HHS Public Access

Author manuscript

Neuron. Author manuscript; available in PMC 2018 May 03.

Published in final edited form as:

Neuron. 2017 May 03; 94(3): 550–568.e10. doi:10.1016/j.neuron.2017.04.022.

The Dynamic Epigenetic Landscape of the Retina During Development, Reprogramming, and Tumorigenesis

Issam Aldiri[1,*], **Beisi Xu**[2,*], **Lu Wang**[1], **Xiang Chen**[2], **Daniel Hiler**[1], **Lyra Griffiths**[1], **Marc Valentine**[3], **Abbas Shirinifard**[4], **Suresh Thiagarajan**[4], **Andras Sablauer**[4], **Marie-Elizabeth Barabas**[1], **Jiakun Zhang**[1], **Dianna Johnson**[5], **Sharon Frase**[6], **Xin Zhou**[2], **John Easton**[2], **Jinghui Zhang**[2], **Elaine R. Mardis**[7,8,9], **Richard K. Wilson**[7,8,10], **James R. Downing**[11], and **Michael A. Dyer**[1,5,12,#] **for the St. Jude Children's Research Hospital – Washington University Pediatric Cancer Genome Project**

[1]Department of Developmental Neurobiology, St. Jude Children's Research Hospital, Memphis, Tennessee, 38105, USA

[2]Department of Computational Biology, St. Jude Children's Research Hospital, Memphis, Tennessee, 38105, USA

[3]Department of Cytogenetics Shared Resource, St. Jude Children's Research Hospital, Memphis, Tennessee, 38105, USA

[4]Department of Diagnostic Imaging, St. Jude Children's Research Hospital, Memphis, Tennessee, 38105, USA

[5]Department of Ophthalmology, University of Tennessee Health Science Center, Memphis, Tennessee 38163, USA

[6]Department of Cell and Tissue Imaging Shared Resource, St. Jude Children's Research Hospital, Memphis, Tennessee, 38105, USA

[7]The Genome Institute, Washington University School of Medicine in St. Louis, St. Louis, Missouri 63108, USA

[8]Department of Genetics, Washington University School of Medicine in St. Louis, St. Louis, Missouri 63108, USA

Correspondence and requests for materials should be addressed to: Michael A. Dyer, PhD, Department of Developmental Neurobiology, MS 323, St. Jude Children's Research Hospital, 262 Danny Thomas Place, Memphis, TN, 38105-3678, USA, Phone: (901) 595-2257, Fax: (901) 595-3143, michael.dyer@stjude.org.

*These authors contributed equally.

#Lead Contact

AUTHOR CONTRIBUTIONS

Conceptualization, I.A., B.X., X.C., M.V., and M.A.D.; Methodology, B.X., X.C., M.V., A.S., S.T., D.J., S.F., X.Z., J.E., and M.A.D.; Software, B.X., X.C., X.Z., and A.S. ; Formal Analysis, B.X., X.C., and A.S.; Investigation, I.A., L.W., D.H., L.G., M.V., M-E. B., and J.Z.; Resources, J.Z., E.R.M., R.K.W. and J.R.D.; Writing-Original Draft, I.A. and M.A.D.; Writing-Review and Editing I.A. and M.A.D.; Visualization, I.A., B.X., X.C., A.S., X.Z., and M.A.D.; Supervision, M.A.D.; Project Administration, M.A.D.; Funding Acquisition, M.A.D. and J.R.D.

[9]Department of Medicine, Washington University School of Medicine in St. Louis, St. Louis, Missouri 63108, USA

[10]Siteman Cancer Center, Washington University School of Medicine in St. Louis, St. Louis, Missouri 63108, USA

[11]Department of Pathology, St. Jude Children's Research Hospital, Memphis, Tennessee, 38105, USA

[12]Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, USA

SUMMARY

In the developing retina, multipotent neural progenitors undergo unidirectional differentiation in a precise spatiotemporal order. Here we profile the epigenetic and transcriptional changes that occur during retinogenesis in mice and humans. Although some progenitor genes and cell cycle genes were epigenetically silenced during retinogenesis, the most dramatic change was derepression of cell type–specific differentiation programs. We identified developmental stage–specific super-enhancers and showed that most epigenetic changes are conserved in humans and mice. To determine how the epigenome changes during tumorigenesis and reprogramming, we performed integrated epigenetic analysis of murine and human retinoblastomas and induced pluripotent stem cells (iPSCs) derived from murine rod photoreceptors. The retinoblastoma epigenome mapped to the developmental stage when retinal progenitors switch from neurogenic to a terminal patterns of cell division. The epigenome of retinoblastomas was more similar to that of normal retina than was that of retina-derived iPSCs, and we identified retina-specific epigenetic memory.

INTRODUCTION

Changes in gene-expression programs mark progression from proliferating multipotent progenitor cells to terminally differentiated neurons. Recent studies of neurogenesis of human and mouse cortices (Lister et al., 2013), mouse photoreceptors (Mo et al., 2016), other mature neuronal classes (Mo et al., 2015), and neurons produced from stem cells in organoid cultures (Ziller et al., 2015) shed light on the changes that occur in the epigenome with the combination of transcriptome analysis, DNA methylation, and (in some studies) histone modification. The cell type–specific epigenome of differentiated cells is thought to be relatively stable once established during development and is thought to be a major barrier to reprogramming differentiated cells, such as neurons, into induced pluripotent stem cells (iPSCs) (Orkin and Hochedlinger, 2011). For some cell types, the resulting iPSCs retain an epigenetic memory of their cellular origins (Hiler et al., 2015; Kim et al., 2010), which can influence subsequent lineage-specific differentiation.

Developmental changes in the epigenome are also central to human disease. For example, childhood cancers are developmental tumors that arise during crucial periods of development and genomic characterization of more than 2000 childhood cancers revealed that virtually every class of epigenetic regulator is mutated in developmental tumors (Huether et al., 2014). Neuroblastomas arise of the sympathoadrenal lineage (Cheung and Dyer, 2013); rhabdomyosarcomas emerge from the muscle lineage (Kashi et al., 2015); and

osteosarcomas form during the period of rapid bone growth in adolescence (Kansara et al., 2014). Genomic characterization of more than 2000 childhood cancers revealed that virtually every class of epigenetic regulator is mutated in developmental tumors (Huether et al., 2014). In this study, we performed a comprehensive analysis of the epigenomic and transcriptional changes that occur during retinogenesis and retinoblastoma in humans and mice, and iPSCs derived from murine rod photoreceptors to elucidate their epigenetic memory. We found that epigenetic changes play a more important role in activating differentiation genes than in silencing progenitor or proliferation genes during retinal maturation. Several retinal progenitor genes were sequestered in the domain of facultative heterochromatin (f-heterochromatin) in rod nuclei, suggesting an alternative mechanism of silencing developmental genes in neurons. Changes in the epigenome were evolutionarily conserved from mice to humans with retinoblastomas matching a narrow window of normal development consistent with their developmental origins. Finally, the genes most likely to be retained as epigenetic memory in iPSCs were not necessarily those that undergo the most dramatic epigenetic changes during differentiation. Together, these data show how a comprehensive profile of the changes in the epigenome during development can provide insight into the developmental stage–specific and cellular origins of pediatric cancer and the relations among the epigenomes of progenitors, stem cells, and cancer cells.

DNA-Methylation Changes Associated with Neurogenesis in the Retina

Previous studies have shown changes in DNA methylation that correlate with changes in gene expression in the developing CNS (Lister et al., 2013; Mo et al., 2015; Ziller et al., 2015). Here we extend those studies to retina, an ideal tissue in which to study the dynamics of the epigenome during development. Retina growth has been extensively characterized (Fig. 1B) (Young, 1984, 1985a, b), and the birth order and birth dates of the 7 classes of retinal cell types are evolutionarily conserved across vertebrate species (Fig. 1C) (Young, 1985a, b). To characterize the epigenetic landscape during mouse and human retinogenesis, we analyzed 8 developmental stages that span key developmental transitions (Fig. 1C) (Young, 1985a, b). To profile DNA methylation changes, we performed whole-genome bisulfite sequencing (WGBS) and RNA sequencing (RNA-seq) analyses for each stage of mouse [embryonic day (E) 14.5, E17.5, postnatal day (P) 0, P3, P7, P10, P14, and P21] and human [developmental week (FW) 10, 14, 17–21, and 23] retinal development and compared the changes in DNA methylation with those in gene expression. In the developing mouse retina, the DNA methyltransferases (Dnmt1, Dnmt3a, and Dnmt3b) have dynamic expression (Fig. S1), but there was no global change in DNA methylation during retinal development in mice or humans (Fig. S2). We identified 12% (473/3918) of genes with decreased expression and 26% (1143/4313) with increased expression during murine retinogenesis with significantly ($p<0.05$) inversely correlated changes in DNA methylation (Fig. 1D,E and Data S1, Table S1). Similarly, 9% (208/2407) of the genes had decreased expression and 13% (445/3330) had increased expression during human retinogenesis with significant ($p<0.05$) inversely correlated changes in DNA methylation (Fig. 1F,G and Data S1, Table S1). As expected, most downregulated genes encoded proteins that regulate proliferation, and most upregulated genes encoded proteins that contribute to neuronal differentiation (Data S1, Table S2).

To complement our unbiased analysis of the entire epigenome during retinogenesis, we performed a similar analysis of 4 validated retinal gene sets (Cherry et al., 2009; Macosko et al., 2015; Mo et al., 2016; Roesch et al., 2008; Roesch et al., 2012; Siegert et al., 2012; Trimarchi et al., 2008) which included those important for the G2/M transition in the cell cycle of proliferating retinal progenitors, expressed in retinal progenitors that are not cell cycle genes, expressed in rods during differentiation, and ubiquitously expressed housekeeping genes (Data S1, Table S3 and Fig. 2A–D). Early in development (E14.5), most cells are proliferating retinal progenitors (Young, 1985a, b), so their epigenetic profile should be apparent in embryonic retinae (Figs. 1C and 2A,B), while in late postnatal retinae (P10, P14, and P21), the majority (80%) of cells are rods, so their epigenetic profile should be apparent (Figs. 1C and 2C) in contrast to control housekeeping genes whose expression does not change (Fig. 2D).

DNA methylation and gene expression were inversely correlated over the course of murine retinal development in 3% (7/235) of the progenitor genes, 0% (0/43) of the G2/M genes, 38% (73/190) of the rod genes, and 4% (1/22) of the housekeeping genes (Data S1, Table S4). These trends were similar for other cell type–specific differentiation programs and in human retina (Data S1, Table S4) and show that changes in DNA methylation accompany critical changes in gene expression during retinogenesis, but only accounts for a small fraction (3%–38%) of the total number of developmentally regulated genes.

Developmental Transitions during Retinogenesis Are Reflected in the Epigenome

To determine if other epigenetic changes accompany the gene-expression changes during retinogenesis, we performed ChIP-seq of 8 histone marks (H3K4me1, H3K4me2, H3K4me3, H3K27me3, H3K27Ac, H3K36me3, H3K9/14Ac, and H3K9me3), as well as for CTCF, Brd4, and RNA PolII across all 8 developmental stages in mice. Due to the limited availability of human retinal tissue, we analyzed 4 developmental stages in human retina (FW13/14, FW15/16, FW18/20, and FW23/24) corresponding to E17.5, P0, P3, and P7, respectively, in the mouse (Fig. 1C) (Workman et al., 2013). Temporal changes in the epigenetic landscape of the promoter's proximal regions [± 2 kb from the transcription start site (TSS)] closely mirrored that of gene expression. For example, *Uhrf1* is a progenitor gene encoding an E3 ubiquitin ligase, which regulates gene expression in proliferating cells via epigenetic mechanisms that involve interactions with Dnmt1 (Berkyurek et al., 2014). During embryonic stages of retinal development, when the proportion of proliferating progenitors is highest, *Uhrf1* had H3K4me1/2/3, H3K27Ac, and H3K9/14Ac histone modifications spanning the promoter, H3K36me3 in the gene body and RNA PolII and Brd4 at the promoter (Fig. 2E), consistent with active transcription and the stages when *Uhrf1* expression is highest (Fig. 2E). As development progressed and *Uhrf1* expression was extinguished, the H3K36me3 modification was lost from the gene body; PolII, Brd4, and H3 acetylation were lost at the promoter; and H3K4me1/2/3 modifications were reduced (Fig. 2E). Repressive histone marks (H3K27me3 and H3K9me3) and DNA methylation were absent from the TSS throughout development, suggesting that *Uhrf1* is not actively silenced by H3K27me3, H3K9me3, or DNA methylation in mature retinal neurons (Fig. 2E).

In contrast to *Uhrf1*, some progenitor genes showed accumulation of the H3K27me3 repressive mark as they were silenced during retinal differentiation. Ascl1 is a proneural bHLH protein expressed in retinal progenitors during murine development (Brzezinski et al., 2011). The pattern of histone marks and PolII/Brd4 binding for the *Ascl1* gene during early development (E17.5) is similar to *Uhrf1* (Fig. 2F). However, when retinal progenitors are depleted (~P7), *Ascl1* expression was silenced and H3K27me3 accumulated at the promoter and throughout the gene body (Fig. 2F). The promoter and gene body DNA remained hypomethylated in the differentiated retina.

Rod-specific genes that are transcriptionally activated during later stages of development as the cells differentiate show the opposite pattern of epigenomic changes. For example, the *Stx3* gene encodes a protein that is required for rod synaptic transmission (Mazelova et al., 2009) and was marked by insulator protein CTCF ChIP-seq peaks from early in development (Fig. 2G). There were H3K4me1/2/3 marks and H3 acetylation (H3K27Ac and H3K9/14Ac) marks at the promoter but no Brd4 or PolII present between E14.5 and P0 (Fig. 2G). The colocalization of H3K27me3 and H3K4me3/H3K27Ac at the *Stx3* promoter during early development suggested that it is a bivalent or poised promoter (Fig. 2G) (Voigt et al., 2013). Consistent with the lack of expression and Brd4/PolII binding early during development, there was no H3K36me3 in the gene body. As rods differentiated and gene transcription increased, H3K27me3 was depleted at the promoter; PolII and Brd4 increased; H3K36me3 accumulated across the gene body and the active histone marks (H3K4me1/2/3, H3K27Ac, and H3K9/14Ac) spread 3′ into the body of the gene and a corresponding reduction in DNA methylation in this region downstream of the TSS (Fig. 2G).

Mapping Chromatin-State Dynamics during Retinal Development

To define the chromatin states and more efficiently analyze the transitions thereof across the genome, we performed chromatin Hidden Markov Modeling (chromHMM) using all 176 murine ChIP-seq datasets (Ernst and Kellis, 2012). ChromHMM enables systematic annotation of the epigenetic states across the human and mouse genomes during retinogenesis. Based on computational methods and manual review, we selected 11 chromHMM chromatin states for analysis (Fig. 3A). State 1 had epigenetic marks consistent with actively transcribed genes; states 2 and 3 were predominantly enhancers; and state 4 identified bivalent promoters (Voigt et al., 2013). State 5 was primarily defined by PolII binding, and states 6 and 7 were consistent with gene bodies (H3K36me3). State 8 represented polycomb-repressed chromatin (H3K27me3) outside of the promoter or enhancers; state 9 was empty chromatin; and state 10 was the H3K9me3-repressed chromatin. Finally, state 11 was marked by the insulator protein CTCF.

The transitions in chromHMM state at progenitor, G2/M, rod, and housekeeping genes across development were consistent with changes in gene expression and individual ChIP-seq data (Figs. 3B and S4A,B). Overall, the active promoter/enhancer marks (states 1–3) were increased for rod genes and decreased for progenitor and G2/M genes during development; states 1–3 were unchanged for the housekeeping genes (Fig. S4C–F) and was conserved across human and mouse retinal development (Figs. 3C,D, S4E–H, and Data S1, Table S5). Analysis of the chromHMM-state transition for the 133 rod genes with the most

robust changes in gene expression during retinal development (Data S1, Table S3, cluster 1) identified 26 (20%) genes that had H3K27me3 marks (states 4 and 8) early in development and then transitioned to active promoters (states 1–3) as the rods differentiated (Data S1, Table S6 and Fig. S5A–C). We identified 19% (33/174) of the progenitor genes with the most robust changes (Data S1, Table S3, cluster 1) that had the opposite pattern (Data S1, Table S6 and Fig. S5D,E). Similar trends were observed for human retina (Data S1, Table S6 and Fig. S5F–I).

