79834 | 1.44217 | 0.646311 |
| MGA | 23269 | 1.441326 | 0.277095 |
| AMD1 | 262 | 1.4321 | 0.058278 |

| | | | |
|---|---|---|---|
| TESK1 | 7016 | 1.421965 | 0.442114 |
| PBX4 | 80714 | 1.412995 | 0.261229 |
| HOPX | 84525 | 1.412908 | 0.082224 |
| MTF1 | 4520 | 1.40846 | 0.470873 |
| MC3R | 4159 | 1.383325 | 0.006231 |
| DUSP6 | 1848 | 1.376909 | 0.041479 |
| TACR2 | 6865 | 1.364175 | 0.016701 |
| RBFOX2 | 23543 | 1.278169 | 0.49135 |
| HIRA | 7290 | 1.269069 | 0.148656 |
| CBFA2T3 | 863 | 1.26127 | 0.021579 |
| LTB4R | 1241 | 1.238326 | 0.246107 |
| CRK7 | 51755 | 1.227744 | 0.131756 |
| ZIC2 | 7546 | 1.212511 | 0.14526 |
| RBMX | 27316 | 1.105143 | 0.009703 |
| VHL | 7428 | 1.071208 | 0.045267 |
| WBP11 | 51729 | 1.046851 | 0.004631 |

Appendix 5. List of hits for the bFGF pathway inhibition condition

| Gene Symbol | Entrez Gene ID | Mean Z-score | Mean Cell Viability Score |
|---|---|---|---|
| PIK3R2 | 5296 | 7.619988 | 0.003103 |
| CDC2L2 | 728642 | 5.322273 | 0.006151 |
| PSMC5 | 5705 | 4.994384 | 0.020127 |
| KAT5 | 10524 | 4.862223 | 0.286129 |
| PVR | 5817 | 4.860288 | 0.022833 |
| RBMX | 27316 | 4.284094 | 0.417923 |
| MC3R | 4159 | 3.997967 | 0.015429 |
| EP400 | 57634 | 3.810639 | 0.141473 |
| TRRAP | 8295 | 3.630418 | 0.191103 |
| NUP62 | 23636 | 3.429893 | 0.005683 |
| GPR147 | 64106 | 3.306154 | 0.020687 |
| CIC | 23152 | 3.295524 | 0.62247 |
| SKIL | 6498 | 3.288141 | 0.754046 |
| WBP11 | 51729 | 3.255237 | 0.057564 |
| SETD5 | 55209 | 3.215081 | 0.840927 |
| NFE2 | 4778 | 3.178993 | 0.791428 |
| ILF2 | 3608 | 3.030825 | 0.433237 |
| ZNF48 | 197407 | 2.945386 | 0.725935 |
| PLK1 | 5347 | 2.875995 | 0.00659 |
| PRPF6 | 24148 | 2.812959 | 0.53907 |
| EIF4A3 | 9775 | 2.799675 | 0.045809 |
| PSMC2 | 5701 | 2.738937 | 0.030004 |
| FANCD2 | 2177 | 2.669228 | 0.571213 |
| NFKB2 | 4791 | 2.649079 | 0.595565 |
| ZNF140 | 7699 | 2.645623 | 0.708664 |
| HSF1 | 3297 | 2.645199 | 0.492832 |
| ZBTB12 | 221527 | 2.612779 | 1.025685 |
| ZNF354C | 30832 | 2.522182 | 0.601728 |
| HDAC3 | 8841 | 2.504701 | 0.869307 |
| ILF3 | 3609 | 2.474084 | 0.627961 |
| LSM7 | 51690 | 2.466111 | 0.429191 |
| SFPQ | 6421 | 2.441514 | 0.023494 |
| TFDP2 | 7029 | 2.435719 | 0.625072 |
| GABPA | 2551 | 2.409886 | 0.219587 |
| SNRPE | 6635 | 2.364563 | 0.07831 |
| DMAP1 | 55929 | 2.362587 | 0.148349 |
| LEF1 | 51176 | 2.362138 | 1.053546 |
| TICRR | 90381 | 2.327715 | 0.35001 |
| CDK11A | 728642 | 2.313449 | 0.065421 |
| GMEB1 | 10691 | 2.289803 | 0.724297 |
| SNRPD1 | 6632 | 2.222764 | 0.044145 |
| CCDC6 | 8030 | 2.210449 | 0.843268 |

