Retinal Cell Type DNA Methylation and Histone Modifications Predict Reprogramming Efficiency and Retinogenesis in 3D Organoid Cultures

Authors
Lu Wang, Daniel Hiler, Beisi Xu, ..., Dianna Johnson, Sharon Frase, Michael A. Dyer

Correspondence
michael.dyer@stjude.org

In Brief
Wang et al. reprogram retinal cell types into iPSCs and test their ability to make retinal organoids. They discover an inverse correlation between reprogramming efficiency and retinal differentiation linked to DNA/chromatin modifications and nuclear organization. These data identify molecular markers of iPSC lines that are efficient at producing retina.

Data and Software Availability
GSE87064

Highlights
- Retinal reprogramming efficiency is cell type and developmental stage specific
- Reprogramming efficiency is inversely correlated with retinal differentiation
- DNA/chromatin modifications and nuclear organization affect reprogramming efficiency
- Meis1 expression is a predictive biomarker of retinal differentiation of iPSCs
Retinal Cell Type DNA Methylation and Histone Modifications Predict Reprogramming Efficiency and Retinogenesis in 3D Organoid Cultures

Lu Wang,1,8 Daniel Hiler,1,8 Beisi Xu,1,8 Issam AlDiri,1 Xiang Chen,2 Xin Zhou,2 Lyra Griffiths,1 Marc Valentine,4 Abbas ShiriniFard,1 András Sablauer,3 Suresh Thigagarajan,1 Marie-Elizabeth Barabas,1 Jiakun Zhang,1 Dianna Johnson,6 Sharon Frase,5 and Michael A. Dyer1,6,7,9,*

1Department of Developmental Neurobiology
2Department of Computational Biology
3Department of Diagnostic Imaging
4Cytogenetics Shared Resource
5Cell and Tissue Imaging Shared Resource
St. Jude Children’s Research Hospital, Memphis, TN 38105, USA
6Department of Ophthalmology, University of Tennessee Health Science Center, Memphis, TN 38163, USA
7Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA
8These authors contributed equally
9Lead Contact
*Correspondence: michael.dyer@stjude.org
https://doi.org/10.1016/j.celrep.2018.01.075

SUMMARY

Diverse cell types can be reprogrammed into pluripotent stem cells by ectopic expression of Oct4 (Pou5f1), Klf4, Sox3, and Myc. Many of these induced pluripotent stem cells (iPSCs) retain memory, in terms of DNA methylation and histone modifications (epigenetic memory), of their cellular origins, and this may bias subsequent differentiation. Neurons are difficult to reprogram, and there has not been a systematic side-by-side characterization of reprogramming efficiency or epigenetic memory across different neuronal subtypes. Here, we compare reprogramming efficiency of five different retinal cell types at two different stages of development. Retinal differentiation from each iPSC line was measured using a quantitative standardized scoring system called STEM-RET and compared to the epigenetic memory. Neurons with the lowest reprogramming efficiency produced iPSC lines with the best retinal differentiation and were more likely to retain epigenetic memory of their cellular origins. In addition, we identified biomarkers of iPSCs that are predictive of retinal differentiation.

INTRODUCTION

Somatic cells can be reprogrammed to multipotent stem cells by ectopic expression of defined factors (Oct4, Klf4, Sox2, and Myc), which holds great promise for patient-specific disease modeling and regenerative medicine (Chen et al., 2015; Dyer, 2016; Singh et al., 2015). In addition to the fibroblasts that were used in the first successful reprogramming experiments (Takahashi et al., 2007; Yu et al., 2007), a myriad of other cell types have been reprogrammed into induced pluripotent stem cells (iPSCs) (Aoi et al., 2008; Lowry et al., 2008; Park et al., 2008). Reprogramming efficiency is cell type specific and is thought to be stochastic for any homogeneous population of cells (Hanna et al., 2009). Bone-marrow-derived hematopoietic stem cells have some of the highest rates of reprogramming (28%) (Eminli et al., 2009), and mature differentiated neurons are among the most difficult to reprogram (Hiler et al., 2015, 2016; Kim et al., 2011). Indeed, early attempts to reprogram murine cortical neurons failed to produce iPSCs, unless the p53 gene was inactivated (Kim et al., 2011). More recently, an alternative approach was developed to reprogram neurons with wild-type p53 (Hiler et al., 2015, 2016).

iPSCs derived from diverse cell types have been shown to harbor epigenetic memory of their cellular origins that makes them more or less likely to differentiate along particular lineages (Bar-Nur et al., 2011; Kim et al., 2010; Polo et al., 2010). In some iPSC lines, this epigenetic memory is gradually depleted with passage in culture, but in other examples, it is stably maintained (Kim et al., 2010, 2011; Nishino et al., 2011; Polo et al., 2010). The majority of studies on epigenetic memory in iPSCs have focused on DNA methylation, but recent evidence suggests that it may also extend to other epigenomic factors such as histone modifications at promoters and gene bodies and higher order chromatin organization with topologically associated domains (TADs) mediated by CTCF (Beagan et al., 2016; Krijger et al., 2016). It is not known how reprogramming efficiency relates to epigenetic memory, nor is it known how the dynamic changes in the epigenome, which occur as cells differentiate, relate to epigenetic memory and cellular reprogramming.

In this study, we compare the reprogramming efficiency of 5 cell types in the retina at two stages of development and relate that to the ability of these retinal-derived iPSCs (r-iPSCs) to subsequently differentiate into retina. The cells that were most difficult to reprogram made the best retina, as determined by STEM-RET scoring (Hiler et al., 2015, 2016), and this was
reflecting their epigenetic memory. Moreover, characterization of a series of lines that failed to produce retina from diverse sources identified epigenetic features of several genes, including Meis1 target genes that are predictive of retinogenesis for stem cells. This work will have implications for the selection of cell populations for cell-based therapy and for using reprogramming of purified cell populations to advance our understanding of the role of the epigenome in normal differentiation.

RESULTS

Cell Type Specification and Developmental Stage Influence Reprogramming Efficiency in the Retina

We have previously demonstrated the feasibility of reprogramming rod photoreceptors using the Nr1-GFP;Col1a1-OKSM;Rosa26-M2rtTA mouse strain (Hiler et al., 2015, 2016) (Figure 1A). To extend our previous studies and compare the reprogramming efficiency across retinal cell types, we generated 4 additional mouse lines by crossing GFP transgenic mice with the Col1a1-OKSM;Rosa26-M2rtTA strain (Stadtfeld et al., 2010). The Chnb4-GFP transgene labels cone photoreceptors (Figure 1B) (Siegrist et al., 2009), Rlbp1-GFP labels Müller glia (Figure 1C) (Vázquez-Chona et al., 2009), Gad1-GFP labels horizontal and a subset of amacrine interneurons (Figure 1D) (Huckfeldt et al., 2009), and Grm6-GFP labels ON bipolar cells (Figure 1E) (Dhingra et al., 2008). Next, we used a serial limiting dilution procedure in retinal pellets to quantitate reprogramming efficiency (Hiler et al., 2015, 2016). Briefly, GFP+ cells from the different reprogrammable mouse strains were mixed with 1.5 x 10^6 C57BL/6 post-natal day (P)0–P5 retinal cells in retinal pellets and cultured on polycarbonate filters in retinal explant culture medium (Hiler et al., 2015, 2016). This procedure is essential, because purified mature retinal neurons do not survive when cultured on plastic or on irradiated mouse embryonic fibroblasts (MEFs). The reprogrammable cells were mixed with the normal retinal cells in these retinal pellets in limiting dilutions (20,000, 2,000, or 200 cells per pellet) (Supplemental Information).

Doxycycline (2 μg/ml) was added to the pellet cultures to induce expression of Oct3/4 (Pou5f1), Kif4, Sox2, and Myc (OKSM) in the reprogrammable neurons for 10 days. Then, individual pellets were dissociated and plated in limiting dilution on irradiated MEFs in the presence of leukemia inhibitor factor (LIF) to support the growth of iPSC colonies (Supplemental Information). We used mature (P21) and immature (P2–P5) cells for each cell type, and 2–3 independent reprogramming experiments were performed for each cell type at each stage of development. We also generated fibroblast-derived iPSCs from the same mouse lines as those used for retinal reprogramming for comparison to the retinal derived iPSCs.

We calculated the reprogramming efficiency (see Supplemental information) for MEFs and each of the 5 cell types at each developmental stage and calculated the mean and SD across multiple independent experiments (Figures 1F–1J). Mature rod photoreceptors had a significantly lower reprogramming efficiency (p < 0.01) relative to cones, Müller glia, amacrine/horizontal interneurons, and immature rod photoreceptors (Figures 1F–1J). The most efficient reprogramming was achieved with Müller glia and interneurons (amacrine/horizontal neurons), and the lowest reprogramming efficiency was achieved with bipolar neurons (Figures 1F–1J).

We isolated and grew lines to passage 20 prior to characterization. Using molecular and cellular markers, we identified lines that underwent complete reprogramming and maintained a stable karyotype as described previously (Figures 1K–1O; Supplemental Information) (Hiler et al., 2015, 2016). We selected 59 lines for subsequent analyses, including the 8 rod-derived iPSC lines that were previously described (Table S1) (Hiler et al., 2015).

Nuclear Organization and Localization of Pluripotency Genes in Retinal Cell Types

One of the key steps to cellular reprogramming is reversing the epigenetic marks that accumulate during differentiation and activation of the endogenous pluripotency genes (Apostolou and Hochmedinger, 2013). This is particularly important in the retina because of the cell-type-specific 3-dimensional (3D) organization of the chromat in the nucleus (Chen et al., 2015; Dyer, 2016; Solovei et al., 2009) (Figures 2A and 2B). To determine whether there was an inverse correlation between the proportion of heterochromatin in the nuclei and reprogramming efficiency, we performed serial block face imaging on the FEI Teneo electron microscope to produce 3D datasets on each of the 5 classes of retinal cell types used in this study in triplicate (rods, cones, bipolar cells, Müller glia, and amacrine cells) (Figure 2B). Next, we

Figure 1. Retinal Cell-Type-Specific Differences in Reprogramming

(A–E) Micrographs of GFP expression alone (arrows in upper panels) and overlaid with DAPI (blue) in the lower panels. (A) Nr1-GFP rod; (B) Chnb4-GFP cone; (C) Rlbp1-GFP Müller glia; (D) Gad1-GFP amacrine/horizontal cells; (E) Grm6-GFP bipolar cells.

(F–J) Bar charts of reprogramming efficiency for immature (gray) and mature retinal cell types. Each bar represents the mean and SD of biological replicate experiments as indicated. (F) Rod; (G) cone; (H) Müller glia; (I) amacrine/horizontal cells; (J) bipolar cells. The dashed lines indicate the reprogramming efficiency of mouse embryonic fibroblasts (MEFs). Each of the retinal-derived cell types was significantly different from MEFs (p < 0.001), and there were significant differences between immature and mature rods (p < 0.0001) and bipolar neurons (p < 0.0001). The reprogramming efficiency is a growth-adjusted percentage of input (Supplemental Information), with number of experiments indicated on the bars.

(K) Micrograph of iPSC colonies after 20 passages on irradiated MEF feeder cells.

(L) Micrograph of alkaline-phosphatase-stained iPSC colonies (red).

(M) Micrograph of individual iPSC colonies immunostained for SSEA1, Oct4, and Nanog (red) with DAPI nuclear counterstain (blue).

(N) Bar chart of qPCR for 7 pluripotency genes plotted relative to the positive control cell line (EB5:Rx-GFP). Data are normalized to Gapdh and are plotted relative to EB5:Rx-GFP (blue dashed line). Each bar represents the mean and SD of technical replicate experiments for the iPSC clone shown in (K).

