<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Deterministic Restriction on Pluripotent State Dissolution by Cell-Cycle Pathways</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Gonzales, Kevin Andrew Uy; Liang, Hongqing; Lim, Yee-Siang; Chan, Yun-Shen; Yeo, Jia-Chi; Tan, Cheng-Peow; Gao, Bin; Le, Beilin; Tan, Zi-Ying; Low, Kok-Yao; Liou, Yih-Cherng; Bard, Frederic; Ng, Huck-Hui</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>Gonzales, K., Liang, H., Lim, Y.-S., Chan, Y.-S., Yeo, J.-C., Tan, C.-P., et al. (2015). Deterministic Restriction on Pluripotent State Dissolution by Cell-Cycle Pathways. Cell, 162(3), 564-579.</td>
</tr>
<tr>
<td><strong>Date</strong></td>
<td>2015</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10220/38785">http://hdl.handle.net/10220/38785</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>© 2015 Elsevier. This is the author created version of a work that has been peer reviewed and accepted for publication by Cell, Elsevier. It incorporates referee’s comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at: [<a href="http://dx.doi.org/10.1016/j.cell.2015.07.001">http://dx.doi.org/10.1016/j.cell.2015.07.001</a>].</td>
</tr>
</tbody>
</table>
Deterministic restriction on pluripotent state dissolution by cell cycle pathways

Kevin Andrew Uy GONZALES\textsuperscript{1,2,*}, Hongqing LIANG\textsuperscript{1,*}, Yee-Siang LIM\textsuperscript{1}, Yun-Shen CHAN\textsuperscript{1}, Jia-Chi YEO\textsuperscript{1,6}, Cheng-Peow TAN\textsuperscript{1}, Bin GAO\textsuperscript{1}, Beilin LE\textsuperscript{1,3}, Zi-Ying TAN\textsuperscript{1,5}, Kok-Yao LOW\textsuperscript{6}, Yih-Cherng LIOU\textsuperscript{2,5}, Frederic BARD\textsuperscript{3,4} and Huck-Hui NG\textsuperscript{1,3,5,6,#}

\textsuperscript{1}Stem Cell and Regenerative Biology, Genome Institute of Singapore, 60 Biopolis St, Singapore 138672
\textsuperscript{2}NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore, 28 Medical Drive, Singapore 117456
\textsuperscript{3}Dept of Biochemistry, National University of Singapore, 8 Medical Drive, Singapore 117597
\textsuperscript{4}Institute of Molecular and Cellular Biology, 61 Biopolis Drive, Singapore 138673
\textsuperscript{5}Dept of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117597
\textsuperscript{6}School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 639798

*Corresponding author:
Address: Genome Institute of Singapore, 60 Biopolis St, Singapore 138672
Tel. #: +65 68088145
E-mail: nghh@gis.a-star.edu.sg

\*These authors contributed equally to this work.
Abstract

During differentiation, human embryonic stem cells (hESCs) shut down the regulatory network conferring pluripotency in a process we designated as pluripotent state dissolution (PSD). In a high-throughput RNAi screen using an inclusive set of differentiation conditions, we identify centrally important and context-dependent processes regulating PSD in hESCs, including histone acetylation, chromatin remodelling, RNA splicing and signalling pathways. Strikingly, we detected a strong and specific enrichment of cell cycle genes involved in DNA replication and G2 phase progression. Genetic and chemical perturbation studies demonstrate that the S and G2 phases attenuate PSD as they possess an intrinsic propensity towards the pluripotent state that is independent of the G1 phase. Our data thus functionally establishes that pluripotency control is hardwired to the cell cycle machinery, where S and G2 phase-specific pathways deterministically restrict PSD while the absence of such pathways in the G1 phase potentially permits the initiation of differentiation.
Introduction

The human pluripotent stem cell state is facilitated by an intricate regulatory network chiefly controlled by master transcription factors (Boyer et al., 2005). These master regulators form multiple regulatory connections with other transcription factors, epigenetic modifiers, signal transduction pathways, non-coding RNAs and other regulators that together maintain self-renewal and pluripotency (Ng and Surani, 2011; Young, 2011). Sustenance of this internal regulatory network is dependent on external cues from the cell culture environment. Human embryonic stem cells (hESCs) mainly rely on the basic fibroblast growth factor (bFGF) and Activin/transforming growth factor beta (TGFβ) pathways for self-renewal (Beattie et al., 2005; Xu et al., 2005). Withdrawal of these signalling pathways causes the shutting down of the complex hESC regulatory network; we call this process pluripotent state dissolution (PSD). Since many applications of hESCs require their complete and efficient differentiation, it is necessary to obtain detailed knowledge of how the hESC regulatory network is dissolved during differentiation. However, as the maintenance of the pluripotent state requires multiple interactions between regulatory pathways and factors, the dissolution of such a state is likely to be an equally complex process, with multiple routes from which it can be enforced. While several studies have identified regulators of PSD in mouse embryonic stem cells (mESCs) (Betschinger et al., 2013; Guo et al., 2011), there is a lack of knowledge about how PSD is regulated in hESCs.

Systematic studies like high-throughput functional genomics assays have greatly advanced the knowledge about the regulatory networks of ESCs. The unbiased nature of functional genomics makes it a powerful discovery tool for identification of key protein complexes and pathways by detecting multiple crucial hits from the same pathway or complex. However, for
hESCs, most studies rely on expanding the hESC regulatory network using previously known factors, and to date, only one arrayed high-throughput functional genomics study has been performed in hESCs (Chia et al., 2010). Thus, there is a lack of knowledge in the molecular understanding of the regulatory network that governs PSD of hESC.

To address this deficiency, we undertook a large-scale high-throughput RNAi screen in differentiating hESCs for the de novo identification of the molecular pathways regulating PSD. To be comprehensive in dissecting PSD, we probed into 5 differentiation conditions to discover both context-dependent and universal gatekeepers. Strikingly, we found a strong enrichment of cell cycle hits, specifically clustering in the S and G2 phases but not in other phases of the cell cycle. Genetic and chemical manipulations of cell cycle progression established an intrinsic propensity towards pluripotency maintenance in the S and G2 phases, enacted by the ATM/ATR-mediated replication checkpoint and Cyclin B1 pathways. Our study reveals a link that hardwires the hESC pluripotency network to the cell cycle machinery, where S and G2 phase-specific pathways deterministically restrict PSD and the absence of such pathways at the G1 phase could permit PSD.
Results

A high-throughput RNAi screen in multiple conditions identifies genes important for PSD in hESCs

The process of ESC differentiation can be conceptually subdivided into two major steps: PSD where the transcriptional network of pluripotency is shut down, and lineage specification where a new transcriptional program corresponding to a specific somatic lineage is assembled. While regulation of lineage specification could greatly differ between different lineages, PSD is an early event for the ESC differentiation process. To obtain evidence that PSD is distinct from lineage specification, we looked at the expression kinetics of pluripotency and lineage-specific genes upon withdrawal of the self-renewal factors bFGF and TGFβ. 48 hours upon induction of differentiation, the downregulation of many hESC-specific genes (Assou et al., 2007) has begun (Figure S1A), but the upregulation of lineage-specific factor expression is only evident at 96 hours (Figure S1B). Examination of gene expression of single cells showed similar results (Figure S1C). Thus, these indicate that PSD occurs immediately at the onset of differentiation, distinct from lineage specification (Figure S1D). Because NANOG is significantly decreased at both early and late time points (Figure S1A), we assigned NANOG downregulation as a marker for PSD and created a NANOG-GFP hESC line as a reporter of PSD (Figure S1E-H).

To achieve robust and unbiased identification of universal and specific factors required for PSD, we carried out a high-throughput RNAi screen under five differentiation conditions (Figure 1A). In the first condition, we removed the two important cytokines, bFGF and TGFβ, from hESC medium (Beattie et al., 2005; Xu et al., 2005). In the second to fourth conditions, we individually perturbed signalling of the TGFβ pathway as well as the mitogen-
activated protein kinase kinase (MEK) and phosphoinositol-3-kinase (PI3K) branches of the bFGF pathway. In the fifth condition, we introduced retinoic acid (RA), a potent inducer of differentiation. All conditions led to PSD as indicated by the efficient loss of hESC identity (Figure S1I-K). Cells were transfected with siRNAs 24 hours before induction of differentiation. The degree of preservation of hESC identity was then measured through the average $NANOG$-GFP fluorescence intensity per cell, as hESCs depleted of PSD effectors are expected to retain higher GFP signals. We screened a total of 4,558 genes in triplicate for all five conditions, summing up to 68,370 data points (Table S1).