| | | | |
|---|---|---|---|
| ZBTB46 | 140685 | 2.198587 | 0.651346 |
| DDX18 | 8886 | 2.194596 | 0.366266 |
| KIAA1310 | 55683 | 2.179639 | 0.392174 |
| AURKA | 6790 | 2.177545 | 0.465421 |
| DHX58 | 79132 | 2.173781 | 0.966479 |
| CDC25A | 993 | 2.169595 | 0.357658 |
| IKZF4 | 64375 | 2.15197 | 0.666883 |
| TRIM8 | 81603 | 2.145776 | 0.780171 |
| SNRNP200 | 23020 | 2.137963 | 0.193514 |
| DDX6 | 1656 | 2.130468 | 0.662929 |
| ACTL6A | 86 | 2.124824 | 0.181861 |
| SF3B1 | 23451 | 2.121826 | 0.05737 |
| TGIF1 | 7050 | 2.116383 | 0.836561 |
| DMRTB1 | 63948 | 2.096732 | 0.481029 |
| UBE2L3 | 7332 | 2.067143 | 0.543246 |
| NDN | 4692 | 2.044973 | 0.680173 |
| VPS72 | 6944 | 2.038468 | 0.84245 |
| RFC2 | 5982 | 2.001052 | 0.481757 |
| MCM5 | 4174 | 1.97192 | 0.784626 |
| DPF2 | 5977 | 1.968835 | 0.677575 |
| SAST | 22983 | 1.932438 | 0.773356 |
| MDM4 | 4194 | 1.905858 | 0.485109 |
| DHX16 | 8449 | 1.859229 | 0.125329 |
| ERG | 2078 | 1.852405 | 1.163113 |
| SRRM1 | 10250 | 1.850282 | 0.689112 |
| JMJD1A | 55818 | 1.83454 | 1.051515 |
| ARID2 | 196528 | 1.811834 | 1.170976 |
| RYBP | 23429 | 1.810047 | 0.6106 |
| GPS2 | 2874 | 1.791849 | 0.38294 |
| BAT1 | 7919 | 1.778783 | 0.215005 |
| KMT2C | 58508 | 1.776583 | 0.94451 |
| EPHB2 | 2048 | 1.760731 | 0.488943 |
| ZNF782 | 158431 | 1.752058 | 1.01651 |
| ARID1A | 8289 | 1.7448 | 0.604572 |
| ATF6B | 1388 | 1.743112 | 0.79474 |
| CHAF1A | 10036 | 1.738658 | 0.0624 |
| ETV5 | 2119 | 1.732212 | 0.523448 |
| TFAP2B | 7021 | 1.725075 | 0.9827 |
| CREB3L2 | 64764 | 1.718285 | 0.381564 |
| ZNF22 | 7570 | 1.71809 | 0.746755 |
| JARID2 | 3720 | 1.706206 | 0.207434 |
| NUP153 | 9972 | 1.703549 | 0.100312 |
| CSNK2A1 | 1457 | 1.698509 | 0.907252 |
| MPND | 84954 | 1.676003 | 0.887471 |