(O) Micrographs of immunohistochemical staining (upper panels) and H&E staining (lower panels) of representative teratomas made from r-iPSCs. ONL, outer nuclear layer; INL, inner nuclear layer; GCL, ganglion cell layer. Scale bars, 25 μm in (A)–(E) and (M) and 100 μm in (K) and (L); and 200 μm in (O). See also Table S1 and Supplemental Information.
developed an algorithm to segment the heterochromatin and euchromatin based on image features in the 3D electron microscopy (EM) datasets (Supplemental Information). Rod photoreceptors have the smallest nuclei with the largest amount of heterochromatin, while the bipolar neurons have the largest nuclei with the least heterochromatin in our analysis (Figures 2B and 2C).

Next, to determine where in the nucleus the reprogramming genes were located, we performed fluorescence in situ hybridization (FISH) of Oct4, Klf4, Sox2, and Myc on adult retina and performed 3D reconstruction of individual nuclei using confocal microscopy. To identify the euchromatin in each nucleus, we performed 3D reconstruction of individual nuclei using confocal microscopy. To identify the euchromatin in each nucleus, we performed immunostaining for H3K4me3 prior to hybridization of the FISH probes (Figures 2D and 2E). The facultative heterochromatin (f-het) and constitutive heterochromatin (c-het) can be distinguished by DAPI staining intensity. In mature rods, Oct4 was localized in the euchromatin, and Myc was in the f-heterochromatin domain (Figures 2D and 2E). Sox2 and Klf4 were also in the f-heterochromatin domain in rods, but they were at the boundary with euchromatin (Figure 2F). In the other cell types, all 4 genes were localized to the euchromatin domain (Figures 2G and 2H).

To relate their 3D position within the nucleus to their 1D epigenetic state, we analyzed chromatin immunoprecipitation sequencing (ChIP-seq) for 8 histone marks, CTCF, Brd4 and RNA polymerase II (Pol II), DNA methylation, assay for transposable-accessible chromatin sequencing (ATAC-seq), and RNA sequencing (RNA-seq) data for adult retina (data available at https://pecan.stjude.org/proteinpaint/study/retina2017; Aldiri et al., 2017). To integrate the individual ChIP-seq tracks, we performed chromatin hidden Markov modeling (chromHMM) using 11 chromHMM states, as done previously for murine retina (Figure 2I) (Aldiri et al., 2017).

In the adult retina, Oct4, Klf4, and Myc are not expressed (fragments per kilobase per million reads [FPKM] < 1.0), and Sox2 is expressed at low levels (FPKM < 10) (Figures S1A–S1D). Sox2 is not expressed in purified rod photoreceptors (FPKM < 1.0), so we performed co-immunolocalization with molecular markers of the other cell types and found that Sox2 is expressed in Müller glia and a subset of amacrine cells (Figures S1E–S1H).

All 4 genes (OKSM) have H3K27me3 at their promoters and are hypomethylated in those regions (Figures 2J–2M). The promoter region for Klf4 is the only one of the 4 genes with the H3K27Ac mark (Figure 2J). All 4 genes have H3K9me3 overlapping with the H3K27me3, and the ATAC-seq data are consistent with their epigenetic repression during differentiation (data available at https://pecan.stjude.org/proteinpaint/study/retina2017; Aldiri et al., 2017).

We have previously performed ChIP-seq analysis of purified rod photoreceptors (Aldiri et al., 2017); here, we extended those data to purified ON bipolar neurons from the Grm6-GFP transgenic mice (Dhingra et al., 2008) to gain insight into the epigenetic barrier to reprogramming. We performed ChIP-seq using previously validated antibodies for H3K27me3, H3K27Ac, H3K4me2, and H3K4me1. We compared the ChIP-seq for the 4 reprogramming genes between rods, bipolar neurons, and MEFs (Figures S2A–S2C) and found that the MEFs had more active histone marks (H3K27Ac, H3K4me2, and H3K4me1) and lacked the repressive histone mark (H3K27me3). The rods and bipolar neurons had much more H3K27me3 than the MEFs and fewer active marks (Figures S2A–S2C).

**Rod- and Bipolar-Derived iPSCs Have More Efficient Retinal Differentiation**

To quantitate retinogenesis from each of our validated iPSC lines, we performed STEM-RET differentiation and scoring for 59 iPSC lines, as described previously (Figure 3A; Table S2) (Hiler et al., 2015, 2016). As a positive control, we used the EB5:Rx-GFP murine embryonic stem cell (ESC) line (Eiraku et al., 2011), and we also included fibroblast-derived iPSCs for our analysis (Table S2).

For eye field specification, we scored eye field induction efficiency (EFE), eye field induction specificity (EFS), and eye field proliferation (EFP) at day 7 (Figures 3A–3C; Table S2; Supplemental Information). For optic cup formation, we scored optic cup efficiency (OCE) and optic cup frequency (OCF) at day 10 (Figures 3A–3C; Table S2; Supplemental Information). To measure retinal differentiation (days 10–28), we analyzed 16 retinal differentiation genes by qPCR by using TaqMan probes and normalized them to Gapdh (retinal differentiation qPCR [RDQ]).

---

**Figure 2. Differences in Nuclear Organization across Retinal Cell Types**

(A) Electron micrograph of an adult mouse retina.

(B) Magnified regions from (A) in the upper panel and tracing of heterochromatin and euchromatin from those nuclei (blue). The mean and SD of the volume of heterochromatin and euchromatin was calculated from 3 individual nuclei for each cell type.

(C) Stack bar chart of nuclear volume distribution by cell type.

(D and E) Representative fluorescent in situ hybridization for (D) Oct4 and (E) Myc in mature rod nuclei (red) with the DAPI nuclear counterstain (blue) and immunofluorescence for H3K4me3 (green).

(F) Stack bar chart of the localization of each of the 4 reprogramming factors to rod nuclei, with a minimum of 50 nuclei scored per probe.

(G) Representative fluorescent in situ hybridization for Oct4 and Myc in mature inner nuclear layer (INL) nuclei (red) with the DAPI nuclear counterstain (blue) and immunofluorescence for H3K4me3 (green). The lower panel is the segmentation of each region using the machine learning algorithm developed for these studies.

(H) Stack bar chart of the localization of each of the 4 reprogramming factors to INL nuclei with a minimum of 50 nuclei scored per probe.

(I) Heatmap of the chromatin hidden Markov modeling (chromHMM) states. The chromHMM states are color coded below with the association to specific regions in the genome.

(J–M) Trace of DNA methylation, chromHMM, and ChIP-seq for H3K9me3, H3K27Ac, and H3K27me3 for each of the reprogramming genes in adult retina. The scale for the ChIP-seq is indicated for each mark. (J) Klf4; (K) Myc; (L) Oct4 (Pou5f1); and (M) Sox2.

olm, outer limiting membrane; ONL, outer nuclear layer; INL, inner nuclear layer; GCL, ganglion cell layer; c-het, constitutive heterochromatin; f-het, facultative heterochromatin; euc, euchromatin. Scale bars, 10 μm in (A) and 1 μm in (B), (D), and (E).

See also Figures S1 and S2.
Figure 3. R-iPSCs Have an Inverse Correlation between Reprogramming Efficiency and Retinal Differentiation

(A) Outline of the STEM-RET protocol.

(B) Representative micrographs from several stages of STEM-RET differentiation.

(C) Bar chart of each scoring metric for a representative interneuron-derived iPSC line that has relatively poor performance in retinal differentiation. Each bar represents the mean and SD of at least 2 biological replicate experiments. The integrated scores for eye field, optic cup, and retinal differentiation are shown in the boxes. The gray bars represent the scores for the positive control EB5:Rx-GFP line. The proportion of organoid that was retina based on triplicate experiments is shown.

(D) Micrograph of immunofluorescent staining of recoverin (red) and DAPI nuclear counterstain (blue).

(E) qPCR for 15 retinal differentiation genes relative to E14.5 retina. Each bar represents the mean and SD of 2 independent experiments, and normal P12 retina is shown in red.

(legend continued on next page)
were very similar, because they are all fully reprogrammed multi-
As expected, the RNA-seq profiles for each of the iPSC lines
Epigenetic Memory of r-iPSCs
See also Tables S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, and S11.
across individual lines (Figures 3C–3H; Table S2). We also quan-
titated the proportion of retinal area in the organoids for retina
produced from each retinal cell type and discovered that the
retinal area was higher for lines that had higher STEM-RET
scores (Figures 3C and 3F; Table S8; Supplemental Information).
The best STEM-RET scores were achieved with the iPSC lines
that were derived from the retinal cell types that were the most
difficult to reprogram (rods and bipolar cells) relative to those
that had more efficient reprogramming, such as amacrine/ horizontal interneurons (Figures 3I–3K; Table S2).

Epigenetic Memory of r-iPSCs
As expected, the RNA-seq profiles for each of the iPSC lines
were very similar, because they are all fully reprogrammed multi-
potent stem cells (Table S3). The lines did not cluster separately in principal-component analysis (PCA) of their gene expression profiles based on their cell of origin (Figures 4A and 4B) or their ability to make retina (Figure 4C). We identified the genes with significant differential gene expression (Table S3) across the 3 types of comparisons (retinal cell of origin, fibroblasts versus retinal origin, and ability to differentiate into retina) and performed pathway analysis. Most of the comparisons did not produce significant pathway enrichment (q < 0.05) (Table S4). However, the genes that were upregulated in rod-derived iPSCs relative to the other iPSC lines were enriched in chromatin binding, chromatin modification, and transcriptional regulation (Table S4).

To determine whether r-iPSCs retain an epigenetic memory of their cellular origins that is not necessarily reflected in their transcrip-
tomes, we selected 33 lines for detailed epigenetic profiling (Table S1). We included iPSCs from each of the 5 retinal cell types, fibroblast-derived iPSCs, and the positive control EB5:Fx-GFP line. We also profiled 11 iPSC lines from diverse sources (retinal and fibroblast derived) that failed to make retina.

We performed integrated analysis, including RNA-seq, DNA methylation analysis, and ChIP-seq of 8 histone marks (active chromatin: H3K4me1, H3K4me2, H3K4me3, H3K27Ac, H3K9/14Ac; repressed chromatin, H3K27me3 and H3K9me3; gene bodies: H3K36me3), as well as for CTCF, Brd4, and Pol II (Supplemental Information). All ChIP-seq was performed on biological replicates, and all antibodies were independently validated for ChIP (protocols are available online at http://stjude.org/CSTN) (Aldiri et al., 2017). In total, 396 ChIP-seq libraries were sequenced for the 33 stem cell lines, and we compared their epigenetic profiles to normal developing mouse retina spanning 8 stages of development using the same approach (Aldiri et al., 2017). All data from ChIP-seq, DNA methylation, and RNA-seq are freely available in a searchable browser (https://pecan.stjude.org/proteinpaint/study/retinalIPSC2017).

Previous analysis of changes in DNA methylation with retinal differentiation identified upregulated and downregulated genes that had correlated and anticorrelated changes in DNA methyl-
ation in whole retina and in individual retinal cell types (Aldiri et al., 2017). Most of those genes were reset to the ESC state across iPSC lines from diverse sources (Table S5). For example, retinal progenitor genes such as Pax6, Lhx2, Six3, Six6, Nes, Meis1, and Vax2 had DNA methylation profiles that were reset to the ESC state in more than 95% of the iPSC lines (Figure 4D; Table S5). Similarly, many of the cell-type-specific retinal differ-
tentiation genes were efficiently reset to the ESC state in iPSCs. The rod-specific genes Gnat1, Tulp1, Sstx3, Prph, and Rpgr were reset in over 95% of the iPSC lines (Table S5). However, there were some notable exceptions. Nr2f was reset in 84% of the lines; Prom1 and Aip1, in 73% of the lines; Rpl1, in 67% of the lines; Pde6a, in 50% of the lines; Rho and Pde6b, in 30% of the lines; and Rcvr1, Guca1b, and Hk1, in less than 16% of the lines (Figure 4E; Table S5).

