

# Chapter 1

## Cell Type-Specific Epigenetic Signatures Accompany Late Stages of Mouse Retina Development

Evgenya Y. Popova, Colin J. Barnstable and Samuel Shao-Min Zhang

**Abstract** We have used ChIP-seq to map the distribution of two important histone H3 modifications, H3K4me2 and H3K27me3, over the whole genome at multiple time points during late mouse retina development. We merged these data with our previous retina developmental expression profiles and show that there are several epigenetic signatures specific for different functional groups of genes. The main conclusion from our study is that epigenetic signatures defined by H3K4me2 and H3K27me3 can distinguish cell-type specific genes from widespread transcripts and may be reflective of cell specificity during retina maturation. Rod photoreceptor-specific genes have a striking signature, a de novo accumulation of H3K4me2 and a complete absence of H3K27me3. We were able to use this signature in an unbiased search of the whole genome and identified essentially all the known rod photoreceptor genes as well as a group of novel genes that have a high probability of being rod photoreceptor specific. Comparison of our genome-wide chromatin signature maps with available data sets for Polymerase-II (Pol-II) and CRX binding sites and DNaseI Hypersensitive Sites (DHS) for retina shows great agreement. Because our approach is not dependent on high levels of gene expression, it provides a new way of identifying cell type-specific genes, particularly genes that may be involved in retinal diseases.

**Keywords** Retina development · Epigenetic · Chromatin · Histone

The retina begins as an early compartment of the forebrain and has frequently served as a model of CNS development [1, 2]. Retinal cell types are formed in a characteristic sequence from E12 to PN5 with ganglion cells, amacrine cells, and

---

C. J. Barnstable (✉) · E. Y. Popova · S. Shao-Min Zhang  
Department of Neural and Behavioral Sciences, Pennsylvania State University College of Medicine, Hershey, PA 17033, USA  
e-mail: cbarnstable@hmc.psu.edu

E. Y. Popova  
e-mail: eyp1@psu.edu

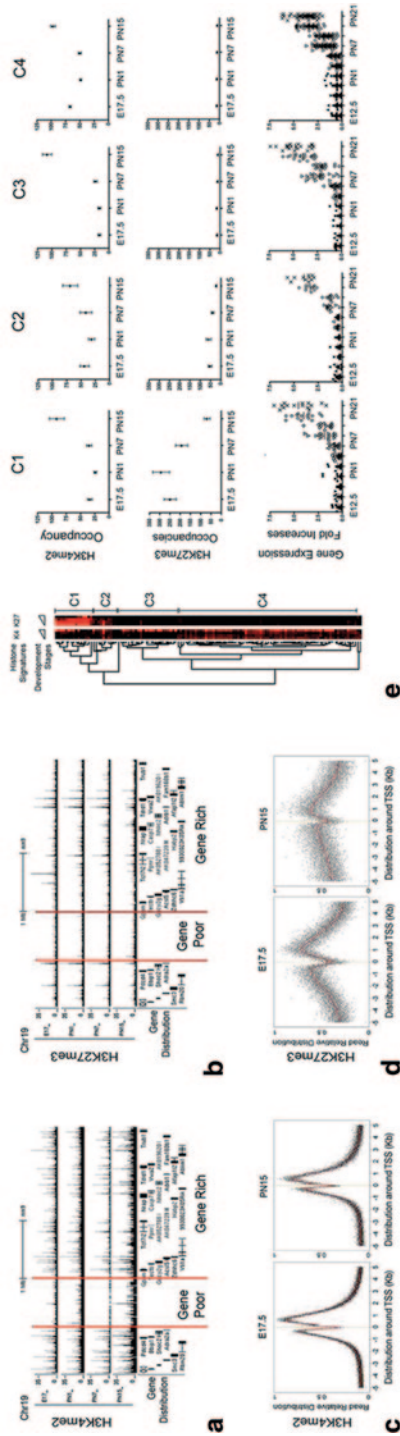
S. Shao-Min Zhang  
e-mail: ssz3@psu.edu

horizontal cells among the early formed types and rod photoreceptors and bipolar cells formed predominantly during the later postnatal period. Control of retinogenesis is a complex process involving changing levels of expression regulation of large groups of genes. Change in chromatin structure around promoters and genes is an important factor in tissue specification, but knowledge about this epigenetic contribution to retina development is just beginning to emerge. Previously, studies have focused on chromatin changes in promoter areas of individual genes, for example rhodopsin [3, 4]. New genome-wide technologies provide the opportunity to study simultaneously developmental changes in large groups of genes during retinogenesis and to better understand how tissue-specific gene expression is established and maintained in retina. Here we have combined and reviewed both our and several other recently published genome-wide data sets for retina development. We have created custom tracks in the UCSC Mouse Genome Browser by using data from our study for H3K4me2 ChIP-Seq [5], Pol-II ChIP-on-Chip data from [6], data for CRX ChIP-Seq [7], and ENCODE project tracks for retina's DHS (University of Washington). This approach provides a new way of identifying cell type-specific genes, particularly genes that may be involved in retinal diseases.

