RNAi Reveals Phase-Specific Global Regulators of Human Somatic Cell Reprogramming

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
RNAi Reveals Phase-Specific Global Regulators of Human Somatic Cell Reprogramming

Graphical Abstract

Highlights

- Key functional regulators of human somatic cell reprogramming are identified
- Combinatorial knockdown of five repressors allows highly efficient reprogramming (85%)
- SFRS11 regulates splicing of genes that are critical for reprogramming
- Splicing of downstream ZNF207 isoform B to A and C affects reprogramming efficiency

Authors
Cheng-Xu Delon Toh, Jun-Wei Chan, Zheng-Shan Chong, ..., Yoon-Pin Lim, Frederic A. Bard, Yuin-Han Loh

Correspondence
li.hu@mayo.edu (H.L.), bchlyp@nus.edu.sg (Y.-P.L.), yhloh@imcb.a-star.edu.sg (Y.-H.L.)

In Brief
Toh et al. identified genome-wide functional regulators of the early stages of human somatic cell reprogramming in a phase-specific RNAi screen. Combinatorial depletion of the top repressors allows reprogramming to proceed unhindered at a near-deterministic efficiency. SFRS11, an mRNA splicer, blocks reprogramming partially through the splicing of ZNF207 isoforms.

Accession Numbers
GSE68234
RNAi Reveals Phase-Specific Global Regulators of Human Somatic Cell Reprogramming

Cheng-Xu Delon Toh,1,2,24 Jun-Wei Chan,1,24 Zheng-Shan Chong,1,24 Hao Fei Wang,1,3 Hong Chao Guo,1,4 Sandeep Satapathy,1 Dongru Ma,6 Germaine Yen Lin Goh,6 Ekta Khattar,7 Lin Yang,8,9,10 Vinay Tergaonkar,7,11 Young-Tae Chang,12,13 James J. Collins,10,14,15,16 George Q. Daley,10,17,18,20 Keng Boon Wee,21,22 Chadi A. EL Farran,1,3 Hu Li,23,* Yoon-Pin Lim,2,11,22,* Frederic A. Bard,6,11 and Yuin-Han Loh1,2,3,*

1Epigenetics and Cell Fates Laboratory, A*STAR Institute of Molecular and Cell Biology, 61 Biopolis Drive Proteos, Singapore 138673, Singapore
2NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore, 28 Medical Drive, Singapore 117456, Singapore
3Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543, Singapore
4College of Life Sciences, Nankai University, Tianjin 300071, China
5Research and Development Unit (RDU), National Heart Centre Singapore, 5th Hospital Drive, Singapore 169609, Singapore
6Membrane Traffic Laboratory, A*STAR Institute of Molecular and Cell Biology, 61 Biopolis Drive Proteos, Singapore 138673, Singapore
7Division of Cancer Genetics and Therapeutics, Laboratory of NF-κB Signaling, A*STAR Institute of Molecular and Cell Biology, 61 Biopolis Drive Proteos, Singapore 138673, Singapore
8Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA
9Department of Genetics, Harvard Medical School, Boston, MA 02114, USA
10Howard Hughes Medical Institute, Boston, MA 02114, USA
11Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117545, Singapore
12Lab of Bioimaging Probe Development, A*STAR Singapore Bioimaging Consortium (SBIC), 11 Biopolis Way Helios, Singapore 138673, Singapore
13Department of Chemistry and MedChem Program, Life Sciences Institute, National University of Singapore, 3 Science Drive 3, Singapore 117543, Singapore
14Institute for Medical Engineering and Science Department of Biological Engineering, and Synthetic Biology Center, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
15Broad Institute of MIT and Harvard, Cambridge, MA 02139, USA
16Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA 02115, USA
17Stem Cell Transplantation Program, Division of Pediatric Hematology/Oncology, Boston Children’s Hospital and Dana-Farber Cancer Institute, Boston, MA 02115, USA
18Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA
19Harvard Stem Cell Institute, Boston, MA 02115, USA
20Manton Center for Orphan Disease Research, Boston, MA 02115, USA
21A*STAR Institute of High Performance Computing (IHPC), Connexis, Singapore 138632, Singapore
22A*STAR Bioinformatics Institute, Singapore 138671, Singapore
23Center for Individualized Medicine, Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, Rochester, MN 55905, USA
24Co-first author
*Correspondence: li.hu@mayo.edu (H.L.), bchlyp@nus.edu.sg (Y.-P.L.), yhloh@imcb.a-star.edu.sg (Y.-H.L.)
http://dx.doi.org/10.1016/j.celrep.2016.05.049

