

Spring 5-15-2018

# Deciphering mechanisms governing the development of the rod epigenome

Philip Andrew Ruzycski  
*Washington University in St. Louis*

Follow this and additional works at: [https://openscholarship.wustl.edu/art\\_sci\\_etds](https://openscholarship.wustl.edu/art_sci_etds)



Part of the [Genetics Commons](#), and the [Ophthalmology Commons](#)

---

## Recommended Citation

Ruzycski, Philip Andrew, "Deciphering mechanisms governing the development of the rod epigenome" (2018). *Arts & Sciences Electronic Theses and Dissertations*. 1573.  
[https://openscholarship.wustl.edu/art\\_sci\\_etds/1573](https://openscholarship.wustl.edu/art_sci_etds/1573)

This Dissertation is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact [digital@wumail.wustl.edu](mailto:digital@wumail.wustl.edu).

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences  
Molecular Genetics and Genomics

Dissertation Examination Committee:

Shiming Chen, Chair

Joseph Corbo

Kristen Kroll

Eugene Oltz

Ting Wang

Deciphering mechanisms governing the development of the rod epigenome

by

Philip A. Ruzycki

A dissertation presented to  
The Graduate School  
of Washington University in  
partial fulfillment of the  
requirements for the degree  
of Doctor of Philosophy

May 2018  
St. Louis, Missouri



© 2018, Philip A. Ruzycki

# Table of Contents

List of Figures .....	v
List of Tables .....	viii
List of Abbreviations .....	ix
Acknowledgements.....	xi
Abstract.....	xiv
Chapter 1: Introduction.....	1
1.1    The Genome and Its Regulation.....	2
1.1.1    Transcription Factors Control Gene Expression .....	2
1.2    Organizing the Genome.....	4
1.2.1    Nucleosomes and modifications.....	4
1.2.2    Genome organization beyond the nucleosome.....	6
1.3    The retina as model system .....	9
1.3.1    Retinal transcription factors and disease .....	10
1.3.2    CRX.....	11
1.3.3    The photoreceptor epigenome .....	13
1.3.4    Photoreceptor Genomic Organization .....	13
1.4    References .....	18
Chapter 2: <i>Crx-L253X</i> Mutation Produces Dominant Photoreceptor Defects in <i>TVRM65</i> Mice.....	26
2.1    Author Contributions.....	27
2.2    Abstract .....	28
2.3    Introduction .....	29
2.4    Results .....	32
2.4.1 <i>L253X</i> overproduces mutant mRNA/protein in affected retinas .....	32
2.4.2 <i>L253X/X</i> mice have no detectable rod and cone function .....	33
2.4.3 <i>L253X/X</i> photoreceptors degenerate earlier than <i>Crx</i> <sup>-/-</sup> .....	33
2.4.4 <i>L235X/X</i> mice exhibit photoreceptor gene misregulation distinct from <i>Crx</i> <sup>-/-</sup> .....	34
2.4.5    Heterozygous <i>L253X/+</i> mice show mild decreases in rod and cone function .....	36
2.4.6 <i>L253X/+</i> retinas do not degenerate .....	36
2.4.7 <i>L253X/+</i> retinas display dynamic changes of photoreceptor gene expression .....	37
2.4.8    Late PTC-caused C-terminal truncation reduces CRX's transactivation function.....	38
2.5    Discussion .....	39
2.5.1 <i>L253X</i> mouse provides a new Class III model for mild dominant cone-rod dystrophy .....	39
2.5.2    The molecular mechanism(s) underlying <i>L253X</i> pathogenicity.....	40
2.5.3 <i>L253X</i> allele-specific overproduction of mutant mRNA and protein has implications for the PTC position effect .....	41
2.6    Supporting Information .....	55
2.7    Acknowledgements .....	55
2.8    References .....	55

Chapter 3: Graded gene expression changes determine phenotype severity in mouse	
models of <i>CRX</i> -associated retinopathies.....	57
3.1 Author Contributions.....	58
3.2 Abstract .....	59
3.3 Introduction .....	60
3.4 Results .....	63
3.4.1 <i>Crx</i> mutations cause graded expression changes in shared gene sets, correlating with phenotype severity .....	63
3.4.2 <i>Crx</i> mutations specifically affect rod- and cone-enriched genes including down-regulation of phototransduction genes .....	67
3.4.3 Insights from <i>E168d2/+</i> and <i>E168d2neo/+</i> : small changes in gene expression level strongly influence phenotype.....	68
3.4.4 Heterozygous mutants show normal gene expression trends from P10 to P21, but many genes fail to reach the normal level at P21 .....	69
3.4.5 Down-regulated and up-regulated genes in <i>Crx</i> mutants show distinct epigenetic profiles in <i>WT</i> retinas .....	70
3.4.6 Up-regulated genes are characteristic of cone photoreceptors, likely resulting from de-repression in rods.....	76
3.4.7 Testing the effect of gene expression changes on light dependent degeneration.....	80
3.5 Discussion .....	84
3.5.1 Gene expression in animal models for <i>CRX</i> -associated retinopathies .....	84
3.5.2 Different modalities for down-regulated and up-regulated gene expression .....	86
3.5.3 <i>Crx</i> mutations affect rod and cone development and transcriptional integrity .....	87
3.5.4 Gene expression dictates phenotypic thresholds .....	88
3.5.5 Conclusions .....	90
3.6 Supporting Information .....	116
3.7 Data Availability .....	116
3.8 Acknowledgements .....	116
3.9 References .....	117
Chapter 4: Meta-analysis of epigenetic remodeling reveals <i>CRX</i> 's mechanism of action in retinal development.....	120
4.1 Author Contributions.....	121
4.2 Abstract .....	122
4.3 Introduction .....	123
4.4 Results .....	127
4.4.1 <i>CRX</i> binds ATAC sensitive regulatory sites.....	127
4.4.2 <i>Crx</i> <sup>-/-</sup> retinas have an altered photoreceptor epigenome.....	128
4.4.3 <i>Crx</i> <sup>-/-</sup> photoreceptors fail to close and open developmentally-modulated regulatory sites .....	129
4.4.4 <i>CRX</i> binding is strongly enriched at sites that lose activity in <i>Crx</i> <sup>-/-</sup> .....	130
4.4.5 <i>CRX</i> acts within a variety of chromatin environments .....	131
4.4.6 <i>CRX</i> is required to activate a subset of local regulatory regions .....	132
4.4.7 <i>CRX</i> is required to activate a set of distal enhancers .....	133

4.4.8	CRX controls distal regulatory sites in the absence of active histone marks .....	134
4.4.9	Base composition and conservation differentiate CRX sites .....	134
4.4.10	CRX has different affinity for Dependent vs Independent sites.....	135
4.4.11	Other TFs may compensate or be more influential at Independent sites .....	136
4.5	Discussion .....	138
4.6	References .....	163
Chapter 5:	Organization of the photoreceptor genome.....	166
5.1	Author Contributions.....	167
5.2	Abstract .....	168
5.3	Introduction .....	170
5.4	Results .....	173
5.4.1	<i>Rho</i> displays strong interactions with local stretch of ‘A’ compartment DNA.....	173
5.4.2	Distal interactions are dependent on <i>Rho</i> activation .....	174
5.4.3	Distal interactions enriched for highly active chromatin .....	176
5.4.3	Inactivation of <i>Rho</i> affects expression of other genes.....	177
5.5	Discussion .....	179
5.5.1	General principles of genomic organization.....	179
5.5.2	<i>Rho</i> interacts with other active regions of its chromosome.....	180
5.5.3	Promoters do not act independently .....	180
5.6	References .....	193
Chapter 6:	Conclusions and Future Directions .....	196
6.1	A cohesive understanding of CRX related diseases.....	197
6.2	A comprehensive genomic map of CRX regulatory activity .....	198
6.3	The rod photoreceptor as a model for genomic organization.....	201
6.4	Conclusions .....	203
6.5	References .....	204
Appendix 1:	The role of the <i>Rhodopsin</i> enhancers.....	205
Appendix 2:	Methods.....	211

# List of Figures

## Chapter 1

Figure 1.1:	Humans and mice have more non-coding genomic sequence.....	17
-------------	--	----

## Chapter 2

Figure 2.1:	Class III <i>Crx</i> mutations introduce a premature termination codon (PTC), resulting in accumulation of truncated CRX proteins and mutant mRNA with variable extended 3'UTR.....	45
Figure 2.2:	<i>L253X/X</i> retinas lack rod and cone responses to light and degenerate earlier than <i>Crx</i> <sup>-/-</sup> .....	46
Figure 2.3:	<i>L253X/X</i> has measurable gene expression differences relative to <i>Crx</i> <sup>-/-</sup> .....	47
Figure 2.4:	<i>L253X/+</i> retinas exhibit reduced rod and cone light responses.....	48
Figure 2.5:	<i>L253X/+</i> retinas appear morphologically similar to <i>WT</i> up to 3 mo.....	49
Figure 2.6:	<i>L253X/+</i> retinas display dynamic changes in photoreceptor gene expression.....	50
Figure 2.7:	<i>L253X</i> retains minimal ability to transactivate Rhodopsin promoter in presence of NRL.....	51
Supplemental Figure 2.1:	<i>L253X/+</i> retinas display variable excess total <i>Crx</i> mRNA .....	52
Supplemental Figure 2.2:	<i>L253X/X</i> retinas display normal ONL genesis but photoreceptor layer-specific progressive thinning, while all <i>L253X/+</i> retinal layers unaffected.....	53

## Chapter 3

Figure 3.1:	Schematic diagram of WT and mutant CRX proteins made by the indicated mouse models .....	93
Figure 3.2:	RNA-seq analyses describe overlapping sets of affected genes in P10 homozygous <i>Crx</i> mutant retinas.....	94
Figure 3.3:	RNA-seq analyses detect graded changes in gene expression in P10 heterozygous <i>Crx</i> mutant retinas.....	95
Figure 3.4:	<i>Crx</i> mutant retinas show graded changes in photoreceptor-specific gene expression .....	96
Figure 3.5:	Hierarchical clustering and epigenetic data reveal groups of similarly-regulated genes.....	97
Figure 3.6:	<i>Crx</i> mutants lose rod gene expression but up-regulate many phototransduction-unrelated cone gene transcripts .....	98
Figure 3.7:	Model of how <i>Crx</i> mutation-caused gene expression changes affect rod and cone development .....	99
Figure 3.8:	Light-dependent photoreceptor degeneration in <i>E168d2neo/+</i> mice.....	100
Supplemental Figure 3.1:	P10 and P21 biological replicates clustered by Principal Component	

	Analysis .....	102
Supplemental Figure 3.2:	Browser images display RNAseq read depth of rod and cone transcripts .....	103
Supplemental Figure 3.3:	Homozygous and heterozygous mutants show largely overlapping datasets .....	104
Supplemental Figure 3.4:	<i>Crx</i> mutants show little change to other non-photoreceptor phototransduction components .....	105
Supplemental Figure 3.5:	Biological replicates show consistency in expression changes between genotypes .....	106
Supplemental Figure 3.6:	<i>Crx</i> mutants show graded changes in gene expression of overlapping gene sets at P21 .....	107
Supplemental Figure 3.7:	<i>Crx</i> mutants do not abandon developmental program, but many genes fail to reach proper expression levels .....	108
Supplemental Figure 3.8:	Biological replicates show consistent expression changes within each genotype .....	109
Supplemental Figure 3.9:	Quantification of normal epigenetic marks near genes in Groups 1, 2, 3, and 6 distinguishes up- and down-regulated genes in <i>Crx</i> mutants .....	110
Supplemental Figure 3.10:	Loss of expression of key components of phototransduction .....	111
Supplemental Figure 3.11:	Changes in rod and cone enriched transcription factors are consistent with general patterns of rod and cone genes expression in CRX mutants .....	112
Supplemental Figure 3.12:	Retinal function is affected by light damage (LD) in <i>E168d2neo/+</i> but not <i>E168d2/+</i> mutant mice .....	113
Supplemental Figure 3.13:	<i>E168d2/+</i> photoreceptor degeneration is light-independent .....	114
Supplemental Figure 3.14:	Increased <i>Crx</i> expression in mutant lines at P10 .....	115

## Chapter 4

Figure 4.1:	CRX binds a subset of active regulatory sites in the rod photoreceptor .....	144
Figure 4.2:	Loss of CRX affects many developmentally activated regulatory sites .....	145
Figure 4.3:	CRX binds within four chromatin environments .....	146
Figure 4.4:	CRX is only required for activity and remodeling of a subset of local regulatory sites .....	147
Figure 4.5:	CRX activates a subset of distal enhancers over development .....	148
Figure 4.6:	Sequence analyses differentiate Dependent and Independent CRX sites .....	149
Figure 4.7:	Model of CRX mechanism of action .....	150
Supplemental Figure 4.1:	CRX bound ATAC peak signal changes correlate with expression changes of nearby genes .....	151
Supplemental Figure 4.2:	CRX binds in distinct regulatory environments .....	152
Supplemental Figure 4.3:	CRX is required to activate Dependent local regulatory elements and to remodel chromatin .....	153
Supplemental Figure 4.4:	Dependent and Independent sites show different chromatin	

	state dynamics .....	154
Supplemental Figure 4.5:	CRX is required to activate Dependent enhancer elements and remodel chromatin.....	155
Supplemental Figure 4.6:	CRX is required to activate Dependent distal regulatory elements.....	156
Supplemental Figure 4.7:	Sites display different presence of CRX motif and binding.....	157
Supplemental Figure 4.8:	TF motifs explain nature of Independent site activation in absence of CRX .....	158

## Chapter 5

Figure 5.1:	<i>Rho</i> displays strong interactions with local DNA in both <i>WT</i> and <i>Nrl</i> <sup>-/-</sup> retinas .....	182
Figure 5.2:	Disruption of the <i>Rho</i> promoter affects local epigenetic landscape .....	183
Figure 5.3:	FISH confirms <i>Rho</i> interactions with distal fragments of its chromosome.....	184
Figure 5.4:	<i>Rho</i> trans interactions are enriched at active segments of the genome .....	185
Figure 5.5:	Loss of <i>Rho</i> expression affects other genes that interact with <i>Rho</i> .....	186
Figure 5.6:	Model of <i>Rho</i> genomic interactions and effects of loss .....	187
Supplemental Figure 5.1:	4C detects interactions of <i>Rho</i> within its TAD and local region of A compartment .....	188
Supplemental Figure 5.2:	Morphometric analysis of <i>PPR</i> <sup>-/-</sup> and <i>RER</i> <sup>-/-</sup> retinas .....	189
Supplemental Figure 5.3:	BAC selection for FISH analysis .....	190
Supplemental Figure 5.4:	<i>Trans</i> interacting regions show very strong enrichment for active epigenetic state .....	191
Supplemental Figure 5.5:	A Compartment segments show differential effects upon loss of <i>Rho</i> activity .....	192

## Appendix 1

Figure A1.1	Epigenetic landscape of <i>Rhodopsin</i> .....	208
Figure A1.2	Disruption of <i>CBR1</i> affects <i>Rho</i> expression.....	209

# List of Tables

<b>Chapter 1</b>	
Table 1.1:	Summary of Chromosome Conformation Capture Techniques..... 16
<b>Chapter 2</b>	
Table 2.1:	Summary of <i>L253X</i> related phenotypes ..... 44
<b>Chapter 3</b>	
Table 3.1:	Phenotype summary of heterozygous <i>Crx</i> mutant mice ..... 91
Table 3.2:	List of RNA-seq experiments and number of genes with altered expression..... 92
<b>Chapter 4</b>	
Table 4.1:	ATAC-seq peaks lost in <i>Crx</i> <sup>-/-</sup> show enrichment for CRX binding in <i>WT</i> ..... 143
Supplemental Table 4.1:	GREAT analysis of CRX bound regulatory sites ..... 159
Supplemental Table 4.2:	De novo motif analysis identifies sequences enriched in different sets of CRX bound regions ..... 160-162
<b>Appendix 2</b>	
Table A2.1:	Number of FISH experiments complete for each probe set..... 229



# List of Abbreviations

3C	Chromosome conformation capture
4C	Circularized chromosome conformation capture
AD	Activation domain
ATAC-seq	Assay for transposase-accessible chromatin using sequencing
BAC	Bacterial artificial chromosome
ChIP-seq	Chromatin immunoprecipitation with high throughput sequencing
Chr	Chromosome
CpG	CG dinucleotide
CPM	Counts per million
DAPI	4',6-diamidino-2-phenylindole
DBD	DNA binding domain
ddPCR	Digital droplet polymerase chain reaction
DHS	Dnase I hypersensitive site
EMSA	Electrophoretic mobility shift assay
ERG	Electroretinogram
FC	Fold change
FDR	False discovery rate
FISH	Fluorescence <i>in situ</i> hybridization
GCL	Ganglion cell layer
GO	Gene ontology
H&E	Hematoxylin and eosin
HD	Homeodomain
HMM	Hidden Markov model
IHC	Immunohistochemistry
INL	Inner nuclear layer
IPL	Inner plexiform layer
LD	Light damage
mo	Months old
MPRA	Massively parallel reporter assay
ONL	Outer nuclear layer
OPL	Outer plexiform layer
OS	Outer segment
P0, P3, etc.	Post-natal day
PCA	Principal component analysis
PNA	Peanut agglutinin
PTC	Premature termination codon
qRT-PCR	Quantitative real-time polymerase chain reaction
RNA-seq	RNA sequencing
RPE	Retinal pigmented epithelium
SE	Super enhancer

SEM	Standard error of the mean
TAD	Topologically associating domain
TF	Transcription factor
TSS	Transcription start site
UTR	Untranslated region
wk	Weeks old

# Acknowledgments

I am incredibly fortunate to have been mentored during graduate school by Dr. Shiming Chen. Shiming combines many of the best attributes one could hope for in a mentor- she is insightful, caring, and deeply interested in the scientific process. During my time in the lab she has expertly been present but has provided me enough room to learn and explore on my own. I hope to live up to her high example in my future scientific endeavors.

I am very grateful for the help and guidance I have received from the rest of the Chen Lab including present members Xiaodong Zhang, Mingyan Yang, Inez Oh, and Anne Hennig and past members Nick Tran and Diana Brightman. I am also particularly grateful to Courtney Linne, a former Washington University undergraduate, who spent countless hours working with me to finish many of the experiments described in this dissertation.

I am grateful for all of the outstanding guidance I have received from many past mentors and friends, many of whom I can gladly categorize as both. I would specifically like to thank Mark Petrash who gave an unseasoned undergraduate a chance in his lab - and convinced me to re-route toward graduate school. The Department of Ophthalmology and Visual Sciences has also been a very welcoming and collaborative environment. I want to specifically thank Peter Lukasiewicz, Vladimir Kefalov, Didier Hodzik, and Tom Ferguson for treating all of the students as peers, and never hesitating to share advice or a beer.

The Molecular Genetics and Genomics program, including students and faculty have been a strong support network throughout graduate school. I am specifically thankful for the constructive discussions I have had with Mike White and Barak Cohen and the hours of presentation guidance from Jim Skeath. The MGG program would not be where it is without the

leadership of Jim Skeath, Tim Schedl, and Melanie Relich – to all of whom I am extraordinarily grateful.

I am particularly thankful for my thesis committee members for their willingness to sit through my long-winded updates and for the guidance they have provided for the past 6 years.

Lastly, I could never fully explain how thankful I am for the support and love of my family: first to my parents, to whom I can say with the utmost confidence [ $p < 0.0001$ ] that I would not be where I am today without their sacrifices, unwavering motivation, or understanding. To my brother and grandparents who have always been there to get my head out of my science and onto something more important. And finally, to my wife Ann, who has been my advocate, devil's advocate, strongest supporter, and my best friend throughout grad school.

This work was supported by National Institutes of Health grants R01EY012543 and R01EY025272 (SC); T32EY013360 (Washington University [WU]) and P30EY002687 (WU, Department of Visual Science [WU-DOVS]); and unrestricted funds from Research to Prevent Blindness (WU-DOVS).

Philip A. Ruzycki

*Washington University in St. Louis*

*May 2018*

Dedicated to my wife and fellow scientist, Ann

## ABSTRACT OF THE DISSERTATION

Deciphering mechanisms governing the development of the rod epigenome

by

Philip Andrew Ruzycki

Doctor of Philosophy in Biology and Biomedical Sciences

Molecular Genetics and Genomics

Washington University in St. Louis, 2017

Professor Shiming Chen, Chair

Precisely coordinated expression of distinct sets of genes is essential for cellular development and function, especially in complex multicellular organisms. This regulation is achieved by the action of transcription factors (TF), proteins that bind specific genomic locations and alter the activity state and packaging of the DNA to promote or repress gene expression. However, while tremendous effort has defined networks of transcription factors that work together to drive specific phenotypes, little is known about their differential activity at the hundreds or thousands of sites where they bind. There are also many questions regarding the basic principles of the packaging of DNA within the nucleus, its influence on gene expression, and how transcription factors regulate this process.

To address these questions, I have investigated gene regulatory mechanisms that control the development of rod and cone photoreceptors. Photoreceptors are the neurons of the retina responsible for the initial conversion of a visual stimulus into an electrical signal. Photoreceptors are a complex, accessible, and highly disease relevant neuronal cell population, the lessons from which are relevant for many other cell types across the body. A primary TF in photoreceptor

development is CRX. This TF when mutated in humans can cause severe vision loss, and its deletion in mice leads to a severe condition with no functional photoreceptors.

First, I determined the functional consequences of human disease-causing variants when modelled in mice. I functionally classified a new model as causing a previously un-determined dominant disease, and discovered that all CRX diseases are the result of graded differences in gene expression of the same core set of genes. Second, I examined the dependency of all genomic binding sites on CRX activity and identified a core set of regulatory elements that require CRX for activation of the target genes. Third, I interrogated the broader organization of the rod photoreceptor epigenome and demonstrated that the rod packages its DNA according to epigenomic activity state. Perturbations to this state can override this organization, resulting in functional consequences on gene expression beyond the local sequence.

In summary, my in-depth investigation has uncovered new insights into the molecular mechanisms controlling development and maintenance of the rod epigenome and its organization in the nucleus. This new knowledge will provide 1) a new understanding of where and how mutations in a key photoreceptor TF cause gene mis-regulation, and 2) guidance for future human genetic studies to identify new disease-causing mutations affecting photoreceptor integrity, not only in the protein coding sequences but also in specific non-coding regulatory regions.

# **Chapter 1**

## **Introduction**

Author: Philip Ruzycki



## 1.1 The Genome and Its Regulation

The genome is the evolutionary substrate of all life on earth. The precise order, or sequence, of the nucleotides creates the diversity differentiating one individual from the next and one species from the other. The most basic function of this sequence is to transmit instructions for the generation of proteins. To this end, the genome must first be converted into a single stranded RNA molecule, in a process called transcription. This message is subsequently translated into the appropriate peptide units by other machinery in the cell.

Primitive organization of single celled organisms allows for the regulation of transcription to be relatively simple. However, with added complexity through evolutionary time, organisms have adopted highly complex body plans. It is only logical to assume that such an increase in complexity would necessitate a more complex set of DNA instructions. Prior to the sequencing of the human genome, it was thought that this increase in complexity would arise from an increase in gene number and genetic material. While this is the case on the basis of gross genome size, especially when compared to simple bacteria (*E. coli*) or the fly (*D. melanogaster*), there is a surprisingly small increase in gene number between lower vertebrates and mammals (Figure 1.1). The only logical conclusion, therefore, is that higher vertebrates have evolved new methods to control the expression of these same genes to drive the increase in functional complexity.

### 1.1.1 Transcription Factors Control Gene Expression

Gene expression is controlled by regulatory proteins called transcription factors (TFs). These proteins bind to defined sites throughout the genome, and can direct the modification of

the local active state. Together, the activity and specificity of distinct combinations of TFs allow a neuron and a fibroblast to have the exact same genomic sequence, but each only creates its own set of proteins in the correct amount at the proper time.

The canonical TF is comprised of two domains; one binds DNA (DNA binding domain or DBD) while the other regulates transcription (e.g. activation domain, or AD) by interacting with other co-regulatory factors.[Ma & Ptashne, 1987] TFs fall within a number of broad families (homeodomains, zinc-fingers, leucine zippers, among others) based on the structure of their DNA binding domain.[Luscombe et al., 2000] Very subtle variation within a family can have implications on the sequence specificity, precluding *a priori* prediction of the optimal DNA targets. Many experimental methods have been employed to understand and classify sequence specificity including promoter bashing reporter assays, binding assays such as EMSAs, or large scale HT-SELEX experiments.[Chen et al., 1997, 2002; Jolma et al., 2010] However, the knowledge of the optimal motif is not always indicative of *in vivo* activity. The analysis of the genome-wide binding of a TF may discover complex binding features; Many bound sites contain the characteristic cognate motif, but other bound sites lack that sequence, while some DNA regions harboring even strong motifs are not bound.[Boyle et al., 2011; Neph et al., 2012; Verfaillie et al., 2016; Zhu et al., 2009]

TFs rarely bind to a genomic location alone, and their binding is often primed by or directly mediated by another factor (or factors). [Iwafuchi-Doi et al., 2016] This combinatorial nature of TF binding and activity allows for both complexity and redundancy, all of which is encoded in the sequence of the regulatory site by number, orientation and spacing of various TF binding motifs. (reviewed in [Long et al., 2016; Slattery et al., 2014; Spitz & Furlong, 2012] New methods such as Massively Parallel Reporter Assays (MPRAs) have begun to unravel such

regulatory logic,[Inoue & Ahituv, 2015; Santiago-Algarra et al., 2017] but unfortunately, this technique is only currently available as a plasmid-based reporter assay. This is a major deficit in the field, as the function of a TF once bound to the proper site is to recruit other general factors and enzymes to modulate the local DNA regulatory environment that is not applicable on a plasmid. New methods must be developed to understand the differential activity of a TF at its thousands of binding sites in the genomic context.

## 1.2 Organizing the Genome

A single human genome stretches 6 feet in length which must be compacted into the nucleus of every single cell in the body. To achieve this, there are a number of discrete levels at which DNA compaction is organized. The double helix of DNA is first wrapped 1.65X around an octamer of proteins called histones to create the nucleosome structure.[Hewish & Burgoyne, 1973; Richmond et al., 1984] Further compaction by the association and wrapping with other adjacent nucleosomes creates a chromatin fiber. This fiber has been visualized by *in vitro* methods,[Song et al., 2014] but many questions still remain as to the precise association between nucleosomes and coiling *in vivo*.

### 1.2.1 Nucleosomes and modifications

At the level of the nucleosome, there is substantial control on the activity of the local genomic environment.[Han & Grunstein, 1988] Histone proteins have extended N-terminal tails that are well conserved, the loss of which causes the histone to lose its function.[Kayne et al., 1988] Histone tails are highly modified within the nucleus.(reviewed in [Li et al., 2007]) These modifications affect the interaction between DNA and the nucleosome [Zhao et al., 2005] and

determine the activity state of that site in the genome.[Strahl & Allis, 2000] Whole genome methods have been used to map the locations of specific modifications and have determined the regulatory logic that defines conventional promoters, enhancers, and other important elements.[Birney et al., 2007; Dixon et al., 2015; Dunham et al., 2012; Heintzman et al., 2007; Roadmap Epigenomics Consortium et al., 2015] The modification of histone H3 is highly relevant for the active state of promoters and enhancers; tri-methylation of K4 is associated with positive gene expression,[Santos-Rosa et al., 2002] the acetylation of K27 is associated with gene expression and enhancer activity,[Creyghton et al., 2010] while the tri-methylation of the same residue creates a strongly repressed environment.[Plath et al., 2002]

These modifications reflect a major element of the ‘epigenome,’ a layer of gene regulation beyond the individual nucleotide bases. The epigenome is not static, and varies dramatically between cell types and over development.[Birney et al., 2007] The epigenome has been theorized to act as a barrier to ensure the continuous progression toward a mature cell (e.g. Waddington’s Landscape) and an efficient means to ensure proper activity post developmentally.[Moris et al., 2016; Waddington, 1942]

A number of enzymes are necessary to modify the epigenome and remodel promoters and enhancers both over development and in response to environmental changes in adult cells. General enzymes such as CBP/P300 have no sequence specificity, but rather are guided by cell type specific TFs.[Vo & Goodman, 2001] Other chromatin modifying enzymes include lysine methyl transferase enzymes MLL1-4 and the repressive PRC2 complex. All of these enzymes are essential for cellular development and homeostasis and their disruption has been associated with a multitude of diseases.[Attar & Kurdistani, 2017; Hess, 2004; Pasini et al., 2004] The specific

roles of these enzymes and their guidance by cell type specific TFs remain to be elucidated in the development of many tissues.

### **1.2.2 Genome organization beyond the nucleosome**

The notion that DNA is packaged at a scale beyond the molecular level has actually been studied for decades in *D. melanogaster*. Cells within the fly salivary gland produce many copies of their genome which fuse into microscopically visible polytene chromosomes that have a distinct banding pattern. This pattern resembles the R/G bands visible in the mitotic genome of other species such as humans and mice under higher magnification. Microscopically, post-mitotic nuclei also show organizational principles. By DAPI staining, DNA clearly forms densely packed dark clusters in the nucleus, but leaves other DNA less compact.

Immunohistochemical staining of epigenetic modifying enzymes, TFs, and histones further support a stratified organization.[Corbo et al., 2010; Solovei et al., 2009]

The gold standard technique to study the localization of specific DNA fragments is fluorescence in-situ hybridization (FISH). The method was originally developed using radioactive probes to localize genes on polytene chromosomes,[M. L. Pardue & Gall, 1969; M. Lou Pardue et al., 1970] but has since been adapted to use fluorescent probes and has been shown to be effective within many cell types. Advances in molecular biology and high throughput sequencing have only recently allowed for the interrogation of genomic organization at the molecular level.[Dekker et al., 2002] Careful analysis that compared the results of FISH and Chromosome Conformation Capture data suggests that they both faithfully report genome organization, but their results can be complementary and both are still necessary to fully understand the organizational state.[Williamson et al., 2014] There are a variety of Chromosome

Conformation Capture (C) techniques (Table 1.1), each with its own particular benefits and shortcomings.(reviewed in [Davies et al., 2017]) While targeted assays such as 3C can give very high precision or resolution, lower sequencing costs are making genome wide assays such as 4C and HiC accessible to answer global questions.

HiC has been instrumental in defining a new understanding of the packing of interphase chromatin. By analyzing the data of this ‘all vs all’ approach, two primary concepts have emerged; first the Topologically Associating Domain (TAD) has been defined as another unit of genomic packaging, [Dixon et al., 2012] and second that the nucleus is generally organized into A and B compartments.[Lieberman-Aiden et al., 2009; Rao et al., 2014] A TAD is defined as a region of DNA, the sequences of which are more likely to interact with one another than with DNA outside of the TAD. In fact, the analysis of the *Drosophila* polytene chromosomes determined that TAD structure closely mirrored the order and distinction of banding patterns. Although these chromosomes are so unique in size, their molecular organization was virtually the same as that observed in normal diploid cells.[Eagen et al., 2015] TADs are largely conserved between cell types and the boundaries of TADs are enriched for the binding of CTCF,[Dixon et al., 2012; Hansen et al., 2017; Vietri Rudan et al., 2015] the loss of which disrupts these units.[Rao et al., 2017]

The function of TADs is still under investigation, however they may serve to confine a genomic segment and ensure proper promoter-enhancer interactions to regulate gene expression.[Jin et al., 2013; Symmons et al., 2014; H. Zhang et al., 2013] They are largely invariable between cell types with some exceptions; subtle changes in TAD boundary locations are essential for some developmental processes.[Andrey et al., 2013] New genomic editing methods have been employed to disrupt principles of this packaging to test their function. The

HOX gene cluster that defines basic body plan in vertebrates has served as a robust and defined test platform. By the deletion of a CTCF element that defined a TAD boundary, the important strict delineation of active vs inactive chromatin was disrupted [Narendra et al., 2015] and gene mis-regulation resulted in severe developmental abnormalities.[Narendra et al., 2016]

While the structure of TADs is quite stable across cell types, their locations within the nucleus are not. TADs can be generally classified as belonging to either an active (A) or inactive (B) compartment, the classification of which is strongly associated with the composition of repetitive elements, replication timing, lamin association, gene density, and gene function. [Dileep et al., 2015; Solovei et al., 2016] While the boundaries of A/B compartments do align with TAD boundaries, the inclusion of any individual TAD in either compartment is dynamic over development and highly cell type specific.[Dixon et al., 2015] The function of A/B compartments is not well understood. It is likely that this organization maintains the proximity of machinery to proper targets and separation from non-expressed genes.[Brown et al., 2008; Dorier & Stasiak, 2010; Feuerborn & Cook, 2015; Lamond & Spector, 2003; Lawrence & Clemson, 2008; Razin et al., 2011; Shen et al., 2012] However, this segregation and any effects of its disruption have not yet been tested. In fact, the mechanisms of compartment assembly are not well understood. Recent studies that disrupted TAD formation globally discovered enhanced compartmental segregation, suggesting that the formation of TADs and compartments are somewhat opposed.[Rao et al., 2017; Schwarzer et al., 2017] More targeted analyses are necessary to elucidate the functions of and relationship between TADs and A/B Compartments.

### 1.3 The retina as model system

The mouse retina offers a tractable but complex and highly relevant neuronal system to interrogate the principles of TF mediated gene regulation and genomic organization. The retina is the neural tissue that lines the posterior surface of the eye. The function of the retina is to convert a light stimulus into an electrical signal, perform some processing and filtering of the signal, and pass this information to the brain. The retina consists of 6 neural and 1 glial cell type to perform this function. These cells are organized into 3 layers; the outermost layer is comprised exclusively of rod and cone photoreceptors that capture photons of light and initiate the electrical response, the central layer contains 4 types of interneurons that perform basic signal processing, and the third layer contains ganglion cells that relay the signal to the brain.

These cell types have been profiled extensively both morphologically and transcriptionally; many subtypes of the 3 interneurons and ganglion cells have been described.[Baden et al., 2016; Macosko et al., 2015; Shekhar et al., 2016; Siegert et al., 2012] Rods and cones, however, are very well conserved in terms of structure and function, and there is only a single type of rod photoreceptor. The mouse retina is composed of 80% rods,[Jeon et al., 1998; Macosko et al., 2015] making the bulk analysis of the tissue highly representative of their signal.

The retina develops over a 2-week period in the mouse. The cell types develop with a highly stereotyped birth order that has been studied extensively.[Young, 1985] This order of progenitor competence and cell fate specification is driven by a set of TFs that are still under investigation, however, many have been shown to drive certain cell fates.[Livesey & Cepko, 2001; de Melo et al., 2016; Wang et al., 2014]



Rod photoreceptors are born from a set of multipotent progenitors, although recent evidence suggests that evolutionarily, rods were actually co-opted from cones and this derivation is visible in their very early development.[Kim, Yang, Oel, et al., 2016] The structure of a rod is highly specialized for the function of capturing photons of light by the photopigment Rhodopsin, and converting this into an electrical signal by a G-protein coupled cascade called phototransduction.(Reviewed in [Arshavsky et al., 2002; Kefalov, 2012] The signal is amplified and the rod passes the signal to the interneurons. Rhodopsin is stored in specialized membranous disc structures located in a modified cilia structure called the outer segment (OS). These discs are shed nightly, making the cell constantly responsible for renewing the membrane and protein components essential for phototransduction. To accomplish this, rods must produce an extraordinary amount of Rhodopsin on a daily basis; Rhodopsin transcript represent more than 2% of the total RNA in the cell.[Blackshaw et al., 2001].

### **1.3.1 Retinal transcription factors and disease**

Rod and cone cell fate specification and gene expression is determined by a cascade of TF activation.(Reviewed in [Swaroop et al., 2010]) This cascade is initiated by a homeodomain TF OTX2 that is necessary and (in the retina) sufficient for photoreceptor fate specification.[Nishida et al., 2003] OTX2 initiates the expression of cone-rod homeobox (CRX)[Nishida et al., 2003] that works with other TFs NRL and NR2E3 in rod photoreceptors to actively drive rod fate. In mice, the loss of NRL results in a transacted state, where rod cells instead adopt the S-cone cell fate.[Mears et al., 2001] The loss of NR2E3, which is activated downstream of NRL,[Oh et al., 2008] results in a mid-fate, where many rod genes are properly activated, but the cell did not properly repress cone related transcripts.[Cheng et al., 2006; G.-H.

Peng et al., 2005] Both of these TFs also cause vision related disease in human patients.[Jacobson et al., 1990; Milam et al., 2002; Wright et al., 2004]

### 1.3.2 CRX

CRX is a homeodomain TF expressed very early after the final mitotic event in photoreceptors and this expression is maintained throughout adulthood.[Chen et al., 1997; Furukawa et al., 1997] In humans, the mutation of CRX results in a number of diseases with varying functional deficits.(reviewed in [Tran & Chen, 2014]). As a TF, the basic function of CRX is to bind target sites throughout the genome to activate photoreceptor related transcripts. CRX has a very simple domain structure with a N-terminal canonical OTX homeodomain region and the C-terminal half of the protein comprises the activation domain.[Chau et al., 2000; Chen et al., 2002]

Human disease causing mutations to CRX can be classified into 4 groups that are defined by the type of mutation and describe the associated severity of symptoms.[Tran & Chen, 2014] Class I mutations are the result of a point mutation to the homeodomain portion of the protein. This mutation disrupts DNA binding,[Chen et al., 2002] and results in a recessive condition both in humans and when modelled in mice (*Crx*<sup>R90W/W</sup>).[Tran et al., 2014] Class III mutants contain a truncation mutation that causes the translated CRX protein to lack a significant portion of the activation domain. Current models predict that mutant CRX retains the ability to bind DNA but lacks the ability to transactivate. In the heterozygous state, Class III mutations cause severe dominant disease, thought to be the result of mutant protein blocking the action of the *WT* protein.[Ocellini et al., 2016; Tran et al., 2014; Tran & Chen, 2014] With the highly divergent nature of phenotypes as a result of mutation to CRX, experiments in Chapters 2-3 of this

dissertation are designed to better understand Class III mutants and to understand the downstream transcriptional effects of Class I and III mutations.

Many studies have probed the function of CRX in normal photoreceptor development. ChIP experiments have determined sites throughout the genome bound by CRX and those co-bound by NRL and/or NR2E3.[Corbo et al., 2010; Hao et al., 2012; Montana et al., 2011; G.-H. Peng et al., 2005; G.-H. Peng & Chen, 2005] Binding sites of CRX are enriched for the canonical homeodomain motif,[Corbo et al., 2010] and disruption of this site abolishes binding and activity.[Chen et al., 1997; Lee et al., 2010; Montana et al., 2011; White et al., 2016, 2013] Massively parallel reporter assays (MPRAs) have tremendously extended our knowledge of CRX and its interactions with other factors. The assessment of sequences bound by CRX discovered that bound sites activated transcription while unbound sites with a CRX motif did not.[White et al., 2013] This result suggested that activity is encoded in the sequence and depends not only on the presence of a CRX motif, but relies on neighboring sequence and the binding of other factors. This same experiment also proposed that CRX may in some instances act as a repressor; sites that displayed low expression in *WT* retinas expressed higher when tested in *Crx*<sup>-/-</sup> retina.[White et al., 2013] Synthetic arrays of strong CRX sites also displayed repressive activity, but the presence of NRL sites overrode this effect.[White et al., 2016] Such negative activity has not yet been demonstrated to be relevant *in vivo*. Together, these experiments suggest that gene regulation by CRX is not binary, but rather highly context dependent.

### **1.3.3 The photoreceptor epigenome**

These TFs direct a photoreceptor progenitor over a two-week period of terminal differentiation to become a functionally mature rod. This process requires a tremendous change in gene expression, the result of a dramatic remodeling of the epigenome. This remodeling of the *WT* rod epigenome has been characterized previously in terms of histone modifications, DNA methylation, genome accessibility, and TF binding, all of which were shown to correlate well with the changes to the transcriptome.[Aldiri et al., 2017; Hughes et al., 2017; Kim, Yang, Brooks, et al., 2016; Mo et al., 2016] However, this epigenome has not been systematically analyzed in mutant photoreceptors. For instance, very little is known about the state of the chromatin environment upon loss of CRX. There is also a lack of understanding of the activity of CRX in establishing an active chromatin state on a whole or at any individual site in the genome. Chapter 4 in this dissertation aims to extend the knowledge of CRX by *in vivo* analysis.

### **1.3.4 Photoreceptor Genomic Organization**

The organization of the mouse rod photoreceptor genome is distinguishable from other retinal cell types simply by microscopic assessment. Condensed heterochromatin is packed into a dense core, while more active euchromatin is located at the periphery of the nucleus. This organization is not unique to the mouse. In fact, a very thorough analysis of many species discovered that this phenomenon was correlated with the activity state of the animal. Nocturnal animals across the evolutionary tree develop this nuclear organization. It was postulated that this organization acts as a condensing lens for photons of light, in contrast to the normal distributed organization of heterochromatin centers that would scatter light.[Solovei et al., 2009]

Further work suggested that this organization is the result of either the inactivation of Lamin B receptor (LBR) or Lamin A/C (LMNA), proteins that would normally tether heterochromatin to the nuclear membrane.[Solovei et al., 2013] Each cell type in the body normally expresses one of these genes, and in mouse rod photoreceptors, the specific inactivation of *Lmna* has been associated with this architecture change as compared to cones.[Hughes et al., 2017] Genetic ablation of *Lbr* in cells that otherwise lack *Lmna* resulted in inverted nuclei suggesting that this loss is sufficient to induce inversion, and the removal of both caused dramatic expression changes in cell culture systems.[Solovei et al., 2013]

This unique organization is altered in mouse models of photoreceptor disease. Studies have shown mis-organization in mutants of *Crx*, *Nrl*, *Sca7*, *Rb*, and *Cbp/P300*. [Helmlinger et al., 2006; Hennig et al., 2013; Johnson et al., 2006; Roger et al., 2014; Tran et al., 2014; J. Zhang et al., 2004] In the *Sca7* mutant, timing of nuclear chromatin changes was determined to predate functional deficits, suggesting that these changes could have resulted in expression differences.[Helmlinger et al., 2006] The interpretation of cause in other mutants is unclear, as photoreceptor precursor and other retinal cells display the normal chromatin state. It is difficult to decouple a change in fate decision or lack of the cue to repress *Lmna* vs a direct result of the disease.

