Published in final edited form as:
Cell Stem Cell. 2014 May 1; 14(5): 617–631. doi:10.1016/j.stem.2014.01.021.

Two miRNA clusters reveal alternative paths in late-stage reprogramming

Ronald J. Parchem[1,2], **Julia Ye**[1,2], **Robert L. Judson**[1,2], **Marie F. LaRussa**[1,2], **Raga Krishnakumar**[1,2], **Amy Blelloch**[1], **Michael C. Oldham**[1,3], and **Robert Blelloch**[1,2,*]

[1]The Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, Center for Reproductive Sciences, University of California, San Francisco, San Francisco, California, 94143, USA

[2]Department of Urology, University of California, San Francisco, San Francisco, California, 94143, USA

[3]Department of Neurology, University of California, San Francisco, San Francisco, California, 94143, USA

Summary

Ectopic expression of specific factors such as Oct4, Sox2, and Klf4 (OSK) is sufficient to reprogram somatic cells into induced pluripotent stem cells (iPSCs). In this study, we examine the paths taken by cells during the reprogramming process by following the transcriptional activation of two pluripotent miRNA clusters, miR-290 and miR-302, in individual cells in vivo and in vitro using knock-in reporters. During embryonic development and embryonic stem cell differentiation, all cells sequentially expressed miR-290 before miR-302. By contrast, during OSK-induced reprogramming cells activated the miRNA loci in a stochastic, non-ordered manner. However, addition of Sall4 to the OSK cocktail led to a consistent reverse sequence of locus activation (miR-302 then miR-290) and increased reprogramming efficiency. These results demonstrate that cells can follow multiple paths during late stages of reprogramming, and that the trajectory of any individual cell is strongly influenced by the combination of factors introduced.

Introduction

Pluripotent stem cells have tremendous therapeutic potential due to their remarkable ability to self-renew and generate all of the cells of the adult organism. Indeed, reprogramming of

*Correspondence to: blellochr@stemcell.ucsf.edu (RB).

Author Contributions

R.J.P. contributed to Figures 1–7 and Supplementary Figures 1–5; J.Y. to Figure 4 (b, d), Figures 5–7, and Supplementary Figures 4–5; R.L.J. to Figure 3, Figure 4 (a, c) and Supplementary Figure 3 (a–d); M.F.L. to Figure 2 (b, d, e–f) and Supplementary Figure 2 (c); R.K. to Figure 4 (d), Figure 7 and Supplementary Figure 4 (b); A.B. to Supplementary Figure 3 (e); and M.C.O. to Figure 7. R.J.P and R.H.B. wrote the manuscript.

somatic or adult cells into pluripotent stem cells (iPSCs) has already significantly impacted the study of disease and the discovery of potential stem cell-based therapies (Bellin et al., 2012). Although iPSC production has become common, the mechanisms by which reprogramming occurs are still poorly understood. Reprogramming has been depicted as a backwards movement up Waddington's epigenetic landscape along a linear path that traverses intermediate stages of development in reverse (Hochedlinger and Plath, 2009). Consistent with this idea, a number of studies have demonstrated the progressive downregulation of somatic markers and upregulation of pluripotency markers over the course of reprogramming (Brambrink et al., 2008; Stadtfeld et al., 2008; Polo et al., 2012; Chan et al., 2009). Single cell expression analyses for a number of somatic and pluripotency genes suggest an early stochastic, but late hierarchical pattern of gene activation, arguing that after a certain point in the reprogramming process, cells undergo an immutable set of cell fate transitions that lead to the iPSC state (Buganim et al., 2012; Golipour et al., 2012). However, these studies focused only on genes that are expressed either in the starting fibroblast population or in the final pluripotent population, not genes that are uniquely expressed in a stage of development that exists between those two. Therefore, it remains unclear whether reprogramming cells de-differentiate along a path that includes transitions through intervening stages of development on their way to naïve pluripotency. During early murine development, embryonic cells transition from the inner cell mass (ICM) to the early epiblast and then to the late epiblast before differentiating down somatic cell lineages. Mouse embryonic stem cells (ESCs) and iPSCs correspond to the ICM/early epiblast stage of development and collectively represent the naïve pluripotent state (Ying et al., 2008; Evans and Kaufman, 1981; Martin, 1981). Therefore, if iPSC formation represents a progression of development in reverse, one would expect the activation of late epiblast markers before that of naïve markers.

MicroRNAs (miRNAs) are small non-coding RNAs that bind and suppress hundreds of mRNAs simultaneously by destabilizing transcripts and inhibiting their translation (Fabian and Sonenberg, 2012; Bartel, 2009). Like many coding transcripts, miRNA expression is temporally and spatially controlled. In the genome, miRNAs are often arranged in clusters and are expressed as a single primary miRNA transcript that is typically processed by the RNases Drosha and Dicer. Two miRNA clusters in mice, miR-290 and miR-302 (miR-371 and miR-302 in humans, respectively), are highly expressed in pluripotent stem cells (Suh et al., 2004; Houbaviy et al., 2003; Jouneau et al., 2012; Stadler et al., 2010). They have been implicated in the regulation of pluripotent properties and differentiation potential, and they possess the ability to promote de-differentiation of somatic cells to pluripotency (Melton et al., 2010; Judson et al., 2009; Wang et al., 2007; Anokye-Danso et al., 2011; Liao et al., 2011; Subramanyam et al., 2011). Here, we developed reporters for the miR-290 and miR-302 loci, which enabled us to follow individual cells over the course of embryonic development, ESC differentiation, and induced de-differentiation to iPSCs.

Results

MiR-290 and miR-302 define sequential stages of pluripotency

To characterize the expression of miR-290 and miR-302 during development and de-differentiation at the single cell level, both clusters were targeted using homologous recombination to introduce fluorescent reporters into the endogenous miRNA loci, creating a dual reporter system (Fig. 1a, Supplementary Fig. 1a). Constructs were designed such that the miR-290 cluster would be co-expressed with red fluorescent protein (mCherry) and the miR-302 cluster with green fluorescent protein (eGFP). To determine the expression patterns of these miRNA clusters during embryonic development, we examined dual reporter mouse embryos from E3.5 to E8.0, the window of development during which pluripotent cells are present and can be derived (Osorno et al., 2012). We found that all cells of the embryo proper expressed the reporters in a sequential and stereotypical fashion. The ICM cells of E3.5 embryos uniquely expressed miR-290-mCherry (red), followed by expression of both reporters in the epiblast of E5.5 embryos (yellow), and then specifically expressed miR-302-eGFP+ (green) alone in the E7.5 gastrulating embryo (Fig. 1b). MiR-302-eGFP remained broadly expressed through gastrulation until E9.5 (Supplementary Fig. 1d). These data show that miR-290 and miR-302 expression divides the pluripotent cells of the embryo into three consecutive stages: naïve (E3.5, miR-290-mCherry only), pre-primed (E4.5–E6.5, miR-290-mCherry+/miR-302-eGFP+), and primed (E7.5–E8.0, miR-302-eGFP only) (Fig. 1c). Interestingly, miR-290-mCherry was also expressed in the extraembryonic tissues during early development, whereas miR-302-eGFP was limited to the epiblast (Fig. 1b).

To characterize the expression of these two miRNA clusters *in vitro*, ESC and EpiSC lines were derived from dual reporter embryos. Consistent with their origin and like the targeted ESCs (Supplementary Fig. 1b,c), ESCs derived from the ICM of E3.5 blastocysts expressed high levels of miR-290-mCherry under standard mouse ESC culture conditions (FBS+Lif) as determined by fluorescence-activated cell sorting (FACS). However, a small fraction of cells expressed both reporters or eGFP alone (Fig. 2a,b), suggestive of low levels of differentiation and consistent with heterogeneity in Lif alone conditions. Dual inhibition of GSK3β and ERK1/2 using small molecules (CHIR99021 and PD0325901, 2i) resulted in greater homogeneity as measured by a decrease in miR-302-eGFP expression (Fig. 2c) (Marks et al., 2012; Nichols et al., 2009; Ying et al., 2008). In contrast to ESCs, embryonically-derived EpiSCs grown in FGF+Activin expressed high levels of miR-302-eGFP and undetectable levels of miR-290-mCherry (Fig. 2a,b). This expression differed from the E5.5 epiblasts from which they were derived: specifically, during derivation, E5.5 epiblast cells transitioned from expressing both markers to miR-302-eGFP alone as they expanded in culture, suggesting that they had transitioned to a more mature fate (Supplementary Fig. 2a). Similarly, FGF+Activin treatment of naïve miR-290-mCherry+ ESCs led to differentiation, expansion, and long-term maintenance of miR-302-eGFP+ EpiSC cultures (Supplementary Fig. 2b). To further characterize reporter expression during *in vitro* differentiation, ESCs were transferred to differentiation conditions (removal of Lif and 2i, but no FGF+Activin) and analyzed by flow cytometry (Fig. 2d and Supplementary Fig. 2c). Upon differentiation, cells progressed from a naïve (miR-290-mCherry+/miR-302-

eGFP−; red) to a pre-primed (miR-290-mCherry+/miR-302-eGFP+; yellow) and then a primed (miR-290-mCherry−/miR-302-eGFP+; green) state before silencing both loci, consistent with the transitions seen *in vivo*. These changes in miR-290 and miR-302 reporter expression correlated with the differences in gene expression, clonogenicity, and mature miRNA levels observed in naïve and primed pluripotent cells during differentiation (Fig. 2e–g). Importantly, differentiating ESCs showed a significant decrease in colony forming potential, with both miR-290-mCherry+/miR-302-eGFP+ (yellow) and miR-290-mCherry−/miR-302-eGFP+ (green) cells showing a significant reduction in clonogenicity relative to miR-290-mCherry+/miR-302-eGFP− (red) cells. In summary, these data demonstrate that expression of the miR-290 and miR-302 loci can be used to define progressive stages of differentiation through pluripotent states both *in vivo* and *in vitro* (Fig. 2h).