| | | | |
|---|---|---|---|
| ACD | 65057 | 1.670423 | 1.067412 |
| SAP30BP | 29115 | 1.660069 | 0.439613 |
| HIST1H2BB | 3018 | 1.640174 | 0.954668 |
| L3MBTL2 | 83746 | 1.635015 | 0.921015 |
| PLK3 | 1263 | 1.626303 | 0.755686 |
| SFRS3 | 6428 | 1.615708 | 0.216354 |
| SETMAR | 6419 | 1.615242 | 0.548172 |
| H2AFZ | 3015 | 1.612695 | 0.801535 |
| MAX | 4149 | 1.610527 | 0.97528 |
| DUSP6 | 1848 | 1.60157 | 0.311926 |
| GMNN | 51053 | 1.59494 | 0.661588 |
| GPR101 | 83550 | 1.591803 | 0.924384 |
| PKNOX2 | 63876 | 1.585563 | 0.82402 |
| CAMTA2 | 23125 | 1.585317 | 0.61444 |
| CXCR6 | 10663 | 1.56801 | 0.457659 |
| MEAF6 | 64769 | 1.563583 | 0.807328 |
| DDX19A | 55308 | 1.548573 | 0.95648 |
| HKR1 | 284459 | 1.524105 | 1.105771 |
| ZFP36L1 | 677 | 1.520691 | 0.572518 |
| SCML4 | 256380 | 1.515937 | 0.952742 |
| MYST1 | 84148 | 1.506528 | 0.158914 |
| TAX1BP3 | 30851 | 1.500625 | 0.870139 |
| KLHL41 | 10324 | 1.495996 | 0.682487 |
| CHD3 | 1107 | 1.481632 | 0.385722 |
| ZNF619 | 285267 | 1.481501 | 0.530358 |
| NFXL1 | 152518 | 1.463693 | 0.874714 |
| PYGO1 | 26108 | 1.453645 | 0.130134 |
| HMGA1 | 3159 | 1.451076 | 0.499978 |
| TOX4 | 9878 | 1.445928 | 0.941139 |
| NTRK2 | 4915 | 1.440484 | 0.20639 |
| ZFAT | 57623 | 1.435305 | 0.881581 |
| NPY | 4852 | 1.434608 | 0.48115 |
| NEK2 | 4751 | 1.432953 | 1.086315 |
| BIRC5 | 332 | 1.422219 | 0.117054 |
| MARK2 | 2011 | 1.408215 | 0.656586 |
| SNIP1 | 79753 | 1.405112 | 0.094095 |
| JDP2 | 122953 | 1.400855 | 0.585662 |
| ADORA1 | 134 | 1.396769 | 0.520874 |
| ZNF574 | 64763 | 1.392161 | 0.235377 |
| MASS1 | 84059 | 1.38833 | 0.753752 |
| ASCC2 | 84164 | 1.367399 | 0.930258 |
| SNRNP27 | 11017 | 1.354516 | 0.73008 |
| PAX3 | 5077 | 1.339302 | 1.034505 |
| CSNK2B | 1460 | 1.301955 | 0.716346 |

| | | | |
|---|---|---|---|
| ESPL1 | 9700 | 1.299969 | 0.059506 |
| TLK1 | 9874 | 1.292012 | 1.143709 |
| MORF4 | 10934 | 1.281612 | 0.762226 |
| RBBP4 | 5928 | 1.277126 | 0.198751 |
| MPP3 | 4356 | 1.265657 | 0.190298 |
| SNRPB | 6628 | 1.216005 | 0.085709 |
| RPS6KA1 | 6195 | 1.013957 | 0.293602 |
| GPRC6A | 222545 | 0.976607 | 0.36602 |

Appendix 6. List of hits for the + Retinoic acid condition

| Gene Symbol | Entrez Gene ID | Mean Z-score | Mean Cell Viability Score |
|---|---|---|---|
| TAF1 | 6872 | 6.608557 | 0.672588 |
| SKIL | 6498 | 5.961297 | 0.985128 |
| RBMX | 27316 | 5.740771 | 0.681441 |
| WBP11 | 51729 | 4.932392 | 0.131931 |
| NFYC | 4802 | 4.762958 | 0.766345 |
| SETD5 | 55209 | 4.694522 | 0.879323 |
| MLNR | 2862 | 4.679523 | 1.039359 |
| RHO | 6010 | 4.512082 | 1.018085 |
| ASF1A | 25842 | 4.375268 | 0.848312 |
| ZNF707 | 286075 | 4.075265 | 0.985303 |
| FFAR1 | 2864 | 4.026534 | 1.084124 |
| CHAF1B | 8208 | 3.962761 | 0.506476 |
| CDKL3 | 51265 | 3.904794 | 1.142033 |
| IRAK2 | 3656 | 3.818146 | 1.036585 |
| BRD4 | 23476 | 3.726535 | 0.778985 |
| CXADR | 1525 | 3.680633 | 0.893693 |
| NEUROG2 | 63973 | 3.640891 | 1.194415 |
| CDK11A | 728642 | 3.599464 | 0.144949 |
| ZBTB11 | 27107 | 3.537131 | 0.697197 |
| KLHL14 | 57565 | 3.53107 | 0.941069 |
| PPP2R1A | 5518 | 3.521348 | 1.183597 |
| PTPRT | 11122 | 3.513873 | 1.023442 |
| NFYB | 4801 | 3.490615 | 0.725104 |
| ZNF140 | 7699 | 3.458326 | 0.840491 |
| DFFB | 1677 | 3.456652 | 1.114096 |
| FRAP1 | 2475 | 3.433615 | 0.954868 |
| CCDC6 | 8030 | 3.383194 | 0.969125 |
| CEBPA | 1050 | 3.35348 | 0.971558 |
| CIB2 | 10518 | 3.290305 | 0.679691 |
| PLRG1 | 5356 | 3.284123 | 0.371918 |
| TSG101 | 7251 | 3.279604 | 0.178139 |
| CHD8 | 57680 | 3.192836 | 0.974168 |
| GTF2IRD2B | 389524 | 3.180282 | 0.947673 |
| MARK2 | 2011 | 3.155987 | 1.159295 |
| MC3R | 4159 | 3.152512 | 0.341665 |
| WBP11 | 51729 | 3.146758 | 0.083127 |
| ETV5 | 2119 | 3.129495 | 0.881457 |
| ZBTB12 | 221527 | 3.043658 | 0.997157 |
| PKN2 | 5586 | 3.036973 | 1.10247 |
| GABPB1 | 2553 | 2.993346 | 0.332446 |
| ZNF317 | 57693 | 2.982855 | 0.861902 |
| ZNF192 | 7745 | 2.976632 | 0.934138 |