SUMMARY

Incomplete knowledge of the mechanisms at work continues to hamper efforts to maximize reprogramming efficiency. Here, we present a systematic genome-wide RNAi screen to determine the global regulators during the early stages of human reprogramming. Our screen identifies functional repressors and effectors that act to impede or promote the reprogramming process. Repressors and effectors form close interacting networks in pathways, including RNA processing, G protein signaling, protein ubiquitination, and chromatin modification. Combinatorial knockdown of five repressors (SMAD3, ZMYM2, SFRS11, SAE1, and ESET) synergistically resulted in ~85% TRA-1-60-positive cells. Removal of the novel splicing factor SFRS11 during reprogramming is accompanied by rapid acquisition of pluripotency-specific spliced forms. Mechanistically, SFRS11 regulates exon skipping and mutually exclusive splicing of transcripts in genes involved in cell differentiation, mRNA splicing, and chromatin modification. Our study provides insights into the reprogramming process, which comprises comprehensive and multi-layered transcriptional, splicing, and epigenetic machineries.

INTRODUCTION

The overexpression of the transcription factors (TFs) OCT4, SOX2, KLF4, and c-MYC (OSKM) has been shown to convert
Figure 1. Genome-wide siRNA Screen Identifies Regulators of Human Reprogramming

(A) Schematic of reprogramming timeline and temporal cell samples collected for analysis. Reprogramming intermediates on 5 d.p.i. (d5I), 8 d.p.i. (d8I), 12 d.p.i. and 16 d.p.i. were harvested for mRNA-seq. 12 d.p.i. and 16 d.p.i. cells were sorted by the TRA-1-60 marker into positive (d12T+, d16T+) and negative (d12T−, d16T−) populations.

(B) Ranked distribution plot for the genome-wide siRNA screen. 599 genes with Z scores of >1.65 are considered potential candidates that are repressor of reprogramming (red dots). 557 genes with Z scores of <1.3 are considered potential candidate effectors of reprogramming (green dots). The rest of the genes are indicated as gray dots.

(C) Representative images of TRA-1-60 (green) and DNA (blue) for selected hits identified from the genome-wide screen.

(D) Secondary screen of 76 genes with a second marker for pluripotency (TRA-1-81). Each row represents the knockdown of a single repressor, effector or NT control. Each column represents individual replicate experiments.

(legend continued on next page)
somatic cells into induced pluripotent stem cells (iPSCs) (Park et al., 2008; Takahashi and Yamanaka, 2006; Takahashi et al., 2007; Yu et al., 2007). Biochemical and genetic dissection of the molecular mechanisms underlying TF-induced reprogramming has been hampered by its low efficiency, slow kinetics, and the fact that it proceeds via heterogeneous intermediate cell populations (Polo et al., 2012; Stadtfeld and Hochedlinger, 2010). Nonetheless, recent studies have indicated that reprogramming is a stepwise process with at least three distinct phases: initiation, maturation, and stabilization (Samavarchi-Tehrani et al., 2010), with each phase marked by unique epigenomic and transcriptomic changes (Buganim et al., 2012; Polo et al., 2012). Several key cellular events, such as cell-cycle modulations (Kawamura et al., 2009; Qin et al., 2012; Zhao et al., 2008) and mesenchymal-to-epithelial transitions (Li et al., 2010; Samavarchi-Tehrani et al., 2010), were reported to influence reprogramming efficiency. In addition, the activity of epigenetic regulators, such as H3K9 methyltransferases (Chen et al., 2013) (EHMT1, SETDB1) and a H3K79 methyltransferase (Onder et al., 2012) (DOT1L), were found to limit the efficiency of reprogramming.