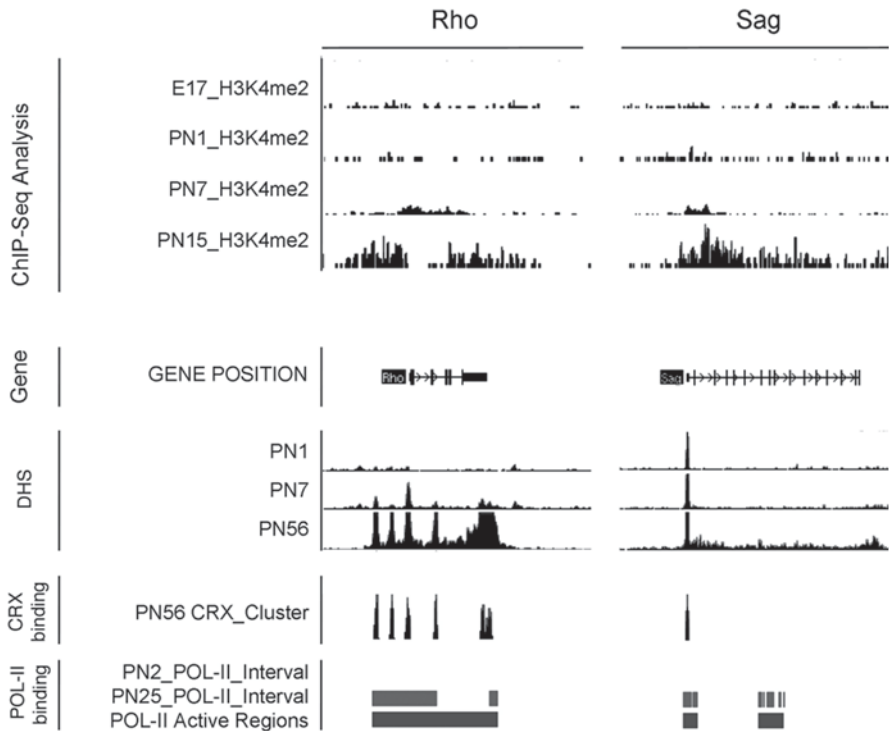
We generated a comprehensive genomic map of H3K4me2 and H3K27me3 using ChIP-seq on mouse retinas of various ages [5]. When the H3K4me2 and H3K27me3 reads were mapped to the genome it was clear that most of these two-histone modifications were in gene-rich regions of the genome (Fig. 1.1a, b). The accumulation of H3K4me2 reads in the genome were primarily localized in the area around the transcription start site (TSS), showing two sharp peaks (Fig. 1.1c) of enrichment approximately  $\pm 1$  Kb surrounding the TSS as defined for the 25,158 genes from the NCBI RefSeq database, including splice variants and alternative TSS. The peaks of H3K27me3 were less pronounced and were even less distinct at PN15 than at E17.5 (Fig. 1.1d). A small set of genes from this large collection was used for confirmation by ChIP-qPCR analysis and the results perfectly matched those of the ChIP-Seq experiments. Analysis of these data has revealed a number of important and surprising results.

*First*, genes that are never expressed in retina, such as erythrocyte-specific globin or olfactory receptor genes, showed no accumulation of either H3K4me2 or H3K27me3. This appears to be a good diagnostic criterion for whether a gene is ever expressed in retina. Interestingly, when we examined available databases of other cell types (ChIP-Seq analysis of C2C12 myogenic cell line, LICR Histone Track, UCSC Genome Informatics) that do not express these genes we found a similar lack of both H3K4me2 and H3K27me3, suggesting that this criterion may be general.

*Second*, we found different histone signatures for retinal genes with the same gene expression pattern. In our earlier developmental gene expression studies we had identified a group of 123 genes whose expression increased in parallel with maturation of rod photoreceptors, and a pool of 119 genes whose expression decreased between E16.5 and PN15. When we carried out a hierarchical cluster analysis on the upregulated genes they fell into four distinct clusters, each with its own epigenetic signature (Fig. 1.1e). Similarly the downregulated genes fell into three distinct clusters.