This unique organization has been associated with large scale changes in the chromatin landscape. Recent ATAC-seq data suggests that there are large regions of the genome that completely lack the normal nominal open regulatory sites.[Hughes et al., 2017] These regions largely correlate with traditional heterochromatin, and the interpretation must be that rod condensation causes these regions to become virtually completely inaccessible. This suggests that a rod nucleus is organized differently, or more strictly by chromatin state than observed in

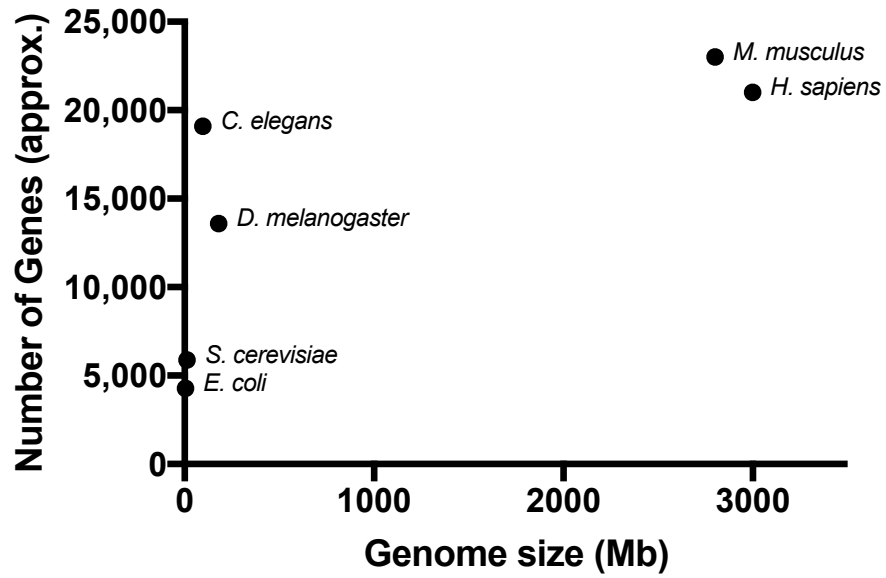
other cell types. Indeed, immunohistochemistry and *in situ* hybridization have shown strict delineation of classes of chromatin. Activating TFs such as CRX are only located at the most peripheral level,[Corbo et al., 2010] while euchromatin vs heterochromatin associated DNA repeat sequences show very little overlap in the nuclear volume.[Solovei et al., 2009] Limited FISH experiments that probed the locations of particular genes also showed active genes are located at the far euchromatin periphery.[Aldiri et al., 2017; Solovei et al., 2009]

Rod genomic organization has only been analyzed locally at the molecular level. 3C experiments described enhancer-promoter interactions and their dependency on cell type specific TFs.[G. Peng & Chen, 2011, 2012] Another recent study used 4C to analyze local compaction of a several Mb region. This work showed that genomic loci in that region clustered within the nuclear volume based on chromatin activity status, irrelevant of linear genomic distance.[van de Werken et al., 2017] Experiments presented in Chapter 5 of this dissertation are designed to interrogate the organization of whole chromosomes in the rod nucleus and to test several fundamental mechanisms and function of this packaging.

**Table 1.1. Summary of Chromosome Conformation Capture Techniques**

<b>Technique</b>	<b>Bait vs Capture</b>	<b>Data</b>
<b>3C</b>	one vs one	PCR
<b>4C</b>	one vs all	Seq
<b>5C</b>	many vs many	Seq
<b>HiC</b>	all vs all	Seq
<b>ChIA-PET</b>	all vs all bound by specific protein	Seq

adapted from [de Wit & de Laat, 2012]



**Figure 1.1: Humans and mice have more non-coding genomic sequence.** Plot displays the approximate size of the genome vs the number of coding genes for each of the labelled species. Adapted from [Cooper, 2000].



## 1.4 References

- Aldiri, I., Xu, B., Wang, L., Chen, X., Hiler, D., Griffiths, L., Valentine, M., et al. (2017). The Dynamic Epigenetic Landscape of the Retina During Development, Reprogramming, and Tumorigenesis. *Neuron*, *94*, 550–568.
- Andrey, G., Montavon, T., Mascrez, B., Gonzalez, F., Noordermeer, D., Leleu, M., Trono, D., et al. (2013). A Switch Between Topological Domains Underlies HoxD Genes Collinearity in Mouse Limbs. *Science*, *340*.
- Arshavsky, V. Y., Lamb, T. D., & Pugh, E. N. (2002). G Proteins and Phototransduction. *Annual Review Of Physiology*, *64*, 153–187.
- Attar, N., & Kurdistani, S. K. (2017). Exploitation of EP300 and CREBBP Lysine Acetyltransferases by Cancer. *Cold Spring Harbor Perspectives In Medicine*, *7*.
- Baden, T., Berens, P., Franke, K., Roman-Roson, M., Bethge, M., & Euler. (2016). The functional diversity of mouse retinal ganglion cells. *Nature*, *529*, 345–350.
- Birney, E., Stamatoyannopoulos, J. A., Dutta, A., Guigó, R., Gingeras, T. R., Margulies, E. H., Weng, Z., et al. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, *447*, 799–816.
- Blackshaw, S., Fraioli, R. E., Furukawa, T., & Cepko, C. L. (2001). Comprehensive analysis of photoreceptor gene expression and the identification of candidate retinal disease genes. *Cell*, *107*, 579–589.
- Boyle, A. P., Song, L., Lee, B. K., London, D., Keefe, D., Birney, E., Iyer, V. R., et al. (2011). High-resolution genome-wide in vivo footprinting of diverse transcription factors in human cells. *Genome Research*, *21*, 456–464.
- Brown, J. M., Green, J., das Neves, R. P., Wallace, H. a C., Smith, A. J. H., Hughes, J., Gray, N., et al. (2008). Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. *The Journal Of Cell Biology*, *182*, 1083–97.
- Chau, K. Y., Chen, S., Zack, D. J., & Ono, S. J. (2000). Functional domains of the cone-rod homeobox (CRX) transcription factor. *The Journal Of Biological Chemistry*, *275*, 37264–70.
- Chen, S., Wang, Q. L., Nie, Z., Sun, H., Lennon, G., Copeland, N. G., Gilbert, D. J., et al. (1997). Crx, a novel Otx-like paired-homeodomain protein, binds to and transactivates photoreceptor cell-specific genes. *Neuron*, *19*, 1017–30.
- Chen, S., Wang, Q., Xu, S., Liu, I., Li, L. Y., Wang, Y., & Zack, D. J. (2002). Functional analysis of cone – rod homeobox ( CRX ) mutations associated with retinal dystrophy. *Human Molecular Genetics*, *11*, 873–884.
- Cheng, H., Aleman, T. S., Cideciyan, A. V, Khanna, R., Jacobson, S. G., & Swaroop, A. (2006). In vivo function of the orphan nuclear receptor NR2E3 in establishing photoreceptor identity during mammalian retinal development. *Human Molecular Genetics*, *15*, 2588–602.
- Cooper, G. M. (2000). *The Cell: A Molecular Approach*. (Sunderland (MA), Ed.) (2nd ed.). Sinauer Associates.
- Corbo, J. C., Lawrence, K. a, Karlstetter, M., Myers, C. a, Abdelaziz, M., Dirkes, W., Weigelt, K., et al. (2010). CRX ChIP-seq reveals the cis-regulatory architecture of mouse photoreceptors. *Genome Research*, *20*, 1512–25.

- Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., Hanna, J., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings Of The National Academy Of Sciences*, *107*, 21931–21936.
- Davies, J. O. J., Oudelaar, A. M., Higgs, D. R., & Hughes, J. R. (2017). How best to identify chromosomal interactions: a comparison of approaches. *Nature Methods*, *14*, 125–134.
- Dekker, J., Rippe, K., Dekker, M., & Kleckner, N. (2002). Capturing chromosome conformation. *Science*, *295*, 1306–11.
- Dileep, V., Rivera-Mulia, J. C., Sima, J., & Gilbert, D. M. (2015). Large-Scale Chromatin Structure-Function Relationships during the Cell Cycle and Development: Insights from Replication Timing. *Cold Spring Harbor Symposia On Quantitative Biology*, *80*, 53–63.
- Dixon, J. R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-bourget, J. E., Lee, A. Y., Ye, Z., et al. (2015). Chromatin architecture reorganization during stem cell differentiation. *Nature*, *518*, 331–336.
- Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., et al. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, *485*, 376–380.
- Dorier, J., & Stasiak, A. (2010). The role of transcription factories-mediated interchromosomal contacts in the organization of nuclear architecture. *Nucleic Acids Research*, *38*, 7410–21.
- Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., Epstein, C. B., et al. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, *489*, 57–74.
- Eagen, K. P., Hartl, T. A., & Kornberg, R. D. (2015). Stable Chromosome Condensation Revealed by Chromosome Conformation Capture. *Cell*, *163*, 934–946.
- Feuerborn, A., & Cook, P. R. (2015). Why the activity of a gene depends on its neighbors. *Trends In Genetics*, *31*, 483–490.
- Furukawa, T., Morrow, E. M., & Cepko, C. L. (1997). Crx, a novel otx-like homeobox gene, shows photoreceptor-specific expression and regulates photoreceptor differentiation. *Cell*, *91*, 531–41.
- Han, M., & Grunstein, M. (1988). Nucleosome loss activates yeast downstream promoters in vivo. *Cell*, *55*, 1137–1145.
- Hansen, A. S., Pustova, I., Cattoglio, C., Tjian, R., & Darzacq, X. (2017). CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *ELife*, *6*, 1–33.
- Hao, H., Kim, D. S., Klocke, B., Johnson, K. R., Cui, K., Gotoh, N., Zang, C., et al. (2012). Transcriptional Regulation of Rod Photoreceptor Homeostasis Revealed by In Vivo NRL Targetome Analysis. *PLoS Genetics*, *8*, e1002649.
- Heintzman, N. D., Stuart, R. K., Hon, G., Fu, Y., Ching, C. W., Hawkins, R. D., Barrera, L. O., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genetics*, *39*, 311–318.
- Helmlinger, D., Hardy, S., Abou-Sleymane, G., Eberlin, A., Bowman, A. B., Gansmüller, A., Picaud, S., et al. (2006). Glutamine-expanded ataxin-7 alters TFTC/STAGA recruitment and chromatin structure leading to photoreceptor dysfunction. *PLoS Biology*, *4*, e67.
- Hennig, A. K., Peng, G.-H., & Chen, S. (2013). Transcription Coactivators p300 and CBP Are Necessary for Photoreceptor-Specific Chromatin Organization and Gene Expression. *PloS One*, *8*, e69721.
- Hess, J. L. (2004). Mechanisms of transformation by MLL. *Critical Reviews In Eukaryotic Gene Expression*, *14*, 235–254.

- Hewish, D. R., & Burgoyne, L. A. (1973). Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. *Biochemical And Biophysical Research Communications*, *52*, 504–510.
- Hughes, A. E. O., Enright, J. M., Myers, C. A., Shen, S. Q., & Corbo, J. C. (2017). Cell Type-Specific Epigenomic Analysis Reveals a Uniquely Closed Chromatin Architecture in Mouse Rod Photoreceptors. *Scientific Reports*, *7*, 43184.
- Inoue, F., & Ahituv, N. (2015). Decoding enhancers using massively parallel reporter assays. *Genomics*, *106*, 159–164.
- Iwafuchi-Doi, M., Donahue, G., Kakumanu, A., Watts, J. A., Mahony, S., Pugh, B. F., Lee, D., et al. (2016). The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation. *Molecular Cell*, *62*, 79–91.
- Jacobson, S. G., Marmor, M. F., Kemp, C. M., & Knighton, R. W. (1990). SWS (blue) cone hypersensitivity in a newly identified retinal degeneration. *Investigative Ophthalmology And Visual Science*, *31*, 827–838.
- Jeon, C. J., Strettoi, E., & Masland, R. H. (1998). The major cell populations of the mouse retina. *The Journal Of Neuroscience*, *18*, 8936–46.
- Jin, F., Li, Y., Dixon, J. R., Selvaraj, S., Ye, Z., Lee, A. Y., Yen, C.-A., et al. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature*, *503*, 290–294.
- Johnson, D. A., Donovan, S. L., & Dyer, M. A. (2006). Mosaic Deletion of Rb Arrests Rod Differentiation and Stimulates Ectopic Synaptogenesis in the Mouse Retina. *The Journal Of Comparative Neurology*, *128*, 112–128.
- Jolma, A., Kivioja, T., Toivonen, J., Cheng, L., Wei, G., Enge, M., Taipale, M., et al. (2010). Multiplexed massively parallel SELEX for characterization of human transcription factor binding specificities. *Genome Research*, 861–873.
- Kayne, P. S., Kim, U. J., Han, M., Mullen, J. R., Yoshizaki, F., & Grunstein, M. (1988). Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mating loci in yeast. *Cell*, *55*, 27–39.
- Kefalov, V. J. (2012). Rod and cone visual pigments and phototransduction through pharmacological, genetic, and physiological approaches. *Journal Of Biological Chemistry*, *287*, 1635–1641.
- Kim, J.-W., Yang, H.-J., Brooks, M. J., Zelinger, L., Karakulah, G., Gotoh, N., Boleda, A., et al. (2016). NRL-Regulated Transcriptome Dynamics of Developing Rod Photoreceptors. *Cell Reports*, *17*, 2460–2473.
- Kim, J.-W., Yang, H.-J., Oel, A. P., Brooks, M. J., Jia, L., Plachetzki, D. C., Li, W., et al. (2016). Recruitment of Rod Photoreceptors from Short-Wavelength-Sensitive Cones during the Evolution of Nocturnal Vision in Mammals. *Developmental Cell*, *37*, 520–532.
- Lamond, A. I., & Spector, D. L. (2003). Nuclear speckles: a model for nuclear organelles. *Nature Reviews. Molecular Cell Biology*, *4*, 605–12.
- Lawrence, J. B., & Clemson, C. M. (2008). Gene associations: True romance or chance meeting in a nuclear neighborhood? *Journal Of Cell Biology*, *182*, 1035–1038.
- Lee, J., Myers, C. a, Williams, N., Abdelaziz, M., & Corbo, J. C. (2010). Quantitative fine-tuning of photoreceptor cis-regulatory elements through affinity modulation of transcription factor binding sites. *Gene Therapy*, *17*, 1390–9.

- Li, B., Carey, M., & Workman, J. L. (2007). The Role of Chromatin during Transcription. *Cell*, *128*, 707–719.
- Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, *326*, 289–93.
- Livesey, F. J., & Cepko, C. L. (2001). Vertebrate neural cell-fate determination: lessons from the retina. *Nature Reviews Neuroscience*, *2*, 109–18.
- Long, H. K., Prescott, S. L., & Wysocka, J. (2016). Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. *Cell*, *167*, 1170–1187.
- Luscombe, N. M., Austin, S. E., Berman, H. M., & Thornton, J. M. (2000). An overview of the structures of protein-DNA complexes. *Genome Biology*, *1*, 1–37.
- Ma, J., & Ptashne, M. (1987). Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell*, *48*, 847–853.
- Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., et al. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*, *161*, 1202–1214.
- Mears, A., Kondo, M., Swain, P. K., Takada, Y., Bush, R. A., Saunders, T. L., Sieving, P. A., et al. (2001). Nrl is required for rod photoreceptor development. *Nature Genetics*, *29*, 447–52.
- de Melo, J., Zibetti, C., Clark, B. S., Hwang, W., Miranda-Angulo, A. L., Qian, J., & Blackshaw, S. (2016). Lhx2 Is an Essential Factor for Retinal Gliogenesis and Notch Signaling. *The Journal Of Neuroscience*, *36*, 2391–405.
- Milam, A. H., Rose, L., Cideciyan, A. V., Barakat, M. R., Tang, W.-X., Gupta, N., Aleman, T. S., et al. (2002). The nuclear receptor NR2E3 plays a role in human retinal photoreceptor differentiation and degeneration. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, *99*, 473–8.
- Mo, A., Luo, C., Davis, F. P., Mukamel, E. A., Henry, G. L., Nery, J. R., Urich, M. A., et al. (2016). Epigenomic landscapes of retinal rods and cones. *ELife*, *5*, 1–29.
- Montana, C. L., Lawrence, K. a, Williams, N. L., Tran, N. M., Peng, G.-H., Chen, S., & Corbo, J. C. (2011). Transcriptional regulation of neural retina leucine zipper (Nrl), a photoreceptor cell fate determinant. *The Journal Of Biological Chemistry*, *286*, 36921–31.
- Moris, N., Pina, C., & Arias, A. M. (2016). Transition states and cell fate decisions in epigenetic landscapes. *Nature Reviews Genetics*, *17*, 693–703.
- Narendra, V., Bulaji, M., Dekker, J., Mazzoni, E. O., & Reinberg, D. (2016). CTCF-mediated topological boundaries during development foster appropriate gene regulation. *Genes And Development*, *30*, 2657–2662.
- Narendra, V., Rocha, P. P., An, D., Raviram, R., Skok, J. A., Mazzoni, E. O., & Reinberg, D. (2015). CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science*, *347*, 1017–1021.
- Neph, S., Vierstra, J., Stergachis, A. B., Reynolds, A. P., Haugen, E., Vernot, B., Thurman, R. E., et al. (2012). An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature*, *489*, 83–90.

- Nishida, A., Furukawa, A., Koike, C., Tano, Y., Aizawa, S., Matsuo, I., & Furukawa, T. (2003). Otx2 homeobox gene controls retinal photoreceptor cell fate and pineal gland development. *Nature Neuroscience*, *6*, 1255–63.
- Occelli, L. M., Tran, N. M., Narfström, K., Chen, S., & Petersen-Jones, S. M. (2016). CrxRdy cat: A large animal model for CRX-associated leber congenital amaurosis. *Investigative Ophthalmology And Visual Science*, *57*, 3780–3792.
- Oh, E. C. T., Cheng, H., Hao, H., Jia, L., Khan, N. W., & Swaroop, A. (2008). Rod differentiation factor NRL activates the expression of nuclear receptor NR2E3 to suppress the development of cone photoreceptors. *Brain Research*, *1236*, 16–29.
- Pardue, M. L., & Gall, J. G. (1969). Molecular hybridization of radioactive DNA to the DNA of cytological preparations. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, *64*, 600–4.
- Pardue, M. Lou, Gerbi, S. A., Eckhardt, R. A., & Gall, J. G. (1970). Cytological localization of DNA complementary to ribosomal RNA in polytene chromosomes of Diptera. *Chromosoma*, *29*, 268–290.
- Pasini, D., Bracken, A. P., & Helin, K. (2004). Polycomb group proteins in cell cycle progression and cancer. *Cell Cycle*, *3*, 396–400.
- Peng, G.-H., Ahmad, O., Ahmad, F., Liu, J., & Chen, S. (2005). The photoreceptor-specific nuclear receptor Nr2e3 interacts with Crx and exerts opposing effects on the transcription of rod versus cone genes. *Human Molecular Genetics*, *14*, 747–64.
- Peng, G.-H., & Chen, S. (2005). Chromatin immunoprecipitation identifies photoreceptor transcription factor targets in mouse models of retinal degeneration: new findings and challenges. *Visual Neuroscience*, *22*, 575–86.
- Peng, G., & Chen, S. (2011). Active opsin loci adopt intrachromosomal loops that depend on the photoreceptor transcription factor network. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, *108*, 17821–6.
- Peng, G., & Chen, S. (2012). Revealing Looping Organization of Mammalian Photoreceptor Genes Using Chromosome Capture (3C) Assays. *Retinal Development: Methods And Protocols*, *884*, 305–318.
- Plath, K., Fang, J., Mlynarczyk-Evans, S. K., Cao, R., Worringer, K. A., Wang, H., de la Cruz, C. C., et al. (2002). Role of Histone H3 Lysine 27 Methylation in X Inactivation. *Science*, *298*, 1039–1043.
- Rao, S. S. P., Huang, S.-C., Hilaire, B. G. S., Engreitz, J. M., Perez, E. M., Kieffer-Kwon, K.-R., Sanborn, A. L., et al. (2017). Cohesin Loss Eliminates All Loop Domains. *Cell*, *171*, 305–320.
- Rao, S. S. P., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., Sanborn, A. L., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell*, *159*, 1–16.
- Razin, S. V., Gavrillov, A. A., Pichugin, A., Lipinski, M., & Iarovaia, O. V. (2011). Transcription factories in the context of the nuclear and genome organization. *Nucleic Acids Research*, *39*, 9085–9092.
- Richmond, T. J., Finch, J. T., Rushton, B., Rhodes, D., & Klug, A. (1984). Structure of the nucleosome core particle at 7 Å resolution. *Nature*, *311*, 532–537.
- Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., et al. (2015). Integrative analysis of 111 reference human epigenomes. *Nature*, *518*, 317–330.

- Roger, J. E., Hiriyanna, A., Gotoh, N., Hao, H., Cheng, D. F., Ratnapriya, R., Kautzmann, M. I., et al. (2014). OTX2 loss causes rod differentiation defect in CRX-associated congenital blindness. *The Journal Of Clinical Investigation*, *124*, 631–643.
- Santiago-Algarra, D., Dao, L. T. M., Pradel, L., España, A., & Spicuglia, S. (2017). Recent advances in high-throughput approaches to dissect enhancer function. *F1000Research*, *6*, 939.
- Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C. T., Schreiber, S. L., et al. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature*, *419*, 407–411.
- Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N. A., et al. (2017). Two independent modes of chromatin organization revealed by cohesin removal. *Nature*, *551*, 51–56.
- Shekhar, K., Lapan, S. W., Whitney, I. E., Tran, N. M., Macosko, E. Z., Kowalczyk, M., Adiconis, X., et al. (2016). Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. *Cell*, *166*, 1308–1323.
- Shen, Y., Yue, F., McCleary, D. F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., et al. (2012). A map of the cis-regulatory sequences in the mouse genome. *Nature*, *488*, 116–120.
- Siebert, S., Cabuy, E., Scherf, B. G., Kohler, H., Panda, S., Le, Y.-Z., Fehling, H. J., et al. (2012). Transcriptional code and disease map for adult retinal cell types. *Nature Neuroscience*, *15*, 487–95.
- Slattery, M., Zhou, T., Yang, L., Machado, A. C. D., Gordân, R., & Rohs, R. (2014). Absence of a simple code: how transcription factors read the genome. *Trends Biochem Sci*, *39*, 381–399.
- Solovei, I., Kreysing, M., Lanctôt, C., Kösem, S., Peichl, L., Cremer, T., Guck, J., et al. (2009). Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. *Cell*, *137*, 356–68.
- Solovei, I., Thanisch, K., & Feodorova, Y. (2016). How to rule the nucleus: divide et impera. *Current Opinion In Cell Biology*, *40*, 47–59.
- Solovei, I., Wang, A. S., Thanisch, K., Schmidt, C. S., Krebs, S., Zwerger, M., Cohen, T. V., et al. (2013). LBR and Lamin A/C Sequentially Tether Peripheral Heterochromatin and Inversely Regulate Differentiation. *Cell*, *152*, 584–598.
- Song, F., Chen, P., Sun, D., Wang, M., Dong, L., Liang, D., Xu, R.-M., et al. (2014). Cryo-EM Study of the Chromatin Fiber Reveals a Double Helix Twisted by Tetranucleosomal Units. *Science*, *344*, 376–380.
- Spitz, F., & Furlong, E. E. M. (2012). Transcription factors: from enhancer binding to developmental control. *Nature Reviews Genetics*, *13*, 613–626.
- Strahl, B. D., & Allis, C. D. (2000). The language of covalent histone modifications. *Nature*, *403*, 41–45.
- Swaroop, A., Kim, D., & Forrest, D. (2010). Transcriptional regulation of photoreceptor development and homeostasis in the mammalian retina. *Nature Reviews. Neuroscience*, *11*, 563–576.
- Symmons, O., Uslu, V. V., Tsujimura, T., Ruf, S., Nassari, S., Schwarzer, W., Ettwiller, L., et al. (2014). Functional and topological characteristics of mammalian regulatory domains. *Genome Research*, *390*–400.
- Tran, N. M., & Chen, S. (2014). Mechanisms of blindness: animal models provide insight into distinct CRX-associated retinopathies. *Developmental Dynamics*, *243*, 1153–66.
- Tran, N. M., Zhang, A., Zhang, X., Huecker, J. B., Hennig, A. K., & Chen, S. (2014). Mechanistically Distinct Mouse Models for CRX-Associated Retinopathy. *PLoS Genetics*, *10*, e1004111.

- Verfaillie, A., Svetlichnyy, D., Imrichova, H., Davie, K., Fiers, M., Atak, Z. K., Hulselmans, G., et al. (2016). Multiplex enhancer-reporter assays uncover unsophisticated p53 enhancer logic. *Genome Research*, 882–895.
- Vietri Rudan, M., Barrington, C., Henderson, S., Ernst, C., Odom, D. T., Tanay, A., & Hadjur, S. (2015). Comparative Hi-C Reveals that CTCF Underlies Evolution of Chromosomal Domain Architecture. *Cell Reports*, 10, 1297–1309.
- Vo, N., & Goodman, R. H. (2001). CREB-binding Protein and p300 in Transcriptional Regulation. *Journal Of Biological Chemistry*, 276, 13505–13508.
- Waddington, C. (1942). Canalization of Development and the Inheritance of Acquired Characters. *Nature*, 150, 563–565.
- Wang, S., Sengel, C., Emerson, M. M., & Cepko, C. L. (2014). A Gene Regulatory Network Controls the Binary Fate Decision of Rod and Bipolar Cells in the Vertebrate Retina. *Developmental Cell*, 30, 513–527.
- van de Werken, H. J. G., Haan, J. C., Feodorova, Y., Bijos, D., Weuts, A., Theunis, K., Holwerda, S. J. B., et al. (2017). Small chromosomal regions position themselves autonomously according to their chromatin class. *Genome Research*, 922–933.
- White, M. A., Kwasnieski, J. C., Myers, C. A., Shen, S. Q., Corbo, J. C., & Cohen, B. A. (2016). A Simple Grammar Defines Activating and Repressing cis-Regulatory Elements in Photoreceptors. *Cell Reports*, 17, 1247–1254.
- White, M. A., Myers, C. A., Corbo, J. C., & Cohen, B. A. (2013). Massively parallel in vivo enhancer assay reveals that highly local features determine the cis-regulatory function of ChIP-seq peaks. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 110, 11952–11957.
- Williamson, I., Berlivet, S., Eskeland, R., Boyle, S., Illingworth, R. S., Paquette, D., & Bickmore, W. A. (2014). Spatial genome organization : contrasting views from chromosome conformation capture and fluorescence in situ hybridization. *Genes And Development*, 28, 2778–2791.
- de Wit, E., & de Laat, W. (2012). A decade of 3C technologies: Insights into nuclear organization. *Genes And Development*, 26, 11–24.
- Wright, A. F., Reddick, A. C., Schwartz, S. B., Ferguson, J. S., Aleman, T. S., Kellner, U., Jurklics, B., et al. (2004). Mutation analysis of NR2E3 and NRL genes in Enhanced S Cone Syndrome. *Human Mutation*, 24, 439–439.
- Young, R. W. (1985). Cell differentiation in the retina of the mouse. *The Anatomical Record*, 212, 199–205.
- Zhang, H., Jiao, W., Sun, L., Fan, J., Chen, M., Wang, H., Xu, X., et al. (2013). Intrachromosomal looping is required for activation of endogenous pluripotency genes during reprogramming. *Cell Stem Cell*, 13, 30–35.
- Zhang, J., Gray, J., Wu, L., Leone, G., Rowan, S., Cepko, C. L., Zhu, X., et al. (2004). Rb regulates proliferation and rod photoreceptor development in the mouse retina. *Nature Genetics*, 36, 351–60.
- Zhao, J., Herrera-diaz, J., & Gross, D. S. (2005). Domain-Wide Displacement of Histones by Activated Heat Shock Factor Occurs Independently of Swi / Snf and Is Not Correlated with RNA Polymerase II Density. *Molecular And Cellular Biology*, 25, 8985–8999.
- Zhu, C., Byers, K. J., McCord, R. P., Shi, Z., Berger, M. F., Newburger, D. E., Saulrieta, K., et al. (2009).

High-resolution DNA binding specificity analysis of yeast transcription factors. *Genome Res*, 19, 556–566.



## **Chapter 2**

### ***Crx-L253X* Mutation Produces Dominant Photoreceptor Defects in *TVRM65* Mice**

## 2.1 Author Contributions

This chapter is adapted from the published manuscript: Philip A. Ruzycski<sup>1,2\*</sup>, Courtney D. Linne<sup>2\*</sup>, Anne K. Hennig<sup>2</sup> and Shiming Chen<sup>1,2,3#</sup> (2017) “*Crx-L253X* Mutation Produces Dominant Photoreceptor Defects in *TVRM65* Mice” *Investigative Ophthalmology and Visual Science*. 58. This work was done in close collaboration with Courtney Linne, who was a co-first author on the paper as well as Anne Hennig and Shiming Chen. Shiming Chen and I conceived of and planned the experiments, Courtney Linne and I performed RNA and morphological experiments, and Anne Hennig performed Electroretinograms. Courtney Linne, Shiming Chen and I wrote the manuscript.

---

\*These authors contributed equally.

#Corresponding Author

<sup>1</sup>Molecular Genetics and Genomics graduate program, Division of Biology & Biomedical Sciences, Washington University, Saint Louis, Missouri, USA

<sup>2</sup>Department of Ophthalmology and Visual Sciences, Washington University, Saint Louis, Missouri, USA

<sup>3</sup>Department of Developmental Biology, Washington University, Saint Louis, Missouri, USA

## 2.2 Abstract

Purpose: The Cone-Rod Homeobox (CRX) transcription factor is essential for photoreceptor gene expression, differentiation and survival. Human *CRX* mutations can cause dominant retinopathies of varying onset and phenotype severity. In animal models, dominant frameshift *CRX* mutations introduce a premature termination codon (PTC), producing inactive truncated proteins that interfere with normal CRX function. Previously, a mutant mouse, *TVRM65*, was reported to carry a recessive late PTC mutation, *Crx-L253X*. More detailed phenotype analysis of *Crx-L253X*'s pathogenicity sheds new light on the variability of *CRX*-linked diseases.

Methods: Homozygous (*L253X/X*), heterozygous (*L253X/+*), *Crx*<sup>-/-</sup> and control *C57BL/6J* (*WT*) mice were analyzed at various ages for changes in retinal function (ERG), morphology (histology) and photoreceptor gene expression (qRT-PCR).

Results: *L253X/X* mice lack visual function at 1 month, show greater reductions in retinal thickness and distinct gene expression changes relative to *Crx*<sup>-/-</sup>, suggesting that *L253X/X*'s phenotype is more severe than *Crx*<sup>-/-</sup>. *L253X/+* mice have reduced rod/cone function, but normal retinal morphology at all ages tested. qRT-PCR assays described a complex phenotype in which both developing and mature photoreceptors are unable to maintain proper gene expression. *L253X* mRNA/protein is overexpressed relative to normal *Crx*, suggesting a pathogenic mechanism similar to early PTC mutations. However, the overexpression is less pronounced, correlating with a relatively mild dominant phenotype.

Conclusions: The *L253X* mouse provides a valuable model for *CRX*-associated retinopathy. The pathogenicity of *CRX* frameshift mutations depends on the position of the PTC, which in turn determines the degree of mutant mRNA/protein overproduction.

## 2.3 Introduction

The Cone-Rod Homeobox protein (CRX) is a transcription factor (TF) that regulates expression of many photoreceptor genes essential for the development and maintenance of both rod and cone photoreceptors.[Freund et al., 1997; Furukawa et al., 1997; Hennig et al., 2008; Swaroop et al., 2010] Human *CRX* mutations are associated with various diseases, including Retinitis Pigmentosa (RP), Cone-Rod Dystrophy (CoRD) and Leber Congenital Amaurosis (LCA).[Rivolta et al., 2001; Swain et al., 1997] *CRX*-linked diseases are largely inherited in an autosomal dominant fashion or arise *de novo*, and vary widely in severity and age of onset.[Jacobson et al., 1998; Koenekoop et al., 2002; Nichols et al., 2010; Rivolta et al., 2001; Sohocki et al., 1998] Since *CRX* mutations can be detected with early genetic testing, we need to be able to predict their effects and to define effective treatment and gene therapy regimens.[Sohocki et al., 1998]

Previous research has divided disease causing *CRX* mutations into four classes.[Tran & Chen, 2014] Classes I and II are missense mutations that fall within or near the region coding for the DNA-binding homeodomain; Class I mutations reduce the binding of CRX to its DNA targets.[Chen et al., 2002] Previous work has shown by a variety of metrics that heterozygous mice carrying the Class I *R90W* mutation suffer from a mild form of CoRD, similar to human patients with such mutations, while homozygotes exhibit a LCA-like condition.[Tran et al., 2014] Class III and IV mutations represent frameshift mutations caused by insertions or deletions in the region coding for CRX's transactivation domains.[Chen et al., 2002] Class IV mutations, modeled by the *RIP* mouse, cause translation of a much longer peptide sequence due to a frameshift and extension of the open reading frame (ORF) into the 3' untranslated region

(3'UTR).[Roger et al., 2014] In contrast to this extended peptide sequence, Class III mutations truncate *CRX*'s ORF with a premature termination codon (PTC) that results in a CRX protein with a shortened transactivation domain and lead to a severe dominant degenerative phenotype (LCA or CoRD).[Tran et al., 2014] This phenotype has been modeled by the *E168d2* mouse [Tran et al., 2014] and *Rdy (A182d1)* cat [Ocelli et al., 2016] (Figure 2.1A). *In vitro* DNA binding and transactivation experiments determined that the E168d2 protein has a similar affinity for CRX target DNA sequences as normal CRX, but is unable to activate transcription on its own or in combination with other retinal TFs that normally synergize with CRX.[Chen et al., 2002; Tran et al., 2014] Furthermore, when tested together with normal CRX, E168d2 protein interfered with normal CRX transactivation in a dominant-negative manner.[Tran et al., 2014]

Class III PTC-causing *CRX* mutations also result in a novel untranslated region between the PTC and normal stop codon, which becomes a part of the mutant mRNA transcript's 3'UTR (Figure 2.1A).[Tran & Chen, 2014] In the *E168d2* mouse and the *Rdy (A182d1)* cat, longitudinal studies showed gradual accumulation of mutant protein and mRNA.[Tran et al., 2014];[Ocelli et al., 2016] The mechanism for this accumulation is unknown, but was postulated to arise from decreased mutant RNA degradation,[Tran et al., 2014] possibly due to the presence of cryptic stabilizing elements within the PTC-expanded mutant 3'UTR. Therefore, the position of the PTC (relative to the normal termination codon) could determine the stability, and thus abundance of mutant RNA.

Previously, another PTC mutant mouse, *TVRM65 (Crx-L253X)*, was discovered in a chemical-induced mutagenesis screen conducted by The Jackson Laboratory[Won et al., 2011] (see Appendix 2 Methods section for nomenclature specifics). The *L253X* mutation (Figure 2.1A) occurs later (closer to the normal termination codon) than *E168d2* and *A182d1*. *L253X*

would be a candidate to model late PTC Class III mutations, such as *L237iI* [Huang et al., 2012] and *Q256X* [Lu et al., 2015] associated with dominant CoRD. However, the initial histological assessments suggested that the *L253X* mouse phenocopies the recessive null (*Crx*<sup>-/-</sup>) condition.[Won et al., 2011] Here, we describe more detailed phenotypic analysis of *L253X* mice, which demonstrates that this mouse models a new dominant but mild form of Class III disease. Our results support the hypothesis that Class III mutation pathogenicity depends on the position of the mutation-induced PTC (a late PTC produces a less severe phenotype than an early PTC), which correlates with the abundance of mutant protein products. These findings have implications for predicting human *CRX* disease severity and progression.

## 2.4 Results

### 2.4.1 *L253X* overproduces mutant mRNA/protein in affected retinas

A hallmark of Class III *CRX* frameshift mutations is the overexpression of the mutant allele.[Ocelli et al., 2016; Tran et al., 2014] To determine if *L253X* retinas accumulate mutant protein, we performed quantitative Western blots on total protein extracts from retinas of *WT* and heterozygous *L253X/+* mice at postnatal day 10 (P10), 1 and 2 months old (mo), and homozygous *L253X/X* mice at P10 only (because photoreceptors degenerate before later ages).[Won et al., 2011] Immunoblotting with an anti-CRX antibody revealed a truncated *L253X* protein that ran faster than the normal full-length CRX protein in samples from both heterozygous and homozygous mutant retinas (Figure 2.1B). We quantified the total amount of CRX protein and of each isoform. Both *L253X/+* and *L253X/X* displayed a 1.5 to 2.2-fold increase in total CRX protein compared to the *WT* retinas (Figure 2.1C). The increases can be attributed to an accumulation of mutant protein (black bar), since normal CRX protein levels (grey bar) in *L253X/+* retinas with one normal allele were roughly half of the total CRX protein present in *WT* retinas with two alleles.

To determine if the increased mutant protein corresponded with an increase in mutant mRNA, we quantified the relative allelic expression of *L253X* and normal *Crx* mRNA using droplet digital PCR (ddPCR). First, using mutant and *WT* mouse DNA, we developed a ddPCR assay that clearly distinguished the *L253X* and normal *Crx* sequences (Figure 2.1D, DNA control). This assay was then used to quantify allele-specific mRNA species. At all ages tested, the mutant transcript was overrepresented, accounting for 57% of total *Crx* mRNA (Figure 2.1D). Total *Crx* transcript levels quantified by conventional qRT-PCR (Supplemental Figure

2.1) showed a statistically significant total *Crx* mRNA increase in 1 mo *L253X/+* retinas relative to *WT*, although levels were comparable to *WT* retinas during development at P10 and by 3 mo. Overall, *L253X/+* retinas develop and maintain an elevated level of mutant mRNA that translates into an imbalance of the two protein forms, but this misregulation is not as severe as observed in other Class III mutants. These results identify *Crx-L253X* as a *bona fide* Class III mutation, but suggest that the position of the PTC determines the toxicity of the allele.

#### **2.4.2 *L253X/X* mice have no detectable rod and cone function**

Homozygous animals carrying Class III or null *Crx* mutations lack measurable photoreceptor function.[Furukawa et al., 1999; Occelli et al., 2016; Tran et al., 2014] To confirm that *L253X/X* photoreceptors are functionally compromised, we quantified rod/cone light responses in *L253X/X* and *WT* control mice at 1 mo using electroretinogram (ERG). Even at the highest light intensity stimuli, *L253X/X* retinas displayed very little response (Figure 2.2A-C). These results confirm that, like other Class III models, homozygous *L253X/X* animals lack photoreceptor visual responses, resembling human LCA.

#### **2.4.3 *L253X/X* photoreceptors degenerate earlier than *Crx*<sup>-/-</sup>**

Other Class III mutations show a greater impact on retina morphology and function than seen in *Crx*<sup>-/-</sup> mice.[Furukawa et al., 1999; Occelli et al., 2016; Tran et al., 2014] The original report noted that degeneration was more rapid in *L253X/X* than *Crx*<sup>-/-</sup> retinas, but suggested differences in mouse background might be a factor.[Won et al., 2011] To directly compare the impact of the *L253X* mutation with that of complete loss of *Crx*, we collected retinas from 1 and 3 mo *WT*, *L253X/X*, and *Crx*<sup>-/-</sup> mice (all backcrossed 10 generations onto *C57BL/6J*) and



assessed morphological changes in H&E stained sagittal sections through the optic nerve head (Figure 2.2D-K). At 1 mo both *L253X/X* and *Crx*<sup>-/-</sup> retinas lacked photoreceptor outer segments (OS) and had thinner Outer Nuclear Layers (ONL) than the *WT* control (Figure 2.2E&F vs. D). However, thinning was more pronounced in *L253X/X* than *Crx*<sup>-/-</sup>: *L253X/X* retina had only 4-5 rows of ONL nuclei, while *Crx*<sup>-/-</sup> maintained 7-8 rows at the same age (Figure 2.2E vs. F). Quantitative morphometry measures showed that the ONL of *L253X/X* was significantly thinner than that of *Crx*<sup>-/-</sup> at multiple positions (Figure 2.2G, black vs. solid grey). By 3 mo, both models had similarly degenerated retinas with only 2-3 rows of ONL cells left (Figure 2.2I vs. J; 2K). Since both mutants showed normal ONL thickness at P7-P10 (Supplemental Figure 2.2A&B and Tran et al.[Tran et al., 2014]), these results suggest that ONL degeneration occurs earlier or at a faster pace in *L253X/X* mice than in *Crx*<sup>-/-</sup>.

In addition to progressive ONL loss, *L253X/X* also underwent age-dependent thinning of the Outer Plexiform Layer (OPL) (Figure 2.2E&I, Supplemental Figure 2.2D, I, N, and Won et al.[Won et al., 2011]). In contrast, non-photoreceptor cell layers (e.g. INL, GCL) were largely unchanged in *L253X/X* compared to *WT* controls at various ages (Supplemental Figure 2.2C-Q). These results suggest that the *L253X* mutation mainly affects photoreceptor structural integrity and survival.

#### **2.4.4 *L235X/X* mice exhibit photoreceptor gene misregulation distinct from *Crx*<sup>-/-</sup>**

To further examine if the more severe morphological phenotype of *L253X/X* retinas reflected greater changes in CRX target gene expression, we collected RNA from *L253X/X* and *Crx*<sup>-/-</sup> retinas at P10 during photoreceptor terminal differentiation prior to cell death.[Tran et al.,

2014] Transcripts of 14 known CRX-dependent genes were quantified by qRT-PCR and compared both with *WT* levels (Supplemental Figure 2.2R) and between the two models (Figure 2.3A). Although most of these genes were dysregulated in both mutants compared to *WT* mice (Supplemental Figure 2.2R), half of them showed statistically significant expression differences between *L253X/X* and *Crx*<sup>-/-</sup> (Figure 2.3A).[Ruzycki et al., 2015; Tran et al., 2014] Three of these, *Gnat1*, *Grk1*, and especially *Rho* differed markedly (>2 fold) between the two models (Figure 2.3A). Such differences in essential photoreceptor gene expression likely contribute to the morphological and phenotype differences between the two models.