We first ensured the quality of the screen by checking for intra-plate layout effects (Figure 1B, S2A), proper inter-plate alignment (Figure S2B) and good correlation between replicates (Figure 1C, S2C). Genes that reproducibly scored above noise ($z>1.25$ or $z>1.5$ in at least 2 replicates) were regarded as hits (Figure 1D). Hits obtained were observed to maintain a visible GFP signal (Figure 1E). We detected no cell number bias in hit selection (Figure S2D), eliminating the possibility that housekeeping genes could be misidentified as hits. We further conducted counter-screens using a hESC line harbouring an $ACTIN$-GFP reporter (Table S2), and found no overlap with the $NANOG$-GFP hits (Figure S2E). Collectively, these results ascertain that the RNAi screen identified genes important for PSD rigorously and reliably.

**Informatics analyses reveal pathways that regulate PSD in a context-specific manner**

The screen was designed to enable the identification of both context-dependent and universally important processes that regulate PSD (Figure 2A). We examined the hits from the five distinct conditions using protein interaction network (Figure 2B), Reactome (Table S3) and gene ontology (Table S4) analyses to look for context-dependent effectors of PSD.
First, we observed that the hits included factors that are associated with the primary pathways transducing the distinct initial differentiation cues. These include repressors SKI and SKIL for TGFβ inhibition (Figure 2B), ERK2 inactivator DUSP6 for MEK inhibition (Table S1), and RA receptor RXRA and transcriptional complexes RNA polymerase II and TFIID/Mediator for +RA (Figure 2B, Table S1). This is reassuring because perturbation of these primary pathways is expected to nullify the differentiation-inducing effect stemming from the same pathway, demonstrating the robustness of the high-throughput screening assay.

Interestingly, the RNA splicing machinery seems to play a role in PSD, especially during RA addition and MEK inhibition (Table S3, S4). This is notable because although multiple RNA splicing factors have been reported to control the pluripotent state of hESCs (Gabut et al., 2011; Lu et al., 2013), it was unknown hitherto whether they also regulate its dissolution. RNA splicing has been shown in other cell types to interact with the MEK and RA pathways (Shilo et al., 2014; Wang et al., 2014). Similar mechanisms might be in place in hESCs, explaining why RNA splicing could affect PSD triggered specifically by MEK inhibition or RA introduction. Our study therefore opens the door for studying the crosstalk between the splicing machinery, PSD, and these developmental pathways.

Besides RNA splicing, studying the functional genetics of PSD further revealed novel nodes of contact between signalling pathways and pluripotency. For example, we observed an enrichment of NuRD complex members in conditions where bFGF-MEK signalling is abolished (Figure 2B, Table S4). The NuRD complex 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. On the other hand, a strong enrichment of histone demethylases, specifically those targeting histone 3 lysine 4 (H3K4), was observed in the PI3K pathway inhibition condition (Figure 2B, Table S3, S4). This points towards an epigenetic link between PI3K pathway and the pluripotency network through the maintenance of activating H3K4 methylation marks.

Finally, we noticed that developmental pathways play highly context-dependent roles in PSD. For instance, Wnt signalling-associated factors, particularly those regulating β-catenin-mediated transcription, were enriched in the hits during inactive bFGF-MEK signalling (Figure 2B, Table S3, S4). This is in agreement with the identified cross-regulation between these two pathways (Ding et al., 2005; Singh et al., 2012) and the pro-differentiation role of nuclear β-catenin in hESCs (Davidson et al., 2012; Dravid et al., 2005). On the other hand, negative regulators of the PI3K pathway are enriched in the -bFGF, -TGFβ condition (Table S3), while positive regulators are enriched in the +RA condition (Figure 2B, Table S3). The PI3K pathway thus seems to help uphold pluripotency in hESCs (Figure S1I-K) in coordination with both bFGF and TGFβ pathways (Singh et al., 2012), but promotes differentiation instead when acting downstream of RA. These emphasize the importance of studying the role of developmental pathways in the proper context.

Thus, separate analyses of the various screening conditions demonstrate the robustness of our screen results, in addition to identifying context-dependent processes that are crucial for PSD.

Combined analysis uncovers epigenetic mechanisms that universally regulate PSD upon withdrawal of self-renewal signals.
Hierarchical clustering of the results from the five conditions revealed that the +RA condition is excluded from a tight cluster comprising the other 4 conditions (Figure 2C). This was conceivable given that the +RA condition introduces a differentiation signal in contrast to the withdrawal of self-renewal signals in the other 4 conditions. Therefore, we next performed a combined analysis of the clustering four conditions (Figure 2D-F) to find central pathways that are important for PSD induced by the removal of self-renewal signals.

The combined analysis identified members of multiple chromatin modifying complexes important for PSD. Histone acetyltransferase (HAT) complex proteins appear to be top hits (Figure 2D-F), most prominently those that belong to SAGA-type (TFTC/STAGA) and TIP60 (NuA4) HAT complexes. Strikingly, the catalytic subunit of the TIP60 complex KAT5 is among the top 5 hits of all four conditions (Table S1), underlining the importance of this complex in PSD. While histone acetylation levels are known to be higher in undifferentiated ESCs (Legartova et al., 2014), a global but transient increase in histone acetylation occurs during PSD (Golob et al., 2008), potentially necessitating the activity of HAT complexes at the onset of PSD. Nucleosome remodelling is also essential, as knockdown of multiple SWI/SNF family members prevented efficient downregulation of NANOG in differentiating hESCs (Figure 2D,F). Notably, certain members of these complexes, such as Trrap of the NuA4 complex and Arid1a of the SWI/SNF complex, have a conserved function in shutting down pluripotency in ground-state mESCs (Betschinger et al., 2013). Thus, our study systematically highlights the importance of epigenetic regulations in the PSD of hESCs.

**The cell cycle machinery deterministically regulates PSD**

Besides epigenetic modifiers, our combined analysis also revealed that genes involved in cell cycle regulation were among the most enriched (Figure 2D-F), particularly those that are
involved in DNA replication during the S phase or in the G2-to-M transition (Figure 2D,E). On the contrary, we found no strong enrichment of processes specific to other phases of the cell cycle, implying that specific cell cycle regulations influence the cell fate transitions of pluripotent cells. While the cell cycle is found to influence cell fate decisions in perspectives other than proliferation (Lee et al., 2014; Rodier et al., 2009), knowledge about the regulation of PSD by the cell cycle is limited, especially so for G1-independent cell cycle states such as the S and G2 phases. To explore this topic, we first functionally validated the effect of cell cycle genes in PSD by impeding DNA replication or prolonging the gap phases using both genetic and chemical approaches (Figure 3A).

Genetic validation of the S- and G2-associated hits confirmed that their depletion (Figure S3A,B) impedes pluripotency gene downregulation after removal of self-renewal signals (Figure 3B,C). Since their depletion concomitantly perturbs progression through the cell cycle (Figure S3C,D), these results suggest that extending the S and G2 phases of the hESC cell cycle may dominantly impede PSD. In contrast, although lengthening of the G1 phase is associated with differentiation, artificially extending the G1 phase by knocking down CDK4/6, depleting Cyclin D or overexpressing CDK inhibitor p21 (Figure S3E-G) did not significantly affect the downregulation of pluripotency markers upon removal of self-renewal signals (Figure 3D). These findings suggest that the lengthening of the G1 phase does not elicit a deterministic effect on PSD of hESCs. NANOG-GFP fluorescence was also preserved (Figure S3H) and differentiation marker upregulation was inhibited (Figure S3I-K) upon knockdown of S and G2 phase progression genes but not upon G1 phase prolongation, corresponding with the observed changes in pluripotency marker expression.
Since depletion of cell cycle-related hits impeded S and G2 phase progression, we next validated if direct manipulation of these cell cycle phases using chemical inhibitors can similarly affect PSD. With proper dosage, we managed to enrich, but not completely arrest, hESCs in specific cell cycle phases (Figure S4A) without inducing extensive apoptosis (Figure S4B,C). In concordance with our screen results, we found that perturbing DNA replication in the S phase as well as delaying mitotic transition from the G2 phase consistently deterred PSD from different hESC lines (Figure 3E, S4D-G). In contrast, inhibitors that led to an elongation of the G1 and M phases did not exhibit PSD attenuation (Figure 3E, S4D-G), proving that restriction of PSD is not conferred by simply locking cell cycle progression. Importantly, these observed changes in marker expression ultimately influenced the functional pluripotency of hESCs (Figure 3F).

