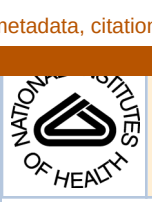

Published in final edited form as:

Nat Genet. 2010 December ; 42(12): 1113–1117. doi:10.1038/ng.710.

Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells

Sabine Loewer[1,2,3,4], **Moran N. Cabili**[5,6], **Mitchell Guttman**[5,7], **Yuin-Han Loh**[1,2,3,4], **Kelly Thomas**[5,8], **In Hyun Park**[1,2,3,4,12], **Manuel Garber**[5], **Matthew Curran**[1,3], **Tamer Onder**[1,2,3,4], **Suneet Agarwal**[1,2,3], **Philip D. Manos**[1,3,4], **Sumon Datta**[1,3,4], **Eric S. Lander**[5,6,7], **Thorsten M. Schlaeger**[1,3,4], **George Q. Daley**[1,2,3,4,8,9], and **John L. Rinn**[5,10,11]

[1]Stem Cell Transplantation Program, Division of Pediatric Hematology/Oncology, Manton Center for Orphan Disease Research, Children's Hospital Boston and Dana Farber Cancer Institute, Boston MA, USA

[2]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston MA, USA

[3]Harvard Stem Cell Institute, Cambridge MA, USA

[4]Stem Cell Program, Children's Hospital Boston, Boston MA, USA

[5]The Broad Institute of Harvard and MIT, Cambridge MA, USA

[6]Department of Systems Biology, Harvard Medical School, Boston MA, USA

[7]Department of Biology, Massachusetts Institute of Technology, Cambridge MA, USA

[8]Division of Hematology, Brigham and Women's Hospital, Boston MA, USA

[9]Howard Hughes Medical Institute, Chevy Chase MD, USA

[10]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge MA, USA

[11]Department of Pathology, Beth Israel and Deaconess Medical Center, Harvard Medical School, Boston MA, USA

Abstract

The conversion of lineage-committed cells to induced pluripotent stem cells (iPSCs) by reprogramming is accompanied by a global remodeling of the epigenome[1-5], resulting in altered patterns of gene expression[2,6-9]. Here we characterize the transcriptional reorganization of large intergenic non-coding RNAs (lincRNAs)[10,11] that occurs upon derivation of human iPSCs, and identify numerous lincRNAs whose expression is linked to pluripotency. Among these, we defined 10 lincRNAs whose expression was elevated in iPSCs compared with embryonic stem cells (ESCs), suggesting that their activation may promote the emergence of iPSCs. Supporting this, our results indicate that these lincRNAs are direct targets of key pluripotency transcription

Correspondence to: George Q. Daley; John L. Rinn.

[12]Present address: Yale Stem Cell Center, Department of genetics, Yale School of Medicine, New Haven CT, USA

Author Contributions Co-direction of the project: G.Q.D. and J.L.R. Study concept and design: G.Q.D., J.L.R., S.L. LincRNA array design: M. Guttman and J.L.R. iPSC generation and characterization: I.H.P., S.L., and P.M. LincRNA array hybridization, lincRNA and protein-coding gene expression analysis: M.N.C., K.T., M. Guttman, S.L., and M. Garber Computational studies: M.N.C., M. Guttman and M. Garber. LincRNA transcriptional regulation: S.L. ChIP assays: Y.-H. L. LincRNA loss-of-function/gain-of-function studies: S.L., M.C. and S.D. T.S. and E.S.L. provided essential ideas and suggestions on the manuscript. Manuscript preparation: G.Q.D., J.L.R., and S.L.

factors. Using loss- and gain-of-function approaches, we found that one such lincRNA (lincRNA-RoR) modulates reprogramming, thus providing a first demonstration for critical functions of lincRNAs in the derivation of pluripotent stem cells.

Cellular reprogramming demonstrates the remarkable plasticity of cell fates, illustrated by the isolation of induced pluripotent stem cells (iPSCs) from fibroblasts[6-9]. Molecular analysis of epigenetic modifications has revealed a near complete remodeling of the epigenome during reprogramming[1-4,12], resulting in the conversion of lineage-specific to uniform protein-coding and miRNA gene expression profiles similar to embryonic stem cells (ESCs)[2,6-9]. We and others have recently discovered a novel class of large intergenic non-coding RNAs (lincRNAs) that are expressed in a cell type-specific manner[13] and can associate with epigenetic regulators[11,14-16] involved in pluripotency and lineage commitment[17,18].