| | | | |
|---|---|---|---|
| EP400 | 57634 | 2.961968 | 0.421063 |
| EFNA3 | 1944 | 2.953257 | 1.245154 |
| ASF1B | 55723 | 2.943198 | 0.725666 |
| EPHB2 | 2048 | 2.92501 | 0.843684 |
| MYF5 | 4617 | 2.909644 | 0.989357 |
| SETD2 | 29072 | 2.899547 | 1.156641 |
| AKT1 | 207 | 2.888259 | 1.010932 |
| ATF7 | 11016 | 2.885865 | 0.89355 |
| PPFIA2 | 8499 | 2.813446 | 0.997352 |
| KIAA2002 | 79834 | 2.811388 | 1.101227 |
| NPAT | 4863 | 2.789561 | 0.55651 |
| CCR6 | 1235 | 2.772246 | 0.94092 |
| RNPS1 | 10921 | 2.767186 | 0.384442 |
| PDK1 | 5163 | 2.760405 | 1.083311 |
| NKAP | 79576 | 2.759491 | 0.126416 |
| NMUR1 | 10316 | 2.758763 | 0.910521 |
| TBX15 | 6913 | 2.758427 | 0.921128 |
| ANKZF1 | 55139 | 2.75236 | 0.982258 |
| MERTK | 10461 | 2.746322 | 0.999175 |
| PVR | 5817 | 2.734018 | 0.536562 |
| FRK | 2444 | 2.733287 | 1.11919 |
| ZNF441 | 126068 | 2.715834 | 1.076417 |
| DHX38 | 9785 | 2.706779 | 0.292804 |
| ACD | 65057 | 2.696886 | 0.938817 |
| BCL11A | 53335 | 2.660337 | 1.044709 |
| CHD3 | 1107 | 2.656699 | 0.343878 |
| HIST1H1A | 3024 | 2.651479 | 0.969417 |
| TFAP2B | 7021 | 2.649261 | 0.969886 |
| DDX1 | 1653 | 2.637643 | 1.158055 |
| SNIP1 | 79753 | 2.618139 | 0.23333 |
| TCEB2 | 6923 | 2.610865 | 0.496504 |
| ASCC3 | 10973 | 2.608188 | 1.081167 |
| NEK1 | 4750 | 2.601714 | 1.145532 |
| GPR101 | 83550 | 2.586515 | 1.032186 |
| TAF2 | 6873 | 2.586196 | 1.163686 |
| ZNF226 | 7769 | 2.578861 | 0.848572 |
| ZNF354C | 30832 | 2.573079 | 1.062939 |
| POLR2I | 5438 | 2.567482 | 0.212402 |
| MAP2K3 | 5606 | 2.566121 | 1.193005 |
| PADI4 | 23569 | 2.559142 | 1.033419 |
| GPR18 | 2841 | 2.557716 | 0.982015 |
| ING5 | 84289 | 2.556089 | 1.043189 |
| MZF1 | 7593 | 2.555891 | 0.813358 |
| NUDT21 | 11051 | 2.554061 | 0.586026 |