Beyond those select rod and progenitor genes, an additional 584 genes had chromHMM-state transitions consistent with H3K27me3 derepression (loss of H3K27me3), and 486 had chromHMM-state transitions consistent with H3K27me3 repression (gain of H3K27me3) during murine retinogenesis with corresponding changes in gene expression (Data S1, Table S6). Similar trends were observed in human retina (Data S1, Table S6), with fewer genes due to the limited retinal stages available for ChIP-seq.

To extend our analysis, we performed hierarchical clustering of all the chromHMM data integrated with gene-expression data. We identified 4 groups of genes with developmental chromHMM and expression patterns similar to those of the rod gene set described above (Fig. 3E). There were 314 genes that showed chromHMM-state transitions during development that were consistent with low or no expression early in development to very high expression in the mature retina, and an additional 206 genes that showed chromHMM-state transitions consistent with low or no expression early during development to intermediate expression during retinal differentiation (Fig. 3E and Data S1, Table S7). We identified 756 genes that had developmental chromHMM and gene-expression patterns similar to those of the retinal progenitor and G2/M genes (Fig. 3F and Data S1, Table S7).

Of the 4313 genes that increased expression during retinal development (Fig. 1E and Data S1, Table S1), 1143 had changes in DNA methylation, 582 had changes in H3K27me3, and 520 had changes in epigenetic state, as measured by chromHMM modeling. In total, 40% (1734/4313) of upregulated genes had at least 1 type of correlated epigenetic change (Fig. 3G and Data S1, Table S7). Of the 3918 genes that decreased expression during retinal development (Fig. 1D and Data S1, Table S1), 472 had changes in DNA methylation, 484 had changes in H3K27me3, and 756 had changes in epigenetic state, as measured by chromHMM modeling. In total, 36% (1402/3918) of downregulated genes had at least 1 type of correlated epigenetic change (Fig. 3H and Data S1, Table S7).

To explore the relations among DNA methylation, gene expression, and chromHMM state during development, we calculated the proportion of each chromHMM state for the genes that are upregulated, hypomethylated, and simultaneously increase expression and downregulate DNA methylation (Fig. S6 and Data S1, Table S8). We performed a similar analysis for the downregulated genes (Fig. S6 and Data S1, Table S8). Overall, the genes that change their expression and DNA methylation have more pronounced differences in their distribution of active and inactive chromHMM states (Fig. S6 and Data S1, Table S8).

Cell Type–Specific Regulation of Epigenetic State in the Retina

The adult retina is made up of 80% rods, and the other cell types account for 0.2%–7% of the tissue. Genes expressed in other cell types may be epigenetically repressed in rods. In our analysis, such genes would increase expression during development but be positively correlated with changes in H3K27me3 and/or chromHMM-repressive states since RNA-seq analysis is more sensitive than ChIP-seq analysis. Therefore, the increase in gene expression that accompanies differentiation could be detected from the less abundant cell types (cones, amacrine, horizontal, bipolar, ganglion, and Müller glia), but the epigenetic profile would be dominated by the rods. To test this, we analyzed the chromHMM and gene-expression data for 136 cone genes, 154 bipolar genes, 65 Müller glial genes, 486 amacrine cell genes, and 131 ganglion cell genes (Data S1, Table S9) (Blackshaw et al., 2001; Blackshaw et al., 2004; Blackshaw et al., 2003; Cherry et al., 2009; Macosko et al., 2015; Mo et al., 2016; Roesch et al., 2008; Roesch et al., 2012; Siegert et al., 2012; Trimarchi et al., 2008). Several of the cell type–specific genes were polycomb repressed in the mature retina (Data S1, Table S9 and Fig. S7). To directly test whether polycomb repression occurred predominantly in rods, we purified those cells at P6/P7 from *Nrl-GFP* transgenic mice (Akimoto et al., 2006) by fluorescence-activated cell sorting and performed RNA-seq and ChIP-seq for H3K27me3 and H3K4me3. 35% (48/136) of the cone genes, 49% (76/154) of the bipolar cell genes, 26% (17/65) of the Müller glia genes, 61% (300/486) of the amacrine cell genes, and 60% (79/131) of the ganglion cell genes had H3K27me3 at their promoters (Data S1, Table S9 and Fig. S7A–D), with similar results in the human retina (Fig. S7E–H and Data S1, Table S9). Moreover, among the 4313 genes with increased expression during development, 1567 were not expressed in the purified rods (FPKM <1 cutoff). Among those genes, 66% (1031/1567) had H3K27me3 in the purified rods, suggesting that they were found in other cell types but were repressed in developing rods (Data S1, Table S9).

Super-Enhancers Are Associated with Important Regulators of Retinogenesis

The mammalian genome is organized into discrete topologically associated domains (TADs), which are stretches of DNA (~1 Mb) that are delimited by sharp boundaries comprising CTCF-insulator sites and constitutively expressed housekeeping genes (Ciabrelli and Cavalli, 2015; Sexton and Cavalli, 2015). TADs are thought to be constant across cell types, but the chromatin landscape within individual TADs varies (Dixon et al., 2012). Individual enhancers and large clusters of enhancers called super-enhancers activate individual genes within TADs by looping-out the intervening DNA (Hnisz et al., 2013). To map each enhancer and super-enhancer to their developmentally regulated genes in individual TADs, we used previously published data (Hi-C Data Browser) for mouse cortex (Dixon et al., 2012). Using our chromHMM data, we classified 1516 TADs in the mouse genome as active, polycomb repressed, null chromatin, or constitutive heterochromatin at each stage of mouse retinogenesis (Data S1, Table S10).

None of the TADs classified as null at E14.5 (228 TADs) or P21 (228 TADs) had any transitions in their classification based on TAD-delineated chromHMM states during development. Similarly, relatively few TADs classified as constitutive heterochromatin at E14.5 (4/68; 5%) or P21 (9/73; 12%) underwent large-scale chromHMM-state transitions during development. Consistent with analysis of individual genes within the genome, the

largest transitions in TAD states were in the active and polycomb-repressed TADs. Of the TADs that were active at E14.5, 22% (17/76) transitioned to repressed chromatin (heterochromatin or polycomb) by P21, and 10% that were active at P21 had been polycomb repressed at E14.5 (Data S1, Table S10).

Next, we identified the putative poised and active enhancers at each stage of retinogenesis in humans and mice by using our ChIP-seq data for H3K4me1, H3K27Ac, and H3K27me3 (Heintzman et al., 2009; Zentner et al., 2011). In addition to the chromatin marks, enhancers have a more open chromatin configuration, as measured by DNAse hypersensitivity or ATAC-seq (Thurman et al., 2012). To identify regions of open chromatin in the developing mouse retina, we performed ATAC-seq on the 8 stages of murine retinal development in biological duplicates (Fig. S8). We annotated 3 classes of enhancers in our data based on previous studies (Zentner et al., 2011). Poised enhancers were at least 500 bp in size, had ATAC-seq peaks and H3K4me1 but lacked H3K27Ac and H3K27me3. *Polycomb*-poised enhancers were at least 500 bp in size, had ATAC-seq peaks, H3K4me1, and H3K27me3 but lacked H3K27Ac. Active enhancers were also at least 500 bp in size and had ATAC-seq peaks, H3K4me1, and H3K27Ac. To avoid overlap with promoters, our enhancer annotation focused on those regions that did not overlap with gene bodies or were within 2 kb of the TSS. We found dynamic enhancer profiles. For example, 128 active enhancers at E14.5 transitioned to poised enhancers at P21, and 46 active enhancers at E14.5 transitioned to polycomb-poised enhancers at P21 (Data S1, Table S11). Similarly, 208 poised enhancers and 29 polycomb-poised enhancers at E14.5 transitioned to active enhancers at P21 (Data S1, Table S11).

Looping within TAD domains, which is mediated by CTCF and cohesion, modulates interactions between enhancers and the TSS for developmentally regulated genes (Ciabrelli and Cavalli, 2015). For this analysis, we focused on super-enhancers defined by clusters of H3K27Ac ChIP-seq peaks (Hnisz et al., 2013). We further refined our classification of super-enhancers by the absence of H3K4me3 typically found at promoters and identified 1768 super-enhancers that were present in at least 2 stages of development (Data S1, Table S12). Among those super-enhancers, 26% (454/1761) increased with retinal differentiation, and 8% (153/1761) decreased with depletion of retinal progenitor cells (Fig. 4A,B and Data S1, Table S12).

Next, we determined whether the distance from each super-enhancer was correlated with the change in gene expression for rod genes, progenitor genes, G2/M genes, and housekeeping genes. The majority of super-enhancers that showed developmental downregulation (75%; 80/107) or upregulation (89%; 231/260) were also within 100 kb (Fig. 4C and Data S1, Table S12), though some downregulated super-enhancers were much farther away (400–500 kb) (Fig. 4D and Data S1, Table S12). Several well-known G2/M and retinal progenitor cell genes (i.e., *Cdk1*, *Top2a*, *Fgf15*, *Ube2c*, *Uhrf1*, and *Ascl1*) were associated with developmentally downregulated super-enhancers (Fig. 4C) while several rod genes (e.g., *Pde6a*, *Guca1b*, *Guca1a*, *Nr2e3*, *Rbp3*, and *Crx*) were associated with developmentally upregulated super-enhancers (Fig. 4D). Among the 82 genes upregulated during development and with corresponding changes in DNA methylation, H3K27me3, and chromHMM state (Fig. 3G), 37 had a corresponding developmentally regulated super-

enhancer (Data S1, Table S7 and Fig. S9A,B). Of the 40 genes downregulated during development (Fig. 3H), 11 had corresponding superenhancers (Data S1, Table S7 and Fig. S9C,D).

We performed a similar analysis using the human retina chromHMM data and identified 938 super-enhancers (Data S1, Table S12), among which 92% (861) overlapped with super-enhancers in the mouse retina, and 420 of those were developmental super-enhancers. When we restricted the super-enhancer analysis to the 686 that were identified using the ROSE algorithm (Whyte et al., 2013), 62% (424/686) were conserved across species (Data S1, Table S12).

To determine if the putative super-enhancers are sufficient to increase expression from a minimal promoter in the developing mouse retina, we cloned a portion of the rod-specific super-enhancers for *Rbp3*, *Tulp1*, *Aipl1*, and *Guca1a* upstream of a minimal promoter and the *Nano-Glo* reporter gene. The cloned region of genome lacking H3K27Ac was used as a negative control. P0 retinae were square-wave electroporated with each plasmid containing the putative super-enhancer element, *Nano-Glo* and a constitutive CMV renilla luciferase reporter plasmid to control for electroporation efficiency. After 5 days in culture to allow time for rod specification, each super-enhancer element had substantial activity in this assay (Fig. 4E,F), including the previously reported *cis*-enhancer elements for *Otx2* and *Vsx2* (Kurokawa et al., 2004).

Super-enhancers are also useful for identifying core transcription factors that contribute to core regulatory circuits (CRCs) (Saint-Andre et al., 2016). Specifically, genes encoding transcription factors that form CRCs often have cell type–specific super-enhancers that contribute to a positive feedback autoregulatory loop (Saint-Andre et al., 2016). For example, the *Otx2* super-enhancer has an Otx2-binding site (Kurokawa et al., 2004) and the ATAC-seq peak of the Crx super-enhancer has a Crx-binding site (Fig. 4G). Using the CRC mapper (Saint-Andre et al., 2016) with our ChIP-seq data, we identified several core transcription factors, such as Vsx2, that are important in retinal progenitors or for photoreceptor differentiation (Otx2 and Crx) during retinal development (Fig. 4H and Data S1, Table S13)(Burmeister et al., 1996; Furukawa et al., 1997; Kurokawa et al., 2004). We also identified several core transcription factors in the CRC analysis that have not been extensively studied in the developing retina. For example, Esrrb is a nuclear orphan receptor that regulates metabolic genes important for photoreceptor homeostasis according to the circadian rhythm(Kunst et al., 2015). This is intriguing because another gene from the CRC analysis was a master circadian regulator called Bhlhe41 (Data S1, Table S13) (Rohleder et al., 2006). The expression of Esrrb and Bhlhe41 increases with photoreceptor differentiation, and their CRCs were identified in mature retina (Data S1, Table S13).

The Retinal Epigenome Is Stable with Aging

We performed the same integrated ChIP-seq and chromHMM analyses in 11-month-old C57Bl/6 mouse retina (Fig. S10). Overall, the retinal epigenome was stable, and a detailed analysis of the chromHMM showed that the biggest difference in the retinal epigenome of aged retina was the expansion of chromHMM state 1, which was found at strong promoters (Fig. S10A). Genes associated with the expansion of state 1 were significantly more likely to

be associated with developmental super-enhancers ($p < 0.05$; Fig. S10B,C), suggesting that the long-distance interactions between super-enhancers and promoters can expand the active chromatin domain with aging in the retina. The genes and pathways associated with a spreading chromHMM state 1 were significantly enriched in photoreceptor differentiation ($p < 0.05$; Data S1, Table S14).

Three-Dimensional Mapping of the Retinal Epigenome

Previous studies have shown that during development, rods dramatically condense their chromatin into a central core and 2 concentric spheres (Solovei et al., 2009). This unusual organization for mammalian nuclei is thought to enhance light transfer through the cellular layers of the retina (Solovei et al., 2009). The central core contains a dense region of constitutive heterochromatin surrounded by a more diffuse region of noncentromeric f-heterochromatin (Solovei et al., 2009). Euchromatin is a thin shell that occupies the outer surface of the rod nuclei. These 3 regions can be distinguished using fluorescence in situ hybridization (FISH) for repetitive sequences characteristic of the C, G, and R bands of mouse chromosomes, immunofluorescence for histone marks, DAPI staining, and electron microscopy (Fig. 5A–D) (Solovei et al., 2009). To measure the volume of the 3 concentric regions of the nucleus, we performed 3D electron microscopy (3D EM) using the Helios dual-beam system. Different chromatin domains were segmented using an algorithm designed to identify regional image features, and the volume of each region was calculated (Fig. 5E).

To determine if there are differences in the 3D localization of genes within the nucleus based on their 2D epigenetic-state transitions during development (DNA methylation, H3K27me3, and chromHMM), we selected 22 genomic regions spanning genes with different epigenetic profiles during retinogenesis. We performed FISH with each probe and immunostained for H3K4me3 on the same tissue sections to mark the euchromatin domain (Fig. 5F). We obtained confocal z-series and scored the colocalization manually and by using a customized automated segmentation and colocalization algorithm (Fig. 5G,H). As expected, rod genes (*Crx*, *Nrl, Nr2e3*, and *Rho*) were localized to the euchromatin domain (Fig. 5H, I and data not shown). Several genes that are expressed in other cell types but are H3K27me3 repressed in rods (*Gad1*, *Pax6*, *Lhx3*, *Rlbp1*, *Grm6*, *Ccnd3*, and *Arr3*) were also localized to the euchromatin domain (Fig. 5I). However, 3 of the 4 progenitor genes (*Ascl1*, *Sfrp2*, *Myc*) that are H3K27me3-repressed in rods were found in the f-heterochromatin domain (Fig. 5I,J), as was the *Cdkn2a* tumor-suppressor gene, the *Car10* bipolar cell–specific gene, and the *Cav1* Müller glia–specific gene (Fig. 5I). The subcellular localization (euchromatin or f-heterochromatin) was consistent across nuclei for each gene, suggesting that the 3D organization of the retinal epigenome is as precisely regulated as the 2D organization of histone marks and DNA methylation at individual genetic loci. Those genes that were primarily localized to the f-heterochromatin but were at the boundary between domains (*Arr3* and *Sfrp2*) had a higher percentage of false-positive calls in euchromatin using the automated algorithm. Genes that are not expressed in the developing retina but are found in other neuronal lineages, such as those restricted to hair cells in the inner ear, were often found in the euchromatin domain (Fig. S11).

Retinal Epigenetic Memory

To determine whether the genes with the most dramatic changes in their epigenetic state during rod differentiation are the ones most likely to retain epigenetic memory in iPSCs derived from rods (r-iPSCs), we profiled a series of r-iPSCs (Hiler et al., 2015; Hiler et al., 2016). Specifically, we developed a new approach for reprogramming differentiated rod photoreceptors and showed that r-iPSCs are better at making retina in 3D organoid cultures using the STEM-RET protocol (Hiler et al., 2015; Hiler et al., 2016) using EB5:Rx-GFP ESC line as our control, which were the source for fibroblast-derived iPSC (f-iPSC) lines (Hiler et al., 2015; Hiler et al., 2016). All iPSC lines used in this study were passaged at least 20 times and were fully reprogrammed pluripotent cells. We performed the same analyses (RNA-seq, WGBS, ChIP-seq, chromHMM), as described above for the developing mouse retina.