MiR-290 and miR-302 are stochastically activated during OSK reprogramming

Having seen highly ordered sequential expression of the two miRNA loci during early embryonic development and ESC differentiation, we next asked if a similar path is followed in reverse during reprogramming of somatic cells to iPSCs (Fig. 3a). We introduced OSK retroviruses into dual reporter mouse embryonic fibroblasts (MEFs) and followed the reprogramming process in several commonly used media conditions every other day for 16 days by fluorescence time-lapse microscopy (Supplementary Fig. 3a). Notably, virtually all colonies that grew to a size detectable at low magnification were miR-290-mCherry+/miR-302-eGFP− (red), indicative of naïve pluripotency (Supplementary Fig. 3a). Interestingly, time-lapsed examination of reprogramming wells at higher magnification revealed evidence of multiple patterns of activation of the miRNA loci preceding colony formation. Many patches of cells showed no activation of miR-302-eGFP during colony formation, while some colonies arose from patches expressing miR-302-eGFP and others from a heterogeneous mixture of miR-290-mCherry+/miR-302-eGFP−, miR-290-mCherry+/miR-302-eGFP+, and miR-290-mCherry−/miR-302-eGFP+ intermediates (Fig. 3b). By the time patches formed a compact growing colony, however, the cells were predominantly miR-290-mCherry+/miR-302-eGFP− (red) (Fig. 3c). After 16 days, colonies were isolated, expanded and tested for pluripotency. Cells from these colonies were miR-290-mCherry+/miR-302-eGFP− and stained positive for other naïve markers such as Nanog and SSEA1 (Supplementary Fig. 3b). Furthermore, they demonstrated properties of fully reprogrammed naïve iPSCs, including silencing of the retroviruses, activation of endogenous pluripotency genes, and production of chimeras that went germline (Supplementary Fig. 3c–e).

To further characterize the activation of these loci over time, we performed flow cytometry to determine the distribution of miR-290-mCherry and miR-302-eGFP expression on a daily basis from days 5–14 post-infection (Fig. 3d,e). Fluorescence was first detected starting around day 7. Surprisingly, there were similar numbers of cells expressing either miR-290-mCherry or miR-302-eGFP, suggesting stochastic activation of these two loci. Two days later, a small number of cells were found to be miR-290-mCherry+/miR-302-eGFP+ (yellow), presumably arising from either miR-290-mCherry+/miR-302-eGFP− (red) or miR-290-mCherry−/miR-302-eGFP+ (green) single positive parents (Fig. 3d,e). Together the microscopy and flow cytometry results suggest that in contrast to normal differentiation,

these miRNA loci are not activated in an ordered manner during late stages of de-differentiation.

The number of activated miRNA loci, not their order of expression, predicts iPSC colony formation potential in OSK reprogramming

Next, we asked whether miR-290 expression, which is uniquely expressed in naïve cells could be used to prospectively identify more fully reprogrammed cells. Cells were sorted at day 12, the first time point at which there were enough cells expressing all potential combinations of reporters to perform the analysis and early enough that it was unlikely that individual cells had passed through expression of both markers. Sorted cells were plated onto irradiated MEFs and evaluated for iPSC colony number 5–6 days later (Fig. 4a). As expected, cells expressing neither miRNA loci (black) formed colonies with the lowest efficiency (0.03%). Suprisingly though, miR-290-mCherry+/miR-302-eGFP− (red) and miR-290-mCherry−/miR-302-eGFP+ (green) sorted cells displayed a similar efficiency, 1.2 and 0.6%, respectively (Fig. 4b). The relatively rare (0.015% of all live cells) double positive (yellow) cells demonstrated by far the highest colony forming efficiency (12%), a 400-fold increase over black cells (Fig. 4b). Similar results were obtained when cells were sorted at day 14 (Fig. 4c). The fidelity of the FACS approach was validated by re-sorting each sorted population prior to plating (Supplementary Fig. 4a). We also performed mRNA profiling of miR-290-mCherry−/miR-302-eGFP− (black), miR-290-mCherry+/miR-302-eGFP− (red), miR-290-mCherry+/miR-302-eGFP+ (yellow), and miR-290-mCherry−/miR-302-eGFP+ (green) cells during de-differentiation to compare their overall molecular constitution. Principal component analysis and hierarchical clustering showed that (1) the three replicate sorted cell populations were most similar within each color group, (2) miR-290-mCherry+/miR-302-eGFP− (red) and miR-290-mCherry−/miR-302-eGFP+ (green) cells were more like each other than they were to miR-290-mCherry+/miR-302-eGFP+ (yellow) cells, and (3) miR-290-mCherry+/miR-302-eGFP+ (yellow) cells were the most similar to iPSCs (Fig. 4d, Supplementary Fig. 4b). Therefore, the global profiles were consistent with the relative potentials of the reprogramming populations. Together, these findings show that in OSK reprogramming conditions, it is the number rather than order of the two miRNA loci that predicts the extent of reprogramming even though one of these loci (miR-302) is not expressed in the final naive state.

Sall4 promotes sequential expression of miR-302 and miR-290 in reprogramming

Next, we asked whether the addition of other reprogramming factors would influence the order of miRNA expression. We focused on Sall4, as it is highly expressed in naïve ESCs, is a major regulator of the naïve state, and is a powerful enhancer of reprogramming (Tsubooka et al., 2009; Sakaki-Yumoto et al., 2006; Lim et al., 2008; Zhang et al., 2006; Elling et al., 2006) To evaluate its effect on the expression of the two miRNA loci, Sall4 was introduced together with OSK into dual reporter MEFs, and cells were sorted at days 6, 9, 12, 15, and 18 for the expression of miR-290-mCherry and miR-302-eGFP. The addition of Sall4 increased the frequency and kinetics of reporter activation (Fig. 5a), and induced a minimal ten-fold increase in iPSC colony number (Fig. 5b, Supplementary Fig. 5a). There was also an approximately 80-fold increase in the number of cells expressing both miRNA loci by day 12 (Fig. 5c). Surprisingly, in contrast to the stochastic activation of the two

miRNA loci observed under OSK-only conditions, the addition of Sall4 resulted in a strong bias toward early activation of the miR-302 locus relative to the miR-290 locus, as demonstrated by the increase in miR-290-mCherry−/miR-302-eGFP+ cells captured by FACS (Fig. 5d). Moreover, the small number of miR-290-mCherry+/miR-302-eGFP− (red) cells arising from Sall4+OSK-transduced MEFs showed a much higher colony forming efficiency than those transduced with OSK alone (Fig. 5e). Indeed, miR-290-mCherry+/miR-302-eGFP− (red) cells formed colonies at a comparable efficiency as double positive miR-290-mCherry+/miR-302-eGFP+ (yellow) cells. We confirmed that the mature miRNA levels corresponded with the color of the sorted cells via RT-qPCR at both the population (Supplementary Fig. 5b) and the single-cell level (Fig. 5f, Supplementary Fig. 5c). MiRNA and fluorescence levels were largely correlated. Occasional detection of small amounts of mature miRNAs in cells that did not express the corresponding fluorescent marker is likely due to differences in the maturation rates and detection methods of miRNAs and fluorescent proteins. These findings suggested that in contrast to OSK reprogramming, Sall4+OSK reprogramming followed a more ordered path of activation of the two miRNA loci and was associated with a higher efficiency in reprogramming.

A limitation of flow cytometry is that it cannot follow individual cells over time. Thus, it remained possible that differences in the order of miRNA expression under the two conditions were due to heterogeneous timing of locus activation among cells rather than within cells. To address this, we performed high-resolution time-lapse microscropy to lineage trace individual cells every 2–4 hours over several days (Figure 6 and Supplementary Movies 1–6). Consistent with our results from flow cytometry, lineage tracing showed that under OSK conditions, comparable numbers of cells went directly from expressing neither marker (black) to expressing either miR-302-GFP+ (green) or miR-290-mCherry+ (red). In contrast to OSK, Sall4+OSK transduced cells mostly transitioned from black to green and then yellow before settling into a final red state. Only a small number of cells transitioned from black to red and from red to yellow. These findings confirm that within developing iPSC colonies, cells take multiple paths to the naïve state under OSK conditions, but a more singular path under Sall4+OSK conditions.