Recent attempts to attain a comprehensive overview of players involved in reprogramming utilized a pooled small hairpin RNA (shRNA) approach to conduct functional genome-wide screens in both mouse (Yang et al., 2014) and human (Qin et al., 2014) fibroblast lines. However, the constitutive knockdown of targeted genes by shRNA renders such screens unsuitable in the identification of key regulators affecting only a specific stage of the multiphasic reprogramming process (Buganim et al., 2012; Polo et al., 2012; Samavarchi-Tehrani et al., 2010).

To systematically identify genetic factors involved in early-phase reprogramming, we developed a high-throughput reprogramming assay adapted to a 384-well format for genome-wide small interfering RNA (siRNA) screening. This screening strategy allowed for transient knockdowns within a specific reprogramming phase, allowing us to identify phase-specific factors whose effects might be otherwise masked due to constitutive expression of the integrated shRNA.

RESULTS

Systematic Genome-wide siRNA Screen Identifies Functional Effectors and Repressors of Reprogramming

To identify the functional switches of early reprogramming, siRNA-mediated gene knockdown was performed on day 5 postinfection (5 d.p.i.) with lentiviral OSKM (Figure 1A). ESET siRNA-mediated gene knockdown was performed on day 5 to identify the functional switches of early reprogramming. Systematic genome-wide siRNA screening. This screening strategy allowed for transient knockdowns within a specific reprogramming assay adapted to a 384-well format for genome-wide small interfering RNA (siRNA) screening. This screening strategy allowed for transient knockdowns within a specific reprogramming phase, allowing us to identify phase-specific factors whose effects might be otherwise masked due to constitutive expression of the integrated shRNA.

Combinatorial Removal of Repressors Allows for High Reprogramming Efficiencies

The strongest repressors—namely SMAD3 (transforming growth factor β [TGF-β] signal transducer), ZMYM2 (epigenetic modifier), SFRS11 (putative splicing factor), SAE1 (Sumo-activating enzyme), and ESET (H3K9 methyltransferase) —function in distinct pathways. Therefore, we hypothesize that each factor is an independent barrier against reprogramming. As expected, the pairwise knockdown of two repressors had an additive effect on reprogramming enhancement (Figure 2A). We then pooled the siRNAs against repressors in groups of five (ESET, ZMYM2, SAE1, SFRS11, and SMAD3) or eight (si5, PAFAH2, FRAS1, and SNA1). The combinatorial knockdown of five or eight repressors enhanced reprogramming several fold as compared to the removal of individual repressors (Figures 2B and 2C). In addition, si5 depletion enhanced the formation of iPSCs in multiple reprogramming systems and parental cell lines (Figures S2A–S2C). As DOT1L and TGF-β signaling pathways

(E) Validation of candidate genes in a retroviral reprogramming system. The area of each marker stained was normalized to cell number, and the Z score was obtained with reference to a siNT2 negative control. n = 4; error bar indicates SD.

(F) Validation of selected candidates assayed at 21 d.p.i. Left: representative images of TRA-1-60 staining. n = 3; error bar indicates SD. Right: graph of the number of colonies stained with TRA-1-60. The gray dashed line indicates the reprogramming efficiency of negative control (siNT). n = 3; error bar indicates SD.
Figure 2. Combinatorial Knockdown of Repressors Synergistically Increases Reprogramming Efficiency

(A) Heatmap of the TRA-1-60 signal at 12 d.p.i., when two candidates were depleted simultaneously. n = 8.

