CRRAO Advanced Institute of Mathematics, Statistics and Computer Science (AIMSCS)

Research Report

Author (s): **Sailu Yellaboina et al.**

Title of the Report:

Integrative framework for identification of key cell identity genes uncovers determinants of ES cell identity and homeostasis

Research Report No.: **RR2014-10**

Date: **April 10, 2014**

Prof. C R Rao Road, University of Hyderabad Campus,
Gachibowli, Hyderabad-500046, INDIA.
www.crraoaimscs.org

Integrative framework for identification of key cell identity genes uncovers determinants of ES cell identity and homeostasis

Senthilkumar Cinghu[a,1], Sailu Yellaboina[a,b,c,1], Johannes M. Freudenberg[a,b], Swati Ghosh[a], Xiaofeng Zheng[d], Andrew J. Oldfield[a], Brad L. Lackford[d], Dmitri V. Zaykin[b], Guang Hu[d,2], and Raja Jothi[a,b,2]

[a]Systems Biology Section and [d]Stem Cell Biology Section, Laboratory of Molecular Carcinogenesis, and [b]Biostatistics Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709; and [c]CR Rao Advanced Institute of Mathematics, Statistics, and Computer Science, Hyderabad, Andhra Pradesh 500 046, India

Edited by Norbert Perrimon, Harvard Medical School and Howard Hughes Medical Institute, Boston, MA, and approved March 17, 2014 (received for review October 2, 2013)

Identification of genes associated with specific biological phenotypes is a fundamental step toward understanding the molecular basis underlying development and pathogenesis. Although RNAi-based high-throughput screens are routinely used for this task, false discovery and sensitivity remain a challenge. Here we describe a computational framework for systematic integration of published gene expression data to identify genes defining a phenotype of interest. We applied our approach to rank-order all genes based on their likelihood of determining ES cell (ESC) identity. RNAi-mediated loss-of-function experiments on top-ranked genes unearthed many novel determinants of ESC identity, thus validating the derived gene ranks to serve as a rich and valuable resource for those working to uncover novel ESC regulators. Underscoring the value of our gene ranks, functional studies of our top-hit Nucleolin (*Ncl*), abundant in stem and cancer cells, revealed *Ncl*'s essential role in the maintenance of ESC homeostasis by shielding against differentiation-inducing redox imbalance-induced oxidative stress. Notably, we report a conceptually novel mechanism involving a Nucleolin-dependent Nanog-p53 bistable switch regulating the homeostatic balance between self-renewal and differentiation in ESCs. Our findings connect the dots on a previously unknown regulatory circuitry involving genes associated with traits in both ESCs and cancer and might have profound implications for understanding cell fate decisions in cancer stem cells. The proposed computational framework, by helping to prioritize and preselect candidate genes for tests using complex and expensive genetic screens, provides a powerful yet inexpensive means for identification of key cell identity genes.

pluripotency | RNA-binding protein | transcription | ROS | computational biology

Cell identity is governed by a set of key regulators, which maintain the gene expression program characteristic of that cell state while restricting the induction of alternate programs that could lead to a new cell state. Identification of cell identity genes is a fundamental step toward understanding the mechanisms that underlie cellular homeostasis, differentiation, development, and pathogenesis. RNAi-based high-throughput screening has become a widely used method for identification of new components of diverse biological processes, including signal transduction, cancer, and host cell responses to infection (1, 2). Genome-scale RNAi screens have led to identification of tumor suppressors (3), oncogenes (4), therapeutic targets (5), and regulators of ES cell (ESC) maintenance (6–10), tissue regeneration (11), viral infection (12), and antiviral response (13).

Despite the success of RNAi screens, false discovery and sensitivity remain a significant and difficult problem to address, with surprisingly small overlap among screen hits from independent but related screens (1). For example, multiple genome-scale RNAi screens for host proteins required for HIV infection/replication resulted in a limited overlap among screen hits at the gene level (1). Similarly, screens performed in mouse ESCs (mESCs) for genes essential for the maintenance of ESC identity resulted in only ∼8% overlap (8, 9), although many of the unique hits in each screen were known or later validated to be real. The lack of concordance suggest that these screens have not reached saturation (14) and that additional genes of importance remain to be discovered.

Motivated by the need for an alternative approach for identification of key cell identity genes, we developed a computational approach for systematic integration of published gene expression data to rank-order genes based on their likelihood of defining a phenotype of interest. We demonstrate the effectiveness of the proposed approach by rank-ordering all mouse genes based on how likely they are to determining ESC identity. Not surprisingly, our analysis correctly ranked known pluripotency-associated genes atop the list. Most importantly, using RNAi-mediated loss-of-function experiments, we were able to unearth many novel determinants of ESC identity including several components of functionally distinct complexes. To ascertain the utility of the gene ranks to serve as a rich and valuable resource for those working to uncover novel ESC regulators, we characterized one of our hits, Nucleolin, as having a mechanistic

Significance

A key step to understanding a phenotype of interest is the identification of genes defining that phenotype. We propose a computational framework for a systematic integration of published gene expression data to identify genes defining a cell identity of interest. We demonstrate the utility of the proposed approach by identifying genes essential for the maintenance of ES cell (ESC) identity. Follow-up functional studies on candidate gene Nucleolin (*Ncl*) reveal *Ncl*'s essential role in the maintenance of ESC homeostasis. Ncl deficiency increases endogenous reactive oxygen species levels and induces p53 activity, resulting in p53-mediated suppression of *Nanog* and subsequent ESC differentiation. These studies uncover a previously unknown regulatory circuitry involving genes associated with traits in both ESCs and cancer.

Author contributions: S.C., S.Y., and R.J. designed research; S.Y. developed the meta-analysis framework; S.C., S.Y., S.G., X.Z., A.J.O., B.L.L., G.H., and R.J. performed research; S.C., S.Y., D.V.Z., G.H., and R.J. contributed new reagents/analytic tools; S.C., S.Y., J.M.F., and R.J. analyzed data; and S.C., S.Y., and R.J. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The microarray gene expression data generated for this study have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE47872).

[1]S.C. and S.Y. contributed equally to this work.

[2]To whom correspondence may be addressed. E-mail: jothi@mail.nih.gov or hug4@niehs.nih.gov.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1318598111/-/DCSupplemental.

role in the maintenance of ESC homeostasis by shielding against differentiation-inducing oxidative stress. Our methodology not only enables identification of key cell identity genes but also provides a powerful, yet inexpensive, framework to preselect or prioritize candidate genes for tests using complex and expensive genetic screens.

Results

Approach to Identify Determinants of ESC Identity. Our approach leverages relative gene expression across various cell types or states from independent perturbation experiments (genetic, treatment, differentiation, etc.) to rank-order genes based on how likely they are to have a role in the maintenance of the cell identity of interest. We used this approach to identify genes with possible roles in the maintenance of ESC identity. ESCs can self-renew indefinitely and can differentiate into all derivatives of the three germ layers, making them an attractive model for regenerative medicine and disease modeling (15–17). Successful development of ESC-based therapies, however, largely depends on our understanding of the genes and pathways that constitute the genetic network governing ESC self-renewal and differentiation. Focused functional studies, over the last two decades, have established *Oct4*, *Sox2*, and *Nanog* as the core transcription factors (18–21), with epigenetic features (22–24), miRNAs (25–27), and telomere maintenance (28) playing key roles in the establishment and the maintenance of the pluripotent state in ESCs. Despite the elucidation of many genes and pathways critical for the maintenance of the pluripotent state, the mechanisms that coordinate the activities of master regulators, key signaling pathways, and epigenetic features remain poorly understood, owing largely to incomplete characterization of the genetic network underlying ESCs.

RNAi-based screens of nearly all genes in mouse and human ESCs have collectively revealed more than 400 genes with roles in ESC maintenance (6–10, 29). However, each screen identified a different set of genes, with limited overlap (Fig. 1*A*). The presence of unique hits in each screen suggest that the screens have not reached saturation (14) and that additional genes essential for ESC self-renewal and pluripotency remain to be discovered.

A large number of genes with critical roles in ESC maintenance are highly expressed in ESCs compared with most other cell types and are significantly down-regulated during the normal course of differentiation. In contrast, differentiation and developmental regulators are either silent or basally expressed in ESCs but get activated or significantly up-regulated, respectively, during differentiation. Exploiting this characteristic feature of cell identity genes, to better characterize the genetic network controlling ESC identity, we developed a bioinformatics framework for systematic integration of published microarray gene expression data from 68 experiments profiling undifferentiated mESCs and various differentiated cell (DC) types (Dataset S1). To get around having to normalize datasets to control for experimental noise and laboratory effects, before integration, expression data for every mESC-DC pairing was analyzed separately using the nonparametric rank-product approach (30), and the genes on the array were rank-ordered based on the *P* values of their associated ranks of expression fold change in DCs vs. mESCs (Fig. 1*B*). Rank-ordered gene lists from 68 mESC-DC comparisons were integrated using a statistical framework to obtain a single rank-ordered gene list in which genes were ordered by their consensus ranks (*Materials and Methods*). The smaller the consensus gene rank, the more severely and consistently that gene was down-regulated during differentiation across experiments, and the more likely that gene has a potential role in the maintenance and/or the establishment of the pluripotent state in mESCs. In contrast, genes with larger ranks are likely to have roles during differentiation and development but not in ESC maintenance.