The iPSC lines were similar in their gene expression (Fig. 6A), which is consistent with data showing that they are fully reprogrammed pluripotent stem cells (Hiler et al., 2015). In principle component analysis of the RNA-seq data, the r-iPSCs were more similar to the EB5:Rx-GFP line than to the f-iPSCs, which is consistent with their efficient formation of retina in STEM-RET (Fig. 6A). There were 444 genes with significantly elevated expression, and 436 genes with significantly decreased expression (2.0-fold cutoff; $p \leq 0.05$) in r-iPSCs relative to f-iPSCs (Data S1, Table S15). The significantly upregulated genes were associated with neuronal differentiation, which is consistent with their retinal origins (Data S1, Table S15). In r-IPSCs, 38% (171/444) of the upregulated genes were expressed in the retina and increased their expression during retinal differentiation; 26% (46/171) of those genes exhibited hypomethylation during retinogenesis; 18% (31/171) showed decreasing H3K27me3; and 5 genes had developmental super-enhancers (Data S1, Table S15).

Analysis of the DNA-methylation data identified 2115 genes that were significantly hypermethylated ($p<0.05$) and 583 genes that were significantly hypomethylated ($p<0.05$) in f-iPSCs relative to r-iPSCs (Fig. 6D and Data S1, Table S16). Pathway analysis of the hypermethylated genes in f-iPSCs identified pathways involved in eye development, photoreceptor outer segment formation, and neurogenesis; that of hypomethylated genes in f-iPSCs identified genes associated with actin cytoskeleton, fibroblast migration, and intracellular signaling (Data S1, Table S16). Several genes with important roles in photoreceptor development and homeostasis, such as *Rcvrn*, *Sag*, *Pde6h*, and *Myo7a*, had differential DNA methylation in r-iPSCs (Fig. 6D and Data S1, Table S16). However, only 7.4% (156/2115) of the hypermethylated genes in f-iPSCs were upregulated during retinal development, and 1.7% (10/583) of the hypomethylated genes in f-iPSCs were downregulated during retinal development with corresponding DNA hypermethylation (Data S1, Table S16).

To determine whether there were any significant differences in the chromHMM between f-iPSCs and r-iPSCs, we analyzed each of the 11 states individually. For the chromHMM states spanning genes or promoters that were significantly increased ($p<0.05$) in f-iPSCs relative to r-iPSCs, the largest changes were in the states with H3K27me3 (states 4 and 8) and the empty state (state 9) (Fig. 6E and Tables S17, S18). For the genes and promoters significantly decreased ($p<0.05$) in f-iPSCs relative to r-iPSCs, the largest changes were in

weak enhancers (state 3), the H3K9me3-repressed state (state 10), and the empty state (state 9) (Fig. 6F and Tables S17, S18).

Pathway analysis showed substantial enrichment in neuronal development, axon morphogenesis, and synaptogenesis for the genes that had decreased H3K9me3 (state 10) and decreased H3K27me3 (states 4 and 8) (Data S1, Table S19). Examples of genes with differential H3K9me3 included *Prkca*, *Rims1*, *Gng4*, *Kcnma1*, *Cacna2d3*, *Adcy2*, and *Ntrk2* (Data S1, Table S18). Examples of genes with differential H3K27me3 included *Vax2*, *Axin1*, *Trpc6*, *Gabrb2*, *Gabra3*, and *Robo2* (Data S1, Table S18). Promoters with differential active states containing H3K4me3 (state 1) were much less common (Fig. 6E–G).

To further refine the differences in the epigenetic landscape in r-iPSCs versus f-iPSCs, we analyzed the individual ChIP-seq data. Consistent with the chromHMM analysis showing that the repressive epigenetic states were among the most different between r-iPSCs and f-iPSCs, the most differential individual epigenetic marks were H3K9me3 and H3K27me3 (Fig. 6H and Data S1, Table S20). The number of genes with differential H3K9me3 and H3K27me3 (6984) was more than 10-fold higher than that of genes with all active marks combined (594) (Fig. 6H,I and Data S1, Table S20). Pathway analysis of the genes with reduced H3K9me3 in the promoter and gene bodies in f-iPSCs relative to r-iPSCs were enriched for nervous system development, neuron differentiation, axon formation, and synaptogenesis (Data S1, Table S21). Analysis of the promoters of the 1103 cell type–specific genes (Data S1, Table S9) identified 41 (4%) genes with differential H3K9me3 or H3K27me3 at their promoters (Data S1, Table S22). Together, this suggests that DNA methylation plays a relatively minor role in contributing to the epigenetic memory of iPSCs derived from neurons, and repressive histone marks are a more prominent feature of epigenetic memory than previously appreciated. Moreover, the genes that undergo the most dramatic epigenetic changes as rods differentiate are more likely to be reset to the embryonic state than to retain epigenetic memory.

Epigenomic Mapping of Retinoblastoma

Retinoblastoma is a childhood cancer of the developing retina that begins in utero and is diagnosed in the first few years of life (Rodriguez-Galindo et al., 2015). We analyzed the gene expression and epigenetic landscape of human and mouse retinoblastomas and compared those data to normal developing retinae from both species. For human retinoblastoma, we used orthotopic patient-derived xenografts (Zhang et al., 2012) and for mouse retinoblastoma, we used a GEMM (*Chx10-Cre;*$Rb^{Lox/Lox}$*;*$p107^{-/-}$$MDMX^{Tg}$) (McEvoy et al., 2011a), the most accurate preclinical model of the human disease among the 6 GEMMs of retinoblastoma available (McEvoy et al., 2011a). Murine retinoblastomas have chromHMM and express G2/M genes matching E14.5–E17.5 retina, and rod and progenitor genes best matching P0–P3 retina (Fig. 7A–F and Data S1, Table S23). FW23/24 was the best match between human retinal development and human retinoblastoma (Data S1, Table S23). Unlike murine retinoblastoma, the correlation between human retinoblastoma and normal human retina for rod genes increased with developmental stage (Data S1, Table S23).

Mouse and human retinoblastomas have molecular features of several different cell types; human retinoblastoma have more photoreceptor features while murine retinoblastoma have

more interneuron features (Mcevoy et al., 2011b). To test this more directly, we identified the genes that were upregulated or downregulated in human and mouse retinoblastomas relative to those in normal retina and had corresponding changes in their epigenomic landscape (Data S1, Table S24). Pathway analysis showed that the photoreceptor genes were significantly downregulated ($p < 0.05$) in murine retinoblastoma relative to later developmental stages, and this was consistent with their interneuron features (Data S1, Table S25). The significantly upregulated ($p < 0.05$) pathways in murine retinoblastoma were all involved in proliferation (Data S1, Table S25). Human retinoblastomas express more photoreceptor genes than do mouse retinoblastomas, and the pathway analysis of differentially expressed genes for human retinoblastoma showed enrichment of synaptic proteins in the inner retina (Data S1, Table S25). Pathways involved in proliferation were upregulated in human retinoblastoma, but the most common upregulated pathways were those involved in RNA processing and metabolism (Data S1, Table S25).

The transcriptional network for photoreceptor maturation has been extensively studied and has led to a hierarchical transcriptional model (Swaroop et al., 2010) (Fig. 7G). We analyzed the chromHMM states and gene expression of human and mouse retinoblastomas for each of the key regulatory genes in the pathway (Figs. 7G and S12). In human retinoblastoma, the transcriptional factor genes that specify photoreceptor precursors (*CRX* and *OTX2*) were both expressed at higher levels in retinoblastoma than in normal retina but had epigenetic states matching normal retina. Genes encoding transcription factors that promote rod differentiation (*NRL* and *NR2E3*) also showed expression and epigenetic states consistent with specification toward this lineage (Fig. 7G). Downstream rod-differentiation genes were not expressed and had inactive epigenetic profiles. There was also evidence of cone specification based on *THRB* expression, but there was no expression of *RXRG* or *ARR3* further downstream in the cone-differentiation pathway (Fig. 7G and data not shown). The S-cone fate is the default state for cone differentiation, and in our analysis, *OPN1SW* was expressed but not *OPN1MW* or *OPN1LW* (Fig. 7G and data not shown). In mouse retinoblastoma, there was little expression of photoreceptor lineage genes (Fig. S12), which was consistent with their progenitor/interneuron phenotype (McEvoy et al., 2011a).

To determine if the upregulated and downregulated genes in retinoblastoma were more likely to be those with dynamic DNA methylation, H3K27me3, chromHMM, or adjacent to a developmental super-enhancer, we overlaid the mouse and human tumor data with normal developmental profiles. Among the 470 genes that were downregulated in the mouse tumor relative to normal retinal development, 76% (357/470) were developmentally upregulated during retinogenesis (Data S1, Table S26 and Fig. S13A,B). Among those genes, 78% (278/357) were epigenetically regulated during development (DNA methylation, chromHMM, or H3K27me3) (Fig. S13C). Only 5 genes overlapped with those that were downregulated during murine retinogenesis. Among the 783 genes that were upregulated in murine retinoblastoma relative to normal retina, 70% (553/783) were normally downregulated during retinogenesis (Data S1, Table S26 and Fig. S13D,E). Among those, 51% (284/553) were epigenetically regulated during development (Fig. S13F). Similar trends were observed for human retinoblastoma (data not shown). Together, this suggests that epigenetically regulated genes during retinal development are most likely to be perturbed in retinoblastoma.

Genes that do not overlap with retinal development were enriched in nuclear transport and epigenetics (upregulated) and metabolism (downregulated) (Data S1, Table S27). More importantly, the epigenetic landscape of murine and human retinoblastoma matched the key transition from proliferative to a neurogenic pattern of cell division, and the human retinoblastomas initiated later during retinogenesis than did the murine retinoblastoma. This was reflected in the poised and active enhancers (Data S1, Table S11) and the CRC analysis (see Data S1, Table S13). Among the 47 core transcription factors identified in the murine CRC analysis, 4 were found only in retinoblastoma (Tcf4, Hif1a, Foxj1, and Egr1). Tcf4 is upregulated 5 fold in retinoblastoma relative to normal retina and is an important regulator of Wnt/β-catenin signaling in cancer (Kypta and Waxman, 2012). Foxj1 is thought to be a key transcriptional regulator of genes that encode proteins implicated in ciliogenesis (Yu et al., 2008). Egr1 is a tumor suppressor (Zwang et al., 2012), and Hif1a is a master regulator of the response to hypoxia often induced in cancers (Denko, 2008).

DISCUSSION

During development, multipotent neural progenitors undergo unidirectional changes in competence to produce differentiated cells in a precise temporal and spatial order (Cepko, 2014). This is required to produce each class of neurons, in the correct place and time, to form neural circuits. Alterations in the balance between proliferation and differentiation in the developing CNS can cause evolutionary change in regions such as the cortex and retina (Dyer, 2016b; Dyer et al., 2009; Otani et al., 2016) or can lead to retinoblastoma (Macpherson and Dyer, 2007). Here we focused on the transcriptional silencing of retinal progenitor cell genes (including those required for proliferation) and the transcriptional activation of differentiation genes in rod photoreceptors and other retinal cell types. Changes in histone modifications at developmentally regulated genes were more prevalent than those in DNA methylation and were more predictive of changes in gene expression. Among the histone modifications, changes in H3K27me3 were the most strongly associated with developmental stage–specific transitions in gene expression. Changes in H3K27me3 and DNA methylation were more prevalent in differentiation genes than in progenitor genes. We used our epigenomic data to identify novel evolutionarily conserved developmental stage–specific super-enhancers and mapped their putative target genes within TADs and the 3D nuclear organization of rods.

Using these data as a backdrop, we mapped the developmental origins of human and mouse retinoblastomas and found they most closely resembled the stage when retinal progenitors transition from neurogenic to terminal patterns of division. The murine retinoblastoma epigenome was more similar to normal developing retina than to stem cells derived from retinal neurons, and epigenetic profiling of r-iPSCs relative to f-iPSCs showed that histone modifications are central to neuronal epigenetic memory.

Dynamic Transitions during Development Are Reflected in the Epigenome

As murine retinal progenitors change competence and exit the cell cycle during retinogenesis, several thousand genes are downregulated, with about one-third showing corresponding changes in their epigenetic state. Most genes transitioned to an empty state

suggesting they are either truly devoid of specific epigenetic marks, or other marks that were not included in our analysis contributed to the repression of proliferation and retinal progenitor cell–specific genes. One intriguing possibility is that the empty genes are sequestered in less accessible regions of the nucleus (heterochromatin) that do not require active epigenetic repression (Solovei et al., 2009). Genes that are not expressed but are localized in regions of the nucleus with active transcription may require epigenetic repression such as H3K27me3. It has been suggested that all genes are localized to the euchromatin domain in rods, irrespective of their expression or epigenetic state (Solovei et al., 2009). However, our data showed that a subset of developmentally regulated retinal genes are sequestered to the f-heterochromatin ring surrounding the constitutive heterochromatin in the center of rod nuclei.

Our 1D epigenomic map may serve as a valuable resource for elucidating the rules that govern 2D looping within TADs and 3D nuclear organization in this class of neurons in future studies. Our data suggest that the localization of genes to heterochromatin or euchromatin domains is cell type specific. For example, *Car10*, *Cav1*, and *Arr3* are localized to the heterochromatin domain in rods but are expressed from the euchromatin domain in other retinal cell types. Moreover, the cell type–specific nuclear localization may extend to genes involved in disease; the *Myc* oncogene and *Cdkn2a* tumor-suppressor gene are both localized to the heterochromatin domain in rods. We have proposed that the 1D, 2D, and 3D organization of the genome during differentiation, a property that we call cellular pliancy (Chen et al., 2015; Dyer, 2016a), contributes to cell type–specific susceptibility to disease. In this model, rods have low pliancy because more of their genome is epigenetically inactive at the 1D, 2D, or 3D level of organization, relative to interneurons such as horizontal cells that have a much more accessible genome (Chen et al., 2015; Dyer, 2016a). Although our data are consistent with the cellular pliancy model of disease susceptibility, a more comprehensive profiling of individual cell types is required to elucidate the underlying molecular mechanisms.

As retinal neurons and glia differentiated during development, the expression of several thousand genes increased. Approximately half of those genes showed a corresponding change in at least 1 epigenetic parameter measured, which is probably an underestimate because of the mixed cell types in the adult retina. We discovered cell type–specific H3K27me3 repression of non-rod genes in rods, suggesting that the dynamics of H3K27me3 extends to cell type–specific programs. The portion of differentiation genes with corresponding epigenetic changes may exceed 50% based on estimates from the rod genes. Indeed, a major limitation of our study is the heterogeneity of the samples analyzed. Not only are the 7 major classes of retinal cell types present in different proportions in the adult retina, but also they are born in an evolutionarily conserved order, in a center-to-periphery gradient. Therefore, at any given stage of development (E14.5–P10), the tissue analyzed may be a mixture of proliferating retinal progenitor cells, newly postmitotic committed cells, and differentiating cells. Developmental analysis of individual cell types will be required to elucidate the key epigenetic programs within each cell type.

The epigenetic data were useful for identifying *cis*-regulatory elements, in particular super-enhancers, for developmentally regulated genes. Identification of stage-specific formation of

super-enhancers and genes within the same TAD that are up or downregulated with the same kinetics, provided insight into the complex regulation of retinogenesis. Moreover, the CRC analysis identified master transcription factors that are central to mouse and human retinal development and retinoblastoma. It will be important to validate the role of those transcription factors and the function of their associated super-enhancers in vivo.

Epigenomic Changes Are Conserved across Species

Epigenetic changes were strongly conserved across mouse and human retinal development with DNA-methylation changes being the least conserved (7%–8%); chromHMM more highly conserved (45%–56%); and the developmental stage–specific super-enhancers the most highly conserved (62%). This may be an underrepresentation because we had access to fewer developmental stages for human retinal development. Therefore, some genes may have changes in DNA methylation that accompany their dynamic changes in gene expression but were missed because the changes occurred later in development than we analyzed in our study. Also, species-specific difference may occur in the epigenomic regulation of retinogenesis based on the organization of the genome. Changes in histone modifications were more predictive of changes in gene expression across species. For example, nearly 90% of the genes with changes in H3K27me3 during development also had corresponding changes in their expression.