(B) Removal of multiple repressors allows high reprogramming efficiency. Reprogramming was assayed at 16 d.p.i. All combinations of siRNA were reverse transfected with equimolar amounts of different siRNA. n = 8; error bars indicate SD.

(C) Representative images of TRA-1-60 (green) and DNA (blue) for the knockdown of repressors singly or in combination.

(D) Repressors function independently from the DOT1L pathway. Left: representative images of TRA-1-60+ colonies on 12-well plates after knockdown of the indicated genes. Right: bar graph indicating the number of TRA-1-60+ colonies obtained at 21 d.p.i. n = 3; error bar indicates SD.

(E) Repressors function independently of the TGF-β pathway. Left: representative images of TRA-1-60+ colonies on 12-well plates after knockdown of the indicated genes. Right: bar graph indicating the area of TRA-1-60 staining obtained at 21 d.p.i. n = 3; error bar indicates SD.

(F) si5 in combination with OSKML (LIN28) induction and TGF-β inhibition increase reprogramming yield. Bar graph of the percentage of cells that stained positive for TRA-1-60 at 28 d.p.i. The highest reprogramming efficiency (82.75%) was observed with the inhibition of TGF-β and si5, with the simultaneous overexpression of LIN28. n = 3; error bar indicates SD.

(legend continued on next page)
were previously implicated as critical barriers of reprogramming, we examined the effect of inhibiting these pathways upon si5 depletion. The inhibition of DOT1L and TGF-β signaling further enhanced reprogramming, suggesting that the repressors’ effect on reprogramming is at least partially independent of DOT1L and TGF-β signaling (Figures 2D and 2E). The combined inhibition of DOT1L or TGF-β with si5 allowed most cells to undergo homogenous morphological changes, resulting in rapid development of reprogrammed colonies (Figures 2D, 2E, and S2D). The ectopic expression of OSKM+L (LIN28), coupled with the depletion of si5 and TGF-β signaling, resulted in the generation of ~85% TRA-1-60+ cells at 28 d.p.i. (Figures 2F and S2E). Moreover, si5 depletion could replace c-MYC in the reprogramming cocktail (Figure 2G). iPSC lines obtained from si5 knockdown share similar characteristics with human embryonic stem cells (hESCs) and have normal chromosomal karyotypes (Figures 2H, 2I, and S2F–S2I). Taken together, these data demonstrate that the simultaneous removal of these repressors allows reprogramming to proceed in a more deterministic manner.

**Knockdown of Regulators Affects Dynamics of Reprogramming**

To understand the underlying mechanisms of the identified reprogramming regulators, we performed mRNA sequencing (mRNA-seq) on cells undergoing reprogramming (Figure 1A; Table S2). Differential GO analysis shows the shift of biological functions associated with fibroblasts toward processes that define pluripotency (Figure 3A). Additionally, transcriptome re-wiring is accompanied by a progressive epigenetic shift toward bivalency on lineage-associated genes and the activation of pluripotency genes (Figures S3A and S3B). The ectopic expression of OSKM+L (LIN28), coupled with the depletion of si5 and TGF-β signaling, resulted in the generation of ~85% TRA-1-60+ cells at 28 d.p.i. (Figures 2F and S2E). Moreover, si5 depletion could replace c-MYC in the reprogramming cocktail (Figure 2G). iPSC lines obtained from si5 knockdown share similar characteristics with human embryonic stem cells (hESCs) and have normal chromosomal karyotypes (Figures 2H, 2I, and S2F–S2I). Taken together, these data demonstrate that the simultaneous removal of these repressors allows reprogramming to proceed in a more deterministic manner.