Previous studies on the temporal changes in the epigenetic landscape during reprogramming have suggested that the first changes with reprogramming are the silencing of the cell-type-specific transcription program, followed by poised enhancers and heterochromatic regions enriched for the H3K9me3 repressive histone mark (Apostolou and Hochedlinger, 2013; Chronis et al., 2017). Using the previously published data on the dynamic epigenetic landscape during retinal development (Aldiri et al., 2017), we analyzed the efficiency of resetting of retinal cell-type-specific promoters and genes in our iPSC lines to the pluripotent state found in ESCs (Tables S6 and S7). Specifically, we performed Bayesian analysis of the chromHMM for each gene and promoter in each iPSC line to determine whether its

**(F)** STEM-RET scores for a bipolar-derived iPSC line.

**(G)** Micrograph of recoverin staining for the m0401 line shown in red, with DAPI nuclear counterstain (blue).

**(H)** qPCR for 15 retinal differentiation genes relative to E14.5 retina, as described above for (E).

**(I)** Combined eye field score, optic cup score, and retinal differentiation score for each cell type. Each bar represents the mean and SD of at least 4 lines per cell type in biological duplicates.

**(J)** Retinal area in organoids from each retinal-derived iPSC line. The mean is plotted from 3 independent experiments, and the SD is indicated in the supplemental tables Tables S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, and S11.

**(K)** Plot of integrated retinal score including retinal area versus reprogramming efficiency. The dashed line separates the groups of iPSC lines that have signific-
antly different reprogramming efficiency and retinal differentiation. The STEM-RET scores include retinal area calculations.

EFE, eye field efficiency; EFS, eye field specificity; EFP, eye field proliferation; OCE, optic cup efficiency; OCF, optic cup frequency; RDQ, retinal differentiation qPCR; RDEM, retinal differentiation electron microscopy; RDIF, retinal differentiation immunofluorescence. Scale bars, 200 μm in (B) and 25 μm in (D) and (G). See also Tables S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, and S11.
Figure 4. Epigenetic Memory in R-iPSCs

(A–C) Principal-component analysis of gene expression from RNA-seq for iPSCs derived from the different retinal cell types (A), comparing r-iPSCs to fibroblast-derived iPSCs (B), or comparing lines that made retina to those that did not (C).

(D) Representative examples of a progenitor gene that resets its DNA methylation profile to the EB5:Rx-GFP ESC state (Pax6).

(E) Representative example of a differentiation gene that resets the DNA methylation profile to the EB5:Rx-GFP ESC state (Aipl1) and one that did not (Guca1b). The arrow represents the region of differential DNA methylation.

(F) Heatmap of the chromatin hidden Markov modeling (chromHMM) states. The chromHMM states are color coded below with the association to specific regions in the genome.

(G–I) ChIP-seq tracks and chromHMM state for the Aipl1 gene for the p21 retina (G), the EB5:Rx-GFP ESC line (H), and a representative retinal-derived iPSC line (I). The Aipl1 gene was properly reset to the ESC state in the iPSC line.

(J–L) ChIP-seq tracks and chromHMM state for the Stx3 gene in the p21 retina (J), the EB5:Rx-GFP ESC line (K), and a representative retinal-derived iPSC line (L). The retinal-derived iPSC line did not properly reset the H3K27me3 at the promoter of Stx3.

(M–O) Bar plot of the difference in percentage of iPSC lines that make retina versus those that failed to make retina for genes of each indicated category along the x axis (M). A similar plot is shown for the promoters of each gene class (N) or enhancer class (O). Each bar is the average of all genes across all lines.

(P) Bar plot of the percentage of genes or chromatin regions that reset to the ESC state for the lines that made retina versus those that failed to make retina. Each bar represents the average of all genes across all lines.

SE, superenhancer.

See also Tables S3 and S4.
epigenetic state was fully reset to the ESC state or whether it was a better match for one of the 9 developmental stages (embryonic day [E]14.5, E17.5, P0, P3, P7, P10, P14, P21, 11 months) with chromHMM data in the retina. There were retinal cell-type-specific genes such as Aipl1 that were reset to the ESC pluripotent state in all 33 iPSC lines analyzed, regardless of their cellular origins (Figures 4F–4I; Tables S6 and S7). However, we also identified retinal-specific genes such as Stx3 that failed to completely reset to the ESC pluripotent state in our iPSC lines and were a better match for the developing retina (Figures 4J–4M; Table S6). As noted earlier, Stx3 reset to the ESC state in over 95% of iPSC lines based on DNA methylation, but there were differences in the resetting of the epigenome of Stx3 to the ESC state that were revealed by the ChIP-seq analysis that was not evident in the DNA methylation analysis. Overall, the majority of cell-type-specific genes for rods (63%), cones (60%), Müller glia (69%), amacrine cells (62%), bipolar cells (64%), and ganglion cells (63%) reset to the ESC pluripotent state (Tables 1 and S6). Similar results were obtained for cell-type-specific promoters (Tables 1 and S7). For those that did not match the ESC state as well as the developing retina, the differences were often subtle, with discrete regions of the gene or promoter having an epigenetic state that was a better match to retina than EB5:Rx-GFP ESCs (Figures 4J–4M).

Next, we analyzed the retinal enhancers and superenhancers identified previously (Aldiri et al., 2017), using the same Bayesian approach. The majority of retinal enhancers and superenhancers reset to the ESC pluripotent state (Figure S3; Tables 2 and S8). We analyzed the H3K9me3 regions of the genome, and they were reset at a similar efficiency (70%) as enhancers, superenhancers, genes, and promoters. However, retinal-specific heterochromatin predicted from the chromHMM (see Supplementary Information) was much less likely (20%–23% reset) to reset to the ESC state than active enhancers (72%–74% reset) or other regulatory regions (Table 2).

To determine whether the retinal-specific epigenetic memory was correlated with the ability of iPSCs to make retina, we compared the resetting efficiency of each class of genes, promoters, and enhancers in those iPSC lines that efficiently produce retina in 3D organoid cultures to the resetting efficiency in those lines that do not (Tables S6, S7, and S8). The retinal cell-type-specific genes, promoters, and enhancers were reset more efficiently in the iPSC lines that made retina relative to those that did not (Figures 4N and 4O), but there were no significant differences in the resetting of the broader chromatin domains (Figure 4P).

### Identification of a Biomarker for Retinal Differentiation in iPSCs

Genes encoding transcription factors that form core regulatory circuits (CRCs) often have cell-type-specific super-enhancers that contribute to a positive-feedback autoregulatory loop, and the same is true in stem cells (Saint-André et al., 2016). Using the CRC mapper (Saint-André et al., 2016) with our ChIP-seq data, we identified the core transcription factors in the iPSC lines used in our study. We also included EB5:Rx-GFP cells as a reference ESC line and several pre-iPSC lines from a previous publication (Chronis et al., 2017). Unsupervised hierarchical clustering revealed that 5 of the iPSC lines (rCh143.02, rG79.04, rCh257.4, rCr33.05, and FCR02) that did not make retina were more similar to the pre-iPSC lines in the CRC analysis than the EB5:Rx-GFP line (Figures 5A and 5B). Thus, by performing RNA-seq and H3K27Ac ChIP, and analyzing the CRCs, half of the iPSC lines that failed to make retina could have been

### Table 1. Proportion of Genes and Promoters that Reset to the ESC State in iPSC Lines

| Group          | Gene Reset | Promoter Reset |
|----------------|------------|----------------|
|                | Total (%)  | Retina (%)     | No Retina (%)  | Total (%)  | Retina (%)     | No Retina (%)  |
| Rod            | 63         | 63             | 63             | 64         | 64             | 64             |
| Cone           | 60         | 61             | 57             | 62         | 62             | 62             |
| Bipolar        | 64         | 66             | 60             | 68         | 70             | 63             |
| Müller         | 69         | 71             | 64             | 57         | 58             | 54             |
| Amacrine       | 62         | 64             | 57             | 63         | 65             | 60             |
| Ganglion       | 63         | 67             | 56             | 64         | 67             | 57             |
| G2/M           | 57         | 55             | 61             | 70         | 68             | 74             |
| Progenitor     | 63         | 64             | 61             | 61         | 61             | 61             |
| Housekeeping   | 59         | 58             | 61             | 68         | 68             | 68             |
| Stress         | 54         | 54             | 53             | 64         | 64             | 64             |
| Cancer         | 56         | 57             | 54             | 64         | 63             | 63             |
| Dev. spread    | 69         | 71             | 66             | 72         | 73             | 71             |
| Dev. shrink    | 62         | 64             | 58             | 62         | 63             | 60             |
| Retina         | 61         | 63             | 57             | 64         | 65             | 62             |
| All            | 60         | 60             | 55             | 62         | 63             | 60             |

See also Tables S5, S6, and S7.

aThe percentage is indicated for all 33 iPSC lines.
bThe percentage is indicated for the iPSC lines that make retina.
cThe percentage is indicated for the iPSC lines that did not make retina.
promoters and proteins from our ChIP-seq profiling at the H3K27Ac were useful in identifying iPSC lines that are defective for all the remaining marks combined (Tables S9 and S10). For comparison, there were only 82 genes and 127 promoters that had significantly elevated H3K4me2 in the lines that made retina relative to those that did not. Specifically, there were 186 genes and 739 promoters that had significantly elevated H3K4me2 in the lines that made retina relative to those that did not (Tables S9 and S10). For comparison, there were only 82 genes and 127 promoters that had differences in individual protein marks from ChIP-seq that correlated with the ability of lines to make retina for all the remaining marks combined (Tables S9 and S10). This was particularly striking because of previous studies that have highlighted the importance of H3K4me2 in reprogramming (Barrero et al., 2013; Koche et al., 2011; Liang and Zhang, 2013). For example, studies on the temporal order of changes in histone modifications during reprogramming had indicated that there is a rapid genome-wide increase in H3K4me2 at thousands of loci early in reprogramming, such as the pluripotency genes and lineage-specific differentiation genes (Barrero et al., 2013; Koche et al., 2011; Liang and Zhang, 2013). While H3K4me2 is a histone modification associated with euchromatin, changes in H3K4me2 do not usually result in changes in gene expression. This is consistent with the lack of a strong transcriptional signature for iPSC lines that made retina versus those that did not in our analysis.

For each of the 739 promoters with elevated H3K4me2 in the iPSC lines that made retina, we further filtered to the top 121 promoters with an area under the curve (AUC) greater than 0.8 from receiver operating characteristic (ROC) curve analysis (Supplemental Information). We performed ChIP enrichment analysis (ChiEA) (Lachmann et al., 2010) for the promoters and genes with elevated H3K4me2, and the top pathway was Meis1 target genes involved in optic cup formation (Table S11). The Meis1 target gene classification was also the top classification of the extended promoter list from the 739 genes with elevated H3K4me2 for the lines that made retina versus those that did not (Table S11). This was particularly interesting, because Meis1 is a homeodomain protein that is important for retinal development (Heine et al., 2008). Meis1 expression was lower in iPSCs that made retina versus those that did not (Figures 5C and 5D), and this was reflected in the epigenetic profile (Figures 5E–5H). Specifically, H3K27me3 is present at the promoter of the Meis1 gene in the EB5:Rx-GFP line and iPSC lines that efficiently make retina (Figures 5F and 5G), and the expression of Meis1 in those lines was essentially undetectable (FPKM < 1.0). However, for a subset of lines that failed to make retina, they expressed Meis1 and lacked the H3K27me3 histone modification at the promoter (Figure 5H). These data suggest that resetting of the Meis1 gene to the ESC epigenetic state revealed by ChIP-seq was not complete in all lines, and this had an impact on their ability to make retina.