Relating the iPSC Epigenome to Retinal Development

There was no evidence to support our initial hypothesis that the genes with the most dynamic changes in their epigenetic landscape were most likely to retain epigenetic memory in r-iPSCs. At every level of analysis, there was more evidence of epigenetic memory in r-iPSCs than in f-iPSCs with significant enrichment in genes important for neurogenesis. In addition, complementing the DNA-methylation analysis with ChIP-seq analysis provided a more complete landscape of retinal epigenetic memory. The most surprising result was the 10-fold greater contribution of H3K9me3 to epigenetic memory than other epigenetic marks. During rod differentiation, changes in H3K27me3 are the most numerous and the dramatic, but those changes are rapidly reversed in r-iPSCs. During reprogramming, the mechanisms that reset the H3K27me3 epigenome may be more efficient than those that reset the H3K9me3 epigenome. Indeed, large chromatin domains with H3K9me3 are impediments to reprogramming (Soufi et al., 2012). Also, a substantial proportion of the epigenetic memory was active repression of genes important for neurogenesis. This may indicate that rather than retaining a more permissive or active epigenetic state, the neurogenic program is more actively repressed in r-iPSCs as part of the reprogramming process, as seen with chromatin changes that occur during 3 key stages of reprogramming in MEFs (Apostolou and Hochedlinger, 2013). With MEF reprogramming, active lineage-specific promoters are silenced during the initial stage, followed by silencing of the permissive enhancers/promoters that transition to a poised chromatin state (Apostolou and Hochedlinger, 2013). During the final stage of reprogramming, the repressed genes (H3K9me3 and H3K27me3) such as *Sox2* and *Nanog* are reset to their pluripotent state (Apostolou and Hochedlinger, 2013). According to this model, the retinal neuron/rod genes would be the most likely to be reset in r-iPSCs, and the subset of those genes that are H3K27me3- or H3K9me3-repressed in MEFs would be least likely to be reset in f-iPSCs. Our data are consistent with this model.

On the basis of these data, we hypothesize that genes that undergo the most dramatic changes in epigenetic state during development are those most likely to be properly reset in iPSCs derived from those cell types. As a result, cells are more likely to produce the tissue they were derived from because the epigenome of the developmental program has been properly reset. Whether several genes are required for retinal differentiation or only a few are involved in early eye-field specification during neurogenesis remains unknown. Analyses of many iPSC lines from different neuronal types will be required to test this hypothesis.

Mapping the Retinoblastoma Epigenome to Retinal Development

Previous studies suggest that retinoblastoma expresses multiple programs of photoreceptors, progenitor cells, and interneurons in single cells (Johnson et al., 2007; McEvoy et al., 2011a). This may result from an uncoupling of the tightly coordinated epigenomic transitions during development, or it could reflect a developmental transition state that is fixed in the tumor. Our data favor the latter possibility with the epigenomes in human and murine retinoblastomas matching the developmental transition from neurogenic to terminal patterns of cell division. The human retinoblastomas had an epigenetic program consistent with photoreceptor specification, while the murine retinoblastomas were more consistent with amacrine cells and horizontal interneurons, potentially reflecting subtle differences in the developmental stage when the retinoblastoma forms. As peak birthdates for amacrine/horizontal neurons precedes that of rods, this suggests that murine retinoblastoma intiate earlier than human retinoblastoma. Most GEMMs of retinoblastoma use retinal progenitor cell–specific *Cre* transgenes to inactive Rb family members during retinal development, so it is not surprising that these tumors may initiate earlier in mice than in humans. Detailed lineage tracing experiments will be required to identify the cell of origin for retinoblastoma.

Of the 783 genes that were upregulated in murine retinoblastoma relative to E14.5 mouse retina, approximately one-third were epigenetically regulated during development. Of the 470 genes that were downregulated in murine retinoblastoma relative to P21 mouse retina, two-thirds were epigenetically regulated during retinogenesis. This suggests that a very specific, powerful selection for a particular epigenomic state occurs during retinogenesis, when malignant transformation is initiated as a result of Rb-pathway inactivation. Tumors preserved the earliest epigenetic state for their cell cycle genes (E14.5) but were further along in development (P0–P3) for the progenitor and rod genes. This concept of epigenomic vulnerability during particular stages of development may apply to other pediatric solid tumors that are arrested at an intermediate stage of development.

STAR METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Michael A. Dyer (michael.dyer@stjude.org).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human Retina and Retinoblastoma—Human retina was provided by Advanced Bioscience Resources (ABR). The samples included: 6 fetal week 13, 7 fetal week 14, 6

fetal week 15, 2 fetal week 16, 4 fetal week 18, 8 fetal week 23 and 1 fetal week 24 pairs. Only fetal week 13 and 14 samples were combined to obtain enough tissue for ChIP. This is designated as FW13/14 in the manuscript. The project was reviewed and approved by the St. Jude Institutional Review Board. All samples were de-identified and only developmental stage was provided by ABR. Human patient-derived orthotopic xenografts were generated and expanded as described previously (Zhang et al., 2012).

Murine Retina and Retinoblastoma—All animal procedures and protocols were approved by the St. Jude Laboratory Animal Care and Use Committee. All studies conform to federal and local regulatory standards. Mice were housed on ventilated racks on a standard 12 hour light-dark cycle. Wild type C57BL/6J mice were purchased from the Jackson Laboratory (Bar Harbor, ME). For timed pregnancy, individual male mice were housed with 4 females in a single cage. Plugged/pregnant females (identified by visual examination and/or palpation) were isolated and embryos or pups were harvested at the appropriate time. Both males and females were combined for this study. The genetically engineered mouse model of retinoblastoma (*Chx10-Cre;*$Rb^{Lox/Lox}$*;*$p107^{-/-}$$MDMX^{Tg}$) was described previously (McEvoy et al., 2011a).

iPSCs and ESCs—The generation and growth of the rod-derived iPSCs and f-iPSCs has been reported previously and is also provided below (Hiler et al., 2015; Hiler et al., 2016). The EB5:Rx-GFP ESC line was purchased from the RIKEN cell line repository. It was validated by whole genome sequencing and comparison to the EB5 parental line that was also sequenced. All the lines used in this study are from males.

Generation of fibroblast-derived iPSCs: RtOK MEFs were dissociated with 0.25% Trypsin (Gibco), and seeded at 10,000 cells/well on a gelatin-coated, 6-well tissue culture plate with an irradiated wild-type MEF feeder layer. Doxycycline (2 μg/ml) was added to the MEF maintenance media and cultured for two days with daily media exchanges. After two days, the MEF maintenance media was exchanged with ESC maintenance media with 2 μg/ml doxycycline added to maintain colony formation. On the sixth day after initiating reprogramming, the MEF lines were dissociated with 0.25% trypsin and re-seeded at 10,000 cells/well on a fresh irradiated wild-type MEF feeder layer. Reprogrammed MEFs were reseeded to reduce density of iPSC-like colonies per well, which made the isolation of individual colonies for sub cloning possible. After 10 days, individual iPSC colonies were isolated, dissociated with 0.25% trypsin for 5 minutes at room temperature, and dispersed on a gelatin-coated 24-well plate with an irradiated MEF feeder layer. Clones were expanded for twenty passages at which point, individual sub clone lines were frozen for further characterization (see below).

Generation of retinal derived iPSCs: Rod photoreceptor neurons from Nrl^{wt}-rtOK mice were harvested at two different developmental stages to elicit the reprogramming profile of a post-mitotic, photoreceptor committed retinal neurons. Nrl-GFP^{+} cells isolated on postnatal day 2–4 were considered “immature” rod photoreceptor neurons while Nrl-GFP^{+} cells isolated on postnatal day 21–23 were considered “mature” rod photoreceptor neurons.

At these two developmental time points, retinas from littermate animals were harvested and digested with 1 mg/mL trypsin (Sigma) in PBS for 5 minutes at 37°C. Retinas were triturated with a p1000 pipette 3–5 times to create a single cell suspension before an additional 3 minute incubation at 37°C. Soybean-trypsin inhibitor (STI, Sigma) and 0.2 mg/mL DNAse (Sigma) was added to each cell suspension and incubated for an additional 5 minutes at 37°C.

To remove cellular debris, each cell suspension was filtered through a bovine serum albumin (BSA, Sigma) gradient prior to collecting the population of Nrl-GFP+ cells via *Fluorescence Activated Cell Sorting (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. We determined this procedure was optimal in ensuring the population of GFP+ cells collected had minimal extracellular contamination.

To facilitate the reprogramming of retinal neurons, the FACs sorted population of Nrl-GFP+ neurons were cultured as organotypic explant cultures with a C57B6/J wild-type retina acting as the substrate medium to maintain the cells in vitro. Sorted cells were seeding in a two-step limiting dilution series with a 1.5 e^6 retinae cell suspension from postnatal day 0–6 C57B6/J mice. A 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. Organotypic explant cultures were created by spinning the mixed cell suspension at 14,000 × g for 20 seconds in 1.5 mL tubes (Eppendorf). Organotypic explants were transferred with a p1000 pipette to a nuclepore track-etch membranes (Whatman) floating on 2 mL of retinal explant media. Retinal explant media contains DMEM/F12 (Gibco) supplemented with 10% FBS, 5 ug/mL Insulin (Gibco), Pen/Strep) with 2 ug/mL doxycycline and 2,000 units/mL of LIF. A 10 uL drop of retinal explant media was added to the explant cultures daily while reprogramming to feed the cells with a half media exchange of the media every third day. After 10 days of reprogramming, the remaining retinal pellets were collected and individually isolated and digested with trypsin as previously above. For these cell suspensions, the BSA cushion gradient was not used in order to preserve the number of cells harvested. Each pellet was then seeded on gelatin-coated 6-well culture dishes with an irradiated MEF feeder layer in a second limiting dilution series of 20,000, 2,000, 200 cells per well. Colony forming events were recorded for each pellet prior to isolation of individual iPSC colonies for subclone expansion. At this point the Nrl-GFP retinal neurons were renamed r-IPSC and expanded for twenty passages at which point, individual sub clone lines were frozen for further characterization.

Characterization of iPSC lines: riPSC and fiPSC subclones established from reprogramming were initially characterized based on colony morphology and rate of growth. 24-subclones were established for each iPSC line and passaged five times after which the number of lines was reduced to 12 candidate lines, which were selected for further characterization. These 12 candidate lines were banked every five passages up until passage 20 and karyotyped to identify lines that had minimal chromosomal abnormalities. Of these 12 initial subclones, 4 lines were chosen for a complete characterization of cellular pluripotency and retinal differentiation.

Each iPSC line was passaged onto gelatin-coated, tissue culture plates containing inactivated feeder cells with or without exogenous LIF. iPSC cells were grown for several days in these culture conditions to monitor differentiation and loss of pluripotency via alkaline phosphatase (AP, Millipore) staining. iPSC lines grown on gelatin-coated chamber slides were fixed with 4% paraformaldehyde for 30 minutes, rinsed 3× with PBS, and incubated 3–4 hours at room temperature with an appropriate blocking serum supplemented with 0.5% Trition X-100. Slides were incubated with primary antibodies overnight at 4°C, 1–2 hours with secondary antibody (Vector) at room temperature, and 30 minutes with an ABC Kit (Vectastain) at room temperature. Immunostaining was visualized by a 10 minute incubation in cyanine 3 conjugated-tyramide reagent (PerkinElmer) prior to DAPI counter staining. Day 28 retinal spheres were fixed with 4% paraformaldehyde overnight at 4°C before immersion overnight in a 4% sucrose solution. To test the pluripotency of each iPSC line to produce all three germ layers, 200,000 cells were harvested and injected into each flank of adult CD1-Nude mice. Teratoma growth was monitored in these animals for up to 12 weeks after injection and removed before the tumor reached 20% of the animal's body weight. Teratoma tumors were resected and fixed in 4% paraformaldehyde at 4°C for several days prior to paraffin sectioning. Individual germ layers were identified by hematoxylin and eosin staining or with the following primary antibodies CK Oscar (endoderm), Vimentin and Desmin (mesoderm), or Synaptophysin (ectoderm).

METHOD DETAIL

Sequencing

RNA Preparation: RNA was isolated from individual TRIzol (Life Technologies) preparations via a phenol–chloroform extraction. Samples were first homogenized at 17,000 rpm for 30 s with a tissue homogenizer (Polytron, PT10-35GT). A 1:4 volume of chloroform (Sigma) was then added to each sample and incubated at room temperature for 3 min followed centrifugation at 12,000 ×*g* at 4°C for 15 min. The aqueous layer was then transferred to a siliconized Eppendorf tube followed by the addition of 2.0 μL glycogen (Roche) and 500 μL isopropanol (Fisher Scientific). Samples were incubated at room temperature for 10 min followed by centrifugation at 12,000 ×*g* at 4°C for 15 min. Samples were then washed twice with ice-cold 80% EtOH (Fisher) to remove salts, resuspended in DEPC H_2O, and the concentration was determined with a NanoDrop (Thermoscientific).

Whole-Genome Bisulfite Sequencing: Genomic DNA was extracted from retinal tissues by using the DNeasy Blood & Tissue Kit (QIAGEN #69504) according to the manufacturer's protocol. Typically, 2 μg genomic DNA is used to make WGBS libraries. WGBS was conducted by the Hudson Alpha Research Institute for Biotechnology using the Cambridge Epigenetix TrumethylSeq kit (BS workflow) per the manufacturer's instructions.

ChIP antibody validation: All antibodies were independently validated and biological replicates were performed for each mark at each stage of development. For most antibodies, we also validated the data using an independent antibody. All antibody validation data and SOPs for ChIP are freely available through the CSTN (www.stjude.org/cstn). Specifically, for the H3K27me3 mark, we have two antibodies with two replicates each at every stage. To estimate the consistency between antibodies, we first called domains using SICER as above,

then at each stage for each antibody domains overlapping common regions between two replicates have been merged, at last we merged the domains from two antibodies and counted extend reads for each sample overlapping these merged domains. At each stage, we draw the correlation plot between each pair of samples as (Fig. S1), as most of the correlation coefficient is > 0.9 (except one outlier at P7 as replicate 1 for AB3), we think both antibodies are working consistently and we used AB1 in the final analysis since it had better signal to noise profile.

ChIP-Seq: All antibodies were independently validated, and those validation data and protocols are available through the Childhood Solid Tumor Network website (https://www.stjude.org/CSTN/) (Stewart et al., 2016). Freshly isolated retinae or cells were cross-linked for 10 min in 1% ChIP-seq–grade formaldehyde (Thermo scientific # 28906) in 1X PBS at room temperature. Glycine was added to a final concentration of 0.125 M to stop the cross linking reaction. The retinal tissue was washed in 1X PBS and then dispersed into a single cell suspension using a dounce homogenizer with 7 ml pestle (Fisher scientific # 06-435A). 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). Ten percent of chromatin from each ChIP reaction was used as an input. After de-crosslinking, DNA was extracted using MinElute PCR-purification kit (QIAGEN #28006), quantified using the Quant-iT PicoGreen ds DNA assay (Life Technologies #Q33120) and subjected to qPCR analysis and library construction for sequencing. The RNA-Pol II, Brd4 and H3K9me3 ChIP was performed by Active Motif.

Libraries were prepared from 5–10 ng DNA by using the NEBNext ChIP-Seq Library Prep Reagent Set for Illumina with NEBNext High-Fidelity 2× PCR Master Mix according to the manufacturer's instructions (New England Biolabs) with the following modifications: a 1:1 Ampure cleanup was added after adaptor ligation a total of 2 times. The Ampure size-selection step prior to PCR was eliminated. The 72 °C extension step of the PCR was lengthened to 45 s. Completed libraries were analyzed for insert-size distribution on a 2100 BioAnalyzer High Sensitivity kit (Agilent Technologies) or Caliper LabChip GX DNA High Sensitivity Reagent Kit (PerkinElmer). Libraries were quantified using the Quant-iT PicoGreen ds DNA assay (Life Technologies) and Kapa Library Quantificaiton kit (Kapa Biosystems) or low-pass sequencing on a MiSeq nano kit (Illumina). Fifty-cycle single-end sequencing was performed on an Illumlina HiSeq 2500.

Retina dissociation: For each sample, retinae were dissociated with Trypsin (Sigma-Aldrich, T9935), and then incubated with soybean-trypsin inhibitor (Sigma-Aldrich, T6522) and DNaseI (Sigma-Aldrich, D4513). After dissociation, explant culture medium (25 ml of FBS, 2.5 ml of HEPES, 2.5 ml of penicillin-streptomycin and 125 μl of insulin in 219.9 ml of DMEM/F12 with GlutaMAX) was added to the cell suspension, and then it was filtered through a 40 μm cell strainer (Falcon, 352340) and pelleted at 1,200 rpm for 10 min at 4 °C.