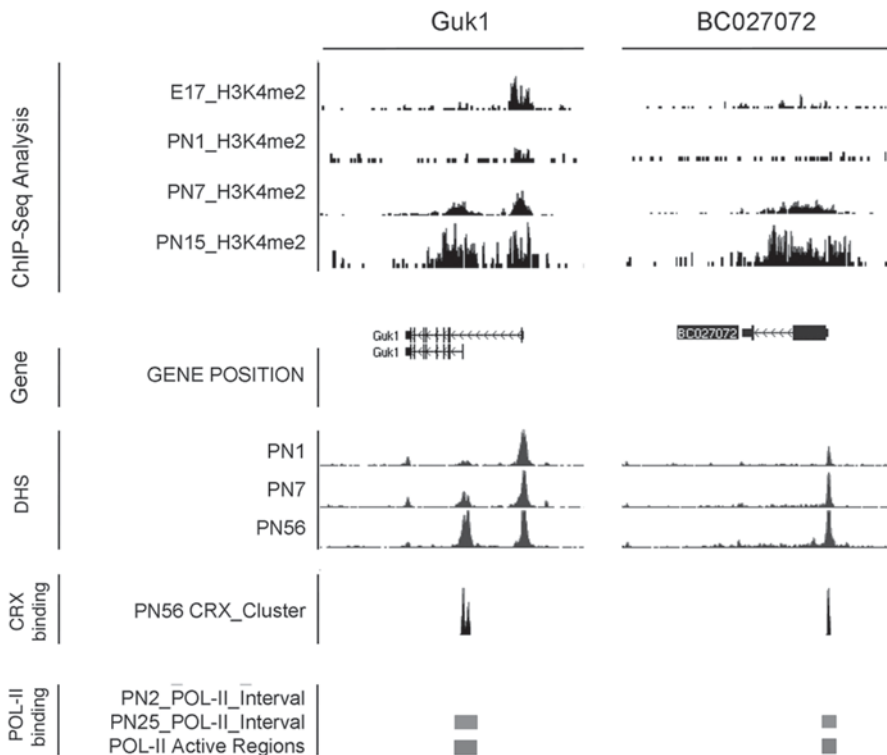
**Fig. 1.1** H3K4me2 and H3K27me3 profiles during retina organogenesis. **a, b** Histone modifications patterns at part of chromosome 19. Peaks of normalized sequenced tags from ChIP-Seq analysis of mouse retina at four developmental stages. **c, d** Histone modifications for all NCBI RefSeq genes around TSS. **e** Cluster analysis of H3K4me2 and H3K27me3 occupancy around TSS ( $\pm 2.5$  Kb) for genes upregulated in mature retina. Tree-view shows 4 clusters (C1 —C4) with distinct epigenetic patterns for H3K4me2 (*upper panel*) and H3K27me3 (*middle panel*), but with same expression patterns (*lower panel*). Adapted from [5]



**Fig. 1.2** Combine genome-wide tracks of H3K4me2 ChIP-Seq analysis, DNase1 Hypersensitive Sites (DHS) and CRX- and Pol-II-binding sites for known rod-specific genes

*Third*, we found that a unique signature of H3K4me2 and H3K27me3 marks rod photoreceptor-specific genes. Our initial analysis identified a cluster that contained a group of known rod-specific genes. The signature of this cluster (C3 on Fig. 1.1e) was a de novo increase in H3K4me2 and no H3K27me3 at any stage. To test whether this was a signature for rod-specific genes we did an unbiased cluster analysis of all 25,158 genes in the RefSeq database. This analysis identified 107 genes with the “rod signature”. This cluster contained all known rod-specific genes, including *rho*, *Nrl*, and *Nr2E3*, mouse orthologues of recently described disease genes such as *BC027072*, and genes such as *Ppef2*, *Lrit1*, and *Lrit2* that have been implicated in phototransduction but whose exact function is unknown. Many other genes in this cluster are predominantly expressed in the retina (data sets from [8]; Stanford SOURCE; MGI from Jackson Lab), but exact localization has yet to be carried out. Since the epigenetic signatures are unrelated to the level of gene expression, we have suggested that they may be a novel way of identifying cell type-specific genes particularly those expressed at low levels.

*Fourth*, the accumulation of H3K4me2 at rod-specific genes was over the whole gene, not just the TSS, as indicated for the example genes in Fig. 1.2. We have suggested that this histone modification is involved in the maintenance of rod gene expression not its initial onset. This is in agreement with other recent genome-wide



**Fig. 1.3** Combine genome-wide tracks of H3K4me2 ChIP-Seq analysis, DNase1 Hypersensitive Sites (DHS), and CRX- and Pol-II-binding sites for newly predicted rod-specific genes

studies, that genes specially expressed in blood cells displayed high levels of H3K-4me2 over the whole gene bodies [9, 10].