**SFRS11 Regulates Alternative Splicing during Reprogramming**

Within the si5 combination, siSFRS11 contributed most of the increased reprogramming efficiency (Figures S4A and S4B). Moreover, SFRS11 is one of the strongest repressors (Figures 1E, 1F, and 2B–2E) in the majority of validation assays. We therefore focused on deciphering the mechanism through which SFRS11 represses reprogramming. siSFRS11-derived iPSC clones had characteristics similar to hESC lines (Figures S4C–S4E) and the capacity to form teratomas in immunodeficient mice (Figure S4F). Interestingly, SFRS11 depletion had no effect on pluripotency (Figures S4G and S4H) (Chia et al., 2010). These data suggest that SFRS11 specifically regulates the acquisition, but not maintenance, of pluripotency. Additionally, clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9-mediated knockout of SFRS11 also enhanced reprogramming efficiency, further confirming the repressive effects of SFRS11 on reprogramming (Figure S4I).

To gain mechanistic insights into the role of SFRS11, we analyzed deep mRNA-seq of the SFRS11- and SMAD3-depleted (non-silencing repressor control) reprogramming cells at 8 d.p.i. Exon inclusion and de novo transcript assembly analyses showed that SFRS11 knockdown resulted in drastic changes in differential isoform expression (Figures 4A, 4B, and S4J). Reads were mapped to exon-spline junction sites and analyzed to identify different splicing events. Depletion of SFRS11 resulted in a change in 696 splicing events (489 genes), whose inclusion ratio indicated significant splicing (false discovery rate [FDR] < 0.01) when compared to the non-targeting siRNA (siNT) control (Table S3). In contrast, SMAD3 knockdown resulted in a change of 43 splicing events (33 genes, FDR < 0.01). SFRS11 regulates the alternative splicing of genes that were enriched with several biological processes, including cell differentiation, cell proliferation, programmed cell death, mRNA splicing, and chromatin modifications (Figure 4C).

Out of these splicing events, four splicing patterns can be observed (Figures 4D, 4K, and S4L). Exon skipping (62.4%) and mutually exclusive (16.4%) events account for the majority of events detected upon SFRS11 depletion (Table S3). The depletion of SFRS11 predominantly increased the inclusion of skipped exons (Figure 4E) and of downstream mutually exclusive exons (Figure 4F). To validate these alternatively spliced events, we performed PCR for regions flanking skipped exons or mutually exclusive exons. Gel electrophoresis and band intensity analysis enable us to obtain the relative proportion of each alternatively spliced exon that differed in size. We validated four exon-skipping and four mutually exclusive splicing events at different time points of the reprogramming process. Interestingly, SFRS11 depletion resulted in splicing isoforms that are closer to the pluripotent state (Figures 4G and S4M). These
Figure 3. Perturbation of Repressors and Effectors Affects Reprogramming Kinetics
(A) Differential GO analysis for different stages of reprogramming. Darker red represents higher significance.
(B) siRepressor knockdown accelerate reprogramming kinetics. Immunostaining of TRA-1-60 (green), SSEA4 (red), and DNA (blue) at different days of reprogramming as indicated.
results suggest that specific splicing events could repress or promote reprogramming.

**SFRS11 Regulates Splicing of Reprogramming Effector ZNF207**

From exon-skipping events, we identified ZNF207 to be a putative downstream splicing target of SFRS11 during the early phase of reprogramming (Figures 4E and 4G). The mouse homolog zfp207 was reported to undergo alternative splicing during mouse reprogramming (Ohta et al., 2013). We first validated that the splicing of ZNF207 was specific to SFRS11 depletion and not due to the accelerated reprogramming kinetics (Figure 4H). Depletion of SFRS11 increased levels of ZNF207 isoform A and C transcripts while decreasing isoform B transcripts, corresponding to the increased retention of exon 9 (Figure 4H).