High iPSC potential is associated with activation of a naïve cell expression signature

Next we asked whether the differences in iPSC potential between early miR-290-mCherry+/miR-302-eGFP− (red) cells under OSK versus Sall4+OSK could be correlated with a more global transcriptional signature. In particular, we hypothesized that the rare early mCherry+/miR-302-eGFP− (red) cells in Sall4+OSK conditions, which presumably arose directly from mCherry−/miR-302-eGFP− (black) cells and yet have a very high iPSC colony forming potential, would have a transcriptional profile more like that of miR-290-mCherry+/miR-302-eGFP+ (yellow) cells than of mCherry−/miR-302-eGFP+ (green) cells. However, both PCA and hierarchical clustering of mRNA profiles from the day 9 Sall4+OSK cell populations showed similar results to OSK (Figure 7a,b). That is, despite the dramatic differences in iPSC potential, the overall mRNA profiles of miR-290-mCherry+/miR-302-eGFP− (red) and miR-290-mCherry−/miR-302-eGFP+ (green) cells remained most similar to each other than to that of miR-290-mCherry+/miR-302-eGFP+ (yellow) cells, regardless of the reprogramming factors used. In both conditions, miR-290-mCherry+/miR-302-eGFP+

(yellow) cells continued to have a profile closest to that of iPSCs. Therefore, we asked if there was a signature common to cells with the greatest iPSC potential (OSK yellow, Sall4+OSK red and yellow) versus those with lower potential (OSK red and greeen, Sall4+OSK green). Specifically, using a p-value cutoff of 0.01 and log2 fold change greater than 0.4, we identified differentially-expressed genes in Sall4+OSK red versus OSK red, Sall4+OSK red and yellow versus green, and OSK yellow versus red and green, yielding a total of 312 differentially expressed genes. Of these genes, 14 were found to overlap in these comparisons (Fig. 7c,d). We next asked if these 14 genes were enriched in the naïve versus primed state by profiling mRNA expression of dual reporter cells during ESC differentiation from red to yellow to green as well as derived green EpiSCs. Indeed, the 14 gene signature was downregulated during the naïve to primed pluripotent transition (Fig. 7e). In contrast, the non-overlapping genes showed little change. Consistent with differential expression in cell populations with the highest iPSC potential, a number of the genes in the list have been previously identified as regulators of naive pluripotency, including Dnmt3l, Dnmt3b, Esrrb, Dppa5, Dppa4, Ooep/Moep19 and Rhox5 (Fig. 7d) (Okano et al., 1999; Amano et al., 2006; Madan et al., 2009; Tashiro et al., 2010; Festuccia et al., 2012; Martello et al., 2012; Neri et al., 2013). Thus, these profiling data identify a set of genes that are (1) co-regulated with the miR-290, but not the miR-302 locus during Sall4+OSK reprogramming, (2) are enriched in naïve versus primed cells, and (3) correlate with iPSC colony forming potential.

Discussion

MicroRNAs are evolutionarily conserved regulators of cellular identity (Christodoulou et al., 2010; Heimberg et al., 2008). Therefore, changes in miRNA expression can be used to follow the progression of development and differentiation. Our data show that changes in miR-290 and miR-302 expression delineate the specific cell fate transitions separating naïve from primed pluripotency during differentiation both *in vivo* and *in vitro*. Previous work had shown enriched expression of miR-290 and miR-302 in ESC and EpiSC populations, respectively (Stadler et al., 2010; Jouneau et al., 2012; Card et al., 2008), but it was unclear how these expression patterns were reflected at the single cell level. Here, we show that pluripotent cells of the embryo transition through three states marked by miR-290 alone, miR-290 and miR-302 concurrently, and miR-302 alone, before silencing both loci in most if not all somatic populations. These transitions correlate with important developmental events such as implantation (activation of miR-302) and gastrulation (downregulation of miR-290).

Using the activation of miR-290 and miR-302 as readouts of the developmental status of individual cells, we found that de-differentiation during OSK reprogramming does not require activation of genes in a definite order, but instead appears to take varying paths to the naïve state. Moreover, the ability of a reprogramming cell to acquire naïve pluripotency is best predicted by the expression of both miRNA loci. Reprogramming cells that have activated both miR-290 and miR-302 showed greater molecular similarity to naïve ESCs than either single positive population when using either OSK or Sall4+OSK. Functionally, double positive cells had the greatest ability to form naïve colonies, which stands in contrast with normal differentiation, in which double positive cells showed reduced colony forming ability relative to miR-290-mCherry single positive cells. Together, the colony forming

assay and molecular profiling of intermediates demonstrate that activation of miR-290-mCherry alone does not necessarily mark the completion of reprogramming. Instead, while some of the miR-290-mCherry single positive cells are likely fully reprogrammed iPSCs, others represent reprogramming intermediates that fail to form colonies. However, the addition of Sall4, an enhancer of reprogramming, biased the process such that most cells activated miR-302 before miR-290. Moreover, miR-290-mCherry single positive cells showed a higher colony forming efficiency than miR-302-eGFP single positive cells, which is evidence that a larger fraction of Sall4+OSK-treated miR-290-mCherry+ (red) cells are functionally closer to the iPSC state than their OSK-treated counterparts. The ability of Sall4 to activate miR-302 early cannot be simply explained by direct activation of the locus by this transcription factor for two reasons: first, there remains a long delay from time of introduction of factors and miR-302 activation. Second, Sall4 is more highly expressed in the naïve (miR-302−) state than in the primed (miR-302+) state (Fig. 2e). Interestingly, it was previously shown that Oct4 and Sox2 bind to the promoter of miR-302 and are required for its expression, but clearly they are not sufficient since Oct4 and Sox2 are also highly expressed in the naïve (miR-302−) state (Card et al., 2008; Marson et al., 2008). Future studies aimed at understanding the transcriptional network upstream of miR-302 should provide important insight into the drivers of the primed state.

Together, our results show that the path taken by cells during reprogramming is influenced by the combination of factors used. Adding Sall4 promoted a defined trajectory, possibly by reducing stochasticity in the later stages of reprogramming. Two other recent papers have addressed the transcriptional events during reprogramming at the single cell level. By performing single cell RT-qPCR analysis and single molecule FISH for various fibroblast and naïve pluripotency markers, Buganim et al., 2012 provided evidence for early stochastic and late hierarchical events in reprogramming. Similarly, using population and single cell profiling following activation of OSK+Myc (OSKM), Polo et al., 2012 found that reprogramming can occur in two phases: an early phase driven by cMyc and Klf4 and a later phase driven by Oct4, Sox2, and Klf4. Interestingly, using population-based profiling for miRNAs during reprogramming, Polo et al., 2012 showed early miR-290 and later miR-302 expression. How this correlates with single cell expression of the loci is unclear; however the population bias toward miR-290 early in reprogramming contrasts to our single cell findings of either equal activation using OSK or a bias toward early miR-302 when Sall4 was included. Once again, this may reflect the combination of factors used. Indeed, Myc has been shown to bind to the miR-290 promoter (Judson et al., 2009); therefore its addition to the reprogramming cocktail could drive early miR-290 expression. Importantly, our results demonstrate that any analysis of reprogramming, including the path taken and the functional analysis of genes involved, can be dependent upon the specific factors used. For example, although miR-302 has been studied extensively in the context of reprogramming, our results suggest its expression is not required during reprogramming to naïve pluripotency.

Both single cell studies described above follow genes expressed in the initial fibroblast population and the terminal naïve pluripotent state. The miR-302 locus functions as a unique marker because it is not expressed in either fibroblasts or ESCs, and yet, as we show here, cells pass through a miR-302+ intermediate during ESC differentiation and embryonic development. Using two cell surface markers to perform population-based profiling on cells

enriched for reprogramming potential, another group recently showed a transient upregulation of a number of epidermis-related genes during reprogramming (O'Malley et al., 2013). However, the expression of these genes during embryonic development or ESC differentiation is not well-characterized. Therefore, it remains unclear whether these markers represent intermediate developmental states. In contrast, our studies show that miR-302 is an obligatory intermediate stage during embryonic development but not during reprogramming. Similar to the Buganim et al., 2012 and Polo et al., 2012 studies, the O'Malley et al., 2013 study uses an inducible OSKM secondary MEF system in which it is expected that every cell will express the four factors at similar levels. Interestingly, previous work has found that the relative expression of each factor is affected by the order in which they are cloned into the inducible cassette, influencing the final molecular state of the iPSCs (Carey et al., 2011). In our study, we use a retroviral system, which is likely to lead to a degree of heterogeneity in the relative expression of factors and hence timing of activation of downstream networks between cells. For this reason, it was especially surprising to uncover the relative homogeneous transitions with the addition of Sall4. In summary, it is unlikely that any particular iPSC intermediate, either stable or transient, marks a required common trajectory taken by all cells that fully reprogram. Instead, there are likely multiple paths, and the choice of trajectory is influenced by the combination of factors introduced and their relative expression.

Functional roles for the miR-290 and miR-302 clusters have been surmised through a number of *in vitro* approaches both in mouse and human ESCs (Melton et al., 2010; Rosa et al., 2009; Kim et al., 2011; Wang et al., 2008; Medeiros et al., 2011; Sinkkonen et al., 2008; Benetti et al., 2008; Qi et al., 2009). In particular, both clusters produce multiple miRNAs that share a common seed sequence forming a large family of miRNAs called the ES cell-specific cell cycle regulating (ESCC) miRNAs (Wang et al., 2008). These miRNAs have been shown to functionally regulate the cell cycle and suppress differentiation of mouse pluripotent stem cells (Wang et al., 2008; Melton et al., 2010) and suppress neural differentiation during hESC differentiation (Kim et al., 2011). In most of these roles, the various members of the ESCC family appear to act redundantly. Quite remarkably, the forced expression of the miR-302 cluster alone has been shown to de-differentiate somatic cells to naïve mouse iPSCs (Anokye-Danso et al., 2011), even though as we show here, the cluster itself is not expressed in fully reprogrammed mouse iPSCs. Introduction of exogenous miR-290/302 ESCC miRNAs along with OSK leads to increased reprogramming efficiency (Judson et al., 2009). In this context, it appears miR-290/302 enhances the mesenchymal-to-epithelial transition and leads to an overall reduction in stochastic gene activation that typifies the early phase of reprogramming (Subramanyam et al., 2011; Judson et al., 2013).