To confirm qRT-PCR results at the protein level, we performed immuno-fluorescent (IHC) staining for RHO on retinal sections of 1 mo *L253X/X*, *Crx*<sup>-/-</sup> and *WT* control mice. RHO was chosen because it displayed the largest difference in the RNA level between *L253X/X* and *Crx*<sup>-/-</sup>. As expected, in *WT* retina, RHO was predominantly localized to the rod outer segments (OS) with faint staining seen in ONL cell bodies (Figure 2.3B). In contrast, *L253X/X* retina displayed no detectable RHO immunostaining (Figure 2.3C), while *Crx*<sup>-/-</sup> retina still had weak and spotty RHO signals localized to individual ONL cell bodies and inner segments (Figure 2.3D). These results supported the RNA expression data and suggest that *L253X* mutant protein negatively affects transcription.

Overall, the comparison between *L253X/X* and *Crx*<sup>-/-</sup> indicates that *L253X/X* disease is more severe than that resulting from loss of *Crx*, supporting a dominant effect of the mutation.[Jacobson et al., 1998; Tran et al., 2014] The more severe photoreceptor deficits in *L253X/X* thus presumably result from antimorphic activity of the truncated CRX protein.

### **2.4.5 Heterozygous *L253X/+* mice show mild decreases in rod and cone function**

The *L253X* (*TVRM65*) mutation was originally reported to cause recessive retinal disease.[Won et al., 2011] However, if the mutant protein possesses antimorphic activity, we would hypothesize that one copy would produce a photoreceptor phenotype. To test this, we investigated photoreceptor function in *L253X/+* retinas by assessing electrical responses to whole retina light stimulation. ERGs recorded at 1, 2, 3, and 5 mo revealed *L253X/+* mice exhibit mild but significant reductions in ERG amplitudes compared to *WT* mice (Figure 2.4). At all ages tested, dark-adapted A-waves were reduced in *L253X/+* mice, especially at high stimulus intensities (Figure 2.4A-D, black line vs. grey line). Dark-adapted B-waves of *L253X/+* mice were comparable to *WT* at 1 mo (Figure 2.4E), but significantly reduced at 2, 3, and 5 mo (Figure 2.4F-H, black line vs. grey line). The degree of A-wave or B-wave reductions appeared to be constant, suggesting that the defects were unlikely a result of progressive degeneration. Reduction in cone function was also observed: Light-adapted B-waves were significantly decreased in *L253X/+* mice at 2, 3 and 5 mo (Figure 2.4J-L), but not at 1 mo (Figure 2.4I). These results suggest that *L253X/+* mice have deficits in both rod and cone function, resembling a mild retinopathy.

### **2.4.6 *L253X/+* retinas do not degenerate**

To test whether the functional defects correspond to morphological changes, we measured retinal thickness in H&E stained retinal sections, comparing *L253X/+* and *L253X/X* with *WT* mice at three ages (Figure 2.5). Despite the progressive retinal degeneration seen in *L253X/X* mice (Figure 2.5C, D, G, H, K & L; Supplemental Figure 2.2C-Q), *L253X/+* showed

no degeneration or morphological abnormalities compared to *WT* controls up to 3 mo of age (Figure 2.5B, D, F, H, J, & L; Supplemental Figure 2.2C-Q).

#### **2.4.7 *L253X/+* retinas display dynamic changes of photoreceptor gene expression**

To examine developmentally regulated photoreceptor gene expression in *L253X/+* vs. *WT* mice, we performed qRT-PCR analyses at two early ages: P10, when rods/cones are still terminally differentiating, and 1 mo when rods/cones are fully mature in normal retinas. The genes tested were previously shown to dramatically change in expression during *WT* rod and cone development.[Kim et al., 2016] Figure 2.6A shows expression changes of these genes in *L253X/+* relative to *WT* mice at both ages. Gene expression conformed to three general trends in addition to the slight increase in total *Crx* expression (dashed black line) noted above: *Gnat1*, *Rho*, and *Sop* (solid black lines), markers of mature photoreceptors, were significantly decreased in *L253X/+* at P10 but attained *WT* levels by 1 mo. *Rxrg* (dotted grey line), which in *WT* retina loses expression over development,[Kim et al., 2016] was elevated in *L253X/+* at P10 but decreased closer to *WT* levels by 1 mo. The other genes tested, which increase over development in *WT* retina, showed similar relative expression levels at both ages that did not vary considerably from *WT* (solid grey lines). Together, these gene expression trends suggest that *L253X/+* retinas exhibit altered gene expression early in photoreceptor development, similar to other *Crx* mutant models.[Ruzycki et al., 2015]

To determine if *L253X/+* retinas maintain photoreceptor gene expression after differentiation, we used qRT-PCR to examine expression of 17 essential rod and cone genes in *WT* and *L253X* mutants at other adult ages. The results at 1, 2, 3 and 5 mo (presented in a

heatmap in Figure 2.6B) suggest a complex phenotype. Genes coding for transcription factors *Crx*, *Rxrg*, *Otx2*, and *Nrl* were upregulated at many of the time points tested. The expression of essential rod and cone phototransduction genes showed a less clear pattern across ages, but in general all exhibit phases of reduced expression relative to *WT*.

#### **2.4.8 Late PTC-caused C-terminal truncation reduces CRX's transactivation function**

As an initial step to understand the molecular mechanism underlying mis-regulation of gene expression in *L253X* mutant retinas, we asked whether a short C-terminal truncation caused by a late PTC affects CRX's ability to transactivate target gene expression. This was tested using transient transfection assays in HEK293T cells with a Rhodopsin (*Rho*) promoter-luciferase reporter (*BR130-luc*) [Chen et al., 2002]. Vectors expressing either the full-length (normal) or truncated CRX<sup>1-254</sup> were co-transfected with a NRL vector to measure their ability to transactivate the *BR130-luc* reporter. CRX<sup>1-254</sup> did show some dose-dependent ability to activate the *Rho* promoter, suggesting that it does bind DNA and interacts with NRL, but the maximum activity was greatly reduced compared to the normal CRX protein (~30% of normal activity) (Figure 2.7).

## 2.5 Discussion

### 2.5.1 *L253X* mouse provides a new Class III model for mild dominant cone-rod dystrophy

Although *L253X* (*TVRM65*) was originally identified as a recessive *Crx* mutation [Won et al., 2011] several pieces of evidence suggest that *L253X* is antimorphic with dominant inheritance. First, *L253X* retinas overexpress the mutant gene product, a hallmark of Class III mutations that amplifies the mutation's effects. Second, the photoreceptors in homozygous *L253X/X* mice degenerate earlier than *Crx*<sup>-/-</sup>, and show patterns of photoreceptor gene dysregulation that are distinct from *Crx*<sup>-/-</sup>. Thus, the presence of the mutant *L253X* protein causes more severe pathology than absence of CRX altogether, which is consistent with the antimorphic effect of other Class III mutations [Ocelli et al., 2016; Tran et al., 2014; Tran & Chen, 2014] Furthermore, heterozygous *L253X/+* mice have early gene expression differences that largely resolve by 1 mo, suggesting that full maturation may be delayed in these retinas. Although gene expression changes in adults were modest, rod and cone functional deficits were detectable as early as 1 mo. Previous work on Class III mutants has emphasized the effects that small changes in gene expression over a large number of genes can have on photoreceptor function. [Ruzycski et al., 2015] Collectively, these findings identify *L253X* as a Class III *Crx* mutation with the phenotypes summarized in Table 2.1.

Compared to previously reported Class III mutations, *L253X*'s pathogenicity is relatively mild. Unlike *E168d2/+* photoreceptors, which lose their OS and undergo progressive degeneration, [Tran et al., 2014] *L253X/+* photoreceptors show normal morphology and do not degenerate. The ERG deficits do not worsen with age relative to *C57BL/6J* controls. The gene expression changes are also relatively modest with a multiphasic nature. This evolving

expression pattern with increased levels of many TFs, but modest depletion of phototransduction cascade components, likely reflects an inability of the cells to maintain a proper homeostasis. This data would suggest that unknown feedback mechanisms are operational but unable to achieve the necessary precision. In adults, *L253X/+* photoreceptors are alive and functional through 5 mo, but it remains to be determined if the phenotype worsens at older ages or if the retina is more susceptible to environmental insults as reported with other Class III mutants. Nevertheless, given the huge phenotypic variability of human *CRX* diseases, the *L253X* mouse provides a valuable animal model for understanding the pathogenesis of Class III mutations and the underlying molecular mechanisms that determine phenotype severity, age of onset and disease progression.

### **2.5.2 The molecular mechanism(s) underlying *L253X* pathogenicity**

The pathogenicity of Class III mutations is generally determined by two factors, the activity and dose of the mutant protein. *L253X* protein contains an intact DNA binding domain but lacks a portion of the C-terminal activation domain (Figure 2.1A). Thus, it likely retains the ability to bind DNA, but loses some aspect of transactivation function. Transient transfection assays showed that  $CRX^{1-254}$ , a truncated *CRX* protein similar to *L253X*, retained only 30% of the normal *CRX* activity, very similar to the published reports.[Chen et al., 2002; Roger et al., 2014] By comparison, the Class III mutant protein E168d2 showed 25-30% of the transactivation activity under similar conditions.[Chen et al., 2002; Tran et al., 2014] Previous reports have also shown  $CRX^{1-254}$  binds appropriate DNA sequence elements very similar to full-length *CRX*. [Roger et al., 2014] It is conceivable that, as previously reported for E168d2, [Tran et al., 2014] *L253X* protein may compete with normal *CRX* to bind to target genes, interfering with

normal CRX function in heterozygous mice. However, transient transfection assays testing this hypothesis did not show this effect on the *Rho*-promoter luciferase reporter (data not shown). The antimorphic competition may not be limited to CRX but could also affect other homeodomain transcription factors that bind to similar DNA targets, and as such transient cell transfection assays on a single reporter that lacks the native chromatin context and regulatory architecture may not be the proper experiment. Supporting this possibility, the retinas of *L253X/X* mutants, like *E168d2/d2*, [Tran et al., 2014] showed more severe degeneration and biochemical defects than *Crx*<sup>-/-</sup> retinas. Other potential targets could include OTX2, [Samuel et al., 2014] which is upregulated in *L253X* retinas (Figure 2.6B). Additional experiments are required to demonstrate *L253X*'s dominant-negative activity on CRX and OTX2. Despite the similarity between *L253X* and *E168d2*, *L253X/+* mice have a remarkably milder phenotype than *E168d2/+*; it is unlikely that the <10% difference in transactivation activity can account entirely for the significant difference between the two models. We therefore propose that the relative amounts of mutant and normal CRX protein must also play a role in phenotype severity.

### **2.5.3 *L253X* allele-specific overproduction of mutant mRNA and protein has implications for the PTC position effect**

Careful analysis of *Crx* mRNA and protein in *L253X/+* and *L253X/X* retinas clearly shows an increased ratio of mutant/normal protein and mRNA. This finding supports the hypothesis that, like other Class III models, *L253X* retinas accumulate mutant transcripts, leading to overproduction of the CRX mutant protein. Interestingly, however, the degree of mutant RNA accumulation in *L253X/+* is lower than that of the other Class III models carrying early PTC mutations. *L253X* mice only display a modest accumulation (57% mutant / 43% normal),



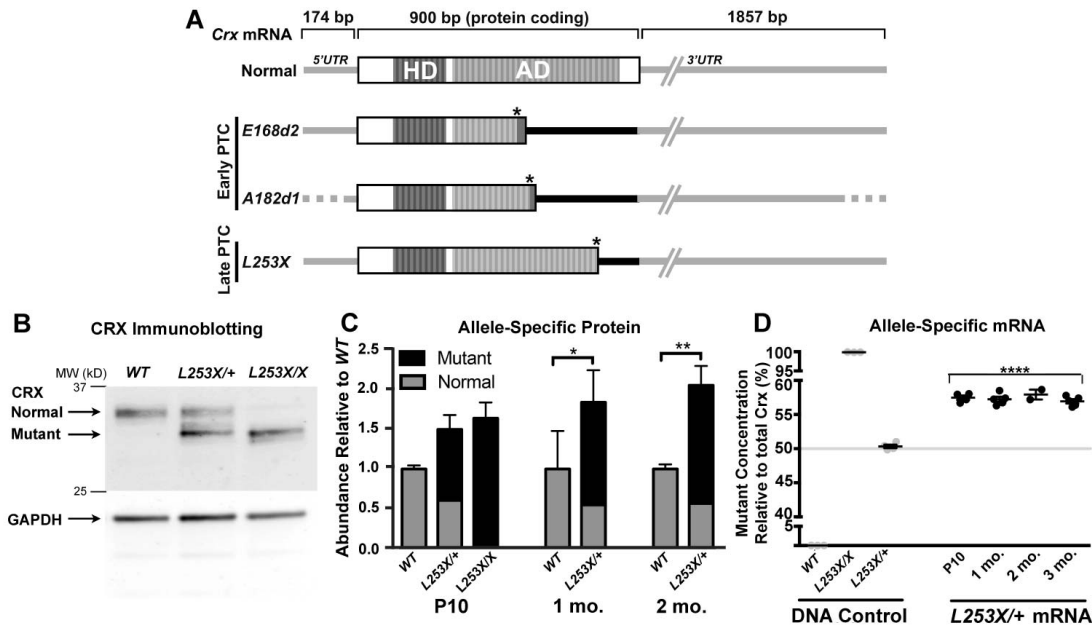
especially when compared to that observed in *E168d2/+* mice at the same age (85% mutant / 15% normal).[Tran et al., 2014] Also, when considered with the quantification of total *Crx* mRNA, the *L253X/+* must result in a downregulation of the normal allele as there is no aggregate overproduction of *Crx*. This suggests that the *L253X* mRNA is still ‘sensed’ by the feedback regulatory mechanism, which functions normally in the mutant retina. Future studies could probe this feedback mechanism by comparing to the *E168d2/+* retinas where although there is a gross increase in total *Crx* mRNA, the normal allele maintains normal expression levels.[Tran et al., 2014] Both mouse models show overproduction of mutant protein relative to *WT*, but *L253X* accumulates at a slower rate than *E168d2*.

The mechanism for allele-specific overproduction of mutant *Crx* in Class III models is not well understood, but likely stems from increased RNA stability due to the presence of the PTC in the last exon. This hypothesis further predicts that the position of the PTC influences the rate of mutant RNA accumulation. Indeed, *L253X* with a later PTC showed less RNA accumulation than the two early PTC alleles, *E168d2* and *A182d1* (Figure 2.1). Thus, the earlier the PTC is, the more mutant product accumulates. These differences between early vs. late PTCs may be attributed to the length of PTC-determined 3’*UTR*, within which multiple discrete elements could independently and additively contribute to mRNA hyperstability. Future identification of novel RNA regulatory elements shared among different PTC mutations may shed light on the mechanisms for mutant transcript hyperstability and phenotype variability in Class III *CRX* disease. It is also notable that the accumulation of mutant protein in *L253X* retinas is more prominent than mutant RNA, raising the new possibility that altered protein turnover rate could also contribute to mutant protein accumulation.

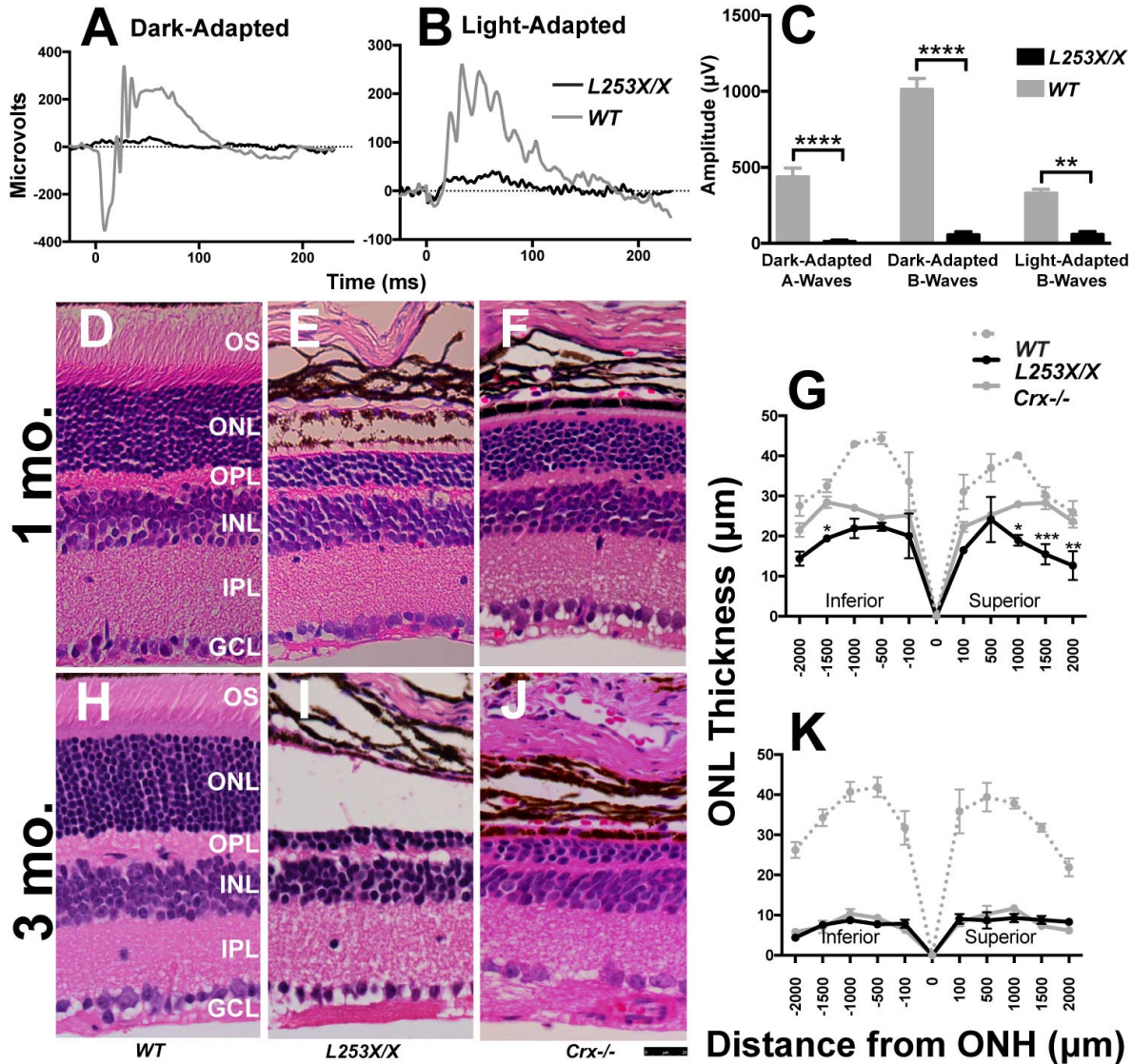
In conclusion, our in-depth characterization of the *Crx-L253X (TVRM65)* mouse supports the classification of *L253X* as a dominant Class III *Crx* mutation and provides a new animal model for understanding *CRX*-associated dominant retinopathies. The insights gained will hopefully lead to new treatment strategies for this complex disease.

**Table 2.1. Summary of *L253X* related phenotypes**

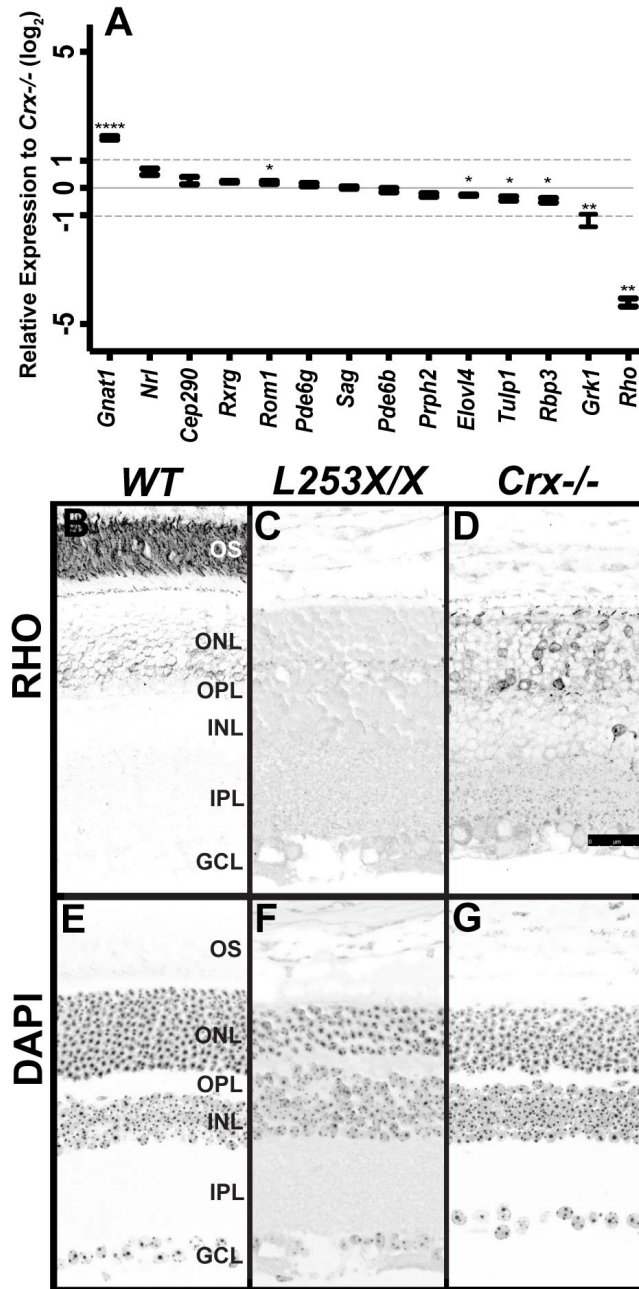
<b>Phenotype</b>	<b><i>Crx-L253X/X</i></b>	<b><i>Crx-L253X/+</i></b>
<b>Morphology (histology)</b>	ONL degeneration earlier than <i>Crx</i> <sup>-/-</sup>	No detectable abnormality up to 5 mo
<b>Function (ERG)</b>	Lack of rod and cone responses to light	Minor reductions in rod and cone function
<b>Gene expression (qRT-PCR)</b>	Altered gene expression distinct from <i>Crx</i> <sup>-/-</sup>	Dynamic changes from <i>WT</i>
<b>CRX protein/RNA levels</b>	<i>L253X</i> protein/RNA overproduction	Increased ratio of <i>L253X</i> vs Normal



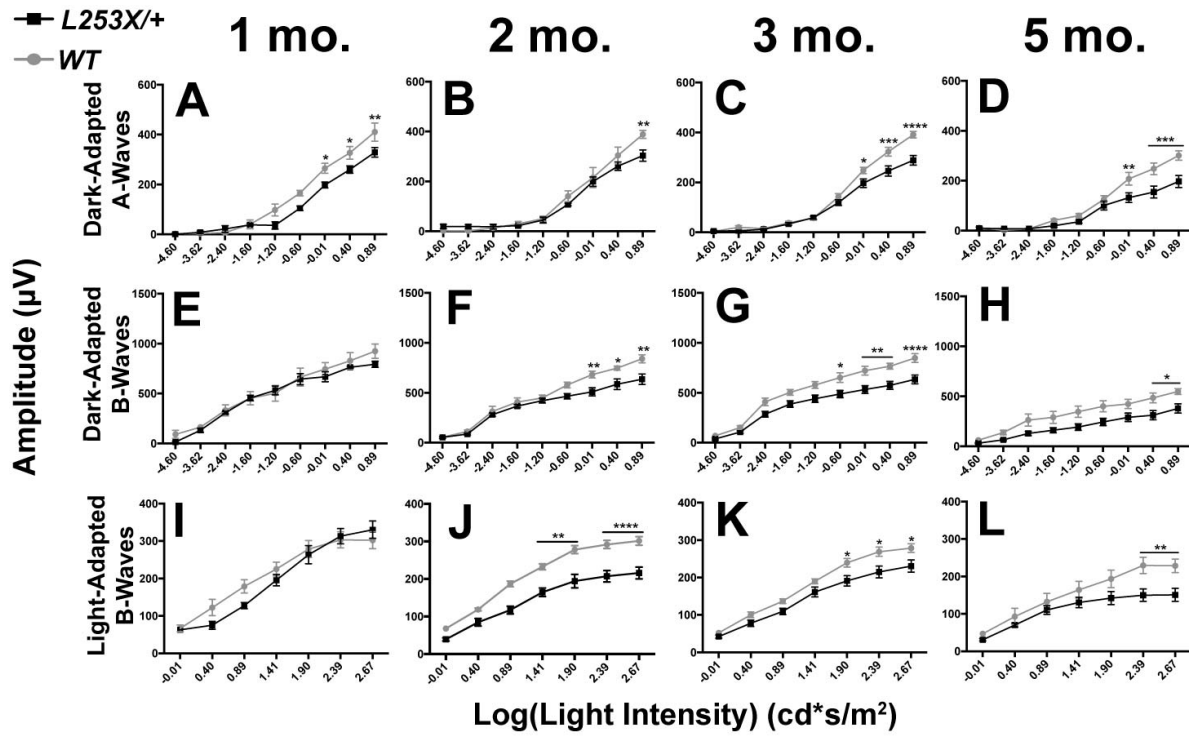
**Figure 2.1: Class III *Crx* mutations introduce a premature termination codon (PTC), resulting in accumulation of truncated CRX proteins and mutant mRNA with variable extended 3'UTR.** (A) Schematic representation of normal and mutant CRX mRNA species, showing untranslated (*UTR*, grey/black lines) and protein-coding (open/striped boxes) with their relative sizes (bp, base pairs), based on mouse (Normal, *E168d2*[Tran et al., 2014] and *L253X*) and cat (*A182d1*)[Occelli et al., 2016] models. The DNA-binding homeodomain (HD, dark strips) and the activation domain (AD, light strips) sequences[Chen et al., 2002] are indicated. All three mutant mRNAs encode the complete HD, but ADs are truncated at different positions due to their PTC (indicated as the right-hand end of the protein coding box) induced by the mutation (marked by \*). These PTCs also expand the 3'UTR to include different lengths of the original coding region (indicated by black line) in front of the normal 3'UTR (grey line). The position of the PTC relative to the normal stop codon determines the length of expanded 3'UTR: Early PTC mutants, *E168d2* and *A182d1* (*Rdy* cat mutation) produce a longer 3'UTR than the late PTC mutant *L253X*. Since 5'UTR and 3'UTR sequences of feline *Crx* mRNA are not available, these regions of the *A182d1* transcript are indicated by dotted lines. (B) Immunoblot of CRX in the retinas of *WT*, *L253X/+*, and *L253X/X* mutants at 1 month old (mo) with GAPDH as a protein loading control. Black arrows indicate the running position for the normal and mutant (truncated) forms of CRX. (C) Quantification of relative amounts of normal and mutant CRX proteins in the indicated retinas at three ages (mean  $\pm$  SEM,  $n \geq 3$ ; \*  $p < 0.05$ , \*\* $p < 0.01$ ; 2-Way Anova with Tukey's Multiple Comparisons Test). (D) Droplet digital PCR (ddPCR) quantification of Mutant (*L253X*) and Normal *Crx* mRNA as percentage of total *Crx*. The specificity of the two allele-specific ddPCR assays was established using triplicate tail DNA samples (grey circles) from the indicated genotypes (DNA control). *L253X/+* mRNA results are presented as percent of total *Crx* mRNA (Mutant plus Normal transcripts) (mean  $\pm$  SEM,  $n \geq 3$ ; \*\*\*\* $p < 0.0001$ ; Unpaired t-Test for mutant mRNA level relative to *L253X/+* DNA control;  $n \geq 3$ ).



**Figure 2.2: *L253X/X* retinas lack rod and cone responses to light and degenerate earlier than *Crx<sup>-/-</sup>*.** Electroretinogram (ERG) was used to measure rod (dark-adapted) and cone (light-adapted) responses to various light stimuli in 1 mo *L253X/X* and *WT* mice. (A, B) Comparison of representative ERG traces at the highest light intensity: 0.89 log cd\*s/m<sup>2</sup> for dark-adapted responses, and 2.67 log cd\*s/m<sup>2</sup> for light-adapted responses. (C) Bar graphs compare mean A-wave and/or B-wave amplitude differences between the two genotypes at the highest light intensity tested (mean ± SEM, n ≥ 3; \* p≤0.05, \*\*\* p≤0.0005, \*\*\*\* p<0.0001; 2-Way Anova with Sidak's Multiple Comparisons Test). (D-F, H-J) H&E-stained paraffin-embedded sagittal retinal sections from the indicated mice at 1 and 3 mo. Images were taken approximately 500μm from the Optic Nerve Head (ONH) at a 40X magnification (scale bar = 25μm). (G,K) Morphometry measures at five set distances (in microns) on either side of the ONH (mean ± SEM, n ≥ 3; \* p≤0.05, \*\* p<.01, \*\*\* p≤.0003; 2-Way Anova with Sidak's Multiple Comparisons Test).

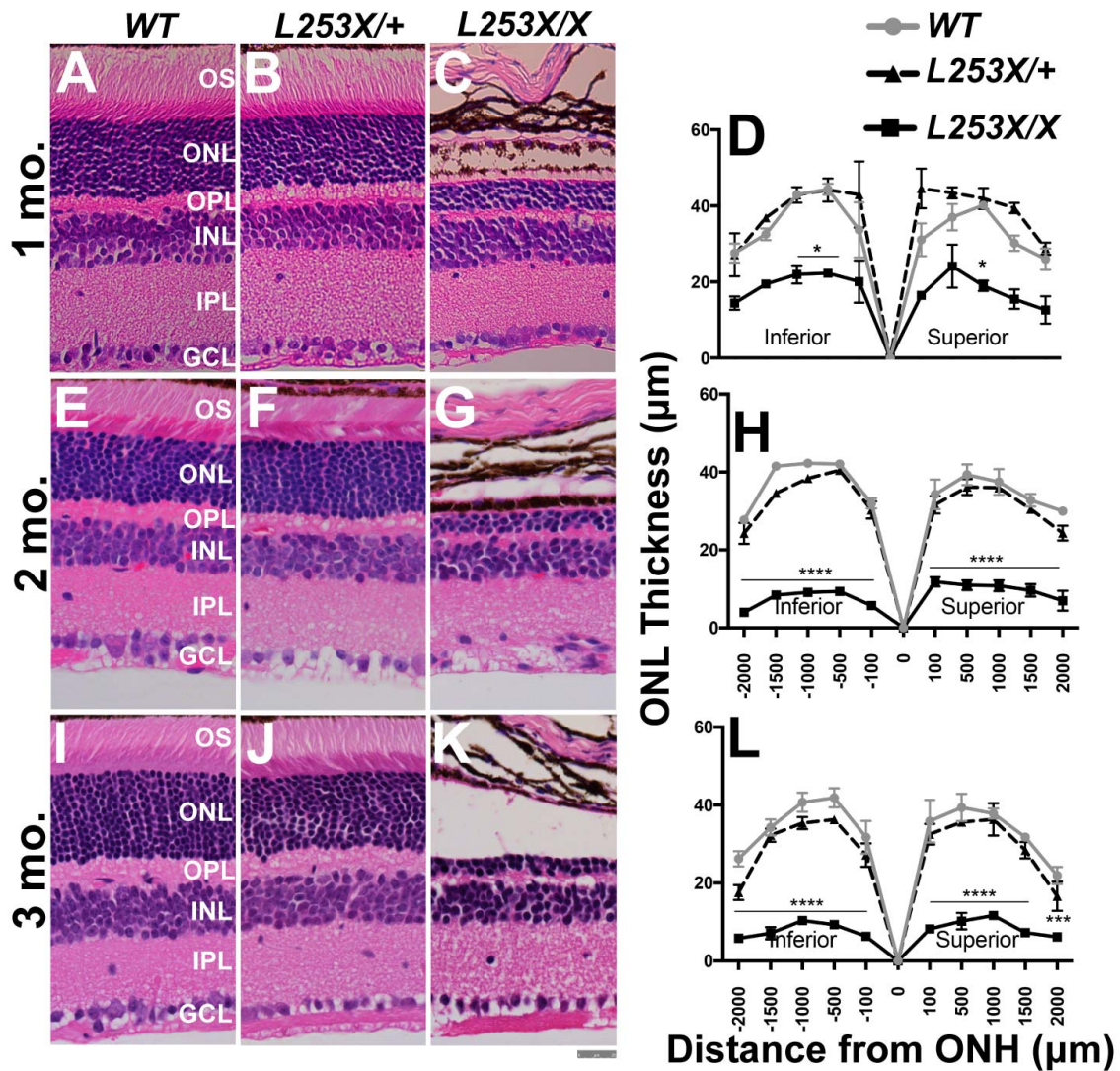


**Figure 2.3: *L253X/X* has measurable gene expression differences relative to *Crx*<sup>-/-</sup>.** (A) Expression at P10 of the indicated photoreceptor genes was measured by qRT-PCR and compared between *L253X/X* and *Crx*<sup>-/-</sup> at P10. The grey dash lines mark 2-fold differences ( $\pm 1$  log<sub>2</sub> relative expression) between the two genotypes (mean  $\pm$  SEM, n  $\geq$  3; \* p<0.05, \*\* p<0.01, \*\*\*\* p<0.0001; Unpaired t-Test with Welch's Correction). (B-D) RHO immunostaining of retinal sections of the indicated mice at 1 mo revealed less RHO expression in *L253X/X* retinas than *Crx*<sup>-/-</sup> (signal in black, scale bar 25um). (E-G) DAPI counterstaining of the above sections (signal in black).



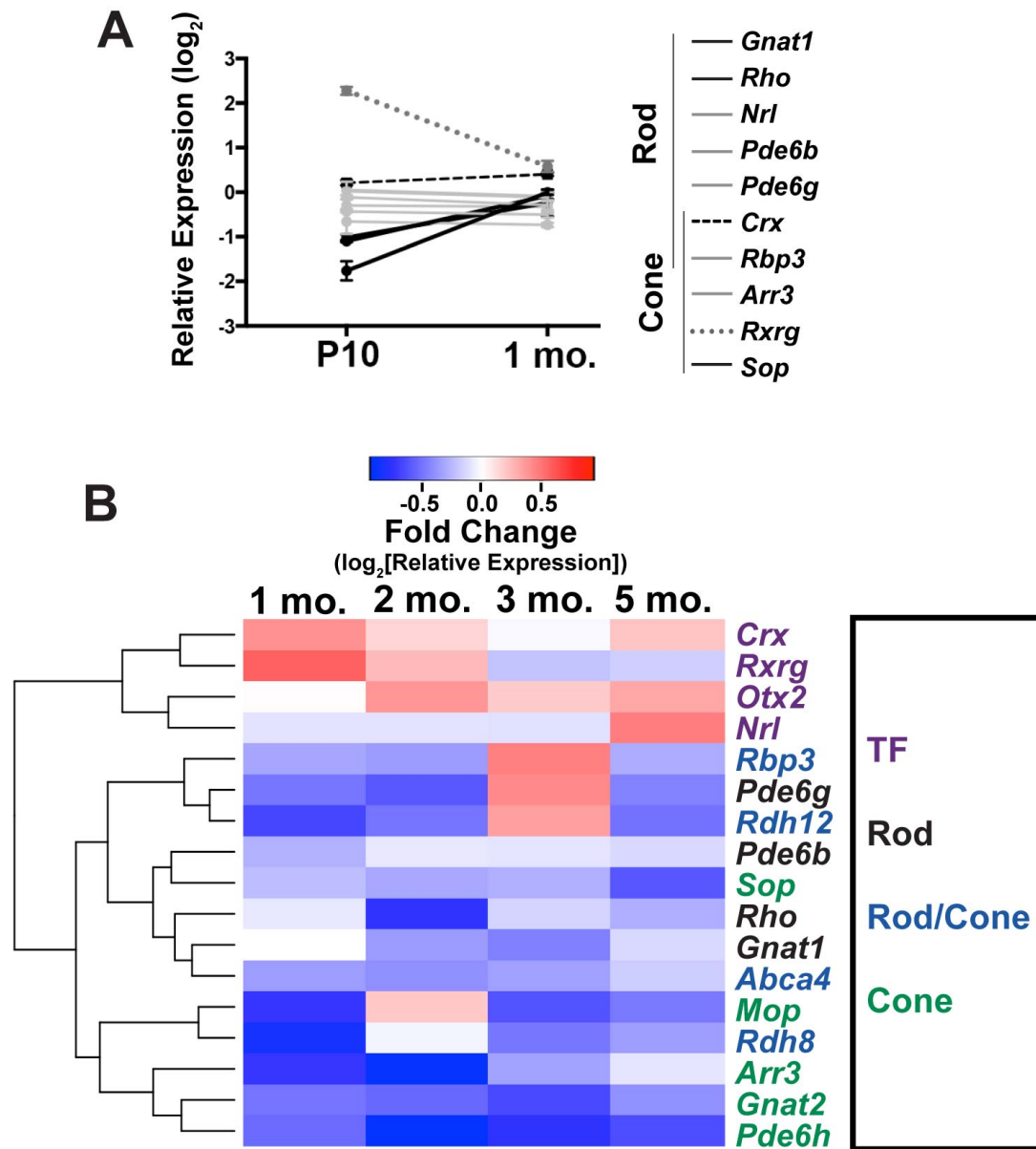
**Figure 2.4: *L253X/+* retinas exhibit reduced rod and cone light responses.** ERG responses of *L253X/+* mice (black lines) and *WT* (gray lines) were tested at 4 adult ages, and dark-adapted and light-adapted A- and B-wave amplitudes quantified. 1 mo (A, E, I), 2 mo (B, F, J), 3 mo (C, G, K), and 5 mo (D, H, L) (mean  $\pm$  SEM,  $n \geq 3$ ; \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.0006$ , \*\*\*\*  $p < 0.0001$ ; 2-Way Anova with Sidak's Multiple Comparisons Test).





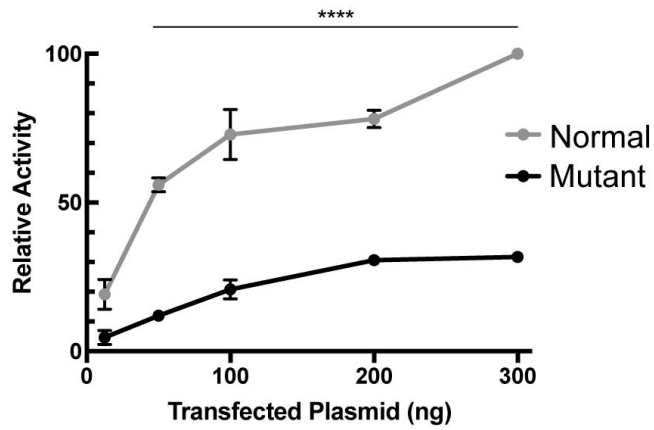
**Figure 2.5: *L253X/+* retinas appear morphologically similar to *WT* up to 3 mo.** H&E stained retinal sections from *L253X/+*, *L253X/X*, and *WT* mice at the indicated ages (**A-C**, 1 mo; **E-G**, 2 mo; **I-K**, 3 mo) were imaged at 40X magnification (Scale bar 25µm). Images were taken approximately 500 microns from the ONH, where differences in layer thickness are less significant. (**D,H,L**) ONL thickness was quantified by morphometry at five set distances (in microns) on either side of the ONH (mean  $\pm$  SEM,  $n \geq 3$ ; \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ ; 2-Way Anova with Tukey's Multiple Comparisons Test).