RNA-Seq: RNA was extracted from freshly isolated retinae by using RNeasy Plus Mini kit (Qiagen #74134). 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.)

ATAC-seq: ATAC-seq library preparation was performed as described (Buenrostro et al., 2013) with minor modifications. Briefly, for each experiment, 2×10^5 – 5×10^5 dissociated retinal cells were washed twice with PBS (Lonza, 17–516F) containing Proteinase Inhibitor Cocktail (Sigma-Aldrich P8340), incubated with 50 µl of lysis buffer (10 mM Tris-HCl (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$ and 0.3% IGEPAL CA-630) on ice for 10 min and centrifuged at 500 *g* for 10 min at 4 °C. After centrifugation, the pellet was immediately incubated with 50 µl of Tn5 transposition reaction mix (25 µl TD buffer, 2.5 µl TDE1 (Nextera DNA Library Prep Kit, Illumina, FC-121-1030) and 22.5 µl nuclease-free water) at 37 °C for 30 min. After purification with Qiagen MinElute PCR Purification Kit (Qiagen, 28006), transposed DNA fragments were amplified with NEBNext High-Fidelity 2× PCR Master Mix (New England Biolabs, M0541) for 5 cycles using the following PCR conditions: 72 °C for 5 min; 98 °C for 30s; and thermocycling at 98°C for 10s, 63°C for 30s and 72°C for 1 min. 5 µl of each library was amplified in a 25-µl qPCR reaction with KAPA SYBRFAST qPCR Kit (Kapa Biosystems, KK4603) to estimate the optimum number of additional cycles to finish amplification of the remaining 45-µl library.

Computational Analysis

RNA-Seq Analysis: FASTQ sequences were mapped to the mouse mm9 (MGSCv37 from Sanger) or human genome hg19(GRCh37) by StrongARM, developed for the PCGP (Downing et al., 2012). To estimate the FPKM values, we first downloaded GTF files(mouse vM7, human v23) from GENCODE (Downing et al., 2012) website and liftOver them to mm9 or hg19 by CrossMap version 0.1.5 (Zhao et al., 2014). 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 ".

Four gene sets (Rod, Progenitor, G2/M, Housekeeping) that were previously identified and validated by single-cell gene expression analysis (Cherry et al., 2009, Trimarchi et al., 2007 and Trimarchi et al., 2008) were clustered by their correlation of log2 (FPKM) across mouse development and the largest cluster has been used for Fig.2.

WGBS Analysis: Whole genome bisulfite sequencing (WGBS) data were mapped to GRCh37-lite (human) or mm9 (mouse) by BSMAP2.74 with the following parameters: -z 33 -f 5 -g 3 -r 0 -m 17 -x 600 –u and raw methylation status was called by Bis-SNP (Version 0.82.2). We developed a regression tree based approach to identify regional methylation pattern of a sample using m-values ($m=\log_2\frac{\#Methylated\ C+0.5}{\#Unmethylated\ C+0.5}$). Adjacent segments with insignificant (p > 1E-6, Student t-test) or small difference between average β-values ($\beta=\frac{\#Methylated\ C}{\#Methylated\ C+\#Unmethylated\ C}$) (< 0.1) were recursively merged.

Segmentation breakpoints from individual samples were pooled to generate an initial list of regions for testing correlation between methylation and gene expression. Correlation between regions with variable methylation (difference between highest and lowest average

m-values >= 0.3) and adjacent genes with variable expression (methylation region overlapping 10 kb upstream of TSS to TES, at least 2 fold difference between highest and lowest FPKM, highest FPKM >= 1) were calculated. Gene/region pairs with FDR <= 0.5 were retained for merging, where adjacent regions for the same gene (no more than 1 kb apart) with the same orientation (positive or negative) of correlation were merged. Correlations were re-evaluated and pairs with final FDR <= 0.1 were used as differentially methylated regions that were significantly correlated with gene expression.

ChIP-Seq Analysis: 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) or human genome hg19(GRCh37), 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(details of QC matrix could be found in supplementary table S1).

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 (median percentage 96.8%). 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.

ATAC-seq data analysis: Nextera adapter were trimming by cutadapt (version 1.9, paired-end mode, default parameter with " –m 6 –O 20 ") for 2×100bp paired-end reads obtained from all samples and aligned to mouse genome mouse genome mm9 (MGSCv37 from NCBI) by BWA (Bailey et al., 2009; Bailey and Gribskov, 1998; Li and Durbin, 2009; Quinlan and Hall, 2010; Robinson et al., 2011)(version 0.5.9-r26-dev, default parameter), duplicated reads were then marked with Picard(version 1.65(1160)) and only proper paired non-duplicated reads have been kept by samtools (parameter "-q 1 -F 1804" version 0.1.18 (r982:295)). After adjustment of Tn5 shift (reads were offset were offset by +4 bp for the sense strand and −5 bp for the antisense strand) we separated reads into nucleosome free, mononucleosome, dinucleosome, trinucleosome by fragment size and generated bigwig files by using the center 80bp of fragments and scale to 75M reads that fragments' length less than 2kb. Observed reasonable nucleosome free peaks and pattern of mono-, di-, trinucleosome on IGV (version 2.3.40). All data have more than 20M nucleosome free reads and after manual review, we performed peaks calling for nucleosome free reads by MACS2(version 2.1.0.20150603 default parameters with "--extsize 200 --nomodel", merged by bedtools if within 100bp) for individual samples. For each sample, we called twice individually as high confidence peaks(q-value cutoff 0.05) and low confidence peaks(q-value cutoff 0.5). To assure the reproducibility, we first finalized nucleosome free regions for each stages as only retained a peak if it called in one of the replicates and also overlap low confidence peaks from the other replicates and merged while keep the high confidence peak edges. Then we counted nucleosome free reads from each replicates and draw correlation plots, the pearson correlation coefficient all > 0.95 indicated our data were highly reproducible. Then we merged finalized nucleosome free regions from all 8 stages and annotated by whether they overlap each samples for each stage.

For comparison to the data in the previously published ATAC-seq analysis (Mo et al., 2016) we performed a liftOver to build mm9. All 88 WT special ATAC peaks were identified in our P14 and P21 ATAC-seq for both replicates. We also identified 1355 out of 1491 NrlKO specific peaks in our data. In total, there was very good agreement between our data and the previously published data. Specifically, 98.2% of the 2463 peaks identified previously in the WT and NrlKO were identified in our data.

Transcription Factor Analysis: For footprinting, bigwig files of the ATAC-seq data were normalized by autosomes reads to 200M reads(e.g. sample with 100M autosome reads would be scaled to double the bigwig profile). Then average bigwig files were generated by mean of replicates at each bp for each sample and motif matches within nucleosome free region were used for footprinting. The average profile across all motif matches at each bp from −100bp from motif match centers to +100bp. Finally, the footprinting profiles were smoothed with 10bp bins.

Enhancer Analysis: Active enhancers were identified from the H3K27Ac ChIP-seq and the poised enhancers were identified from the H3K4me1 ChIP-seq. The H3K27Ac high confidence peaks (q-value cutoff = 0.05) and low confidence peaks(q-value cutoff = 0.5) were identified using MACS2 for each sample. For the biological replicates, if there was overlap with the high confidence peaks or a high confidence peak in one replicate and a low

confidence peak in the other sample, it was selected as a putative active enhancer. For samples without replicates, we just used high confidence peaks from MACS2.

For H3K4me1, we first identified the domains using SICER for each replicate and extracted the common regions between replicates. These regions were selected as putative poised enhancers. For samples without replicates, we just used domains identified using SICER.

To determine if the enhancers were conserved between mouse and human, we used the liftOver algorithm with a 50% cutoff for conserved domains. If there was not any overlapping peak across species, we indicated the percentage of conservation between humans and mice.

For the analysis of iPSCs, we used the H3K27Ac high confidence peaks across f-iPSC and r-iPSC samples. To compare between the iPSC samples, we analyzed read counts using VOOM and those data are presented as Counts Per Million(CPM) in the tab “H3K27Ac-Rod_vs_Fibr”. Similar analyses were performed for H3K4me1 using domains identified with SICER. Mouse developmental enhancers were annotated in the column “H3K27Ac-Rod_vs_Fibr” and “H3K4me1-Rod_vs_Fibr” by Up2 (significant up that passed FDR 5% and > 2 fold in riPSC), Up (significant up that passed FDR 5% but not > 2 fold in riPSC), Down2 (significant down that passed FDR 5% and > 2 fold downregulated in riPSC), Down (significant down that passed FDR 5% but not > 2 fold down in riPSC), Other (overlap any peak in either fiPSC or riPSC but not significant).

Chromatin State Modeling: Non-duplicated aligned reads were extend by fragment size define 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 modeling 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 these 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.

ChromHMM Grouping: We first calculated the mahalanobis distance between each pair of variable genes(max FPKM across stages > 2 * min FPKM across stages) based on the normalized FPKM value and the percentage of 11 HMM states covered the gene flanking 2kb region along all development stages. Then we clustered genes from rod, progenitor, G2M, housekeeping groups by expression and identified subclusters for some groups as follows:

Rod_N2H : no active marks To highly active marks

Rod_N2M: no active marks To moderately active marks

Rod_L2H: low level of active marks To highly active marks

Rod_L2M: low level of active marks To moderately active marks

Progenitor_L2N: low level of active marks To no active marks

Progenitor_L2L: low level of active marks To low level of active marks

Progenitor_M2L: moderately active marks To low level of active marks

Progenitor_M2N: moderately active marks To no active marks

The G2/M genes and housekeeping genes did not separate into subgroups. For each variable gene, we calculated the average mahalanobis distances to each of these 10 groups and assign the group to a gene by the group with smallest average distance (Table S4, "Group_HMM_Exp"). We then repeat the group assignment by using mahalanobis distances for log2 FPKM (Table S4, "Group_Exp"). Since a gene could be assigned to one of these 10 groups but actually not similar to any of them, we indicated which 10 percentile the gene ranked across all genes to that group accordingly in Table S4 by the columns started "HMMQ10"(by distance of HMM and Expression) or "ExpQ10"(by distance of log2 FPKM).

Super-enhancer Analysis: 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. In additional to the SEs called by ROSE (cutoff

chose by slope 1, 1353 out of 1354 overlapping BRD4 peaks) we find many regions ranked top consistently in many stages and the percentage of these regions overlapping BRD4 peaks (1698 out of 1708) are high indicated those are SE regions. So 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.

To identify the dynamically regulated superenhancers during retinal development, we used the ROSE identification combined with analysis of the H3K27Ac reads at E14.5 and P21. This was done to reduce the false positive calling of developmentally regulated superenhancers. Specifically, we scored the number of H3K2Ac reads for each SE region at E14.5 and P21 for both replicates, then we use Kolmogorov–Smirnov to determine if the number of reads was significantly different across those two stages (column “New_in_P21_by_Reads_Tests” or “Disappear_in_P21_by_Reads_Tests” in TableS8).

After get the development regulated genes, we annotated SE in human and mouse if the development regulated genes is within 500kb of SE and tried to match the SE to development regulated genes both in mouse and human and result in tab “Dev_Regu_both”, SE and gene pair in this tab indicated gene has been identified as development regulated genes in both human and mouse meanwhile the SE in human is within 500kb of the gene and the conserved region in mouse from the same SE is within 500kb of gene in mouse(some might be development regulated in different direction but could be easily filtered out). “Human_Dev_Regu” and “Mouse_Dev_Regu” columns indicated how strong they have been regulated along development(suffix _fc2 indicated 2 fold change, suffix _trend indicated identified as development regulated genes above) and direction(“up” indicated higher expressed in latest stage). The tab include every potential development SE in human could be liftover(cutoff 20%) to mouse so it is recommend to filter by column “Overlap_Mm9_Development_SE” or “Enhancer_Loci_Mouse” that the first one assures after liftOver it’s in active SE region, the second one could help avoid regions between enhancers but within SE.

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.

H3K27me3 Analysis: Manually reviewed genes overlapping H3K27me3 domains called by SICER, we found many genes overlapping H3K27me3 domains at early stage but not later only have weak H3K27me3 signal. After HMM modeling, we calculated the percentage of a gene(+/− 2kb) or promoter(TSS-2kb) occupied by different chromatin state, genes with at least 20% promoter covered by state 4 or state 8 at P0 and twice larger than P21 or with at least 10% gene region covered by state 4 or state 8 at P0 and twice larger than P21 have been considered as genes active derepressed by H3K27me3. Similar, genes got larger state 4 and state 8 at P21 compared to P0 are considered as active repressed. To identify other genes sharing similar pattern, we further filtering genes meeting this criterion by 2 fold expression

level change comparing P0 to P21 but opposite to direction of percentage of HMM state(e.g. higher expression level at P21 and smaller percentage of state 4 and state 8 together defines active derepressed genes). Finally, we counted H3K27me3 reads in Human at earliest stage and latest stage for these candidate genes and annotated the gene by whether the normalized reads number increase or decrease 2 fold(at least 20 reads for the bigger one) either at promoter or gene flanking 2kb region.

TAD Annotation: Mouse Cortex TAD bed file have been downloaded from Bing Ren's lab webpage. For each TAD we counted genes (The TSS and TES of a gene should both within the TAD to be counted) in each of the pre-defined groups (rod, progenitor, G2/M and housekeeping) and use Fisher's exact test to calculate the p-value of how each those groups is enriched in this TAD compared to the background probability of observing genes in those groups within or not within a TAD. The group with lowest p-value have been assigned to the TADs. We then cluster TADs by the average mahalanobis distance of log2 FPKM of genes within one TAD to genes within another TAD. Coloring TADs by the above group assignment, we found one TAD cluster is enriched for rod genes and another one is enriched for progenitor genes. Finally, we calculated the percentage of a TAD have been covered by different HMM states, and annotated a TAD as "active" TAD if the total percentage of state 1,2,3,5,6,7 is top 5 percentile or "polycomb" if the total percentage of state 4,8 is top 10 percentile or "heterochromatin" if the percentage of state 10 is top 5 percentile or "null" if the percentage of state 9 is top 5 percentile. If a TAD was enriched for both polycomb and heterochromatin, it was labeled heterochromatin.

Retinoblastoma Analysis: To identify the genes that were most significantly upregulated in human retinoblastoma, we started with the hg19 up gene list in Table S11. If the variability between xenograft and primary tumor was greater than 2-fold, the gene was eliminated from this analysis. Next, we filtered out those genes that had tumor values below 10 and then eliminated those that had a tumor/normal ratio below 2-fold. Finally, we sorted the list by the multiplying (tumor/normal)x(tumor-normal). In total 408 genes met this criteria. A similar analysis was carried out to identify the genes that were most significantly downregulated in human retinoblastoma. The only difference was that we filtered out the genes that had levels below 10 for the normal and we sorted based on a 2-fold cutoff for normal/tumor rather than tumor/normal. The final list of 500 genes was sorted by (normal/tumor)x(normal-tumor). The same approach was used for mouse retinoblastoma upregulated (784) and downregulated (470) genes. The only difference was that we only analyzed one MDMX tumor so eliminating those tumors with high variability was not necessary.

Imaging

3D Electron Microscopy: The samples for 3D EM were stained with a modified heavy-metal staining method and processed through a graded series of alcohol and propylene oxide washes, and then embedded in Epon hard resin. Thick (0. 5-μm) sections were cut to determine the correct area and then coated with iridium in a Denton Desk II sputter coater. The 3D EM images were collected on a Helios Nanolab 660 Dualbeam system.

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.

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.

Enhancer Validation

Cloning Enhancers: DNA fragments that are likely to be a “hotspot” for activity were amplified from mouse genomic DNA. All primers used are listed in table. Enhancer regions were chosen based on the following criteria:

1. Strong overlapping peaks of BRD4 and H3K27Ac.
2. Strong ATAC-seq signal (P21) overlapping with BRD4 and H3K27me3 peaks
3. Regions within 2 kb from TSS were excluded.
4. Super enhancers that lie outside TADs were excluded.
5. Patterns of BRD4, H3K27Ac strongly resemble related gene expression.
6. If there are multiple regions, the one closest to the TSS (beyond 2 kb of TSS) were picked.
7. In selected SEs, multiple “hotspots” of interest were picked.