Since cell cycle manipulations in ESCs inevitably lead to a certain degree of cell death (Ruiz et al., 2011) (Figure S4B,C), we wanted to ascertain that the effects of cell cycle perturbation on PSD are not just a secondary effect of cell death. We observed that the levels of induced cell death did not correlate with the effects of cell cycle perturbation on PSD (Figure S4H). Additionally, inhibition of apoptosis during perturbation of the S or G2 phases did not alter the observed delay in PSD (Figure S4I). These findings collectively indicate that the effects of cell cycle perturbation on PSD are independent of cell death.

Altogether, these data demonstrate that the cell cycle can dominantly influence PSD. More importantly, these results also indicate that resistance to PSD is potentially mediated by specific cell cycle events in the S and G2 phases.
The propensity of the S and G2 phases towards pluripotent state maintenance is intrinsic and G1 phase-independent.

The G1 phase of the cell cycle has been correlated to susceptibility 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 G1 cyclins that contribute to lineage specification (Li et al., 2012; Pauklin and Vallier, 2013). As the reduction in proportion of cells in the G1 phase was observed when we delayed progression in the S and G2 phases (Figure S3C,D, S4A), it can be argued that the resulting PSD block could be an indirect effect of an inaccessibility to the G1 phase or G1-associated factors that initiate differentiation.

To test this hypothesis, we firstly measured the actual time cells spend in the G1 and S/G2 phases using the FUCCI cell cycle reporter (Sakaue-Sawano et al., 2008). While hESCs spent more time in the S/G2 phase when we interfere with DNA replication and G2 phase progression, the absolute time in the G1 phase remained largely unchanged (Figure 4A). However, the collective time spent in the G1 phase is still decreased if we allow the cells to differentiate for the same period of time (Figure 4B). Therefore, we measured pluripotency marker expression after an equal number of cell divisions in untreated and cell cycle-perturbed hESCs (Figure 4C). In these experiments, we observed that PSD was still attenuated in hESCs with delayed S or G2 phase progression (Figure 4C), ruling out the possibility that the restriction of PSD could be a result of G1 phase inaccessibility.

We next checked if the expression levels of G1 phase-specific factors that are implicated in differentiation were decreased. Neither CDK inhibitors such as p21 and p27 nor Cyclin D was downregulated upon knockdown of S or G2 phase-related hits (Figure 4D). In fact, the protein levels of these G1-associated factors were slightly increased, indicating that the
consequent prevention of PSD is definitely not due to an unavailability of G1-associated lineage specification factors. As delayed PSD could not be explained via the length of the G1 phase or by its associated factors, this phenomenon might stem from a direct effect of S- and G2-specific pathways, which render these phases of the cell cycle intrinsically inclined towards pluripotent state maintenance.

We also looked at the transcription of pluripotency factors to look for potential cell cycle phase-specific regulation. The steady-state levels of pluripotency genes were previously found to remain similar across the hES cell cycle (Singh et al., 2013). However, when we examined _de novo_ transcription levels of pluripotency regulators in different cell cycle phases, genes such as _NANOG_ and _PRDM14_ were preferentially transcribed during the S and G2/M phases compared to the G1 phase (Figure 4E). This suggests that the S and G2 phases may be intrinsically wired into maintaining pluripotency, thus delaying PSD when progression through these cell cycle phases is altered.

**Activation of the ATM/ATR-mediated checkpoint at the S phase attenuates PSD**

To look for the specific cell cycle machineries that uphold the pluripotency network, we first checked pathways that get activated when the S phase is perturbed. Perturbation of DNA replication generates replication stress which can lead to DNA damage accumulation, and subsequent activation of the ATM/ATR-mediated checkpoint (Bakkenist and Kastan, 2003; Zou and Elledge, 2003). Indeed, we observed elevated γH2AX foci (Figure S5A) and CHEK1/2 phosphorylation (Figure 5A, S5B) after perturbation of DNA replication. To test whether this activation of the checkpoint is causative for blocking PSD, we subdued ATM/ATR-mediated checkpoint signalling in hESCs using various checkpoint inhibitors (Figure 5B). Genetic and chemical inhibition of sensor kinases ATM and ATR or of effector
kinase CHEK2 reversed the block on PSD by Aphidicolin, as evidenced by the downregulation of pluripotency markers (Figure 5C,D, S5C-F). These changes in gene expression translate to functional pluripotency as observed in teratoma formation assays (Figure 5E). In addition, inhibition of the checkpoint similarly abolished the PSD delay resulting from knockdown of replication-associated hits (Figure 5F). These evidences from both genetic and chemical approaches demonstrate that signalling through the ATM/ATR-CHEK2 axis directly contributes to pluripotency. Interestingly, in all these cases wherein the checkpoint activated by Aphidicolin treatment was abolished, DNA replication remained largely delayed (Figure 5G). This verifies that the PSD block is not a result of a simple S phase lock or the physical state of DNA during replication, but is a direct result of ATM/ATR-mediated checkpoint activation. Finally, the activation of the ATM/ATR pathway by DNA damage-inducing reagents similarly attenuated PSD (Figure 5H, S5G), confirming that checkpoint activation can directly block PSD.

The activity of the ATM/ATR pathway on PSD is specifically downstream of the S phase replication perturbation, as increased CHEK2 phosphorylation was not observed when perturbing other cell cycle phases (Figure S5H). Besides, inhibiting checkpoint signalling did not change PSD kinetics when perturbing the other cell cycle phases (Figure S5I). Interestingly, we observed higher activity of the ATM/ATR pathway in unperturbed hESCs in the S phase compared to other cell cycle phases (Figure 5I), suggesting that this pathway may also contribute to the intrinsic propensity of the S phase towards pluripotent state maintenance. In summary, these results demonstrate the ATM/ATR-mediated checkpoint is the dominant PSD inhibitor downstream of DNA replication perturbation and also likely during normal S phase progression.
The ATM/ATR-mediated checkpoint activates p53 to enhance TGFβ signalling and uphold pluripotency

To obtain a mechanistic connection between ATM/ATR-CHEK2 activation and pluripotent state retention in hESCs, we performed a time-course microarray analysis in hESCs treated with DNA replication inhibitor Aphidicolin and checkpoint inhibitor AZD7762 (Figure 6A). Upon Aphidicolin treatment, we observed an upregulation of genes involved in active TGFβ signalling and a simultaneous downregulation of BMP4 pathway genes and TGFβ pathway antagonists (Figure 6B,C, S6A,B). The changes in TGFβ pathway gene expression led to enhanced TGFβ signalling as indicated by increased SMAD2 phosphorylation (Figure 6D, S6C), and can be mimicked by knockdown of replication-related hits (Figure 6E, S6D). Importantly, the augmentation of TGFβ signalling can be reversed by checkpoint inhibitor treatment (Figure 6B-D, S6E), confirming that the effect on TGFβ signalling is directly from checkpoint activation. These results therefore establish that the ATM/ATR-CHEK2 axis augments TGFβ signalling.

The TGFβ pathway has a well-known role in promoting the human pluripotent state (James et al., 2005), which we observe to occur in a dose-dependent manner (Figure S6F). Thus, it can be inferred that heightened TGFβ signalling is responsible for the PSD inhibition during replication perturbation. However, it is also possible that the block in PSD occurs upstream of the changes in TGFβ-related gene expression. We therefore compared time-course expression profiles upon the withdrawal of self-renewal signals and observed that TGFβ-related gene expression changes occur earlier and to a greater degree compared to pluripotency marker expression changes (Figure 6C). This implies that the changes in TGFβ-related gene expression are under closer control of the ATM/ATR-mediated checkpoint compared to pluripotency genes. In addition, these changes were detected even when cells are in hESC
medium (Figure S6E), and altered TGFβ-related expression does not simply correlate with pluripotency status (Figure S6G). These data rule out the possibility that changes in TGFβ-related gene expression could be a secondary effect from delayed PSD, and supports the notion that pluripotent state preservation occurs as a result of the initial augmentation of TGFβ signalling.