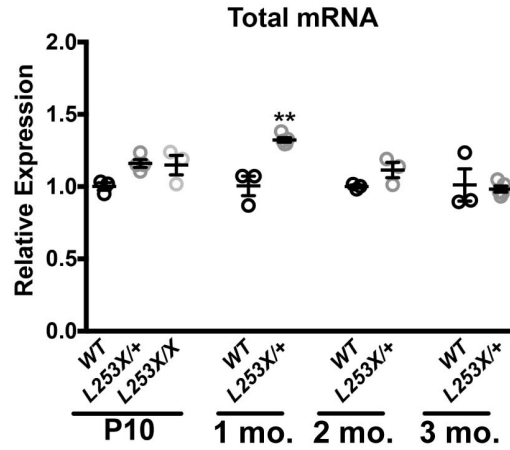


**Figure 2.6: *L253X/+* retinas display dynamic changes in photoreceptor gene expression.**

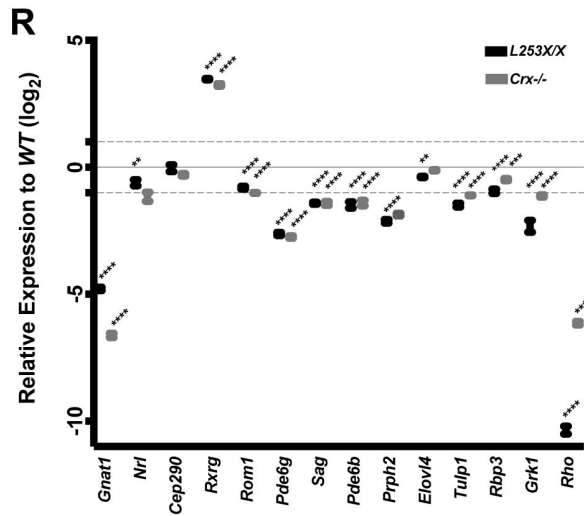
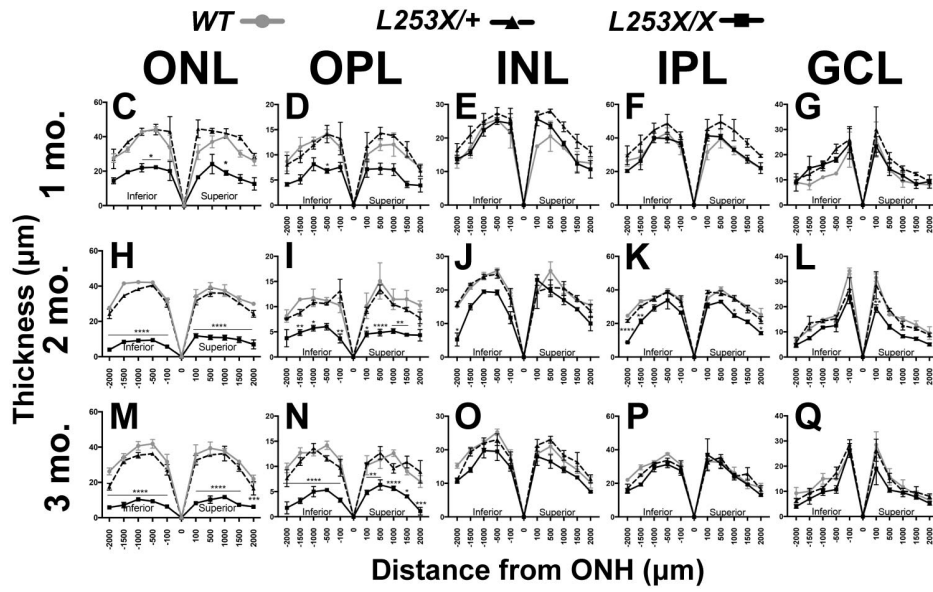
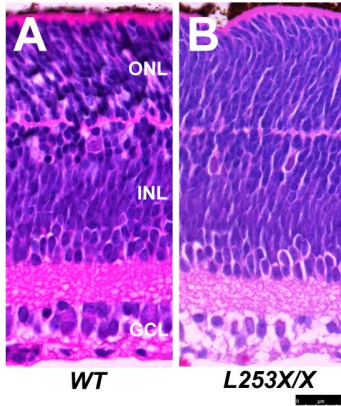
qRT-PCR analyses of *L253X/+* expression levels (relative to *WT* mice) of the indicated rod/cone genes are shown at P10 and 1 mo (mean  $\pm$  SEM,  $n \geq 3$ ). Lines connecting expression levels of each gene at the two ages show the dynamic changes in expression, and reveal four different expression trends indicated here by line color and solidity, as described in Results. **(B)** Heatmap depicting qRT-PCR-detected expression changes of the listed photoreceptor genes in *L253X/+* relative to *WT* at the indicated ages. Data are clustered based on gene expression patterns ( $\log_2$  [mean fold-change],  $n \geq 3$ ). Colored gene names indicate expression in specific photoreceptor subtype(s). “TF” indicates transcription factors expressed by rods and/or cones that are essential for cell fate specification.



**Figure 2.7: L253X retains minimal ability to transactivate Rhodopsin promoter in presence of NRL .** Dual luciferase assays showing dose-dependent increase in CRX transactivation activity of normal or mutant (CRX<sup>L253X</sup>) in the presence of constant amount of NRL. (mean  $\pm$  SEM, n = 3; \*\*\*\* p<0.0001; 2-Way Anova with Bonferroni's Multiple Comparisons Test)



**Supplemental Figure 2.1: *L253X/+* retinas display variable excess total *Crx* mRNA.** qRT-PCR assays were used to quantify total *Crx* mRNA from mutant and *WT* control retinas at the indicated ages. The results are presented as relative expression to *WT*. (mean ± SEM, n=3, circles represent individual biological replicates; \*\* p< .01; 1-Way Anova with Tukey's Multiple Comparisons Test).



**Supplemental Figure 2.2: *L253X/X* retinas display normal ONL genesis but photoreceptor layer-specific progressive thinning, while all *L253X/+* retinal layers unaffected.** H&E stained retinal sections from P7 *WT* (A) and *L253X/X* (B) mice were imaged at 40X magnification (scale bar 25um). Morphometry quantification of five retinal layers in *L253X/X* and *L253X/+* mutants along with *WT* controls at 1 mo (C-G), 2 mo (H-L) and 3 mo (M-Q). These layers include ONL-outer nuclear layer, OPL-outer plexiform layer; INL-inner nuclear layer; IPL-inner plexiform layer; and GCL- ganglion cell layer. (mean  $\pm$  SEM,  $n \geq 3$ ; \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ ; 2-Way Anova with Tukey's Multiple Comparisons Test). (R) Expression levels of 14 essential photoreceptor genes in P10 *L253X/X* and *Crx*<sup>-/-</sup> mutants and *WT* controls were determined by qRT-PCR. The results are presented as expression changes relative to *WT* in a  $\log_2$  scale with 2-fold differences ( $\pm 1$ ) marked by grey dash lines (mean  $\pm$  SEM,  $n \geq 3$ ; \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*\*  $p < 0.0001$ ; 1-Way Anova with Tukey's Multiple Comparisons Test)

## 2.6 Supporting Information

Supporting information is available at:

<http://iovs.arvojournals.org/article.aspx?articleid=2654114>

## 2.7 Acknowledgements

The authors wish to thank Dr. Connie Cepko for providing *Crx*<sup>-/-</sup> mice, GuangYi Ling and Belinda Dana of the Immunology core for technical assistance in histology and immunohistochemistry assays; and Mingyan Yang for mouse genotyping and other technical assistance. The authors also want to thank Dr. Brandon McKethan of BioRad for designing ddPCR assays to quantify mutant vs. normal transcripts in *L253X*<sup>+</sup> retinas.

## 2.8 References

- Chen, S., Wang, Q.-L., Xu, S., Liu, I., Li, L. Y., Wang, Y., & Zack, D. J. (2002). Functional analysis of cone-rod homeobox (CRX) mutations associated with retinal dystrophy. *Human Molecular Genetics*, *11*, 873–84.
- Freund, C. L., Gregory-Evans, C. Y., Furukawa, T., Papaioannou, M., Looser, J., Ploder, L., Bellingham, J., et al. (1997). Cone-rod dystrophy due to mutations in a novel photoreceptor-specific homeobox gene (CRX) essential for maintenance of the photoreceptor. *Cell*, *91*, 543–53.
- Furukawa, T., Morrow, E. M., & Cepko, C. L. (1997). *Crx*, a novel *otx*-like homeobox gene, shows photoreceptor-specific expression and regulates photoreceptor differentiation. *Cell*, *91*, 531–541.
- Furukawa, T., Morrow, E. M., Li, T., Davis, F. C., & Cepko, C. L. (1999). Retinopathy and attenuated circadian entrainment in *Crx*-deficient mice. *Nature Genetics*, *23*, 466–470.
- Hennig, A. K., Peng, G. H., & Chen, S. (2008). Regulation of photoreceptor gene expression by *Crx*-associated transcription factor network. *Brain Research*, *1192*, 114–133.
- Huang, L., Xiao, X., Li, S., Jia, X., Wang, P., Guo, X., & Zhang, Q. (2012). CRX variants in cone-rod dystrophy and mutation overview. *Biochemical And Biophysical Research Communications*, *426*, 498–503.
- Jacobson, S. G., Cideciyan, A. V., Huang, Y., Hanna, D. B., Freund, C. L., Affatigato, L. M., Carr, R. E., et al. (1998). Retinal degenerations with truncation mutations in the cone-rod homeobox (CRX) gene. *Investigative Ophthalmology And Visual Science*, *39*, 2417–2426.

- Kim, J. W., Yang, H. J., Brooks, M. J., Zelinger, L., Karak?lah, G., Gotoh, N., Boleda, A., et al. (2016). NRL-Regulated Transcriptome Dynamics of Developing Rod Photoreceptors. *Cell Reports*, *17*, 2460–2473.
- Koenekoop, R. K., Loyer, M., Dembinska, O., & Beneish, R. (2002). Visual improvement in Leber congenital amaurosis and the CRX genotype. *Ophthalmic Genet*, *23*, 49–59.
- Lu, Q.-K., Zhao, N., Lv, Y.-S., Gong, W.-K., Wang, H.-Y., Tong, Q.-H., Lai, X.-M., et al. (2015). A novel CRX mutation by whole-exome sequencing in an autosomal dominant cone-rod dystrophy pedigree. *International Journal Of Ophthalmology*, *8*, 1112–1117.
- Nichols, L. L., Alur, R. P., Boobalan, E., Sergeev, Y. V., Caruso, R. C., Stone, E. M., Swaroop, A., et al. (2010). Two novel CRX mutant proteins causing autosomal dominant leber congenital amaurosis interact differently with NRL. *Human Mutation*, *31*, 1472–1483.
- Occelli, L. M., Tran, N. M., Narfström, K., Chen, S., & Petersen-Jones, S. M. (2016). Retinal Cell Biology Cat: A Large Animal Model for CRX-Associated Leber Congenital Amaurosis. *Invest Ophthalmol Vis Sci*, *57*, 3780–3792.
- Rivolta, C., Berson, E. L., & Dryja, T. P. (2001). Dominant Leber congenital amaurosis, cone-rod degeneration, and retinitis pigmentosa caused by mutant versions of the transcription factor CRX. *Human Mutation*, *18*, 488–498.
- Roger, J. E., Hiriyanna, A., Gotoh, N., Hao, H., Cheng, D. F., Ratnapriya, R., Kautzmann, M. A. I., et al. (2014). OTX2 loss causes rod differentiation defect in CRX-associated congenital blindness. *Journal Of Clinical Investigation*, *124*, 631–643.
- Ruzycki, P. A., Tran, N. M., Kefalov, V. J., Kolesnikov, A. V., & Chen, S. (2015). Graded gene expression changes determine phenotype severity in mouse models of CRX-associated retinopathies. *Genome Biology*, *16*, 171.
- Samuel, A., Housset, M., Fant, B., & Lamonerie, T. (2014). Otx2 ChIP-seq reveals unique and redundant functions in the mature mouse retina. *PLoS ONE*, *9*.
- Sohocki, M. M., Sullivan, L. S., Mintz-Hittner, H. A., Birch, D., Heckenlively, J. R., Freund, C. L., McInnes, R. R., et al. (1998). A range of clinical phenotypes associated with mutations in CRX, a photoreceptor transcription-factor gene. *American Journal Of Human Genetics*, *63*, 1307–15.
- Swain, P. K., Chen, S., Wang, Q.-L., Affatigato, L. M., Coats, C. L., Brady, K. D., Fishman, G. A., et al. (1997). Mutations in the Cone-Rod Homeobox Gene Are Associated with the Cone-Rod Dystrophy Photoreceptor Degeneration. *Neuron*, *19*, 1329–1336.
- Swaroop, A., Kim, D., & Forrest, D. (2010). Transcriptional regulation of photoreceptor development and homeostasis in the mammalian retina. *Nature Reviews. Neuroscience*, *11*, 563–576.
- Tran, N. M., & Chen, S. (2014). Mechanisms of blindness: Animal models provide insight into distinct CRX-associated retinopathies. *Developmental Dynamics*, *243*, 1153–1166.
- Tran, N. M., Zhang, A., Zhang, X., Huecker, J. B., Hennig, A. K., & Chen, S. (2014). Mechanistically Distinct Mouse Models for CRX-Associated Retinopathy. *PLoS Genetics*, *10*.
- Won, J., Shi, L. Y., Hicks, W., Wang, J., Hurd, R., Naggert, J. K., Chang, B., et al. (2011). Mouse Model Resources for Vision Research. *Journal Of Ophthalmology*, *2011*, 1–12.

## **Chapter 3**

**Graded gene expression changes determine phenotype severity in mouse models of *CRX*-associated retinopathies**



### 3.1 Author Contributions

This chapter is adapted from the published manuscript: Philip A. Ruzycki<sup>1,3\*</sup>, Nicholas M. Tran<sup>1,3\*\$</sup>, Alexander V. Kolesnikov<sup>1</sup>, Vladimir J. Kefalov<sup>1</sup> and Shiming Chen<sup>1,2,3#</sup> (2015) “Graded gene expression changes determine phenotype severity in mouse models of CRX-associated retinopathies” *Genome Biology*. 16. This work was done in close collaboration with Nicholas Tran, who was a co-first author on the paper as well as Vladimir Kefalov, Alexander Kolesnikov, and Shiming Chen. Shiming Chen and Nick Tran conceived of the original study, while all authors participated in experimental design. Nick performed morphological and light damage experiments, Alexander performed Electroretinograms, and I analyzed all RNAseq data and performed immunohistochemistry experiments. Nick Tran, Shiming Chen and I wrote the manuscript.

---

\*These authors contributed equally.

#Corresponding Author

<sup>1</sup>Department of Ophthalmology and Visual Sciences, Washington University School of Medicine, Saint Louis, Missouri, USA

<sup>2</sup>Department of Developmental Biology, Washington University School of Medicine, Saint Louis, Missouri, USA

<sup>3</sup>Molecular Genetics and Genomics graduate program, Division of Biology & Biomedical Sciences, Washington University School of Medicine, Saint Louis, Missouri, USA

## 3.2 Abstract

Background: Mutations in the cone-rod-homeobox protein CRX are typically associated with dominant blinding retinopathies with variable age of onset and severity. Five well-characterized mouse models carrying different *Crx* mutations show a wide range of disease phenotypes. To determine if the phenotype variability correlates with distinct changes in CRX target gene expression, we perform RNA-seq analyses on three of these models and compare the results with published data.

Results: Despite dramatic phenotypic differences between the three models tested, graded expression changes in shared sets of genes are detected. Phenotype severity correlates with the down-regulation of genes encoding key rod and cone phototransduction proteins. Interestingly, in increasingly severe mouse models, the transcription of many rod-enriched genes decreases decrementally, whereas that of cone-enriched genes increases incrementally. Unlike down-regulated genes, which show a high degree of CRX binding and dynamic epigenetic profiles in normal retinas, the up-regulated cone-enriched genes do not correlate with direct activity of CRX, but instead likely reflect a change in rod cell-fate integrity. Furthermore, these analyses describe the impact of minor gene expression changes on the phenotype, as two mutants showed marginally distinguishable expression patterns but huge phenotypic differences, including distinct mechanisms of retinal degeneration.

Conclusions: Our results implicate a threshold effect of gene expression level on photoreceptor function and survival, highlight the importance of CRX in photoreceptor subtype development and maintenance, and provide a molecular basis for phenotype variability in *CRX*-associated retinopathies.

### 3.3 Introduction

Rod and cone photoreceptors are the two primary light detecting cell types of the retina and are essential for vision. Each cell type preferentially expresses a set of genes critical for development and maintenance of its specialized function. This cell type-specific gene expression is regulated by a network of transcription factors (reviewed in [Hennig et al., 2008; Hsiao et al., 2007]). Disruptions in this regulatory network can have a dramatic effect on rod [J. Chen et al., 2005; Corbo & Cepko, 2005; Daniele et al., 2005; Mears et al., 2001] and cone [Roberts et al., 2005, 2006] development, function, cell-fate integrity and survival. The homeodomain transcription factor CRX is expressed in both rods and cones and plays a central role in mediating photoreceptor transcription. CRX works with rod-specific, cone-specific and general transcription factors to control photoreceptor gene expression [S. Chen et al., 1997; Hennig et al., 2013; Mitton et al., 2000; Onishi et al., 2009; G.-H. Peng et al., 2005; Tran & Chen, 2014]. In particular, CRX and the rod-specific transcription factor NRL cooperatively regulate rod gene transcription and have highly overlapping DNA-binding patterns [Brooks et al., 2011; Corbo et al., 2010; Hao et al., 2012]. Loss of CRX dramatically impairs both rod and cone photoreceptor gene transcription leading to failed photoreceptor maturation and rapid degeneration [Furukawa et al., 1999], while the loss of NRL converts rods into cells with cone-like transcription and functional properties [Brooks et al., 2011; Daniele et al., 2005].

Mutations in the human *CRX* gene have been associated with dominant forms of retinal degenerative diseases such as Retinitis Pigmentosa (RP), Cone-Rod Dystrophy (CoRD) and Leber Congenital Amaurosis (LCA), with varied age of onset and severity (reviewed in [Huang et al., 2012; Rivolta et al., 2001; Tran & Chen, 2014]). Disease-causing mutations fall into at

least four classes, based on mutation type, pathogenic mechanism and molecular properties of the mutant protein: I) hypomorphic substitution mutations with reduced DNA-binding activity; II) antimorphic substitution mutations with variable DNA-binding; III) antimorphic frameshift mutations with intact DNA-binding and IV) antimorphic frameshift mutations with variable DNA-binding (reviewed in [Tran & Chen, 2014]). While each class of disease-causing mutations is associated with distinct clinical phenotypes, the underlying transcriptional changes mediating these phenotypes are poorly understood.

In this study, we used mRNA sequencing (RNA-seq) to examine gene expression in the developing and adult retinas of three *Crx* mutation *Knock-IN* mouse models that have distinct retinal phenotypes [Tran et al., 2014]: *Crx*<sup>R90W</sup> (*R90W*), *Crx*<sup>E168d2</sup> (*E168d2*) and *Crx*<sup>E168d2neo</sup> (*E168d2neo*) (Figure 3.1, Table 3.1). The *R90W* mice carry a Class I mutation and their phenotypes resemble mild late-onset dominant CoRD and recessive LCA. *E168d2* and *E168d2neo* mice carry the same Class III mutation but *E168d2neo* mice express the mutant protein at a lower level, due to the retention of an intronic *neomycin* cassette. *E168d2* mouse phenotypes resemble severe dominant LCA, while *E168d2neo* mouse phenotypes resemble a less severe dominant CoRD phenotype, due to reduced mutant protein expression. We assessed how each *Crx* mutation impacted the expression of retinal cell-type specific genes and investigated cellular pathways that may contribute to disease. We correlated these transcriptional changes with the direct DNA-binding activity of CRX [Corbo et al., 2010] and NRL [Hao et al., 2012] and the epigenetic landscape of rods and cones [Popova et al., 2012; Vierstra et al., 2014]. Additionally, we compared the expression profiles of these *Crx Knock-IN* mice to the previously characterized *Crx*<sup>-/-</sup> and *Crx*<sup>Rip</sup> (*Rip*) mouse models. *Rip* mice carry a Class IV mutation and have a more severe dominant LCA phenotype than *E168d2* mice [Roger et al., 2014]. Our

results demonstrate that all five mouse models have graded changes in photoreceptor-specific gene expression that correlate with the severity of their phenotypes. These graded changes include reduced expression of direct CRX target genes required for rod/cone photoreceptor function and survival, and derepression of ‘cone’ genes in rods through an indirect mechanism, suggesting that rod cell-fate integrity is compromised in the more severe models.

Lastly, we identified threshold effects of gene expression changes on retinal phenotype. Despite having only small differences in gene expression changes, heterozygous *E168d2* and *E168d2neo* mice showed drastically different phenotypes, including light-independent vs. light-dependent photoreceptor degeneration, respectively. These phenotypic and mechanistic differences between the two models are likely attributed to slight shifts in gene expression throughout several photoreceptor-specific pathways, particularly phototransduction and the retinoid (visual) cycle. This highlights the delicate balance between photoreceptor gene transcription, function and cellular integrity. These results demonstrate that the transcriptional landscape in models of retinal degeneration can dramatically affect disease pathology. Effective therapeutic design may therefore be highly context specific.

## 3.4 Results

### 3.4.1 *Crx* mutations cause graded expression changes in shared gene sets, correlating with phenotype severity

To assess the effects of *Crx* mutations on retinal gene expression we performed RNA-seq on retinas of the *Crx* mutant mouse models *R90W*, *E168d2* and *E168d2neo*, and age-matched wild-type (*WT*) controls, in triplicate. As listed in Table 3.2, we analyzed heterozygous mutants at both P10 and P21, but homozygous mutants at P10 only, as their retinas are severely degenerated at later ages. Because they lack WT CRX to antagonize the antimorphic mutant CRX protein, homozygous *E168d2* and *E168d2neo* mice show essentially the same severe phenotype at morphological, functional and gene expression levels [Tran et al., 2014] (Table S1A). Thus, we only performed RNA-seq on *E168d2/d2* to provide a reference for an extremely severe phenotype. RNA-seq libraries were sequenced on the Illumina HiSeq2000 and each generated more than 28 million mapped reads (Table S2). Sample quality was assessed by principal component analysis (PCA) of the expression values of all genes that passed the 5 counts per million (CPM) threshold in any genotype (Supplemental Figure 3.1), and by visual inspection of mapped reads at individual gene loci using Integrated Genomics Viewer (IGV) [Robinson et al., 2011] (Supplemental Figure 3.2A,B). PCA plots (Supplemental Figure 3.1) show that the genotypes clustered as expected with the *E168d2/d2* samples showing the most variation from *WT* controls at both ages tested. At P10 *E168d2/+*, *E168d2neo/+* and *R90W/W* samples clustered between *WT* and *E168d2/d2*, corresponding to their intermediate phenotypes. In contrast, *R90W/+* samples clustered with *WT* samples, consistent with their normal phenotype at this age. At P21, despite increased replicate variability, differences between heterozygous

mutant genotypes and *WT* are consistent with their phenotypes. Since the PCA analysis only described sample differences as a whole, we further determined the relationship of gene expression changes and phenotype differences, using multiple independent analyses of gene expression. First, to determine differential expression, EdgeR [Nikolayeva & Robinson, 2014] was used to compare the triplicate samples. Changes considered “significant” were 2-fold or greater ( $FC \geq 2$ ) with a false discovery rate ( $FDR \leq 0.05$ ) from the appropriate *WT* control, unless otherwise noted. Additional filtering and analysis details can be found in the Appendix 2 Methods section. Expression changes in subsets of RNA-seq identified genes were further validated using qRT-PCR (Table S1 A-C).

### **- Homozygous *Crx* mutant mice show drastically changed expression of a large number of genes at P10 before photoreceptor degeneration**

We first compared retinal gene expression in the homozygous mutants *E168d2/d2* and *R90W/W* at P10 to that in age-matched *WT* controls. As shown by the scatterplots (Figures 3.2A,B), both mutants displayed  $\geq 2$ -fold expression changes for a large number of genes. However, the number of genes affected in *E168d2/d2* was much larger than *R90W/W* (Table 3.2). Among the changed genes, down-regulated genes outnumbered up-regulated genes (Table 3.2), consistent with the established role of CRX in the activation of transcription. Some of the affected genes (white highlighted in Figures 3.2A,B) encode well-characterized proteins essential for rod and cone identity, function and survival. Next, we compared the changed gene sets between the two mutants. Even though *E168d2/d2* and *R90W/W* have mechanistically distinct mutant proteins and phenotypes [Tran et al., 2014], the significantly affected ( $FC \geq 2$  &  $FDR \leq 0.05$ ) genes showed a high degree of overlap for both down-regulated and up-regulated gene sets

(Figures 3.2C,D). In addition, for these shared genes the degree of change was generally greater in *E168d2/d2* than *R90W/W*, as seen by visually inspecting the positions of the highlighted genes in Figure 3.2A vs B. These changes were confirmed to be consistent between replicates by comparing the raw mapped reads for several of these genes (Supplemental Figure 3.2A). Homozygous mouse retina RNA-seq results for several genes were also consistent with previous qRT-PCR data [21] (Table S1B). These data suggest that photoreceptor gene expression in *R90W/W* is less disrupted than in *E168d2/d2*, consistent with the phenotype differences between the two models.

**- Heterozygous *Crx* mutant mice have moderate expression changes in fewer genes than the homozygous mutants at P10, with few independently affected genes**

To determine if heterozygous mutants also share gene expression changes between genotypes, we compared RNA-seq data from heterozygous mutants *E168d2/+*, *R90W/+*, as well as *E168d2neo/+*. At P10, the heterozygous mutants showed fewer gene changes than the respective homozygous counterparts (Table 3.2). Similar to the homozygous mutants, the number of significantly affected genes in heterozygotes correlated with phenotype severity in the order *E168d2/+* > *E168d2neo/+* > *R90W/+*. Figures 3.3A-C and Supplemental Figure 3.2A show that, qualitatively, the white highlighted photoreceptor transcripts were most severely affected in the *E168d2/+* mutant, less affected in the *E168d2neo/+* and showed no change greater than 2-fold in the *R90W/+* line. Furthermore, affected transcripts in heterozygotes also showed a high degree of overlap between the genotypes (Figures 3.3D,E).

Next, we analyzed whether the degree of expression changes correlates with phenotype severity. To gain quantitative results for the entire datasets, we compared the overall fold change



(relative to *WT*) for the union of all affected genes (see Table 3.2 for total numbers) in *E168d2/+*, *E168d2neo/+* and *R90W/+* mutants (Figures 3.3F,G). Both down-regulated (Figure 3.3F) and up-regulated (Figure 3.3G) groups showed a stepwise pattern of expression fold changes: *E168d2/+* > *E168d2neo/+* > *R90W/+*. Thus, in addition to the number of affected genes, the degree of expression changes also correlates with phenotype severity in heterozygous mutants. While this difference in expression was statistically significant (Figures 3.3F,G), the differences between the median log<sub>2</sub> fold change for individual genes expressed in the different heterozygotes were rather small. This is in contrast to the large phenotype differences between these genotypes (see Table 3.1), demonstrating the importance of precisely regulated gene expression for photoreceptor integrity. The expression changes of several genes in heterozygous mouse retinas as determined by RNA-seq results were reproducible (Supplemental Figure 3.2A) and consistent with previous qRT-PCR data [21] (Table S1B,C).

We also directly compared the data for overlapping and independently affected genes in age-matched (P10) heterozygote and homozygote *Crx* mutant animals. Significantly fewer down-regulated and up-regulated genes were seen in the heterozygous *E168d2/+* and *R90W/+* mutants than in their homozygous counterparts (Supplemental Figure 3.3). There was only a small fraction of genes that was independently changed in either heterozygote mutant but unaffected in the homozygotes (Supplemental Figure 3.3). These data are consistent with the differences in phenotype severity between the heterozygous and homozygous mutants.

### 3.4.2 *Crx* mutations specifically affect rod- and cone-enriched genes including down-regulation of phototransduction genes

To ensure that decreased retinal function and impaired development of the *Crx* mutant mice are not a result of perturbations to other non-photoreceptor cell types, we analyzed the P10 RNA-seq data for changes in expression of genes representing different types of retinal neurons. For each cell type, we chose the ten genes with the highest specificity ratio [Siegert et al., 2012] present in our data. The heatmap in Figure 3.4A represents the fold change from *WT* for each of these genes in the P10 mutants, organized by the cell type they represent. The cell types that showed the greatest gene expression changes in all mutants were rod and cone photoreceptors, whereas large-scale changes in gene expression were not observed in other retinal cell types at the age of P10.

We also specifically analyzed expression of genes important for retinal visual function, as changes in this pathway in cell types other than photoreceptors could account for the decreased function as measured by ERG [Tran et al., 2014]. Supplemental Figure 3.4 shows that non-photoreceptor specific genes encoding proteins involved in visual function had little to no change in expression in any of the mutants. In contrast, expression changes were seen in many known photoreceptor-enriched components of the phototransduction cascade (Figure 3.4B). Nearly every constituent was negatively impacted in *E168d2/+*, but remained virtually unaffected in *R90W/+*. *E168d2neo/+* adopted an intermediate expression level, again emphasizing the stepwise and graded changes between the three genotypes.

### 3.4.3 Insights from *E168d2/+* and *E168d2neo/+*: small changes in gene expression level strongly influence phenotype

Since the above analysis detected little difference between the two phenotypically distinct mouse lines that both carry the same *Crx* mutation, *E168d2/+* and *E168d2neo/+*, we expanded our analyses to directly compare these two datasets. A scatterplot comparing *E168d2/+* and *E168d2neo/+* shows only subtle expression differences (Figure 3.4C), further confirming the analyses presented in Table 3.2 and Figure 3.3. Differences in gene expression levels were within the 2-fold margin with only a few exceptions. To decipher the genes that are likely to impact phenotype severity, we analyzed the 118 genes that passed the statistical threshold in the direct comparison of *E168d2/+* vs. *E168d2neo/+* ( $FDR \leq 0.05$ , no FC cut-off, represented in white in Figure 3.4C). To confirm the validity of the results of this comparison, we also calculated the Z-Score of each biological replicate from the mean expression level (Supplemental Figure 3.5), which showed reproducible differences between the two genotypes. The list of down-regulated genes in *E168d2/+* was highly enriched for those relevant to photoreceptor biology and function by Gene Ontology (GO) analyses (Figure 3.4D). In contrast, the up-regulated genes showed only a modest enrichment for a single photoreceptor-relevant GO category (Figure 3.4E). These findings suggest threshold effects of expression level changes on photoreceptor phenotype, especially in those pathways represented in the down-regulated gene set illustrated in Figure 3.4D.

### 3.4.4 Heterozygous mutants show normal gene expression trends from P10 to P21, but many genes fail to reach the normal level at P21

*Crx* mutant mice show early deficits in photoreceptor morphology and function [Furukawa et al., 1999; Tran et al., 2014]. In heterozygous *Crx knockout and R90W* mice, morphology and function recover at later ages, suggesting a transient developmental delay. *E168d2neo/+* mice also recover rod morphology but have abnormal cone morphology and only partially recover rod and cone function in adulthood. *E168d2/+* mice remain impaired in adulthood with shortened outer segments and severely impaired retinal function [Tran et al., 2014], suggesting a blockade in photoreceptor maturation. To determine if differences in morphological and functional recovery are related to gene expression changes over time, we also performed RNA-seq analyses on heterozygous *R90W*, *E168d2* and *E168d2neo* mutants at P21 when photoreceptors are mature. The P21 gene expression data described a very similar scenario seen in the P10 datasets: The three heterozygous genotypes displayed a graded degree and number of genes affected in the mutants relative to the P21 *WT* control (Table 3.1, Supplemental Figures 3.6A-E).

Next, we compared expression changes in the heterozygous mutants between P10 and P21 relative to P10 *WT* expression levels. In *WT* mice, 678 genes showed expression changes between P10 and P21 ( $FC \geq \pm 2$ ,  $FDR \leq 0.5$ ), as shown in a heatmap (Supplemental Figure 3.7A). These age-dependent gene expression changes likely reflect retina terminal differentiation: Gene Ontology analyses showed that the top 100 up-regulated genes were enriched for those important for developing mature photoreceptor structure and function, including visual perception, detection of light stimulus, photoreceptor outer segment formation and monovalent inorganic cation transport (Supplemental Figure 3.7B). In contrast, the down-regulated gene set was

composed of genes that are important for neurogenesis during development, including system development, cell adhesion and regulation of cell proliferation (Supplemental Figure 3.7B). We next analyzed age-dependent gene expression changes in the *Crx* mutant lines, primarily focusing on *E168d2/+* and *E168d2neo/+* models, because of their distinct phenotype from *WT* mice. Unexpectedly, both mutants showed a trend of changes similar to *WT* mice for P21 up-regulated and down-regulated gene sets (Supplemental Figures 3.7A,C). However, in mutants, expression of many members of these two gene sets did not reach the *WT* level at P21 (Supplemental Figures 3.7A,C). As expected, this defect was more prominent in *E168d2/+* than *E168d2neo/+* mice. This suggests that the *E168d2/+* mutants continue to develop after P10 and recover some gene expression, but the degree of recovery is insufficient to achieve normal photoreceptor maturation. Interestingly, *R90W/+* also displayed slightly altered expression levels of many dynamically changed genes. These subtle differences may eventually contribute to the minor functional deficits observed for *R90W/+* at 6 months of age [Tran et al., 2014]. Taken together, these data suggest that *Crx* mutant mice fail both in repressing developmental genes and in activating genes required for photoreceptor maturation. These age-related gene expression changes support the morphological and functional observations that *Crx* mutant mice vary in their photoreceptor maturation rates.

### **3.4.5 Down-regulated and up-regulated genes in *Crx* mutants show distinct epigenetic profiles in *WT* retinas**

To determine the modality of CRX's regulation of differentially expressed genes, we further investigated their expression patterns and the epigenetic landscape of their proximal *cis*-regulatory regions in *WT* mice. We first used hierarchical cluster analysis on P10 and P21

datasets to find sets of genes that were similarly affected in all mutants. Figure 3.5A shows a heatmap representing  $\log_2$  fold change relative to age matched *WT* samples for any gene that displayed significant change from *WT* ( $FC \geq \pm 2$ ,  $FDR \leq 0.05$ ) in any single genotype; the data is arranged by hierarchical clustering (clustering branches are shown to the left of the heatmap). By visual inspection of clustered data, we further subdivided affected genes into eight groups based on similarity of altered expression patterns. These are designated as Groups 1-8 (shown to the right of the heatmap) (see Tables S3 and S4 for lists and order of genes). Further analyses focused on Groups 1, 2, 3 and 6, as these represented the largest and most consistent clusters. Visual inspection of biological replicate data also confirmed the consistency of the expression changes in these groups (Supplemental Figure 3.8). Group 1 genes were the most down-regulated genes across all genotypes. Group 2 genes were decreased compared to *WT* levels, but across the board were less affected than those in Group 1. Groups 3 and 6 were composed of genes that were up-regulated in many of the genotypes. Group 6 genes generally were up-regulated to a greater extent.

### **- The proximal cis-elements of down-regulated genes are enriched for CRX and NRL binding**

To determine which of the above four major gene groups are directly regulated by CRX and its interacting rod-specific transcription factor NRL, we analyzed previously published CRX and NRL ChIP-seq data [Corbo et al., 2010; Hao et al., 2012] obtained from adult mouse retinas for all of the genes in these four clusters. Since CRX is expressed by both rods and cones, CRX ChIP-seq data from both *WT* (predominantly rod) and *Nrl*<sup>-/-</sup> (predominantly cone) retinas were included in our analyses. The *WT* NRL ChIP-seq data represents rod data, since NRL is not

expressed in cones. The data are presented in heatmaps (Figure 3.5B, Columns 1-3), where each gene within the group is represented by a single line. The data within each column reports the average read depth of the indicated experiment +/- 1kb in 40bp bins centered on the transcription start site (TSS). These results presented interesting contrasts between the various groups of genes: First, Groups 1 and 2 displayed a significant amount of CRX binding in rods (Column 1) and cones (Column 2) and NRL binding in rods (Column 3) around the TSS, while this binding was virtually absent in Groups 3 and 6 (Figure 3.5B, Columns 1-3; Supplemental Figures 3.9A-C). This suggests that the down-regulated genes in Groups 1 and 2 are enriched for direct CRX/NRL targets, but not the up-regulated genes in Groups 3 and 6, consistent with the primary role of CRX and NRL in transactivation. Second, the more extensively down-regulated genes in Group 1 showed more CRX (Columns 1-2) and NRL (Column 3) binding than the less-severely down-regulated genes in Group 2 (Figure 3.5B, Columns 1-3; Supplemental Figures 3.9A-C), raising the possibility that Group 1 genes are more dependent on CRX/NRL transactivation activity than Group 2 genes. Third, comparison of CRX binding in rods (Column 1) vs. cones (Column 2) for Group 1 genes showed a high degree of CRX binding in both photoreceptor subtypes, suggesting that this group contains genes that are activated by CRX in both rods and cones. In contrast, Group 2 genes showed CRX binding largely in rods, suggesting these genes are activated by CRX mainly in rods (Supplemental Figures 3.9A,B).

#### **- Down-regulated genes become more “open” during postnatal retinal development**

To determine which of the four major groups undergo epigenetic landscape changes during photoreceptor development, we analyzed three different sets of previously published and

publicly available epigenetic signature data (Figure 3.5B, columns 4-10). Results are presented as heatmaps in a similar format as CRX and NRL ChIP-seq data.

We first analyzed retinal DNase I hypersensitivity (DHS) datasets at the age of P1, P7 and 8 weeks (8wk) (Figure 3.5B, Columns 4-6) from the ENCODE project [22]. Enriched DHS is an indicator of ‘open’ chromatin and is a predictor of active transcription. Both down-regulated and up-regulated genes showed dynamic changes in their epigenetic landscapes with age. First, Group 1 genes showed increases in DHS signal from P1 to 8wk of age. This pattern is not seen in Group 2 genes, in which high DHS signal was largely stable across the ages tested. Group 3 and 6 genes, on the other hand, showed different patterns, in which DHS signal decreased from P1 to 8wk of age. Second, comparing across groups at 8wk of age, DHS data described Groups 1 and 2 as having more ‘open’ chromatin (higher DHS signal) around their TSS than Groups 3 and 6 (Figure 3.5B, Column 6). These data suggest that the chromatin of genes down-regulated in *Crx* mutants (Groups 1 or 2) have ‘open’ chromatin in the adult *wild-type* retina and the mostly strongly down-regulated genes (Group 1) are genes that develop a more ‘open’ chromatin conformation postnatally. In contrast, up-regulated genes (Groups 3, 6) tended to have ‘open’ chromatin in the early postnatal retina that became more ‘closed’ with age. To determine if these dynamic DHS changes represented specific events for particular groups, we included a random control group in our analyses. A random set of genes was chosen from the UCSC gene list to match the size of the largest Group from the analysis (Group “Random”). As illustrated by Supplemental Figures 3.9D-F, the Random group (grey lines) also showed a trend of reduction in DHS during postnatal retinal development. Importantly, when compared to Random genes, Group 1 genes were less ‘open’ at P1 and P7 (Supplemental Figures 3.9D,E, blue line vs. grey line) but more ‘open’ at 8 weeks (Supplemental Figure 3.9F, blue line vs. grey line),



verifying the trend visible by eye in the Figure 3.5B heatmaps. These results were consistent with the observation that Group 1 genes largely increased their expression from P2 to P21, during *WT* retinal development (Supplemental Figure 3.9K, based on published RNA-seq data by [Roger et al., 2014]). Interestingly, Group 2 genes were more ‘open’ than Random genes at all ages tested (Supplemental Figures 3.9D,F, dashed blue line vs. grey line), consistent with an overall modest increase in their expression from P2 to P21 in *WT* mice (Supplemental Figure 3.9L). In contrast, Group 6 genes showed overall lower DHS than Random genes at all three ages analyzed, while Group 3 genes showed similar DHS patterns as Random. Both Group 3 and 6 genes normally showed no change or a slight decrease in expression from P2 to P21 (Supplemental Figures 3.9M,N). This data suggests the DHS changes identified in Group 1 genes are dynamically regulated during retina development.

### **- Down-regulated genes undergo histone modification changes during postnatal retinal development**

Next, we analyzed these gene groups for the presence of active (H3K4me2) (Columns 7,8) and repressive (H3K27me3) (Columns 9,10) histone marks as determined by ChIP-Seq in P1 and P15 *WT* retinas [23] within the 2 kb window centered on the TSS of each gene. Consistent with gaining a more ‘open’ chromatin configuration, Group 1 genes gained the active mark H3K4me2 and lost the repressive mark H3K27me3 between P1 and P15 (Figure 3.5B, Columns 7-10; Supplemental Figures 3.9G-J, blue line vs. grey line). Group 2 genes, showed similar changes in histone marks (increase in H3K4me2 and decrease in H3K27me3) from P1 to P15. Group 2 genes also showed a higher level of the active mark H3K4me2 than Random genes at P1, while the level of the repressive mark H3K27me3 was similar to Random at this age

(Figure 3.5B, Columns 7-10; Supplemental Figure 3.9G-J, dashed blue line vs. grey line).

Overall, the data is consistent with a postnatal constitutively open chromatin configuration for Group 2 genes as measured by DHS. Finally, Groups 3 and 6 showed no difference relative to the Random group in either H3K4me2 or H3K27me3 occupancy through retina development (Figure 3.5B, Columns 7-10; Supplemental Figures 3.9G-J, red lines vs. grey lines), consistent with their overall constant low level of gene expression.

Together, the above analyses suggest that Group 1 genes undergo dynamic epigenetic changes during retina maturation associated with their substantial transcriptional activation during postnatal photoreceptor development. The expression pattern of these genes correlated with CRX and NRL expression, and their regulatory regions were directly bound by CRX and NRL in *WT* mice. These results suggest that Group 1 is enriched for genes that are inactive in precursor cells and actively turned on by key photoreceptor transcription factors during development. Thus, mutations in CRX have a profound negative impact on the expression of these genes. Group 2 genes were similar to Group 1 in that their epigenetic landscape is supportive of high transcriptional levels in the *WT* retina and are likely regulated directly by CRX and NRL. Interestingly, these genes do not show the time dependent chromatin ‘opening’, suggesting that CRX and NRL may not be necessary to initiate chromatin remodeling and could account for the fact that Group 2 genes lose expression to a lesser degree than Group 1 genes. Finally, Group 3 and 6 data suggest that these genes undergo chromatin remodeling in the *WT* retina that results in a less permissive state. This is consistent with their low or even decreasing expression over time. However, this process is not directly controlled by CRX or NRL, as ChIP-seq binding of these two proteins was not observed in these Groups.

### 3.4.6 Up-regulated genes are characteristic of cone photoreceptors, likely resulting from de-repression in rods

Since CRX and NRL binding as well as epigenetic data implicate down-regulated genes as being active CRX targets in rods, we investigated if differentially-expressed genes normally have rod or cone cell-type-specific expression patterns. Using published data [Roger et al., 2014] to classify these genes as either rod- or cone-enriched (or non-specific – N.S.), we found a significant overrepresentation of rod genes in Groups 1 and 2, although a number of cone genes were also found in Group 1 (Figure 3.6A and Table S3). This was expected considering the mouse retina is rod-dominant and CRX acts as a transcription activator for genes critical for photoreceptor structure and function. However, surprisingly, up-regulated Groups 3 and 6 contained a significant enrichment of cone transcripts (Figure 3.6A and Table S3).