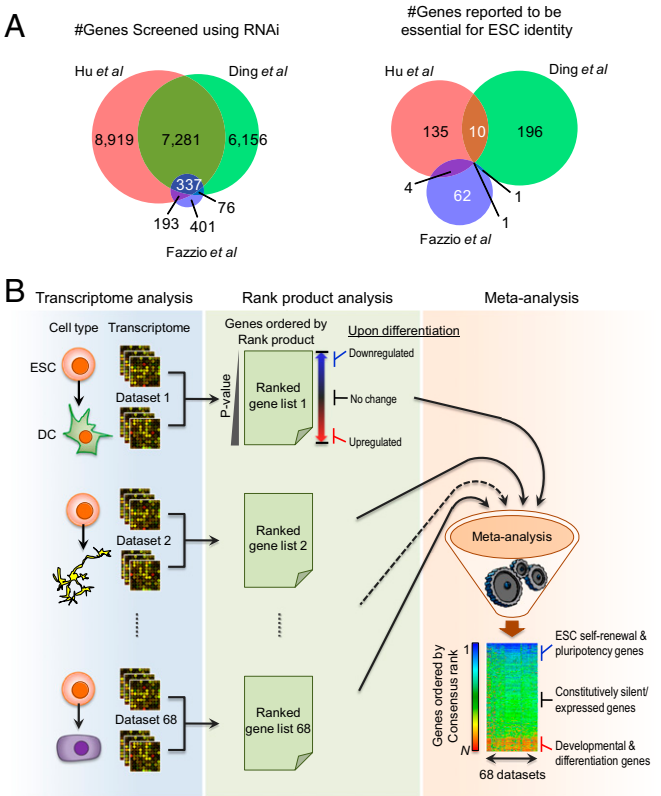

Fig. 1. An integrative approach to identify determinants of ESC identity. (*A*) Venn diagrams summarizing results from RNAi screen experiments in mESCs by Hu et al. (8), Fazzio et al. (7), and Ding et al. (9). (*B*) A schematic of the meta-analytic framework that integrates published gene expression microarray data in mESCs and differentiated cell types (DC) to generate a ranked gene list in which genes are rank-ordered by decreasing likelihood of their role in the maintenance of mESC identity.

ESC Identity Genes Are Enriched Atop the Ranked List. Integrating evidence from multiple datasets using the proposed framework proved to be extremely effective: master ESC regulator *Oct4* is ranked number one, followed by *Nanog* and *Sox2* (Fig. 2*A* and Dataset S2). Moreover, several other regulators that have been implicated in ESC maintenance including *Nr0b1*, *Trim28*, *Pcgf6*, *Sall4*, *Tet1*, *Esrrb*, *Tcl1*, *Klf2/4/5*, and *Smc2* were ranked within the top 1%, along with a number of genes that have not been previously implicated in ESC biology (Fig. 2*A* and Dataset S2). Remarkably, several components of functionally distinct biochemical complexes, with known roles in the maintenance of the pluripotent state in ESCs, were ranked in the top 10% including members of the Tip60-p400 chromatin remodeling complex (7), the Ino80 chromatin remodeling complex (7, 8, 10), the Paf1 complex (9), the transcription factor IID (TFIID) complex (31), the ubiquitin-proteosome system (32), the spliceosome complex (10), the mediator complex (33), the COP9 signalosome (10), and the condensin complex (7) (Fig. 2*B* and Fig. S2*A*). Notably, the key components of the ESC-specific BAF complex (esBAF) (34, 35), defined by specialized subunit composition of Brg1 but not Brm, Smarcc1 (BAF155) but not Smarcc2 (BAF170), Arid1a (BAF250a) but not Arid1b (BAF250b), and Smarcd1 (BAF60a) and Actl6a (BAF53a) but not their homologs Smarcd3 (BAF60c) and Actl6b (Baf53b), respectively, were correctly ranked high on the list (Fig. 2*C*). In contrast, components that replace ESC-specific subunits of this complex during differentiation were appropriately ranked at the bottom (Fig. 2*C*).

Gene ontology (GO) analysis of the genes at the top of this ranked gene list revealed enrichment for genes with roles in stem cell maintenance, mRNA processing, and cell cycle (Fig. S1).

Consistent with this analysis, genes that were reported to be essential for mESC maintenance by RNAi studies (7–9) were enriched at the top of the list (Fig. 2*D*). In contrast, genes at the bottom of the list included developmental regulators such as *Meis1*, *Gata3*, *Hand1*, and *Actl6b* (Fig. 2*A*) and were enriched in the processes of differentiation and development (Fig. S1). Furthermore, binding sites for transcriptional regulators of mESCs including Oct4, Sox2, Nanog, Stat3, and Brg1 were preferentially enriched at the promoters of genes ranked atop the list, whereas the Polycomb member Suz12, known to repress developmental and differentiation genes in ESCs (22), preferentially bound to promoters of genes at the bottom of the list (Fig. 2*E*). Consistent with Polycomb occupancy, genes near the bottom of the list were marked by the repressive histone modification H3K27me3 in mESCs and lacked signs of gene activity, including elongating RNA Polymerase II and the H3K36me3 modification (Fig. S2*B*). Together, these data illustrate the effectiveness of our approach in identifying genes and complexes defining the pluripotent state and suggest that genes ranked atop the list are good candidates for identification of hitherto unknown regulators of mESC identity.

RNAi Validation Uncovers Novel ESC Self-Renewal Genes. To identify novel determinants of mESC identity, we performed RNAi-mediated loss-of-function experiments on 49 candidate genes including 38 chosen from the top 2% of the list (31 from the top 1%) and 11 others that form complexes with candidate genes ranked in the top 2% (Table S1). We examined the colony morphology and alkaline phosphatase (AP) staining 96 h after transfection of mESCs with siRNAs targeting each candidate. Depletion of 17 candidates showed morphological changes and loss of AP staining consistent with mESC differentiation, suggesting that these genes are essential for mESC maintenance (Fig. 3*A* and Fig. S3). Although the depletion of the remaining 32 genes did not exhibit obvious/consistent self-renewal maintenance defects, we cannot rule out the possibility that at least some of them are essential for ESC differentiation [e.g., Utf1 (36) and Eras (37)] and/or the establishment of the pluripotent state, attributes not assessed by our self-renewal assay.

The 17 positive hits included genes coding for Nucleolin (Ncl), all three members of the trimeric transcription factor complex NFY (NfyA, NfyB, and NfyC), components of the DNA replication machinery including the origin recognition complex (ORC) and the minichromosome maintenance (MCM) complex, bromodomain-containing protein Brd4, Sin3/HDAC1-associated protein Fam60a, and an uncharacterized protein Fam169a. The fact that multiple components of the same complex exhibit the same unique knockdown (KD) phenotype (e.g., NFY, ORC, and MCM) not only demonstrates the legitimacy of our hits, but also the value of our ranked gene list in identifying new protein complexes and regulatory pathways with possible roles in ESC maintenance.

Quantitative RT-PCR (RT-qPCR) experiments confirmed that the observed morphological changes are consistent with the molecular changes associated with the differentiation phenotype (Fig. 3*B*). RNAi-mediated depletion of all hits but *Orc1l* led to a significant down-regulation of key pluripotency regulators including *Oct4* and *Nanog*. Consistent with their role in the maintenance of mESCs, all 17 hits are expressed at much higher levels in mESCs compared with mouse embryonic fibroblasts (MEFs) (Fig. 3*C*) and are largely down-regulated during differentiation induced by embryoid body (EB) formation, retinoic acid (RA) treatment, or leukemia inhibitory factor (LIF) withdrawal (Fig. 3 *D*–*F*). Together, these data support the conclusion that the 17 hits identified from the screen are essential for mESCs to maintain their self-renewal characteristics.

Nucleolin Is Essential to Maintain ESC Identity. To demonstrate how valuable the derived gene ranks would be for those looking to uncover novel ESC regulators, as a proof of concept, we focused on Nucleolin (*Ncl*), the top-ranked hit, whose depletion resulted in the most severe cellular and molecular phenotype (Fig. 3 *A* and *B*). Ncl is a highly conserved protein primarily localized in the nucleolus of nearly all cell types, but is highly abundant in stem and cancer cells (38, 39). To elucidate the role for Ncl in ESC maintenance, we performed microarray analysis of control and *Ncl* KD mESCs 96 h after siRNA transfection. Two siRNAs targeting *Ncl* were used to ensure that the observed expression changes are due to *Ncl* depletion and not due to siRNA off-target effects (Fig. 4*A*). RT-qPCR and immunostaining confirmed gene expression changes observed in the microarray analysis (Fig. 4*B* and Fig. S4 *A* and *B*). Pluripotency genes including *Oct4*, *Nanog*, *Sox2*, *Tcl1*, *Tet1*, and nodal antagonists *Lefty1* and *Lefty2*, which are among the earliest to be down-regulated during mESC differentiation, were significantly down-regulated in *Ncl* KD cells. Additionally, several markers of early differentiation including *Cdx2*, *Gata3*, *Gata6*, and *Sox17* were significantly up-regulated in *Ncl*-depleted cells (Fig. 4*B*), consistent with the observed differentiation phenotype. Based on these data, we conclude that *Ncl* is essential to maintain mESCs in an undifferentiated pluripotent state and that depletion of *Ncl* in mESCs induces expression of early differentiation markers.