The conclusion from our study is that epigenetic signatures defined by H3K-4me2 and H3K27me3 can distinguish cell type-specific genes from widespread transcripts and may be reflective of cell specificity during retina maturation.

We then compared our genome-wide chromatin signature maps with available data sets for retina for ChIP-on-Chip study of Polymerase-II (Pol-II) binding sites [6], ChIP-Seq study for CRX binding sites [7], and DHS (University of Washington data, available on UCSC mouse Genome Browser as a part of ENCODE project) with a focus on the cluster of genes with a “rod signature”. In this cluster of 107 genes, 95 genes (or 89%) have CRX binding site around the TSS and gene, with average ~2 CRX binding sites per gene. ENCODE project retina’s DHS tracks are available for three developmental stages: PN1, PN7, PN56. During development DHS could spread widely, for example at the promoter or TSS of developmentally upregulated genes, or could disappear at the TSS and promoter of downregulated genes. We monitored developmentally upregulated DHS for our rod-specific cluster genes and 105 of them (or 98%) have such DHS around gene or its promoter with average three upregulated DHS per gene. Pol-II binding data is available for retina

at PN2 and PN25 developmental stages with a list of genes that have ratio of Pol-II PN25/ Pol-II PN2 more than 1.8. 67 genes (or 63 %) from our cluster of rod-specific genes are in this list with the average level of Pol-II PN25/Pol-II PN2 around 4.2. When compared, combination of tracks (CRX and Pol-II binding sites, DHS and H3K4me2 accumulation) for known rod-specific genes (Fig. 1.2) looks very similar to the tracks of newly predicted genes (Fig. 1.3).

The agreement between our whole-genome data and other data sets shows the predictive power of epigenetic signatures and the importance of studying changes in the epigenome and chromatin structure of promoters and genes during retina development. Rod photoreceptors also provide us with an almost unique model to address the question of whether epigenetic changes are controlled by the same transcription factors that regulate gene expression or by some as yet unknown mechanism.

**Acknowledgment** We thank K.P. Mitton for providing an access to genome tracks for the Pol-II binding sites. This work was supported by NIH grant EY013865 and Macular Vision Research Foundation to C.J.B.

**Funding/Support:** Pfizer Ophthalmology External Research Unit, The Foundation Fighting Blindness CDA (MEP), Research to Prevent Blindness (Unrestricted grant to Casey Eye Institute, CDA to MEP), NIH grant P51OD011092 (MN), K08 EY021186-01 (MEP).

## References

1. Barnstable CJ (1987) A molecular view of vertebrate retinal development. *Mol Neurobiol* 1:9–46
2. Dyer MA, Cepko CL (2001) Regulating proliferation during retinal development. *Nat Rev Neurosci* 2:333–342
3. Peng GH, Chen S (2007) Crx activates opsin transcription by recruiting HAT-containing co-activators and promoting histone acetylation. *Hum Mol Genet* 16:2433–2452
4. Peng GH, Chen S (2011) Active opsin loci adopt intrachromosomal loops that depend on the photoreceptor transcription factor network. *Proc Natl Acad Sci U S A* 108:17821–17826
5. Popova EY, Xu X, Dewan AT, Salzberg AC, Berg A, Hoh J, Zhang SS, Barnstable CJ (2012) Stage and gene specific signatures defined by histones H3K4me2 and H3K27me3 accompany mammalian retina maturation in vivo. *PLoS One* 7:e46867
6. Tummala P, Mali RS, Guzman E, Zhang X, Mitton KP (2010) Temporal ChIP-on-Chip of RNA-polymerase-II to detect novel gene activation events during photoreceptor maturation. *Mol Vis* 16:252–271
7. Corbo JC, Lawrence KA, Karlstetter M, Myers CA, Abdelaziz M, Dirkes W, Weigelt K, Seifert M, Benes V, Fritsche LG et al (2010) CRX ChIP-seq reveals the cis-regulatory architecture of mouse photoreceptors. *Genome Res* 20:1512–1525
8. Su AI, Cooke MP, Ching KA, Hakak Y, Walker JR, Wiltshire T, Orth AP, Vega RG, Sapinoso LM, Moqrich A et al. (2002) Large-scale analysis of the human and mouse transcriptomes. *Proc Natl Acad Sci U S A* 99:4465–4470
9. Wong P, Hattangadi SM, Cheng AW, Frampton GM, Young RA, Lodish HF (2011) Gene induction and repression during terminal erythropoiesis are mediated by distinct epigenetic changes. *Blood* 118:e128–138
10. Pekowska A, Benoukraf T, Ferrier P, Spicuglia S (2010) A unique H3K4me2 profile marks tissue-specific gene regulation. *Genome Res* 20:1493–1502