To determine whether this up-regulation of cone gene expression is a general trend beyond Group 3 and 6 genes, we expanded our analysis to all genes enriched in either rods or cones in adult mice. Furthermore, to ensure the results are applicable to other classes of *Crx* mutants, we also analyzed three previously published RNA-seq datasets from other *Crx* mutant models, *Crx*<sup>-/-</sup>, *Rip*<sup>+/+</sup> and *Rip/Rip*. The results, presented by heatmaps (Figures 3.6B,C), showed a broad switch in the global pattern of gene expression in all mutants, with loss of expression of a large set of rod genes (Figure 3.6B) and increased expression of transcripts normally enriched in cones (Figure 3.6C). There were exceptions: expression of a small number of cone genes, indicated with an asterisk at the bottom of Figure 3.6C, was consistently decreased in all the *Crx* mutants. These represented key cone-specific phototransduction genes including *Pde6h*, *Arr3*, *Opn1mw*, *Opn1sw*, *Gnat2* and others (Supplemental Figure 3.10B). Their loss of expression was similar to rod-specific phototransduction components that also

consistently decreased, including *Rho*, *Gnat1*, *Cngb1*, and *Pde6g* (Figure 3.6B marked by asterisk, Supplemental Figure 3.10A). Again, rod-enriched genes and cone-enriched genes displayed graded changes that reflected phenotype severity in our models (*E168d2/d2*: most severe; *R90W/+*: least severe). These results also place the *Rip* model (*Rip/+* and *Rip/Rip*) in line with other *Crx* models, demonstrating even more severe expression changes than *E168d2/d2*, although the RIP protein is thought to cause disease by a distinct molecular mechanism [Roger et al., 2014].

Because RNA-seq was performed on whole-retina samples, the down-regulation of rod genes and up-regulation of cone genes could represent a decrease in the number of rods and an increase in cones. However, previously published histology showed that in fact the proportions of these cell populations shift in the opposite direction, as the *E168d2/+* retina has a very severe and early depletion of cone photoreceptors, prior to any loss of rod photoreceptors [Tran et al., 2014]. In fact, all *Crx* mutant mouse models including *Crx*<sup>-/-</sup> and *Rip/+* show a similar trend of rapid cone loss followed by slower rod degeneration. This result was also unexpected considering the role of CRX in trans-activating cone as well as rod genes [S. Chen et al., 2004; Tran et al., 2014]. This raised the possibility that the up-regulated cone genes were abnormally derepressed in the mutant rods. To further understand the molecular mechanism underlying mis-expression of cone genes in mutant rods, we assessed all rod- and cone-enriched transcription factors for expression changes in the mutants. Consistent with overall expression changes, we observed a general loss of rod transcription factors and enhanced expression of many cone-enriched transcription factors (Supplemental Figures 3.11A,B). However, the reduced expression of rod-specific factors that are essential for maintaining rod cell fate, such as *Nrl* and *Nr2e3*, was rather minor or even absent in the heterozygous *E168d2* and *R90W* mutants, unlike

that reported for *Rip* mice [Roger et al., 2014]. Instead, these results suggest an unexpected role for CRX in rods to repress cone-enriched genes, including a number of cone-enriched transcription factors. To test whether CRX could play a direct role in repressing these cone-enriched transcription factors, we examined the CRX ChIP-seq data (Supplemental Figures 3.11A-C) for evidence to support a direct interaction of CRX with up-regulated cone transcription factors in rod cells. Some up-regulated cone TFs showed low ChIP signal in *WT* retina, suggesting a potential repressive binding in rods, but the pattern was not consistent (Supplemental Figure 3.11C). Thus, loss of repression of cone genes in mutants is likely a secondary effect. In contrast, ~33% of rod genes were bound by CRX, significantly enriched over random genes (~8%). Furthermore, when compared with NRL ChIP-seq data, >73% of CRX-bound rod genes were also bound by NRL, but no enrichment of NRL binding was observed for cone genes (data not shown), consistent with CRX's interaction with NRL to activate the expression of rod genes, but not cone genes.

To verify that these cone transcripts are being de-repressed in mutant photoreceptors that normally would adopt a rod cell fate, we performed immunohistochemistry for several cone targets on retinal sections from three *E168d2* sublines and *WT* control mice at P10. The nuclear receptor  $RXR\gamma$  is preferentially expressed by cones in P10 *WT* retina, and is up-regulated in most *Crx* mutants by RNA-seq and qRT-PCR [Tran et al., 2014] (Table S1A,B; biological replicate RNAseq raw mapped sequencing reads displayed in Supplemental Figures 3.2A,B). As shown in Figures 3.6D1-4, instead of the cone-expression pattern in normal retina (Figure 3.6D1),  $RXR\gamma$  immunoreactivity was detected in most outer nuclear layer (ONL) cells of the homozygous and heterozygous *E168d2* mutant retinas (Figures 3.6D2,D3), with a perinuclear distribution similar to the pattern reported for rod transcription factors [Roberts et al., 2005].

The ONL RXR $\gamma$  staining appeared stronger and more widely spread in *E168d2/d2* than *E168d2/+* retinas. The enhanced RXR $\gamma$  immunoreactivity was not seen in *E168d2neo/+* ONL where the rods were not disturbed. Immunostaining of the second cone marker CNGB3 displayed similar patterns of enhanced ONL expression in affected mutant rods (Figures 3.6D5-D8). We next assessed peanut agglutinin (PNA) binding (Figures 3.6D9-D12), which normally stains the cone sheath including cone outer segments and pedicles. Figure 3.6D shows that, despite the lack of cones and rod outer segments in *E168d2/d2* retinas, a strong and uniform PNA staining was seen across the outer margin of the entire retina (Figure 3.6D6), suggesting that mutant rods adopt cone-like characteristics of their sheath. *E168d2neo/+* retinas did not show enhanced PNA binding at this age (Figure 3.6D12), while *E168d2/+* showed an intermediate PNA staining that was higher than in *WT* but much lower than in *E168d2/d2* retinas. Taken together, all three selected cone markers showed enhanced expression in mutant rods in the order *E168d2/d2* > *E168d2/+* > *E168d2neo/+*, thus confirming RNA-seq findings of the impact of *Crx* mutations on rod cell fate.

In summary, our RNA-seq analyses identified graded expression changes of shared gene sets in seven available mouse models for *Crx*-associated disease. Using the data presented and referenced in this paper [Hsiau et al., 2007; Roger et al., 2014; Tran et al., 2014], these models can be ranked as illustrated in Figure 3.7, with the lightest bars representing the model with most severely affected rod and cone gene expression. This order correlates with phenotype severity by morphological and electrophysiological standards, and with changes in gene expression in rods and cones. This correlation was seen not only for down-regulated rod and cone genes encoding phototransduction components, but also for those up- and down-regulated gene sets that highlight the partial rod to cone conversion of the developing photoreceptors. The schematic also

predicts the level of photoreceptor identity and function in various models. Most importantly, the different gene expression changes between some phenotypically distinct *Crx* mutant models (such as *E168d2/+* vs. *E168d2neo/+*) were rather modest, in contrast to their substantial impact on the disease phenotype.

### **3.4.7 Testing the effect of gene expression changes on light dependent degeneration**

#### **- *E168d2neo/+* but not *E168d2/+* mice are sensitive to light damage**

To validate the phenotypic significance of the small differences in gene expression between *E168d2/+* and *E168d2neo/+* mice (Figure 3.4C), we analyzed expression of genes involved in the photoreceptor retinoid (visual) cycle that is responsible for the recycling of visual pigment chromophore 11-*cis*-retinal by RNA-seq (Figures 3.8a and Supplemental Figure 3.2A,B) and qRT-PCR (Table S1C). This discrete set of visual cycle genes expressed in photoreceptors also showed a similar pattern of expression changes across the mutant lines (Figure 3.8A): *R90W/+* showed little or no changes relative to *WT*, while *E168d2/+* and *E168d2neo/+* showed reduction in gene expression levels. We validated the expression patterns of three visual cycle genes: *Rbp3*, *Rdh12* and *Abca4* by qRT-PCR and results were consistent with RNA-seq findings (Table S1C). We hypothesized that, as a result of these changes combined with slight shifts in expression levels of phototransduction genes (Figure 3.4B), the two *E168d2* sublines may have different sensitivity to light-induced damage (LD). Exposing mice to high intensity light for an extended period of time puts stress on the retina and can lead to photoreceptor degeneration [Noell et al., 1966]. LD-related degeneration depends on three interconnected pathways (Figure 3.8B): 1) the phototransduction pathway, 2) visual cycle in the

retinal pigment epithelium (RPE) and 3) visual cycle in photoreceptor cells. Most mouse models with disruptions to components of phototransduction are insensitive to LD [Grimm et al., 2000], while mice with an impaired visual cycle in photoreceptors generally have increased sensitivity to LD [A. Maeda et al., 2006, 2007, 2008; T. Maeda et al., 2009; Radu et al., 2003]. To test if the transcriptional defects in *E168d2/+* and *E168d2neo/+* mice alter their sensitivity to LD, we exposed 6 week old (wo) *WT*, *E168d2/+* and *E168d2neo/+* mice to a high intensity light (12-13.5 KLUX; ~10-20x brighter than ambient light conditions) for 8 hrs. Following LD exposure, mice were kept in 12 hr ambient light/dark cycle for 7 days before retinal function and morphology were assessed. All the mice tested were backcrossed to the *C57BL/6J* background for >10 generations, which is resistant to LD [Wenzel et al., 2001]. As expected, *WT* mice were highly resistant to LD and displayed normal retinal morphology (Figure 3.8D vs. 3.8C, left panels), normal ONL thickness as determined by morphometry (Figure 3.8E, left panel), and normal retinal function as measured by ERG (Supplemental Figure 3.12A; black dashed lines vs. black solid lines) compared to *WT* controls that were not exposed to LD. *E168d2/+* mice exposed to LD did not show differences in retinal morphology (Figure 3.8D vs. 3.8C, middle panels) or ONL thickness (Figure 3.8E, middle panel) and showed only minor ERG differences compared to normal light exposed *E168d2/+* controls (Supplemental Figure 3.12A; red dashed lines vs. red solid lines). In contrast, *E168d2neo/+* mice exposed to LD showed shortened outer segments and a loss of ONL nuclei, ~3-4 nuclei compared to ~10-12 nuclei in normal light *E168d2neo/+* mice (Figure 3.8D vs. 3.8C, right panels). Morphometry revealed significant reduction in ONL thickness at the inferior -100  $\mu\text{m}$  and superior 100  $\mu\text{m}$  and 500  $\mu\text{m}$  positions, with the thickness at 500  $\mu\text{m}$  being most affected (~58% reduced) (Figure 3.8E, right panel). Retinal function was also affected with dark-adapted a-waves, dark-adapted b-waves and light-



adapted b-waves all showing reduced maximal response amplitudes compared to *E168d2neo/+* controls kept in normal light (Supplemental Figure 3.12A; blue dashed lines vs. blue solid lines). These data suggest that *E168d2neo/+* mice are more susceptible to LD than either *E168d2/+* or *WT* mice, providing further insight into the distinct pathobiology of *E168d2/+* and *E168d2neo/+* mice.

### **- *E168d2neo/+* sensitivity to light damage is linked to abnormal visual cycle in photoreceptors**

To determine if the increased sensitivity to LD in *E168d2neo/+* mice reflects changes in expression of visual cycle genes, mice were pre-treated with 13-*cis* retinoic acid (13-*cis*-RA; Accutane) before LD exposure. 13-*cis*-RA interferes with the regeneration of visual chromophore (11-*cis*-retinal) and is a strong antagonist of the RPE-driven phase of the visual cycle [Sieving et al., 2001]. Inhibiting the visual cycle at the earliest stage reduces the production of toxic retinoid intermediates in the retinoid pathway during LD (Figure 3.8B). This treatment strategy was previously shown to effectively ameliorate retinal degeneration in mouse models with impaired visual cycle [T. Maeda et al., 2009; Radu et al., 2003; Sieving et al., 2001]. Retinas of *WT* mice pretreated with 13-*cis*-RA and exposed to LD were morphologically indistinguishable from DMSO-injected controls (Figure 3.8F, two left panels). In contrast, retinas of *E168d2neo/+* animals pretreated with 13-*cis*-RA were almost completely protected from the rapid degeneration observed in DMSO-injected *E168d2neo/+* controls (Figure 3.8F, two right panels). Retinas of 13-*cis* RA-pretreated *E168d2neo/+* mice revealed significant improvements in ONL thickness at the superior 500  $\mu\text{m}$  and 1000  $\mu\text{m}$  positions (Figure 3.8G, right panel) compared to those DMSO treated. Retinal function was also better preserved in 13-

*cis*-RA pretreated mice. Amplitudes of dark-adapted ERG a-wave, dark-adapted b-wave and light-adapted b-wave were all significantly restored in 13-*cis*-RA pretreated *E168d2neo/+* mice, compared to DMSO injected controls (Supplemental Figure 3.12B; green lines vs. blue dashed lines). To validate that LD sensitivity is linked to intrinsic defects in the visual (retinoid) pathway, we tested rod dark adaptation in *WT* and *E168d2neo/+* mice. Following >90% rhodopsin photobleach, the dark adaptation of *E168d2neo/+* mouse rods was substantially delayed, as compared to that in *WT* mice (Supplemental Figure 3.12C, blue line vs. black line). Pretreatment of *E168d2neo/+* mice with 13-*cis*-RA strongly delayed rod dark adaptation further (Supplemental Figure 3.12C, red line vs. blue line), indicating that 13-*cis*-RA effectively blocked regeneration of visual pigment by the RPE visual cycle. Together, these data suggest that the increased sensitivity of *E168d2neo/+* mouse retinas to LD is linked to intrinsic defects in the photoreceptor visual cycle.

Finally, to determine if the degeneration of photoreceptors in *E168d2/+* mice is at all affected by ambient light conditions, we reared *E168d2/+* mice under either normal 12 hr light/dark cycle or in constant darkness. Retinal morphology of *E168d2/+* mice raised under either light condition was indistinguishable at 3 and 6 months of age (Supplemental Figure 3.13), suggesting that photoreceptor degeneration in *E168d2/+* mice is independent of light conditions. These results illustrate that photoreceptor degeneration in *E168d2/+* and *E168d2neo/+* mice is mediated by discrete light-independent and light-dependent mechanisms, respectively.

## 3.5 Discussion

### 3.5.1 Gene expression in animal models for *CRX*-associated retinopathies

Our RNA-seq results for mutant *Crx E168d2* and *R90W* mouse retinas improve and expand upon previously published genomic expression microarray studies and qRT-PCR on these mouse lines [Tran et al., 2014]. The increased sensitivity of this platform allowed for the detection of more differentially expressed genes, especially for genes with low expression level, and also for the detailed analysis of the modest expression level differences between heterozygous animals. In combination with RNA-seq data from *Crx*<sup>-/-</sup> and *Rip* mice [Roger et al., 2014], a comprehensive dataset of retinal gene expression now exists for *Crx* mutant models with a range of phenotypes that reflect human disease. We have used these datasets to identify key changes in gene expression that correlate with disease severity.

Our results suggest that even relatively small changes in gene expression level can have a profound effect on the delicate balance of cellular pathways critical for photoreceptor function and survival, especially phototransduction and the visual cycle. Our RNA-seq analysis of retinal gene expression in mutant *Crx* mouse models by several methods indicates graded changes in photoreceptor gene expression at every level from overall expression patterns down to specific cellular pathways. First, homozygous mice were always more severely affected than their heterozygous counterparts (Supplemental Figure 3.3). Among homozygous and heterozygous mice, expression changes from least to most severe consistently followed the pattern: *R90W*, *E168d2neo*, *E168d2*. These patterns were observed for both down-regulated and up-regulated genes. A high degree of overlap existed in the differentially expressed genes among models (Figures 3.2, 3.3, and Supplemental Figure 3.6) and few genes displayed opposite expression

patterns between models (Figure 3.5). This suggests that few unique gene targets exist between models, and non-overlapping genes largely result from genes with similar trends in expression that simply do not pass the significance threshold of the *in silico* analysis. Graded changes were also observed for genes involved in retinal development; *Crx* mutants showed impairment in both the activation and repression of genes that shift expression patterns from P10-P21 (Supplemental Figure 3.7). Despite having collected mRNA from the whole retina, we were able to determine that the differentially expressed genes mostly occurred in rod and cone photoreceptors by using known patterns of cell-type specific retinal gene expression (Figure 3.4A). This was consistent with CRX's expression pattern and role in photoreceptor maturation. Within photoreceptor-specific pathways involved in light response, namely rod and cone phototransduction and the visual cycle, all *Crx* mutant models displayed graded down-regulation (Figure 3.4B, Figure 3.8A). Aside from these pathways, down-regulated genes tended to be rod genes, while up-regulated genes tended to be cone genes de-repressed in rods (Figure 3.6, Supplemental Figure 3.11). We investigated genes that did not fall into either category (Table S3) but did not find any other major trends beyond this. Expansion of this analysis by assessing expression of all known rod or cone-enriched genes again showed a graded pattern of expression changes in these genes (Figures 3.6B,C), including rod and cone-specific transcription factors (Supplemental Figures 3.11A,B). Re-analysis of previously published microarray data [Tran et al., 2014] showed the same trends of gene expression changes (data not shown). Finally, by analyzing RNA-seq results from two previously published animal models, *Crx*<sup>-/-</sup> and *Rip*, we were similarly able to detect graded expression changes in a common gene set (Figure 3.6). In all, our analyses show graded changes in gene expression at several levels in *Crx* mutant mice for both down-regulated and up-regulated genes.

### 3.5.2 Different modalities for down-regulated and up-regulated gene expression

While *Crx* mutant mice showed both down-regulated and up-regulated expression (Table 3.2), more genes were down-regulated in every mouse model tested. This is consistent with CRX's established role as a transcriptional activator, though CRX does act as a repressor in certain contexts [White et al., 2013]. Utilizing available CRX [Corbo et al., 2010] and NRL ChIP-seq [Hao et al., 2012] data sets and epigenetic data sets including DnaseI hypersensitivity from the ENCODE project [Vierstra et al., 2014] and active (H3K4me2) and repressive (H3K27me3) histone marks [Popova et al., 2012], we identified patterns in the proximal *cis*-regulatory elements of differentially expressed genes. CRX and NRL binding were enriched at genes that were down-regulated in *Crx* mutant models, suggesting these genes are direct targets, while genes that were up-regulated were not enriched for CRX or NRL binding, suggesting these genes are indirect targets (Figure 3.5). DNase hypersensitivity data showed that a group of strongly down-regulated genes in mutant models shifted from 'closed' to 'open' chromatin from P1 to 8 wk in *WT* rods, suggesting these genes are normally activated during postnatal retinal development (Figure 3.5, Supplemental Figures 3.9D-F). In contrast, the data described that a group of up-regulated genes became more 'closed', suggesting these genes are being repressed during postnatal development (Figure 3.5B, Supplemental Figures 3.9D-F). The group of down-regulated genes also showed an increase in the active histone mark H3K4me2 and a decrease in H3K27me3 during postnatal development in *WT* retina (Figure 3.5B, Supplemental Figures 3.9G-J), indicative of active chromatin remodeling. Changes in these histone marks were not observed for up-regulated genes, suggesting these genes are not subject to this type of active chromatin remodeling. Together, these results suggest that genes down-regulated in *Crx* mutant

models are direct targets of CRX and NRL that normally are activated during postnatal development through chromatin remodeling, while up-regulated genes are indirect targets of CRX and NRL that are normally repressed during postnatal development through an unknown mechanism. CRX is known to recruit co-activators with positive chromatin remodeling capacity, such as CBP/p300 and the STAGA complex [Palhan et al., 2005; G. Peng & Chen, 2007]. Thus, CRX and its interacting co-activators may play an essential role in active chromatin remodeling during development for the direct target genes represented by the down-regulated groups in *Crx* mutants. Future profiling of epigenetic changes in these *Crx* mutant models will provide additional support for this possibility.

### **3.5.3 *Crx* mutations affect rod and cone development and transcriptional integrity**

The graded expression changes in *Crx* mutant models have profound effects on the development of their rods and cones as illustrated by Figure 3.7. Homozygous *Crx*<sup>-/-</sup>, *E168d2*, *R90W*, *Rip* and *Rip*<sup>+/+</sup> mice do not form physiologically functional photoreceptors, due to a strong reduction in both rod and cone phototransduction gene expression. Other models show gradations in both retinal function and phototransduction gene expression. In all models studied, the development and survival of cones are affected earlier and more severely than rods, implicating CRX in the terminal differentiation of cones. Interestingly, in *WT* rods, CRX and NRL appear to mediate both the activation of rod-specific genes and the repression of cone-specific genes. The dual reduction in rod-gene expression and de-repression of cone genes in *Crx* mutant rods coincides with rods adapting more cone-like properties including less condensed chromatin and shorter outer segments [Roger et al., 2014; Tran et al., 2014]. In *Rip*<sup>+/+</sup> mice, this

shift was found to coincide with loss of NRL expression by P21, indicating a complete loss of rod cellular identity. The loss of NRL expression does not occur in *E168d2/+* mice by this age, but de-repression of cone-genes in rods is still observed. While E168d2 protein does not affect the function of NRL *in vitro* [Tran et al., 2014], its effect on NRL *in vivo* is uncharacterized. Our results provide evidence that even though RIP and E168d2 mutant proteins act through different pathological mechanisms, the resulting pathologies arise from scaled changes in similar sets of genes.

### 3.5.4 Gene expression dictates phenotypic thresholds

The animal models used in this study have wide-ranging phenotypes that match those observed in human patients. The magnitude of overall gene expression changes correlated with phenotype severity in every model tested, but not in a linear manner. Instead, there are thresholds in gene expression that determine phenotypic presentation. An example of this threshold effect is evident in the *E168d2/+* and *E168d2neo/+* mice. These mice have drastically different phenotypes but have only small differences in gene expression. As we have previously shown, the *neomycin* cassette retained in *E168d2neo/+* mice suppresses an accumulation of the mutant transcript by an unknown mechanism [Tran et al., 2014]. qRT-PCR results (Table S1A) of the same biological samples used for RNA-seq confirm that the *WT* allele showed no compensation and maintained approximately 50% of its normal expression. In contrast, the results of RNA-seq (Supplemental Figure 3.14) and qRT-PCR (Tables S1A,B), combined with previously published Western blots and immunohistochemistry [Tran et al., 2014], all support a significant overexpression of mutant *Crx* mRNA and protein in *E168d2/+* mice. This mutant allele-specific overexpression was not evident in *E168d2neo/+* mutants. This difference in

expression of the dominant negative form of CRX resulted in very small difference in fold change of down-regulated genes (median: -1.8 in *E168d2/+* compared to -1.7 in *E168d2neo/+* mice; Figures 3.3F,G). Despite this slight difference, *E168d2/+* mice had much more severe deficits in retinal function and photoreceptor degeneration. Under normal light conditions, *E168d2neo/+* mice had little photoreceptor degeneration in early adulthood. However, *E168d2neo/+* rods were highly sensitive to light damage while *WT* and *E168d2/+* rods were largely resistant (Figure 3.8). Targeting the visual cycle blocked sensitivity of *E168d2neo/+* photoreceptors to LD, implicating this pathway in the degeneration phenotype. In contrast, the degeneration of *E168d2/+* photoreceptors was independent of light and the visual cycle. Sensitivity to LD requires phototransduction and is mediated by the visual cycle, both of which are affected in the *E168d2* and *E168d2neo* mouse models. However, these results suggest that the balance in function between these two pathways could be critical for determining sensitivity to damaging light. Our data support the conclusion that phototransduction is too impaired and OS structural changes are already too severe in *E168d2/+* mice to allow for further light-mediated degeneration, while *E168d2neo/+* photoreceptors largely preserve their structure and substantial levels of phototransduction components to reveal visual cycle defects under intense light. These results highlight that even minor tuning differences in photoreceptor gene expression have a dramatic effect on the mechanisms of disease pathology. These findings could have significant clinical importance as patients with *CRX*-associated retinopathies might have different responses to environmental factors like damaging light levels, which would affect their clinical outcome. These results also provide clues for potential therapeutic intervention to target the visual cycle in patients with late-onset *CRX*-associated retinopathy.



### 3.5.5 Conclusions

Linking genotype to phenotype for *CRX* retinopathies remains imprecise [45]. Genomic assessment of retinal gene expression suggests that the range of phenotypes of *Crx* mutant mouse models are driven by graded changes in photoreceptor gene expression. Since *CRX* has such wide-ranging function in photoreceptor transcription, mutations that sometimes just slightly alter the expression of many genes can have a profound effect on the resultant phenotype. We relate retinal gene expression to phenotypic thresholds in rod and cone photoreceptor development, cellular integrity, function and degeneration in several *Crx* mutant mouse models. We have utilized publicly available genomic datasets to gain insight into cell-type specific expression changes and the different modalities for gene down-regulation and up-regulation in these mice. We have demonstrated how slight differences in gene expression can alter retinal susceptibility to light-dependent degeneration in *E168d2/+* and *E168d2neo/+* mice. These findings provide evidence that *CRX*-retinopathies stem from graded changes in photoreceptor gene expression, which could significantly contribute to phenotypic variability.

**Table 3.1: Phenotype summary of heterozygous *Crx* mutant mice**

Mouse (1)	Mutation class (2)	CRX expression (3)	Rod		Cone		Disease model	Phenotype severity (5)
			Function (4)	Degeneration	Function (4)	Degeneration		
<i>WT</i>	n/a	+	++++	Undetectable	++++	Undetectable	n/a	n/a
<i>R90W</i>	I	+	++++	Undetectable	++++	Undetectable	CoRD	mild
<i>E168d2neo</i>	III	+	+++	Undetectable	++	≥1yr	CoRD	moderate
<i>E168d2</i>	III	++	++	1-6mo	+	1mo	LCA	severe
<i>Rip</i>	IV	+	-	1-18mo	-	Undetectable	LCA	very severe

- (1) Heterozygous mice harboring the indicated *Crx* mutations [Roger et al., 2014; Tran et al., 2014] are used for phenotype comparisons. "*WT*": *C57BL/6J* wild-type control; "n/a": not applicable.
- (2) Classification described by [Tran & Chen, 2014].
- (3) Grading based on quantitative Western blots [Tran et al., 2014]. Note a 2-fold increase in *E168d2* but normal level in others.
- (4) Grading based on reduction of ERG peak amplitudes [Tran et al., 2014]. "-": undetected ERG signals.
- (5) Severity based on combined morphological and functional deficits. "n/a": not applicable

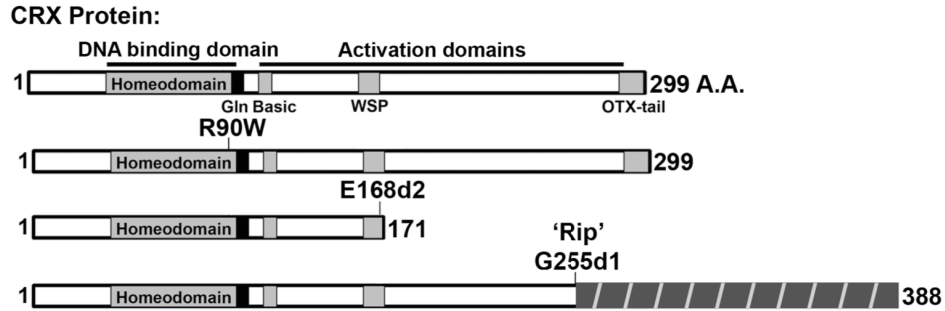
**Table 3.2: List of RNA-seq experiments and number of genes with altered expression**

Mouse (1)	Experiments & replicates (2)		Down-regulated genes (3)		Up-regulated genes (3)	
	P10	P21	P10	P21	P10	P21
<i>WT</i>	3	3	-	-	-	-
<i>E168d2/d2</i>	3	Not tested	425	-	248	-
<i>R90W/W</i>	3	Not tested	195	-	70	-
<i>E168d2/+</i>	3	3	136	150	61	43
<i>E168d2neo/+</i>	3	3	85	83	38	19
<i>R90W/+</i>	3	3	20	27	12	27

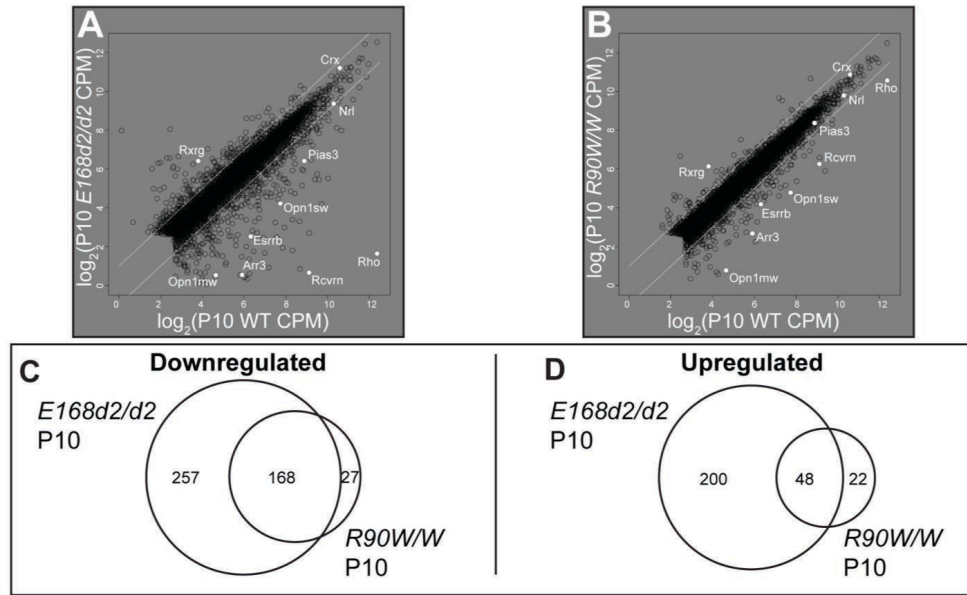
(1) All mutant mice were backcrossed to *C57BL/6J* (*WT*) control for >10 generations and genotyped for common variants [Tran et al., 2014].

(2) Numbers represent biological replicates. Each replicate contains 4 pooled retinas from a pair of male and female

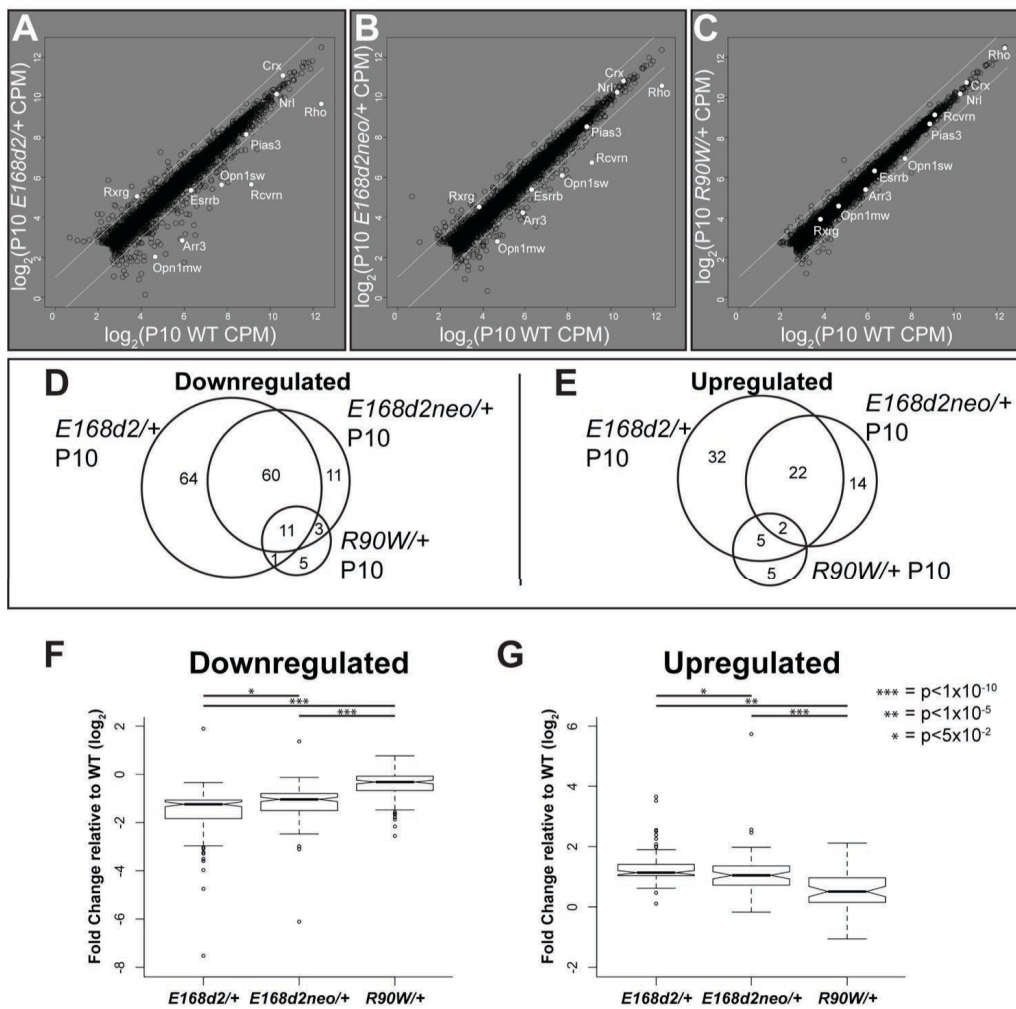
(3) Numbers represent transcripts significantly altered ( $FC \geq 2$ ,  $FDR \leq 0.05$ ) relative to *WT* control. "-": not applicable



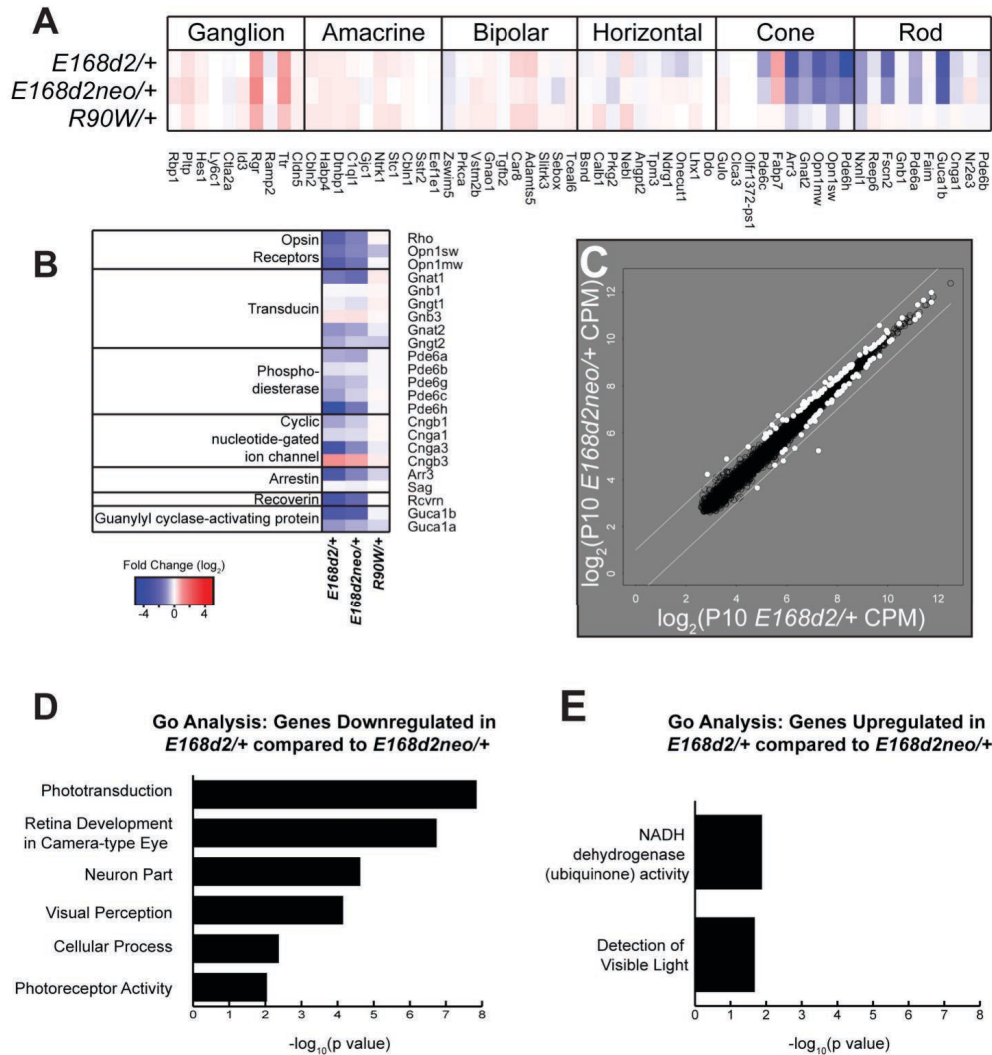
**Figure 3.1: Schematic diagram of WT and mutant CRX proteins made by the indicated mouse models.** The WT CRX protein shown on the top is 299 amino acids (A.A.) in length and contains the indicated DNA-binding and transactivation domains (indicated by bars above diagram) and several conserved motifs (marked by solid grey and black boxes). The substitution mutation *R90W* lies within the homeodomain and reduces DNA-binding [Tran et al., 2014]. The frameshift mutation *E168d2* results in a C-terminus truncated CRX protein that retains DNA-binding capability but fails to activate transcription, and therefore is antimorphic [Tran et al., 2014]. The frameshift mutation *G255d1* '*Rip*' results in a non-homologous C-terminal extension (dark grey hashed box) creating an antimorphic protein that no longer binds DNA [Roger et al., 2014]. The phenotypes of heterozygous mice carrying each of these mutations are summarized in Table 1.



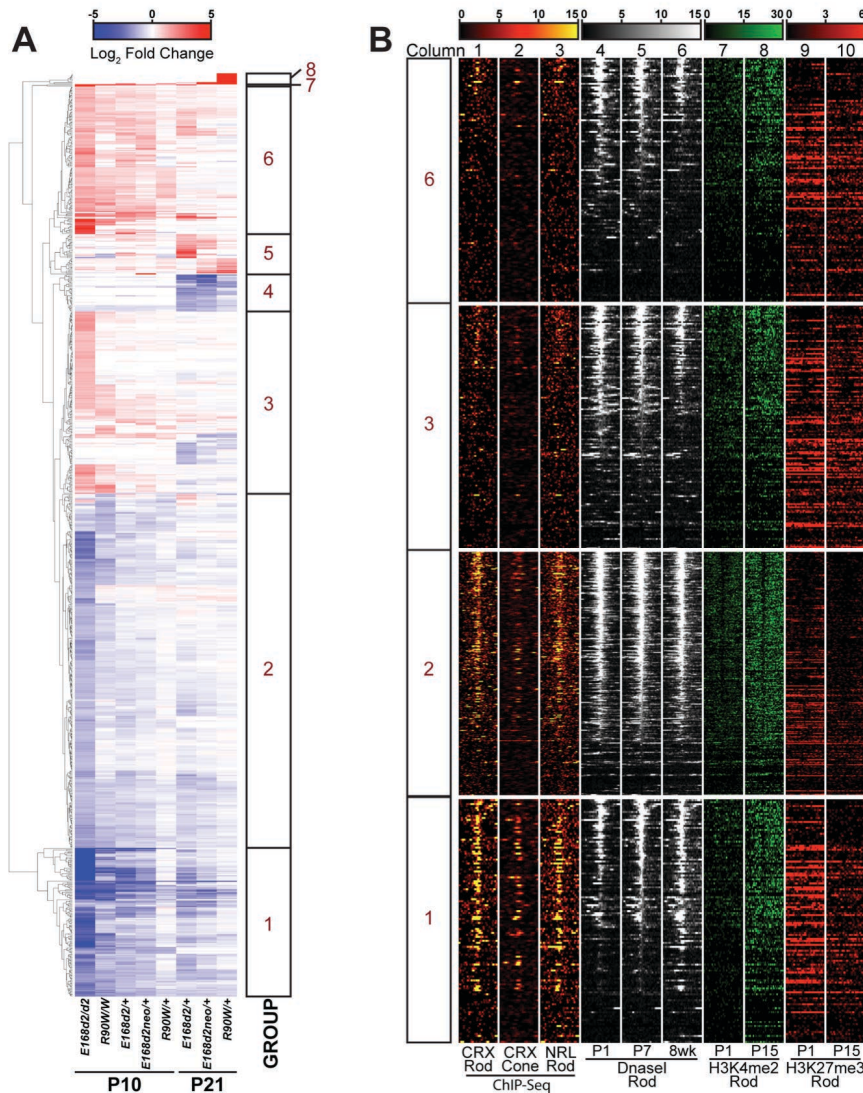
**Figure 3.2: RNA-seq analyses describe overlapping sets of affected genes in P10 homozygous *Crx* mutant retinas.** (A-B) Gene expression [ $\text{Log}_2$  Count Per Million (CPM)] in the indicated homozygous mutants (y-axes) is compared with the *WT* control (*C57BL/6J*, x-axes) (A-B). White highlighting indicates prototypical photoreceptor transcripts. White diagonal lines represent  $\pm 2$  Fold Change. (C-D) Venn diagrams illustrate the numbers of significantly affected genes that are shared or are uniquely changed in indicated mutants [fold change (FC)  $\geq 2$ , false discovery rate (FDR)  $\leq 0.05$ ].



**Figure 3.3: RNA-seq analyses detect graded changes in gene expression in P10 heterozygous *Crx* mutant retinas.** (A-C) Gene expression [ $\text{Log}_2$  CPM] in the indicated heterozygotes (y-axes) are compared with the *WT* control (*C57BL/6J*, x-axes). White highlighting indicates prototypical photoreceptor transcripts. White diagonal lines represent  $\pm 2$  Fold Change. (D-E) Venn diagrams illustrate the numbers of significantly affected genes that are shared or are uniquely changed in the indicated mutants ( $\text{FC} \geq 2$ ,  $\text{FDR} \leq 0.05$ ). (F-G) Analysis of the union of genes affected in heterozygous mutants ( $\text{FC} \geq 2$ ,  $\text{FDR} \leq 0.05$ ) presented as FC relative to *WT* control. Significance calculated by Wilcoxon Rank-Sum Test with Bonferroni correction. Notched Box Whisker plot describes median and quartiles of data.