To extend our analysis of the relationship between Meis1 expression and the ability of the retinal-derived iPSC line to differentiate into retina, we analyzed a group of iPSC lines that had not been previously analyzed. We found that 4 of the 8 iPSC lines in this group that failed to make retina had high levels of expression of Meis1, as measured by qPCR (Figure 5I). In total, 41 lines that made retina had levels of Meis1 expression that were less than 50% of the levels found endogenously in the adult retina. Among the 16 lines that failed to make retina, 31% (5/16) had levels of Meis1 expression similar to that of adult retina (Figure 5I).

The remaining 5 lines that failed to make retina and clustered with normal ESCs in the CRC analysis were outliers for individual epigenetic marks. Line rCh209.4 is an outlier for H3K36me3 (Figure 5J); rM0408, rCr263.4, and fCR01 are outliers for Pol II; and fG140.02 is an outlier for H3K4me1 (Figure 5J; data not shown). DNA methylation was less informative for identifying iPSC lines that failed to make retina in our analysis (Table S5; data not shown). Taken together, our data highlight the importance of performing integrated analysis with RNA-seq, DNA methylation profiling, and ChIP-seq to study epigenetic memory and to prospectively identify lines that will efficiently produce retina.

**DISCUSSION**

In this study, we reprogrammed retinal neurons and Müller glia and discovered that there are cell-type-specific differences in reprogramming efficiency. While this may be due in part to the epigenetic landscape and nuclear organization of individual classes of retinal neurons, there are likely to be additional factors that contribute to differences in reprogramming efficiency. We also compared two different developmental stages for each of the 5 cell types. For those cell types that were more difficult to reprogram (rods and bipolar neurons), the immature cells were more...
readily reprogrammed than their mature fully differentiated counterparts. In order to relate the reprogramming efficiency to differentiation, we carried out retinal differentiation using a quantitative assay called STEM-RET (Hiler et al., 2015, 2016). Using this approach, we discovered that the lines derived from those cells that were more difficult to reprogram had better STEM-RET scores than iPSC lines derived from cells that had higher reprogramming efficiencies. Subsequent epigenetic profiling indicated that this was due to retention of epigenetic memory primarily made up of repressive marks. Taken together, these data suggest that there is an inverse correlation between reprogramming efficiency and epigenetic memory. Our study also identified a set

Figure 5. Prospective Identification of iPSC Lines with the Potential to Make Retina in 3D Organoids
(A) Heatmap of unsupervised hierarchical cluster of core regulatory circuit (CRC) analysis for all iPSC lines in this study, EB5:Rx-GFP ESCs, MEFs, and pre-iPSCs from MEFs. Blue indicates that the gene is a CRC in that cell line. The red box indicates a separate cluster with iPSC lines that fail to make retina (red).
(B) Representative ChIP-seq for the superenhancer for Pou5f1 from the CRC heatmap in (A) showing that Pou5f1 is a CRC gene in the ESCs, and two lines that make retina but did not reset to the ESC state in two lines that failed to make retina. Those lines had the CRC and ChIP-seq profile of MEFs and pre-iPSCs.
(C and D) Bar chart of RNA-seq (C) and qPCR for Meis1. A subset of the lines that failed to make retina (gray) had elevated levels of Meis 1 expression (D). Each bar in (D) represents the mean and SD of triplicate PCR reactions for the same cDNA sample. The dashed lines indicate the expression of Meis 1 in adult mouse retina.
(E–H) ChIP-seq and chromHMM of the Meis1 gene for E14.5 mouse retina when the gene is expressed at its peak (E), the EB5:Rx-GFP ESC line (F), an iPSC line that makes retina and has low expression of Meis 1 (G), and an iPSC line that fails to make retina (H). Arrows indicate a major peak of H3K27me3 at the promoter of the Meis1 gene.
(I) Bar chart of qPCR for Meis1 for 40 iPSC lines that make retina and 16 lines that fail to make retina (gray box). The blue arrows indicate independent lines that were used as a validation cohort for Meis 1 expression to predict lines that fail to make retina. The dashed line indicates the expression of Meis 1 in adult mouse retina.
(J) PCA plot of H3K36me3 in genes across 29 iPSC lines. The light red dots indicate samples that are outliers in this PCA, and the dark red dot indicates a sample that was not identified in previous CRC analysis or Meis 1 expression analysis. The yellow/orange dots indicate the individual iPSC lines that are similar in this PCA.
(K) ChIP-seq for H3K36me3 and chromHMM for EB5:Rx-GFP ESCs, an iPSC line that makes retina (rCh257.4), and an iPSC line that fails to make retina (rCh209.4).
See also Tables S9, S10, and S11.
of biomarkers, including Meis1 and CRCs, that can be used to prospectively identify stem cell lines with the ability to make retina. Additional studies will be required to determine whether this extends to human iPSCs that may, one day, be used for cell-based therapy to treat retinopathies.

Reprogramming Efficiency in the Retina

In order to quantitate the reprogramming efficiency of retinal cell types at different stages of development, we took advantage of the Col1a1-OKSM;Rosa26-M2rtTA reprogrammable mouse strain (Stadtfeld et al., 2010). The advantage of this mouse strain is that OKSM can be induced with the addition of doxycycline in a uniform manner with temporal control. We purified individual reprogrammable retinal cell types by sorting and then mixed them with an excess of non-reprogrammable retinal cell types to ensure that they survive during the 10 days of reprogramming in retinal pellet cultures, as described previously (Hiler et al., 2015, 2016). This is important, because previous studies on cortical neurons showed that they could not be reprogrammed when plated on plastic without the inactivation of p53. Our chimeric pellet cultures allowed us to reprogram retinal neurons without the inactivation of p53 and to provide a direct comparison across cell types and developmental stages.

Our data suggest that there are differences across cell types that are consistent with reprogramming efficiency. For example, rods are difficult to reprogram, and they have the highest proportion of heterochromatin (Figure 2C), lack endogenous expression of any of the reprogramming genes (Figure S1), and have Myc sequestered in the f-heterochromatin domain (Figures 2E and 2F). In contrast, cell types that are easy to reprogram, such as Müller glia and a subset of amacrine interneurons, have a lower proportion of their nuclear volume sequestered into heterochromatin (Figure 2C) and express the Sox2 protein (Figure S1). However, these data do not explain all aspects of the differences in reprogramming efficiency. Specifically, bipolar neurons have the highest proportion of euchromatin in their nuclei, but they have the lowest reprogramming efficiency. It is possible that there are other aspects of genomic organization or other cellular features of bipolar neurons that make them difficult to reprogram using our mosaic pellet culture approach.

Retinal Differentiation from r-iPSCs

In this study, we showed that virtually all r-iPSCs were equivalent in terms of the cell types produced, whether they were from immature or mature neurons or glia. In our previous study, we showed that the retina produced from iPSCs made from fibroblasts had specific defects in production of some retinal cell types (Hiler et al., 2015, 2016). While the r-iPSCs made intact retina, there were differences in the efficiency of retinal differentiation, as measured using the STEM-RET scoring system that incorporates quantitative measures of molecular and cellular features characteristic of eye field specification, optic cup formation, and retinal differentiation (Hiler et al., 2015, 2016). The mature rods and immature bipolar neurons had the lowest reprogramming efficiency, yet they had the highest STEM-RET score. Taken together, these data suggest that there is both positive and negative epigenetic memory in iPSCs. The positive epigenetic memory is present in the rod- and bipolar-derived iPSCs, and the negative epigenetic memory is present in the f-iPSCs.

The implication of this study is important in considering the cell source for reprogramming. If whole retina were used, those cells with the greatest propensity for reprogramming would be the ones most likely to produce iPSCs clones. As a result, the subsequent retinal differentiation would be somewhat less efficient. Instead, by performing comprehensive profiling of the individual cell types, it is possible to identify those cell types that produce the iPSCs that make the best retina. In the retina, the cell types that are most difficult to reprogram made the best retina, but additional studies on other regions of the CNS or other tissue would be required to determine whether the inverse relationship between reprogramming efficiency and differentiation is a general property of iPSCs or unique to the retina.

Retinal Epigenetic Memory

We found that those retinal-specific genes and enhancers that undergo changes in their epigenetic state during development (Aldiri et al., 2017) were most likely to be properly reset to the native state found in ESCs. In contrast, the regions of the genome that were in more repressed epigenetic states and sequestered in heterochromatin in a cell-type-specific manner were the regions most likely to be retained in the iPSCs from retinal cell types. This was more prevalent in the iPSCs that were derived from cells that were most difficult to reprogram (rods and bipolar cells). These data are consistent with previous studies on the OKSM-mediated reprogramming of fibroblasts into iPSCs that suggested that there is a stepwise process that involves silencing of the cell-type-specific differentiation genes, resetting of poised enhancer elements, and reorganization of the epigenetic state can lead to epigenetic memory of the cellular origin for individual iPSC lines that, in turn, can impact subsequent lineage-specific differentiation. DNA methylation changes are thought to occur last during the reprogramming process, and as a result, epigenetic memory is often reflected in the DNA methylation of iPSCs (Lister et al., 2011). Our data are consistent with those mechanisms and show that the histone modifications associated with heterochromatin are also an important component of epigenetic memory. However, we also discovered that there are genes that appear to reset completely to the ESC state by DNA methylation but had incomplete resetting of other epigenetic marks. To gain a more complete understanding of epigenetic memory in iPSCs, it may be important to include ChIP-seq and other methods to profile the epigenome in future studies on epigenetic memory.

Biomarkers to Predict Retinal Differentiation

One of the major challenges in retinal differentiation of iPSCs is the prospective identification of lines that have the highest likelihood of making retina. In this study, we identified a set of stem cell lines that were fully reprogrammed but failed to make retina.
They were indistinguishable from the lines that made retina, using standard assays for measuring pluripotency. However, using our STEM-RET protocol, they failed at the very early stage of eye field specification. We discovered that Meis1 expression and the epigenetic state of Meis1 and its target genes were useful markers of retinal potential in a subset of 30%–50% of iPSC lines. It is important to emphasize that this link between Meis1 and retinal differentiation is counter-intuitive. One might assume that retaining expression of Meis1 would promote retinogenesis, as it is required for early retinal development. However, the opposite was true. One explanation for this observation is that formation of the retina is a stepwise process in iPSC-derived organoids as it is in vivo, and expression of Meis1 before it is required for eye field specification disrupts the entire process. Indeed, the lines that fail to make retina are defective in the earliest stages of retinal differentiation, consistent with the hypothesis that they fail to successfully transition through the requisite anterior neural specification. These data may also inform efforts to directly transdifferentiate fibroblasts or other cell types to retinal neurons. That is, it may be much more difficult to directly produce retinal neurons from fibroblasts if the cells do not transition through the key steps of early neurogenesis before retinal specification.

In addition to Meis1, CRC analysis was very helpful in identifying lines that were not fully reprogrammed from an epigenetic state, even though they met standard molecular and cellular criteria for reprogramming. At a minimum, H3K27Ac ChIP-seq and RNA-seq are required for CRC analysis, so by performing those two assays, the elevated Meis1 and CRC analysis could be performed and would prospectively identify the lines that are less likely to make retina.

**EXPERIMENTAL PROCEDURES**

**Animals**

All procedures were approved by the Institutional Animal Care and Use Committee at St. Jude Children’s Research Hospital. The reprogrammable mouse strain (Col1a1-OKSM; Rosa-26-M2rtTA mice) (Stadtfeld et al., 2010) was crossed with five different GFP transgenic mouse strains, including Nrl-GFP mice (Hiler et al., 2015), Chmbl4-GFP mice (Siegert et al., 2009), Crabp-GFP mice (Vázquez-Chona et al., 2009), Gad1-GFP mice (Huckfeldt et al., 2009), and Grm6-GFP mice (Dhingra et al., 2008), to produce reprogrammable mouse strains with transgenes labeling specific types of retinal cells that could be isolated by fluorescence-activated cell sorting (FACS). C57BL/6J mice (Jackson Laboratory, 000664) were used for supporting cells in retinal explant cultures for retinal reprogramming. CD-1 nude mice (Jackson Laboratory) were used to test the pluripotency of stem cell lines by teratoma formation.