To understand the interaction between SFRS11 and ZNF207, we deleted the two functional domains of SFRS11, namely the RNA-recognition motif (RRM) that confers the ability to bind to single-stranded RNA and an arginine-serine rich domain (RS) that facilitates protein-protein interactions (Lunde et al., 2007) (Figures S4N and S4O). Consistent with the repressive phenotype observed from siRNA-mediated knockdown, overexpression of wild-type (WT) SFRS11 dramatically repressed reprogramming (p < 0.05). In contrast, mutation of either the RRM (ΔRRM) or the RS (ΔRS) domain partially rescued the repressive effects of SFRS11 (Figure 4I). We then asked if ZNF207 transcripts were directly regulated by SFRS11 by performing ultraviolet crosslinking coupled to immunoprecipitation and qPCR (CLIP-qPCR) in cells undergoing reprogramming (8 d.p.i.). We confirmed that SFRS11 binds directly to endogenous ZNF207 transcripts via its RRM domain (Figures 4J, 4K, and S4P). Detection of ZNF207 transcripts was only possible when antibody against WT SFRS11 was used in the immunoprecipitation (Figures 4J and 4K).

We next asked if the overexpression of individual ZNF207 isoforms affects reprogramming efficiency. Interestingly, the overexpression of ZNF207 isoforms A and C consistently enhanced reprogramming efficiency by 1.12- and 1.5-fold, respectively (p < 0.05), while overexpression of the ZNF207 isoform B decreased reprogramming efficiency to 0.7-fold (p < 0.05) (Figure 4L). The knockdown of all three ZNF207 isoforms using siRNA decreased overall reprogramming efficiency dramatically (Figure S4Q). We therefore designed steric hindrance antisense oligonucleotides (AONs) to induce specific skipping of either ZNF207 exons 6 or 9 (Figure 4H), resulting in the specific switching of ZNF207 isoform C to A (AON 1 and 2) or switching of ZNF207 isoform A and C to B (AON 3–5) (Figures S4R and S4S; Table S4). The introduction of these AONs reduced ZNF207 isoform C (Figures S4R and S4S) during reprogramming, resulting in a decrease in reprogramming efficiency (p < 0.001). As AON 3–5 promotes the switching of ZNF207 isoform C to isoform B, the increased of ZNF207 isoform B expression further reduces reprogramming efficiency as compared to AON 1 and 2 (p < 0.001) (Figures S4R and S4S). These data support the results of the overexpression studies, showing that an increase in ZNF207 isoforms A and C is crucial for reprogramming whereas an increase in ZNF207 isoform B hinders reprogramming.

**DISCUSSION**

Somatic cell reprogramming proceeds through ordered and distinct phases of initiation, maturation, and stabilization (Goll-pour et al., 2012; Samavarchi-Tehrani et al., 2010). We have developed a sensitive high-throughput screen assay, which enabled us to systematically unravel 599 repressors and 557 effectors of the early phase of reprogramming, including established and previously undescribed regulators (Onder et al., 2012; Qin et al., 2014; Rais et al., 2013; Samavarchi-Tehrani et al., 2010; Yang et al., 2014). Expectedly, the importance of OSM during early reprogramming was highlighted by our screen, consistent with a previous focused screen performed in the mouse system (Samavarchi-Tehrani et al., 2010). Interestingly, analysis of mRNA expression indicates that a significant proportion of reprogramming regulators remain unchanged during reprogramming (Table S2). This supports a previous study reporting that ~53% to 70% of the identified regulators remain transcriptionally stable during mouse reprogramming (Yang et al., 2014).

Among the repressors identified, the effect on reprogramming was strongest when SMAD3, SFRS11, ZMYM2, ESET, or SAE1 were depleted. These repressors represent functionally distinct regulators of unique pathways. The simultaneous removal of these repressors had a cumulative and maximal effect on improving reprogramming efficiency and kinetics. In combination with the overexpression of LIN28 and small-molecule inhibition of TGF-β, up to ~85% TRA-1-60+ cells can be detected at 28 d.p.i. Thus, we established that the simultaneous elimination of these functionally distinct repressors allows reprogramming to proceed synchronously in a near-deterministic manner with accelerated kinetics.