To date, it is not known whether large-scale transcriptional changes induced by reprogramming apply to lincRNAs, and if these changes have any functional relevance. To test this, we compared the transcriptional profiles of human lincRNAs alongside protein-coding genes across fibroblasts, their derivative iPSCs, and ESCs. We reprogrammed four primary fibroblast lines[7] and validated the functionality of resulting iPSC lines (Supplementary Figure 1, 2, and data not shown). We then performed DNA microarray analysis of the parental fibroblasts, seven of their derivative iPSC lines, and two ESC lines. Consistent with previous studies, analysis of gene expression profiles revealed that all iPSCs were similar to ESCs[19,20], and distinct from fibroblasts (Figure 1A, Supplementary Figure 3). We detected 3694 and 3283 genes up- and downregulated, respectively, in iPSCs and ESCs compared with fibroblasts (>2fold, $P<0.05$; Figure 1B). Taken together, our fibroblast-derived iPSCs fulfill functional criteria of *bona fide* iPSCs[20] and exhibit a uniform protein-coding gene expression profile similar to ESCs.

To explore the expression of lincRNAs, we designed a microarray probing ~900 lincRNAs in the human genome[11] and analyzed their expression in the above cell lines. The global lincRNA expression profiles of iPSCs were very similar to ESCs and distinct from fibroblasts (Figure 1C). We observed 133 and 104 lincRNAs that were induced or repressed (>2fold, FWER<0.05) across all iPSCs and ESCs compared with fibroblasts (Figure 1D, E, Supplementary Table 1). Similar to protein-coding genes, direct reprogramming resulted in concomitant activation or repression of numerous lincRNAs consistent with a reactivation of the ESC state.

To exclude the possibility that reprogramming-induced changes in lincRNA expression reflects the opening and closing of chromatin domains of neighboring protein-coding genes, we analyzed the correlation of expression between each reprogrammed lincRNA and their neighboring genes, and found no significant correlation ($P=0.999$, Figure 1F). This indicates an independent and cell type-specific regulation of lincRNA expression.

We sought to identify lincRNAs with potentially important functions in ESCs/iPSCs. Among the many pluripotency-associated lincRNAs (Figure 1D and Supplementary Figure 4), we searched for those that were expressed in both ESCs and iPSCs, but showed elevated levels in iPSCs relative to ESCs, reasoning that their high expression may have conferred a selective advantage on emerging iPSCs. We identified 28 lincRNAs that showed greater expression in fibroblast-iPSCs relative to ESCs (>2fold, FWER<0.05; Figure 2A), and refer to these as “iPSC-enriched” lincRNAs hereafter.

We hypothesized that if iPSC-enriched lincRNAs are important for reprogramming, they should be elevated in iPSCs independent of the cell-of-origin. To test this, we profiled

lincRNA expression in $CD34^+$ hematopoietic stem/progenitor cells, two $CD34^+$-iPSC lines[21] and ESCs, using the same approach as above. Like fibroblast-iPSCs, $CD34^+$-iPSCs had similar global lincRNA expression profiles as ESCs, distinct from that of $CD34^+$ cells (Supplementary Figure 4). 10 of the 28 lincRNAs elevated in fibroblast-iPSCs were also elevated in $CD34^+$-iPSCs (Figure 2B, Supplementary Figure 5). This overlap was statistically significant ($P<0.0001$). We independently validated levels of 8/10 common iPSC-enriched lincRNAs by qRT-PCR (Figure 2B), and detected considerable variation in expression. Positive selection for minimal RNA levels and the absence of counter-selection against higher expression during reprogramming may be the cause for this variability. Collectively, these results demonstrate that numerous lincRNAs are tightly associated with the pluripotent state, including a subset that is consistently enriched in iPSCs independent of the cell-of-origin.

If iPSC-enriched lincRNAs are important for iPSC derivation, we suspected a link with the pluripotency network. To test this, we first intersected published OCT4 binding regions in ESCs[22] with iPSC-enriched lincRNA loci (demarcated by domains of histone H3K4 and H3K36 methylation[10,11], named according to their neighboring 3' gene), and identified three overlapping loci: lincRNA-SFMBT2, lincRNA-VLDLR, and lincRNA-ST8SIA3. We performed independent ChIP-qPCR to validate the binding of OCT4, and probe for SOX2 and NANOG occupancy at these sites. All three transcription factors occupied these regions, coinciding with or in close proximity to lincRNA promoters (peaks of H3K4me[10]; Figure 3A, Supplementary Figure 6).