| | | | |
|---|---|---|---|
| PTPN12 | 5782 | 2.534503 | 1.074874 |
| SCAP1 | 8631 | 2.529485 | 1.252792 |
| PLK4 | 10733 | 2.504691 | 0.834073 |
| PON1 | 5444 | 2.502199 | 0.972084 |
| CHAF1A | 10036 | 2.488923 | 0.539164 |
| GZF1 | 64412 | 2.460614 | 1.039601 |
| MED26 | 9441 | 2.455073 | 0.321475 |
| TRRAP | 8295 | 2.452951 | 0.372737 |
| SFMBT1 | 51460 | 2.428209 | 1.06154 |
| ZFP57 | 346171 | 2.426377 | 0.998124 |
| ZNF79 | 7633 | 2.424132 | 0.606429 |
| ZSCAN32 | 54925 | 2.423906 | 1.095694 |
| DHX9 | 1660 | 2.418265 | 1.024849 |
| ZNF619 | 285267 | 2.412454 | 0.976275 |
| HMGA1 | 3159 | 2.412244 | 0.949968 |
| ILF3 | 3609 | 2.40397 | 1.044045 |
| ADORA2A | 135 | 2.396333 | 0.958711 |
| SIN3A | 25942 | 2.389798 | 0.519032 |
| SP2 | 6668 | 2.345387 | 0.972043 |
| USE1 | 55850 | 2.335257 | 0.548172 |
| PPP1R13B | 23368 | 2.320487 | 1.059457 |
| TCF23 | 150921 | 2.315445 | 1.131012 |
| ZNF532 | 55205 | 2.314937 | 1.102752 |
| ATXN2 | 6311 | 2.31355 | 1.121691 |
| PMVK | 10654 | 2.312961 | 1.165157 |
| TRIB3 | 57761 | 2.307221 | 0.431334 |
| URKL1 | 54963 | 2.305345 | 1.146269 |
| CIAO1 | 9391 | 2.285529 | 0.834677 |
| PSMC2 | 5701 | 2.285375 | 0.059163 |
| HDAC6 | 10013 | 2.269314 | 0.947187 |
| KHK | 3795 | 2.263077 | 1.212515 |
| DDX28 | 55794 | 2.262938 | 0.999503 |
| ZNF207 | 7756 | 2.234563 | 0.349732 |
| MRGPRX4 | 117196 | 2.232668 | 1.044715 |
| DUSP4 | 1846 | 2.224797 | 1.062662 |
| GABPA | 2551 | 2.214518 | 0.327658 |
| ZNF2 | 7549 | 2.214452 | 0.957617 |
| ZNF154 | 7710 | 2.209309 | 0.877216 |
| GTF2H1 | 2965 | 2.207691 | 0.927566 |
| TUBA1B | 10376 | 2.201631 | 0.358028 |
| SNRPN | 6638 | 2.199508 | 0.965584 |
| PHB2 | 11331 | 2.184548 | 0.379605 |
| CIC | 23152 | 2.181712 | 0.838143 |
| SFRS6 | 6431 | 2.166749 | 0.772448 |