Temporal proteomic profiling of cells undergoing reprogramming indicated the robust upregulation of RNA processing factors during the early phase of reprogramming (Hansson et al., 2012). Our genome-wide screen indicates that RNA processing factors could functionally repress reprogramming (Figure S1H). Among these RNA-binding proteins, SFRS11 was the strongest repressor. Analysis of deep mRNA-seq samples at 8 d.p.i. (C) Principal-component analysis (PCA) of RNA-sequencing libraries (refer to Table S2). The knockdown of repressor genes accelerated the switch to a pluripotent state. Day 12 and 16 samples are sorted with TRA-1-60. (D) Dynamic expression of genes that are differentially expressed between fibroblasts and H1 hESCs during reprogramming. d12T+ and d16T+ siRepressor samples demonstrate a higher rate of downregulation of fibroblast-associated genes. siNT2 samples are boxed (light gray). (E) Heatmap demonstrating dynamic expression of the indicated genes clustered into categories consisting of genes expressed at the early (I), intermediate (II), or late stages (III) of reprogramming. Genes belonging to EMT (M) and MET (E) are also indicated. siNT2 samples are boxed (light gray). (F) qRT-PCR quantification of pluripotency (ERRB3, NANOG, ZFP42), mesenchymal (TWIST1), and epithelial (EPCAM, KRT8) markers. Data shown are normalized to H1 hESCs. n = 2; error bars indicate SD.
Figure 4. SFRS11 Regulates the Splicing of the Transcripts for Reprogramming Effector ZNF207

(A) Difference in exon inclusion levels of significant differential splicing events between siSFRS11 versus siNT and siSMAD3 versus siNT. Heatmap color scale ranges from white (no change in inclusion level) to red (large difference in inclusion level).

(legend continued on next page)
revealed that 489 genes are alternatively spliced when SFRS11 was transiently depleted (Table S3). Interestingly, these genes were enriched for biological functions associated with mRNA splicing (MBNL1, SFRS3, U2AF1, SFRS2, and PUF60). MBNL1 was previously shown to regulate the splicing of a pluripotency switch (FOXP1) in hESCs (Han et al., 2013) and was also identified as repressor from our genome-wide screen (Figure 1B). The mouse homologs for SFRS3 and U2AF1 were previously identified as repressor from our genome-wide screen (Figure 1B). The was previously shown to regulate the splicing of a pluripotency splicing factors (Figure 4N).

This strongly suggests that SFRS11 could regulate reprogramming through the splicing of downstream pluripotent-specific splicing factors (Figure 4N).

Our study identified ZNF207 as one of the effectors of human reprogramming. SFRS11 depletion during reprogramming induces the isofrom switch of ZNF207 B to A and C. The identification of ZNF207 and other SFRS11-regulated splicing events highlights an extensive role for the alternative splicing network during reprogramming (Figure 4N). While SFRS11 repression enhances reprograming by at least 5-fold (Figure 1F), overexpression of ZNF207 isoform C enhances reprograming by 1.5-fold. It is highly possible that other SFRS11-regulated splicing events are contributing to its role as a reprogramming repressor. The functional significance of other SFRS11-regulated splicing variants remains to be validated in the reprogramming context.

Our study demonstrates that the regulation of a multiphasic process, such as reprogramming, involve distinct regulators that are involved in multiple pathways. Together, our systematic and transient siRNA screen provides a wealth of information about the functional regulators required specifically at the early phase of reprogramming. These regulators may not have been uncovered with the complete depletion of gene function using gene knockout or constitutive shRNA knockdown (Qin et al., 2014; Yang et al., 2014). Taken together, these findings highlight the need to transiently generate hypomorphic states of gene expression in order to functionally dissect multi-phasic somatic cell reprogramming.