DNA fragments were cloned in XhoI site of Nanluc plasmid containing minimal promoter pNL3.1[Nluc/minP] (Promega N103A) using standard cloning techniques. The primer sequences and coordinates are available upon request.

Luciferase reporter assay: 90 ug of enhancer constructs and 10 ug of firefly luciferase plasmid (pGL4.51[luc2/CMV/Neo] (Promega E132A) (normalization control) were co-electroporated into C57/BL6 retinae that were dissected at P0. Retinae were cultured in explant culture media for 6 days. Then retinae were dissociated and Luciferase assay were performed using Nano-Glo Dual-Luciferase Reporter (NanoDLR) assay system (Promega N1610) according to the manufacturer's protocol. Luminescence measurements were taken as singlets using a SpectraMax M5 (Molecular Devices) plate reader. Experiments were performed in quadruplicates.

Data Processing for Figures—To simplify visualization in Figure 2E–G and Fig. 3B, averaged ChIP-seq data were manually traced and scaled in Adobe Illustrator and shaded according the indicated color scheme.

QUANTIFICATION AND STATISTICAL ANALYSIS

The statistical details of experiments can be found above in the METHODS DETAILS section for each type of analysis.

DATA AND SOFTWARE AVAILABILITY

All of the software for the analysis is freely available and is described above in the METHODS DETAILS section for each type of analysis. All sequence data has been deposited in the GEO database with the following accession number: GSE87064.

ADDITIONAL RESOURCES

We have developed a new website so users can browse the gene expression and chromHMM for any genomic region. The site is: https://pecan.stjude.org/proteinpaint/study/retina2017

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal Anti-H3K4me3	Diagenode	Cat# C15410003-50
Rabbit polyclonal Anti-H3K4me2	Abcam	Cat# ab7766, RRID:AB_2560996
Rabbit polyclonal Anti-H3K4me1	Abcam	Cat# ab8895, RRID:AB_306847
Rabbit polyclonal Anti-H3K9/14Ac	Diagenode	Cat# C15410200
Rabbit polyclonal Anti-H3K27Ac	Abcam	Cat# ab4729, RRID:AB_2118291
Rabbit polyclonal Anti-H3K27me3	Millipore	Cat# 07-449, RRID:AB_310624
Rabbit polyclonal Anti-H3K27me3	Active Motif	Cat# 39155, RRID:AB_2561020
Rabbit polyclonal Anti-CTCF	Active Motif	Cat# 61311, RRID:AB_2614975
Rabbit polyclonal Anti-RNA pol II	Abcam	Cat# ab5095, RRID:AB_304749

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Rabbit polyclonal Anti-BRD4	Bethyl Laboratories	Cat# A301-985A, RRID:AB_1576498
Rabbit polyclonal Anti-H3K9me3	Active Motif	Cat# 39161, RRID:AB_2532132
Rabbit polyclonal Anti-H3K36me3	Active Motif	Cat# 61101, RRID:AB_2615073
Biological samples		
Healthy fetal retinal tissue stage FW13	Advanced Bioscience Resources, INC	N/A
Healthy fetal retinal tissue stage FW14	Advanced Bioscience Resources, INC	N/A
Healthy fetal retinal tissue stage FW15	Advanced Bioscience Resources, INC	N/A
Healthy fetal retinal tissue stage FW16	Advanced Bioscience Resources, INC	N/A
Healthy fetal retinal tissue stage FW18	Advanced Bioscience Resources, INC	N/A
Healthy fetal retinal tissue stage FW20	Advanced Bioscience Resources, INC	N/A
Healthy fetal retinal tissue stage FW23	Advanced Bioscience Resources, INC	N/A
Healthy fetal retinal tissue stage FW24	Advanced Bioscience Resources, INC	N/A
Retinoblastoma Patient-derived xenografts (PDX)	The Childhood Solid Tumor Network (CSTN), St Jude Children's Research Hospital	www.Stjude.org/cstn
Chemicals, Peptides, and Recombinant Proteins		
Critical Commercial Assays		
TruChIP chromatin shearing kit	Covaris	Cat# 520127
iDeal ChIP-seq kit	Diagenode	Cat# C01010051
Nano-Glo Dual-Luciferase Reporter (NanoDLR)	Promega	Cat# N1610
Deposited Data		
Raw and analyzed data	This paper	GEO: GSE87064
Mouse ES cell topological domains	Dixon, et al. 2012	http://chromosome.sdsc.edu/mouse/hic/download.html
ATAC-Seq data for WT and NrlKO in mouse	Mo, et al. 2016	GEO: GSE72550
Crx ChIP-Seq in mouse for WT and NRLKO	Corbo, et al. 2010	GEO: GSE20012
Otx ChIP-Seq in mouse for retina and RPE	Samuel, et al. 2014	GEO: GSE54084
Human reference genome NCBI build 37, GRCh37(hg19)	Genome Reference Consortium	https://www.ncbi.nlm.nih.gov/grc/human
Mouse reference genome9(C57BL/6J) MGSC build 37, MGSCv37(mm9)	Mouse Genome Sequencing Consortium	https://www.ncbi.nlm.nih.gov/grc/mouse
Reference genome annotation mouse vM7, human v23	Harrow, et al. 2012	http://www.gencodegenes.org
Experimental Models: Cell Lines		
ESC line EB5:Rx-GFP	Riken Cell Bank	AES0145
Rod derived iPSC lines	Hiler et al. 2015	8601, 8602, 7602, 3302
Fibroblast derived iPSC lines	Hiler et al. 2015	FNR07, FMG02
Experimental Models: Organisms/Strains		

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mouse: C57/BL6	The Jackson Laboratory	JAX: 000664
Mouse: Chx10-Cre;RbLox/Lox;p107/MDMXTg	McEvoy et al, 2011	N/A
Oligonucleotides		
Recombinant DNA		
pGL4.51[luc2/CMV/Neo]	Promega	E132A
pNL3.1[Nluc/minP]	Promega	N103A
pNL1.1 CMV[Nluc/CMV]	Promega	N109A
Software and Algorithms		
BWA version 0.5.9-r26-dev	Li, et al. 2010	http://bio-bwa.sourceforge.net
Picard version 1.65(1160)	https://broadinstitute.github.io/picard/	https://broadinstitute.github.io/picard
samtools version 0.1.18 (r982:295)	Li, et al. 2009	http://www.htslib.org/
Bedtools version 2.17.0	Quinlan, et al. 2010	https://github.com/arq5x/bedtools2
IGV version 2.3.40	Robinson, et al. 2011	http://software.broadinstitute.org/software/igv
cutadapt version 1.9	Martin, et al. 2011	https://github.com/marcelm/cutadapt
Liftover 031411	Kent, et al. 2002	https://genome-store.ucsc.edu
ROSE 2015-02-10	Whyte, et al. 2013 and Lovén, et al. 2013	https://bitbucket.org/young_computation/rose
CRCMapper 2016-10-14	Saint-André, et al. 2016	https://bitbucket.org/young_computation/crcmapper
Bis-SNP version 0.82.2	Liu, et al. 2012	https://sourceforge.net/projects/bissnp
BSMAP version 2.74	Xi, et al. 2009	https://github.com/genestack-open/bsmap
ChromHMM version 1.10	Ernst J, et al 2012.	http://compbio.mit.edu/ChromHMM
MACS2 version 2.1.0.20150603	Zhang, et al. 2008	https://github.com/taoliu/MACS
SICER version 1.1	Zang, et al. 2009	http://home.gwu.edu/~wpeng/Software.htm
CrossMap version 0.1.5	Zhao, et al. 2013	http://crossmap.sourceforge.net/
Cufflinks package version 2.1.1	Trapnell, et al. 2010	http://cole-trapnell-lab.github.io/cufflinks
R version 2.14.0	R Development Core Team, 2011	https://www.r-project.org
SPP version 1.11	Kharchenko, et al. 2008	http://compbio.med.harvard.edu/Supplements/ChIP-seq
caTools version 1.17	R package	https://cran.r-project.org/web/packages/caTools/index.htm
Bitops version 1.0–6	R package	https://cran.r-project.org/web/packages/bitops/index.html
Voom function from edgeR 3.12.1, limma 3.26.9	Law, et al. 2014	http://dx.doi.org/10.1186/gb-2014-15-2-r29
ggplot2 version 2.2.1	Wickham, et al. 2009	http://ggplot2.org
StrongARM	Downing, et al. 2012	https://www.ncbi.nlm.nih.gov/pubmed/22641210
Other		
chromHMM, tracks and expression level	This paper	https://pecan.stjude.org/proteinpaint/study/retina2017

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

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 a grant from Alex's Lemonade Stand Foundation for Childhood Cancer, the Tully Family Foundation, and the Peterson Foundation. Most of this research was supported by HHMI. The human retina and retinoblastoma RNA-Seq and WGBS was performed as part of the St. Jude Children's Research Hospital—Washington University Pediatric Cancer Genome Project. The murine retina, retinoblastoma, and iPSC sequencing were performed with assistance from the Hartwell Center at St. Jude. 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.

References

Akimoto M, Cheng H, Zhu D, Brzezinski JA, Khanna R, Filippova E, Oh EC, Jing Y, Linares JL, Brooks M, et al. Targeting of GFP to newborn rods by Nrl promoter and temporal expression profiling of flow-sorted photoreceptors. Proc Natl Acad Sci U S A. 2006; 103:3890–3895. [PubMed: 16505381]

Apostolou E, Hochedlinger K. Chromatin dynamics during cellular reprogramming. Nature. 2013; 502:462–471. [PubMed: 24153299]

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009; 37:W202–208. [PubMed: 19458158]

Bailey TL, Gribskov M. Combining evidence using p-values: application to sequence homology searches. Bioinformatics. 1998; 14:48–54. [PubMed: 9520501]

Berkyurek AC, Suetake I, Arita K, Takeshita K, Nakagawa A, Shirakawa M, Tajima S. The DNA methyltransferase Dnmt1 directly interacts with the SET and RING finger-associated (SRA) domain of the multifunctional protein Uhrf1 to facilitate accession of the catalytic center to hemi-methylated DNA. J Biol Chem. 2014; 289:379–386. [PubMed: 24253042]

Blackshaw S, Fraioli RE, Furukawa T, Cepko CL. Comprehensive analysis of photoreceptor gene expression and the identification of candidate retinal disease genes. Cell. 2001; 107:579–589. [PubMed: 11733058]

Blackshaw S, Harpavat S, Trimarchi J, Cai L, Huang H, Kuo WP, Weber G, Lee K, Fraioli RE, Cho SH, et al. Genomic Analysis of Mouse Retinal Development. PLoS Biol. 2004; 2:E247. [PubMed: 15226823]

Blackshaw S, Kuo WP, Park PJ, Tsujikawa M, Gunnersen JM, Scott HS, Boon WM, Tan SS, Cepko CL. MicroSAGE is highly representative and reproducible but reveals major differences in gene expression among samples obtained from similar tissues. Genome Biol. 2003; 4:R17. [PubMed: 12620102]

Brzezinski, JAt, Kim, EJ., Johnson, JE., Reh, TA. Ascl1 expression defines a subpopulation of lineage-restricted progenitors in the mammalian retina. Development. 2011; 138:3519–3531. [PubMed: 21771810]

Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods. 2013; 10:1213–1218. [PubMed: 24097267]

Burmeister M, Novak J, Liang MY, Basu S, Ploder L, Hawes NL, Vidgen D, Hoover F, Goldman D, Kalnins VI, et al. Ocular retardation mouse caused by Chx10 homeobox null allele: impaired retinal progenitor proliferation and bipolar cell differentiation. Nat Genet. 1996; 12:376–384. [PubMed: 8630490]

Cepko C. Intrinsically different retinal progenitor cells produce specific types of progeny. Nat Rev Neurosci. 2014; 15:615–627. [PubMed: 25096185]

Chen X, Pappo A, Dyer MA. Pediatric solid tumor genomics and developmental pliancy. Oncogene. 2015; 34:5207–5215. [PubMed: 25639868]

Cherry TJ, Trimarchi JM, Stadler MB, Cepko CL. Development and diversification of retinal amacrine interneurons at single cell resolution. Proc Natl Acad Sci U S A. 2009; 106:9495–9500. [PubMed: 19470466]

Cheung NK, Dyer MA. Neuroblastoma: developmental biology, cancer genomics and immunotherapy. Nat Rev Cancer. 2013; 13:397–411. [PubMed: 23702928]

Ciabrelli F, Cavalli G. Chromatin-driven behavior of topologically associating domains. J Mol Biol. 2015; 427:608–625. [PubMed: 25280896]

Denko NC. Hypoxia, HIF1 and glucose metabolism in the solid tumour. Nat Rev Cancer. 2008; 8:705–713. [PubMed: 19143055]

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012; 485:376–380. [PubMed: 22495300]

Downing JR, Wilson RK, Zhang J, Mardis ER, Pui CH, Ding L, Ley TJ, Evans WE. The Pediatric Cancer Genome Project. Nature genetics. 2012; 44:619–622. [PubMed: 22641210]

Dyer MA. Lessons from Retinoblastoma: Implications for Cancer, Development, Evolution, and Regenerative Medicine. Trends Mol Med. 2016a; 22:863–876. [PubMed: 27567287]

Dyer MA. Stem Cells Expand Insights into Human Brain Evolution. Cell Stem Cell. 2016b; 18:425–426. [PubMed: 27058930]

Dyer MA, Martins R, da Silva Filho M, Muniz JA, Silveira LC, Cepko CL, Finlay BL. Developmental sources of conservation and variation in the evolution of the primate eye. Proc Natl Acad Sci U S A. 2009; 106:8963–8968. [PubMed: 19451636]

Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. Nat Methods. 2012; 9:215–216. [PubMed: 22373907]

Furukawa T, Morrow EM, Cepko CL. Crx, a novel otx-like homeobox gene, shows photoreceptor-specific expression and regulates photoreceptor differentiation. Cell. 1997; 91:531–541. [PubMed: 9390562]

Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, Ye Z, Lee LK, Stuart RK, Ching CW, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature. 2009; 459:108–112. [PubMed: 19295514]

Hiler D, Chen X, Hazen J, Kupriyanov S, Carroll PA, Qu C, Xu B, Johnson D, Griffiths L, Frase S, et al. Quantification of Retinogenesis in 3D Cultures Reveals Epigenetic Memory and Higher Efficiency in iPSCs Derived from Rod Photoreceptors. Cell Stem Cell. 2015; 17:101–115. [PubMed: 26140606]

Hiler DJ, Barabas ME, Griffiths LM, Dyer MA. Reprogramming of mouse retinal neurons and standardized quantification of their differentiation in 3D retinal cultures. Nat Protoc. 2016; 11:1955–1976. [PubMed: 27658012]

Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, Hoke HA, Young RA. Super-enhancers in the control of cell identity and disease. Cell. 2013; 155:934–947. [PubMed: 24119843]

Huether R, Dong L, Chen X, Wu G, Parker M, Wei L, Ma J, Edmonson MN, Hedlund EK, Rusch MC, et al. The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. Nat Commun. 2014; 5:3630. [PubMed: 24710217]

Johnson DA, Zhang J, Frase S, Wilson M, Rodriguez-Galindo C, Dyer MA. Neuronal differentiation and synaptogenesis in retinoblastoma. Cancer Res. 2007; 67:2701–2711. [PubMed: 17363591]

Kansara M, Teng MW, Smyth MJ, Thomas DM. Translational biology of osteosarcoma. Nat Rev Cancer. 2014; 14:722–735. [PubMed: 25319867]

Kashi VP, Hatley ME, Galindo RL. Probing for a deeper understanding of rhabdomyosarcoma: insights from complementary model systems. Nat Rev Cancer. 2015; 15:426–439. [PubMed: 26105539]

Kim K, Doi A, Wen B, Ng K, Zhao R, Cahan P, Kim J, Aryee MJ, Ji H, Ehrlich LI, et al. Epigenetic memory in induced pluripotent stem cells. Nature. 2010; 467:285–290. [PubMed: 20644535]

Kunst S, Wolloscheck T, Grether M, Trunsch P, Wolfrum U, Spessert R. Photoreceptor cells display a daily rhythm in the orphan receptor Esrrbeta. Mol Vis. 2015; 21:173–184. [PubMed: 25737630]

Kurokawa D, Kiyonari H, Nakayama R, Kimura-Yoshida C, Matsuo I, Aizawa S. Regulation of Otx2 expression and its functions in mouse forebrain and midbrain. Development. 2004; 131:3319–3331. [PubMed: 15201224]

Kypta RM, Waxman J. Wnt/beta-catenin signalling in prostate cancer. Nat Rev Urol. 2012; 9:418–428. [PubMed: 22710668]

Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25:1754–1760. [PubMed: 19451168]

Lister R, Mukamel EA, Nery JR, Urich M, Puddifoot CA, Johnson ND, Lucero J, Huang Y, Dwork AJ, Schultz MD, et al. Global epigenomic reconfiguration during mammalian brain development. Science. 2013; 341:1237905. [PubMed: 23828890]

Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell. 2015; 161:1202–1214. [PubMed: 26000488]

Macpherson D, Dyer MA. Retinoblastoma: from the two-hit hypothesis to targeted chemotherapy. Cancer Res. 2007; 67:7547–7550. [PubMed: 17699756]

Mazelova J, Ransom N, Astuto-Gribble L, Wilson MC, Deretic D. Syntaxin 3 and SNAP-25 pairing, regulated by omega-3 docosahexaenoic acid, controls the delivery of rhodopsin for the biogenesis of cilia-derived sensory organelles, the rod outer segments. J Cell Sci. 2009; 122:2003–2013. [PubMed: 19454479]

McEvoy J, Flores-Otero J, Zhang J, Nemeth K, Brennan R, Bradley C, Krafcik F, Rodriguez-Galindo C, Wilson M, Xiong S, et al. Coexpression of normally incompatible developmental pathways in retinoblastoma genesis. Cancer Cell. 2011a; 20:260–275. [PubMed: 21840489]

Mcevoy J, Flores-Otero J, Zhang J, Nemeth K, Brennan R, Rodriguez-Galindo C, Wilson M, Xiong S, Lozano G, Sage J, et al. Coexpression of Normally Incompatible Developmental Pahtways in Retinoblastoma Genesis. Cancer Cell. 2011b in press.