**Statistical Analysis**

For differential analysis of RNA-seq, gene-based raw counts were analyzed by Voom after TMM. Log2 fold change, t values, p values, and FDR (false discovery rate)-corrected p values were extracted from Voom results and provided with FPKM values and AUC (area under the ROC curve) values. For ChIP-seq, we extended reads to the estimated fragment size and counted read numbers for promoters (transcription start site [TSS] ± 2 kb) or genes (TSS – 2 kb to TES + 2 kb) and performed statistical analysis with Voom. Instead of FPKM, we provided counts per million (CPM). For gene enrichment analysis, p values were determined using Fisher’s exact test requested from the Enrichr server, while q values (i.e., FDR-corrected p value) were determined using the Benjamini-Hochberg method for correction for multiple-hypothesis testing.

**DATA AND SOFTWARE AVAILABILITY**

The accession number for the data reported in this paper is GEO: GSE87064. All the data are also hosted on a free online viewer at https://pecan.stjude.org/proteinpaint/study/retinaliPSC2017.

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes Supplemental Experimental Procedures, three figures, and 11 tables and can be found with this article online at https://doi.org/10.1016/j.celrep.2018.01.075.

**ACKNOWLEDGMENTS**

We thank Angela McArthur for editing the manuscript. This work was supported, in part, by Cancer Center Support (CA21765) from the NCI, grants to M.A.D. from the NIH (EY014867, EY018599, and CA168875), and ALSAC. M.A.D. was also supported by the Tully Family Foundation, the Peterson Foundation, and HHMI. The 3D EM was performed with assistance from the Cell and Tissue Imaging Shared Resource at St. Jude, and the FISH analysis was performed with assistance from the Cytogenetics Shared Resource at St. Jude.

**AUTHOR CONTRIBUTIONS**

Conceptualization, M.A.D., D.H., and L.W.; Methodology, B.X., X.C., M.V., A. Shirinifard, and S.T.; Formal Analysis, M.A.D., D.H., B.X., X.C., X.Z., and A. Shirinifard; Investigation, M.A.D., D.H., L.W., M.V., L.G., J.Z., D.J., S.F., I.A., and M.-E.B.; Writing, M.A.D.; Supervision, M.A.D., A. Sablauer, and X.C.; Project Administration, M.A.D.; Funding Acquisition, M.A.D.

**DECLARATION OF INTERESTS**

The authors declare no competing interests.

Received: December 15, 2017 
Revised: December 15, 2017
Accepted: January 24, 2018
Published: March 6, 2018

**REFERENCES**

Aldiri, I., Xu, B., Wang, L., Chen, X., Hiler, D., Griffiths, L., Valentine, M., Shirinifard, A., Thiagarajan, S., Sablauer, A., et al. (2017). The dynamic epigenetic landscape of the retina during development, reprogramming, and tumor-igenesis. Neuron 94, 550–568.e10.

Aoi, T., Yae, K., Nakagawa, M., Ichisaka, T., Okita, K., Takahashi, K., Chiba, T., and Yamanaka, S. (2008). Generation of pluripotent stem cells from adult mouse liver and stomach cells. Science 321, 699–702.

Apostolou, E., and Hochedlinger, K. (2013). Chromatin dynamics during cellular reprogramming. Nature 502, 462–471.

Bar-Nur, O., Russ, H.A., Efrat, S., and Benvenisty, N. (2011). Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet beta cells. Cell Stem Cell 9, 17–23.

Barrera, M.J., Sese, B., Kuebler, B., Blic, J., Boue, S., Marti, M., and Izpisua Belmonte, J.C. (2013). Macrohistone variants preserve cell identity by preventing the gain of H3K4me2 during reprogramming to pluripotency. Cell Rep. 3, 1005–1011.

Beagan, J.A., Gilgenast, T.G., Kim, J., Plona, Z., Norton, H.K., Hu, G., Hsu, S.C., Shields, E.J., Lyu, X., Apostolou, E., et al. (2016). Local genome topology proteinpaint/study/retinaliPSC2017.

Chen, X., Pappo, A., and Dyer, M.A. (2015). Pediatric solid tumor genomics and developmental pancytology. Oncogene 34, 5207–5215.
Cells derived from rod photoreceptors. Cell Stem Cell 17, 101–115.

Kim, J., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M.J., Ji, Z.J., and Wong, R.O. (2009). Transient neurites of retinal horizontal cells exhibit columnar tiling via homotypic interactions. Nat. Neurosci. 12, 1197–1204.

Saint-André, V., Federation, A.J., Lin, C.Y., Abraham, B.J., Reddy, J., Lee, T.I., Bradner, J.E., and Young, R.A. (2016). Models of human core transcriptional regulatory circuits. Genome Res. 26, 385–396.

Siegert, S., Scherf, B.G., Del Punta, K., Didkovsky, N., Heintz, N., and Roska, B. (2009). Genetic address book for retinal cell types. Nat. Neurosci. 12, 1197–1204.

Singh, V.K., Kalsan, M., Kumar, N., Saini, A., and Chandra, R. (2015). Induced pluripotent stem cells: applications in regenerative medicine, disease modeling, and drug discovery. Front. Cell Dev. Biol. 3, 2.

Solovei, I., Kreysing, M., Lanctôt, C., Kösem, S., Peichl, L., Cremer, T., Guck, J., and Joffe, B. (2009). Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. Cell 137, 356–368.

Soufi, A., Donahue, G., and Zaret, K.S. (2012). Facilitators and impediments of the pluripotency reprogramming factors’ initial engagement with the genome. Cell 157, 994–1004.

Stadtfeld, M., Maherali, N., Borkent, M., and Hochedlinger, K. (2010). A reprogrammable mouse strain from gene-targeted embryonic stem cells. Nat. Methods 7, 53–65.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 137, 861–872.

Vázquez-Chona, F.R., Clark, A.M., and Levine, E.M. (2009). Rbp1 promoter drives robust Müller glial GFP expression in transgenic mice. Invest. Ophthal. Mol. Vis. Sci. 50, 3996–4003.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Franse, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science 318, 1917–1920.
Supplemental Information

Retinal Cell Type DNA Methylation and Histone Modifications Predict Reprogramming Efficiency and Retinogenesis in 3D Organoid Cultures

Lu Wang, Daniel Hiler, Beisi Xu, Issam AlDiri, Xiang Chen, Xin Zhou, Lyra Griffiths, Marc Valentine, Abbas Shirinifard, András Sablauer, Suresh Thiagarajan, Marie-Elizabeth Barabas, Jiakun Zhang, Dianna Johnson, Sharon Frase, and Michael A. Dyer
SUPPLEMENTAL ITEM TITLES AND LEGENDS

Figure S1 related to Figure 2. Epigenetic landscape of the 4 reprogramming genes in adult retina. A-D) ChromHMM of Oct4, Sox2, Klf4 and Myc across development in the mouse retina. The heatmap for the chromHMM states is shown in Fig. 2 and is available online at https://pecan.stjude.org/#/proteinpaint/study/retina2017. The expression of each gene at each stage is shown in the histogram to the right in each panel. Below the chromHMM is the ChIP-seq for H3K27me3 and DNA methylation using whole genome bisulfite sequencing. E-H) Micrographs of co-immunolocalization of Sox2 protein with Cralbp in Müller glia (E), and lack of localization with calbindin in horizontal neurons (F), PKC-alpha in bipolar cells (G) and HPC1 in amacrine cells (H). Abbreviations: FPKM, fragments per kilobase per million reads. Scale bars, 10 µm.
Figure S2 related to Figure 2. Epigenetic landscape of the 4 reprogramming genes in rods, bipolars and MEFs. All data are available online at

https://pecan.stjude.org/#/proteinpaint/study/retina2017. ChIP-seq tracks with scale indicated for each track is shown for purified bipolar cells, rods and MEFs. Dashed lines indicate the start and end of transcription for each of the reprogramming genes (Klf4, Myc, Oct4 and Sox2).
Bipolar cells

- H3K4me1
- H3K4me2
- H3K27Ac
- H3K27me3

Myc

Oct4 (Pou5f1)

Sox2

Rod photoreceptors

- H3K4me1
- H3K4me2
- H3K27Ac
- H3K27me3

Myc

Oct4 (Pou5f1)

Sox2

MEFs

- H3K4me1
- H3K4me2
- H3K27Ac
- H3K27me3

Myc

Oct4 (Pou5f1)

Sox2
Figure S3 related to Table 2. Resetting the epigenetic landscape of enhancers to the ESC state in retinal derived iPSCs. A) Heatmap of the chromatin Hidden Markov Modeling (chromHMM) states. The chromHMM states are color coded below with the association to specific regions in the genome. B-D) ChIP-seq tracks and chromHMM for adult retina for the superenhancer downstream of the Crx gene (dashed box) in adult retina (B). The same region is not active in the ESCs (C) and it was reset to the ESC state in retinal derived iPSCs (D). E-G) ChIP-seq tracks and chromHMM for adult retina (E) and EB5 ESCs (F) and an iPSC line that failed to reset the enhancer properly.
A

B

C

D

E

F

G

H3K9/14Ac
H3K27Ac
H3K4me1
H3K4me2
H3K4me3
BRD4
RNAPolII
H3K36me3
H3K27me3
H3K9me3
input
ChromHMM

1.5 kb

ChromHMM state

1 2 3 4 5 6 7 8 9 10 11

promoter/enhancer
full
gene
body
polycomb
empty
repressed
insulator

P21 Retina (FPKM=500)
Crx

EB5 ESC (FPKM<1.0)
Crx

iPSC 208.04 (FPKM<1.0)
Crx

P21 Retina (FPKM=1.5)
Sdk2

EB5 ESC (FPKM<1.0)
Sdk2

iPSC 3302 (FPKM<1.0)
Sdk2
Table S1 related to figure 1. Stem cell lines used for epigenetic profiling and retinal differentiation (provided as additional excel file).

Table S2 related to figure 3. STEM-RET differentiation of iPSCs (provided as additional excel file).

Table S3 related to figure 4. Transcriptome analysis of iPSC lines (provided as additional excel file).

Table S4 related to figure 4. Pathway analysis of differentially expressed genes (provided as additional excel file).

Table S5 related to Table 1. Resetting of DNA methylation to ESC state (provided as additional excel file).

Table S6 related to Table 1. Resetting of epigenetic landscape of genes to ESC state (provided as additional excel file).

Table S7 related to Table 1. Resetting of epigenetic landscape of promoters to ESC state (provided as additional excel file).

Table S8 related to Table 2. Resetting of enhancers to ESC state (provided as additional excel file).

Table S9 related to Table 2. ChIP-seq analysis of reset genes (provided as additional excel file).

Table S10 related to Table 2. ChIP-seq of reset promoters (provided as additional excel file).