Nucleolin Inhibits p53-Mediated Suppression of Nanog. To probe the mechanisms underlying Ncl's essential role in the maintenance of the pluripotent state, we performed Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of differentially expressed genes in *Ncl*-depleted mESCs. Our analysis revealed a significant enrichment for p53 and MAPK/Erk signaling pathway genes among those that were up-regulated (Fig. 4*C*), suggesting a potential role for *Ncl* in suppressing these differentiation-inducing signaling pathways. Activation of p53 and MAPK/Erk signaling in mESCs has previously been shown to promote differentiation via their suppression of *Nanog* (40, 41). Interestingly, we found that global gene expression changes on *Ncl* depletion are similar to those observed after *Nanog* depletion (Fig. 4*D*), suggesting that *Ncl* activity in mESCs may be critical to maintain Nanog expression. This observation prompted us to hypothesize that *Ncl* might be inhibiting endogenous factors that would otherwise induce p53 and Erk activity to suppress *Nanog*.

Toward testing this hypothesis, we first evaluated the expression of total and activated forms of p53 in control and *Ncl* KD mESCs and observed a strong increase in p53 levels on *Ncl* depletion (Fig. 4 *E* and *F*). Established p53 target genes such as *p21*, *Mdm2*, *Gadd45g*, and *Noxa* were significantly up-regulated in the *Ncl*-depleted mESCs, indicating that the activated p53 is functional (Fig. S4*C*). Consistent with the activation of the p53 pathway genes, as previously reported (42), proliferation defects and impaired cell cycle progression were evident in *Ncl*-depleted cells (Fig. S4 *D* and *E*). Notably, we detected an increase in p53 protein but not mRNA levels on Ncl depletion in mESCs (Fig. 4 *E* and *F* and Fig. S4*C*), indicating that *Ncl*'s negative regulation of p53 is likely posttranscriptional. Ncl has previously been shown to regulate gene expression through direct interactions with RNA (38, 39). Therefore, we tested whether Ncl's repression of p53 in mESCs is direct via its RNA-binding activity. Indeed, RNA immunoprecipitation (RNA-IP) using an antibody against Ncl in mESCs revealed Ncl binding to *p53* mRNA and the known Ncl target *Bcl2* (43), but not the control *Actin* (Fig. 4*G*). These findings, consistent with Ncl's repression of p53 in breast cancer cells (44), led us to postulate that p53 activation could be a contributing factor to differentiation in *Ncl*-depleted mESCs.

Because p53 plays an active role in promoting differentiation of human ESCs (45) and acts as a barrier during reprogramming of somatic cells into induced pluripotent stem cells (iPSCs) (46), we next investigated whether inactivating p53 can rescue the *Ncl* depletion phenotype. Indeed, depletion of *p53* in combination with *Ncl* in mESCs largely rescued the cellular and molecular changes observed in mESCs depleted with *Ncl* (Fig. 4*H* and Fig. S4*F*). The incomplete rescue, however, suggested that there may

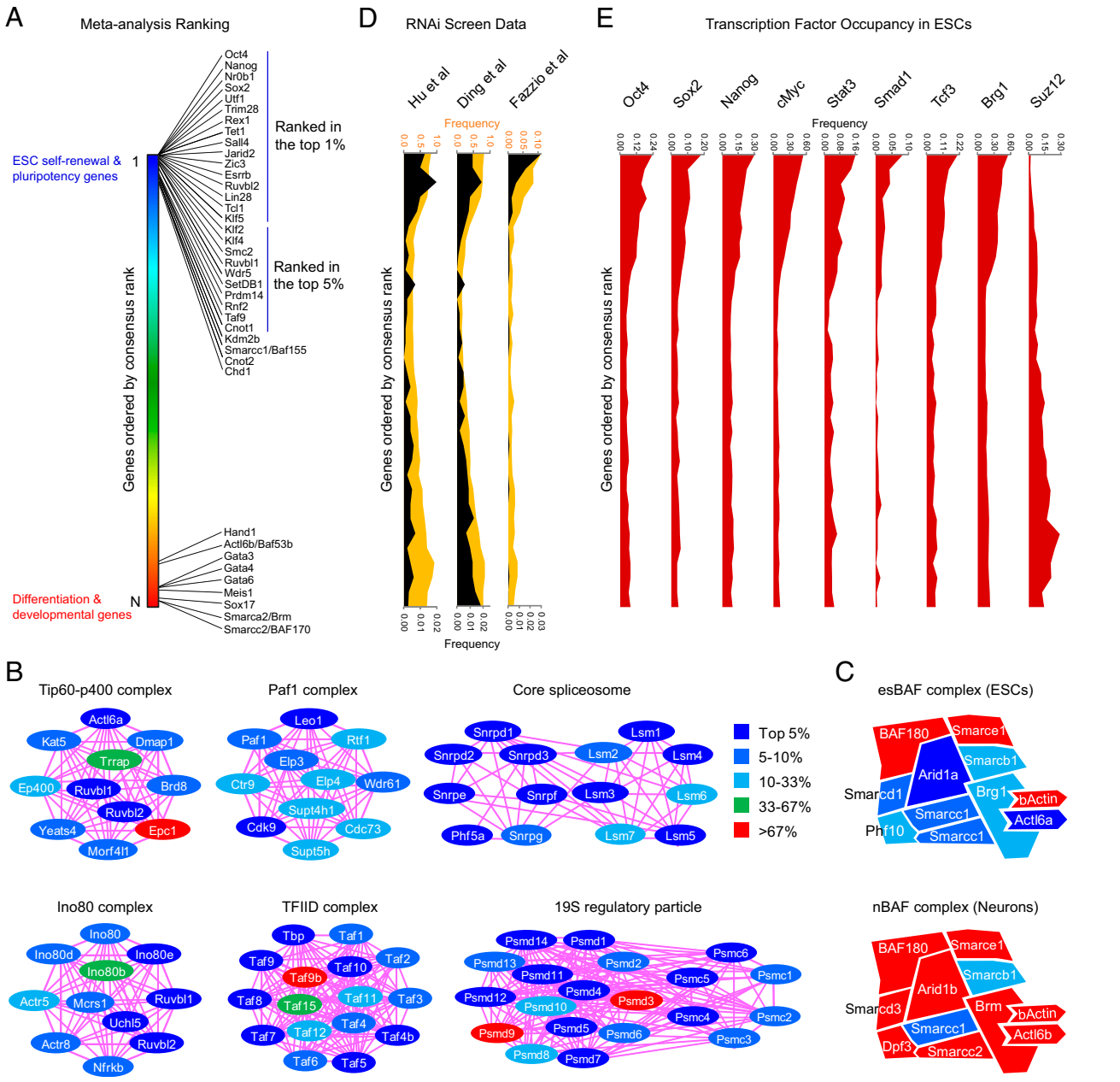

Fig. 2. Gene ranks are good indicators of pluripotency-associated genes. (*A*) Enrichment for known pluripotency-associated factors among genes ranked atop the rank-ordered gene list obtained from meta-analysis in Fig. 1*B*. Known developmental and differentiation regulators are among genes ranked at the bottom. (*B*) Components of the Tip6-p400 complex, Paf1 complex, core spliceosome, Ino80 complex, TFIID complex, and the 19S regulatory particle and their ranking on the list shown in *A*. (*C*) Components of the ESC-specific BAF complex (esBAF) and the neuronal BAF (nBAF) complex and their ranking on the list shown in *A*. Color key same as in *B*. (*D*) Frequency distribution of genes screened (orange) and genes reported to be essential (black) in RNAi screen experiments by Hu et al. (8), Fazzio et al. (7), and Ding et al. (9) in relation to the consensus gene ranks. (*E*) Frequency distribution of binding occupancy of pluripotency-associated transcription regulators Oct4, Sox2, Nanog, cMyc, Stat3, Smad1, Tcf3, Brg1, and Suz12 in relation to the consensus gene ranks.

be other affected pathways in *Ncl*-depleted cells that p53 depletion is unable to rescue. We thus focused on the MAPK/Erk pathway, which becomes activated on *Ncl* depletion (Fig. 4*C* and Fig. S4*G*). Inhibition of the MAPK/Erk signaling cascade using selective small-molecule inhibitor PD0325901 (1 μM) partially restored the morphological changes associated with the differentiation phenotype, but not the expression of differentiation markers (Fig. S4 *G–I*), suggesting a metastable state due to neutralizing conflicts between the self-renewal and the differentiation programs (41). Alternatively, the outcome could be due to impairment of differentiation commitment in the absence of MAPK/Erk signaling. Interestingly, a combination of Erk inhibition and p53 depletion fully restored the morphological and molecular changes in Ncl-depleted mESCs (Fig. S4 *H* and *I*). Collectively, these results support the conclusion that differentiation due to *Ncl* depletion is dependent on p53.