The existence of the two clusters appears to be unique to placental mammals (Houbaviy et al., 2005). Our findings demonstrate that the two clusters possess distinct temporal and spatial expression patterns along with a partially redundant pattern in the early epiblast. Genetic deletion of the miR-290 cluster results in a variable phenotype with incomplete embryonic lethality (Medeiros et al., 2011), which could in part be explained by early partial rescue from the miR-302 cluster. However, such a conclusion awaits the production of a

miR-302 knockout model. Since miR-290 is also expressed in extraembryonic structures during early development, an ancestral miR-302-like cluster may have duplicated and evolved to function in the extra-embryonic lineage that gave rise to the placenta. The two clusters have maintained largely redundant functions, therefore the ability to independently regulate ESCC miRNA expression from two clusters in embryonic versus extra-embryonic lineages may have played a role in placental evolution.

The finding that an enhancer of reprogramming leads to biased activation of miR-302 before miR-290 in a manner that would be expected for developmental reversal leaves open the possibility that, while not required, a sequential unfolding of the differentiation program in reverse increases the efficiency of iPSC generation. Such an idea is not unreasonable, considering the fact that directed differentiation of therapeutic cell types from pluripotent stem cells often involves a recapitulation of developmental events. Indeed, the high iPSC potential-associated genes that we identified are expressed differentially in naive and primed pluripotency and are upregulated with increasing iPSC potential in reprogramming intermediates in a manner consistent with a reversal of normal differentiation. Recently, it was shown that removal of Mbd3, a NuRD complex member, allows OSKM to drive reprogramming in a highly efficient and deterministic manner (Rais et al., 2013). It will be interesting to see how timing of expression of the miR-290 and miR-302 loci are influenced by the loss of Mbd3 in the context of different reprogramming factors and means of introduction.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Nicole Moore for technical support, Grace Wei for assistance with the BioStation CT and Diana Laird for critical reading of the manuscript. This work was supported by funds to R.H.B from the National Institutes of Health (R01: GM101180), the Helmsley Trust (09PG-T1D002), and the California Institute of Regenerative Medicine (New Faculty Award: RN2-00906-1). R.J.P. was supported by NIH post-doctoral fellowships (T32 HD 007263 -27, F32 HD 070572 -01.); R.L.J. by a UCSF Julius R. and Patricia A. Krevans Fellowship; and R.K. by an A.P. Giannini Foundation post-doctoral fellowship.

Abbreviated Experimental Procedures (See Supplemental Information for more detailed methods)

Targeting strategy — *MiR-290-mCherry/miR-302-eGFP* (dual reporter) ES cells were generated by inserting coding regions for eGFP and mCherry downstream of the transcriptional start sites of miR-302 and miR-290, respectively

ESC and EpiSC Derivation — For ESCs, timed matings of dual reporter mice were used to isolate blastocysts at E3.5. Blastocyst outgrowths were trypsinized and passaged every 3–4 days until ESC lines were established. For EpiSCs, timed matings of dual reporter mice were used to isolate embryos between E5.5–E7.5. Individual epiblasts were dissected

	and plated onto MEF feeder layers. The undifferentiated portions of outgrowths were picked and passaged every 3–4 days until homogeneous cultures were obtained
Reprogramming/De-differentiation	MEFs were treated with retroviruses expressing Oct4, Sox2, Klf4 (OSK) +/− Sall4 and cultured in standard reprogramming media. Colony formation was assessed using high throughput imaging and high content analysis
iPS Colony Formation Assay	Over the course of reprogramming, individual wells were trypsinized, sorted by FACS, and plated in 96-well plates with reprogramming medium for daily imaging to follow colony formation. The INCell Developer software (GE) was used to count the number of colonies and to quantify the area of miR-290-mCherry and miR-302-eGFP expression in each well. The ratio of colonies formed to the number of cells plated was defined as the iPSC colony forming potential
mRNA Profiling	Expression analysis was done on Illumina Bead Arrays. Array data can be found at GEO accession # …… (request submitted and will add # to proof). Total RNA was isolated from sorted populations. Data were preprocessed, and quality control was performed using BeadArray software and the SampleNetwork R function (Oldham et al., 2012). This analysis revealed no sample outliers among the biological replicates. Following quantile normalization, a batch effect corresponding to the microarray ID was corrected using ComBat (Johnson et al., 2007). Significant changes in gene expression between sets ($P < 0.01$, log2 fold change > 0.4) were determined using Limma. For principal component analysis, the prcomp command in R was used with the default S3 method. For nearest neighbor analysis, hierarchical clustering using the hclust command in R was performed with the ‘single’ method
RT-qPCR	Messenger RNA was reverse transcribed with oligo-dT primers (Invitrogen Superscript Kit). Gene-specific primer sets (500 nM) and Power SYBR Green PCR master mix (ABI) were used for qPCR. Endogenous and exogenous Oct4, Sox2, and Klf4 primers were previously described (Judson et al., 2009). Population-level RT-qPCR for miRNAs was performed using the polyA/SYBRGreen method as previously described (Shi and Chiang, 2005). Primer sequences are listed in Supplementary Experimental Procedures
Single cell miRNA RT-qPCR	Single cells obtained from reprogramming dual reporter MEFs at day 9 (Sall4+OSK) or day 12 (OSK) were sorted into 96-well PCR plates, and amplified cDNA was generated using a method

	modified from Moltzahn et al., 2011. qPCR was performed using the TaqMan approach
ESC Differentiation	ESCs were maintained in FBS+Lif or FBS+Lif+2i conditions, and differentiation was induced by removing Lif and 2i
ESC Clonogenecity Assay	ESCs differentiating in FBS-Lif-2i conditions were sorted for mCherry and eGFP expression into standard ES medium and imaged over several days to count the number of colonies that formed
High-resolution time-lapse microscopy	OSK and Sall4+OSK reprogramming wells were imaged at 10X magnification every 2–4 hours beginning at day 9 after retroviral transduction using a BioStation CT (Nikon). CL-Quant software (Nikon) was used to analyze the data and prepare movies
Animal Use	All animal experiments were approved by the Institutional Animal Care and Use Committee of the University of California San Francisco

References

Amano H, Itakura K, Maruyama M, Ichisaka T, Nakagawa M, Yamanaka S. Identification and targeted disruption of the mouse gene encoding ESG1 (PH34/ECAT2/DPPA5). BMC Dev Biol. 2006; 6:11. [PubMed: 16504174]

Anokye-Danso F, Trivedi CM, Juhr D, Gupta M, Cui Z, Tian Y, Zhang Y, Yang W, Gruber PJ, et al. Highly Efficient miRNA-Mediated Reprogramming of Mouse and Human Somatic Cells to Pluripotency. Cell Stem Cell. 2011; 8:376–388. [PubMed: 21474102]

Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell. 2009; 136:215–233. [PubMed: 19167326]

Bellin M, Marchetto MC, Gage FH, Mummery CL. Induced pluripotent stem cells: the new patient? Nat Rev Mol Cell Biol. 2012; 13:713–726. [PubMed: 23034453]

Benetti R, Gonzalo S, Jaco I, Muñoz P, Gonzalez S, Schoeftner S, Murchison E, Andl T, Chen T, et al. A mammalian microRNA cluster controls DNA methylation and telomere recombination via Rbl2-dependent regulation of DNA methyltransferases. Nat Struct Mol Biol. 2008; 15:268–279. [PubMed: 18311151]

Brambrink T, Foreman R, Welstead GG, Lengner CJ, Wernig M, Suh H, Jaenisch R. Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. Cell Stem Cell. 2008; 2:151–59. [PubMed: 18371436]

Buganim Y, Faddah DA, Cheng AW, Itskovich E, Markoulaki S, Ganz K, Klemm SL, van Oudenaarden A, Jaenisch R. Single-Cell Expression Analyses during Cellular Reprogramming Reveal an Early Stochastic and a Late Hierarchic Phase. Cell. 2012; 150:1209–222. [PubMed: 22980981]

Card DA, Hebbar PB, Li L, Trotter KW, Komatsu Y, Mishina Y, Archer TK. Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells. Mol Cell Biol. 2008; 28:6426–438. [PubMed: 18710938]

Carey BW, Markoulaki S, Hanna JH, Faddah DA, Buganim Y, Kim J, Ganz K, Steine EJ, Cassady JP, et al. Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells. Cell Stem Cell. 2011; 9:588–598. [PubMed: 22136932]

Chan EM, Ratanasirintrawoot S, Park IH, Manos PD, Loh YH, Huo H, Miller JD, Hartung O, Rho J, et al. Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. Nat Biotechnol. 2009; 27:1033–37. [PubMed: 19826408]

Christodoulou F, Raible F, Tomer R, Simakov O, Trachana K, Klaus S, Snyman H, Hannon GJ, Bork P, Arendt D. Ancient animal microRNAs and the evolution of tissue identity. Nature. 2010; 463:1084–88. [PubMed: 20118916]

Elling U, Klasen C, Eisenberger T, Anlag K, Treier M. Murine inner cell mass-derived lineages depend on Sall4 function. Proc Natl Acad Sci U S A. 2006; 103:16319–324. [PubMed: 17060609]

Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. Nature. 1981; 292:154–56. [PubMed: 7242681]

Fabian MR, Sonenberg N. The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC. Nat Struct Mol Biol. 2012; 19:586–593. [PubMed: 22664986]

Festuccia N, Osorno R, Halbritter F, Karwacki-Neisius V, Navarro P, Colby D, Wong F, Yates A, Tomlinson SR, Chambers I. Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. Cell Stem Cell. 2012; 11:477–490. [PubMed: 23040477]

Golipour A, David L, Liu Y, Jayakumaran G, Hirsch CL, Trcka D, Wrana JL. A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network. Cell Stem Cell. 2012; 11:769–782. [PubMed: 23217423]