Table S11 related to Table 2. Pathway analysis of reset promoters and genes (provided as additional excel file).
### Supplemental Experimental Methods

#### Key Resources Table

| Reagent or Resource | Source | Identifier  |
|---------------------|--------|-------------|
| Antibodies          |        |             |
| Anti-H3K4me3        | Diagenode | C15410003-50 |
| Anti-H3K4me2        | Abcam  | ab7766      |
| Anti-H3K4me1        | Abcam  | ab8895      |
| Anti-H3K9/14Ac      | Diagenode | C15410200 |
| Anti-H3K27Ac        | Abcam  | ab4729      |
| Anti-H3K36me3       | Active Motif | 61101 |
| Anti-H3K27me3       | Active Motif | 39155 |
| Anti-CTCF           | Active Motif | 61311 |
| Anti-RNA Pol II     | Abcam  | ab5095      |
| Anti-BRD4           | Bethyl Laboratories | A301-985A |
| Anti-H3K9me3        | Active Motif | 39161 |
| cMyc                | Santa Cruz | SC-40 |
| Nanog               | Cosmo Bio | RCAB0002P-F |
| Oct3/4              | BD Biosciences | 611202 |
| SSEA1               | Millipore | MAB4301 |
| Calbindin-D-28K     | Sigma-Aldrich | C9848 |
| Syntaxin            | Sigma-Aldrich | S0664 |
| PKC-alpha           | Millipore | 05-154 |
| Antibody                      | Manufacturer               | Catalog Number |
|-------------------------------|----------------------------|----------------|
| cone arrestin                 | Millipore                  | AB15282        |
| Glutamine synthetase          | BD Biosciences             | 610518         |
| Pax6                          | Developmental Studies      |                |
| Chx10                         | Exalpha Biologicals        | X1180P         |
| recoverin                     | Millipore                  | AB5585         |
| rhodopsin                     | Gift from R. Molday        | N/A            |
| HPC-1                         | Sigma Aldrich              | S0664          |
| cytokeratin                   | Biolegends                 | 908203         |
| desmin                        | Thermo Fisher              | RB9014         |
| synaptophysin                 | Spring Bioscience          | E2172          |

**Method details**

**Generation of iPS cells derived from different types of retinal cells**

Each class of the five retinal neurons investigated in this study was isolated through FACS from corresponding GFP transgenic mouse strain at two distinct developmental stages as follows:

| Neuron Type       | Immature | Mature |
|-------------------|----------|--------|
| Rod               | P2       | P21    |
| Bipolar           | P6       | P21    |
| Cone              | P2       | P21    |
| Muller            | P6       | P21    |
| Amacrine/horizontal | P4     | P21    |
After sorting, each sample was reprogrammed into iPS cells according to the protocol described as (Hiler et al., 2015; Hiler et al., 2016). Briefly, the GFP positive retinal cells were co-cultured with $1.5 \times 10^6$ dissociated retinal cells from postnatal day 0-5 C57BL/6J mice as a retinal pellet on polycarbonate filters in serial dilutions. The limiting dilution series (60,000, 6,000, 600, 60 or 30,000, 3,000, 300, 30 cells per pellet) was used to determine the minimum number of neurons required for a reprogramming to occur. After the incubation for 10 days with doxycycline (2 mg/ml), each pellet was dissociated and seeded in a serial dilution (20,000, 2,000, 200 cells per well) onto gelatin-coated 6-well plates with a feeder cell layer of irradiated mouse embryonic fibroblasts (IRMEFs) in the presence of leukemia inhibitory factor (LIF). In order to calculate the reprogramming efficiency of each cell line, colony-forming events were recorded for each pellet prior to isolation of individual iPS cell lines for subclone expansion. Established iPSC lines were expanded for at least twenty passages for further characterization.

**Scoring colony formation after reprogramming**

The ability of retinal neurons to reprogram was quantified with a two-step limiting dilution series followed by scoring of the number of colonies that grew out. Reprogrammed retinal iPSCs were seeded at 200, 2,000, and 20,000 cells per well on an IR-MEF feeder layer in a 6-well tissue culture plate. Retinal explants were seeded in triplicate at each dilution and were maintained for 4 days with daily media exchanges. Four days after seeding, bright-field image montages were taken for each well with a 4x Plan Fluor 0.13 NA objective and Coolsnap 4F camera (Photometerics) on a TE2000 confocal microscope (Nikon). The image montages were automatically
stitched together with a 10% image overlap with Elements imaging software (Nikon).

Image montages for individual wells or entire tissue culture plates were reduced in size by 50%, converted to a black and white image, and auto-scaled to adjust the brightness using Photoshop CS 5.1 (Adobe). The entire well within each image was digitally scored by identifying a colony and marking it with a red dot in a separate layer to preserve the original image. After each well was counted, the original image was then converted to a pure black image and merged with the layer indicating individual colonies to create a red-on-black binary image. These processed images were then automatically counted using Elements (Nikon) by setting the threshold to the red channel and using the automatic cell counting feature to identify individual colonies. The total number of colonies for each well was recorded and averaged to determine the number of colonies present at each dilution.

**Reprogramming efficiency calculations**

The equation used for calculating reprogramming efficiency is:

\[
\frac{\left( \frac{N}{P_t/P_p} \right) \times C}{R}
\]

N=number of iPSC colonies

P_t=total number of cells in pellet

P_p=total number of cell plated on feeders

C=proliferation constant (32)

R=number of retinal cells put in pellet for reprogramming
For example, if 60,000 reprogrammable MEFs were put into a retinal pellet of $1.5 \times 10^6$ non-reprogrammable cells and after 10 days of reprogramming, 20,000 of those cells are plated in a single well to give rise to 250 iPSC colonies, then the reprogramming efficiency is $\frac{(250(1,560,000/20,000)/32)}{60,000}=0.0102$ or 1.02%. This is in good agreement with direct reprogramming efficiency of MEFs from the OKSM mouse. The proliferation constant is included because each reprogrammed iPSC will divide within the retinal pellet before they are plated on feeders. If this constant is not included, it will artificially elevate the apparent reprogramming efficiency. The constant assumes 5 rounds of cell division within the pellet before the cells are plated. The doubling time for murine iPSCs has been estimated at between 5-10 hours so it is reasonable to assume 5 doublings within the retinal pellet before plating. This constant was used for all calculations so the reprogramming efficiency can be directly compared across individual cell types.

If there were multiple dilutions with scores, we used the data from wells that had the largest number of colonies to obtain the most reliable scores and we compared biological replicates that had the same number of cells put into the pellets and plated when possible.

**Characterization of iPSC lines**

The characterization of each established iPSC lines was carried out as described previously (Hiler et al., 2015; Hiler et al., 2016). For karyotyping, iPSC cell lines were cultured overnight on gelatin-coated 12-well plates with IRMEFs as previously described
and then incubated with 20 μl/ml colcemid solution (10 μg/ml, Life Technologies, 15212-012) for 4-5 hours to arrest the cells in metaphase. The cells were dissociated with trypsin (Cellgrow, 25052CI), lysed in 0.075 M potassium chloride, fixed and plated onto microscope slides. After air-drying, the chromosomes were stained with the traditional Giemsa-trypsin method. A fifteen-cell analysis was performed on each clone as cell morphology allowed. Metaphase cell analysis was performed using a Nikon Labphot light microscope under a PlanApo100X objective and images captured with a Leica CytoVision Imaging system equipped with software version 7.2.

To demonstrate that the lines were doxycycline-independent, iPS cell lines were cultured on gelatin-coated 12-well plates with IRMEFs for 5 passages in ESC maintenance medium that contained both doxycycline and LIF and medium that contained LIF only (no doxycycline). After 5 passages, doxycycline-independence was assessed by comparing colony number and morphology with and without doxycycline.

For alkaline phosphatase staining, iPS cell lines were seeded at 5,000-10,000 cells/well onto gelatin-coated 12-well plates with IRMEFs in ESC maintenance medium with or without exogenous LIF. The corresponding medium was exchanged daily for 5 days. The loss of pluripotency without LIF was assessed by alkaline phosphatase staining (Millipore).

For quantitative PCR (qPCR), Direct-zol (Zymo Research, R2050) and High-Capacity RNA-to-cDNA (Thermo Fisher Scientific, 4387406) kits were used to isolate RNA and
synthesize cDNA from stem cell lines \((1\times10^6 \text{ cells/sample})\). Real time qPCR was performed using TaqMan Custom Array Cards (Hiler et al., 2016). Expression data were normalized to GAPDH. For stem cells, stem cell marker expression was compared to Eb5-RxGFP murine ES cells. Retinal marker expression data for the retinal spheres (RDQ) were compared to normal postnatal day 0 and 12 murine retinas. The RDQ score was integrated from the mean and standard deviation of the integrated normalized relative scores for each biological replicate.

For immunofluorescence, iPS cell lines were grown on gelatin-coated chamber slides overnight and then fixed with 4% paraformaldehyde for 30 minutes prior to immunostaining. Retinal spheres were fixed at 4°C overnight with 4% paraformaldehyde and then incubated at 4°C overnight in a 4% sucrose solution. The spheres were embedded in OCT compound (Tissue-Tek), frozen on dry ice, sectioned into 10 μM thick sections using a cryostat and allowed to air dry on slides prior to immunofluorescence staining. Stem cell, cortical and retinal samples were then incubated for 3-4 hours at room temperature with an appropriate blocking serum with 0.5% Triton X-100. Stem cells were incubated overnight at 4°C with the following antibodies: cMyc (Santa Cruz; 1:200), Nanog (Repro Cell; 1:500), Oct3/4 (BD Biosciences; 1:500) and SSEA1 (Millipore; 1:500). Retinal spheres were incubated with these antibodies: Calbindin-D-28K (Sigma-Aldrich; 1:100), Syntaxin (Sigma-Aldrich; 1:500), PKC-alpha (Millipore; 1:5000), cone arrestin (Millipore; 1:5000), Glutamine synthetase (BD Biosciences; 1:100), Pax6 (Developmental Studies Hybridoma; 1:100), Chx10 (Exalpha Biologicals; 1:500), and recoverin (1:5000). The samples were incubated for 1-2 hours with secondary antibodies
at room temperature and 30 minutes with an ABC Kit (Vectastain) at room temperature. Immunostaining was visualized by a 10-minute incubation with cyanine 3-conjugated-tyramide reagent (Perkin Elmer) followed by a DAPI counterstain.

For the teratoma formation assay, iPS cell lines were resuspended in matrigel (Corning) and injected into each flank of adult CD1-nude mice (200,000 cells/injection). Teratoma growth was monitored for up to 12 weeks after injection and removed before the tumor reached 20% of the animal’s body weight. Tumors were fixed in 4% paraformaldehyde at 4°C for several days prior to paraffin sectioning. Individual germ layers were identified by hematoxylin & eosin staining and by IHC with primary antibodies against: cytokeratin (BioLegend; 1:100), desmin (Thermo Scientific; 1:500), or synaptophysin (Spring Bioscience; 1:400).

**Cell culture**

Established iPS cell lines were maintained on gelatin-coated culture plates with a feeder cell layer of IRMEFs in ESC maintenance medium and incubated at 37°C with 5% CO₂. The ESC maintenance medium contained Glasgow Minimum Essential Medium (GMEM; Sigma-Aldrich) with 10% Knockout Serum Replacement (KSR; Thermo Fisher Scientific), 1% FBS, 1X Non-Essential Amino Acids (NEAA; Thermo Fisher Scientific), 1 mM sodium pyruvate (Thermo Fisher Scientific), 1X GlutaMax (Thermo Fisher Scientific), and 0.1 mM 2-mercaptoethanol (Sigma-Aldrich). Prior to use, 2,000 units/ml of Leukemia Inhibitory Factor (LIF, Millipore) and 2 μg/ml doxycycline was added to the medium. ESC maintenance medium was exchanged daily, and cell lines were passaged
every other day for twenty passages. Stem cells were then prepared for FAC sorting for SSEA1+ cells to produce a cell stock enriched in undifferentiated cells. Stem cells were isolated from IRMEFs, blocked in 25% human beta-globin in ESC maintenance medium, and incubated on ice for 15-20 minutes with SSEA1-PE antibody (BD Biosciences; 1:175) and rat IgG isotype control (BD Biosciences; 1:1500) prior to FACS. The threshold for PE+ cells was set against the isotype control. All SSEA1+ cells were collected and cryopreserved for use in characterization and retinal differentiation experiments.