Mo A, Luo C, Davis FP, Mukamel EA, Henry GL, Nery JR, Urich MA, Picard S, Lister R, Eddy SR, et al. Epigenomic landscapes of retinal rods and cones. Elife. 2016:5.

Mo A, Mukamel EA, Davis FP, Luo C, Henry GL, Picard S, Urich MA, Nery JR, Sejnowski TJ, Lister R, et al. Epigenomic Signatures of Neuronal Diversity in the Mammalian Brain. Neuron. 2015; 86:1369–1384. [PubMed: 26087164]

Ollion J, Cochennec J, Loll F, Escude C, Boudier T. TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization. Bioinformatics. 2013; 29:1840–1841. [PubMed: 23681123]

Orkin SH, Hochedlinger K. Chromatin connections to pluripotency and cellular reprogramming. Cell. 2011; 145:835–850. [PubMed: 21663790]

Otani T, Marchetto MC, Gage FH, Simons BD, Livesey FJ. 2D and 3D Stem Cell Models of Primate Cortical Development Identify Species-Specific Differences in Progenitor Behavior Contributing to Brain Size. Cell Stem Cell. 2016; 18:467–480. [PubMed: 27049876]

Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010; 26:841–842. [PubMed: 20110278]

Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. Nat Biotechnol. 2011; 29:24–26. [PubMed: 21221095]

Rodriguez-Galindo, C., Wilson, M., Dyer, M. Retinoblastoma. In: Orkin, S., editor. Hematology and Oncology of Infancy and Childhood. Philadelphia, PA: Elsevier; 2015. p. 1747-1778.

Roesch K, Jadhav AP, Trimarchi JM, Stadler MB, Roska B, Sun BB, Cepko CL. The transcriptome of retinal Muller glial cells. J Comp Neurol. 2008; 509:225–238. [PubMed: 18465787]

Roesch K, Stadler MB, Cepko CL. Gene expression changes within Muller glial cells in retinitis pigmentosa. Mol Vis. 2012; 18:1197–1214. [PubMed: 22665967]

Rohleder N, Langer C, Maus C, Spiwoks-Becker I, Emser A, Engel L, Spessert R. Influence of photoperiodic history on clock genes and the circadian pacemaker in the rat retina. Eur J Neurosci. 2006; 23:105–111. [PubMed: 16420420]

Saint-Andre V, Federation AJ, Lin CY, Abraham BJ, Reddy J, Lee TI, Bradner JE, Young RA. Models of human core transcriptional regulatory circuitries. Genome Res. 2016; 26:385–396. [PubMed: 26843070]

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012; 9:676–682. [PubMed: 22743772]

Sexton T, Cavalli G. The role of chromosome domains in shaping the functional genome. Cell. 2015; 160:1049–1059. [PubMed: 25768903]

Siegert S, Cabuy E, Scherf BG, Kohler H, Panda S, Le YZ, Fehling HJ, Gaidatzis D, Stadler MB, Roska B. Transcriptional code and disease map for adult retinal cell types. Nat Neurosci. 2012; 15:487–495. S481–482. [PubMed: 22267162]

Singh RK, Mallela RK, Hayes A, Dunham NR, Hedden ME, Enke RA, Fariss RN, Sternberg H, West MD, Nasonkin IO. Dnmt1, Dnmt3a and Dnmt3b cooperate in photoreceptor and outer plexiform layer development in the mammalian retina. Exp Eye Res. 2016

Solovei I, Kreysing M, Lanctot C, Kosem S, Peichl L, Cremer T, Guck J, Joffe B. Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. Cell. 2009; 137:356–368. [PubMed: 19379699]

Soufi A, Donahue G, Zaret KS. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell. 2012; 151:994–1004. [PubMed: 23159369]

Stewart E, Federico S, Karlstrom A, Shelat A, Sablauer A, Pappo A, Dyer MA. The Childhood Solid Tumor Network: A new resource for the developmental biology and oncology research communities. Dev Biol. 2016; 411:287–293. [PubMed: 26068307]

Swaroop A, Kim D, Forrest D. Transcriptional regulation of photoreceptor development and homeostasis in the mammalian retina. Nat Rev Neurosci. 2010; 11:563–576. [PubMed: 20648062]

Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B, et al. The accessible chromatin landscape of the human genome. Nature. 2012; 489:75–82. [PubMed: 22955617]

Trimarchi JM, Stadler MB, Cepko CL. Individual retinal progenitor cells display extensive heterogeneity of gene expression. PLoS One. 2008; 3:e1588. [PubMed: 18270576]

Voigt P, Tee WW, Reinberg D. A double take on bivalent promoters. Genes Dev. 2013; 27:1318–1338. [PubMed: 23788621]

Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell. 2013; 153:307–319. [PubMed: 23582322]

Workman AD, Charvet CJ, Clancy B, Darlington RB, Finlay BL. Modeling transformations of neurodevelopmental sequences across mammalian species. J Neurosci. 2013; 33:7368–7383. [PubMed: 23616543]

Young RW. Cell death during differentiation of the retina in the mouse. J Comp Neurol. 1984; 229:362–373. [PubMed: 6501608]

Young RW. Cell differentiation in the retina of the mouse. Anat Rec. 1985a; 212:199–205. [PubMed: 3842042]

Young RW. Cell proliferation during postnatal development of the retina in the mouse. Brain Res. 1985b; 353:229–239. [PubMed: 4041905]

Yu X, Ng CP, Habacher H, Roy S. Foxj1 transcription factors are master regulators of the motile ciliogenic program. Nat Genet. 2008; 40:1445–1453. [PubMed: 19011630]

Zentner GE, Tesar PJ, Scacheri PC. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome Res. 2011; 21:1273–1283. [PubMed: 21632746]

Zhang J, Benavente CA, McEvoy J, Flores-Otero J, Ding L, Chen X, Ulyanov A, Wu G, Wilson M, Wang J, et al. A novel retinoblastoma therapy from genomic and epigenetic analyses. Nature. 2012; 481:329–334. [PubMed: 22237022]

Zhao H, Sun Z, Wang J, Huang H, Kocher JP, Wang L. CrossMap: a versatile tool for coordinate conversion between genome assemblies. Bioinformatics. 2014; 30:1006–1007. [PubMed: 24351709]

Ziller MJ, Edri R, Yaffe Y, Donaghey J, Pop R, Mallard W, Issner R, Gifford CA, Goren A, Xing J, et al. Dissecting neural differentiation regulatory networks through epigenetic footprinting. Nature. 2015; 518:355–359. [PubMed: 25533951]

Zwang Y, Oren M, Yarden Y. Consistency test of the cell cycle: roles for p53 and EGR1. Cancer Res. 2012; 72:1051–1054. [PubMed: 22315347]

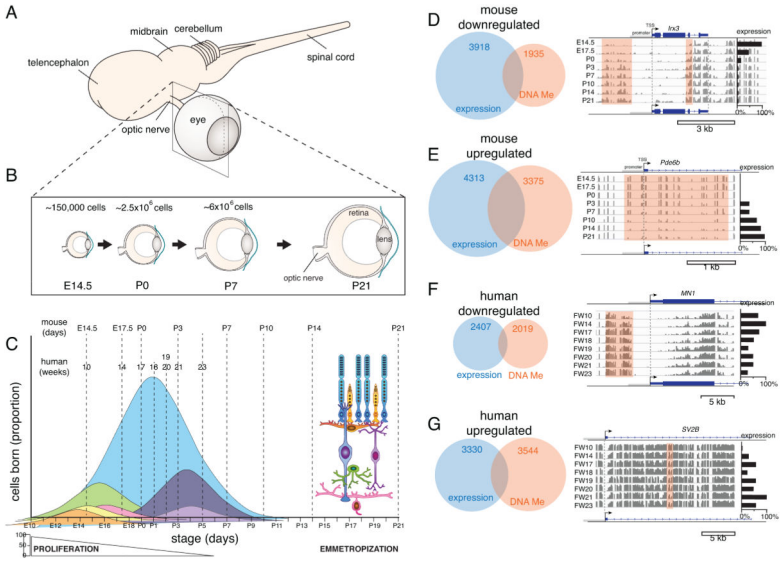

Figure 1. Cell Growth and Differentiation during Retinal Development

(A) Schematic of the developing murine CNS showing the autonomous development of the eye and retina.

(B) Drawing of murine retinal development highlighting proliferative growth from E14.5 to P7.

(C) Simplified diagram of the developmental birth order of retinal cell types; the area under each curve is proportional to the fraction of cells in the adult retina. The proportion of proliferating progenitor cells is represented along the bottom of the plot. Each stage of murine (Mu) and human (Hu) retinal development analyzed in this study is indicated by dashed lines.

(D–G)Venn diagrams of the genes that are up or downregulated (blue) and those that are hyper or hypomethylated (orange) during mouse and human retinal development. Examples are shown, and regions with changing DNA methylation are highlighted in orange. Normalized relative expression during development is shown in the histograms.

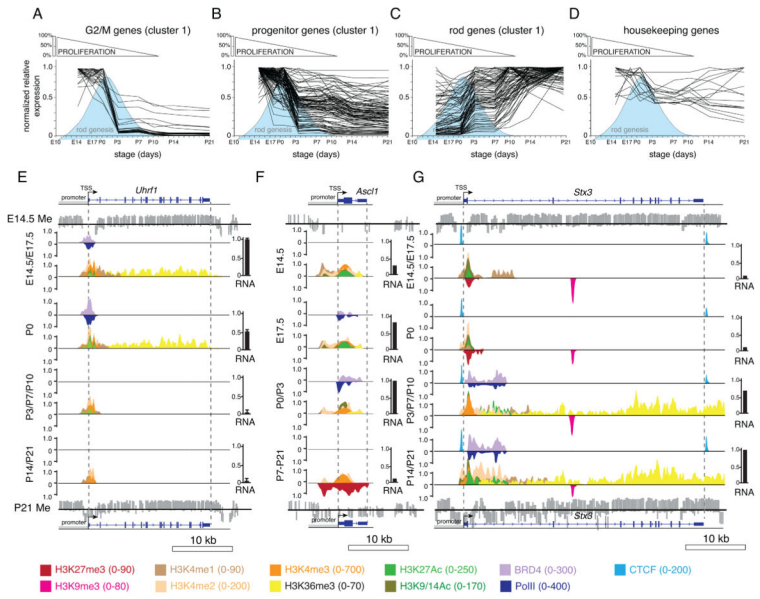

Figure 2. Epigenetic Changes Correlate with Gene Expression Changes during Retinal Development

(A–D) Normalized relative expression of representative genes from 8 stages of mouse retinal development, as determined by RNA-seq analysis. A subset of the genes identified by hierarchical clustering (cluster 1 in Data S1, Table S3) are shown for simplicity. The G2/M genes are required for retinal progenitor cell proliferation, (A) and the progenitor genes are expressed in retinal progenitors but are not necessarily involved in cell cycle control (B). The rod genes are upregulated as rods differentiate (C), and the housekeeping genes are constitutively expressed throughout retinal development (D).

(E–G) Representative traces of tracks of averaged ChIP-seq data for the retinal progenitor gene *Uhrf1*, which is silenced during retinal development (E). DNA-methylation data are available for all stages but only the representative E14.5 and P21 DNA-methylation data are presented here. β-value of 1.0 indicates the site is fully methylated, 0 indicates it is unmethylated. The normalized relative gene expression from RNA-seq analysis is shown in the histograms to the right of each ChIP-seq plot. Some retinal progenitor cell genes (e.g., *Ascl1*) are repressed by H3K27me3 and are silenced during retinal differentiation (F). A representative series of ChIP-seq traces is shown for the rod gene *Stx3*, which is upregulated as rods differentiate (G). Individual marks and their scales in the plots are indicated in different colors at the bottom of the figure.

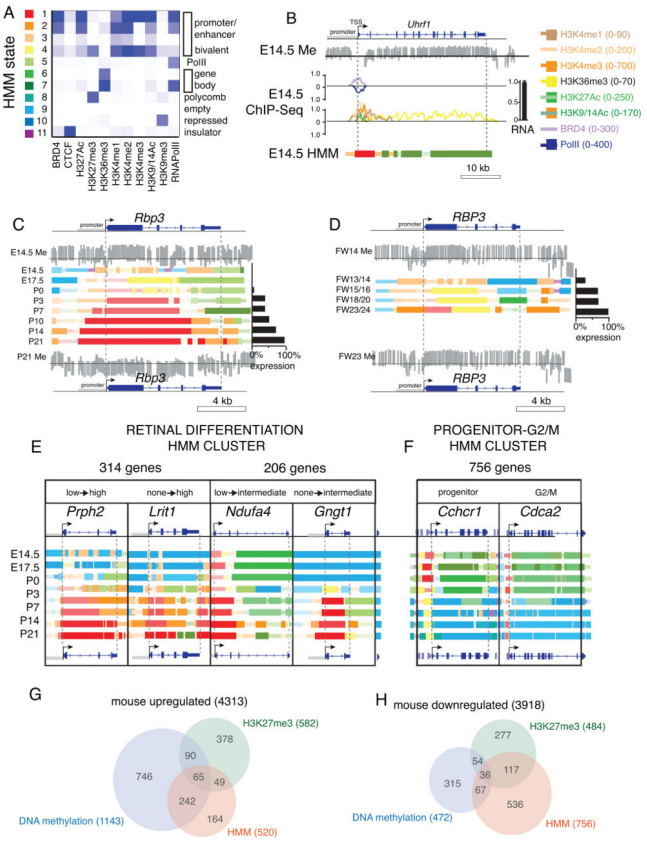

Figure 3. Hidden Markov Modeling of the Retinal Epigenome

(A) Heat map of the 11 chromHMM states used in this analysis. The darker blue represents more abundance of that ChIP-seq mark in the particular HMM state.

(B) Each state is color coded and then used to represent the chromHMM states for the retinal progenitor gene *Uhrf1*. The 2 states that are the most abundant are full-height bars, and the remaining 9 states are half the height. The intensity of each bar is proportional to the percentage of each state across all stages for that gene. For the bars that are half the height, the intensity is scaled starting at 50% of maximum intensity. Normalized relative-fold expression of *Uhrf1* is shown in the histogram on the right, and the relative β-value for DNA methylation is shown for E14.5. β-value of 1.0 means the site is fully methylated, and 0 is unmethylated. The ChIP-seq traces are shown as a reference for the chromHMM states.

(C, D) HHM states, DNA methylation, and gene expression of a representative rod-differentiation gene during mouse (C) or human (D) retinal development. The human retinal stages are aligned with the corresponding murine stages based on comprehensive developmental, morphological, and neuroanatomical modeling across mammalian species (http://www.translatingtime.net).

(E) Representative chromHMM modeling and DNA-methylation data of the 4 distinct clusters of rod genes identified by the chromHMM data analysis. The clusters were associated with expression during development, as indicated.