Heimberg AM, Sempere LF, Moy VN, Donoghue PCJ, Peterson KJ. MicroRNAs and the advent of vertebrate morphological complexity. Proc Natl Acad Sci USA. 2008; 105:2946–950. [PubMed: 18287013]

Hochedlinger K, Plath K. Epigenetic reprogramming and induced pluripotency. Development. 2009; 136:509–523. [PubMed: 19168672]

Houbaviy HB, Dennis L, Jaenisch R, Sharp PA. Characterization of a highly variable eutherian microRNA gene. RNA. 2005; 11:1245–257. [PubMed: 15987809]

Houbaviy HB, Murray MF, Sharp PA. Embryonic stem cell-specific MicroRNAs. Dev Cell. 2003; 5:351–58. [PubMed: 12919684]

Jouneau A, Ciaudo C, Sismeiro O, Brochard V, Jouneau L, Vandormael-Pournin S, Coppée JY, Zhou Q, Heard E, et al. Naive and primed murine pluripotent stem cells have distinct miRNA expression profiles. RNA. 2012; 18:253–264. [PubMed: 22201644]

Judson RL, Babiarz JE, Venere M, Blelloch R. Embryonic stem cell-specific microRNAs promote induced pluripotency. Nat Biotechnol. 2009; 27:459–461. [PubMed: 19363475]

Judson RL, Greve TS, Parchem RJ, Blelloch R. MicroRNA-based discovery of barriers to dedifferentiation of fibroblasts to pluripotent stem cells. Nat Struct Mol Biol. 2013; 20:1227–235. [PubMed: 24037508]

Kim H, Lee G, Ganat Y, Papapetrou EP, Lipchina I, Socci ND, Sadelain M, Studer L. miR-371-3 expression predicts neural differentiation propensity in human pluripotent stem cells. Cell Stem Cell. 2011; 8:695–706. [PubMed: 21624813]

Liao B, Bao X, Liu L, Feng S, Zovoilis A, Liu W, Xue Y, Cai J, Guo X, et al. MicroRNA cluster 302-367 enhances somatic cell reprogramming by accelerating a mesenchymal-to-epithelial transition. J Biol Chem. 2011; 286:17359–364. [PubMed: 21454525]

Lim CY, Tam WL, Zhang J, Ang HS, Jia H, Lipovich L, Ng HH, Wei CL, Sung WK, et al. Sall4 regulates distinct transcription circuitries in different blastocyst-derived stem cell lineages. Cell Stem Cell. 2008; 3:543–554. [PubMed: 18804426]

Madan B, Madan V, Weber O, Tropel P, Blum C, Kieffer E, Viville S, Fehling HJ. The pluripotency-associated gene Dppa4 is dispensable for embryonic stem cell identity and germ cell development but essential for embryogenesis. Mol Cell Biol. 2009; 29:3186–3203. [PubMed: 19332562]

Marks H, Kalkan T, Menafra R, Denissov S, Jones K, Hofemeister H, Nichols J, Kranz A, Stewart AF, et al. The transcriptional and epigenomic foundations of ground state pluripotency. Cell. 2012; 149:590–604. [PubMed: 22541430]

Marson A, Levine SS, Cole MF, Frampton GM, Brambrink T, Johnstone S, Guenther MG, Johnston WK, Wernig M, et al. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell. 2008; 134:521–533. [PubMed: 18692474]

Martello G, Sugimoto T, Diamanti E, Joshi A, Hannah R, Ohtsuka S, Göttgens B, Niwa H, Smith A. Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. Cell Stem Cell. 2012; 11:491–504. [PubMed: 23040478]

Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci USA. 1981; 78:7634–38. [PubMed: 6950406]

Medeiros LA, Dennis LM, Gill ME, Houbaviy H, Markoulaki S, Fu D, White AC, Kirak O, Sharp PA, et al. Mir-290-295 deficiency in mice results in partially penetrant embryonic lethality and germ cell defects. Proc Natl Acad Sci U S A. 2011; 108:14163–68. [PubMed: 21844366]

Melton C, Judson RL, Blelloch R. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. Nature. 2010; 463:621–26. [PubMed: 20054295]

Moltzahn F, Hunkapiller N, Mir AA, Imbar T, Blelloch R. High throughput microRNA profiling: optimized multiplex qRT-PCR at nanoliter scale on the fluidigm dynamic arrayTM IFCs. J Vis Exp. 2011

Neri F, Krepelova A, Incarnato D, Maldotti M, Parlato C, Galvagni F, Matarese F, Stunnenberg HG, Oliviero S. Dnmt3L Antagonizes DNA Methylation at Bivalent Promoters and Favors DNA Methylation at Gene Bodies in ESCs. Cell. 2013; 155:121–134. [PubMed: 24074865]

Nichols J, Silva J, Roode M, Smith A. Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. Development. 2009; 136:3215–222. [PubMed: 19710168]

Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell. 1999; 99:247–257. [PubMed: 10555141]

O'Malley J, Skylaki S, Iwabuchi KA, Chantzoura E, Ruetz T, Johnsson A, Tomlinson SR, Linnarsson S, Kaji K. High-resolution analysis with novel cell-surface markers identifies routes to iPS cells. Nature. 2013; 499:88–91. [PubMed: 23728301]

Osorno R, Tsakiridis A, Wong F, Cambray N, Economou C, Wilkie R, Blin G, Scotting PJ, Chambers I, Wilson V. The developmental dismantling of pluripotency is reversed by ectopic Oct4 expression. Development. 2012; 139:2288–298. [PubMed: 22669820]

Polo JM, Anderssen E, Walsh RM, Schwarz BA, Nefzger CM, Lim SM, Borkent M, Apostolou E, Alaei S, et al. A molecular roadmap of reprogramming somatic cells into iPS cells. Cell. 2012; 151:1617–632. [PubMed: 23260147]

Qi J, Yu JY, Shcherbata HR, Mathieu J, Wang AJ, Seal S, Zhou W, Stadler BM, Bourgin D, et al. microRNAs regulate human embryonic stem cell division. Cell Cycle. 2009; 8:3729–741. [PubMed: 19823043]

Rais Y, Zviran A, Geula S, Gafni O, Chomsky E, Viukov S, Mansour AA, Caspi I, Krupalnik V, et al. Deterministic direct reprogramming of somatic cells to pluripotency. Nature. 2013; 502:65–70. [PubMed: 24048479]

Rosa A, Spagnoli FM, Brivanlou AH. The miR-430/427/302 family controls mesendodermal fate specification via species-specific target selection. Dev Cell. 2009; 16:517–527. [PubMed: 19386261]

Sakaki-Yumoto M, Kobayashi C, Sato A, Fujimura S, Matsumoto Y, Takasato M, Kodama T, Aburatani H, Asashima M, et al. The murine homolog of SALL4, a causative gene in Okihiro syndrome, is essential for embryonic stem cell proliferation, and cooperates with Sall1 in anorectal, heart, brain and kidney development. Development. 2006; 133:3005–013. [PubMed: 16790473]

Shi R, Chiang VL. Facile means for quantifying microRNA expression by real-time PCR. Biotechniques. 2005; 39:519–525. [PubMed: 16235564]

Sinkkonen L, Hugenschmidt T, Berninger P, Gaidatzis D, Mohn F, Artus-Revel CG, Zavolan M, Svoboda P, Filipowicz W. MicroRNAs control de novo DNA methylation through regulation of transcriptional repressors in mouse embryonic stem cells. Nat Struct Mol Biol. 2008; 15:259–267. [PubMed: 18311153]

Stadler BM, Ivanovska I, Mehta K, Song S, Nelson A, Tan Y, Mathieu J, Darby GC, Blau CA, et al. Characterization of microRNAs Involved in Embryonic Stem Cell States. Stem Cells Dev. 2010

Stadtfeld M, Maherali N, Breault DT, Hochedlinger K. Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. Cell Stem Cell. 2008; 2:230–240. [PubMed: 18371448]

Subramanyam D, Lamouille S, Judson RL, Liu JY, Bucay N, Derynck R, Blelloch R. Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells. Nat Biotechnol. 2011; 29:443–48. [PubMed: 21490602]

Suh MR, Lee Y, Kim JY, Kim SK, Moon SH, Lee JY, Cha KY, Chung HM, Yoon HS, et al. Human embryonic stem cells express a unique set of microRNAs. Dev Biol. 2004; 270:488–498. [PubMed: 15183728]

Tashiro F, Kanai-Azuma M, Miyazaki S, Kato M, Tanaka T, Toyoda S, Yamato E, Kawakami H, Miyazaki T, Miyazaki J. Maternal-effect gene Ces5/Ooep/Moep19/Floped is essential for oocyte cytoplasmic lattice formation and embryonic development at the maternal-zygotic stage transition. Genes Cells. 2010; 15:813–828. [PubMed: 20590823]

Tsubooka N, Ichisaka T, Okita K, Takahashi K, Nakagawa M, Yamanaka S. Roles of Sall4 in the generation of pluripotent stem cells from blastocysts and fibroblasts. Genes Cells. 2009; 14:683–694. [PubMed: 19476507]

Wang Y, Baskerville S, Shenoy A, Babiarz JE, Baehner L, Blelloch R. Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation. Nat Genet. 2008; 40:1478–483. [PubMed: 18978791]

Wang Y, Medvid R, Melton C, Jaenisch R, Blelloch R. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat Genet. 2007; 39:380–85. [PubMed: 17259983]

Ying QL, Wray J, Nichols J, Batlle-Morera L, Doble B, Woodgett J, Cohen P, Smith A. The ground state of embryonic stem cell self-renewal. Nature. 2008; 453:519–523. [PubMed: 18497825]

Zhang J, Tam WL, Tong GQ, Wu Q, Chan HY, Soh BS, Lou Y, Yang J, Ma Y, et al. Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. Nat Cell Biol. 2006; 8:1114–123. [PubMed: 16980957]

Highlights

- miR-290 and miR-302 are sequentially activated from naïve to primed pluripotency.
- They are stochastically activated during OSK reprogramming, often bypassing miR-302.
- Addition of Sall4 results in sequential expression in reverse during reprogramming.
- The path taken during reprogramming is influenced by the cocktail of factors used.