| | | | |
|---|---|---|---|
| ZNF668 | 79759 | 2.165083 | 1.152706 |
| SMARCD2 | 6603 | 2.163007 | 0.995018 |
| DHX58 | 79132 | 2.162601 | 1.131214 |
| GUK1 | 2987 | 2.157346 | 0.480052 |
| SOX7 | 83595 | 2.156989 | 1.02067 |
| AK5 | 26289 | 2.156125 | 0.933279 |
| PHF7 | 51533 | 2.154314 | 1.012255 |
| CAMTA2 | 23125 | 2.152327 | 0.954164 |
| POLR2C | 5432 | 2.14222 | 0.1184 |
| ZNF467 | 168544 | 2.128767 | 0.847746 |
| DHX33 | 56919 | 2.127794 | 0.827236 |
| AGTR1 | 185 | 2.125032 | 1.076853 |
| GPR44 | 11251 | 2.121728 | 0.857756 |
| CCRK | 23552 | 2.09963 | 1.225904 |
| CNOT2 | 4848 | 2.094046 | 0.748952 |
| RXRA | 6256 | 2.091618 | 1.049529 |
| KDM4D | 55693 | 2.090811 | 1.052768 |
| MED31 | 51003 | 2.087189 | 0.732073 |
| TFDP2 | 7029 | 2.086706 | 0.754933 |
| DHX8 | 1659 | 2.085239 | 0.113291 |
| MMP9 | 4318 | 2.083827 | 0.318174 |
| CKMT2 | 1160 | 2.08025 | 1.200875 |
| ZNF212 | 7988 | 2.079028 | 1.033086 |
| ATF6B | 1388 | 2.077408 | 0.909917 |
| VEGFA | 7422 | 2.076966 | 1.047638 |
| GAK | 2580 | 2.072499 | 0.562777 |
| PRKCL2 | 5586 | 2.070978 | 0.970255 |
| HMGA2 | 8091 | 2.06746 | 0.848229 |
| BRD4 | 23476 | 2.06372 | 0.934712 |
| EEF1A1 | 1915 | 2.063124 | 0.220679 |
| ARID1A | 8289 | 2.062662 | 0.907839 |
| CALCOCO1 | 57658 | 2.062361 | 1.044775 |
| EIF5B | 9669 | 2.058996 | 0.72722 |
| NFATC3 | 4775 | 2.058429 | 1.077119 |
| HMGXB3 | 22993 | 2.057697 | 0.296634 |
| KUB3 | 91419 | 2.057408 | 1.226778 |
| PRPF6 | 24148 | 2.055437 | 0.507785 |
| ANKRD6 | 22881 | 2.055124 | 0.953723 |
| CDK5 | 1020 | 2.049546 | 1.010251 |
| SMC2 | 10592 | 2.048696 | 0.69066 |
| TTBK1 | 84630 | 2.043131 | 1.179032 |
| ZNF782 | 158431 | 2.040877 | 1.025315 |
| RFXAP | 5994 | 2.036599 | 1.068578 |
| JAK1 | 3716 | 2.029144 | 1.082776 |

| | | | |
|---|---|---|---|
| ZFP42 | 132625 | 2.019272 | 1.049466 |
| RUVBL2 | 10856 | 2.016043 | 0.160336 |
| ZBTB6 | 10773 | 2.010123 | 0.893377 |
| RAP1GAP | 5909 | 2.00957 | 0.802694 |
| CDKN2C | 1031 | 2.008508 | 1.284162 |
| DNMT1 | 1786 | 2.004839 | 1.043459 |
| SNRNP200 | 23020 | 2.000804 | 0.137003 |
| POLR3E | 55718 | 1.994337 | 0.637038 |
| HMGN1 | 3150 | 1.991516 | 0.799106 |
| GTF2H1 | 2965 | 1.988693 | 0.95328 |
| TAL2 | 6887 | 1.977656 | 1.029222 |
| NKX3-2 | 579 | 1.96051 | 1.031935 |
| H1F0 | 3005 | 1.959344 | 0.986053 |
| KCNH8 | 131096 | 1.951063 | 1.046875 |
| GTF2H4 | 2968 | 1.948488 | 0.984776 |
| PAK7 | 57144 | 1.937819 | 0.964427 |
| ZFAND3 | 60685 | 1.937552 | 0.840348 |
| EVX2 | 344191 | 1.935099 | 0.912368 |
| BUD31 | 8896 | 1.932619 | 0.098543 |
| MAP3K10 | 4294 | 1.92729 | 1.116724 |
| CDK9 | 1025 | 1.92487 | 0.625604 |
| FHL1 | 2273 | 1.923802 | 0.766709 |
| POLR2H | 5437 | 1.918986 | 0.180686 |
| DIP2B | 57609 | 1.918699 | 1.029573 |
| HDAC11 | 79885 | 1.910072 | 1.120867 |
| CBFA2T3 | 863 | 1.899169 | 0.387933 |
| POLR2L | 5441 | 1.897246 | 0.340649 |
| POLR2D | 5433 | 1.896288 | 0.070567 |
| CDC42BPA | 8476 | 1.875172 | 1.15843 |
| ESPL1 | 9700 | 1.873263 | 0.375823 |
| PHF19 | 26147 | 1.87238 | 0.23382 |
| SETD8 | 387893 | 1.868008 | 0.183135 |
| PYGO1 | 26108 | 1.867475 | 0.472932 |
| TAF1 | 6872 | 1.86728 | 0.857286 |
| NFATC2 | 4773 | 1.865802 | 0.654839 |
| GPR157 | 80045 | 1.863016 | 1.03157 |
| KIAA1310 | 55683 | 1.858318 | 0.305745 |
| RNF113A | 7737 | 1.850884 | 1.114593 |
| ZFP36 | 7538 | 1.850168 | 0.828363 |
| ZNF215 | 7762 | 1.84555 | 1.069404 |
| SUGP1 | 57794 | 1.844163 | 1.11175 |
| YY1 | 7528 | 1.841999 | 1.000052 |
| SMARCA5 | 8467 | 1.839925 | 1.137309 |
| MAF1 | 84232 | 1.833559 | 1.225161 |