Changes in TGFβ-related gene expression upon checkpoint activation indicate an underlying transcriptional regulatory mechanism. To find out the responsible transcription factor, we first looked at the canonical transcription factor p53, as it is activated by the ATM/ATR-CHEK2 axis during DNA damage response in the S phase (Banin et al., 1998; Hirao et al., 2000; Tibbetts et al., 1999) (Figure 6D, S6C,D,H) and was observed to have binding sites near TGFβ-related genes in hESCs (Akdemir et al., 2014). We thus manipulated p53 levels in hESCs to observe its effect on the prevention of PSD by checkpoint activation (Figure 6F). *TP53* knockdown (Figure S6I) in checkpoint-activated hESCs not only abolished the delay in PSD (Figure 6G) but also reversed the altered expression of TGFβ-related genes (Figure 6H), proving its necessity in impeding PSD. Moreover, stabilization of p53 without DNA damage using Nutlin-3 (Figure S6J,K) is sufficient to delay pluripotency marker downregulation (Figure 6I) and alter TGFβ-related gene expression (Figure 6J) in the absence of self-renewal factors. Importantly, Nutlin-3 treatment did not induce significant changes in cell cycle profile of hESCs (Figure S6L), confirming that p53 can directly delay PSD, independent of cell cycle changes. These data demonstrate both the necessity and sufficiency of p53 in upholding the human pluripotent state. Finally, in line with the elevated ATM/ATR activity during the S phase (Figure 5I), we observed higher p53 protein and SMAD2 phosphorylation levels in S phase-sorted hESCs (Figure S6M), implying that the mechanisms discussed similarly work during normal progression through the cell cycle.
Collectively, our data strongly entails for a detailed mechanism behind how S phase perturbation upholds pluripotency (Figure 6K). In the presence of replication stress or DNA damage, p53 stabilization by ATM/ATR-CHEK2 signalling enhances TGFβ pathway signalling, which consequently sustains \textit{NANOG} expression (Vallier et al., 2009; Xu et al., 2008) and preserves the pluripotent state.

**High levels of Cyclin B1 in the G2 phase attenuate PSD**

We finally looked at how the enrichment of hESCs at the G2 phase could delay PSD (Figure 7A). Genetic and chemical perturbations of G2 phase progression resulted in a consistent elevation of Cyclin B1 expression (Figure 7B, S7A), but not other cyclins (Figure 7B). Furthermore, Cyclin B1 was higher in hESCs sorted in the G2 phase (Figure S7B). Therefore, we hypothesized that Cyclin B1 might be the underlying cause for promoting the pluripotent state during the G2 phase. Interestingly, when we knocked down Cyclin B1 in hESCs (Figure S7C), a dramatic downregulation of pluripotency marker expression was observed prior to cell death (Figure 7C, S7D), indicating a tight linkage between Cyclin B1 and pluripotency. Cyclin B1 knockdown enriched cells in the G2 phase instead of G1 (Figure 7D), strikingly demonstrating a decoupling of the expected cell cycle profile associated with differentiation.

We furthermore overexpressed Cyclin B1 in hESCs (Figure S7E) and demonstrated that overexpression of a single cell cycle factor can delay PSD (Figure 7E, S7F). Moreover, Cyclin B1 overexpression did not cause significant enrichment of cells in the G2 phase (Figure 7F), indicating that Cyclin B1 prevents PSD downstream of G2 phase prolongation. Importantly, these observations extend to protein levels of pluripotency markers (Figure S7G), are replicable in other hESC lines (Figure S7H,I), and ultimately affect functional pluripotency (Figure 7G). Altogether, these results indicate that Cyclin B1 is a connecting
node between the G2 phase and its ability to prevent PSD. 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 initial upregulation of TGFβ agonists (Figure S7J,K), suggesting that Cyclin B1 might also work through TGFβ to prevent PSD.
Discussion

The context-dependent regulation of PSD

Differentiation necessitates the disintegration of the pluripotency network. Yet how differentiation cues lead to the breakdown of this network is ill-defined, especially in hESCs. To identify genes required for PSD, we performed a systematic large-scale RNAi screen under multiple conditions to examine the regulation of PSD given various cues initiating differentiation. Firstly, we have observed that the regulation of PSD upon RA introduction is starkly divergent from the conditions where hESCs were deprived of self-renewal signals (Figure 2C). 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 influences PSD regulation. Secondly, our multi-conditional screening approach identified both general and context-specific regulators of PSD. We found that the withdrawal of self-renewal signals is associated with several common PSD effectors. 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 (Figure 2D-F). Conversely, other PSD regulators like RNA splicing and Wnt signalling tend to function in a more context-dependent manner (Figure 2B). Finally, we found that multiple members of certain complexes were detected to play a role in PSD (Figure 2B,D), providing a starting point to sift out the function of different members of protein complexes by revealing the ones which are specifically important for PSD in hESCs. Altogether, our large-scale RNAi study provides a unique resource for dissecting the mechanisms behind PSD, by providing new insight on the universal roles of various factors and pathways on PSD while raising important
notions about the inherent context-specificity of the regulatory network governing PSD in hESCs.

The deterministic effect of cell cycle states on PSD

During proliferation, cells experience dramatic biochemical and physical changes by going through the cell cycle. This results in distinct cell cycle states, which cells have evolved to utilize in order to prime and regulate other events that are not immediately related to proliferation such as immune response, metabolism and lineage specification (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. Here, we demonstrate that specific cell cycle states dominantly block PSD in hESCs. Specifically, when progression through the S or G2 phases of the hES cell cycle is perturbed, cells are enriched at specific cell cycle states with machineries that trigger a selective preference towards pluripotency maintenance and delay PSD (Figure 3). Particularly, the ATM/ATR-CHEK2-mediated activation of p53 and Cyclin B1 upregulation during arrest at the S and G2 phases, respectively, defines their respective cell cycle states and exerts control over the pluripotency network. Moreover, these pathways are not only activated at the perturbed states but are also involved in natural cell cycle progression (Figure 5I, S6M, S7B). Thus, these pathways and their effects in upholding pluripotency can be applied to normal progression through the S and G2 phases. Overall, our investigation provides the first evidence that cell cycle machineries corresponding to specific states can directly and dominantly regulate the human pluripotent state.
Furthermore, the employment of the said pathways by the S and G2 phases establish their active role in boosting the pluripotent state. This differs from the current perspective that the S and G2 phases passively retain the pluripotent state, given that only the G1 phase harbours a propensity to receive extracellular differentiation signals and express lineage specific factors (Scott et al., 1982; Singh et al., 2013). In fact, we hypothesize that the absence of such pathways in the G1 phase may underlie its responsiveness to differentiation cues. Hence, we propose that the S/G2 and G1 phases shift the weights between pluripotency maintenance and PSD priming across the hES cell cycle (Figure 7H). This model advocates that a balance between cell cycle phases is critical for ESC fate determination unlike the previously recognized G1-centric model.

Consequences of fate choices upon DNA damage in pluripotent stem cells

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. Here, we demonstrate that pluripotent stem cells resist cell fate changes in the presence of replication-induced DNA damage (Figure 5). We further show that the DNA damage-induced checkpoint stabilizes p53 to regulate TGFβ pathway gene expression and retain pluripotency marker expression (Figure 6). As a hub transcription factor, p53 was known to enact diverse functions under different cellular contexts and activation dynamics (Aylon and Oren, 2007; Chang et al., 2011; Ubil et al., 2014). While p53 activation was thought to trigger apoptosis or differentiation in hESCs (Qin et al., 2007), chromatin binding of p53 differs depending on whether it is activated by differentiation signal or DNA damage (Akdemir et al., 2014). Here, we demonstrate a new perspective where p53 acts to maintain the human pluripotent state upon the withdrawal of
self-renewal factors (Figure 6F-J). By retaining the pluripotent cell fate in the face of DNA damage, cells benefit from an increased efficiency of DNA damage repair as ESCs are known to express higher levels of homologous recombination and damage repair proteins (Momcilovic et al., 2010; Tichy et al., 2010). Furthermore, if the damage incurred is beyond the capacity to repair, hESCs have the choice to undergo apoptosis more effectively due to a lower apoptotic threshold in hESCs compared to differentiated cells (Dumitru et al., 2012; Liu et al., 2013). As a result, pluripotent stem cells can more efficiently resolve damage before committing to differentiation, or if the damage is beyond repair, alternatively terminate themselves to avoid the potential detriment of giving rise to lineage progenitors with damaged genomes.