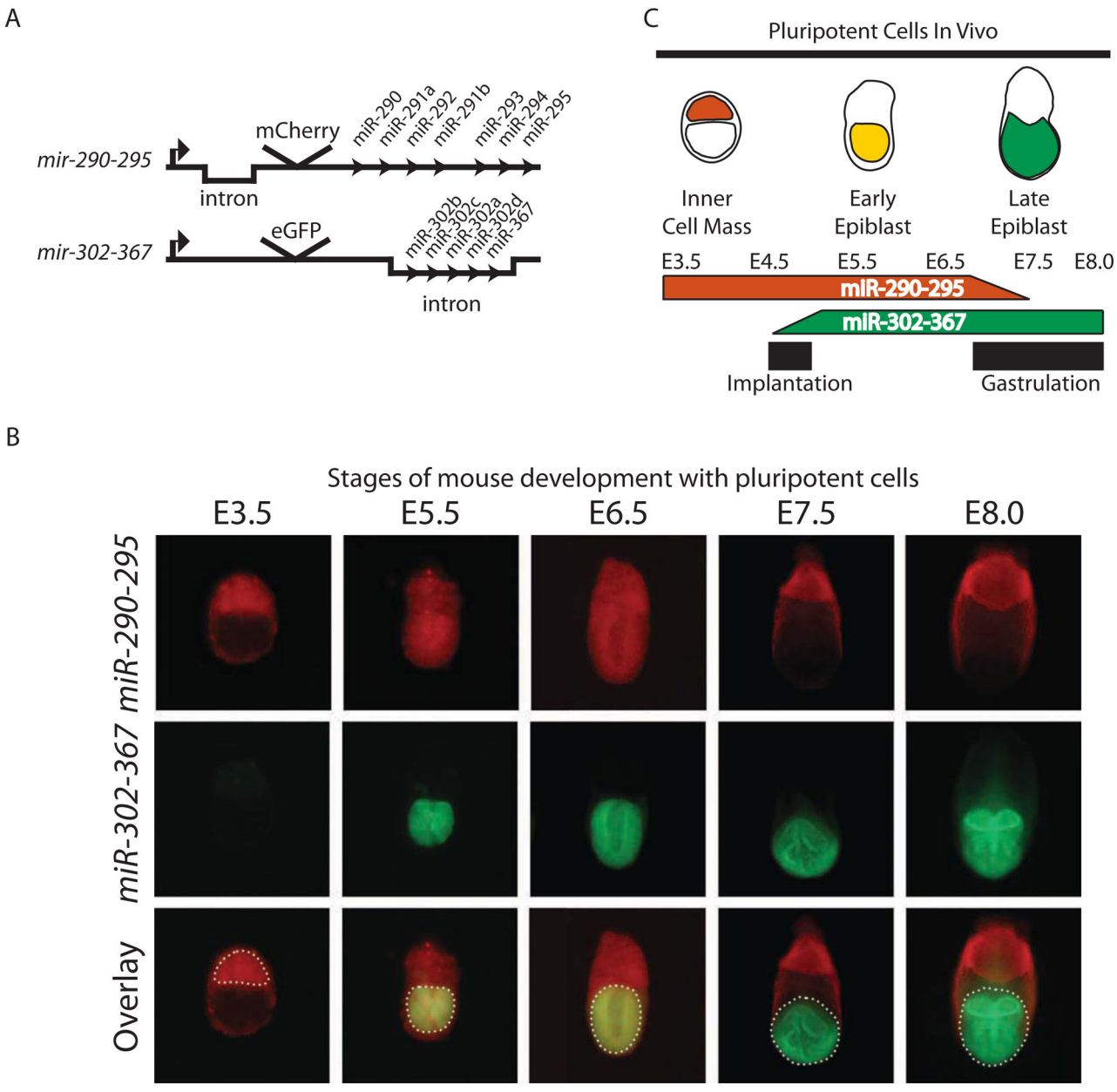

Figure 1. The miR-290 and miR-302 clusters define three transient and distinct states of pluripotency in early development

(a) Schematic of targeting strategy to introduce fluorescent reporters under the control of endogenous miRNA promoters. (b) Fluorescent images of embryos following timed matings showing the expression pattern of miR-290 and miR-302 reporters from embryonic days E3.5–E8.0 of mouse development. Embryological structures from which pluripotent stem cells are derived are outlined in white dashed lines. (E3.5, ICM, ESCs; E5.5–E6.5, early epiblast, EpiSCs; E7.5–8.5, late epiblast, EpiSCs). Fluorescent cells outside dashed lines represent extra-embryonic tissue. (c) Schematic summary of miR-290 and miR-302 expression during mouse embryonic development. See also Figure S1.

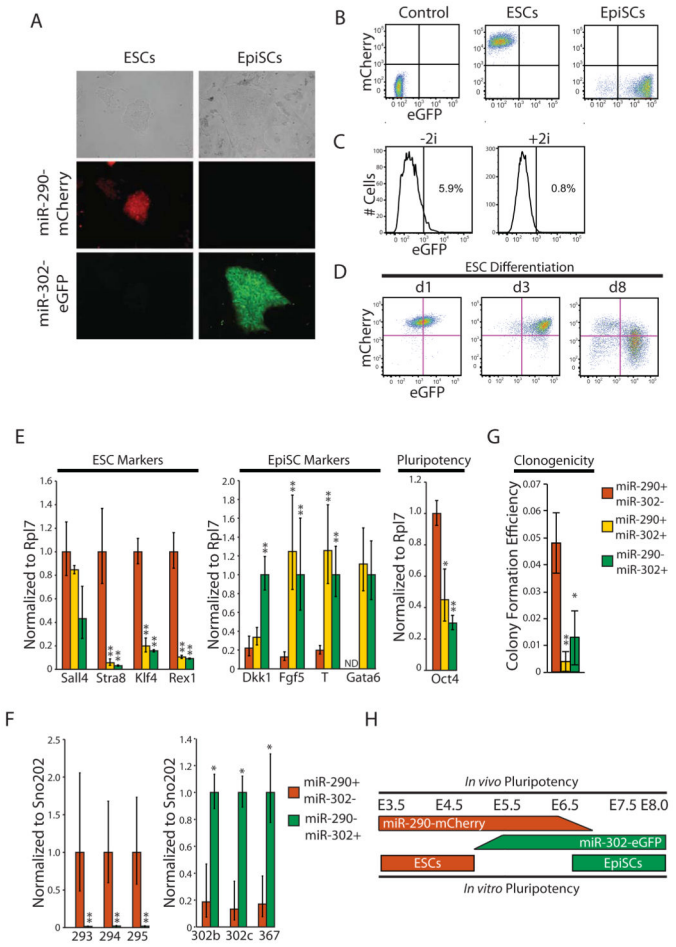

Figure 2. ESCs differentiate through miR-302+ intermediates, recapitulating transitions seen *in vivo*

(a) Representative images of ESCs grown in FBS+Lif+2i conditions (left column) and EpiSCs grown in N2B27 medium with FGF/Activin (right column). (b) FACS plots showing expression of mCherry or eGFP in dual reporter ESCs, EpiSCs, and untargeted V6.5 ESCs (Control). (c) Flow cytometry evaluating expression of eGFP in dual reporter ESCs grown with or without 2i in FBS+Lif medium. (d) Flow cytometry of mCherry and eGFP expression in ESCs cultured in FBS without Lif conditions for 1–8 days. (e) RT-qPCR of ESC and EpiSC markers in ES cells differentiated in FBS without Lif conditions for 5 days and sorted for cells expressing neither miR-290 nor miR-302 (black), miR-290 only (red), miR-302 only (green), or both loci (yellow) (n=3). (f) RT-qPCR analysis of miRNAs in ESCs that were differentiated in FBS without Lif conditions for 2 days or 6 days and sorted for miR-290-mCherry+/miR-302-eGFP− or miR-290-mCherry−/miR-302-eGFP+ cells, respectively (green and red bars, respectively) (n=3). (g) Colony forming efficiency of ESCs differentiated for 5 days and sorted as shown in (d). Colony number was assessed after 7 days of culture (n=3). h) Schematic summary comparing reporter expression *in vitro* and *in vivo*. Error bars represent SD. ND = not determined, *p < 0.05, **p<0.005. See also Figure S2 and Table S1.