| | | | |
|---|---|---|---|
| SNRPD1 | 6632 | 1.830657 | 0.254661 |
| SAP18 | 10284 | 1.830294 | 0.885024 |
| AEBP2 | 121536 | 1.818966 | 1.069476 |
| TA-PP2C | 160760 | 1.81555 | 1.165046 |
| PIK3CD | 5293 | 1.813079 | 1.219739 |
| UBTF | 7343 | 1.812966 | 0.841507 |
| PTGS2 | 5743 | 1.811386 | 1.049819 |
| TESK1 | 7016 | 1.809663 | 1.022506 |
| PIK3C2A | 5286 | 1.809448 | 1.083858 |
| GPRC5C | 55890 | 1.796335 | 1.13993 |
| EVI1 | 2122 | 1.795074 | 1.093359 |
| CASR | 846 | 1.790111 | 0.999492 |
| PIAS1 | 8554 | 1.787799 | 0.947417 |
| ZNF396 | 252884 | 1.785548 | 0.232572 |
| SKIV2L | 6499 | 1.782419 | 1.138005 |
| HIST1H2AH | 85235 | 1.771539 | 0.959616 |
| SYK | 6850 | 1.769253 | 1.062341 |
| ZNF75D | 7626 | 1.768371 | 0.937129 |
| ZNF259 | 8882 | 1.767814 | 0.990059 |
| ZFP112 | 7771 | 1.761642 | 1.087045 |
| SSX2 | 6757 | 1.760966 | 0.961949 |
| ZNF709 | 163051 | 1.758862 | 0.922181 |
| DGUOK | 1716 | 1.752543 | 1.004862 |
| ZBTB46 | 140685 | 1.751421 | 0.94292 |
| UBE2K | 3093 | 1.748684 | 1.102903 |
| AGAP3 | 116988 | 1.747745 | 1.077434 |
| PPAP2B | 8613 | 1.747704 | 1.045103 |
| MPP3 | 4356 | 1.747008 | 0.822777 |
| HELZ2 | 85441 | 1.746732 | 1.094994 |
| PKMYT1 | 9088 | 1.746249 | 1.048993 |
| VHL | 7428 | 1.731148 | 0.984836 |
| SNURF | 0 | 1.730748 | 1.044642 |
| TUBA1C | 84790 | 1.720925 | 1.096472 |
| MMS19 | 64210 | 1.717978 | 1.128329 |
| GPR65 | 8477 | 1.716955 | 1.019322 |
| HELT | 391723 | 1.714972 | 0.892647 |
| DLG2 | 1740 | 1.71305 | 1.061361 |
| MARK4 | 57787 | 1.710763 | 0.894343 |
| PRKCB | 5579 | 1.710599 | 1.092583 |
| NFXL1 | 152518 | 1.70753 | 0.985756 |
| RFC2 | 5982 | 1.707029 | 0.78658 |
| MPND | 84954 | 1.705249 | 1.076496 |
| KDM5B | 10765 | 1.695557 | 0.985859 |
| POLR1E | 64425 | 1.68921 | 1.019065 |