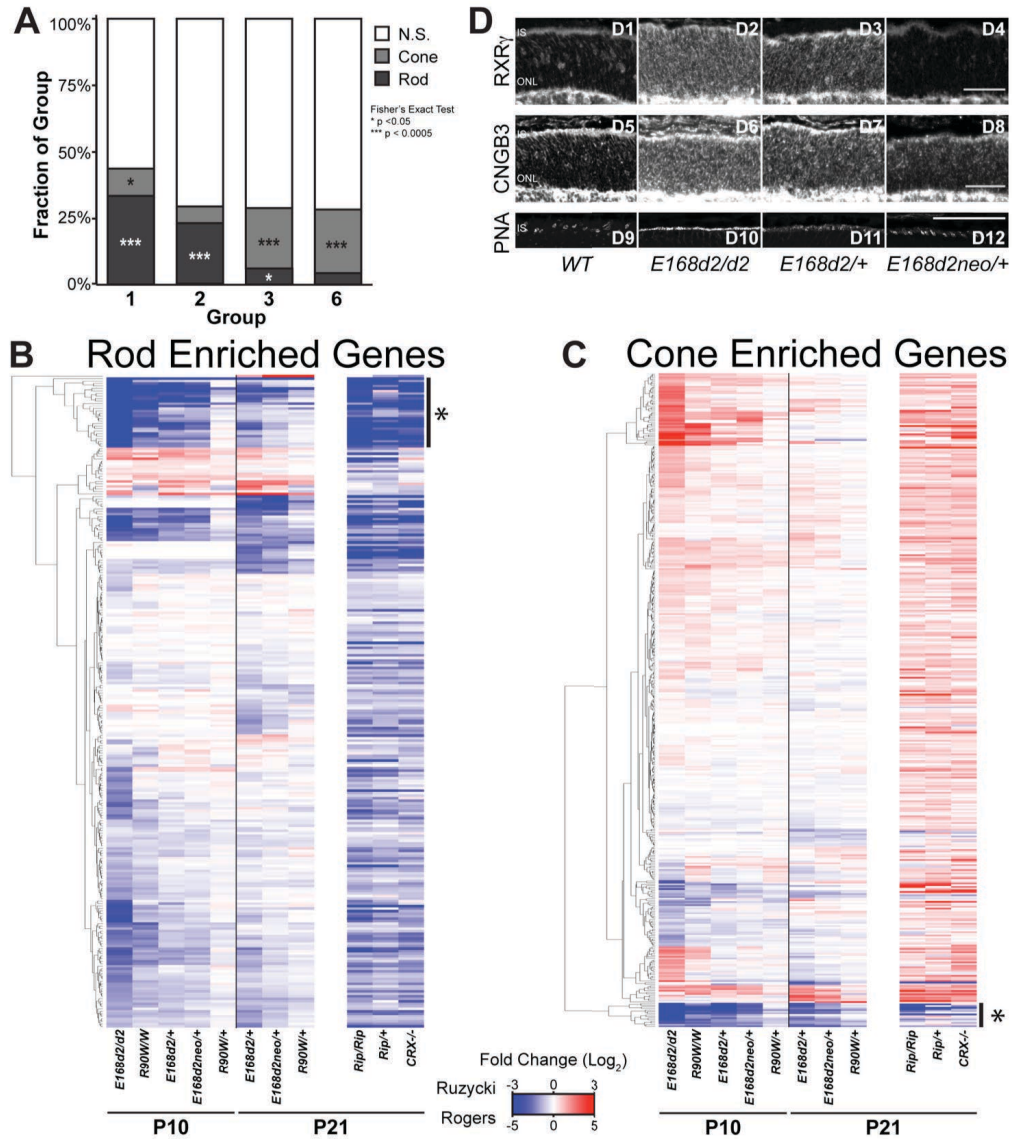


**Figure 3.4: *Crx* mutant retinas show graded changes in photoreceptor-specific gene expression.** (A-B) Heatmaps present FC of gene expression in P10 *Crx* mutants relative to that of *WT* age-matched controls in subsets of genes with highest cell type specificity ratios for the 6 retinal neurons (A) and photoreceptor-enriched genes involved in phototransduction (B). (C) Changes in gene expression levels [ $\text{Log}_2$  CPM] are compared for *E168d2/+* (x-axis) vs. *E168d2neo/+* (y-axis). Highlighted genes represent 118 transcripts significantly differentially expressed between the mutants ( $\text{FDR} \leq 0.05$ , no FC cut-off; see Supplemental Figure 3.5 and Table S4 for designations). (D-E) Top GO terms of down-regulated (D) and up-regulated (E) gene sets in *E168d2/+* compared to *E168d2neo/+*.

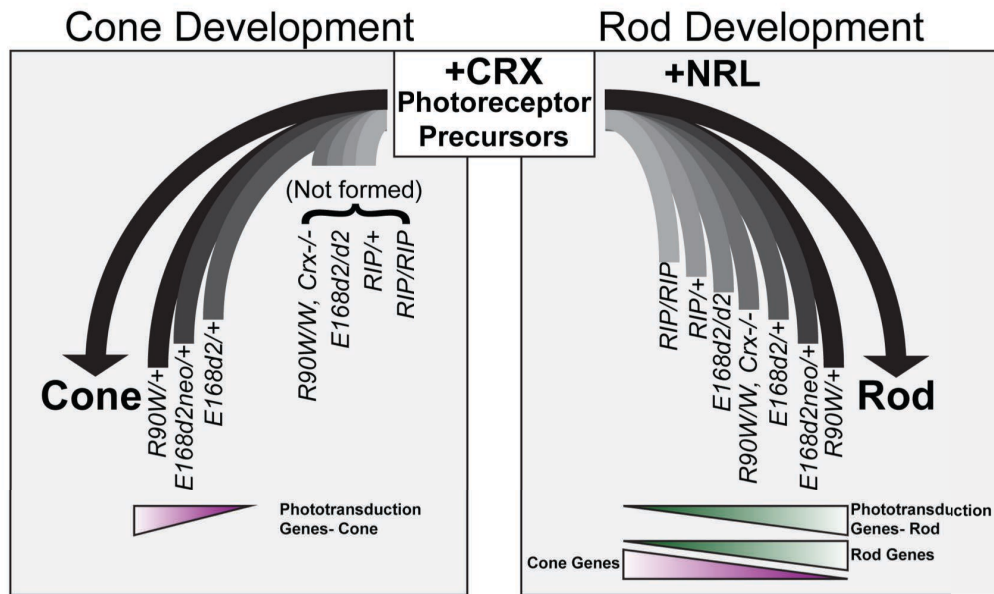


**Figure 3.5: Hierarchical clustering and epigenetic data reveal groups of similarly-regulated genes.** (A) Hierarchical clustering analysis of all genes that showed significantly altered expression ( $FC \geq 2$ ,  $FDR \leq 0.05$ ) in at least one mutant genotype relative to age-matched *WT* expression. Expression levels in the indicated genotypes at the indicated ages are indicated by the blue-red heatmap. Eight groups of genes (indicated by the bars on the right) are clearly defined by the results (see Table S4 for designations). (B) Heatmaps showing the epigenetic landscape near the transcription start site (TSS) of genes in groups 1, 2, 3, and 6. Each row represents +/- 1kb from the TSS (at center of panel) of individual genes contained within the groups as noted on the left; rows are ordered within each group by decreasing Dnase I hypersensitivity (DHS) at the TSS in the 8wk dataset. Columns 1 and 2: CRX binding determined by ChIP-seq in adult *WT* and *Nrl*<sup>-/-</sup> retinas, respectively; Column 3: NRL ChIP-seq in adult *WT* retinas; Columns 4-6: DHS of *WT* retinas at 3 indicated ages, Columns 7-8: H3K4me2 ChIP-seq in *WT* retinas at P1 and P15, Columns 9-10: H3K27me3 ChIP-seq in *WT* retinas at P1 and P15. Quantification is presented in Supplemental Figure 3.9.

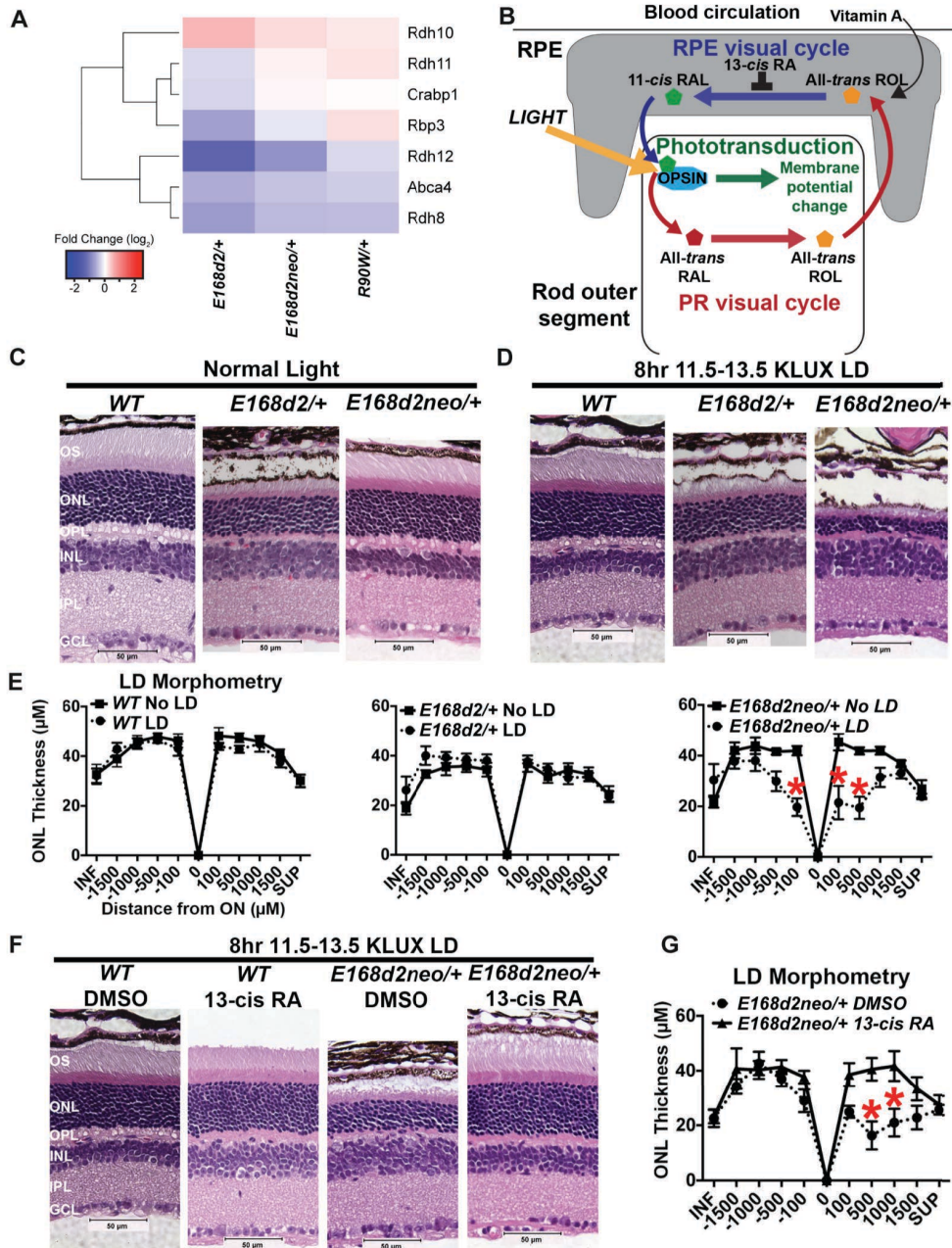




**Figure 3.6: *Crx* mutants lose rod gene expression but up-regulate many phototransduction-unrelated cone gene transcripts.** (A) Classification of genes within groups 1, 2, 3, and 6 as rod-enriched (dark grey), cone-enriched (light grey), or not specific to a particular cell type (N.S., white) in the normal adult retina. Rod and cone-enriched genes are defined based on the comparison of published P21 *WT* to *NRL*<sup>-/-</sup> RNA-seq data [Roger et al., 2014]. (Rod: FC  $\leq 2$ , FDR  $\leq 0.05$ ; Cone: FC  $\geq 2$ , FDR  $\leq 0.05$ ; Fishers Exact Test, \* p < 0.05, \*\*\* p < 0.0005). (B-C) Heatmaps depict hierarchical clustering of FC relative to *WT* for all rod-enriched (B) and cone-enriched (C) genes. Expression data from a published RNA-seq study describing the *Crx Rip*<sup>+/+</sup>, *Rip/Rip*, and *Crx*<sup>-/-</sup> mice are also separately presented in the order determined by the aforementioned clustering. Regions noted in (B) and (C) with asterisks are presented in larger format in Supplemental Figure 3.10 and represent many down-regulated genes involved in phototransduction in rods and cones. (D) Immunohistochemical staining for RXR $\gamma$ , CNGB3, and PNA in P10 WT and 3 CRX mutant retina sections. Scale bar: 100 $\mu$ m.

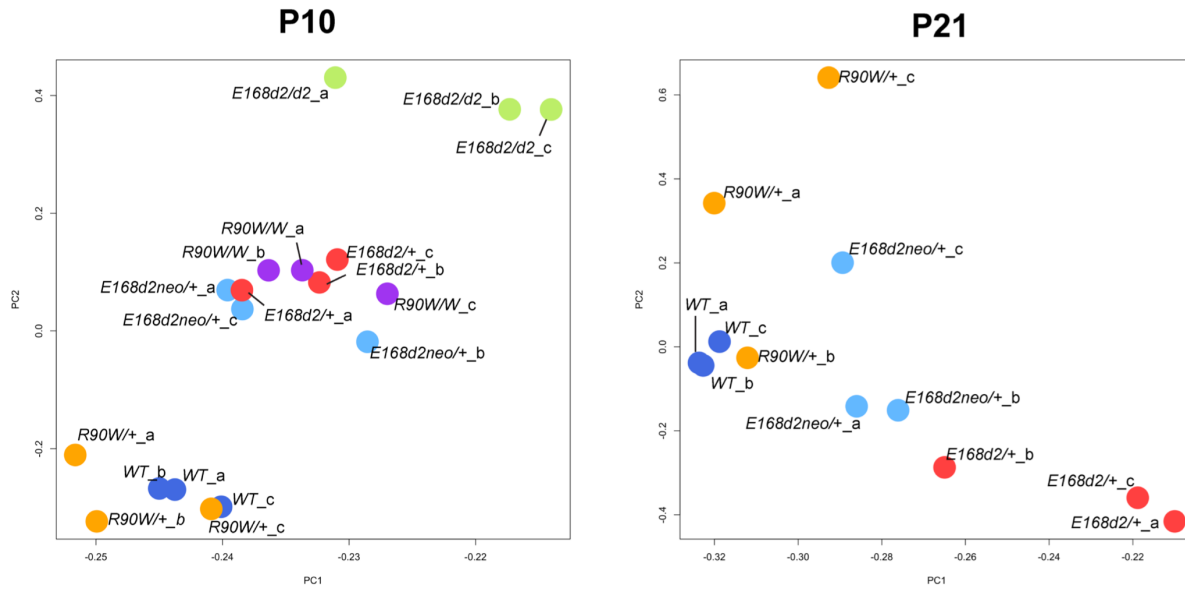


**Figure 3.7: Model of how *Crx* mutation-caused gene expression changes affect rod and cone development.** Left panel describes the formation of cones in a subset of the *Crx* mutants and variable levels of those cells' expression of phototransduction genes. Right panel shows how development of rods in all models is related to their gene expression changes. It also emphasizes the novel findings that *Crx* mutant rods display a graded phenotype of both the decreased expression of proper rod genes, and the mis-expression of cone genes.

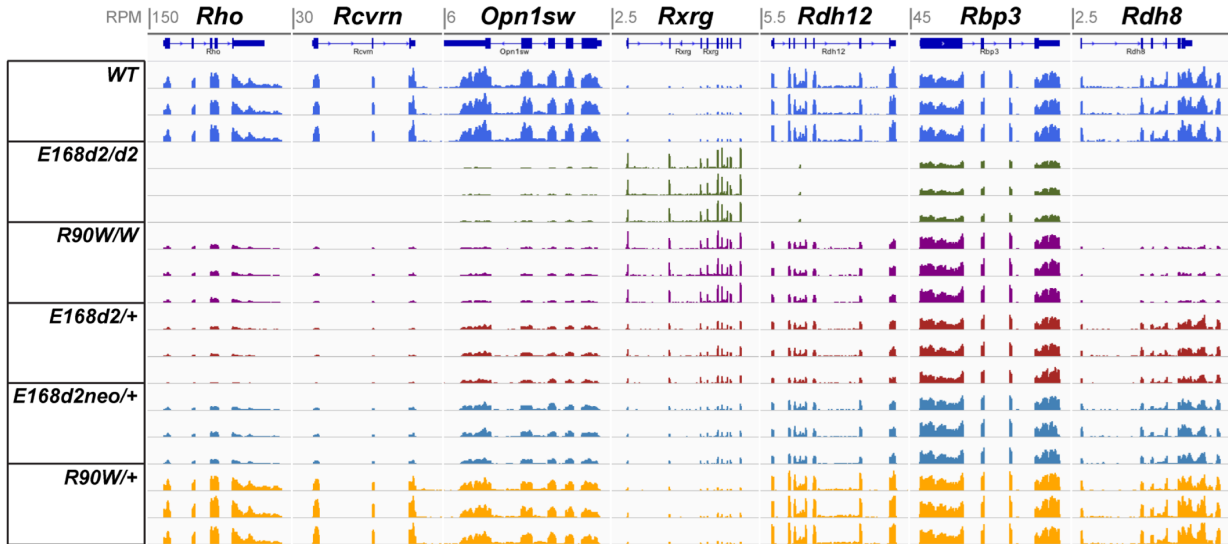
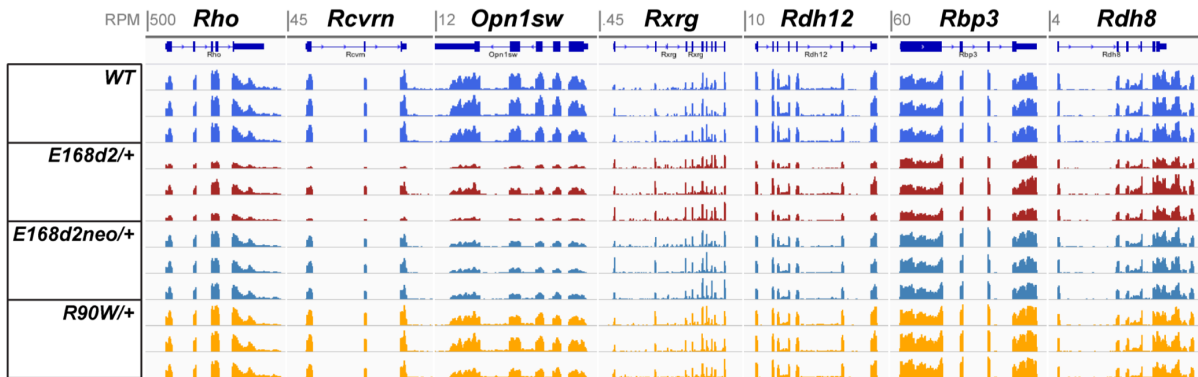


**Figure 3.8: Light-dependent photoreceptor degeneration in *E168d2neo/+* mice.** (A) Heatmap presents P21 *Crx* mutants FC relative to *WT* for genes involved in photoreceptor visual cycle. (B) Model for the RPE visual cycle: visual chromophore [11-*cis*-retinal (RAL)] is generated in the retinal pigmented epithelium (RPE) and delivered to the photoreceptor; stimulation of 11-*cis* RAL bound to an opsin protein activates the phototransduction cascade leading to photoreceptor membrane hyperpolarization; this process isomerizes 11-*cis* RAL into all-*trans* RAL, which is then recycled back into 11-*cis*-RAL through a series of enzymatic steps, known as the visual cycle, in the photoreceptor and RPE. The drug 13-*cis* retinoic acid (13-*cis*-RA) blocks the

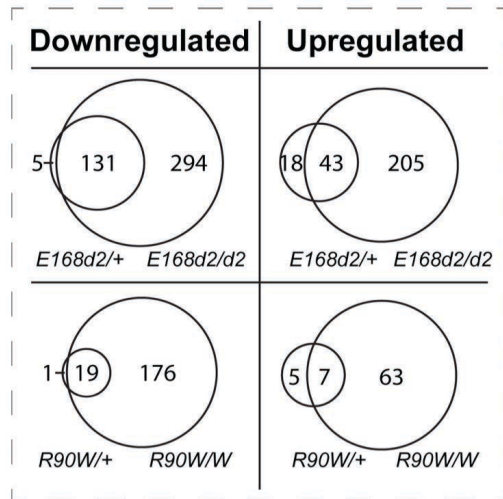
synthesis of 11-*cis*-RAL, effectively reducing the burden on the retinoid cycle. **(C-D)** Retinal morphology of 6 wo mice with indicated genotypes under normal light conditions: **(C)** (12 hr room light-dark cycle) or **(D)** light damage (LD) treatments [8 hr 11.5-13.5 kilolux (KLUX) bright light followed by 7 days of normal light-dark cycles]. Retinal morphology was assessed by hematoxylin and eosin (H&E) staining of sagittal retinal sections through the optic nerve (ON). Images were taken in the central superior retina (~500  $\mu\text{m}$  from the ON). **(E)** Morphometry quantification of ONL thickness for the samples presented in (C-D). Error bars represent Standard error of mean (SE) from  $\geq 3$  biological replicates and significance was calculated using two-way ANOVA. Data points that significantly differ from the control ( $p \leq 0.05$ ) are marked by red asterisks. **(F-G)** Effect of 13-*cis*-RA pretreatment on LD in *E168d2neo/+* mutant mice: Retinal morphology **(F)** and morphometry **(G)** at 7 days following LD where mice were pretreated with 13-*cis*-RA or DMSO as control. Note the significant improvements of ONL thickness and photoreceptor OS length in the mice treated with 13-*cis*-RA. OS- outer segment, ONL- outer nuclear layer, OPL- outer plexiform layer, INL- inner nuclear layer, IPL- inner plexiform layer, GCL- ganglion cell layer. Scale bar: 50  $\mu\text{m}$ .



**Supplemental Figure 3.1: P10 and P21 biological replicates clustered by Principal Component Analysis.** PCA plot shows expected distribution based on functional and morphological phenotypes. PC1 and PC2 represent greater than 98% of the variance of the data in both analyses.

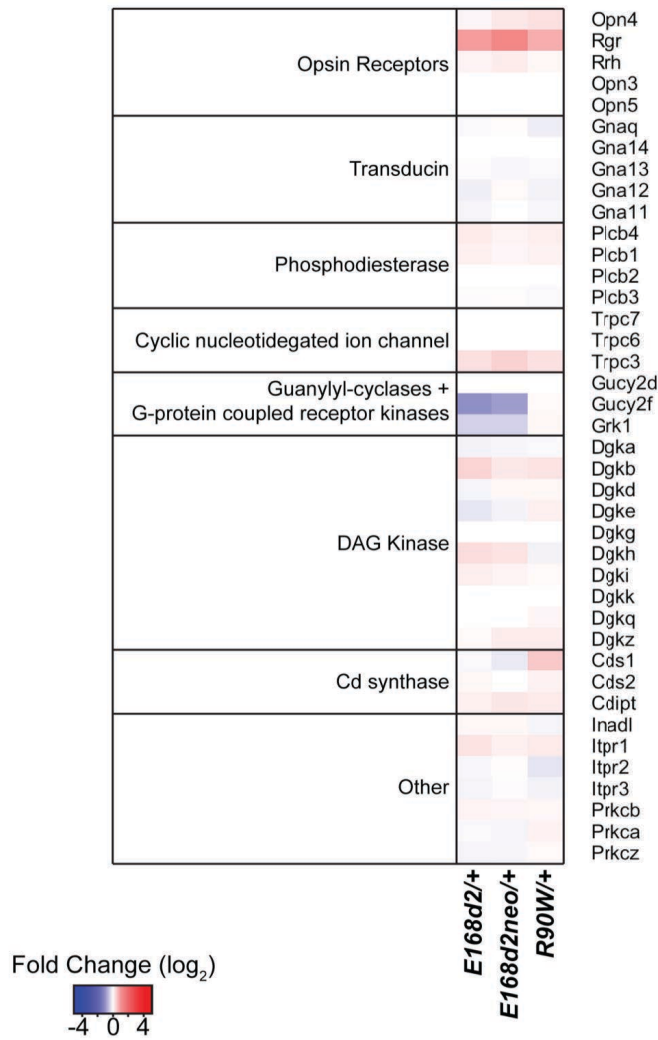
**A****P10****B****P21**

**Supplemental Figure 3.2: Browser images display RNaseq read depth of rod and cone transcripts.** P10 (A) and P21 (B) browser shots of RPM normalized mapped read count for selected genes for each biological replicate show reproducibility of sequencing results between samples.



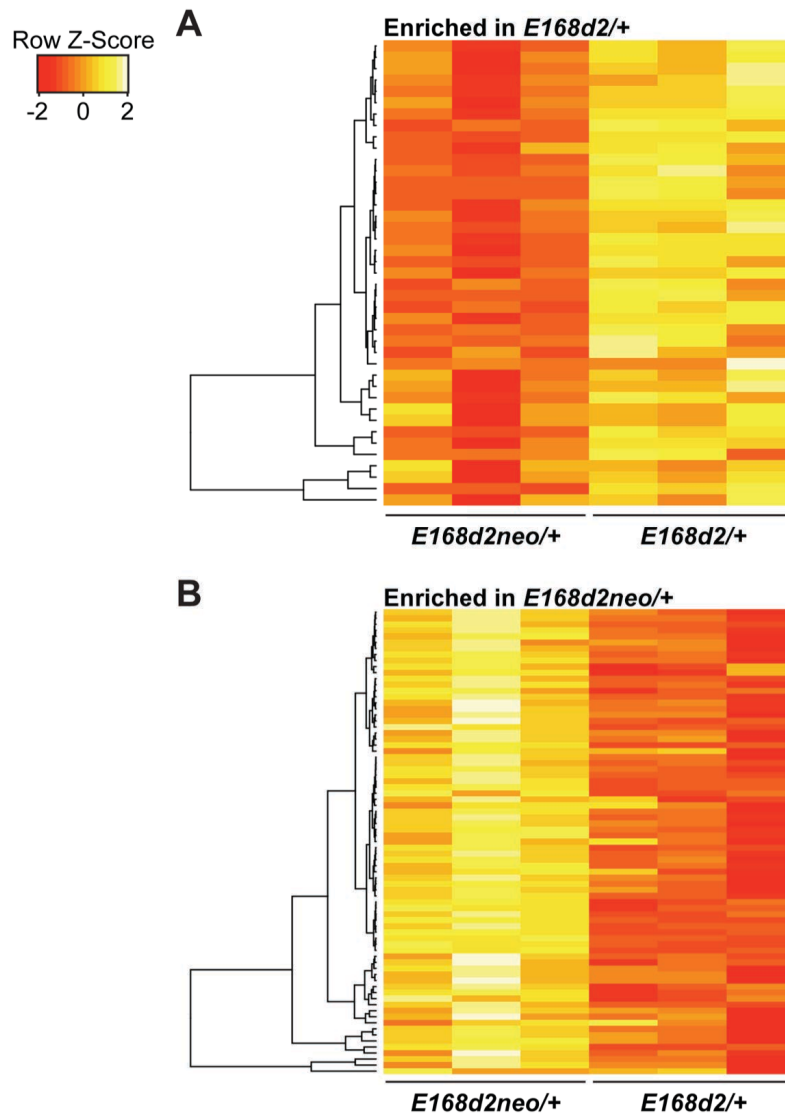
**Supplemental Figure 3.3: Homozygous and heterozygous mutants show largely overlapping datasets.** Venn diagrams comparing significantly affected transcripts ( $FC \geq 2$ ,  $FDR \leq 0.05$  relative to *WT*) in the indicated heterozygous and homozygous *Crx* mutant retinas.



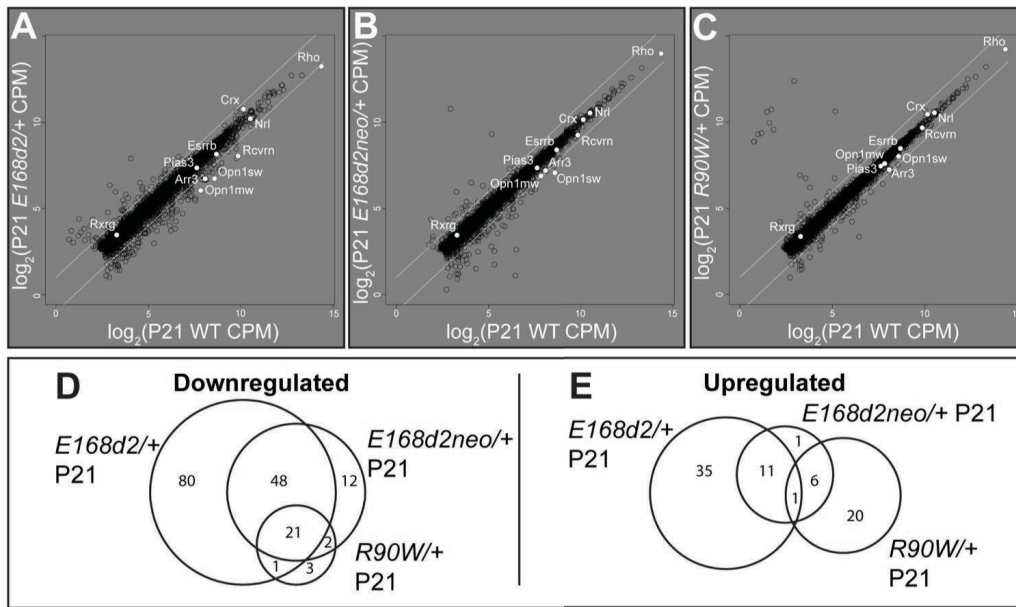


**Supplemental Figure 3.4: *Crx* mutants show little change to other non-photoreceptor phototransduction components.** Heatmap depicts P10 FC relative to *WT* for genes involved in phototransduction enriched in non-photoreceptor retinal cell types in the heterozygous mutants.

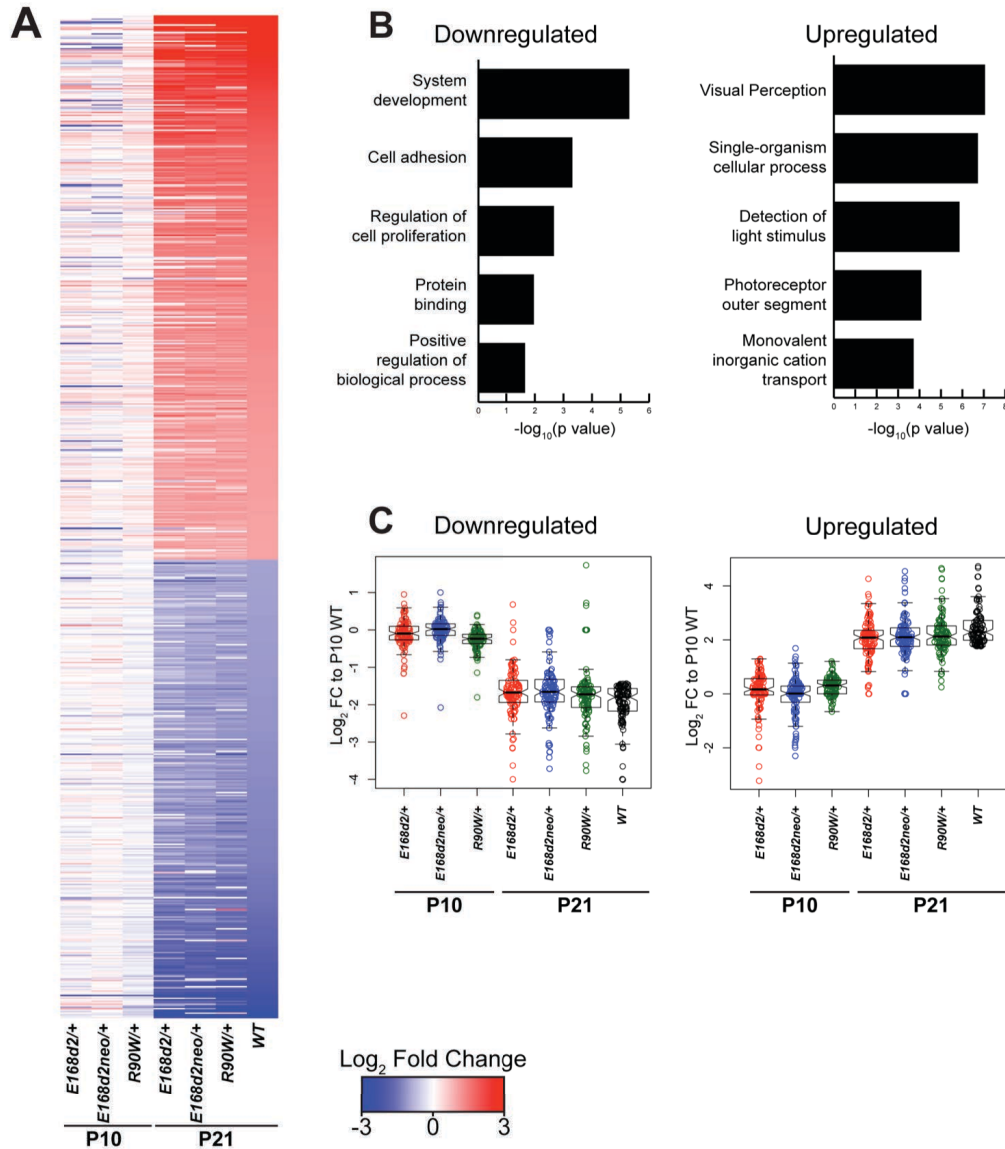




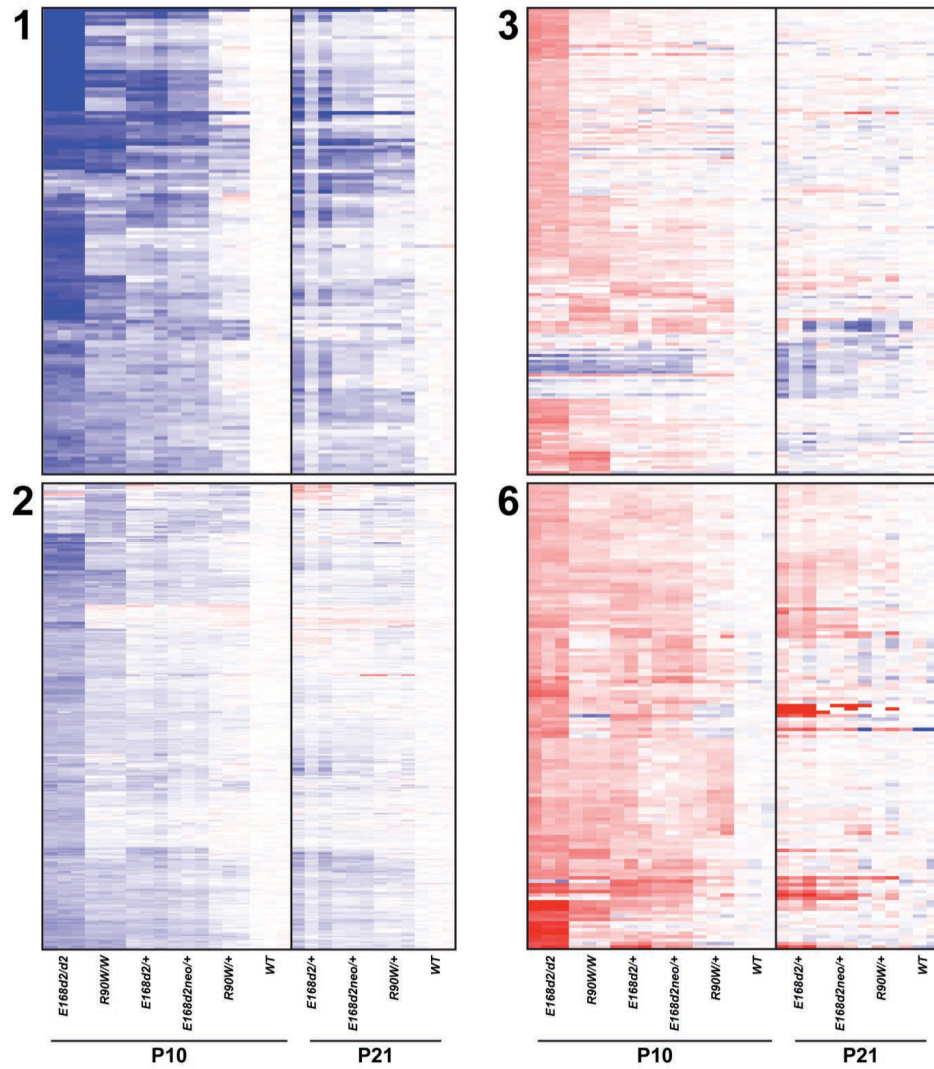
**Supplemental Figure 3.5: Biological replicates show consistency in expression changes between genotypes.** Heat map showing the Z-scores comparing 3 biological replicates of genes determined to be enriched in either *E168d2/+* (**A**) or *E168d2neo/+* (**B**) retinas. Genes are arranged in rows, clustering dendrogram is shown on the left. Patterns of expression across biological replicates are consistent, supporting the statistical significance of the difference calculated by EdgeR RNA-seq analysis.



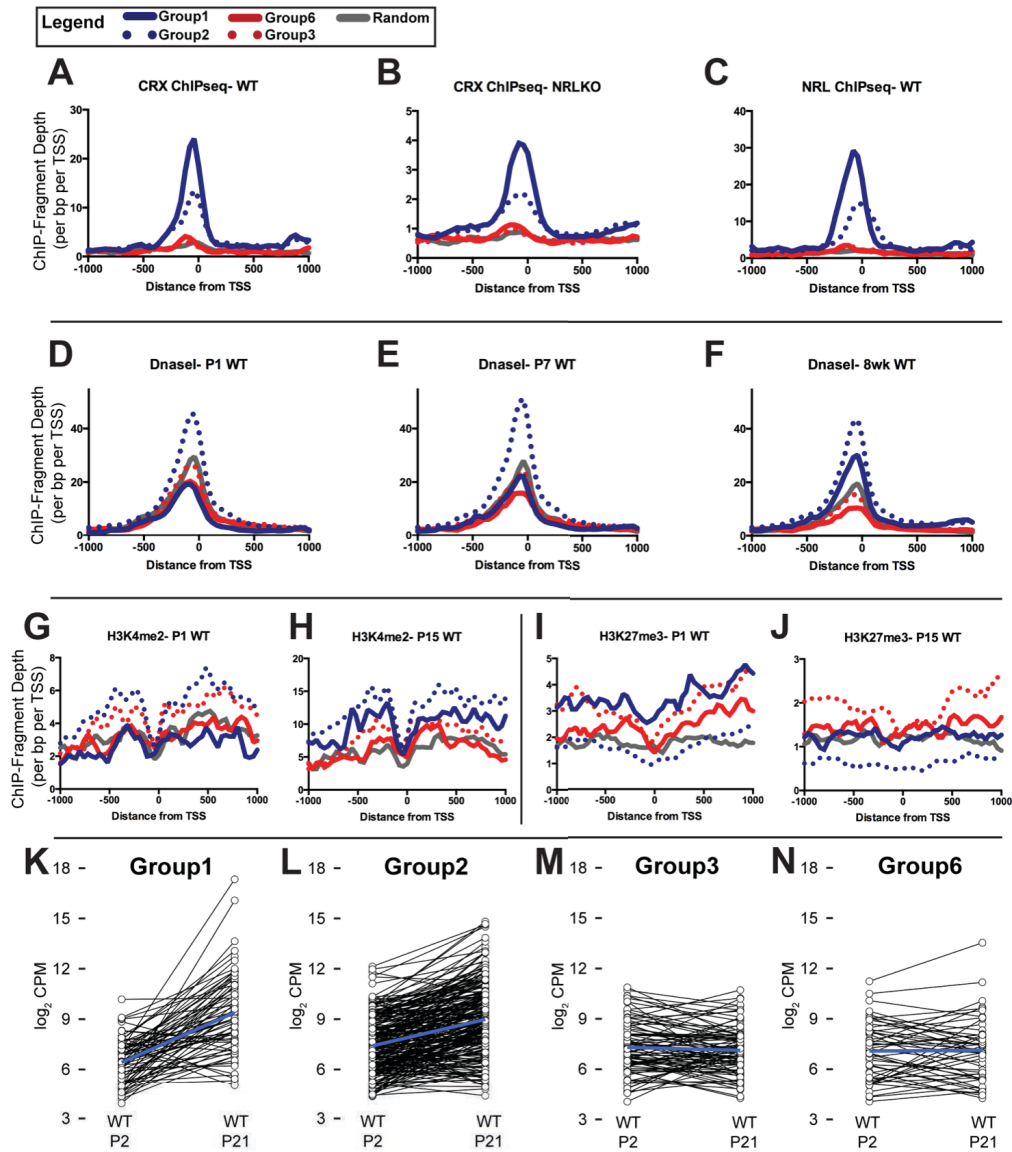
**Supplemental Figure 3.6: *Crx* mutants show graded changes in gene expression of overlapping gene sets at P21.** (A-C)  $\log_2$  CPM from P21 heterozygous *E168d2/+* (A), *E168d2neo/+* (B), and *R90W/+* (C) mice (y-axes) are compared to  $\log_2$  CPM from age-matched *WT C57Bl/6J* (x-axes) mice. White letters highlight prototypical photoreceptor transcripts. White diagonal lines represent  $\pm 2$  Fold Change. (D-E) Venn diagrams illustrate the numbers of overlapping and distinct significantly affected genes ( $FC \geq 2$ ,  $FDR \leq 0.05$ ).