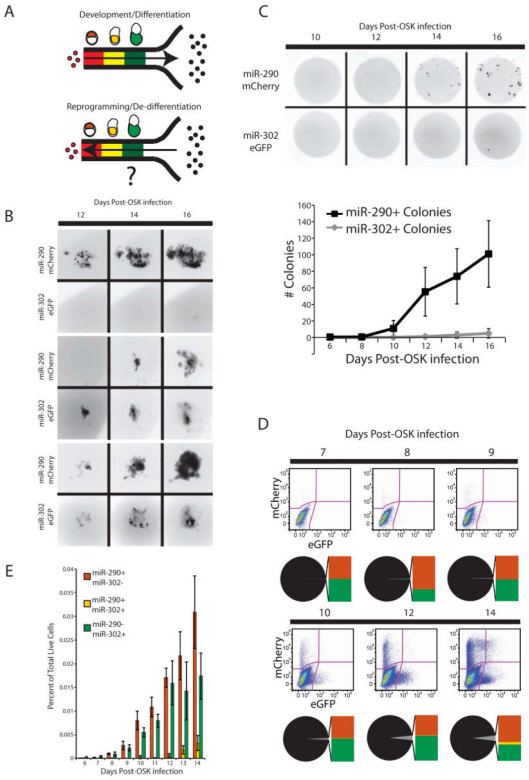

Figure 3. OSK reprogramming leads to late stage stochastic activation of miR-290 and miR-302

(a) Schematic overview of miR-290 and miR-302 expression during normal differentiation toward somatic tissues (top panel) and hypothetical reversal during reprogramming (bottom panel). (b) Representative images (10x magnification) of individual colonies followed through time using high resolution microscopy showing multiple patterns of loci activation in small patches of cells prior to the formation of naïve miR-290-mCherry+/miR-302-eGFP − colonies. (c) Representative images (2x magnification) of reprogramming wells comparing miR-290-mCherry and miR-302-eGFP expression in colonies. Quantification of colony number over 16 days of reprogramming. MiR-290-mCherry+ and miR-302-eGFP+ colonies were counted on days 6, 8, 10, 12, 14, and 16 following OSK infection. Values represent the average of seven different OSK viral preps and twelve technical replicates for each OSK prep. Error bars represent SD. (d) Representative FACS plots and corresponding bar of pie graphs showing the distribution of cells expressing miR-290-mCherry and/or miR-302-eGFP during reprogramming. Black pie slices represent cells that have not activated either miR-290 or miR-302 (black). The distribution of fluorescent cells (gray pie slices) is represented by the bar graphs. (e) Quantification by flow cytometry showing the percentage of total fluorescent cells expressing miR-290-mCherry and/or miR-302-eGFP for days 6–14 following OSK infection. Error bars represent SD of 3 biological replicates. See also Figure S3.

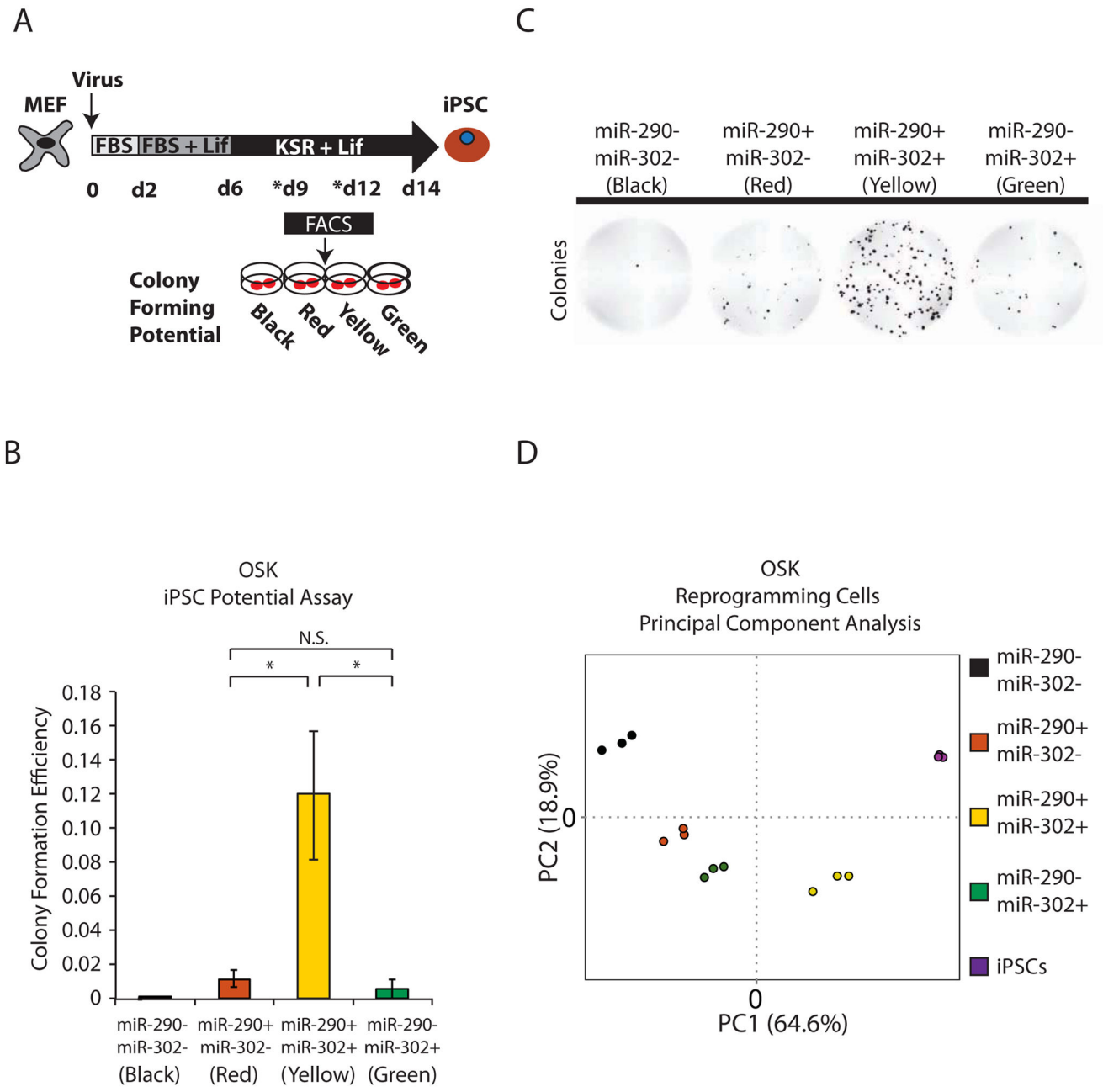

Figure 4. MiR-290-mCherry+/miR-302-eGFP+ double-positive cells show the greatest potential for generating iPSC colonies

(a) Schematic overview of the iPSC potential assay. Reprogramming wells were sorted on either day 9 (Sall4+OSK) or day 12 (OSK) of reprogramming and plated onto irradiated MEFs in reprogramming medium. Sorted intermediates were imaged daily and counted 5–6 days after FACS to determine the efficiency of colony formation. (b) Representative images comparing miR-290-mCherry+ colony formation of different fluorescent cells sorted at day 14 (1000 sorted cells/well) and imaged 5 days later. (c) Colony forming efficiency (number of colonies formed/number of cells plated) of different cellular populations sorted at day 12 of reprogramming. *$p < 0.01$, N.S. = not significant, as determined by Student's t-test.

Values represent the average of 3 OSK virus batches and 2 technical replicates for each prep. Error bars represent SD of the 3 OSK batches. (d) Principal component analysis (PCA) of Illumina bead array expression profiles from sorted day 12 OSK reprogramming cells expressing neither miR-290 nor miR-302 (black), miR-290 only (red), miR-302 only (green), or both loci (yellow), along with the resulting iPSCs (n=3). Percentages reflect the proportion of variance assigned to the principal components. See also Figure S4