To determine whether expression of iPSC-enriched lincRNAs is dependent on pluripotency transcription factors, we depleted OCT4 in iPSCs and ESCs using siRNAs and monitored the levels of iPSC-enriched lincRNAs. We verified OCT4 knock-down and induction of the differentiation marker *LMNA* (Figure 3B, Supplementary Figure 7). Levels of all three iPSC-enriched lincRNAs fell within 72h (Figure 3B, Supplementary Figure 7C). To further verify that downregulation of iPSC-enriched lincRNAs is caused by perturbation of the pluripotency network, we induced embryoid body (EB) formation as a distinct pathway of differentiation. Again, levels of all three iPSC-enriched lincRNAs fell within two days (Figure 3C, Supplementary Figure 7D). The expression of these lincRNAs thus appears controlled by pluripotency transcription factors in ESCs and iPSCs.

We turned to investigate the functional roles of iPSC-enriched lincRNAs in the reprogramming process. To this end, we generated shRNA-expressing lentiviruses targeting lincRNA-ST8SIA3 and lincRNA-SFMBT2, which showed the strongest response to EB differentiation and OCT4 knock-down, and validated each knock-down relative to a non-targeting control shRNA (Figure 4A, Supplementary Figure 8A). To test the effect of lincRNA depletion on reprogramming, we infected dH1f fibroblasts[7] with both the shRNA-expressing and the reprogramming viruses[7,9], and scored emerging iPSC colonies based on Tra-1-60 marker expression (day 21)[20]. Interference with lincRNA-SFMBT2 did not affect iPSC colony formation (Supplementary Figure 8B, C), suggesting that lincRNA-SFMBT2 is not essential, or alternatively, that its moderate reduction was insufficient to perturb reprogramming. In contrast, knock-down of lincRNA-ST8SIA3 resulted in a significant 2 to 8-fold decrease of iPSC colonies relative to the control, where as progenitor cells were unaffected ($P<0.01$; Figure 4B, C, D, Supplementary Table 2). As expected, resulting iPSC colonies fulfilled additional criteria of fully reprogrammed cells (Supplementary Figure 9). These results demonstrate a functional requirement of lincRNA-ST8SIA3 expression for iPSC derivation.

Several studies have established critical roles of cell proliferation and bypass of senescence during the early stages of reprogramming[23-27]. We therefore examined if knockdown of

lincRNA-ST8SIA3 compromised cell growth of fibroblasts or cells during this window, and failed to detect significant differences in cells infected with the lincRNA-ST8SIA3-targeting virus compared with the control (Figure 4C, Supplementary Figure 10). In addition, the kinetics of reprogramming upon knock-down of lincRNA-ST8SIA3 was similar to the control (Supplementary Figure 11). Collectively, these findings point to a specific inhibition of the reprogramming process, rather than a delay of iPSC formation upon loss of lincRNA-ST8SIA3.

Intrigued by this phenotype, we used 5' and 3' rapid amplification of cDNA ends to clone the full-length transcript of lincRNA-ST8SIA3 (Figure 4E), which recovered a 2.6kb long RNA comprised of four exons (Figure 4E, red). We did not detect any clones that were spliced to protein-coding genes nor intact open reading frames, and confirmed the presence of a single transcript of expected length by Northern Blotting (Figure 4E, Supplementary Figure 12).

We next used a complementary gain-of-function approach to test whether elevated lincRNA-ST8SIA3 expression might enhance reprogramming. We infected dH1fs with empty pBabe-puro retrovirus, GFP-expressing, or lincRNA-ST8SIA3-expressing virus, selected transgenic cells, and documented 25 to 70-fold overexpression of lincRNA-ST8SIA3 relative to the levels in H9 ESCs (Figure 4F). We induced reprogramming in these stable cell lines and consistently observed a more than 2-fold increase in iPSC colony formation (day 28 +/-2 days) ($P<0.001$; Figure 4G). This was not associated with significant changes in cell growth of fibroblasts or cells at early stages of reprogramming (Figure 4H, Supplementary Figure 10). Thus, overexpression of lincRNA-ST8SIA3 positively affects the establishment of iPSCs during reprogramming (Figure 4G, I), in addition to possible functions in iPSC maintenance. Supporting the latter, transient knock-down of lincRNA-ST8SIA3 in ESCs and established iPSCs resulted in a growth deficiency linked with elevated apoptosis (Supplementary Figure 13).