Conclusions

In conclusion, our systematic functional screening of hESCs under multiple differentiation conditions enabled the identification of both universal and specific gatekeepers for PSD (Figure 7I). We unravel a universal regulatory role for HAT complexes and the SWI/SNF complex in PSD, while the NuRD complex, RNA splicing and signalling pathways regulate PSD in a context-specific manner. Our screen unexpectedly discovered a strong enrichment of cell cycle hits, specifically clustering in the S and G2 phases but not in other phases of the cell cycle. Our mechanistic studies demonstrate that the S and G2 phases possess an intrinsic propensity towards the pluripotent state, mediated by the ATM/ATR-CHEK2-p53 and Cyclin B1 pathways, respectively. Thus, we introduce a new paradigm for the coupling of cell cycle and pluripotency, wherein the S and G2 phases employ pathways that inhibit PSD and maintain the ESC identity, while the absence of such pathways in the G1 and M phases potentially contributes to their amenability to PSD (Figure 7H).
Figure Legends

Figure 1. A high-throughput RNAi screen for regulators of PSD. (A) Schematic representation of siRNA screen. siNT = non-targeting siRNA; siOCT4 = siRNA against POU5F1 (OCT4); siGFP = siRNA against GFP. (B) Representative plate heatmap from the -bFGF, -TGFβ condition. (C) Scatter plot showing the correlation between screen replicates in the -bFGF, -TGFβ condition. (D) Representative dot plot of the results of the transcription factor and epigenetic modifier subset from the -bFGF, -TGFβ condition. Gray line indicates the cut-off z-score > 1.5. (E) Representative images for NANOG-GFP fluorescence (green) and Hoechst staining (blue) for hits in the five differentiation conditions. Scale bar = 200um. See also Figure S1-S2, Tables S1-S2.

Figure 2. Identification of context-dependent and universally important regulators of PSD. (A) Diagram outlining the major findings from the RNAi screen. (B) 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. See also Tables S3-S4. (C) Heatmap depicting hierarchical clustering by Euclidian distance between different screening conditions. (D) 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. (E) Overrepresented pathways (p<0.05) from the combined hits as determined using the web resource Reactome. (F) Representative terms (p<0.05) for enriched functional annotation clusters from the combined hits.
Figure 3. Deterministic regulation of the cell cycle on PSD. (A) Schematic outlining the cell cycle perturbations at specific phases. (B-E) qPCR for pluripotency genes upon genetic perturbations of (B) DNA replication, (C) G2 phase progression and (D) G1 phase progression, or upon (E) treatment with small molecules enriching for specific cell cycle phases in the -bFGF, -TGFβ condition. Triplicate data represented as mean +/- SD. (F) Images for embryoid bodies and immunofluorescence staining of differentiated cells derived from H1 hESCs treated with various cell cycle inhibitors in the -bFGF, -TGFβ condition. N.A. = no cells survived. Scale bar = 100um. See also Figure S3-S4.

Figure 4. The propensity for promoting pluripotency in the S and G2 phases is intrinsic and G1 phase-independent. (A) Boxplots depicting time spent in the G1 and S/G2 phases as calculated using live cell imaging of FUCCI H1 hESCs after various cell cycle perturbations. * = p-value < 0.05 compared to empty vector or DMSO control. (B) Diagram depicting the difference in lengths of time-controlled and cell cycle number-controlled experiments. (C) qPCR for pluripotency genes in the -bFGF, -TGFβ condition upon cell cycle perturbations, with data collection time points normalized to the number of cell cycles. Triplicate data represented as mean +/- SD. (D) Western blot for p21, p27 and Cyclin D1 levels upon knockdown of S- and G2-associated hits. (E) qPCR for nascent transcripts of pluripotency genes collected from H1 hESCs sorted into different cell cycle phases according to DNA content.

Figure 5. ATM/ATR-CHEK2 prevents PSD upon replication arrest. (A) Western blot of phosphorylated CHEK2 and CHEK1 proteins upon treatment with cell cycle inhibitors. (B) Outline of manipulations performed on the ATM/ATR pathway. (C-D) qPCR for pluripotency genes in the -bFGF, -TGFβ condition upon treatment with Aphidicolin
concomitant with (C) treatment with shRNAs against ATM/ATR pathway members, or (D) inhibition of the ATM/ATR pathway with small molecules. Triplicate data represented as mean +/- SD. (E) Table describing teratoma formation efficiencies of H1 hESCs treated with Aphidicolin, AZD7762 and Caffeine in the -bFGF, -TGFβ condition, and teratoma formation assay for Aphidicolin-treated cells. (F) qPCR for pluripotency genes upon knockdown of S phase-associated hits together with DMSO or AZD7762 treatment in the -bFGF, -TGFβ condition. Triplicate data represented as mean +/- SD. (G) FACS quantification of Hoechst staining indicating the cell cycle status of H1 hESCs after ATM/ATR pathway inhibition in hESC medium. (H) qPCR for pluripotency genes in the -bFGF, -TGFβ condition upon treatment with DNA damage-inducing reagents. Triplicate data represented as mean +/- SD. (I) Western blot for phosphorylated CHEK2 levels in H1 hESCs sorted according to cell cycle phases based on DNA content. See also Figure S5.

**Figure 6. Augmentation of the TGFβ and p53 pathways by the ATM/ATR-mediated checkpoint.** (A) Schematic diagram of the time-course microarray experiment. (B) Microarray heatmap for differentially expressed genes upon treatment with Aphidicolin and AZD7762 in the -bFGF, -TGFβ condition. (C) Time-course qPCR for pluripotency and TGFβ-related genes upon treatment with Aphidicolin and AZD7762 in the -bFGF, -TGFβ condition. Triplicate data represented as mean +/- SD. (D) Western blot of phospho-SMAD2 and p53 proteins upon treatment of H1 hESCs with Aphidicolin and AZD7762. (E) qPCR for TGFβ agonists (color) and antagonists (grayscale) in the -bFGF, -TGFβ condition upon knockdown of replication-related genes. (F) Outline of experimental manipulations performed on the ATM/ATR-CHEK2-p53 axis. (G-J) qPCR of H1 hESCs in the -bFGF, -TGFβ condition for (G) pluripotency genes and (H) TGFβ agonists (color) and antagonists (grayscale) upon knockdown of TP53 concomitant with treatment with Aphidicolin, or (I)
pluripotency genes and (J) TGFβ agonists (color) and antagonists (grayscale) upon stabilization of p53 with Nutlin-3. Triplicate data represented as mean +/- SD. (K) Schematic flowchart outlining the mechanism behind the regulation of pluripotency by the ATM/ATR-mediated checkpoint. See also Figure S6.