Given the mutually exclusive expression patterns of p53 and Nanog in Control and *Ncl* KD mESCs (Fig. 4*F*) and that stress-induced p53 in mESCs is known to bind to the *Nanog* enhancer and suppress its transcription (40, 47) (Fig. 4*I*), we asked whether p53, activated on *Ncl* depletion, can bind to and suppress *Nanog*. To test this, we took advantage of a luciferase-based reporter assay. As a positive control for p53 activation on *Ncl* depletion, we used a reporter construct with a minimal p21 promoter

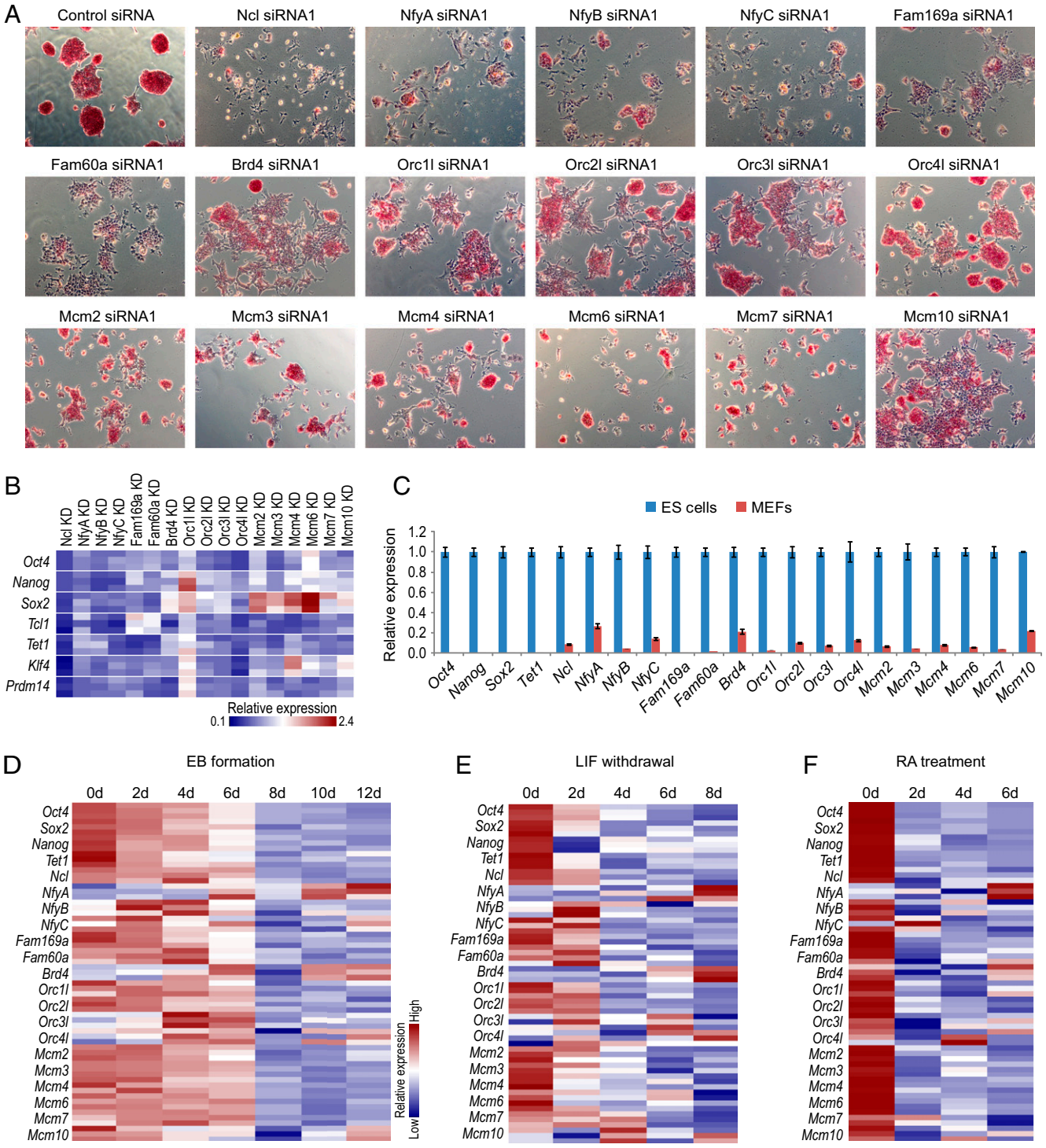

Fig. 3. Validation of candidate self-renewal genes. (*A*) Morphology and alkaline phosphatase staining of mESCs 96 h after transfection with control siRNA or a siRNA targeting a candidate gene. Representative images from at least three experiments are shown. (*B*) RT-qPCR analysis of relative mRNA levels of selected stem cell marker genes in control KD and *Ncl* KD mESCs 96 h after siRNA transfection. The mRNA level in control mESCs is set as 1. Expression changes from three experiments are shown. (*C*) RT-qPCR analysis of relative mRNA levels of candidate genes in mESCs and mouse embryonic fibroblasts (MEFs). The mRNA level in mESCs is set as 1. Data are normalized to *Actin*. Error bars represent SEM of three experiments. (*D–F*) RT-qPCR analysis of relative mRNA levels of candidate genes during embryoid body (EB) formation (*D*) and during differentiation induced by leukemia inhibitory factor (LIF) withdrawal (*E*) and retinoic acid (RA) treatment (*F*). Data are normalized to *Actin*. Data from three experiments are shown.

containing the p53 binding motif. *Ncl* depletion in mESCs drove the luciferase activity of the minimal *p21* promoter, thus confirming p53 activation (Fig. 4*J*). Using a reporter construct containing the *Nanog* promoter and enhancer, we observed ∼10-fold reduction in the luciferase activity in *Ncl*-depleted mESCs (Fig. 4*J*), indicating that p53 activated on *Ncl* depletion can suppress *Nanog* expression. If the differentiation caused by *Ncl* depletion is due to p53-mediated suppression of *Nanog*, we