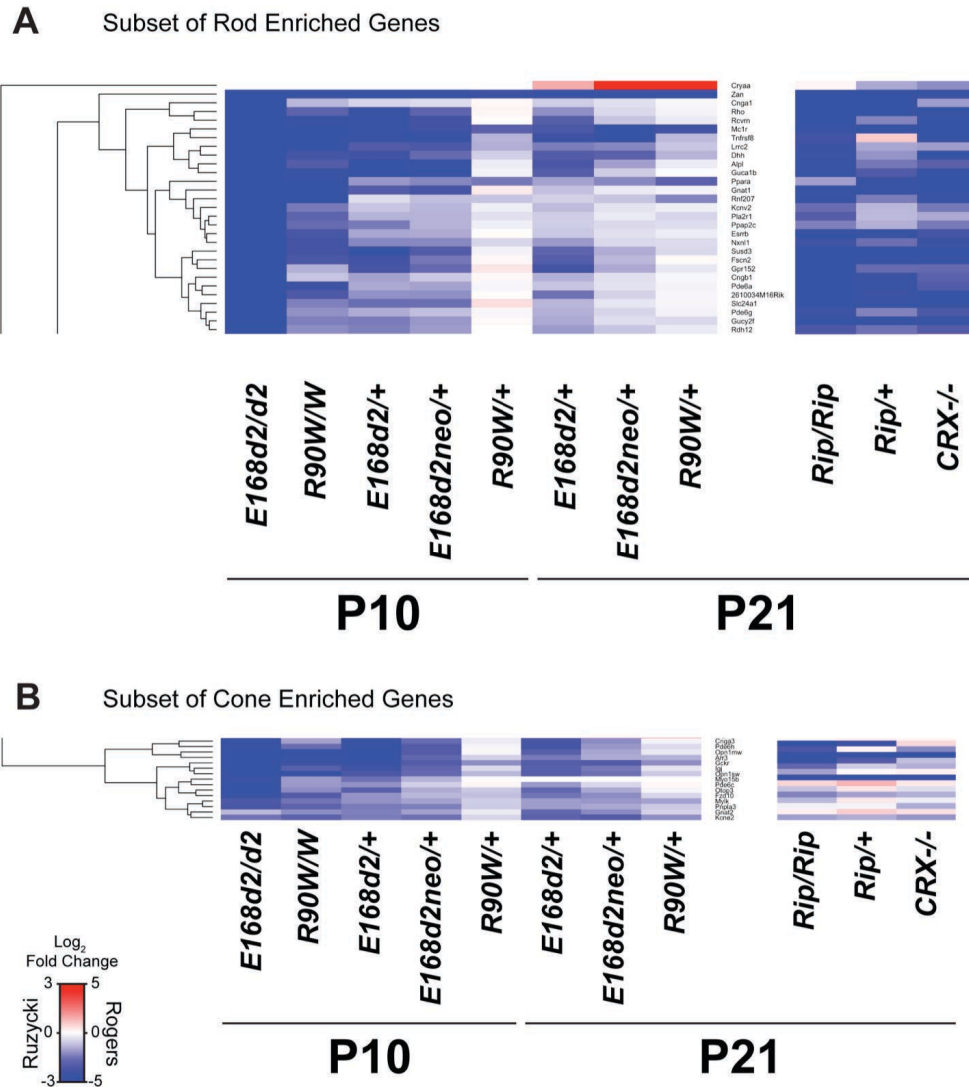
**Supplemental Figure 3.7: *Crx* mutants do not abandon developmental program, but many genes fail to reach proper expression levels.** (A) Heatmap, ordered by the magnitude of expression change between P10 and P21 in *WT* retina, depicts FC compared to P10 *WT* to analyze developmental dynamics. (B) Top GO terms from analysis of top 100 genes that are down- (left panel) and up- (right panel) regulated normally during late postnatal retinal development (P21 *WT* vs. P10 *WT*; FC  $\leq$  or  $\geq$  2, FDR  $\leq$  0.05). (C) Violin and boxplots quantifying median FC from P10 *WT* of each mutant.



**Supplemental Figure 3.8: Biological replicates show consistent expression changes within each genotype.** Groups 1, 2, 3, and 6 genes (defined in Figure 5) presented as a heatmap with calculated fold-change values for each biological replicate compared to the mean age-appropriate *WT* value. All genes are ordered exactly as presented in Figure 5.

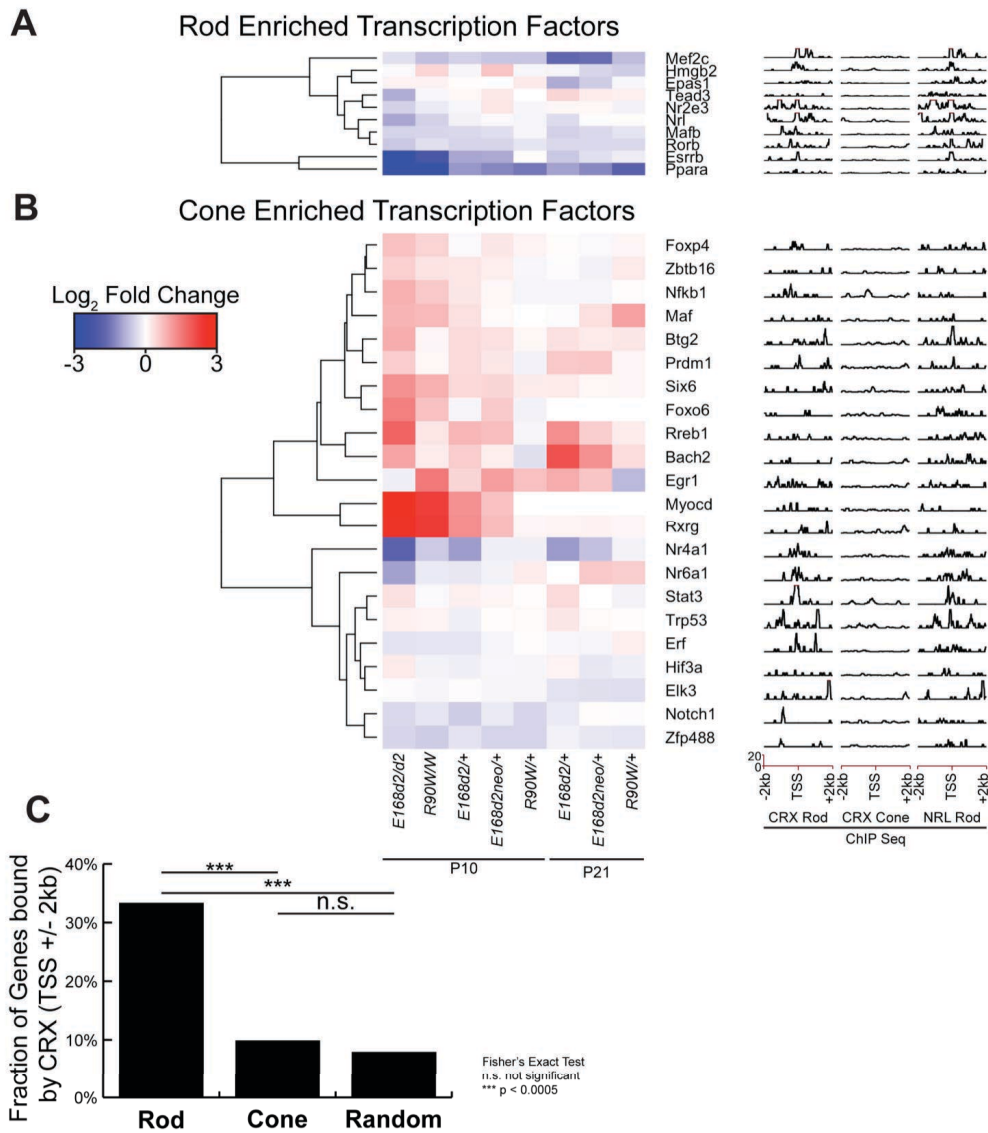


**Supplemental Figure 3.9: Quantification of normal epigenetic marks near genes in Groups 1, 2, 3, and 6 distinguishes up- and down-regulated genes in *Crx* mutants.** Control set of equal sized random group of mm9 genes used as background control (grey in all). **(A-C)** Quantification of CRX ChIP-seq [Corbo et al., 2010] in rod dominant *WT* retina **(A)** and cone dominant *Nrl*<sup>-/-</sup> retina **(B)**, and NRL ChIP-seq [Hao et al., 2012] in *WT* retina **(C)**. **(D-F)** DnaseI Hypersensitivity plots from *WT* retina at the three indicated ages [Vierstra et al., 2014]. **(G-H)** H3K4me2 ChIP-seq of *WT* retina at the two ages [Popova et al., 2012]. **(I-J)** H3K27me3 ChIP-seq of *WT* retina at the two ages. All represent normalized mean read depth of each library centered on TSS of genes within Groups 1, 2, 3, and 6. **(K-N)** Analysis of *WT* expression of each gene and average change in expression (blue line) from the age of P2 to P21 (raw data in [Roger et al., 2014]), presented as Log<sub>2</sub> CPM.

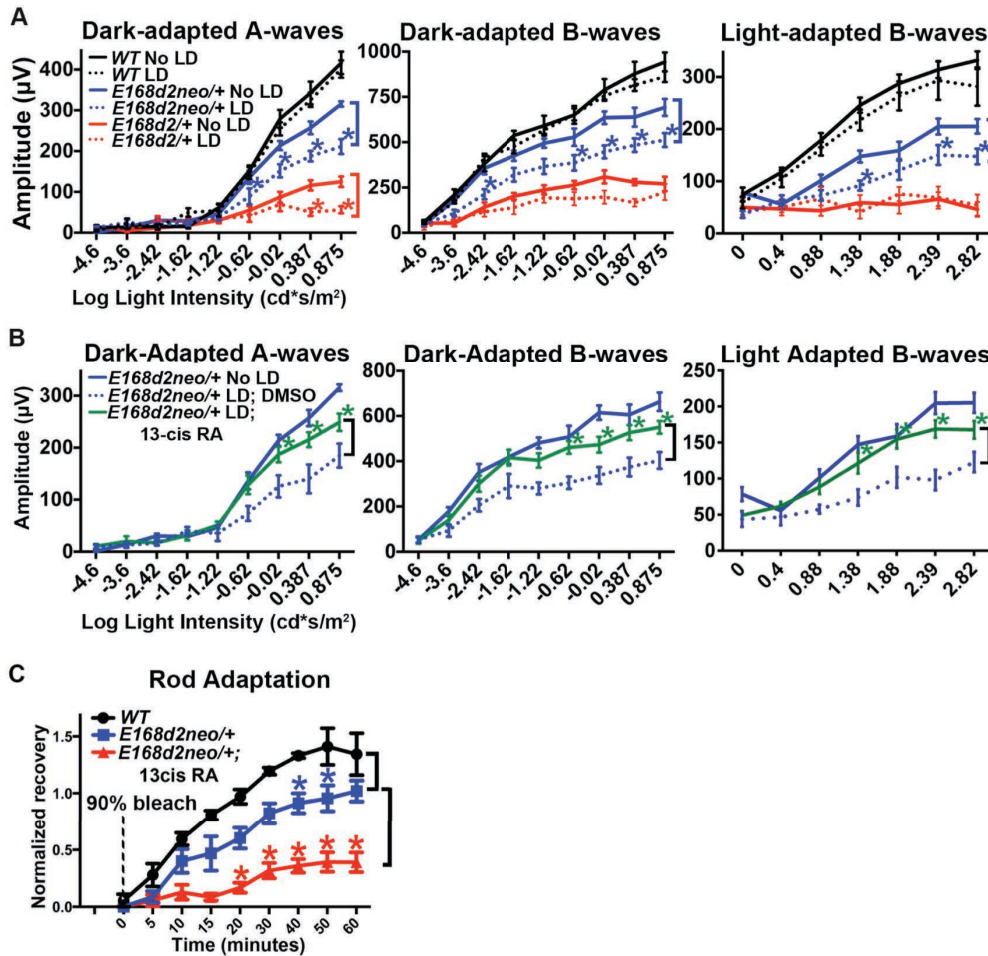


**Supplemental Figure 3.10: Loss of expression of key components of phototransduction.** Expanded view of areas marked with asterisk (\*) in Figure 6. All data and presentation exactly as described previously.



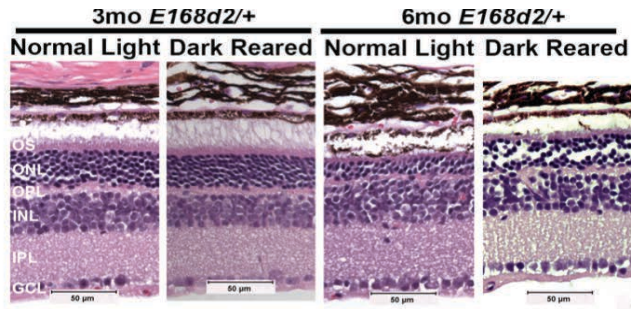


**Supplemental Figure 3.11: Changes in rod and cone enriched transcription factors are consistent with general patterns of rod and cone genes expression in CRX mutants. (A-B)** Heatmap describes FC relative to *WT* in each *Crx* mutant for rod (A) and cone (B) enriched genes annotated as transcription factors (Raw data [Roger et al., 2014]; *WT* vs *Nrl*<sup>-/-</sup> RNA-seq  $FC \leq$  or  $\geq 2$  &  $FDR \leq 0.05$ ; GO: 0003700). ChIP-seq data is presented for 4 kb window surrounding TSS for each gene for CRX (in rods and cones) and NRL (in rods). (C) Analysis detailing percent of rod enriched, cone enriched, or random set of genes with CRX ChIP-seq peak within 2kb of TSS in *WT* retina. Note the significant enrichment of CRX binding in the rod set over cone and random gene sets.

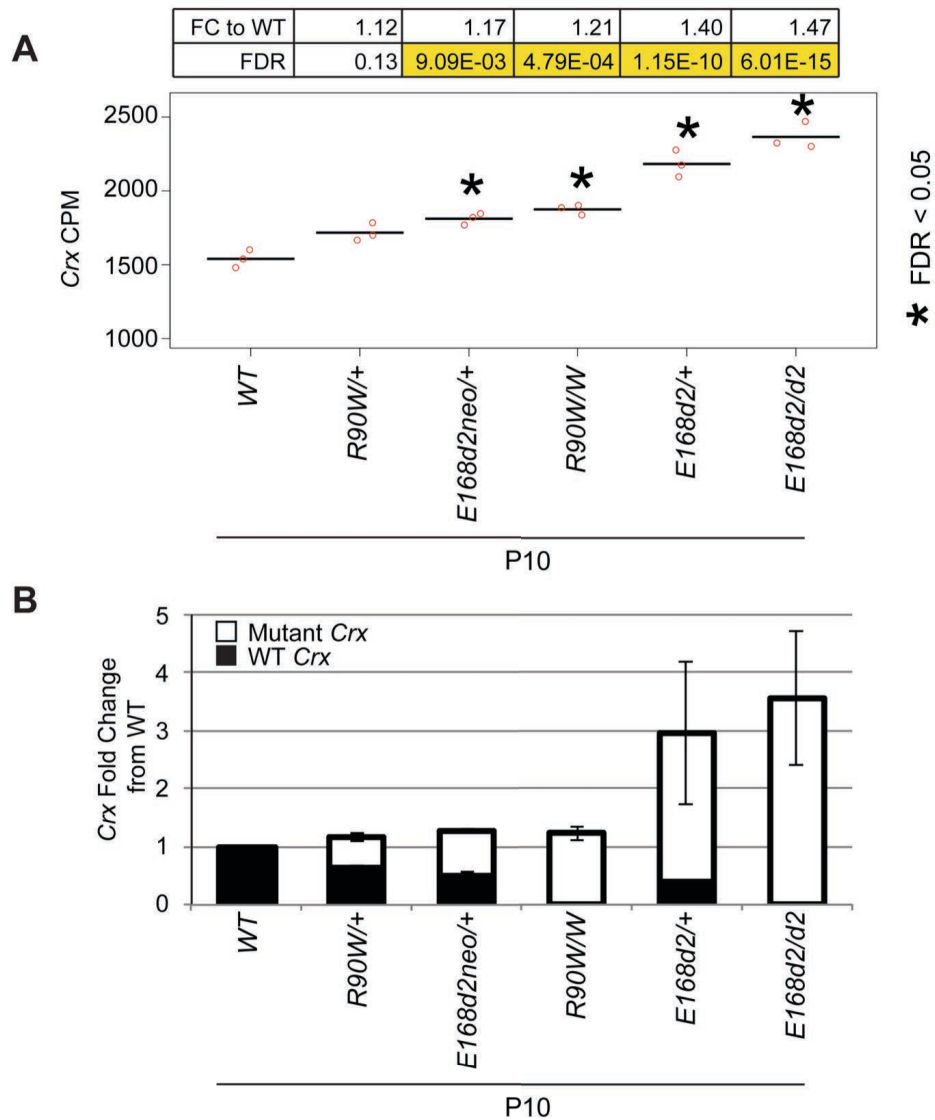


**Supplemental Figure 3.12: Retinal function is affected by light damage (LD) in *E168d2neo/+* but not *E168d2/+* mutant mice.** (A) Intensity-response plots for dark-adapted ERG a-wave, dark-adapted b-wave and light-adapted b-wave from mice with the indicated genotypes, with or without LD treatment. (B) Intensity-response plots for ERG responses following LD for mice pretreated with either 13-*cis*-RA or DMSO. (C) Rod dark adaptation measured by recovery of maximal rod ERG a-wave following >90% rhodopsin photobleach for WT and *E168d2neo/+* mice, with and without drug treatment. Error bars represent Standard error of mean (SE,  $n \geq 3$ ). “\*” marks the data point of significant difference ( $p \leq 0.05$ ) from the control as determined by two-way ANOVA (see Appendix 2 Methods).





**Supplemental Figure 3.13: *E168d2/+* photoreceptor degeneration is light-independent.** Retinal morphology of *E168d2/+* mice raised under 12 hr light-dark cycle (normal light) or constant darkness (dark-reared) for 3 or 6 months. Note that the dark rearing did not improve ONL thinning in the mutant retina. OS- outer segment, ONL- outer nuclear layer, OPL- outer plexiform layer, INL- inner nuclear layer, IPL- inner plexiform layer, GCL- ganglion cell layer. Scale bar: 50 μM.



**Supplemental Figure 3.14: Increased *Crx* expression in mutant lines at P10.** (A) RNA-seq derived raw CPM values for each biological replicate (red circles) and mean (black bar). Top panel of spreadsheet lists the EdgeR calculated FC and FDR for each genotype relative to *WT* control. Asterisk denotes samples where FDR < 0.05. (B) qRT-PCR analysis of transcript levels of *WT* and mutant *Crx* alleles in each genotype. Note the same trend of *Crx* expression changes in these mutants detected by both methods.

## 3.6 Supporting Information

Supporting information including the following is available at:

<https://genomebiology.biomedcentral.com/articles/10.1186/s13059-015-0732-z>

Table S1

Table S2

Table S3

Table S4

Table S5

## 3.7 Data Availability

Raw RNA seq data and primary analysis are available in Gene Expression Omnibus (GEO) under accession GSE65506.

## 3.8 Acknowledgements

The authors wish to thank Xi-Qin Ding from the University of Oklahoma for providing the anti-CNGB3 antibody, Belinda Dana, Frank Schottler, and Guangyi Ling in the Imaging Core (Ophthalmology and Visual Sciences, Washington University), Mingyan Yang in the Chen Lab for technical assistance; Anne Hennig and other Chen Lab members for suggestions and critical review of the manuscript. We thank the Genome Technology Access Center of Washington University for performing Illumina sequencing and Nobish Varghese for initial analysis of RNA-seq results. We thank the Alvin J. Siteman Cancer Center at Washington University School of Medicine and Barnes-Jewish Hospital in St. Louis, MO for the use of the Center for Biomedical Informatics, which provided the in silico analysis service.

## 3.9 References

- Brooks, M. J., Rajasimha, H. K., Roger, J. E., & Swaroop, A. (2011). Next-generation sequencing facilitates quantitative analysis of wild-type and *Nrl(-/-)* retinal transcriptomes. *Molecular Vision*, *17*, 3034–54.
- Chen, J., Rattner, A., & Nathans, J. (2005). The rod photoreceptor-specific nuclear receptor Nr2e3 represses transcription of multiple cone-specific genes. *The Journal Of Neuroscience*, *25*, 118–29.
- Chen, S., Peng, G.-H., Wang, X., Smith, A. C., Grote, S. K., Sopher, B. L., & La Spada, A. R. (2004). Interference of Crx-dependent transcription by ataxin-7 involves interaction between the glutamine regions and requires the ataxin-7 carboxy-terminal region for nuclear localization. *Human Molecular Genetics*, *13*, 53–67.
- Chen, S., Wang, Q. L., Nie, Z., Sun, H., Lennon, G., Copeland, N. G., Gilbert, D. J., et al. (1997). Crx, a novel Otx-like paired-homeodomain protein, binds to and transactivates photoreceptor cell-specific genes. *Neuron*, *19*, 1017–30.
- Corbo, J. C., & Cepko, C. L. (2005). A hybrid photoreceptor expressing both rod and cone genes in a mouse model of enhanced S-cone syndrome. *PLoS Genetics*, *1*, e11.
- Corbo, J. C., Lawrence, K. a, Karlstetter, M., Myers, C. a, Abdelaziz, M., Dirkes, W., Weigelt, K., et al. (2010). CRX ChIP-seq reveals the cis-regulatory architecture of mouse photoreceptors. *Genome Research*, *20*, 1512–25.
- Daniele, L. L., Lillo, C., Lyubarsky, A. L., Nikonov, S. S., Philp, N., Mears, A. J., Swaroop, A., et al. (2005). Cone-like morphological, molecular, and electrophysiological features of the photoreceptors of the *Nrl* knockout mouse. *Investigative Ophthalmology & Visual Science*, *46*, 2156–67.
- Furukawa, T., Morrow, E. M., Li, T., Davis, F. C., & Cepko, C. L. (1999). Retinopathy and attenuated circadian entrainment in *Crx*-deficient mice. *Nature Genetics*, *23*, 466–70.
- Grimm, C., Wenzel, A., Hafezi, F., Yu, S., Redmond, T. M., & Remé, C. E. (2000). Protection of *Rpe65* - deficient mice identifies rhodopsin as a mediator of light-induced retinal degeneration. *Nature Genetics*, *25*, 63–66.
- Hao, H., Kim, D. S., Klocke, B., Johnson, K. R., Cui, K., Gotoh, N., Zang, C., et al. (2012). Transcriptional Regulation of Rod Photoreceptor Homeostasis Revealed by In Vivo NRL Targetome Analysis. *PLoS Genetics*, *8*, e1002649.
- Hennig, A. K., Peng, G.-H., & Chen, S. (2008). Regulation of photoreceptor gene expression by Crx-associated transcription factor network. *Brain Research*, *1192*, 114–33.
- Hennig, A. K., Peng, G.-H., & Chen, S. (2013). Transcription Coactivators p300 and CBP Are Necessary for Photoreceptor-Specific Chromatin Organization and Gene Expression. *PloS One*, *8*, e69721.
- Hsiau, T. H.-C., Diaconu, C., Myers, C. A., Lee, J., Cepko, C. L., & Corbo, J. C. (2007). The cis-regulatory logic of the mammalian photoreceptor transcriptional network. *PloS One*, *2*, e643.
- Huang, L., Xiao, X., Li, S., Jia, X., Wang, P., Guo, X., & Zhang, Q. (2012). CRX variants in cone-rod dystrophy and mutation overview. *Biochemical And Biophysical Research Communications*, *426*, 498–503.
- Maeda, A., Maeda, T., Golczak, M., Imanishi, Y., Leahy, P., Kubota, R., & Palczewski, K. (2006). Effects of potent inhibitors of the retinoid cycle on visual function and photoreceptor protection

- from light damage in mice. *Molecular Pharmacology*, 70, 1220–9.
- Maeda, A., Maeda, T., Golczak, M., & Palczewski, K. (2008). Retinopathy in mice induced by disrupted all-trans-retinal clearance. *The Journal Of Biological Chemistry*, 283, 26684–93.
- Maeda, A., Maeda, T., Sun, W., Zhang, H., Baehr, W., & Palczewski, K. (2007). Redundant and unique roles of retinol dehydrogenases in the mouse retina. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 104, 19565–70.
- Maeda, T., Maeda, A., Matosky, M., Okano, K., Roos, S., Tang, J., & Palczewski, K. (2009). Evaluation of potential therapies for a mouse model of human age-related macular degeneration caused by delayed all-trans-retinal clearance. *Investigative Ophthalmology & Visual Science*, 50, 4917–25.
- Mears, A., Kondo, M., Swain, P. K., Takada, Y., Bush, R. A., Saunders, T. L., Sieving, P. A., et al. (2001). Nrl is required for rod photoreceptor development. *Nature Genetics*, 29, 447–52.
- Mitton, K. P., Swain, P. K., Chen, S., Xu, S., Zack, D. J., & Swaroop, A. (2000). The leucine zipper of NRL interacts with the CRX homeodomain. A possible mechanism of transcriptional synergy in rhodopsin regulation. *The Journal Of Biological Chemistry*, 275, 29794–9.
- Nikolayeva, O., & Robinson, M. D. (2014). EdgeR for Differential RNA-seq and ChIP-seq Analysis: An Application to Stem Cell Biology. *Stem Cell Transcriptional Networks Methods In Molecular Biology*, 1150, 45–79.
- Noell, W. K., Walker, V. S., Kang, B. S., & Berman, S. (1966). Retinal damage by light in rats. *Investigative Ophthalmology & Visual Science*, 5, 450–473.
- Onishi, A., Peng, G.-H., Hsu, C., Alexis, U., Chen, S., & Blackshaw, S. (2009). Pias3-dependent SUMOylation directs rod photoreceptor development. *Neuron*, 61, 234–46.
- Palhan, V. B., Chen, S., Peng, G.-H., Tjernberg, A., Gamper, A. M., Fan, Y., Chait, B. T., et al. (2005). Polyglutamine-expanded ataxin-7 inhibits STAGA histone acetyltransferase activity to produce retinal degeneration. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 102, 8472–7.
- Peng, G.-H., Ahmad, O., Ahmad, F., Liu, J., & Chen, S. (2005). The photoreceptor-specific nuclear receptor Nr2e3 interacts with Crx and exerts opposing effects on the transcription of rod versus cone genes. *Human Molecular Genetics*, 14, 747–64.
- Peng, G., & Chen, S. (2007). Crx activates opsin transcription by recruiting HAT-containing co-activators and promoting histone acetylation. *Human Molecular Genetics*, 16, 2433–52.
- Popova, E. Y., Xu, X., Dewan, A. T., Salzberg, A. C., Berg, A., Hoh, J., Zhang, S. S., et al. (2012). Stage and Gene Specific Signatures Defined by Histones H3K4me2 and H3K27me3 Accompany Mammalian Retina Maturation In Vivo. *PloS One*, 7, e46867.
- Radu, R. a, Mata, N. L., Nusinowitz, S., Liu, X., Sieving, P. A., & Travis, G. H. (2003). Treatment with isotretinoin inhibits lipofuscin accumulation in a mouse model of recessive Stargardt’s macular degeneration. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 100, 4742–7.
- Rivolta, C., Berson, E. L., & Dryja, T. P. (2001). Dominant Leber congenital amaurosis, cone-rod degeneration, and retinitis pigmentosa caused by mutant versions of the transcription factor CRX. *Human Mutation*, 18, 488–98.
- Roberts, M. R., Hendrickson, A., McGuire, C. R., & Reh, T. A. (2005). Retinoid X receptor (gamma) is necessary to establish the S-opsin gradient in cone photoreceptors of the developing mouse retina.

*Investigative Ophthalmology & Visual Science*, 46, 2897–904.

- Roberts, M. R., Srinivas, M., Forrest, D., Morreale de Escobar, G., & Reh, T. A. (2006). Making the gradient: thyroid hormone regulates cone opsin expression in the developing mouse retina. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 103, 6218–23.
- Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature Biotechnology*, 29, 24–26.
- Roger, J. E., Hiriyan, A., Gotoh, N., Hao, H., Cheng, D. F., Ratnapriya, R., Kautzmann, M. I., et al. (2014). OTX2 loss causes rod differentiation defect in CRX-associated congenital blindness. *The Journal Of Clinical Investigation*, 124, 631–643.
- Siegert, S., Cabuy, E., Scherf, B. G., Kohler, H., Panda, S., Le, Y.-Z., Fehling, H. J., et al. (2012). Transcriptional code and disease map for adult retinal cell types. *Nature Neuroscience*, 15, 487–95.
- Sieving, P. A., Chaudhry, P., Kondo, M., Provenzano, M., Wu, D., Carlson, T. J., Bush, R. a., et al. (2001). Inhibition of the visual cycle in vivo by 13-cis retinoic acid protects from light damage and provides a mechanism for night blindness in isotretinoin therapy. *Proceedings Of The National Academy Of Sciences*, 98, 1835–1840.
- Tran, N. M., & Chen, S. (2014). Mechanisms of blindness: animal models provide insight into distinct CRX-associated retinopathies. *Developmental Dynamics*, 243, 1153–66.
- Tran, N. M., Zhang, A., Zhang, X., Huecker, J. B., Hennig, A. K., & Chen, S. (2014). Mechanistically Distinct Mouse Models for CRX-Associated Retinopathy. *PLoS Genetics*, 10, e1004111.
- Vierstra, J., Rynes, E., Sandstrom, R., Zhang, M., Canfield, T., Hansen, R. S., Stehling-Sun, S., et al. (2014). Mouse regulatory DNA landscapes reveal global principles of cis-regulatory evolution. *Science*, 346, 1007–1012.
- Wenzel, A., Reme, C. E., Williams, T. P., Hafezi, F., & Grimm, C. (2001). The Rpe65 Leu450Met Variation Increases Retinal Resistance Against Light-Induced Degeneration by Slowing Rhodopsin Regeneration. *The Journal Of Neuroscience : The Official Journal Of The Society For Neuroscience*, 21, 53–58.
- White, M. A., Myers, C. A., Corbo, J. C., & Cohen, B. A. (2013). Massively parallel in vivo enhancer assay reveals that highly local features determine the cis-regulatory function of ChIP-seq peaks. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 110, 11952–11957.

## **Chapter 4**

### **Meta-analysis of epigenetic remodeling reveals CRX's mechanism of action in retinal development**

## 4.1 Author Contributions

This chapter is based on a project that will be submitted for publication prior to the oral defense: Philip A. Ruzycki<sup>1,2\*</sup>, Xiaodong Zhang<sup>2</sup> and Shiming Chen<sup>1,2,3#</sup> (*in preparation*) “*Meta-analysis of epigenetic remodeling reveals CRX’s mechanism of action in retinal development.*” Shiming Chen and I conceived of and planned the experiments, Xiaodong Zhang performed the ChIP and I performed all other experiments and analyses. Shiming Chen and I wrote the manuscript.

---

#Corresponding Author

<sup>1</sup>Molecular Genetics and Genomics graduate program, Division of Biology & Biomedical Sciences, Washington University, Saint Louis, Missouri, USA

<sup>2</sup>Department of Ophthalmology and Visual Sciences, Washington University, Saint Louis, Missouri, USA

<sup>3</sup>Department of Developmental Biology, Washington University, Saint Louis, Missouri, USA



## 4.2 Abstract

Transcription factors (TFs) acting during development often direct chromatin remodeling to allow for and promote gene expression. However, our knowledge of TFs is largely limited to their regulatory activity in plasmid-based reporter assays. To overcome this limitation, we investigated the importance of the photoreceptor TF CRX in remodeling the rod epigenome over development. By performing ATAC-seq in *WT* and *Crx*-deficient retinas, and analyzing the data alongside other genome-wide datasets, we determined that CRX is only required for remodeling or sustaining activity of <math><1/3</math> of its binding sites. These CRX-Dependent sites are defined by developmental epigenetic remodeling and the expression of nearby genes that rely on CRX. Collectively, CRX acts as a master TF to execute rod differentiation by exerting distinct and overlapping functions in chromatin rearrangement. Our study serves as a model for investigating other cell type-specific TFs and highlights the importance of analyzing TF activities at multiple levels.

## 4.3 Introduction

The retina is the highly specialized portion of the central nervous system responsible for initiating and processing visual signals before they are transmitted to the brain. The retina consists of six major classes of neurons and one of glia, although numerous subtypes of these classes have been described morphologically, functionally, and molecularly.[Baden et al., 2016; Macosko et al., 2015; Shekhar et al., 2016; Siegert et al., 2012]

Rods and cones are two types of photoreceptors responsible for the initial conversion of a photon of light into an electrical signal. Mouse retinas are rod dominant; rods constitute 80% of the retinal cells while cones comprise only 2%. [Jeon et al., 1998; Macosko et al., 2015] Retinal neurogenesis follows a stereotyped developmental program with overlapping waves of birthdates of specific cell types.[Young, 1985] In mice, rods are born over a long time window that peaks at postnatal day 0 (P0) and continues until P2.[Young, 1985] Postmitotic rod precursors undergo differentiation over an extended 2-week period, during which the cells establish a rod-specific gene expression profile, develop unique subcellular structures (OS and presynaptic terminals), and eventually can perform phototransduction.

Precisely regulated gene expression is essential for rod structural/functional development and survival, as even subtle perturbations can result in blinding diseases.[Lem et al., 1999; Ruzycski et al., 2015] Rod gene expression is tightly regulated by a number of transcription factors (TFs), acting in a cascade during development. [Reviewed in Swaroop, et al.[Swaroop et al., 2010]] The homeodomain (HD) TF OTX2 specifies the photoreceptor lineage by turning on the expression of Cone Rod homeobox (CRX) and its downstream TFs. CRX is an OTX-like HD TF, whose expression coincides with the final mitotic event in rod and cone photoreceptors and

is maintained into adulthood.[S. Chen et al., 1997; Furukawa et al., 1997] CRX binds to the promoter of rod/cone genes and activates their expression via its transactivation domain.[S. Chen et al., 1997, 2002] Two rod-specific TFs, NRL and NR2E3, act together with CRX to direct rod differentiation by activating rod and silencing cone genes.[J. Chen et al., 2005; Mears et al., 2001; Peng et al., 2005; Swaroop et al., 2010] General TFs involved in chromatin remodeling including MEF2D, CBP/P300, and the STAGA complex are also a part of the CRX regulatory network.[Andzelm et al., 2015; Hennig et al., 2013; La Spada et al., 2001] Together, these factors direct the proper establishment of the rod epigenome and transcriptome.

CRX is essential for photoreceptor differentiation and functional development. A *Crx* null mutation (*Crx*<sup>-/-</sup>) produces a recessive phenotype in the mouse retina where the immature photoreceptor cells fail to differentiate and begin to degenerate at ~4 weeks of age.[Furukawa et al., 1999] Mutations in human CRX have been associated with dominant blinding retinopathies with varying severity and etiology [Reviewed in [Tran & Chen, 2014]]. Interestingly, in the corresponding mouse models, distinct *Crx* mutations all affect the expression of a common set of genes, but the degree of dysregulation correlates with phenotype severity.[Ruzycki et al., 2015] The mechanism for this misregulation remains to be determined, but a better understanding of CRX's mechanism of action during photoreceptor development would provide insights into the pathogenicity of CRX mutations.

Toward this goal, a number of studies have investigated the activity of CRX using plasmid-based assays. Luciferase reporter assays in heterologous systems have been employed to map the activation domain, demonstrate the synergy between CRX and NRL, and determine the functional effect of CRX mutations.[S. Chen et al., 2002; Roger et al., 2014; Ruzycki et al., 2017; Tran et al., 2014] Plasmid based systems have been adapted to report the activity of

enhancers *in vivo* using GFP,[Corbo et al., 2010; Hsiao et al., 2007; Montana et al., 2011] but these experiments are limited in scope as each construct must be tested individually. Recent technological advances have overcome these throughput limitations. Massively Parallel Reporter Assays (MPRAs) performed in the retina confirmed that hundreds of CRX bound regulatory elements positively regulate transcription, while unbound regions did not.[White et al., 2013] Additional experiments suggested that motif affinity directly contributed to this regulatory potential.[White et al., 2016] However, a major limitation of all plasmid-based systems is the lack of the genomic chromatin context. Studies have shown that sites known not to be bound in the genome, when cloned in a plasmid reporter system, are bound by the TF simply based on the raw sequence and motif.[Grossman et al., 2017] This suggests that chromatin plays a major role in the specificity of TF binding. Additionally, plasmid-based systems only measure steady state activity of the complement of TFs, but are unable to report developmental epigenomic regulatory potential of key TFs.

These concepts have been studied separately. Chromatin immunoprecipitation with high-throughput sequencing (ChIP-seq) for CRX was performed in the adult retina.[Corbo et al., 2010] Binding sites were enriched for a HD binding motif, and nearby genes represented many known photoreceptor specific elements. However, use of this data alone assumes TF binding is a binary control. Previous analysis of gene expression shows that many genes that lose expression in *Crx* mutant mice are in close proximity to CRX binding sites.[Tran et al., 2014] However, while this result displayed a very significant enrichment, it failed to explain why many more genes that are also ‘bound’ by CRX are not affected.

Other experiments have sought to understand normal epigenetic development of rods and cones.[Aldiri et al., 2017; Hughes et al., 2017; Mo et al., 2016; Yue et al., 2014] Dramatic

developmental remodeling was described by ChIP-seq of histones and regulatory proteins and by profiling of DNA methylation.[Aldiri et al., 2017] Other studies have profiled DNA accessibility through development or compared mature rods vs cones.[Hughes et al., 2017; Mo et al., 2016; Yue et al., 2014] These studies matched regulatory sites to CRX ChIP-seq data and predicted cofactors based on motif enrichment. However, the adult profiling of normal cells limited the interpretation to a static prediction and was unable to directly attribute the activation of any site to a particular factor.

To bridge this gap in the understanding of TF binding vs regulatory potential, here we use ATAC-seq to show that CRX is only responsible for the epigenomic rearrangement of a subset of its binding sites. These ‘Dependent’ sites reside within a variety of chromatin environments and are highly correlated with the genes affected in mutant retinas. By applying this technique, we clarify the role of this important TF in retinal development and disease and provide a model for future studies of other TFs essential in the development of other cell types.

## 4.4 Results

### 4.4.1 CRX binds ATAC sensitive regulatory sites

To determine active regulatory regions in the genome of mouse rod photoreceptors, we performed whole retina ATAC-seq,[Buenrostro et al., 2015] a technique that profiles open chromatin regions. We chose to profile at postnatal day 14 (P14), because all retinal cell types are born by this age [Young, 1985] and photoreceptor specification is completed. Replicate experiments on whole retina of *C57BL/6J* (WT) mice, where rods comprise ~80% of the cells,[Jeon et al., 1998; Macosko et al., 2015] were consistent with available DNase1 data profiling of adult (8wk) whole retina (Figure 4.1a), where areas enriched for ATAC data also showed enrichment of signal by ENCODE DNase1 hypersensitivity datasets.[Yue et al., 2014]

We overlapped these regulatory sites (ATAC-seq peaks) with those sites bound by CRX (detected by ChIP-seq).[Corbo et al., 2010] The majority (>98%) of CRX binding sites were contained within ATAC-sensitive genomic regions, while many ATAC-sensitive sites showed no CRX enrichment (Figure 4.1a, b). This comparison both supports the specificity of the CRX ChIP and the unique role of CRX in binding a subset of regulatory sites to control gene expression.

To understand the role of these CRX binding sites in the regulation of gene expression, we first analyzed their distribution around the transcription start site (TSS) of each gene expressed in the retina, in the context of that gene's dysregulation in *Crx*<sup>-/-</sup> cells (Figure 4.1c, left panel). While all ATAC-sensitive regulatory sites showed a strong preference to be located near the TSS, there was no preference for genes that were determined to be up- or down-regulated in the *Crx*<sup>-/-</sup> retina (top or bottom of plot as genes ordered by fold-change in *WT/Crx*<sup>-/-</sup>

; Figure 4.1c, right panel). We also plotted in the same manner the subset of ATAC regions bound by CRX (Figure 4.1d). While these sites showed a similar distribution near the TSS of expressed genes, they displayed a preference for genes that lose expression in the *Crx*<sup>-/-</sup> retina. However, to our surprise, many CRX bound sites were near genes that had no transcriptional changes. This data suggested that while CRX is a strong transcriptional activator,[S. Chen et al., 1997] not every binding site has the same regulatory potential or dependency upon CRX activity.

#### **4.4.2 *Crx*<sup>-/-</sup> retinas have an altered photoreceptor epigenome**

To determine the functional implications of the loss CRX on the epigenome of photoreceptors, we also performed duplicate ATAC-seq experiments on P14 *Crx*<sup>-/-</sup> retinas and compared the results to those of *WT* retinas. Again, *Crx*<sup>-/-</sup> replicates were highly consistent with one another (Figure 4.1a). Although many of the ATAC peaks qualitatively resembled *WT* signal, some (e.g. those spanning *Gnat1*) appeared much weaker in *Crx*<sup>-/-</sup> than *WT* (Figure 4.1a). Since no photoreceptor degeneration or cell death is detected in *Crx*<sup>-/-</sup> at this age,[Furukawa et al., 1999; Tran et al., 2014] differential ATAC signals between the two mouse lines are indicative of changes in the photoreceptor epigenome. Indeed, by quantitative comparison of *WT* and *Crx*<sup>-/-</sup> datasets, *Crx*<sup>-/-</sup> retinas display a widely disturbed epigenome with both increased (“up”) or decreased (“down”) activity at potential regulatory sites (Figure 4.2a). Roughly 25% of ATAC peaks were altered, with virtually equivalent proportions “up” as those that were “down” (Figure 4.2b).

### 4.4.3 *Crx*<sup>-/-</sup> photoreceptors fail to close and open developmentally-modulated regulatory sites

Previous DNase1 hypersensitivity data analyzed three stages of retinal development: P1, P7 and Adult.[Yue et al., 2014] We can generalize these samples based on the over-representation of rods in the mouse retina to represent early progenitor, immature, and mature rod photoreceptors, respectively. At a gross level, comparison to these datasets delineated distinct patterns at our classified genomic sites (Figure 4.2c-e). Sites that were not changed between *WT* and *Crx*<sup>-/-</sup> were generally open at all three stages (Figure 4.2d). However, sites that are affected display contrasting dynamics over the course of development: The set that loses activity in the *Crx*<sup>-/-</sup> retina normally would be activated, while the set that gains activity in the mutant retina would be closed over normal development (Figure 4.2c vs 2e).