**Figure 7. Cyclin B1 attenuates PSD upon G2 phase prolongation.** (A) Outline of manipulations performed on Cyclin B1. (B) Western blot of cyclin levels upon knockdown of G2-associated hits. (C) qPCR for pluripotency genes upon knockdown of Cyclin B1 in hESC medium. Triplicate data represented as mean +/- SD. (D) FACS quantification indicating the cell cycle status of H1 hESCs based on DNA content after Cyclin B1 knockdown in hESC medium. (E) qPCR for pluripotency genes in the -bFGF, -TGFβ condition upon overexpression of Cyclin B1. Triplicate data represented as mean +/- SD. (F) FACS quantification indicating the cell cycle status of H1 hESCs based on DNA content after Cyclin B1 overexpression in hESC medium. (G) Images for embryoid bodies and immunofluorescence staining of differentiated cells derived from H1 hESCs overexpressing Cyclin B1 in the -bFGF, -TGFβ condition. Scale bar = 100um. See also Figure S7. (H) Model for the deterministic regulation of the cell cycle on PSD. The G2 and S phases are inclined towards pluripotent state maintenance due to active pluripotent state-promoting pathways. (I) Summary of the findings from the high-throughput RNAi screen for PSD. Both core and context-specific gatekeepers were uncovered, providing a mechanistic model of PSD regulation in hESCs. Color scheme for core gatekeepers follow Figure 2D, while color scheme for context-specific gatekeepers follow Figure 2B.
Methods

Cell Culture

The hESC lines H1 (WA-01), HES2, HES3, \textit{NANOG-GFP} H1 reporter and \textit{ACTIN-GFP} H1 reporter 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 1U/ml Dispase in DMEM/F12 (STEMCELL Technologies) every 4–5 days. For experiments, hESCs were passaged using TrypLE™ Express (Life Technologies) before treatment. See Extended Experimental Procedures for experimental treatment protocols.

High-throughput RNAi screening

Pooled siRNAs (2.5uL of 1uM; Dharmacon, Ambion) were printed on 384-well plates (Greiner) coated with 30x-diluted Matrigel (BD) and frozen at -80°C before use. Lipofectamine RNAi Max (Invitrogen) diluted 200x in 5uL OptiMEM (Invitrogen) was added per well, and incubated at room temperature for 20 mins. 3,000 \textit{NANOG-GFP} H1 hESCs in 45uL mTeSR1 (STEMCELL Technologies) containing approximately were seeded into each well. 24hrs after seeding, mTeSR1 was replaced with differentiation media: 1) -bFGF, -TGFβ condition: mTeSR1 without select growth factors (STEMCELL Technologies), 2) TGFβ pathway inhibition: mTeSR + 1uM A8301 (Stemolecule), 3) MEK pathway inhibition: mTeSR1 + 2.5uM PD0325901 (Sigma), 4) PI3K pathway inhibition: mTeSR1 + 20uM LY294002, 5) +RA: mTeSR1 + 20uM RA (Sigma). Cells were incubated in differentiation media for 120hrs for condition 1, and 48 hrs for conditions 2-5. Media were then replaced with mTeSR1 and incubated for another 24 hrs. Cells were then fixed with 4% paraformaldehyde (Sigma), and stained with Hoechst 3342 (1:4000, Invitrogen). Microscope
images were acquired using ImageXpress Ultra (Research Instruments) at 20x magnification, and quantified using MetaXpress Image Acquisition and Analysis software. See Extended Experimental Procedures for informatics analysis.

RT-qPCR
RNA extraction using TRIZol® (Invitrogen), reverse transcription using Superscript II (Invitrogen) and qPCR using SYBR Green (KAPA) were all performed via standard procedures. Measured transcripts were normalized to GAPDH levels.

Immunoblotting and immunostaining
Immunoblotting and immunostaining were performed as conventional procedures. See Extended Experimental Procedures for details and antibodies used.

Microarray
mRNAs derived from hESCs were reverse transcribed, labelled and analyzed on Illumina microarray platform (HumanHT-12 v4 Expression BeadChips) according to manufacturer’s instructions. Rank invariant normalization was applied. The Gene Expression Omnibus (GEO) accession numbers for the microarray data are GSE62062 and GSE63215.
References


Author contributions:
K.A.G. designed research, performed experiments, analysed data, wrote the paper, and supervised the overall project. H.L. designed research, performed experiments, analysed data and wrote the paper. Y-S.L., Y-S.C., J-C.Y., C-P.T., B.G., B.L., Z-Y.T. and K-Y.L. performed experiments. F.B. and Y-C.L supervised the overall project. H-H.N. designed research, wrote the paper, and supervised the overall project.

Acknowledgements:
This work was funded by the Agency for Science, Technology and Research, Singapore and the National University of Singapore. We thank G. Goh, J. Chia, P. Kumar, G. Periyasamy, H. M. Seah, P. Cheng, P. Kraus, F. Sachs and I. Liu for assistance in experimental work. We also thank J. H. Ng, X. Lu, C. Chu and Y. Hu for comments on the manuscript, and U. Surana for intellectual discussion. The Tol2 cassette was a gift from K. Kawakami in the National Institute of Genetics, Japan.
Supplemental Information
Deterministic restriction on pluripotent state dissolution by cell cycle pathways (Gonzales, Liang et al.)

Supplemental Figure 1. Characterization of the NANOG-GFP H1 hESC line and differentiation media used for the RNAi screens, Related to Figure 1.

Supplemental Figure 2. Quality controls for the RNAi screens, Related to Figure 1.

Supplemental Figure 3. Genetic validations for the role of the cell cycle in pluripotent state dissolution, Related to Figure 3.

Supplemental Figure 4. Chemical validations for the role of the cell cycle in pluripotent state dissolution, Related to Figure 3.

Supplemental Figure 5. Validations for the role of the ATM/ATR-mediated checkpoint in pluripotent state dissolution, Related to Figure 5.

Supplemental Figure 6. Validations for the role of the TGFβ pathway and p53 in the restriction of pluripotent state dissolution by the ATM/ATR-mediated checkpoint, Related to Figure 6.

Supplemental Figure 7. Validations for the role of Cyclin B1 in pluripotent state dissolution, Related to Figure 6.

Supplemental Table 1. Gene list for the RNAi screen on NANOG-GFP H1 cells under the 5 differentiation conditions, Related to Figure 1.

Supplemental Table 2. Gene list for the RNAi screen on ACTIN-GFP H1 cells under the -bFGF, -TGFβ and TGFβ pathway inhibition conditions, Related Figure 1.

Supplemental Table 3. Reactome analysis for the 5 individual screening conditions, Related to Figure 2.

Supplemental Table 4. Gene ontology analysis for the 5 individual screening conditions, Related Figure 2.
Figure S1. Characterization of the NANOG-GFP H1 hESC line and differentiation media used for the RNAi screens, Related to Figure 1. (A-B) Expression of (A) hESC-specific genes and (B) lineage specification genes after incubation in the -bFGF, -TGFβ condition for 48 and 96 hrs relative to H1 hESCs in hESC medium as measured by microarray analysis. Triplicate data represented as mean +/- SD. (C) Heatmap depicting expression levels of pluripotency and lineage-specific genes of single H1 hESCs after incubation in the -bFGF, -TGFβ condition for 48 and 96 hrs relative to H1 hESCs in hESC medium. (D) Diagram depicting pluripotent state dissolution at the initiation of the differentiation process. (E) Images for NANOG-GFP fluorescence (green) and Hoechst staining (blue) upon transfection with control siRNAs. Scale bar = 400um. (F) Quantitative PCR for NANOG upon transfection of NANOG-GFP H1 cells with control siRNAs. NANOG expression levels correspond with GFP fluorescence intensity. Triplicate data represented as mean +/- SD. (G) Cytogenetic analysis of NANOG-GFP H1 cells. The cells show a normal karyotype with 46XY chromosomes. (H) Teratoma formation assay for NANOG-GFP H1 cells. Teratoma consisted of tissues from ectoderm, mesoderm and endoderm lineages 6 weeks after injection into SCID mice. (I) Images for NANOG-GFP fluorescence (green) and Hoechst staining (blue) after incubation in the differentiation media and recovery in hESC medium. Scale bar = 200um. (J) Immunofluorescence staining for pluripotency markers after incubation in the differentiation media. Scale bar = 100um. (K) Quantitative PCR for pluripotency and lineage-specific markers after incubation in the differentiation media. Triplicate data represented as mean +/- SD.

Supplemental Figure 2. Quality controls for the RNAi screens, Related to Figure 1. (A) Representative plate heatmaps from the TGFβ pathway inhibition, MEK pathway inhibition, PI3K pathway inhibition and + Retinoic acid conditions. No plate layout effects are evident. (B) Representative box plot of plate alignment from the -bFGF, -TGFβ condition. z-scores are properly aligned between plates after normalization. Similar results were observed for the other 4 conditions. (C) Scatter plots showing the correlation between replicates in the TGFβ pathway inhibition, MEK pathway inhibition, PI3K pathway inhibition and + Retinoic acid conditions. (D) Representative dot plot of z-scores (x-axis) versus cell count (y-axis) from the -bFGF, -TGFβ condition. Hits are acquired across low to high cell counts, and no bias is observed towards low cell numbers. Similar results were observed for the other 4 conditions. (E) Venn diagram showing the overlap between hits using ACTIN-GFP H1 cells versus
NANOG-GFP H1 cells from the kinase, phosphatase and GPCR subsets under the -bFGF, -TGFβ and the TGFβ pathway inhibition conditions.