To gain insight into which cellular pathways are affected by lincRNA-ST8SIA3 knock-down, we performed microarray gene expression analysis. Consistent with its apoptotic phenotype, knock-down of lincRNA-ST8SIA3 led to upregulation of genes involved in the p53 response, the response to oxidative stress and DNA damage-inducing agents, as well as cell death pathways (Supplementary Table 3). Interestingly, simultaneous knock-down of p53 partially rescued the apoptotic phenotype caused by ablation of lincRNA-ST8SIA3 (Supplementary Figure 14). Taken together, these results suggest that lincRNA-ST8SIA3 plays a role in promoting survival in iPSCs and ESCs, likely by preventing the activation of cellular stress pathways including the p53 response.

Our transcriptional profiling approach has revealed numerous lincRNAs that are part of the transcriptional repertoire of human ESCs and are induced during reprogramming of different cell types. We have identified several iPSC-enriched lincRNAs that appear to be directly regulated by the pluripotency network. Interestingly, we found no direct syntenic correlates of the 10 iPSC-enriched lincRNAs expressed in mouse ESCs (with the exception of lincRNA-VLDLR). Similar to what has been described for protein-coding genes[28], the transcriptional networks of lincRNAs in ESCs may have become rewired, conferring species-specific regulation.

The modulation of reprogramming by lincRNA-ST8SIA3 provides the first functional example of a lincRNA in establishing iPSCs, we therefore name it lincRNA-RoR for Regulator of Reprogramming. Future studies will be required to decipher the molecular mechanism by which lincRNA-RoR acts, and to gain a global understanding of lincRNA function in the establishment and maintenance of pluripotency. One possibility is that

pluripotency-associated lincRNAs interface with chromatin modifying complexes to assist in the regulation of the distinct epigenetic architecture in pluripotent cells. Supporting this, previous studies have demonstrated critical roles for chromatin-modifying complexes in the establishment and maintenance of pluripotency, and numerous lincRNAs can interact with these complexes to impart target specificity[11,15,16]. Here we demonstrate the modulation of reprogramming by a large non-coding RNA, supporting the notion that lincRNAs represent an additional layer of complexity in the networks controlling cellular identity.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Hongguang Huo, Alexander DeVine for technical assistance; Sutheera Ratanasirintrawoot and Erin McLoughlin for reagents; Mathew W. Lensch for teratoma interpretation,. J.L.R. is a Damon Runyon-Rachleff, Searle, Smith Family Foundation and Richard Merkin Foundation Scholar. S.L. was supported by a Human Frontier Science Program Organization long-term fellowship. Y.-H.L. is supported by the Agency of Science, Technology and Research International Fellowship and the A*Star Institute of Medical Biology, Singapore. GQD is an investigator of the Howard Hughes Medical Institute. Research was funded by grants from the National Institutes of Health (NIH) to G.Q.D (1 RC2-HL102815) and J.L.R. (1DP2OD00667-01).

References

1. Ball MP, et al. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat Biotechnol 2009;27:361–8. [PubMed: 19329998]
2. Chin MH, et al. Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. Cell Stem Cell 2009;5:111–23. [PubMed: 19570518]
3. Doi A, et al. Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. Nat Genet 2009;41:1350–3. [PubMed: 19881528]
4. Maherali N, et al. Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell Stem Cell 2007;1:55–70. [PubMed: 18371336]
5. Mikkelsen TS, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 2007;448:553–60. [PubMed: 17603471]
6. Lowry WE, et al. Generation of human induced pluripotent stem cells from dermal fibroblasts. Proc Natl Acad Sci U S A 2008;105:2883–8. [PubMed: 18287077]
7. Park IH, et al. Reprogramming of human somatic cells to pluripotency with defined factors. Nature 2008;451:141–6. [PubMed: 18157115]
8. Takahashi K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 2007;131:861–72. [PubMed: 18035408]
9. Yu J, et al. Induced pluripotent stem cell lines derived from human somatic cells. Science 2007;318:1917–20. [PubMed: 18029452]
10. Guttman M, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 2009;458:223–7. [PubMed: 19182780]
11. Khalil AM, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A 2009;106:11667–72. [PubMed: 19571010]
12. Mikkelsen TS, et al. Dissecting direct reprogramming through integrative genomic analysis. Nature 2008;454:49–55. [PubMed: 18509334]
13. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. Cell 2009;136:629–41. [PubMed: 19239885]
14. Nagano T, et al. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science 2008;322:1717–20. [PubMed: 18988810]