**Retinal differentiation**

The established iPSC lines were differentiation into retina using a 3-dimensional culture method described as (Eiraku et al., 2011), and we quantified the progression of retinal morphogenesis using STEM-RET protocol described as (Hiler et al., 2015; Hiler et al., 2016). Stem cells were seeded onto 96-well Nunclon Sphera plates at 300 cells/well (Thermo Scientific Nunc) in retinal differentiation medium and incubated at 37°C with 5% CO₂. Retinal differentiation medium contained GMEM with 1.5% KSR, 1X NEAA, 1 mM sodium pyruvate, 1X GlutaMax, and 0.1 mM 2-mercaptoethanol. The following day, matrigel (1-2 mg/ml in retinal differentiation medium; Corning) was applied to the spheres. Cultures were incubated at 37°C with 5% CO₂ for an additional 6 days. After seven days, retinal spheres were scored such that spheres with one or more optic vesicle outcroppings with thick, laminated edges scored an “A”, spheres with moderate to high amounts of non-retinal tissue scored “B”, spheres that differentiated into non-retinal anterior neural tissue or had retinal features obscured by non-retinal tissue scored “C”,
and spheres with thinly-laminated edges and stalled retinal morphogenesis scored “X”. The A- and B-scoring spheres were transferred to maturation medium 1 (MM1) and incubated at 37°C with 40% O₂ for the remainder of retinal differentiation. MM1 contained DMEM/F-12 with 1X GlutaMax, 1,000 units/ml penicillin, 1,000 µg/ml streptomycin, and 1X N2 supplement (Thermo Fisher Scientific). On day 10, optic cup outcroppings were excised and transferred to maturation medium 2 (MM2) with 0.5 µM retinoic acid and 1 mM taurine (Sigma-Aldrich). MM2 consisted of DMEM/F-12 with 1X GlutaMax, 1X N2 supplement, 1,000 units/ml penicillin, 1,000 µg/ml streptomycin, and 10% FBS. On day 14, the culture medium was completely exchanged with MM2 with 1 mM taurine (no retinoic acid) and continued to the end of the STEM-RET protocol with half media exchanges every three days. Retinal tissue was harvested on day 28 after plating and assessed for mature retinal differentiation by 1) immunofluorescence of retinal markers (RDIF), 2) qPCR for retinal markers (RDQ), and 3) electron microscopy analysis (RDEM). The scoring of STEM-RET, RDIF, RDQ and RDEM are described as (Hiler et al., 2015; Hiler et al., 2016). To integrate retinal area, the combined STEM-RET score was multiplied by the percentage of organoid spheres that were retina.

**ChIP-seq for iPS cells**

After removal of IRMEFs, iPS cells were collected after washing with PBS. Then cells were subjected to crosslinking in 1% formaldehyde (Thermo scientific, 28906) for 10 minutes at room temperature, followed by the addition of glycine at the final concentration of 0.125 M to quench the reaction. After centrifuge, cells were washed with cold PBS, counted and pelleted as 10 million cells for each nuclei preparation. The nuclei
were isolated and prepared for shearing using TruChIP chromatin shearing kit (Covaris, 520127) according to manufacturer’s protocol. After shearing, ChIP was performed using the iDeal ChIP-seq kit (Diagenode, C01010051). The antibodies used were: anti-H3K4me3 (Diagenode, C15410003-50), anti-H3K4me2 (Abcam, ab7766), anti-H3K4me1 (Abcam, ab8895), anti-H3K9/14ac (Diagenode, C15410200), anti-H3K27ac (Abcam, ab4729), anti-H3K36me3 (Active Motif, 61101), anti-H3K27me3 (Active Motif, 39155), and anti-CTCF (Active Motif, 61311). After de-crosslinking, DNA was extracted using MinElut PCR-purification kit (Qiagen, 28006) or Agencourt AMPure XP beads (Beckman Coulter, A63881) and subjected to library preparation using NEBNext ChIP-Seq Library Prep Master Mix Set for illumina (NEB, E6240L). Each library was sequenced by HiSeq 2000 using single-end 50-bp reactions.

The ChIP experiments and library preparation for antibodies including anti-RNA polymerase II (Abcam, ab5095), anti-Brd4 (Bethyl Laboratories, A301-985A) and anti-H3K9me3 (Active Motif, 39161), were performed by Active Motif.

**ChIP-Seq for purified retinal cells**

Mature rods and bipolar cells were collected from *Nrl*-GFP and *Grm6*-GFP transgenic mice at P21-23 through FACS. Briefly, retinae from each strain were washed twice with PBS after dissection, cross-linked for 10 min in 1% formaldehyde (Thermo Fisher Scientific, 28906) at room temperature, quenched with Glycine at a final concentration of 0.125 M. Then the retinal tissue was washed with PBS and dispersed into a single cell suspension using a dounce homogenizer with 7 ml pestle (Thermo Fisher Scientific, 06-435A). The cell suspension was filtered through a bovine serum albumin (BSA, Sigma)
gradient to remove cellular debris, prior to collecting the population of GFP positive cells via FACS. Each cell suspension was brought to a total volume of 3 mL PBS and layered on a column of chilled BSA media (DMEM/F12 media with 4% BSA). The BSA column was centrifuged at 500 rcf for 10 minutes at 4°C.

After sorting, purified retinal neurons were pelleted and the nuclei were isolated and prepared for shearing using TruChIP chromatin shearing kit (Covaris, 520127) using microTUBE (Covaris, 520052) according to manufacturer’s protocol. After shearing, ChIP was performed using the iDeal ChIP-seq kit (Diagenode, C01010051) with minor modifications. Briefly, protein A magnetic beads were incubated with the antibody in a 200 μl reaction system at 4 °C on a rotator for at least 1 hour. Then 100 μl diluted sheared chromatin were added to the system and incubated overnight at 4 °C. For washing steps, add one more time for Buffer 1 and Buffer 4 respectively. After de-crosslinking, DNA was purified using Agencourt AMPure XP beads (Beckman Coulter, A63881) and subjected to library preparation using NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, E7645S). Each library was sequenced by HiSeq 2000 using single-end 50-bp reactions.

**Methyl-Seq**

Genomic DNA was extracted using DNeasy Blood & Tissue kit (Qiagen, 69504) for each iPS cell lines. Then DNA was processed using the SureSelectXT Mouse Methyl-Seq Reagent kit (Agilent, 931052). After capture, the bisulfite conversion of the eluted DNA was performed by EZ DNA Methylation-Gold kit (Zymo Research, D5005), and then the
converted DNA was amplified using the PCR reagents from SureSelectXT Mouse Methyl-Seq Reagent kit (Agilent, 931052) according to the manufacturer’s recommendations. Sequencing was performed using paired-end 100 base chemistry on an Illumina HiSeq2000 system.

**RNA-Seq**

RNA was extracted from each iPS cell lines by using RNeasy Plus Mini kit (Qiagen, 74134) or Direct-zol kit (Zymo Research, R2050). Libraries were prepared from ~500 ng total RNA with the TruSeq Stranded Total RNA Library Prep Kit according to the manufacturer’s directions (Illumina). Paired-end 100-cycle sequencing was performed on HiSeq 2000 or HiSeq 2500 sequencers according to the manufacturer’s directions (Illumina.)

**Real Time QPCR Validation**

RNA was isolated from each iPS cell lines by using RNeasy Plus Mini kit (Qiagen, 74134) and DNA was removed by TURBO DNA-free kit (Thermo Fisher Scientific, AM1907) according to the manufacturer’s instructions. CDNA was synthesized using SuperScript III Reverse Transcriptase (Thermo Fisher Scientific, 18080044) and then subjected to real time PCR using SYBR Select Master Mix (Thermo Fisher Scientific, 4472908). The primers used for each gene are listed below.

| Gene Symbol | Forward Primer       | Reverse Primer                  |
|-------------|----------------------|---------------------------------|
| *Meis1*     | CATGATAGACCAGTCCAA   | ATTGGCTGTCCATCAGGGTTA           |
|       |     |     |
|-------|-----|-----|
| **CCGA** | CCGA | ACCAGTATCGGCTATTGATCT |
| **Lyz2** | ATGGAATGGCTGGCTACT | ACCAGTATCGGCTATTGATCT |
|        | ATGG | GA   |
| **Boc** | CGAAGAGAGCGGCTATA | CCCAGGATCAGAGGTCC |
|        | CTT  | CTCC |
| **Smarca2** | CTCCTGGACCAATTCTGGG | CATCGTTGACAGAGGTG |
|        | G    | AG   |
| **Gapdh** | CTCCACTCACGGCAAA | CGCTCCTGGAAGA |
|        | TTCA | TGGT |

**Retinal Area Measurement**

The samples were fixed in 2.5% glutaraldehyde and 2% paraformaldehyde in 0.1M sodium cacodylate buffer pH 7.4 and post fixed in 2% osmium tetroxide in 0.1M sodium cacodylate buffer with 0.3% potassium ferrocyanide for 1.5 hours. After rinsing in same buffer the tissue was dehydrated through a series of graded ethanol to propylene oxide, infiltrated and embedded in epoxy resin and polymerized at 70°C overnight. Semithin sections (0.5 micron) were stained with toluidine blue for light microscope examination. Ultrathin sections (80nm) were cut and imaged using an FEI Tecnai F 20 TEM FEG Electron Microscope with A ATM XR41 Camera. Based on the light microscopy images for each sphere from the stained semithin sections, the retinal area was classified manually according to the retinal lamination and the rosette formation, and the percentage of retinal area was calculated as the ratio of retina area to the whole tissue area within the specific image.
**3DEM fixation and processing**

The samples for 3DEM were fixed in 2% paraformaldehyde and 2.5% glutaraldehyde in 0.1M cacodylate buffer overnight and rinsed in same buffer. The tissue was stained using a modified heavy metal staining method (Ellisman et al.) and processed through a graded series of alcohol, propylene oxide and infiltrated in propylene oxide /epon gradients. The tissue was then infiltrated overnight in 100% Epon 812 resin and polymerized for 48 hours in a 70°C oven.

**Preparation steps for focused ion beam scanning electron microscopy imaging**

The sample block is mounted on an aluminum pin stub with a conductive silver epoxy, and sputter-coated with a thin layer (~ 60nm) of iridium to electrically ground the sample reducing charging. The block face is scanned to locate the ROI. The block is trimmed to the ROI using an ultramicrotome with a diamond knife. Next a relief is milled into the block face using a suitable ion beam current, and a protective cap is deposited using the carbon gas injection system. Fiducials are created to aid the computer vision pattern placement algorithm.

**Focused ion beam scanning electron microscopy**

The samples are imaged and processed for 3D data collection on a Helios G3 system. Imaging is normally at 2kV, 100pA, 5x5x5nm, ICD. Regions of interest are imaged utilizing Auto Slice and View 4 software package to automate serial sectioning and data collection processes.
**Preparation steps for serial blockface scanning electron microscopy imaging**

The sample block is mounted on an aluminum pin stub with a conductive silver epoxy, and sputter-coated with a thin layer (~ 30nm) of iridium to electrically ground the sample reducing charging. The block face is scanned to locate the ROI. The block is trimmed to the ROI using an ultramicrotome with a diamond knife.

**Serial Blockface Scanning Electron Microscopy**

The samples are imaged and processed for 3D data collection on a Teneo Volumescope. Imaging is normally at 2kV, 100pA, 25x25x50nm, LV-BSD.

**Scanning Electron Microscopy**

The samples are imaged on a Helios G3 system. The blockface is normally scanned at an acceleration voltage of 10kV and a current of 800pA on an SEM using a concentric backscatter detector.

**Image Processing and Segmentation**

Image stacks are aligned, filtered, cropped, segmented, and rendered using the Amira software package. Rendered volumes are false colored. Scale bars are only valid in one plane of three dimensional volume projections.