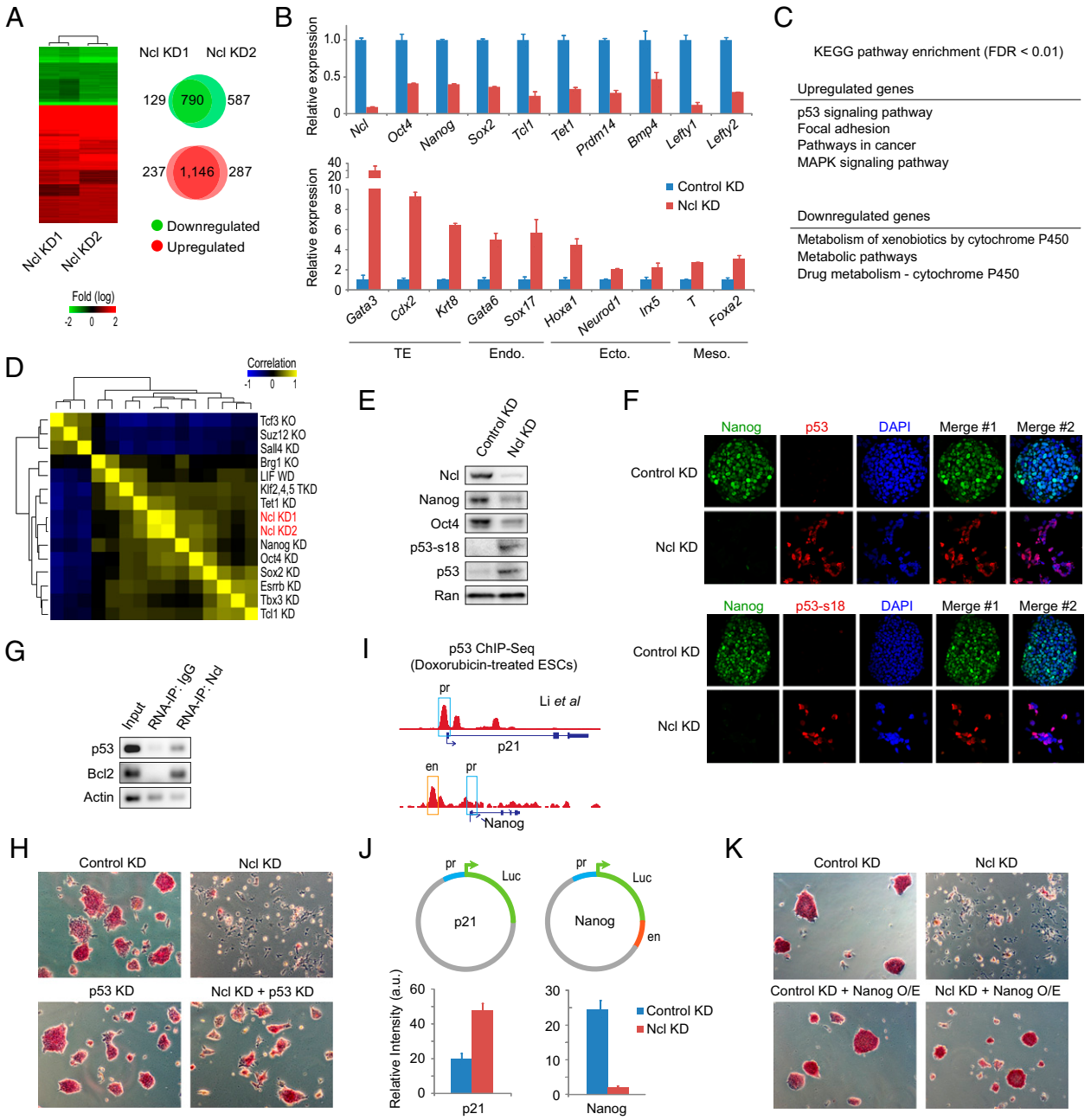

Fig. 4. Nucleolin inhibits differentiation-inducing p53-mediated suppression of Nanog to maintain mESCs in the undifferentiated pluripotent state. (*A*) Heatmap representation (*Left*) of microarray gene expression fold changes due to *Ncl* knockdown (KD), measured 96 h after transfection of two different siRNAs (KD1 and KD2). Only genes that were differentially expressed (FDR $\leq$ 0.01 and fold-change $\geq$ 2) in KD1 and/or KD2 are represented. Venn diagrams (*Right*) show the overlap between the number of up- and down-regulated in KD1 vs. KD2. (*B*) RT-qPCR analysis of relative mRNA levels of selected stem cell and differentiation marker genes in control KD and *Ncl* KD mESCs 96 h after siRNA transfection. The mRNA level in control KD cells is set as 1. Data are normalized to *Actin*. Error bars represent SEM of three experiments. TE, trophectoderm; Endo., endoderm; Ecto., ectoderm; Meso., mesoderm. (*C*) KEGG pathway enrichment analysis of genes up- and down-regulated on *Ncl* KD. (*D*) Correlation between global gene expression changes due to *Ncl* KD and those observed after KD or KO of other pluripotency-associated factors, as reported in other studies. Rows/columns are ordered based on unsupervised hierarchical clustering. TKD, triple KD; WD, withdrawal. (*E*) Western blot analysis of Ncl, total and phosphorylated p53 (p53 and p53-s18, respectively), and selected stem cell factors (Nanog, Oct4) in control KD and *Ncl* KD mESCs 96 h after siRNA transfection. Ran is used as a loading control. Representative blots from three experiments are shown. (*F*) Coimmunostaining of Nanog and total p53 (*Upper*), and Nanog and phosphorylated p53 (*Lower*) in control KD and *Ncl* KD mESCs 96 h after siRNA transfection. Nuclei were counterstained by DAPI. Merge #1, Nanog+p53; merge #2, Nanog+p53+DAPI. Representative images from three experiments are shown. (*G*) PCR analysis of mESC RNA, immunoprecipitated with Ncl or control IgG antibody. *Bcl2* and *Actin* are used as positive and negative controls, respectively. Representative gel images from three experiments are shown. (*H*) Morphology and alkaline phosphatase (AP) staining of control KD, *Ncl* KD, *p53* KD, and *Ncl* KD + *p53* KD mESCs 96 h after siRNA transfection. Representative images from three experiments are shown. (*I*) ChIP-Seq data from Li et al. (47) showing p53 binding at *p21* promoter and *Nanog* enhancer in doxorubicin-treated mESCs. pr, promoter; en, enhancer. (*J*) (*Upper*) Luciferase reporter constructs cloned with *p21* promoter, and *Nanog* promoter and enhancer regions, as highlighted in *I*. (*Lower*) Luciferase activity of the reporter constructs in control KD and *Ncl* KD mESCs 96 h after siRNA transfection. Error bars represent SEM of three experiments. (*K*) Morphology and AP staining of control KD and *Ncl* KD mESCs, with and without exogenous *Nanog* overexpression (O/E), 96 h after siRNA transfection. Representative images from three experiments are shown.

reasoned that Nanog overexpression should rescue the differentiation phenotype. As postulated, overexpression of exogenous *Nanog* largely rescued the phenotype observed in *Ncl*-depleted mESCs (Fig. 4*K* and Fig. S4*J*). Together, these results confirmed our hypothesis that the differentiation due to *Ncl* deficiency is at least in part due to the activation of the p53 signaling, and subsequent suppression of *Nanog*.

Nucleolin Shields ESCs from Differentiation-Inducing Oxidative Stress. Given that the p53 levels in cells are tightly regulated, the nuclear accumulation of activated p53 in *Ncl*-depleted mESCs prompted us to investigate mechanisms underlying p53 activation. Typically, p53 is primarily localized in the cytosol. In response to stress stimuli, p53 is phosphorylated and translocated to the nucleus (40). In the absence of exogenous stress-inducing agents, we speculated that endogenous reactive oxygen species (ROS), natural byproducts of cellular respiration, might be activating p53 in *Ncl*-depleted cells. Elevated levels of ROS, caused by the cell's inability to efficiently detoxify ROS, have been well documented to induce p53 activity (48), and ROS levels are thought to be low in mESCs grown under normoxic conditions (~20% oxygen). Notably, the proliferation and growth of ESCs are known to be sensitive to oxygen levels, with propagation maximized under hypoxic conditions (~5% oxygen) or in the presence of antioxidants and free-radical scavengers (49), but mechanisms that maintain the normal redox state in ESCs is poorly understood.

Quantification of endogenous ROS levels revealed a significant increase in the *Ncl*-depleted mESCs (Fig. 5*A*). To probe the link between *Ncl* deficiency and elevated ROS levels, we turned our attention to endogenous antioxidant proteins that are known to scavenge ROS. In particular, we focused on selenoproteins, a subset of which were recently shown to be positively regulated by *Ncl* in cancer cells through posttranscriptional recoding of the mRNA during translation (43). RNA-IP experiments in mESCs confirmed Ncl binding to mRNAs of Ncl-dependent selenoprotein-coding genes *Txnrd1* (thioredoxin reductase 1) and *Gpx7* (glutathione peroxidase 7), but not the Ncl-independent selenoprotein-coding genes *Gpx1* and *Sep15* (43) (Fig. 5*B*). Consistent with *Ncl*'s role in the posttranscriptional regulation of these proteins, *Ncl* deficiency significantly decreased Txnrd1 and Gpx7 expression at the protein level but not at the mRNA level (Fig. 5*C* and Fig. S5*A*). These data indicated that reduced levels of Ncl-dependent antioxidant selenoproteins could contribute significantly to elevated ROS levels in *Ncl*-depleted mESCs.

If elevated ROS levels were the cause for p53 activation and subsequent differentiation in *Ncl*-depleted mESCs, we hypothesized that adding antioxidants to the culture medium should at least partly restore ROS levels and rescue the differentiation phenotype. We repeated the *Ncl* KD experiments in mESCs with and without antioxidant ascorbic acid (vitamin C; 15 μM) or Trolox (6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid; a water-soluble analog of vitamin E) in the culture medium. As predicted, *Ncl*-depleted cells grown in the presence of antioxidants exhibited a partial restoration of the ROS levels, cellular morphology, AP staining, and molecular changes (Fig. 5 *D* and *E* and Fig. S5 *B* and *C*), indicating that ROS-scavenging antioxidants can attenuate ROS and partially rescue the overall phenotype. These results are consistent with previous reports showing antioxidant addition improving reprogramming efficiency during iPSC generation (50), and indicate that elevated ROS levels are a cause and not a consequence for differentiation in *Ncl*-depleted mESCs.

To establish that elevated ROS levels are capable of inducing differentiation in mESCs, we treated mESCs with buthionine sulfoximine (BSO) to induce high ROS levels. BSO is a potent and selective inhibitor of γ-glutamyl cysteine synthetase, a key enzyme in glutathione biosynthesis, used to block cellular resistance to chemotherapy by reducing the levels of ROS-scavenging glutathione proteins (51). mESCs treated with BSO (2 mM) exhibited ~2.5-fold increase in ROS levels (Fig. S5*D*), similar to that observed in Ncl-depleted mESCs (Fig. 5*D*), accompanied by loss of colony morphology, down-regulation of pluripotency markers, up-regulation of early lineage markers, and p53 activation, all consistent with differentiation (Fig. S5 *E*–*G*). These data support the conclusion that elevated ROS levels in Ncl-depleted mESCs is the primary cause, if not a contributing factor, for the loss of the pluripotent state and that *Ncl* maintains cellular homeostasis in mESCs by shielding against redox imbalance-inducible differentiation.

Nanog Transcriptionally Regulates Nucleolin. Last, to probe how *Ncl*'s expression is regulated in mESCs, we examined published transcription factor ChIP-Seq data in mESCs (18) and found Nanog targeting a site in the 3′UTR region of *Ncl* that is evolutionarily conserved in many mammals (Fig. 6*A* and Fig. S6). ChIP using an antibody against Nanog followed by RT-qPCR analysis confirmed this binding. To determine if Nanog binding at this site has any influence on *Ncl* expression, we evaluated *Ncl* expression in mESCs overexpressing or depleted of *Nanog*. *Nanog* overexpression in mESCs led to increased Nanog occupancy at the Ncl 3′UTR site and a more than threefold increase in *Ncl* expression (Fig. 6 *A* and *B*). In contrast, *Nanog* depletion in mESCs abolished Nanog binding and reduced *Ncl* expression (Fig. 6 *A* and *B*). These data strongly indicate that Nanog transcriptionally regulates *Ncl* in mESCs by binding to an evolutionarily conserved *cis*-acting regulatory element in its 3′UTR.