15. Rinn JL, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 2007;129:1311–23. [PubMed: 17604720]
16. Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science 2008;322:750–6. [PubMed: 18974356]
17. Boyer LA, et al. Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature 2006;441:349–53. [PubMed: 16625203]
18. Lee TI, et al. Control of developmental regulators by Polycomb in human embryonic stem cells. Cell 2006;125:301–13. [PubMed: 16630818]
19. Assou S, et al. A meta-analysis of human embryonic stem cells transcriptome integrated into a web-based expression atlas. Stem Cells 2007;25:961–73. [PubMed: 17204602]
20. Chan EM, et al. Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. Nat Biotechnol 2009;27:1033–7. [PubMed: 19826408]
21. Loh YH, et al. Generation of induced pluripotent stem cells from human blood. Blood 2009;113:5476–9. [PubMed: 19299331]
22. Marson A, et al. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell 2008;134:521–33. [PubMed: 18692474]
23. Hong H, et al. Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. Nature. 2009
24. Kawamura T, et al. Linking the p53 tumour suppressor pathway to somatic cell reprogramming. Nature. 2009
25. Li H, et al. The Ink4/Arf locus is a barrier for iPS cell reprogramming. Nature. 2009
26. Marion RM, et al. A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. Nature. 2009
27. Utikal J, et al. Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. Nature. 2009
28. Kunarso G, et al. Transposable elements have rewired the core regulatory network of human embryonic stem cells. Nat Genet 2010;42:631–4. [PubMed: 20526341]
29. Dinger ME, et al. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. Genome Res 2008;18:1433–45. [PubMed: 18562676]
30. Reich M, et al. GenePattern 2.0. Nat Genet 2006;38:500–1. [PubMed: 16642009]
31. Loh YH, et al. The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat Genet 2006;38:431–40. [PubMed: 16518401]

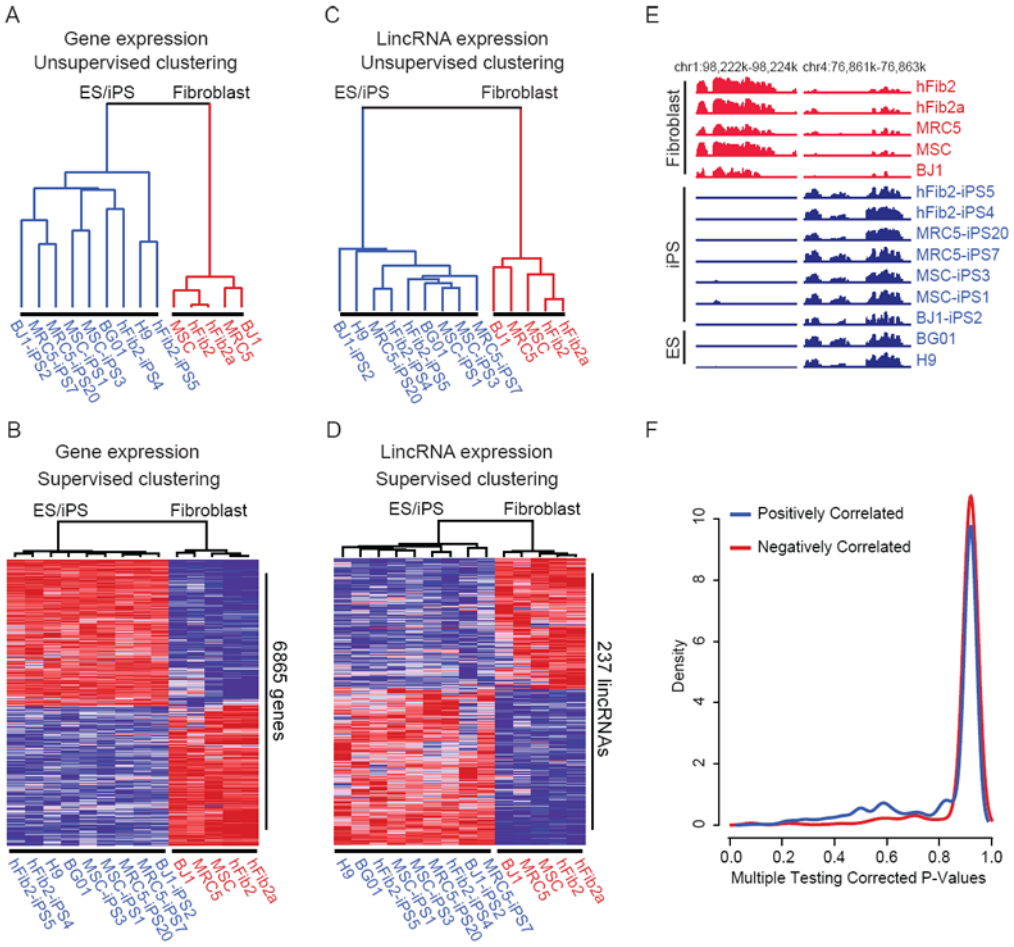

Figure 1.