**RNA-Seq data processing**
FASTQ sequences were mapped to the mouse mm9 (MGSCv37 from Sanger) by StrongARM, developed for the PCGP (http://www.ncbi.nlm.nih.gov/pubmed/22641210). To estimate the FPKM values, we first downloaded GTF files (mouse vM7) from GENCODE website (http://www.ncbi.nlm.nih.gov/pubmed/22955987) and convert the coordinates to mm9 by CrossMap version 0.1.5 (http://www.ncbi.nlm.nih.gov/pubmed/24351709). Then only “transcript” and “exon” records from the GTF files were retained for compatible of feed to cuffdiff (from Cufflinks package version 2.1.1) along with default parameters and “--frag-bias-correct --multi-read-correct -time-series”.

**ChIP-Seq data processing**

We first employ BWA (version 0.5.9-r26-dev, default parameter) to align the ChIP-Seq reads to mouse genome mm9 (MGSCv37 from Sanger, Picard(version 1.65(1160)) then have been used for marking duplicated reads. Then only non-duplicated reads with have been kept by samtools (parameter “-q 1 -F 1024” version 0.1.18 (r982:295)). We followed the ENCODE criterion to quality control (QC) the data that non-duplicated version of SPP (version 1.11) have been used to draw cross-correlation and calculated relative strand correlation value (RSC) under support of R (version 2.14.0) with packages caTools (version 1.17) and bitops(version 1.0-6) and estimated the fragment size. We required > 10M unique mapped reads for point-source factor (H3K4me2/3, H3K9/14ac, H3K27ac, CTCF, RNAPolII, BRD4) and RSC > 1. We required 20M unique mapped reads for broad markers (H3K9me3, H3K27me3, H3K36me3). We required 10M unique mapped reads for INPUTs and RSC < 1. We noticed H3K4me1 is point-source factor in
some stages while broad in other stages so we QC H3K4me1 as broad markers. Then upon manually inspection, the cross-correlation plot generated by SPP, the best fragment size estimated (the smallest fragment size estimated by SPP in all our cases) were used to extend each reads and generate bigwig file to view on IGV (version 2.3.40). All profiles were manually inspected for clear peaks and good signal to noise separation. For mouse data, all data broad markers with RSC < 0.8 have biological replicates except several samples with one of antibody for H3K9me3 (due to availability of outsourcing), we think the quality of those H3K9me3 are good since they show clear peaks and RSC are bigger than most of the H3K9me3 data available in ENCODE or Epigenomics Roadmap.

To estimate the consistence between biological replicates, for point-source factors, MACS2 (version 2.0.9 20111102, option “nomodel” with “extsize” defined as fragment size estimated above) have been used to call peaks with FDR corrected p-value cutoff 0.05 (strong peak set) and 0.5 (weak peak set) separately, peaks within 100bp have been merged by bedtools (version 2.17.0). We required > 80% of strong peak set from either one replicate overlapping the weak peak set from the other replicate. For broad peaks, SICER (version 1.1 redundancy threshold 1, window size 200bp, effective genome fraction 0.86, gap size 600bp, FDR 0.00001 with fragment size defined above) has been used for domain calling and we required > 70% of domains called overlapping the domains called from the other replicate (median percentage 89.9%). After confirmed the consistency between replicates, we pool extended reads and generate the final bigwig track for visualization.

**Target Enrichment Bisulfite Sequencing Analysis**
Target enriched bisulfite sequencing (TEBS, a.k.a Methyl-Seq) data were mapped to mm9 (mouse) by BSMAP2.9 with the following parameters: “-z 33 -f 5 -g 3 -r 0 -m 17 -x 600 -u”. We confirmed the converting rate > 99%, after remove duplicated reads marked by Picards(v2.0.1), methratio.py script from BSMAP have been used to count the number of C or T at each CpG sites.

We first clustered CpG sites with > 5 reads together if they are within 100bp of each other in any of the samples, removed the CpG clusters with <= 5 CpG sites and then merged CpG clusters only if appears in at least two samples. For each of the merged CpG clusters we test each sample whether it’s Beta-value is significant different from Eb5 if they are more than 5 CpG sites available both for the sample and Eb5 by student t-test.

We required both p-value < 0.05 and mean Beta-value different > 0.1 to be considered as not reset to mESC. At last, for each gene region, promoter region, or any predefined region from previously work, we summarized whether they are not reset if there are more non reset CpG clusters than reset CpG clusters.

Combining Immunofluorescence Protein Staining with DNA FISH

Freshly cut unfixed cryosections were fixed in 1% PFA in PBS for 5 minutes, followed by an additional fixation in 1% PFA and 0.05% NP40. Fixed samples were stored in 70% ethanol at -20°C until needed. Following fixation slides were subjected to immunostaining by first blocking in 1% BSA and 2X SSC for 10 minutes followed by protein detection using various antibodies at appropriate dilutions in the same blocking solution. Detections were carried out at RT for 45 minutes followed by washing in PBS for 5 minutes. Appropriate secondary antibodies were applied using the same procedure
as for primary antibodies. Slides were then fixed in 4% PFA and 0.5% Tween 20 and 0.5% NP40 for 10 minutes, followed by treatment in 0.2N HCl containing 0.5% Triton X-100 for 10 minutes. Slides were then denatured in 70% formamide and 2X SSC at 80°C for 10 minutes. Following denaturation, the slides were dehydrated in a graded ethanol series consisting of 70%, 80%, and 100% ethanol for 2 minutes each. Denatured probes were then applied to the slides under 22 X 22 mm coverslips at 37°C overnight in a solution containing 50% formamide, 2X SSC, and 10% dextran. Slides were then washed in 50% formamide and 2X SSC at 37°C for 5 minutes followed by mounting in Vectashield containing DAPI.

| gene name | BAC clone | Genomic coordinates (mm10) | Ch band |
|-----------|-----------|---------------------------|---------|
| Pou5f1    | RP23-38P5 | 35434790-35654260          | 17B1    |
| Sox2      | RP23-423J10 | 34577449-34780793         | 3A3     |
| Klf4      | RP24-230N3 | 55483138-55664485          | 4B3     |
| Myc       | RP23-457I7 | 61969481-62141358          | 15D1    |

Confocal image analysis of FISH probes

Images were taken with Zeiss LSM 700 confocal microscope using 63X lens. To map FISH probe nuclear position, we employed machine-learning methods (Schindelin et al., 2012) to segment different nuclear region types (core: blue channel, DAPI-ring: blue channel and Green ring: green channel, FISH spots: red channel) in each 2D plane of the z-stacks. The machine-learning classifier uses 85 feature images for each plane that are created based on texture, gray-scale intensity and neighborhood information. We then use the 3DROIManager ImageJ plug-in (Ollion et al., 2013) that calculates common surface area between the segmented regions in the 3D segmented image. We use the values of common surface area to identify FISH spots inside each region.
**ChromHMM**

The ChromHMM was performed exactly as for the retinal ChIP-seq so we could directly compare the data (Aldiri et al., 2017). Non-duplicated aligned reads were extend by fragment size defined above and ChromHMM (version 1.10, with “-colfields 0,1,2,5 -center” for BinarizeBed) was used for chromatin state modeling. To choose the state number we first modeled all mouse development stage together from 7 states to 33 states. We want choose the state number most correlated to the biological event as greatest up/down regulation of genes here. So we take the FPKM values from RNA-seq then remove all genes that are FPKM <1 across all stages of mouse development. Of those that remain, genes FPKM is <1 at a particular stage have been normalized to 1. Than for 311 genes down 10 fold from P0 to P21 and 366 genes up 10 fold from P0 to P21, we tested the significance that each state in each model occupied different percentage of gene and flank region (from -2kb of TSS to +2kb of TES) by t-test. Order the state and model by smallest p-value we found larger percentage of state 9 in model 11 at P0 is greatest associated with genes down-regulated 10 fold from P0 to P21 while larger percentage of state 1 in model 11 at P0 is greatest associated with genes up-regulated 10 fold. We repeat similar analysis for genes up/down 10 fold from E14.5 to P14, model 11 is still top in all the models so we choose HMM model with 11 states. For better visualization of the dynamics of HMM state across stages, we normalized color intensity by the max total percentage of a state covered a gene and flanking region across mouse development and tumor. To determine the best region represent a gene, we first filtering annotated isoforms by TSS within 2kb of any
H3K4me3/H3K4me2/H3K27Ac/H3K9-14Ac peaks at any development stages and then we selected the highest expressed isoform at any development stages estimated by cuffdiff or the longest isoform if no expression level estimated by cuffdiff. At last, we reduced the interval for a HMM state to half bar and the intensity to half of the normalized intensity if it didn’t ranked top 2 HMM state for a gene. For HMM states could be assigned by multiple genes, the max total percentage across genes have been used for the normalization.

For the HMM state in Human, after processing of ChIP-seq data similar to mouse, we applied the Mouse 11 states HMM model to human data. The selection of isoform is similar but using human data while the max total percentage used for normalization of intensity is based on mouse if there is mouse ortholog for that gene.

**Calling Super-Enhancer Regions**

At each stage, after called the H3K27ac peaks with MACS2, in addition to ROSE’s distance cutoff parameter, we first combine peaks from replicates as potential enhancer region (this significantly improve consistency of super enhancer called between replicates by avoid false negative “bridging” weak peaks) and excluded peaks overlapping H3K4me3 peaks to make sure the super enhancers called are not due to promoter signal. Then we use ROSE (stitched if within 12.5kb, not counted as stitch if within 2.5kb of TSS) to call super-enhancers (SEs) independently for both replicate and retain the SEs called in both replicates. We finalized SE regions by merging SE called by ROSE (in
both replicates) at any stages or ranked as top 1000 at > 1 stages (in both replicates) and finally filtered by overlapping BRD4 peaks in any stages.

**CRC analysis**

The CRC mapper was run as described previously (Saint-Andre et al., 2016) except we performed two separate analyses to compare data with and without promoter regions. We used our H3K4me3 ChIP-seq data to remove those putative super-enhancers that overlap with promoters. This was done because some very actively expressed genes had large regions of H3K27Ac that are called as super-enhancers and it is difficult to distinguish between these putative super-enhancers and strong promoters at genes.

**References**

Aldiri, I., Xu, B., Wang, L., Chen, X., Hiler, D., Griffiths, L., Valentine, M., Shirinifard, A., Thiagarajan, S., Sablauer, A., et al. (2017). The Dynamic Epigenetic Landscape of the Retina During Development, Reprogramming, and Tumorigenesis. Neuron 94, 550-568 e510.

Eiraku, M., Takata, N., Ishibashi, H., Kawada, M., Sakakura, E., Okuda, S., Sekiguchi, K., Adachi, T., and Sasai, Y. (2011). Self-organizing optic-cup morphogenesis in three-dimensional culture. Nature 472, 51-56.

Hiler, D., Chen, X., Hazen, J., Kupriyanov, S., Carroll, P.A., Qu, C., Xu, B., Johnson, D., Griffiths, L., and Frase, S. (2015). Quantification of retinogenesis in 3D cultures reveals epigenetic memory and higher efficiency in iPSCs derived from rod photoreceptors. Cell Stem Cell 17, 101-115.

Hiler, D.J., Barabas, M.E., Griffiths, L.M., and Dyer, M.A. (2016). Reprogramming of mouse retinal neurons and standardized quantification of their differentiation in 3D retinal cultures. Nat Protoc 11, 1955-1976.

Ollion, J., Cochennec, J., Loll, F., Escude, C., and Boudier, T. (2013). TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization. Bioinformatics 29, 1840-1841.

Saint-Andre, V., Federation, A.J., Lin, C.Y., Abraham, B.J., Reddy, J., Lee, T.I., Bradner, J.E., and Young, R.A. (2016). Models of human core transcriptional regulatory circuitries. Genome Res 26, 385-396.
Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods 9, 676-682.