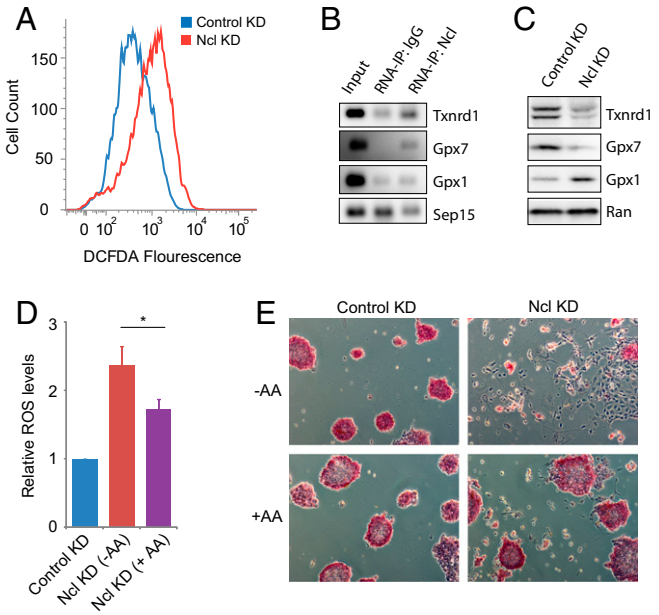

Fig. 5. Elevated ROS levels in Ncl-depleted mESCs contribute to the loss of the pluripotent state. (*A*) Intracellular ROS levels measured by flow cytometry using DCFDA fluorescence in control KD and *Ncl* KD mESCs 96 h after siRNA transfection. Representative data from three experiments are shown. (*B*) PCR analysis of mESC RNA, immunoprecipitated with Ncl or control IgG antibody, probed for selenoproteins *Txnrd1*, *Gpx7*, *Gpx1,* and *Sep15*. Representative gel images from three experiments are shown. (*C*) Western blot analysis of Txnrd1, Gpx7, and Gpx1 in control KD and *Ncl* KD mESCs 96 h after siRNA transfection. Ran is used as a loading control. Representative blots from three experiments are shown. (*D*) Intracellular ROS levels measured by flow cytometry using DCFDA in control KD and *Ncl* KD mESCs, grown in the absence or presence of ascorbic acid (AA), 96 h after siRNA transfection. Error bars represent SEM of three experiments. $^*P < 0.01$. (*E*) Morphology and alkaline phosphatase staining of control KD and *Ncl* KD mESCs, grown in the absence or presence of AA (−AA and +AA, respectively), 96 h after siRNA transfection. Representative images from three experiments are shown.

Discussion

Genome-scale RNAi screens have become part of the standard experimental repertoire for identification of key genes associated with a phenotype of interest. A typical screen involves a primary screen to identify positive hits, followed by a secondary validation screen of a different sort to distinguish the real hits from the false positives. A survey of recent RNAi screening studies revealed that most of the primary hits (~60–80% on average) are false positives, attributable to off-target effects of RNAi reagents and statistical cutoffs used to call high-confidence hits (1). Efforts to minimize false positives increase the number of false negatives and vice versa, making development of appropriate approaches that will minimize false discovery rates (FDRs) a challenge (1).

We developed a computational approach for identification of key cell identity genes via systematic integration of microarray gene expression datasets from independent perturbation experiments. We applied this approach to rank-order genes in the mouse genome based on their likelihood to have a role in the maintenance of ESCs. Identification of genes and pathways that constitute the pluripotency network is critical to the understanding of the molecular mechanisms controlling the balance between self-renewal and differentiation in ESCs. Although genome-wide RNAi-based screens were instrumental in uncovering many key players essential for the maintenance of ESC identity, discrepancies between the reported hits and the presence of unique hits in each case suggest that the screens have not reached saturation perhaps due to their limitations (e.g., sensitivity of the assay/readout, target-gene specificity or nonspecific off-target effects) and that additional genes essential for the maintenance of ESC identity remain to be discovered. Our approach is effective in identification of ESC identity genes, which is evident from the enrichment of known pluripotency-associated genes atop the ranked list. Furthermore, enrichment of several components of functionally distinct complexes within the top 10% illustrates the method's ability to identify not only individual genes but also complexes controlling ESC identity. RNAi-based validation of a number of top-ranked genes with previously unknown roles in ESC maintenance demonstrates that the gene ranks derived from our approach will serve as a rich and valuable resource to those looking to uncover novel regulators of ESCs.

To underscore the value of our analysis to those seeking to discover novel ESC regulators, we performed detailed investigations of our top hit, *Ncl*, and establish its mechanistic role in the maintenance of cellular homeostasis in mESCs. *Ncl* deficiency increases endogenous ROS levels and induces p53 activity, resulting in p53-mediated suppression of *Nanog* and subsequent mESC differentiation. Ncl, which helps maintain Nanog levels, is itself positively regulated by Nanog at the transcriptional level. Collectively, our findings on *Ncl* support a model (Fig. 6*D*) wherein Nanog's positive regulation of *Ncl* provides the means for Nanog to suppress p53 activity in a Ncl-dependent manner to maintain mESC self-renewal. However, when p53 is activated on depletion of *Ncl*, it can suppress *Nanog* to induce mESC differentiation. Together, this would constitute an Ncl-dependent Nanog-p53 bistable switch regulating the homeostatic balance between self-renewal and differentiation in mESCs. Thus, *Ncl* supports self-renewal in mESCs by shielding against ROS-induced p53 activation, which serves as a barrier not only during iPSC generation but also in tumorigenicity (46).

Exceptional genomic stability is one of the hallmarks of ESCs. Despite an unusually short cell cycle, characterized by a truncated G1 phase (52), ESCs have a much lower rate of spontaneous mutation compared with somatic cells (53), indicating that ESCs likely have stringent mechanisms to maintain genomic stability during stress-inducing rapid proliferation. Paradoxically, p53-dependent stress response pathways that can induce cell cycle arrest and senescence are defective in undifferentiated ESCs (54). Instead, p53 in ESCs has an unorthodox role of inducing differentiation after DNA damage by directly suppressing *Nanog*, leading to spontaneous differentiation of ESCs into other cell types, which can then undergo efficient p53-dependent cell cycle arrest or apoptosis (40). Mechanisms that control p53 activation to permit rapid ESC cycling without compromising the genomic stability remained unclear until recently when posttranslational mechanisms have been shown to restrict differentiation-inducing p53 activity (55, 56). Endogenous ROS, natural byproduct of cellular respiration, is a major source of DNA damage and a substantial factor contributing to genomic instability and accumulation of mutations (57). Redox imbalance-induced oxidative stress, caused by the cell's inability to efficiently detoxify ROS, is well documented to trigger p53 activation in somatic and cancer cells (48), but mechanisms that maintain the normal redox state in ESCs are poorly understood. Our findings on Nucleolin's inhibition of p53 activity through its regulation of ROS levels sheds light on a tightly regulated mechanism that permit rapid proliferation of ESCs without compromising genomic stability.

Many cancers have elevated levels of Ncl (38, 39) and Nanog (58, 59) in addition to impaired p53 signaling, traits reminiscent of those ascribed to ESCs. Notably, poorly differentiated tumors including glioblastomas and bladder carcinomas, which generally have the worst prognoses, show preferential overexpression of ESC-associated factors including Nanog (58). Furthermore, subsets of cancer stem cells (CSCs) in some tumors contain lower ROS levels, akin to normal stem cells (60). Given the parallels between the hallmark traits of ESCs–self-renewal and differentiation capacity–and the proliferative capacity and the phenotypic plasticity of tumor and CSCs (58, 61), our findings on *Ncl* have potential implications for understanding not only tumorigenesis but also regulatory circuitry determining cell fate decisions in CSCs, which are widely believed to possess tumor initiating capabilities (58).

Our approach for identification of cell identity genes is based on the assumption that genes with critical roles in the maintenance of a given cell state are highly expressed in that cell state compared with most others and thus are likely to be significantly down-regulated as the cell transitions into a new state. Although this is a reasonable assumption, and is perhaps true for most cell identity genes, there may be genes with essential roles but whose expression is constitutive (e.g., housekeeping genes). Our approach, which relies on expression differences across various cell types or states, may not be as successful in enriching for such genes as they are likely to be essential in many cell types or states. However, many housekeeping genes are ranked high on our ranked list because of their high expression levels in ESCs (Fig. 2*B* and Fig. S1), suggesting that these genes may either have an ESC-specific function in addition to their housekeeping one or contribute to rapid proliferation (ESCs have an unusually short cell cycle, characterized by a truncated G1 phase), which may require high levels of some housekeeping genes.

In addition to identifying genes that determine a cell's identity, our approach will also enrich for those that define a cell's identity but may not necessarily be essential for its maintenance, e.g., Rex1 (Zfp42), a marker of pluripotency in ESCs. This phenomenon could partly explain why depletion of only 17 of the 49 candidates tested exhibit ESC maintenance defects. Some of the remaining 32 genes, whose depletion did not result in obvious self-renewal defects, are known to be essential for ESC differentiation and/or the establishment, but not the maintenance, of the pluripotent state, attributes we did not test for. Nevertheless, the success rate (~35%) of our approach compares well with the percentage of primary hits (~20–40%) that are validated in the secondary screens (1). Thus, our approach provides a powerful yet inexpensive means for identification of key cell identity genes.