Direct reprogramming of fibroblasts converts both protein-coding gene and lincRNA expression to a puripotent cell-specific profile. A, C) Unsupervised hierarchical clustering of protein-coding gene expression (A) and lincRNA expression (C) segregates fibroblasts (red) from ESCs and fibroblast-derived iPSCs (blue). B, D) Supervised hierarchical clustering analysis identifies 6865 protein-coding genes (B) and 237 lincRNAs (D) that are differentially expressed between ESCs/iPSCs and fibroblasts (genes: >2fold, P<0.05; lincRNAs: >2fold, FWER<0.05). Expression values are represented in shades of red and blue relative to being above (red) or below (blue) the median expression value across all samples (log scale 2, from -3 to +3). hFib2 fibroblasts are represented as two replicates (hFib2 and hFib2a). E) Examples of reprogrammed lincRNAs; left: lincRNA expressed in all fibroblasts is repressed in all pluripotent cells, right: a pluripotent cell-specific lincRNA that becomes activated during reprogramming. Expression values for each tiled probe (x-axis) are displayed as normalized hybridization intensity (y-axis). F) Correlation analysis of lincRNAs and neighboring genes. Density plot of multiple testing corrected p-values (x-axis) for lincRNAs that are positively (blue) or negatively (red) correlated with their protein-coding gene neighbors.

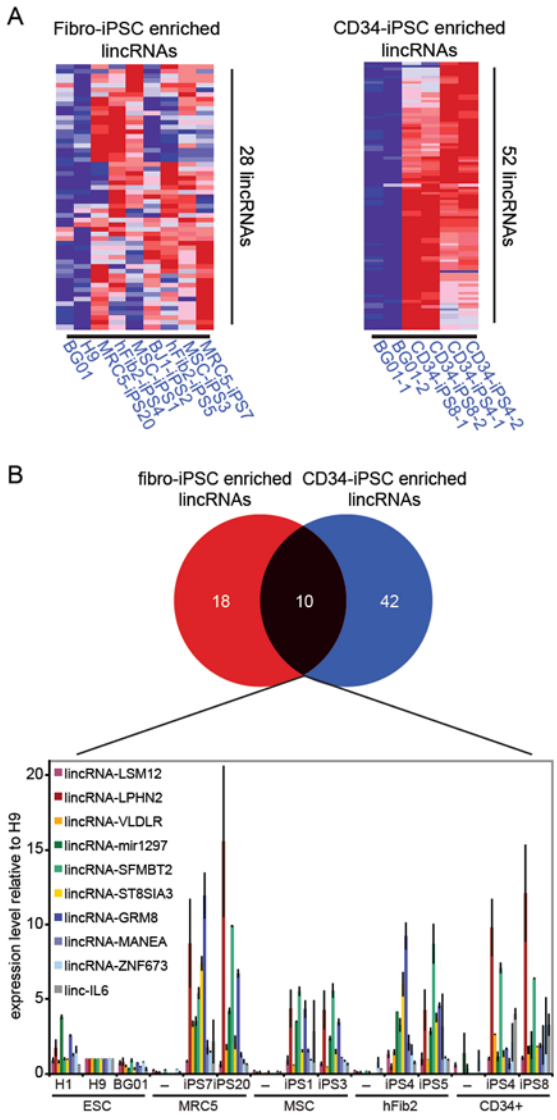

Figure 2.

Several lincRNAs show enriched expression in iPSCs compared with ESCs. A) Heatmap of 28 and 52 lincRNAs that are more highly expressed in fibroblast-derived iPSCs (left) and CD34^{+}-derived iPSCs (right), respectively, compared with ESCs (>2fold, FWER<0.05). Expression values are represented in shades of red and blue relative to being above (red) or below (blue) the median expression value across all samples (log scale 2, from -3 to +3). B) Top: Venn Diagram shows 10 lincRNAs that are commonly enriched in fibroblast- and CD34^{+}-derived iPSCs. Bottom: qRT-PCR validation of the 10 commonly enriched lincRNAs (named according to their 3' protein-coding gene neighbor) across three human ESC lines (H1, H9, BG01), fibroblasts (MRC5, MSC, hFib2) and CD34^{+} cells, and their derivative iPSC lines. Expression values are represented relative to the RNA levels in H9 ESCs.