**Supplemental Figure 3. Genetic validations for the role of the cell cycle in pluripotent state dissolution, Related to Figure 3.** (A-B) Quantitative PCR for (A) S phase- and (B) G2 phase-associated genes upon RNAi knockdown. All RNAi constructs effectively decreased expression of their target gene. Triplicate data represented as mean +/- SD. (C-D) FACS quantification indicating the cell cycle status of H1 hESCs based on DNA content after RNAi knockdown of (C) S phase- and (D) G2 phase-associated hits in hESC medium. (E) Quantitative PCR for G1-associated genes upon shRNA knockdown. All shRNAs effectively decreased expression of their target gene. Triplicate data represented as mean +/- SD. (F) Western blot of p21 levels upon overexpression in H1 hESCs. (G) FACS quantification indicating the cell cycle status of H1 hESCs based on DNA content after RNAi knockdown of G1 phase-associated genes in hESC medium. (H) Images for NANOG-GFP (green) and Hoechst staining (blue) after knockdown of cell cycle-associated genes in the -bFGF, -TGFβ condition. Scale bar = 400um. (I-K) Quantitative PCR for differentiation genes upon genetic perturbations of the (I) S, (J) G2 and (K) G1 phases in the -bFGF, -TGFβ condition. Triplicate data represented as mean +/- SD.

**Supplemental Figure 4. Chemical validations for the role of the cell cycle in pluripotent state dissolution, Related to Figure 3.** (A) FACS quantification indicating the cell cycle status of H1 hESCs after treatment with small molecules known to trigger cell cycle arrest in hESC medium. (B) Images for TUNEL staining (green) and Hoechst staining (blue) after treatment with cell cycle inhibitors. Scale bar = 100um. (C) Quantification of Annexin V-positive H1 hESCs after treatment with cell cycle inhibitors as measured by FACS. (D) Images for NANOG-GFP (green) and Hoechst staining (blue) after treatment with cell cycle inhibitors in the -bFGF, -TGFβ condition. Scale bar = 400um. (E) Quantitative PCR for pluripotency genes upon treatment of HES2 and HES3 hESCs with cell cycle inhibitors in the -bFGF, -TGFβ condition. Triplicate data represented as mean +/- SD. (F) Quantitative PCR for differentiation genes upon treatment of H1 hESCs with cell cycle inhibitors in the -bFGF, -TGFβ condition. Triplicate data represented as mean +/- SD. (G) Immunofluorescence staining for pluripotency markers upon treatment of H1 hESCs with cell cycle inhibitors in the -bFGF, -TGFβ condition. Scale bar = 100um. (H) Scatter plot depicting correlation between the percentage of Annexin V-positive H1 hESCs and degree of PSD prevention.
upon treatment of H1 hESCs with cell cycle inhibitors in the -bFGF, -TGFβ condition. PSD prevention score is the log2 mean of the 5 pluripotency markers in Fig. 3E. No correlation between the levels of PSD and apoptosis was observed. (I) Quantitative PCR for pluripotency genes upon concomitant treatment of H1 hESCs with cell cycle inhibitors and the caspase inhibitor Boc-D-FMK. Apoptotic inhibition had no effect on the PSD prevention by S and G2 phase inhibitors.

Supplemental Figure 5. Validations for the role of the ATM/ATR-mediated checkpoint in pluripotent state dissolution, Related to Figure 5. (A) Hoechst (blue) and anti-γ-histone 2A.X immunofluorescence (red) staining of H1 hESCs after RNAi knockdown of replication-associated hits or treatment with Aphidicolin in hESC medium. Scale bar = 6um. (B) Western blot of phosphorylated CHEK2 upon knockdown of replication-related hits in H1 hESCs. (C) Quantitative PCR for ATM/ATR-mediated checkpoint genes upon shRNA knockdown. All shRNAs effectively decreased expression of their target gene. Triplicate data represented as mean +/- SD. (D) Images for NANOG-GFP fluorescence (green) and Hoechst staining (blue) upon inhibition of the ATR/ATM-mediated checkpoint with shRNAs and small molecules. Scale bar = 400um. (E) Immunofluorescence staining for pluripotency markers upon treatment of H1 hESCs with Aphidicolin and checkpoint inhibitors in the -bFGF, -TGFβ condition. Scale bar = 100um. (F) Quantitative PCR for pluripotency genes upon treatment of HES2 and HES3 hESCs with Aphidicolin and checkpoint inhibitors in the -bFGF, -TGFβ condition. Triplicate data represented as mean +/- SD. (G) Images for NANOG-GFP (green) and Hoechst staining (blue) after treatment with DNA damage-inducing reagents in the -bFGF, -TGFβ condition. Scale bar = 400um. (H) Western blot of phosphorylated CHEK2 upon treatment of H1 hESCs with cell cycle inhibitors. (I) Quantitative PCR for pluripotency genes upon co-treatment of H1 hESCs with cell cycle inhibitors and either DMSO or AZD7762 in the -bFGF, -TGFβ condition. N.A. = no cells survived. Triplicate data represented as mean +/- SD.

Supplemental Figure 6. Validations for the role of the TGFβ pathway and p53 in the restriction of pluripotent state dissolution by the ATM/ATR-mediated checkpoint, Related to Figure 6. (A-B) Gene ontology analysis for (A) upregulated and (B) downregulated genes identified in the microarray analysis upon Aphidicolin treatment for 48 hours in the -bFGF, -TGFβ condition. (C-D) Western blot for phosphorylated SMAD2 and p53 protein levels upon (C) treatment with cell cycle inhibitors or (D) knockdown of
replication-related hits. (E) Microarray heatmap for differentially expressed genes upon treatment with Aphidicolin and AZD7762 in hESC medium. (F) Quantitative PCR for pluripotency genes upon supplementation of H1 hESCs with various concentrations of recombinant TGFβ in the -bFGF, -TGFβ condition. Triplicate data represented as mean +/- SD. (G) Quantitative PCR for pluripotency markers and TGFβ-related genes after incubation in different media. No consistent correlation between pluripotency markers and TGFβ-related genes is observed. Triplicate data represented as mean +/- SD. (H) Hoechst (blue) and anti-p53 (red) staining of H1 hESCs after treatment with Aphidicolin and AZD7762 in the -bFGF, -TGFβ condition. Scale bar = 40um. (I) Quantitative PCR for TP53 upon knockdown in H1 hESCs. Triplicate data represented as mean +/- SD. (J) Western blot of p53 upon treatment of H1 hESCs with Nutlin-3. (K) Hoechst (blue) and anti-p53 (red) staining of hESCs after treatment with Nutlin-3 in the -bFGF, -TGFβ condition. Scale bar = 40um. (L) FACS quantification indicating the cell cycle status of H1 hESCs based on DNA content upon Nutlin-3 treatment in hESC medium. (M) Western blot for pSMAD2 and p53 levels in H1 hESCs sorted according to cell cycle phases based on DNA content.