(F) No distinct subclusters of progenitor or G2/M genes were identified by the chromHMM data analysis. Representative chromHMM and DNA-methylation data are shown for a progenitor gene and a G2/M gene.
(G, H) Venn diagrams of the overlap in genes that have corresponding changes in DNA methylation, chromHMM, and H3K27me3 for the up and downregulated genes in murine retinogenesis.

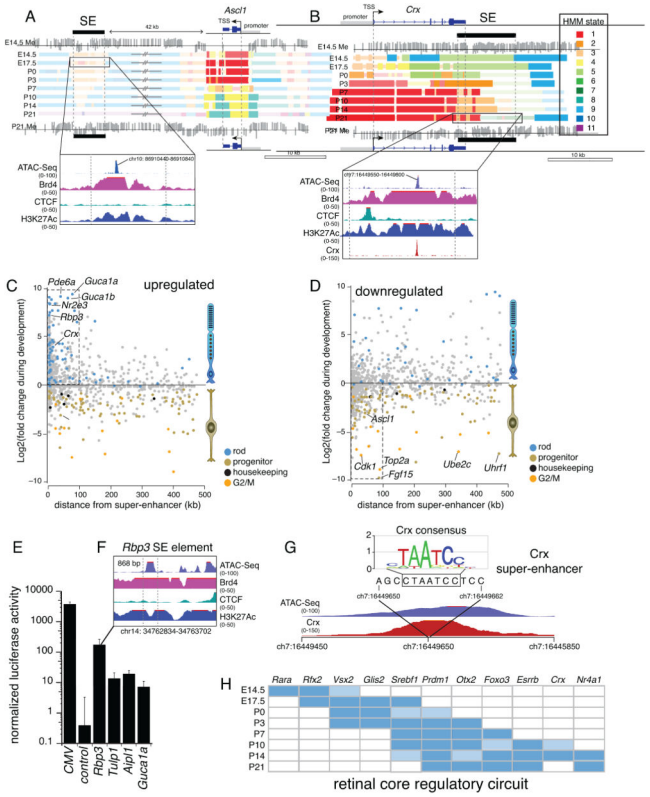

Figure 4. Developmentally Regulated Super-Enhancers during Retinal Development

(A) ChromHMM and DNA methylation of a super-enhancer (SE) 40 kb from the developmentally regulated progenitor gene *Ascl1*.

(B) ChromHMM and DNA methylation of a representative developmentally regulated photoreceptor gene (*Crx*). The super-enhancers are indicated by black boxes. Magnified views of the ATAC-seq and ChIP-seq tracks for Brd4, CTCF, and H3K27Ac are shown in the corresponding boxes.

(C–D) Scatterplot of all super-enhancers in the mouse genome showing chromHMM-state transitions during development, consistent with progenitor cells (i.e., active early and inactive later in development) and with differentiating rods (i.e., those that are inactive early and active later in development). The closest gene within 500 kb was identified for each super-enhancer, and the relative change in expression during development was plotted, relative to the distance from the super-enhancer. Rod, progenitor, G2/M, and housekeeping genes adjacent to super-enhancers are indicated with corresponding colors. Representative progenitor cell and G2/M (C) or rod genes (D) are indicated. The dashed box highlights the super-enhancer/gene pairs that are downregulated during development and fall within 100 kb.

(E) Histogram of the normalized luciferase activity of individual regions of super-enhancers promixal to the indicated genes. Each bar represents the mean (±SD) of 5 biological replicates. CMV=strong CMV enhancer, positive control; control= region devoid of H3K27Ac in the retina, negative control.

(F) Magnified views of the ATAC-seq and ChIP-seq tracks for Brd4, CTCF, and H3K27Ac for the 868-bp region of the *Rbp3* super-enhancer tested in the luciferase assay in (E).

(G) Crx consensus–binding site within and overlapping with the ATAC-seq peak of the *Crx* super-enhancer shown in (B).
(H) Representative core regulatory circuit transcription factors across murine retinal development (blue boxes). Darker boxes indicate transcription factors identified in both biological replicates, lighter boxes are those found in 1 replicate.

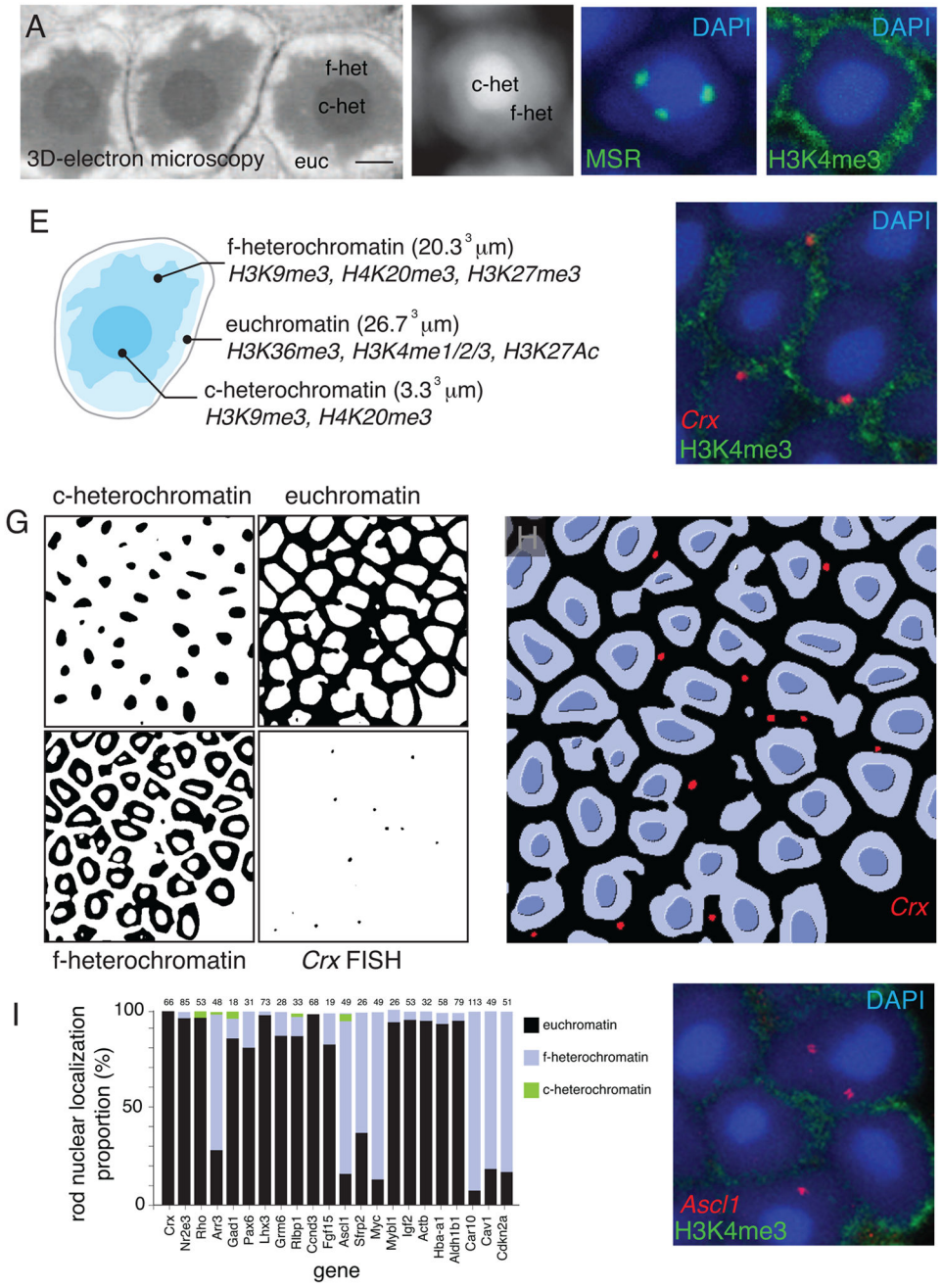

Figure 5. 3D Localization of Genes in Rod Nuclei

(A) Representative electron micrograph of a rod nucleus in the adult mouse retina showing the 3 concentric regions of chromatin organization.

(B) DAPI-stained rod nucleus showing the 3 concentric regions of chromatin organization.

(C) FISH of the microsatellite repeat (MSR) region at the boundary of the heterochromatin domains.

(D) Immunofluorescence of H3K4me3 at the promoters of actively transcribed genes in the euchromatin domain.

(E) 3D reconstruction of rod nucleus showing the corresponding volumes of the 2 heterochromatin domains and the euchromatin domain.
(F) FISH of the *Crx* gene (arrows) in the euchromatin domain overlapping with H3K4me3.
(G, H) Automated segmentation, 3D reconstruction, and localization of *Crx* FISH signal (red) to the euchromatin domain (arrows).
(I) Histogram showing the proportion of nuclei with each gene in the euchromatin or f-heterochromatin domains.
(J) Micrograph showing FISH (arrows) for the *Ascl1* gene in the f-heterochromatin domain that is exclusive from H3K4me3. Scale bars: 1 μm. Abbreviations: c-het, constitutive heterochromatin; euc, euchromatin; f-het, facultative heterochromatin

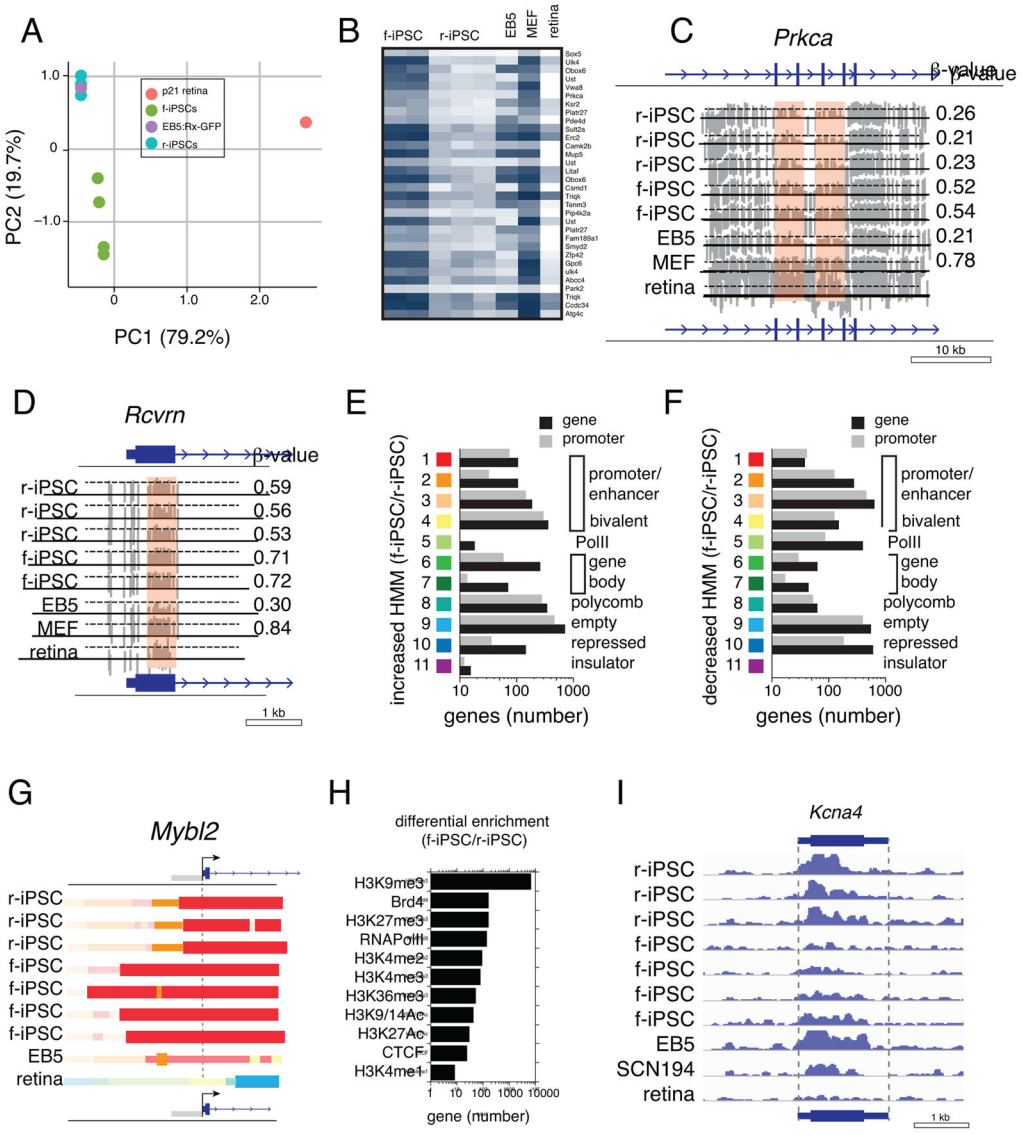

Figure 6. Epigenetic Memory in r-iPSCs

(A) Principal component analysis plot of RNA-seq data for adult mouse retina (p21 retina), 4 f-iPSC lines, 3 r-iPSC lines, and the EB5:Rx–GFP stem cell line.

(B) Heatmap of select genes with differential DNA methylation that also had changes in DNA methylation and expression during retinal development. The blue shading is higher β-values for the DNA methylation, and the white shading is lower values.

(C, D) Example of DNA methylation from WGBS for the coding region of the *Prkca* and the *Rcvrn* genes. Corresponding β-values are indicated for each sample.

(E, F) Histograms of genes and promoters with differential chromHMM states that are increased (E) or decreased (F) in f-iPSCs relative to r-iPSCs.

(G) Representative chromHMM map for *Mybl2* showing an extension of state 1 in the f-iPSCs relative to r-iPSCs.

(H) Histogram of combined genes and promoters with differences in ChIP-seq data between f-iPSCs and r-iPSCs.

(I) Example of ChIP-seq of H3K9me3 from the *Kcna4* gene with differential abundance in f-iPSCs vs r-iPSCs.

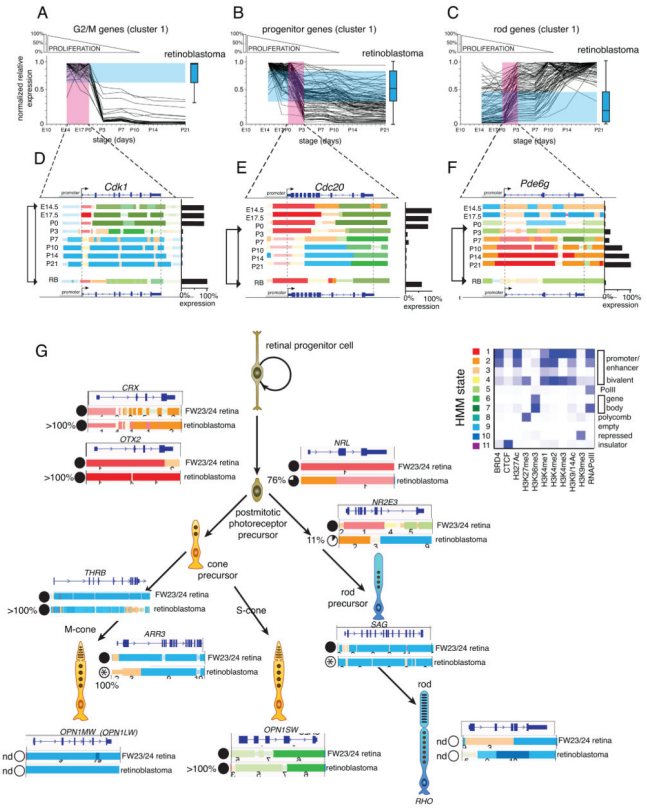

Figure 7. Photoreceptor Pathway in Human Retinoblastoma

(A–C) Changes in expression of G2/M, progenitor, and rod genes during mouse retinal development. Boxplot of expression of the same genes in retinoblastoma is shown to the right, and the intersection of those genes (blue-shaded region) with retinal development (pink-shaded region) is indicated (purple-shaded region).

(D–F) Representative chromHMM, DNA methylation, and gene expression of a G2/M gene, a progenitor gene, and a rod gene are shown.

(G) Simplified transcriptional network that regulates photoreceptor development in the human retina. The chromHMM states are indicated for each gene, and the expression is represented in individual pie charts next to each gene for normal human retina at FW 23/24 and retinoblastoma. Retinoblastoma gene expression is normalized to that in healthy human retina. Abbreviations: nd, none detected; * FPKM less than 10 indicates very low levels of expression in the normal retina and tumor.