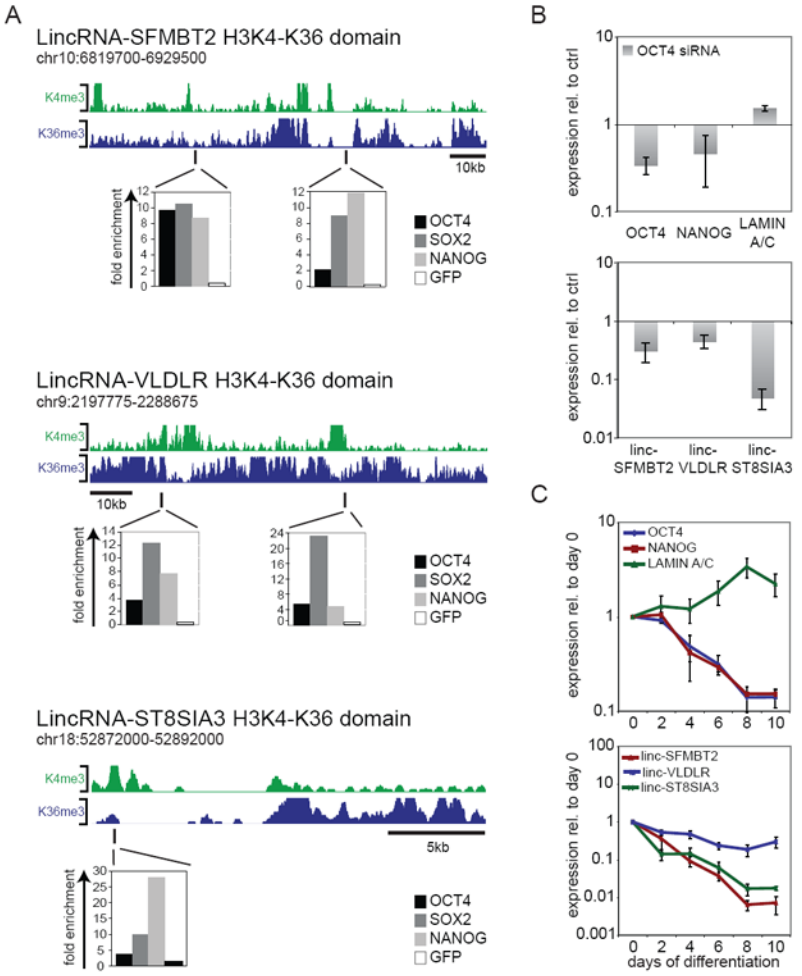

Figure 3.

Transcriptional regulation of iPSC-enriched lincRNAs. A) iPSC-enriched lincRNA loci are bound by pluripotency transcription factors. Top: lincRNA loci demarcated by domains enriched in histone H3K4me3 indicating RNA polymerase II promoters and H3K36me3 indicating regions of transcriptional elongation[10,29] in human ESCs (green and blue, respectively). Bottom: ChIP in hFib2-iPS5 cells followed by quantitative PCR analysis detects binding of OCT4, SOX2, and NANOG within lincRNA-SFMBT2, lincRNA-VLDLR, and lincRNA-ST8SIA3 regions close to lincRNA promoter regions (peaks of H3K4me). ChIP enrichment values are displayed normalized to a control region (chr12: 7,839,777- 7,839,966; hg18); anti-GFP ChIP was used as a negative control. Positions of ChIP-PCR fragments are indicated by black lines. B) Changes in iPSC-enriched lincRNA levels upon siRNA-mediated knock-down of OCT4 in iPSC. Top: qRT-PCR of *OCT4*, *NANOG* and *LMNA* transcript levels upon depletion of OCT4. Bottom: qRT-PCR of iPSC-enriched lincRNA levels upon depletion of OCT4. Transcript levels are displayed relative to non-targeting control siRNAs (ctrl siRNA) (n=3, error bars +/-s.e.m). C) iPSC-enriched lincRNA expression during EB differentiation. Top: qRT-PCR analysis monitoring transcript levels of pluripotency markers (*OCT4* and *NANOG*) and the differentiation marker *LMNA* over a ten day differentiation time-course; bottom: qRT-PCR analysis of iPSC-enriched lincRNAs. RNA levels are depicted relative to undifferentiated cells on day 0 (n=3, error bars +/-s.e.m).