| | | | |
|---|---|---|---|
| HRH2 | 3274 | 1.677022 | 0.998784 |
| STAT1 | 6772 | 1.670454 | 1.025621 |
| MAPKAPK3 | 7867 | 1.668271 | 1.278873 |
| CCDC67 | 159989 | 1.663355 | 1.103498 |
| POLR2F | 5435 | 1.660021 | 0.102977 |
| NFYA | 4800 | 1.658947 | 0.402536 |
| L3MBTL | 26013 | 1.654855 | 0.980413 |
| POU4F3 | 5459 | 1.650159 | 0.729711 |
| ZNF772 | 400720 | 1.642743 | 0.82177 |
| CTNNBL1 | 56259 | 1.637794 | 0.883979 |
| TRIM29 | 23650 | 1.634918 | 1.216643 |
| PROP1 | 5626 | 1.632856 | 0.46534 |
| KAT2B | 8850 | 1.625278 | 1.10763 |
| ZBED2 | 79413 | 1.624771 | 1.005634 |
| GPR141 | 353345 | 1.620593 | 1.15372 |
| ZNF200 | 7752 | 1.616548 | 0.925347 |
| ZNF227 | 7770 | 1.61198 | 1.111626 |
| ZNF485 | 220992 | 1.603201 | 1.144446 |
| GCC2 | 9648 | 1.599156 | 1.024341 |
| SMAD3 | 4088 | 1.596292 | 0.813216 |
| MAFA | 389692 | 1.587636 | 1.099492 |
| ZNF239 | 8187 | 1.586478 | 1.081891 |
| ZNF416 | 55659 | 1.579751 | 0.882693 |
| MEF2B | 0 | 1.574541 | 0.890167 |
| EYA4 | 2070 | 1.574434 | 1.081472 |
| NUDT2 | 318 | 1.572086 | 1.089536 |
| TRAF6 | 7189 | 1.565043 | 1.017594 |
| DDX41 | 51428 | 1.549409 | 0.19565 |
| SOX9 | 6662 | 1.544943 | 0.970958 |
| SUPT20H | 55578 | 1.541065 | 1.142956 |
| POU4F1 | 5457 | 1.539879 | 1.094733 |
| TRAF5 | 7188 | 1.53854 | 1.005842 |
| SCAND2 | 54581 | 1.522487 | 1.04344 |
| MXD3 | 83463 | 1.522485 | 1.012511 |
| ZNF266 | 10781 | 1.509767 | 1.052183 |
| ZSCAN2 | 54993 | 1.509006 | 1.10666 |
| ZNF860 | 344787 | 1.504266 | 1.144657 |
| ZNF273 | 10793 | 1.489337 | 0.963982 |
| DUSP12 | 11266 | 1.488376 | 0.944269 |
| MTA1 | 9112 | 1.488106 | 1.019468 |
| TIMM50 | 92609 | 1.484426 | 1.058261 |
| HINFP | 25988 | 1.459763 | 0.859232 |
| DUSP6 | 1848 | 1.456161 | 0.883148 |
| ZNF776 | 284309 | 1.438796 | 1.034583 |

| | | | |
|---|---|---|---|
| PNN | 5411 | 1.435147 | 0.776129 |
| FOXR1 | 283150 | 1.425075 | 1.143995 |
| TBL1XR1 | 79718 | 1.421482 | 1.082033 |
| POU1F1 | 5449 | 1.421405 | 1.049736 |
| PPP3R2 | 5535 | 1.380098 | 1.104494 |
| SMAD4 | 4089 | 1.368683 | 1.223986 |
| SPEN | 23013 | 1.348423 | 0.884493 |
| E2F5 | 1875 | 1.339401 | 1.011681 |
| SLC2A4RG | 56731 | 1.320798 | 0.894955 |
| TRMT11 | 60487 | 1.316348 | 0.967644 |
| C9ORF67 | 84814 | 1.316073 | 0.97041 |
| ZIC3 | 7547 | 1.298578 | 0.895853 |
| NUP153 | 9972 | 1.294525 | 0.574038 |
| TAF8 | 129685 | 1.279132 | 0.967674 |
| PPP2CZ | 333926 | 1.277532 | 1.164914 |
| TAF9 | 6880 | 1.267752 | 0.93178 |
| ARNT | 405 | 1.203382 | 1.000856 |
| PPAPDC1 | 196051 | 1.174661 | 1.091659 |
| PPM1F | 9647 | 1.162985 | 1.150928 |
| CASP2 | 835 | 1.121582 | 1.061862 |
| XBP1 | 7494 | 1.109072 | 0.896269 |
| EYA3 | 2140 | 1.063191 | 0.943605 |
| STK10 | 6793 | 1.022171 | 1.045349 |