Supplemental Figure 7. Validations for the role of Cyclin B1 in pluripotent state dissolution, Related to Figure 7. (A-B) Western blot of Cyclin B1 protein in H1 hESCs (A) treated with cell cycle inhibitors or (B) sorted according to cell cycle phases based on DNA content. (C) Quantitative PCR of CCNB1 upon shRNA knockdown in H1 hESCs. Triplicate data represented as mean +/- SD. (D) Representative images for NANO-GFP fluorescence (green) and Hoechst staining (blue) after shRNA knockdown of CCNB1 in hESC medium. Scale bar = 400um. (E) Western blot of Cyclin B1 levels upon ectopic expression in H1 hESCs. (F) Representative images for NANO-GFP fluorescence (green) and Hoechst staining (blue) after ectopic expression of Cyclin B1 in the -bFGF, -TGFβ condition. Scale bar = 400um. (G) Immunofluorescence staining for pluripotency markers upon knockdown or overexpression of Cyclin B1. Scale bar = 100um. (H-I) Quantitative PCR for pluripotency genes upon (H) knockdown or (I) overexpression of Cyclin B1 in HES2 and HES3 hESCs in hESC medium and -bFGF, -TGFβ condition, respectively. Triplicate data represented as mean +/- SD. (J) Microarray heatmap for differentially expressed genes upon overexpression of Cyclin B1 in the -bFGF, -TGFβ condition. (K) Time-course quantitative PCR for pluripotency and TGFβ-related genes upon overexpression of Cyclin B1 in the -bFGF, -TGFβ condition.
Supplemental Table 1. Gene list for the RNAi screen on *NANOG-GFP* H1 cells under the 5 differentiation conditions, Related to Figure 1. Genes are sorted by average z-scores.

Supplemental Table 2. Gene list for the RNAi screen on *ACTIN-GFP* H1 cells under the -bFGF, -TGFβ and TGFβ pathway inhibition conditions, Related to Figure 1. Genes are sorted by average z-scores. Only the kinase, phosphatase and GPCR subsets of the library were used.

Supplemental Table 3. Reactome analysis for the 5 individual screening conditions, Related to Figure 2. Top pathways that are overrepresented (p<0.05) among the hits as determined using the web resource Reactome in each of the five conditions are shown.

Supplemental Table 4. Gene ontology analysis for the 5 individual screening conditions, Related to Figure 2. Representative terms for enriched functional annotation clusters in each of the five conditions are shown.
Supplemental Methods

Deterministic restriction on pluripotent state dissolution by cell cycle pathways (Gonzales, Liang et al.)

Small molecule treatment

For treatment with small molecules, chemicals diluted in DMSO were added to cell culture medium as follows: PD0332991 (6uM, Santa Cruz), aphidicolin (75ng/mL, Sigma), 5-hydroxyurea (0.2mM, Sigma), alisertib (65nM, Selleckchem), RO3306 (10uM, Enzo Life Science), nocodazole (18ng/mL, SciMed), barasertib (50nM, Selleckchem), BI2539 (5.5nM, Selleckchem), taxol (1.1nM, Santa Cruz), caffeine (1mM, Sigma), AZD7762 (100nM, Selleckchem), nutlin-3 (4uM, Cayman Chemical), doxorubicin (5ng/mL, Sigma), etoposide (20nM, Sigma).

hESC differentiation assays

For spontaneous differentiation through embryoid body formation, hESCs were dissociated with TrypLE™ Express (Life Technologies) and cultured in suspension inside low attachment 10cm dishes. For growth factor-induced differentiation, cells were seeded onto Matrigel and treated with the respective medium for differentiation along different lineages: 100ng/ml Activin A in advanced RPMI medium (Gibco) containing 2% FBS (Gibco) for definitive endoderm differentiation, 100ng/ml BMP4 and 4ng/ml bFGF in F12 DMEM (Gibco) containing 20% KSR (Gibco) for mesoderm differentiation, 100ng/ml BMP4 and 1 μM PD0325901 in F12 DMEM (Gibco) containing 20% KSR (Gibco) for trophectoderm differentiation, and 10μM SB431542 and 2μM dorsomorphin in neural basal medium supplemented with 0.5x N2 and B27 supplements for neural ectoderm differentiation.
Knockdown of genes 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. 2μL of concentrated viruses was added to culture medium with 4μg/mL polybrene (Sigma). Knockdowns were also performed using siRNAs (Bioneer) at a final concentration of 50nM transfected using Lipofectamine RNAi Max (Invitrogen).

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 hybridization (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).

**Informatics analysis for screen results**

$Z$ 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).

For individual analyses (Fig. 2B, Supp. Fig. 4), 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 (Fig. 2D-F), genes that were considered hits in at least 2 of the 4 conditions (- bFGF, - TGF-β, TGF-β pathway inhibition, MEK pathway inhibition and PI3K pathway inhibition) were used. Hits for the + Retinoic acid condition were excluded in the combined analysis. For the ACTIN-GFP H1 RNAi screens (Supp. Fig. 3E), genes considered hits have $z>1.25$ in both replicates.

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 MCL clustering algorithm.

**Quantification of de novo transcription**

For quantification of de novo transcription in different cell cycle phases (Fig. 4E), cells were labelled with Hoechst dye in culture and pulsed with 5-Ethynyle Uridine (Life Technologies) for one hour and sorted into different cell cycle phases based on DNA content by Hoechst 3342 (Invitrogen) staining. Nascent RNA transcript was enriched using ClickIT® Nascent RNA capture kit (Life Technologies) and subjected to qRT-PCR as described in the Main Experimental Procedures.

**Single cell qRT-PCR**

Single cell gene expression data were collected following the Single-Cell Gene Expression Using SsoFast EvaGreen SuperMix with Low ROX on the Biomark System Advanced Development Protocol 41. Briefly, hESCs were sorted directly with the BD FACS Aria II (BD Bioscience) into 96-well PCR plates containing the RT-STA solution (Life Technologies). Reverse transcription, 20-cycle pre-amplification and exonuclease 1 (NEB) treatment were performed in a thermal cycler prior to 5-fold dilution of pre-amplified products. Amplified samples were quantified with 2X SsoFast EvaGreen Supermix with Low ROX (Bio-Rad) and individual qPCR primers using 48x48 Dynamic Arrays on a Biomark System (Fluidigm). Threshold crossing (Ct) values were calculated using the BioMark Real-Time PCR Analysis software (Fluidigm).
**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 immunostaining for microscopy, cells were fixed with 4% formaldehyde (Sigma) on tissue culture plates (Falcon) and permeabilized using 1% Triton X-100. For γH2AX, cytospinning was performed prior to fixation. All microscope images were taken using Observer Z.1 (Zeiss). To perform immunostaining for FACS, cells were harvested, and then either fixed with 4% formaldehyde (Sigma) and 80% ethanol before staining for cell cycle analysis (pH3, Hoechst), or directly stained for cell death analysis (Annexin V). FACS analyses were done using LSRII flow cytometer (BD). Primary antibodies used are as follows: OCT4 (ab19857, Abcam, 1:5000), NANOG (AF1997, R&D system, 500ng/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), FOXA2 (sc-6554, Santa Cruz, 1:100), p57Kip2 (RB-1637P, NeoMarkers, 1:1000), NES (MAB5326, Millipore, 1:500), GATA4 (sc-25310, Santa Cruz, 1:200), phospho-histone H3 (06-570, Merck, 1:1000), Annexin V (A35110, Life Technologies, 1:20), 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.
Live cell imaging

FUCCI reporter H1 hESCs were generated using lentivirus construct modified from mCherry-hCdt1 and SG2M-Cyan (Clontech). Cells were treated with chemicals or RNAi 24-48hrs prior to live cell imaging. Cells were imaged using Leica SD AF microscope (Leica) for 24hrs.

TUNEL assay

TUNEL assay was performed using the In Situ Cell Death Detection Kit, Fluorescein (Roche) following manufacturer’s instructions.

Teratoma assay and Karyotyping

Assays were performed as previously described (Chia et al., 2010).
Systematic screen in differentiating hESCs under multiple conditions

Identification of gatekeepers of pluripotent state dissolution

Chromatin-modifying complexes
Cell cycle regulators
Cell signalling pathways

Pluripotent state maintenance

Primed for pluripotent state dissolution

ATM/ATR-mediated checkpoint activation
Upregulation of Cyclin B1

M | G1 | S | G2
A Reverse transfection in triplicate 384-well libraries containing 4,558 siRNA pools encompassing:
- Transcription factors
- Epigenetic modifiers
- Kinases
- Phosphatases
- GPCRs

Induction of pluripotent state dissolution
Withdrawal of self-renewal signals
Introduction of differentiation signals

Enrichment for NANOG-GFP positive cells
Recovery in hESC medium
Measurement of total GFP signal over number of nuclei per well (68,370 data points measured)

NANOG promoter
GFP
NANOG-GFP reporter hESC

B

C

D

E

Figure 1
Figure 4
Figure 5
Figure 6
Figure 7
Supp Figure 3
Supp Figure 5