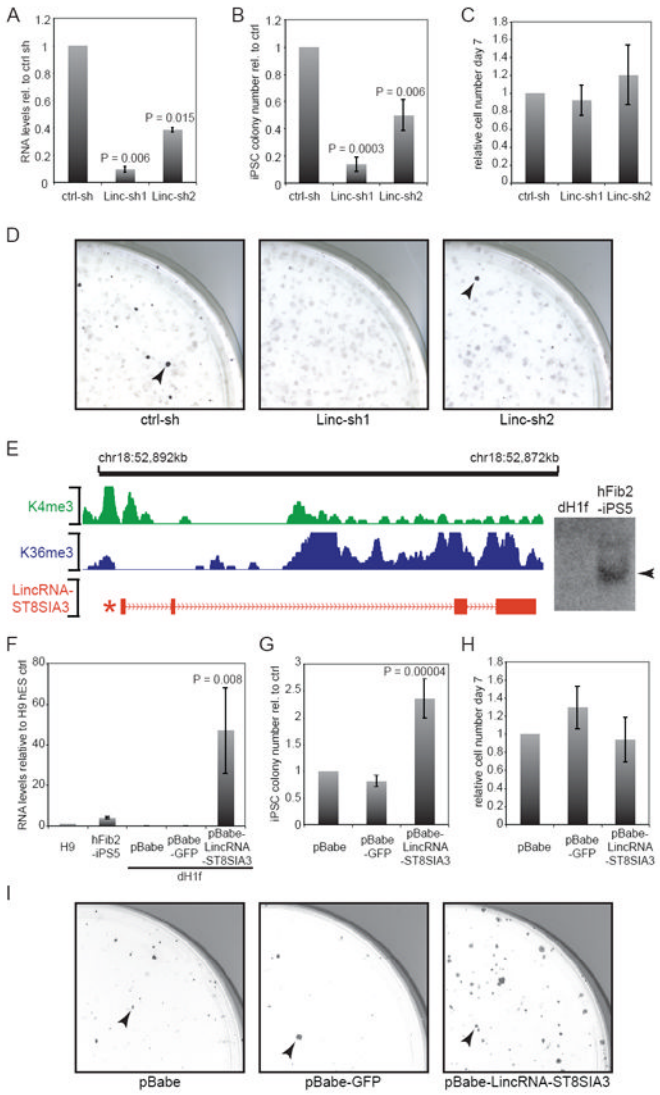

Figure 4.

LincRNA-ST8SIA3 expression modulates reprogramming. A) qRT-PCR verifies lincRNA-ST8SIA3 knock-down with Linc-sh1 and Linc-sh2 in hFib2-iPS5 cells relative to a non-targeting shRNA control (n=2, error bar: +/-s.e.m). B) Quantification of Tra-1-60^{+} iPSC colonies upon knock-down of lincRNA-ST8SIA3 relative to the control (day 21; n=4, error bar: +/-s.e.m). C) Quantification of cell numbers on day 6-7 of reprogramming in lincRNA-ST8SIA3 shRNA samples relative to the control (n=4, error bar: +/-s.e.m). D) Images showing quarters of Tra-1-60 stained reprogramming plates upon infection of a non-targeting control and two lincRNA-ST8SIA3 targeting shRNAs. Arrowheads mark Tra-1-60+ iPSC colonies. E) Structure of the lincRNA-ST8SIA3 locus. Green, Blue: Demarcation of the H3K4me-H3K36me domain in ESCs. Red: Structure of lincRNA-ST8SIA3 RNA; asterisk marks position of OCT4/SOX2/NANOG binding (see Figure 3A). Right: Northern hybridization of lincRNA-ST8SIA3 detects a 2.6kb transcript in hFib2-iSP5, but not dH1f (full-length blot see Supplementary Figure 12). F) qRT-PCR verifies lincRNA-ST8SIA3 overexpression from a retroviral vector (pBabe-lincRNA-ST8SIA3) compared with pBabe-puro and pBabe-puro-GFP vectors in dH1f relative to the levels in H9 ESCs and hFib2-iPS5 (n=2, error bars: +/-s.e.m). G) Quantification of Tra-1-60^{+} iPSC

colonies upon overexpression of lincRNA-ST8SIA3 compared to pBabe and pBabe-GFP controls (n=5, error bar: +/-s.e.m.). H) Quantification of cell numbers on day 6-7 in lincRNA-ST8SIA3 overexpressing cells and controls. Cell numbers are relative to the pBabe control (day 28 +/-2; n=5, error bar: +/- s.e.m.). I) Image of quarter-plates of Tra-1-60 stained colonies (arrowheads) in pBabe, pBabe-GFP, and pBabe-lincRNA-ST8SIA3 infected samples. Statistical analysis was performed using Student's t-test.