Although we used the approach for the identification of determinants of ESC identity, the general framework can readily be used to identify genetic determinants of any cellular type or state of interest, including in studies seeking to identify susceptibility genes in cases vs. controls, as long as sufficient high-quality datasets are available. The proposed computational framework, by helping to prioritize and preselect candidate

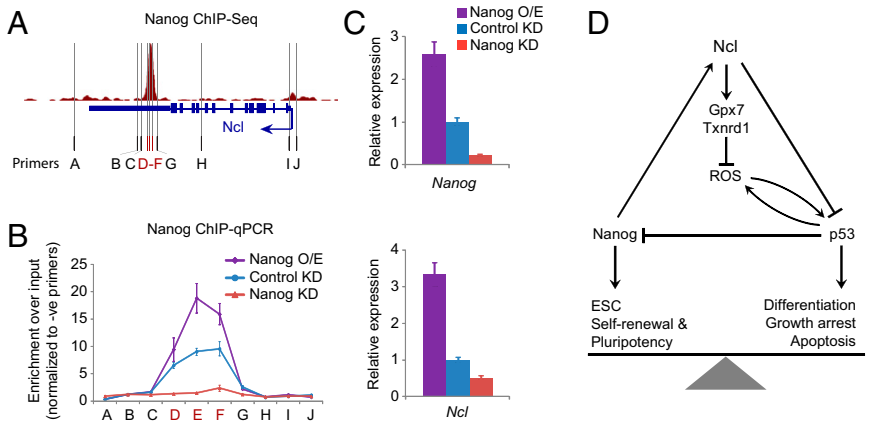

Fig. 6. Nanog regulates Nucleolin. (*A*) ChIP-Seq data from Chen et al. (18) showing Nanog occupancy within the 3′ UTR of *Ncl*. (*B*) ChIP-qPCR analysis using antibody against Nanog in control KD, *Nanog* KD, and *Nanog* overexpressing (O/E) mESCs. Immunoprecipitates were probed with 10 primer pairs (A–J) located across the *Ncl* locus. Error bars represent SEM of three experiments. (*C*) RT-qPCR analysis of relative *Nanog* and *Ncl* levels in control KD, *Nanog* KD, and *Nanog* O/E mESCs. The mRNA level in control KD cells is set as 1. Data are normalized to *Actin*. Error bars represent SEM of three experiments. (*D*) Proposed model of the Ncl-mediated regulation of the homeostatic balance between self-renewal and differentiation in ESCs. Perturbation of this balance by depletion of *Ncl* in mESCs leads to ROS-induced p53 activation and subsequent self-renewal defects and differentiation. Nanog's transcriptional regulation of *Ncl* restrains p53, which when activated transcriptionally suppresses Nanog. We propose this reciprocity in regulation between Nanog and p53 in a Ncl-dependent manner as a Ncl-dependent Nanog-p53 bistable switch controlling the homeostatic balance between self-renewal and differentiation in ESCs.

genes, will complement RNAi-based screens and should be of immense value to those looking to identify genes associated with a specific biological phenotype.

Materials and Methods

Integrative Framework. Published microarray gene expression data sets from undifferentiated mESCs and various DC types (Dataset S1) were downloaded from Gene expression Omnibus (GEO; www.ncbi.nlm.nih.gov/geo/) and Array Express (www.ebi.ac.uk/microarray-as/ae/). To facilitate comparison of data across various microarray platforms, probe sets were mapped to Entrez gene identifiers using annotation files obtained from array manufacturers and databases including National Center for Biotechnology Information (ftp://ftp.ncbi.nih.gov/gene/DATA/), BioMart (www.ensembl.org/info/data/biomart.html), Mouse Genome Informatics (www.informatics.jax.org), and National Institute of Aging Array (http://lgsun.grc.nia.nih.gov/geneindex/).

Data analysis for every ESC-DC comparison was performed using RankProd (62), which rank-orders all genes on the array based on their rank-product *P* values (13). Briefly, rank product is a nonparametric statistic to prioritize genes based on how consistently they are found among the most strongly up- or down-regulated genes from a number of replicate experiments (13). Rank product for every gene is computed as the geometric mean of ranks of expression fold changes from *k* replicates (13). In our case, the smaller the rank product, the more severely and consistently that gene was down-regulated during differentiation (in DCs vs. ESCs). Every gene *g* from a ESC-DC comparison *i* receives an estimated rank-product *P* value $0 < P_{g,i} < 1$, based on 10,000 permutations, representing the probability of observing gene *g* at rank $r_{g,i}$ or smaller under the null hypothesis. For every gene *g*, rank-product *P* values $P_{g,i}$ ($i = 1$–68), obtained from 68 ESC-DC comparisons, were integrated to obtain a combined meta-analysis *P* value

$$P_g = 1 - \Phi\left[\frac{\sum_{i=1}^{N} w_i \Phi^{-1}(1 - P_{g,i})}{\sqrt{\sum_{i=1}^{N} w_i^2}}\right],$$

where, $n = 68$, Φ and Φ^{-1} represent standard normal cumulative distribution function (CDF) and its inverse respectively, and w_i is the weight to be given to the *i*th *P* value $P_{g,i}$, which we set to 1 (63). Informally, the obtained rank-product *P* values were transformed to z-scores using the inverse CDF, and the weighted sum of z-scores was divided by the square root of sum of the weights, which is converted to the meta-analysis *P* value using the CDF. Because we are analyzing data from each microarray experiment independently and combining the derived gene ranks and associated *P* values via the integrative framework described above, data across different array platforms did not need to be normalized.

Mouse ES Cell Culture, RNAi, and AP Staining. Mouse ESC culture, siRNA transfection, and AP staining were performed as previously described (24). Briefly, E14Tg2a and Oct4GiP mESCs and Oct4GiP mESCs overexpressing Nanog (24) were maintained on gelatin-coated plates in the ESGRO complete plus clonal grade medium (Millipore). For siRNA transfections, mESCs were cultured in M15 medium: DMEM (Invitrogen) supplemented with 15% FBS (vol/vol), 10 mM 2-mercaptoethanol, 0.1 mM nonessential amino acids (Invitrogen), 1× EmbryoMax nucleosides (Millipore), and 1,000 U of ESGRO (Millipore). mESCs (~25×10^3) were transfected with siRNAs at 50 nM using lipofectamine 2000 (Invitrogen). AP staining was performed using Alkaline Phosphatase Detection Kits from Millipore (#SCR004) and Stemgent (#00–0055) according to the manufacturer's instructions. See Table S2 for a list of gene specific siRNAs used, and Table S3 for a list of gene-specific primers used for RT-PCR analysis.

RNA-IP. RNA-IP was performed as described previously (64). Briefly, mESCs (1×10^7) were collected, washed twice in ice-cold PBS, and lysed using a lysis buffer containing 50 mM Hepes (pH 7.5), 140 mM NacCl, 1 mM EDTA, 10% (vol/vol) glycerol, 0.5% Igepal CA-630, 0.25% Triton X-100, 25 μM MG132, and 1× Complete protease inhibitor (Roche). The cell pellet was dissolved in a buffer containing 10 mM Tris·HCL, pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.1% sodium deoxycholate, 25 μM MG132, 1× protect RNA, and 1× Complete protease inhibitor. The lysate was precleared by adding 50 μL A/G magnetic beads (Pierce). The antibody-coated beads (5 μg of Ncl/IgG antibody and 100 μL of beads prepared in 1% BSA) were added to the precleared samples and rocked overnight at 4 °C. Subsequent washing and reverse cross-linking steps were performed as described previously (64). Immunoprecipitated RNA was isolated by adding TRIzol (Invitrogen) directly to the beads. The resulting RNA was then precipitated with glycogen (10 μg/reaction), and the pellet was resuspended in a volume appropriate for DNaseI digestion according to the manufacturer's protocol (Roche). cDNAs were generated using the iScript kit (Bio-Rad) according to the manufacturer's instructions. RNA-IP enrichment for a primer set was evaluated by PCR and agarose gel electrophoresis. See Fig. S7 for the original RNA-IP gel images shown in Figs. 4*G* and 5*B*. See *SI Materials and Methods* for RNA-IP primer sequences used.

Detection of Intracellular ROS. ROS levels were determined by incubating cells with 10 mg/mL dichlorfluorescein diacetate (DCFDA; Sigma-Aldrich) for 20 min at 37 °C. The cells were washed twice in PBS, trypsinized, and were quantitated for fluorescence with a FACScan flow cytometer (Becton Dickinson; exitation at 488 nm, emission at 515–545 nm). Data were analyzed using CELLQuest software and plotted as mean ± SEM of three or more biological replicates.

For additional experimental details, please refer to *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank K. Adelman and P. A. Wade for valuable suggestions and K. Adelman, T. K. Archer, A. Barski, M. B. Fessler, T. A. Kunkel, L. Ho, and P. A. Wade for critical comments on the manuscript. We thank H. Kinyamu and J. Yang for guidance on RNA-IP experiments and NIEHS Microarray, Confocal Microscopy, Flow Cytometry, Viral, and Protein cores for support. Nanog and p21 luciferase vectors are kind gifts from J. Huang (National Cancer Institute) and M. Resnick (National Institute of Environmental Health Sciences), respectively. This work was supported by Intramural Research Program of the National Institutes of Health, National Institute of Environmental Health Sciences Grants 1ZIAES102625 (to R.J.), 1ZIAES10274 (to G.H.), and 1ZIAES101866 (to D.V.Z.), and the Department of Science and Technology, India (DST-CMS GoI Project SR/S4/MS: 516/07, to S.Y.).