**EXPERIMENTAL PROCEDURES**

**Transfection, Staining, and Imaging in 384-Well Plates**

Plates (384-well, Grenier) were coated with 15 μl Matrigel (BD) overnight at 37°C. siRNA (2.5 μl 500 nM siGENOME, Dharmacon) was added to the plates, which were then frozen at −20°C before use. For reverse transfection, a master mix of 0.04 μl DharmaFECT1 (Dharmacon) transfection reagent and 7.46 μl OptiMEM (Invitrogen) mix was added to each well of the 384-well plates containing the siRNA. 5 days post-OSKM, the DharmaFECT1 and siRNA mix was incubated for 20 min before 4,000 BJ fibroblasts were seeded in each well in 40 μl MEF medium.

For the genome-wide screen, the whole-genome Dharmacon SMARTpooled siRNA library targeting 21,121 human genes was printed on 67 Matrigel-coated 384-well plates, where each well contained a mixture of four siRNAs targeting a single gene. On each plate, we included negative controls (non-target siRNA) and positive controls (ESET siRNA and OCT4 siRNA) in the designated wells. These controls were chosen because they are known to play an important role in somatic cell reprogramming: OCT4 (Park et al., 2008; Takahashi and Yamanaka, 2006; Takahashi et al., 2007; Yu et al., 2007) is indispensable in most reprogramming assays, and ESET (Chen et al., 2013; Souli et al., 2012) is a well-established barrier of reprogramming. Medium was changed every other day for 6 days. Subsequently, cells were washed once with PBS and fixed with 50 μl 4% paraformaldehyde (Sigma) for 20 min. Cells were then washed twice before 25 ml anti-human TRA-1-60 antibody (1:300, BD #560173) was added, and cells were incubated at 4°C overnight. Hoechst 3342 (120,000, Invitrogen) was added to each well and stained for 30 min. The cells were then washed once with PBS and covered in 50 ml PBS.

Cells were imaged with an IXU ultra plate-scanning confocal microscope (Molecular Devices) at x10 magnification, and nine pictures were taken per well. Images obtained from multiple sites were combined to generate a
montage of different panels. Granule area, integrated fluorescent intensity, and nuclei number were quantified using MetaXpress Image Acquisition and Analysis software v2. The Z score was calculated using the formula $Z = (X - m)/SD$, where $m$ is the mean of negative control and $SD$ is the standard deviation of the whole population. $X$ is the sample value calculated based on TRA-1-60+ area per number of cells. For further analysis, we excluded genes that reduced cell number by <50% of the average cell number per well in each plate. This was done to remove false-positive candidates that affected cell viability.

More details on production of viral supernatant, reprogramming assay, cell culture, differentiation of iPSC clones, immunostaining, RNA extraction, RNA sequencing, ChIP sequencing, AON transfection, microarray data, differential GO analysis, enrichment analysis, GSEA, and other procedures can be found in Supplemental Experimental Procedures.

ACCESSION NUMBERS

The accession number for the raw data from sequencing experiments reported in this paper is GEO: GSE68234.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2016.05.049.

AUTHOR CONTRIBUTIONS

C.-X.D.T., J.-W.C., and Z.-S.C. designed and performed research, analyzed data, and wrote the paper; H.F.W., H.C.G., D.M., S.S., G.Y.L.G., and E.K. designed research, analyzed data, and wrote the paper.

ACKNOWLEDGMENTS

We are grateful to Kenneth Zaret and Samantha Sehay for helpful discussion. We thank Tong Yang-Si, Xin Xiao-Xuan, and Julien Maury for technical assistance. H.L. supported by grants from the National Institutes of Health (CA196631-01A1 and 1U54GM114838-01). Y.-H.L. is supported by the A*Star Investigators research award and grants JCO R09138 and JCO R09125. We are grateful to the Biomedical Research Council, Agency for Science, Technology and Research, Singapore for research funding.

Received: February 19, 2016
Revised: April 3, 2016
Accepted: May 11, 2016
Published: June 9, 2016

REFERENCES