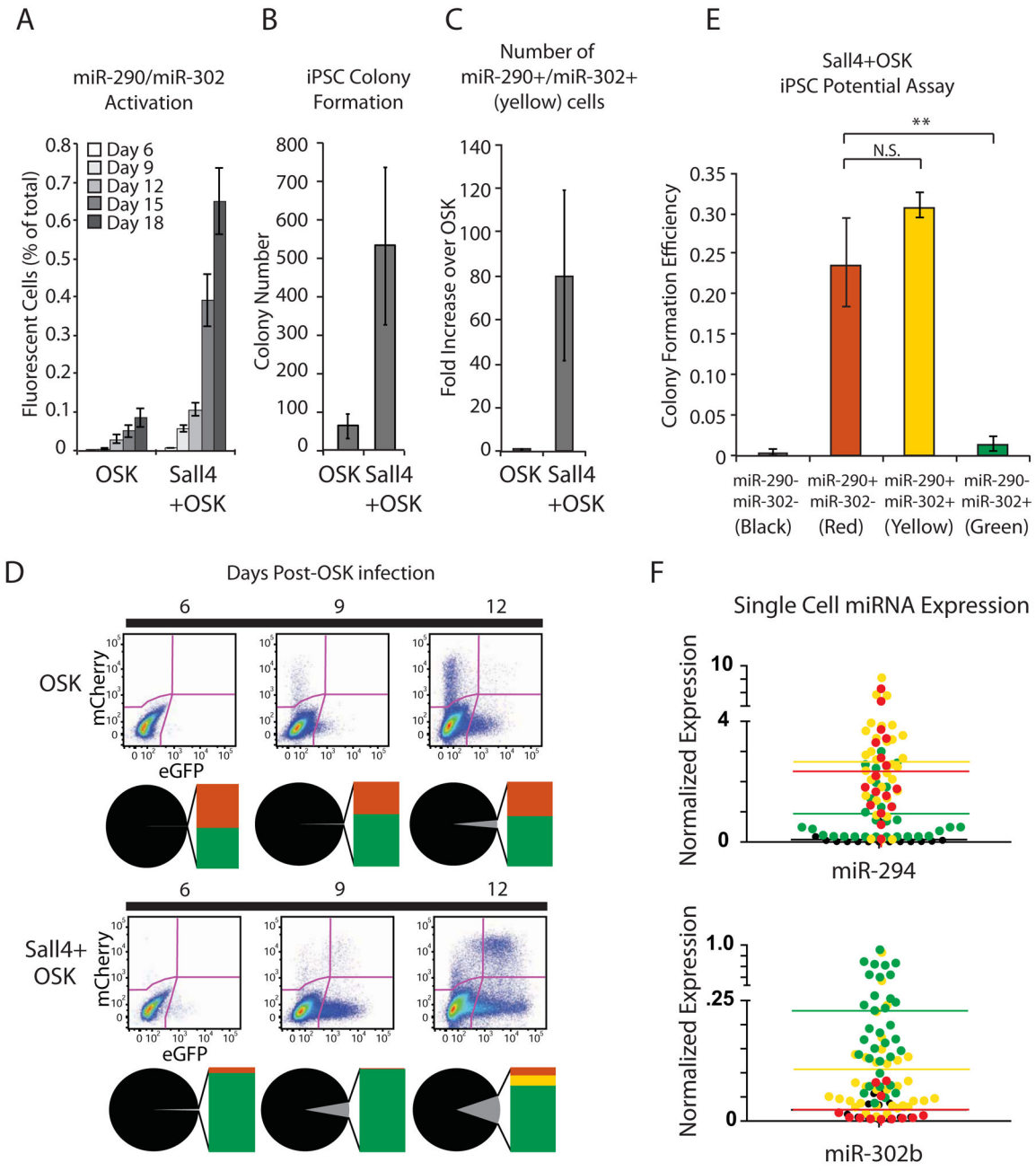

Figure 5. Sall4 enhances reprogramming with bias toward early miR-302 activation

(a) Quantification by flow cytometry of the fraction of total fluorescent cells (miR-290 and/or miR-302 positive) detected over multiple days during OSK or Sall4+OSK reprogramming. (b) Quantification of iPSC colony formation in OSK or Sall4+OSK reprogramming. Colonies were counted at day 15 after viral transduction. In (a) and (b), values represent the average of 3 different OSK viral preps and 2 technical replicates for each batch of OSK. Error bars represent SD of the 3 viral preps. (c) Quantification of the number of miR-290-mCherry+/miR-302-eGFP+ double-positive cells in OSK+Sall4-treated MEFs relative to OSK treated-MEFs at day 12 of reprogramming. (d) Bar of pie graphs

showing the distribution of miR-290 and miR-302 expressing cells from day 6–12 of reprogramming. Pie and bar slices represent the average of 3 OSK preps and 2 technical replicates for each prep. (e) Quantification of colony forming efficiency comparing different sorted populations during Sall4+OSK reprogramming at the first time point when cells could be detected for all populations (day 9). Values represent the average of 3 OSK virus batches and 2 technical replicates per OSK batch. Error bars represent the SD of the 3 OSK batches. **$p < 0.005$, N.S. = not significant, as determined by Student's t-test. (f) Expression of representative miR-290 and miR-302 family miRNAs normalized to sno202 in individual Sall4+OSK reprogramming intermediates as determined by single cell RT-qPCR. Cells were single cell sorted at day 12 of Sall4+OSK reprogramming and RT-qPCR reactions were run in technical triplicates. Each dot represents the mean expression value of that miRNA in a single cell. Horizontal bars indicate the mean of the entire cell population. Only samples with detectable sno202 levels were considered. Samples with miR-294 or miR-302b Ct values too high to be determined (Not Detected) were given a normalized expression value of 0 in the calculation of the population mean. See also Figure S5.

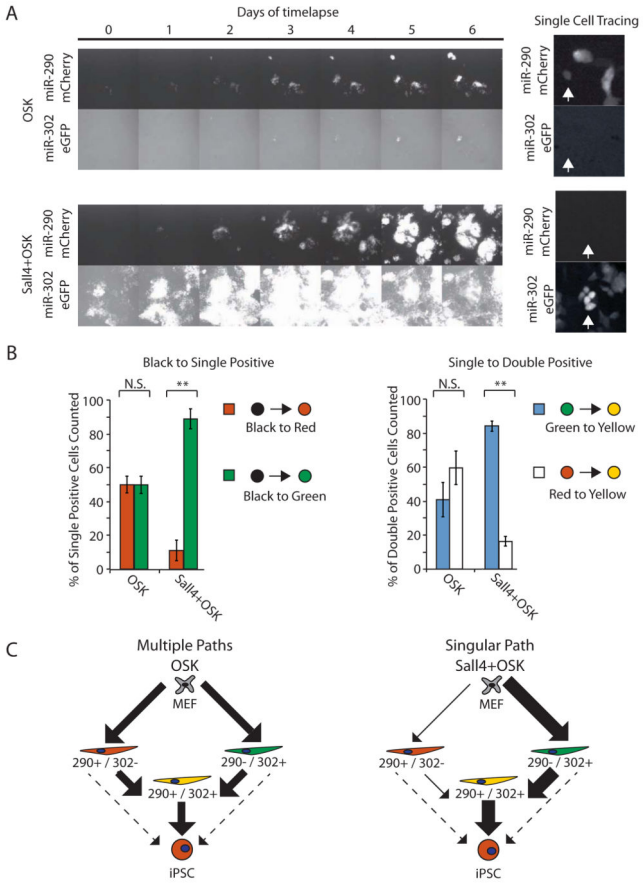

Figure 6. Sall4 biases the path taken during reprogramming

(a) Representative time-lapse images over 6 days during OSK or Sall4+OSK reprogramming showing the evolution of miRNA locus expression beginning with a single cell (10x magnification). Right, representative images demonstrating how individual cells were scored. Arrows point to matching cells in either OSK or Sall4+OSK conditions. (b) Quantification of transitions in reporter expression observed with time-lapse imaging every 2–4 hours. Activation of miR-290-mCherry or miR-302-eGFP was quantified in single cells that went on to form colonies. The number of cells that changed reporter expression before forming a colony is presented as a percentage of the total number of single positive or double positive cells traced. In total, 303 single positive cells were followed in OSK (150 black to red and 153 black to green cells) and 467 single positive cells in Sall4+OSK (45 black to red and 422 black to green cells). 114 double positive cells were traced in OSK (46 green to yellow and 68 red to yellow cells) and 236 double positive cells in Sall4+OSK (198 green to yellow and 38 red to yellow cells). Black to double positive (black to yellow) transitions were not observed. Values represent the average of 3 different wells. Error bars represent the SD of cells traced over 3 wells. **$p < 0.005$, N.S. = not significant, as determined by Student's t-test. (c) Schematic summary showing potential paths taken by reprogramming cells relative to miR-290 and miR-302 reporters. Bold arrows represent the preferred path taken during OSK or Sall4+OSK reprogramming. Dashed arrows represent possible paths that were not quantified in our experiments. See Supplemental Movies 1–6.

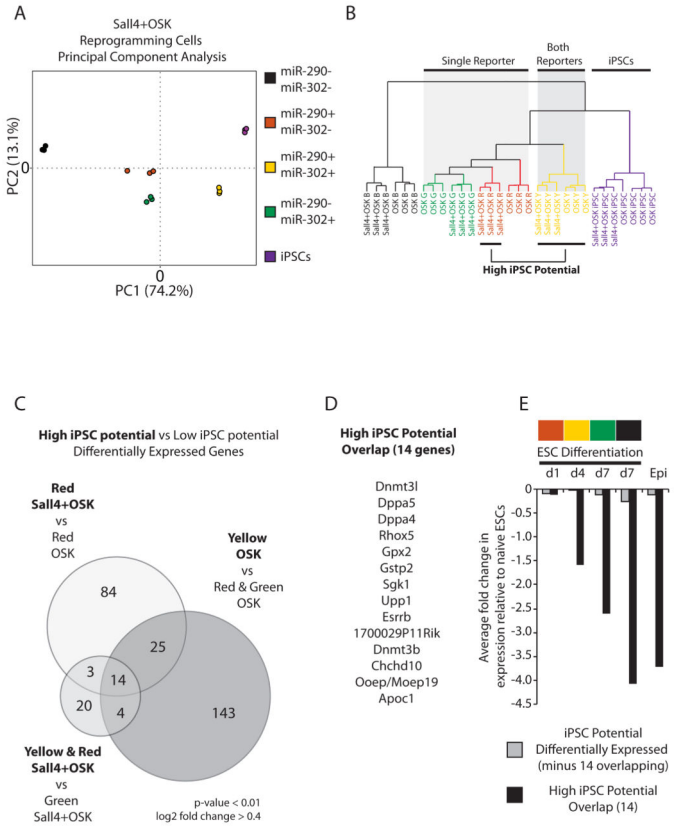

Figure 7. Identification of genes enriched within subsets of cells with greatest iPSC potential

(a) Principal component analysis (PCA) of Illumina bead array expression profiles from sorted day 9 Sall4+OSK reprogramming cells expressing neither miR-290 nor miR-302 (black), miR-290 only (red), miR-302 only (green), or both loci (yellow), along with the resulting iPSCs (n=3). Percentages reflect the proportion of variance assigned to the principal components. (b) Hierarchical clustering of Illumina bead array expression profiles from sorted day 12 OSK and day 9 Sall4+OSK reprogramming cells expressing neither miR-290 nor miR-302 (B), miR-290 (R), miR-302 (G) or both loci (Y) along with the resulting iPSCs (n=3). (c) Venn diagram showing the overlap of genes that are differentially expressed between populations with high versus low iPSC potential. $p < 0.01$ and log2 fold change > 0.4 (d) List of 14 overlapping genes correlated with high iPSC potential. (e) The average expression of all differentially expressed iPSC potential genes versus the 14 overlapping genes during ESC differentiation and in primed EpiSC cultures.