1. Mohr S, Bakal C, Perrimon N (2010) Genomic screening with RNAi: Results and challenges. *Annu Rev Biochem* 79:37–64.
2. Boutros M, Ahringer J (2008) The art and design of genetic screens: RNA interference. *Nat Rev Genet* 9(7):554–566.
3. Zender L, et al. (2008) An oncogenomics-based in vivo RNAi screen identifies tumor suppressors in liver cancer. *Cell* 135(5):852–864.
4. Luo J, et al. (2009) A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene. *Cell* 137(5):835–848.
5. Zuber J, et al. (2011) RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* 478(7370):524–528.
6. Ivanova N, et al. (2006) Dissecting self-renewal in stem cells with RNA interference. *Nature* 442(7102):533–538.
7. Fazzio TG, Huff JT, Panning B (2008) An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. *Cell* 134(1):162–174.
8. Hu G, et al. (2009) A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. *Genes Dev* 23(7):837–848.
9. Ding L, 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(5):403–415.
10. Chia NY, et al. (2010) A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. *Nature* 468(7321):316–320.
11. Wuestefeld T, et al. (2013) A Direct in vivo RNAi screen identifies MKK4 as a key regulator of liver regeneration. *Cell* 153(2):389–401.
12. Krishnan MN, et al. (2008) RNA interference screen for human genes associated with West Nile virus infection. *Nature* 455(7210):242–245.
13. Baril M, et al. (2013) Genome-wide RNAi screen reveals a new role of a WNT/CTNNB1 signaling pathway as negative regulator of virus-induced innate immune responses. *PLoS Pathog* 9(6):e1003416.
14. Subramanian V, Klattenhoff CA, Boyer LA (2009) Screening for novel regulators of embryonic stem cell identity. *Cell Stem Cell* 4(5):377–378.
15. Wu SM, Hochedlinger K (2011) Harnessing the potential of induced pluripotent stem cells for regenerative medicine. *Nat Cell Biol* 13(5):497–505.
16. Robinton DA, Daley GQ (2012) The promise of induced pluripotent stem cells in research and therapy. *Nature* 481(7381):295–305.
17. Suvà ML, Riggi N, Bernstein BE (2013) Epigenetic reprogramming in cancer. *Science* 339(6127):1567–1570.
18. Chen X, et al. (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133(6):1106–1117.
19. Young RA (2011) Control of the embryonic stem cell state. *Cell* 144(6):940–954.
20. Jaenisch R, Young R (2008) Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* 132(4):567–582.
21. Ng HH, Surani MA (2011) The transcriptional and signalling networks of pluripotency. *Nat Cell Biol* 13(5):490–496.
22. Orkin SH, Hochedlinger K (2011) Chromatin connections to pluripotency and cellular reprogramming. *Cell* 145(6):835–850.
23. Ho L, et al. (2011) esBAF facilitates pluripotency by conditioning the genome for LIF/STAT3 signalling and by regulating polycomb function. *Nat Cell Biol* 13(8):903–913.
24. Freudenberg JM, et al. (2012) Acute depletion of Tet1-dependent 5-hydroxymethylcytosine levels impairs LIF/Stat3 signaling and results in loss of embryonic stem cell identity. *Nucleic Acids Res* 40(8):3364–3377.
25. Melton C, Judson RL, Blelloch R (2010) Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. *Nature* 463(7281):621–626.
26. Marson A, et al. (2008) Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* 134(3):521–533.
27. Wang Y, et al. (2008) Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation. *Nat Genet* 40(12):1478–1483.
28. Zalzman M, et al. (2010) Zscan4 regulates telomere elongation and genomic stability in ES cells. *Nature* 464(7290):858–863.
29. Bilodeau S, Kagey MH, Frampton GM, Rahl PB, Young RA (2009) SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes Dev* 23(21):2484–2489.
30. Breitling R, Armengaud P, Amtmann A, Herzyk P (2004) Rank products: A simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett* 573(1-3):83–92.
31. Pijnappel WW, et al. (2013) A central role for TFIID in the pluripotent transcription circuitry. *Nature* 495(7442):516–519.
32. Buckley SM, et al. (2012) Regulation of pluripotency and cellular reprogramming by the ubiquitin-proteasome system. *Cell Stem Cell* 11(6):783–798.
33. Kagey MH, et al. (2010) Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467(7314):430–435.
34. Ho L, et al. (2009) An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proc Natl Acad Sci USA* 106(13):5181–5186.
35. Ho L, et al. (2009) An embryonic stem cell chromatin remodeling complex, esBAF, is an essential component of the core pluripotency transcriptional network. *Proc Natl Acad Sci USA* 106(13):5187–5191.
36. Jia J, et al. (2012) Regulation of pluripotency and self- renewal of ESCs through epigenetic-threshold modulation and mRNA pruning. *Cell* 151(3):576–589.
37. Takahashi K, Mitsui K, Yamanaka S (2003) Role of ERas in promoting tumour-like properties in mouse embryonic stem cells. *Nature* 423(6939):541–545.
38. Mongelard F, Bouvet P (2007) Nucleolin: A multiFACeTed protein. *Trends Cell Biol* 17(2):80–86.
39. Abdelmohsen K, Gorospe M (2012) RNA-binding protein nucleolin in disease. *RNA Biol* 9(6):799–808.
40. Lin T, et al. (2005) p53 induces differentiation of mouse embryonic stem cells by suppressing Nanog expression. *Nat Cell Biol* 7(2):165–171.
41. Silva J, Smith A (2008) Capturing pluripotency. *Cell* 132(4):532–536.
42. Yang A, et al. (2011) Nucleolin maintains embryonic stem cell self-renewal by suppression of p53 protein-dependent pathway. *J Biol Chem* 286(50):43370–43382.
43. Miniard AC, Middleton LM, Budiman ME, Gerber CA, Driscoll DM (2010) Nucleolin binds to a subset of selenoprotein mRNAs and regulates their expression. *Nucleic Acids Res* 38(14):4807–4820.
44. Takagi M, Absalon MJ, McLure KG, Kastan MB (2005) Regulation of p53 translation and induction after DNA damage by ribosomal protein L26 and nucleolin. *Cell* 123(1):49–63.
45. Jain AK, et al. (2012) p53 regulates cell cycle and microRNAs to promote differentiation of human embryonic stem cells. *PLoS Biol* 10(2):e1001268.
46. Krizhanovsky V, Lowe SW (2009) Stem cells: The promises and perils of p53. *Nature* 460(7259):1085–1086.
47. Li M, et al. (2012) Distinct regulatory mechanisms and functions for p53-activated and p53-repressed DNA damage response genes in embryonic stem cells. *Mol Cell* 46(1):30–42.
48. Liu B, Chen Y, St Clair DK (2008) ROS and p53: A versatile partnership. *Free Radic Biol Med* 44(8):1529–1535.
49. Ying QL, et al. (2008) The ground state of embryonic stem cell self-renewal. *Nature* 453(7194):519–523.
50. Esteban MA, et al. (2010) Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. *Cell Stem Cell* 6(1):71–79.
51. Griffith OW (1982) Mechanism of action, metabolism, and toxicity of buthionine sulfoximine and its higher homologs, potent inhibitors of glutathione synthesis. *J Biol Chem* 257(22):13704–13712.
52. Stead E, et al. (2002) Pluripotent cell division cycles are driven by ectopic Cdk2, cyclin A/E and E2F activities. *Oncogene* 21(54):8320–8333.
53. Zhao T, Xu Y (2010) p53 and stem cells: New developments and new concerns. *Trends Cell Biol* 20(3):170–175.
54. Aladjem MI, et al. (1998) ES cells do not activate p53-dependent stress responses and undergo p53-independent apoptosis in response to DNA damage. *Curr Biol* 8(3):145–155.
55. Lee DF, et al. (2012) Regulation of embryonic and induced pluripotency by aurora kinase-p53 signaling. *Cell Stem Cell* 11(2):179–194.
56. Han MK, et al. (2008) SIRT1 regulates apoptosis and Nanog expression in mouse embryonic stem cells by controlling p53 subcellular localization. *Cell Stem Cell* 2(3):241–251.
57. Sablina AA, et al. (2005) The antioxidant function of the p53 tumor suppressor. *Nat Med* 11(12):1306–1313.
58. Ben-Porath I, et al. (2008) An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet* 40(5):499–507.
59. Zbinden M, et al. (2010) NANOG regulates glioma stem cells and is essential in vivo acting in a cross-functional network with GLI1 and p53. *EMBO J* 29(15):2659–2674.
60. Diehn M, et al. (2009) Association of reactive oxygen species levels and radioresistance in cancer stem cells. *Nature* 458(7239):780–783.
61. Palmer NP, Schmid PR, Berger B, Kohane IS (2012) A gene expression profile of stem cell pluripotentiality and differentiation is conserved across diverse solid and hematopoietic cancers. *Genome Biol* 13(8):R71.
62. Hong F, et al. (2006) RankProd: A bioconductor package for detecting differentially expressed genes in meta-analysis. *Bioinformatics* 22(22):2825–2827.
63. Zaykin DV (2011) Optimally weighted Z-test is a powerful method for combining probabilities in meta-analysis. *J Evol Biol* 24(8):1836–1841.
64. Keene JD, Komisarow JM, Friedersdorf MB (2006) RIP-Chip: The isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. *Nat Protoc* 1(1):302–307.