These sites also show differential tissue specificity. When compared to other DNase1 data, unaffected sites are also generally active in the brain and liver (Figure 4.2d). Again, affected sites show contrasting patterns, where those lost in *Crx*<sup>-/-</sup> retinas are largely retina specific (Figure 4.2e) while those that gain accessibility in mutant retinas are also very active in the brain (Figure 4.2c). This data suggests that *Crx*<sup>-/-</sup> retinas maintain a majority of their basic epigenetic state that would be similar in all cell types. However, they fail to turn on a highly photoreceptor specific set of sites and to inactivate many sites that are used generally in unspecified neurons.

We next tested what types of genes were likely regulated by these sites using GREAT Gene Ontology (GO) analysis (Supplemental Table 4.1).[McLean et al., 2010] “Down” sites are highly enriched for a number of categories of genes involved in “phototransduction,” “detection of light,” and microtubule related processes. “Up” sites were enriched for categories that



involved general development or differentiation of oligodendrocytes, somatic motor neurons, and peripheral nervous system. There was also an enrichment for a single retinal related category ‘detection of visible light.’ Further investigation discovered the enriched genes to primarily be cone related (*Cnga*, *Cngb*, and *Gnat2*, among others), which are present in developing “immature” rods or transfated S-cones in the absence of NRL.[Brooks et al., 2011; Mears et al., 2001]

#### **4.4.4 CRX binding is strongly enriched at sites that lose activity in *Crx*<sup>-/-</sup>**

To understand the primary role of CRX in this dynamic chromatin environment, we classified ATAC sites based on the presence of an overlapping CRX ChIP signal (Figure 4.2f, g). Nearly half of the ATAC peaks that are reduced in the *Crx*<sup>-/-</sup> retina are bound by CRX [CRX Dependent] compared to a very small fraction (<1%) of those peaks that show increased signal (Figure 4.2f, g and 4.2c vs e). There was also a significant enrichment of CRX binding within ATAC sites that were unchanged in the *Crx*<sup>-/-</sup> retina [CRX Independent], although to a lesser degree than the “down” group, (Figure 4.2d, f, g; Table 4.1).

We next sought to determine the relationship between changes in ATAC and local gene expression. CRX bound ATAC sites that are decreased in *Crx*<sup>-/-</sup> compared to *WT* show a very strong preference for the TSS of genes that lose expression (Supplemental Figure 4.1f vs d). A similar trend was also observed for non-CRX bound ATAC sites that lose signal in *Crx*<sup>-/-</sup> retina (Supplemental Figure 4.1e). Conversely, ATAC sites that are increased in the *Crx*<sup>-/-</sup>, nearly all of which lack CRX binding, are greatly enriched near genes that increase in expression in the *Crx*<sup>-/-</sup> retina (Supplemental Figure 4.1a, b). Both CRX bound and unbound sites that are unchanged in the mutant retina show no strong association with any changes in gene expression

(Supplemental Figure 4.1c, d). Together, these data show the strong correlation between activity of local regulatory sites and gene expression, but indicate that only a subset of CRX bound sites are affected upon the loss of CRX.

#### **4.4.5 CRX acts within a variety of chromatin environments**

We decided to explore the differences between these CRX binding sites that were either lost [CRX Dependent] or maintained [CRX Independent] in the *Crx*<sup>-/-</sup> retina. Previous studies have classified TF binding sites into promoter or enhancer classes based on distance from the TSS. This approach is much more straightforward, but with new studies showing the breadth of local regulatory domains, [Benayoun et al., 2014; Whyte et al., 2013] it likely does not reflect the true biology or regulatory role of each site. Instead, we decided to distinguish binding sites based on the local epigenetic state. CRX binding sites were located in four distinct chromatin environments defined by ChIP-seq of H3K4me3 and H3K27Ac at P14 in *WT* retinas (Figure 4.3a, b). Groups A and C represent the largest proportion of the CRX binding sites and are located in local regulatory domains and distal enhancers, respectively (Fig 3, Supplemental Figure 4.2a). Group D represents CRX sites where neither active histone mark is present, these are located similar to distal enhancer Group C, far from annotated genes (Supplemental Figure 4.2a). Group B includes a very small number of CRX sites that are near genes expressed at a much lower level than the three other Groups (Supplemental Figure 4.2a, b). We found a relatively equal representation of Dependent and Independent CRX sites within the three main groups (Figure 4.3c), although Group A displayed some overrepresentation of Independent sites.

#### 4.4.6 CRX is required to activate a subset of local regulatory regions

To investigate the differential influence of CRX at local regulatory regions (Figure 4.3; Group A), we first analyzed the temporal changes in the activation state as profiled by DNase1. CRX Dependent local sites show strong retinal specificity and developmental activation (Figure 4.4a (left panel), Supplemental Figure 4.3a). This pattern is in stark contrast to Independent sites that display strong signal also in brain and liver samples and retinal accessibility at all ages (Figure 4.4a (right panel), Supplemental Figure 4.3b).

We next sought to understand whether these ATAC signals correspond to changes in other epigenetic marks. To understand the epigenetic state of these sites over normal retinal development, we analyzed data recently published by Aldiri et al.[Aldiri et al., 2017] By both the analysis of individual ChIP datasets (data not shown) and the aggregate of this data in the HMM classification of the binding sites (Supplemental Figure 4.4a, b), Independent sites were nearly all classified as HMM classes 1-4 at E14.5 while Dependent sites displayed temporal remodeling up to P7.

To determine whether the loss of CRX affected the local histone modifications at Dependent sites, we performed ChIP-seq for H3K4me3 in *WT* and *Crx*<sup>-/-</sup> retinas. Peaks containing CRX Independent sites largely maintained H3K4me3 presence in mutant retinas, while peaks containing CRX Dependent sites display a significant loss of signal in *Crx*<sup>-/-</sup> (Supplemental Figure 4.3c, d).

We sought to determine the effects of these epigenetic changes on gene expression. RNAseq data profiling *WT* development and changes in *Crx*<sup>-/-</sup> was consistent with ATAC and DNase1 data. Genes near CRX independent sites show very little developmental expression changes and are relatively unaffected in mutant retinas. In contrast, those near Dependent sites

increase dramatically over development but largely fail to do so in the *Crx*<sup>-/-</sup> (Figure 4.4b). Gene Ontology (GO) analysis of these genes suggests Dependent genes are largely photoreceptor related, while the Independent set is comprised of some genes with photoreceptor function, but also genes with more general functions (protein transport, RNA processing, protein localization, etc; Supplemental Table 4.1).

#### **4.4.7 CRX is required to activate a set of distal enhancers**

CRX also binds sites marked by H3K27Ac (Figure 4.3; Group C), traditionally thought to be active distal enhancers. Dependent and Independent sets of enhancers both displayed dynamic changes over development and a high degree of retinal specificity (Figure 4.5a and Supplemental Figure 4.5a, b). However, by more careful comparison of the temporal dynamics of this activation, we observed that Independent sites showed earlier activation than Dependent sites (Supplemental Figure 4.5a vs b). We again analyzed other epigenomic data to determine if they also would support this subtle temporal difference. HMM classification of the sites was nearly identical by P14, but comparison of P0 and P3 data showed differences where Independent sites were switching to more active classifications earlier than Dependent (Supplemental Figure 4.4c, d).

To test whether ATAC signal changes correlated with the mutant epigenome, we performed ChIP-seq for H3K27Ac at P14 in *WT* and *Crx*<sup>-/-</sup> retinas. As expected, Dependent sites displayed a significant decrease in H3K27Ac deposition in *Crx*<sup>-/-</sup> retinas while Independent sites largely maintained this mark (Supplemental Figure 4.5c, d).

Expression of the nearest gene to each enhancer also highlights differences between CRX binding sites. Genes near both Dependent and Independent sites increased in expression over

development in *WT* retinas (Figure 4.5b). However, this increase is attenuated for the Dependent set, while there was no decrease observed in the Independent set (Figure 4.5b). GO analysis showed a very clear distinction where Dependent genes are photoreceptor related, while Independent genes have a variety of general functions (Supplemental Table 4.1).

#### **4.4.8 CRX controls distal regulatory sites in the absence of active histone marks**

A significant number of CRX bound sites were not marked by either H3K4me3 or H3K27Ac (Figure 4.3; Group D). When we analyzed the DNaseI and HMM data at these sites, the pattern looked similar to Group C enhancers where CRX dependent sites opened later than Independent and were more retina specific (Supplemental Figure 4.6a, d, e & Supplemental Figure 4.4e, f). One notable difference was that HMM classification highlighted a subset of Independent sites classified throughout development as Class 11 [Insulator] (Supplemental Figure 4.4e).

Interestingly, gene expression data highlighted a different scenario than other groups. The expression of the nearest gene to CRX Dependent sites displayed no significant differences (Supplemental Figure 4.6b). While differences in gene expression were statistically significant in the Independent set, the degree of change was very modest.

#### **4.4.9 Base composition and conservation differentiate CRX sites**

Our epigenome analyses identified distinct types of CRX sites, but raised the question whether there are sequence characteristics ‘coding’ for these differences. As base composition, such as CpG density, can determine the general regulatory role of genomic locations, we

analyzed nucleotide frequency across CRX regulatory sites (Figure 4.6a). This analysis uncovered several very interesting differences. First, Group A Dependent sites show a fairly equal nucleotide distribution, especially within the central 200 bp of the CRX bound regulatory site. This is in dramatic comparison to the Independent Group A sites that are within highly GC rich regions. Group C and D distal regulatory elements did not show any obvious differences between Independent or Dependent sites. Both have balanced content within the regulatory site, but lie within AT enriched domains.

We also analyzed the extent to which the sequences were conserved across vertebrate evolution (Figure 4.6b). All groups display a central region that is more conserved than the surrounding environment (Figure 4.6b), supporting the importance of the ATAC and CRX ChIP specified regulatory sites. Surprisingly, Independent sites of all three classes showed significantly higher conservation than their Dependent counterpart. We also noted that Group D Dependent sites and surrounding genomic regions were less conserved than Dependent Group A or C regions.

#### **4.4.10 CRX has different affinity for Dependent vs Independent sites**

We next sought to determine whether the predicted binding affinity of CRX could distinguish sites. In Group A proximal regulatory regions, significantly more Dependent than Independent sites contained the canonical CRX site (Supplemental Figure 4.7a). This difference was reflected in re-analysis of the CRX ChIP-seq, where quantification shows a higher read depth at Dependent sites (Supplemental Figure 4.7b). Group C and D sites displayed the opposite characteristics, where Independent sites were more likely to contain the CRX motif

(Supplemental Figure 4.7a), although there was no difference in the intensity of the original ChIP data (Supplemental Figure 4.7c, d).

We also predicted CRX occupancy using an algorithm that takes into account the relative ratio of protein to DNA.[White et al., 2016; Zhao et al., 2009] At both low and high ratios ( $\mu$  values), Group A Dependent sites displayed higher binding occupancy than Independent sites (Figure 4.6c, d). However, when Groups C and D were analyzed at a low protein:DNA ratio ( $\mu=0.5$ ), Independent sites have a significantly higher occupancy score, consistent with the traditional motif search data (Figure 4.6c). When analyzed at a high protein:DNA ratio, this relationship inverted and Dependent sites instead showed higher predicted occupancy (Figure 4.6d). This suggests that Dependent distal regulatory sites have more low affinity CRX binding sites that could additively contribute to CRX recruitment and perhaps inform functional differences.

#### **4.4.11 Other TFs may compensate or be more influential at Independent sites**

We also wondered whether binding sites for other TFs could differentiate CRX Dependent vs Independent sites. We analyzed  $\pm 100$ bp from the center of each CRX peak for enrichment of sequences *de novo* (Table S2) and motifs included in the JASPAR database (Supplemental Figure 4.8a). Both methods reported the highest enriched motif in all Groups to be a canonical homeobox (Supplemental Figure 4.8a marked by \*, Supplemental Table 4.2).

The most striking difference we noted by both methods was a unique enrichment of many TF motifs in Group A Independent sites. Known TFs included KLF4, ELK1/4, E2F1, and NFY $\alpha$

(Supplemental Figure 4.8a, marked by \*\*) while *de novo* motifs included promoter elements (GC-box), NFY, NRF1, BHLH among others (Supplemental Table 4.2).

The other five sets shared enrichment of many known factors including GATA, LHX, and NKX family motifs (Supplemental Figure 4.8a), although there were also notable differences. Both methods suggested an enrichment of CTCF specifically in Group D Independent sites (Supplemental Figure 4.8a, marked by \*\*\*; Supplemental Table 4.2), and known motif analysis suggested that ESRRB and FOXC1 are enriched only in Dependent sites (Supplemental Figure 4.8a, marked by \*\*\*\*). *De novo* methods also indicated strong unique enrichment of NEUROD1, MEF2D, and MEIS1 motifs in Group C Independent sites (Supplemental Table 4.2).

To confirm and explore the relevance of the motif analysis, we chose to focus on the enrichment of MEF2D and CTCF sites. The CTCF motif was enriched only in Group D Independent sites by both methods. As previously discussed, this set contains many sites with Insulator function (Supplemental Figure 4.4e, HMM Class 11 purple). We reanalyzed CTCF ChIP-seq data and confirmed that CTCF is enriched at Group D Independent sites throughout development, and this binding was reasonably specific with minimal enrichment at Independent Group A sites but none at any Dependent sites (Supplemental Figure 4.8c, d).

We also explored the enrichment of the MEF2D motif in Group C Independent sites (Supplemental Table 4.2). Analysis of previously published ChIP-seq data [Andzelm et al., 2015] showed that MEF2D binds both Dependent and Independent enhancers in *WT* retina, but this signal is dramatically reduced only at Dependent sites in the *Crx*<sup>-/-</sup> retina (Supplemental Figure 4.8b). Together, these data support the relevance and interpretation of the motif enrichment analyses.



## 4.5 Discussion

By comparing the genome-wide changes of activity in *WT* vs *Crx*<sup>-/-</sup> retinas, we have determined that the loss of CRX causes significant changes in activity at ~25% of all regulatory sites. Even though CRX is an activating TF, sites that increase were as prevalent as those that lose activity (Figure 4.2). By analyzing these sites alongside time course DNaseI hypersensitivity data, we have confirmed previously hypothesized mechanisms of *CRX*-associated disease, that the rod photoreceptors are stuck in a pre-developmental state; sites that were supposed to close remain active, and sites that were supposed to progressively open did not do so (Figure 4.2e-g). This model of *CRX* disease explains previous findings that *Crx* mutants over express ‘cone’ related transcripts.[Ruzycki et al., 2015] Together with this data, recent reports that suggest a more common developmental path of rods and cones than previously recognized[Kim et al., 2016] support the model that *Crx* mutant retinas resemble a premature state prior to the dramatic epigenetic switch toward rod photoreceptor identity.

Many of the peaks that lose activity are normally bound by CRX itself, consistent with these sites being Dependent on its activity. CRX showed virtually no binding at peaks with increased activity, but surprisingly did bind many peaks that were unchanged, or Independent (Figure 4.2c, d). These CRX Dependent and Independent sites were located throughout the genome. Our method that incorporated other aspects of the local epigenetic neighborhood, including two active histones, clearly establishes distinct types of CRX binding sites beyond simply TSS proximal and TSS distal. We propose this classification system is more informative and biologically relevant to the function of individual TFs.

CRX is activated in early born photoreceptors and maintains expression in mature cells. Both developmental DNase1 and HMM classifications showed that all CRX Dependent sites normally undergo dramatic activation over retinal development. By ChIP-seq, we showed that H3K4me3 and H3K27Ac marks were lost at Group A and C Dependent sites respectively, and by RNAseq that nearby genes were not properly activated over development in *Crx*<sup>-/-</sup> retina. Together, this suggests that CRX is involved in this activation likely through its direct binding and recruitment of other general TFs and co-regulators. Previous work has already shown clear roles of such co-factors including CBP/P300, Ataxin7-STAGA, and DNMT1 among others in retinal development.[Hennig et al., 2013; Nasonkin et al., 2013; Palhan et al., 2005; La Spada et al., 2001]

Independent sites showed two very different patterns when data was analyzed (Figure 4.7). Group A (TSS proximal) Independent sites were open constitutively. These regions showed no evidence of developmental remodeling, and were also active in other tissues. This suggests that these sites are generally used regulatory elements, and GO analysis of the nearby genes supports this interpretation as these genes were categorized as having very general cellular functions. Conversely, Independent Group C and D sites did show temporal activation, but quantitative analysis of the DNase1 and HMM data suggests they begin remodeling prior to CRX expression and their respective Dependent group sites. The ability of the photoreceptor to maintain or continue activating these Independent sites and properly express nearby genes suggest that the binding of CRX here is not essential. However, we cannot formally exclude the possibility that other homeodomain TFs (such as OTX2) compensate in the *Crx*<sup>-/-</sup> retina.

We analyzed the affinity of CRX by two distinct methods. Both predict that CRX binds more efficiently to Dependent than Independent Group A sites, which was supported by analysis

of ChIP-seq data (Supplemental Figure 4.7). We were surprised to find that Dependent distal elements (Groups C and D) display less high but more low-quality binding sites. Differences in activity relevant to binding affinity have already been predicted for CRX.[White et al., 2016] This difference could inform a fundamental mechanism for CRX activity at enhancers, but alternative biochemical experiments would be necessary to study binding efficiency, on/off rates, and functional effects of these different types of sequences.

*De novo* motif analysis did not reveal any differences in the homeodomain site between Dependent and Independent sites. This is not unexpected, as this motif is shared among many homeodomain TFs expressed throughout the body, including others expressed in the retina (OTX2 in particular). These results together suggest that CRX binding is directed by the presence of a strong consensus HD motif, but this association cannot blindly be interpreted as an important functional interaction (Figure 4.7). At proximal Independent regulatory sites, CRX binding may simply be the byproduct of accessibility of the HD element used either by another retinal HD TF, or more important in other cell types. Similarly, at distal Independent elements the HD motif may be required only for activation in other neuronal cell types (Group C) or may be recognized by another HD factor in the retina itself; only the precise mutation of these elements in the genomic context could answer this question.

Previous work had noted that CRX bound within many GC rich genomic regions.[Corbo et al., 2010] Our analysis discovered that these constitute proximal Independent regulatory sites. While GC rich areas are normally thought to represent repressed CpG islands, other reports have noted that ubiquitously expressed promoters display high GC content while cell type specific promoters are AT rich.[Landolin et al., 2010] We expected that CRX Dependent sites would be very highly conserved across species. While it was the case that the regulatory elements were

more conserved than flanking sequences, Independent sites were more conserved than Dependent counterparts. *De novo* and known motif analyses discovered that Independent sites, especially Group A promoters, were enriched for a variety of TF motifs. Together, these data suggest that Independent sites are active in other cell types of the body and are regulated by shared or more general TFs.

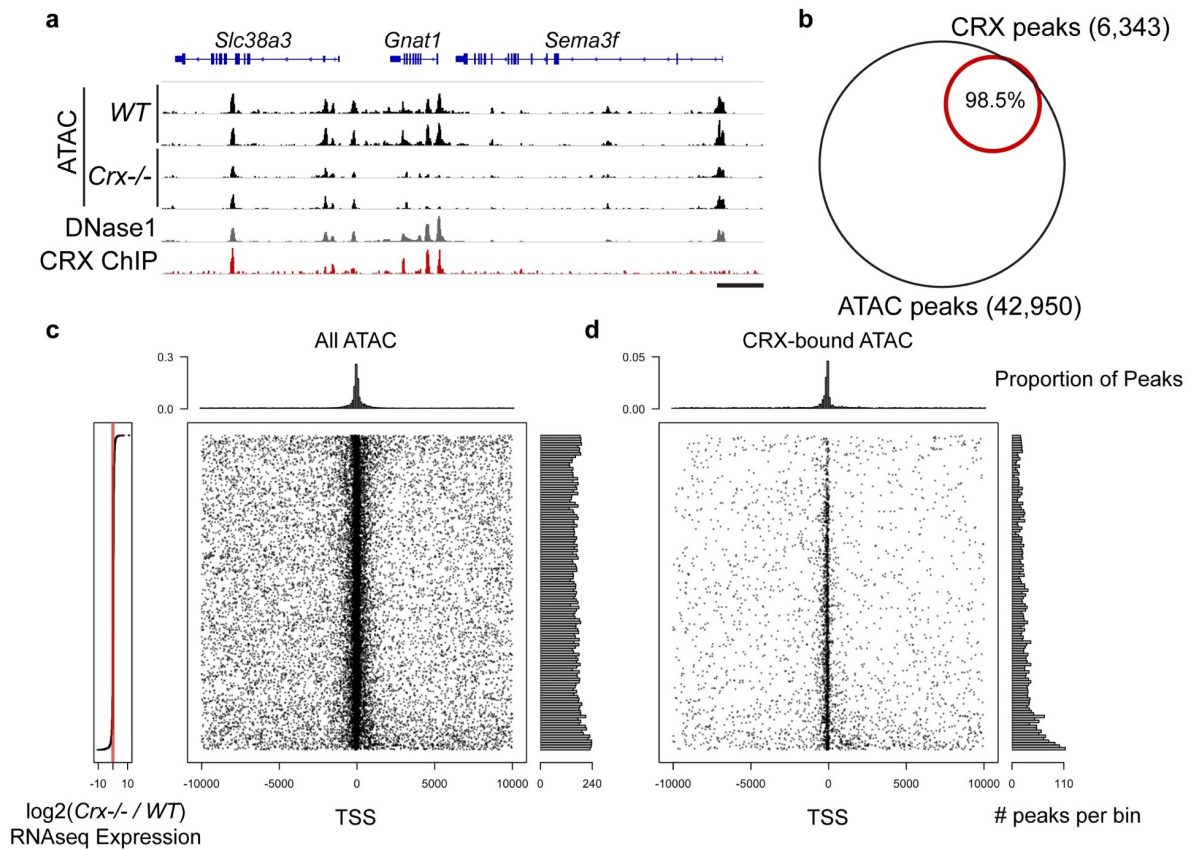
Analysis of Dependent sites did suggest several CRX interacting partners, although motifs for known partners NRL and NR2E3 that synergize or bind with CRX [S. Chen et al., 2002; Peng & Chen, 2005] were not observed. This supports the model that while all three factors are necessary for proper gene expression, CRX acts as the primary sequence specific targeting factor for the three as a complex. Group C sites in particular displayed enrichment for a number of other factors including NKX, ESRRB, and FOXC. These could represent new TFs that either act in coordination with CRX to activate these enhancers, or could represent pioneer factors that opened the sites to allow for CRX to bind and fully activate.

Our analyses have emphasized that every TF binding site is not equal. While CRX or any other homeodomain TF may bind thousands of sites, only a subset of those sites (for CRX, <1/3) require the TF to establish or maintain the local epigenetic state. We propose these elements and the homeodomain motifs within Dependent sites are excellent candidates where non-coding variants may cause human retinal disease. Our findings are based on an embryonic loss-of-function study. In this setting, CRX acts as an “acceleratory factor” required for enhancing Dependent site activity. However, these experiments do not address if CRX is sufficient for achieving the active chromatin state and when CRX’s epigenomic activity is required (plasticity of the system). Ectopic expression of CRX in cultured HEK293 and Y79 retinoblastoma cells failed to produce a rod-like epigenome, even with co-expression of NRL (Ruzycki and Chen,

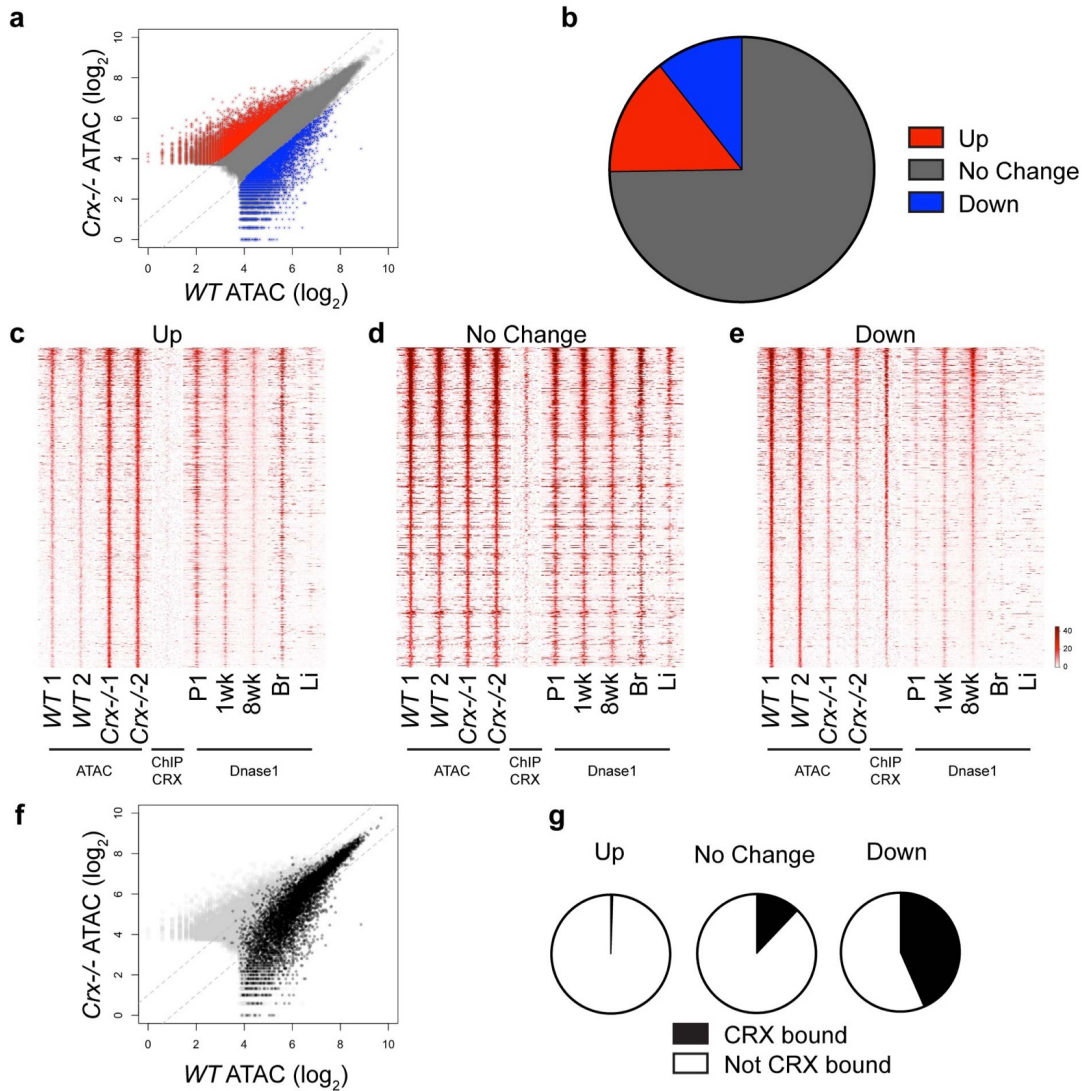
unpublished results), suggesting CRX is not a “pioneer” factor that can bind to fully closed sites to mediate *de novo* chromatin remodeling for cell fate specification. Instead, CRX appears only able to act on those sites that are “primed” for photoreceptor rearrangement in the precursor cells. Indeed, even CRX Dependent sites do show some level of ATAC sensitivity, perhaps evidence for prior opening by another factor, although we cannot exclude that this is the result of compensation. Future studies, such as temporal knockout or ectopic expression of CRX in developing or mature photoreceptors are needed to address the sufficiency and plasticity questions, important for understanding and treating CRX-linked diseases.

**Table 4.1: ATAC-seq peaks lost in *Crx*<sup>-/-</sup> show enrichment for CRX binding in *WT***

	Total	CRX bound	%
Up	6,259	33	0.5
No Change	32,105	3,883	12.1
Down	4,586	1,992	43.4

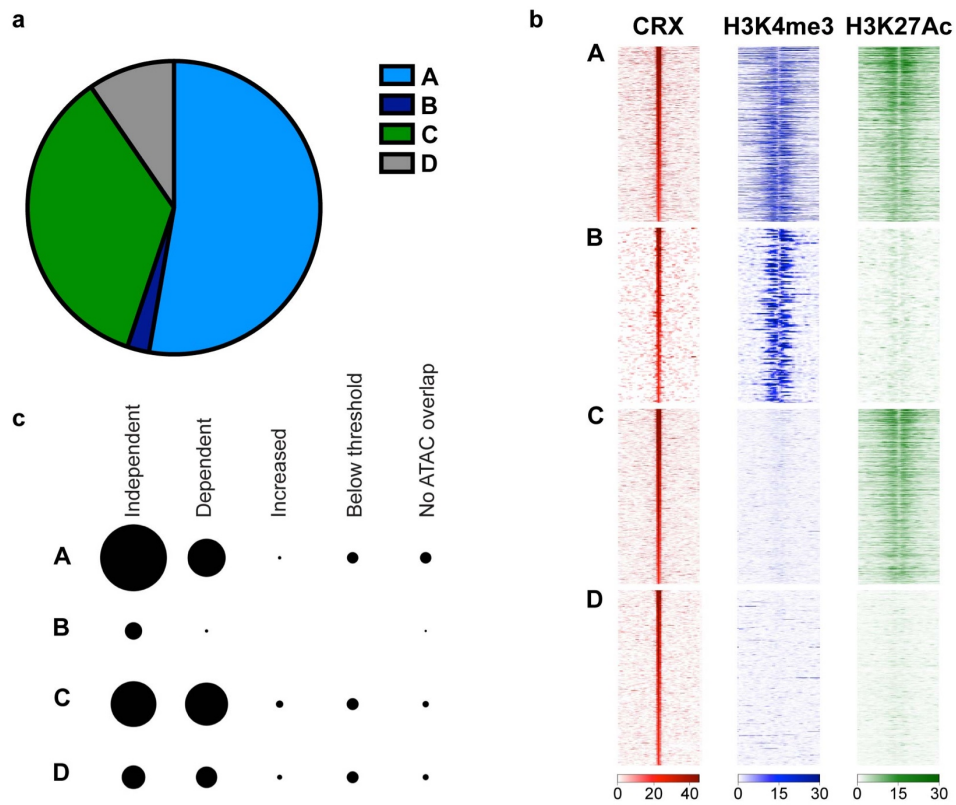


**Figure 4.1: CRX binds a subset of active regulatory sites in the rod photoreceptor. (a)** Browser track displays ATAC-seq, DNase1 and CRX ChIP-seq read depth. (Scale bar 5kb) **(b)** Venn diagram depicting number of CRX ChIP-seq-defined binding sites that overlap with regulatory sites defined by ATAC-seq. **(c, d)** Meta-gene plots of all genes expressed in P21 *WT* and *Crx*<sup>-/-</sup> retinas, ordered by [ $\log_2$ ] fold-change (as depicted in plot on left). Black dots represent the center of ATAC regulatory site relative to TSS of all ATAC peaks. **(c)** and of only the subset bound by CRX. **(d)** Histograms of X and Y axes display density and distribution of the data.

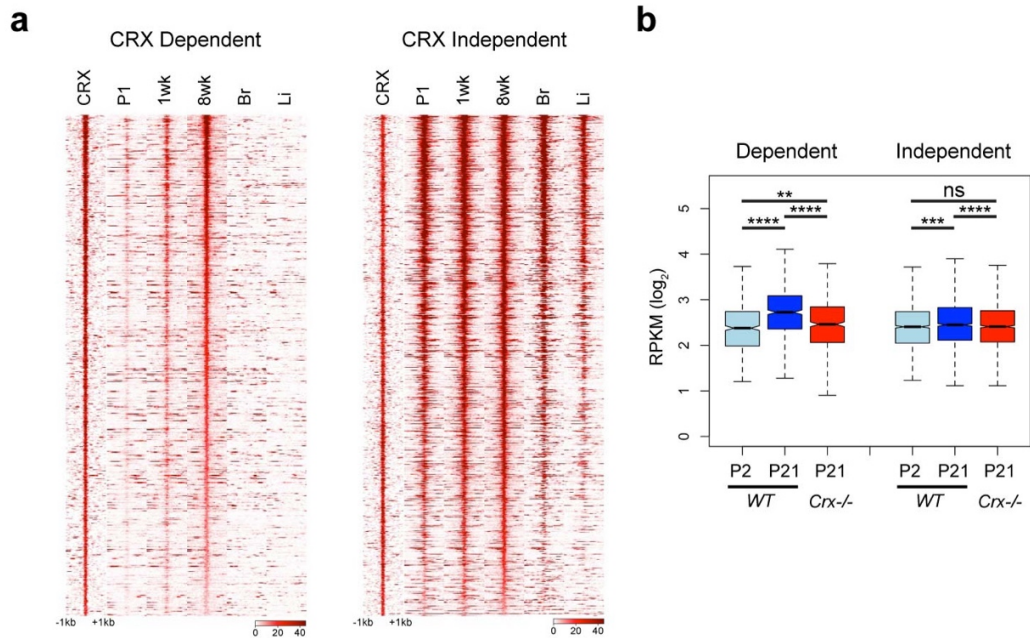


**Figure 4.2: Loss of CRX affects many developmentally activated regulatory sites. (a, b)** Comparison of *WT* and *Crx*<sup>-/-</sup> ATACseq data shows highly disturbed epigenome in *Crx*<sup>-/-</sup> retinas where significant numbers of sites display increased (red) or decreased (blue) signal. **(c-e)** Read density heatmaps display reproducibility of ATAC experiments (*WT*1 & 2 vs *Crx*<sup>-/-</sup> 1 & 2), overlap of CRX binding signal, and regulatory activity as defined by DNase1 in P1, 1wk, and 8wk retina, Brain, and Liver samples. **(f)** Scatterplot displays distribution of the subset of CRX bound ATAC sites over the comparison of *WT* vs *Crx*<sup>-/-</sup> data. **(g)** Proportion of total sites within each ATAC class bound by CRX (black) are displayed in pie charts.

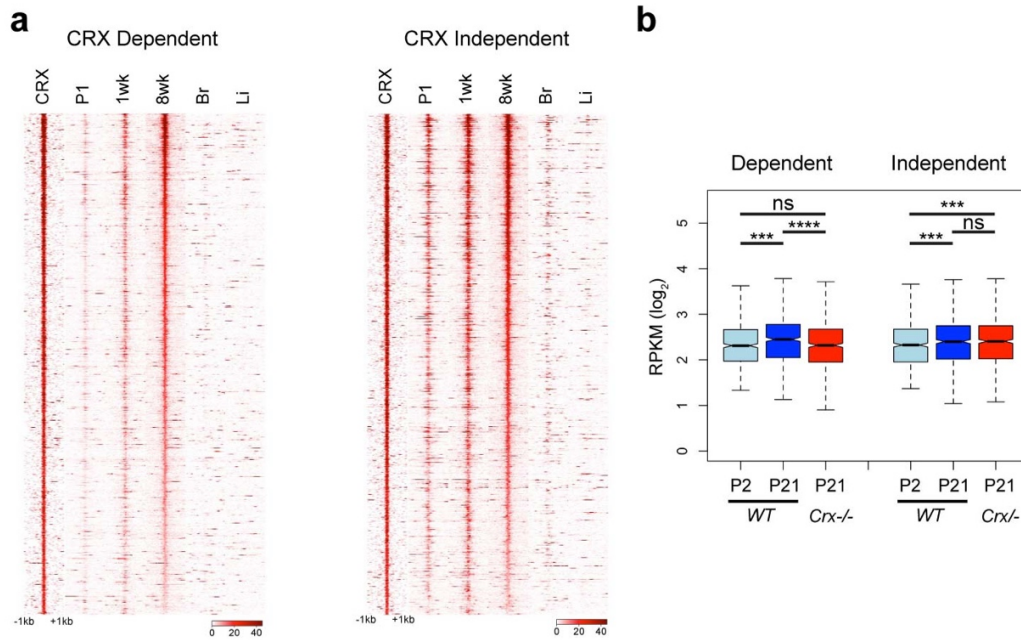




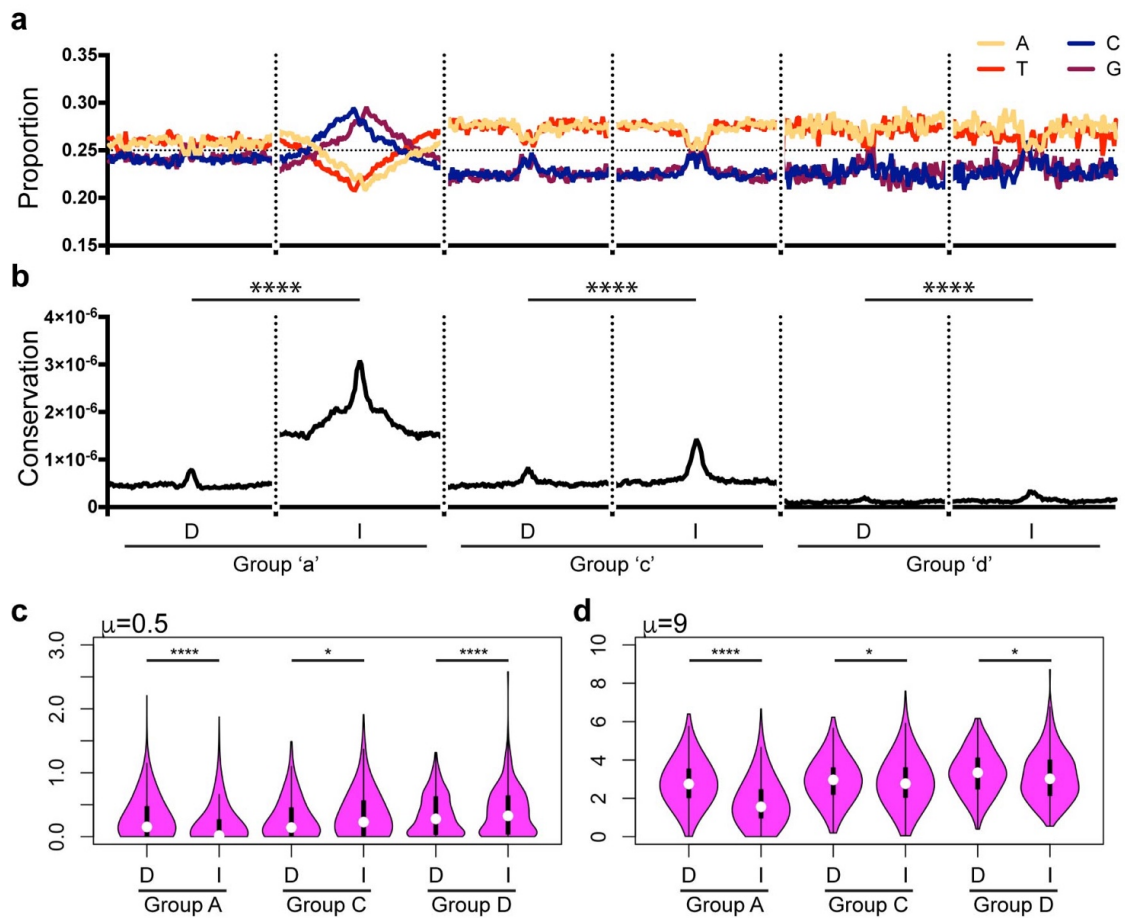
**Figure 4.3: CRX binds within four chromatin environments.** (a) Pie chart depicts number of CRX binding sites that reside within Groups A-D as defined by overlap with (b) ChIP-seq data of H3K4me3 and H3K27Ac datasets in P14 *WT* retina. (c) Size of circle represents number of CRX sites within each group A-‘D,’ and displays distribution relative to ATAC-seq comparison between *WT* and *Crx*<sup>-/-</sup>.



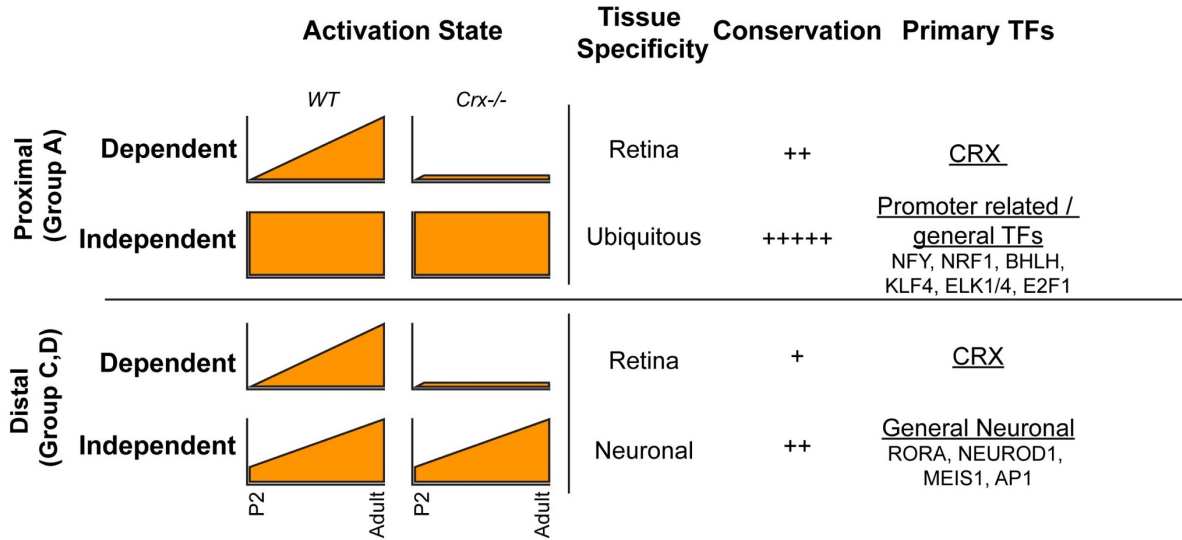
**Figure 4.4: CRX is only required for activity and remodeling of a subset of local regulatory sites.** (a) Plots display read density of DNase1 experiments centered on CRX binding site of Dependent and Independent Group A sites. (b) Analysis of RNAseq of nearest gene to each peak, displayed as boxplot of normalized RPKM values at P2 and P21 in *WT* and P21 in *Crx*<sup>-/-</sup>. (Wilcoxon Rank Sum Test, Paired; \*  $p < 0.05$ , \*\*  $p < 2.2 \times 10^{-5}$ , \*\*\*  $p < 2.2 \times 10^{-10}$ , \*\*\*\*  $p < 2.2 \times 10^{-16}$ )



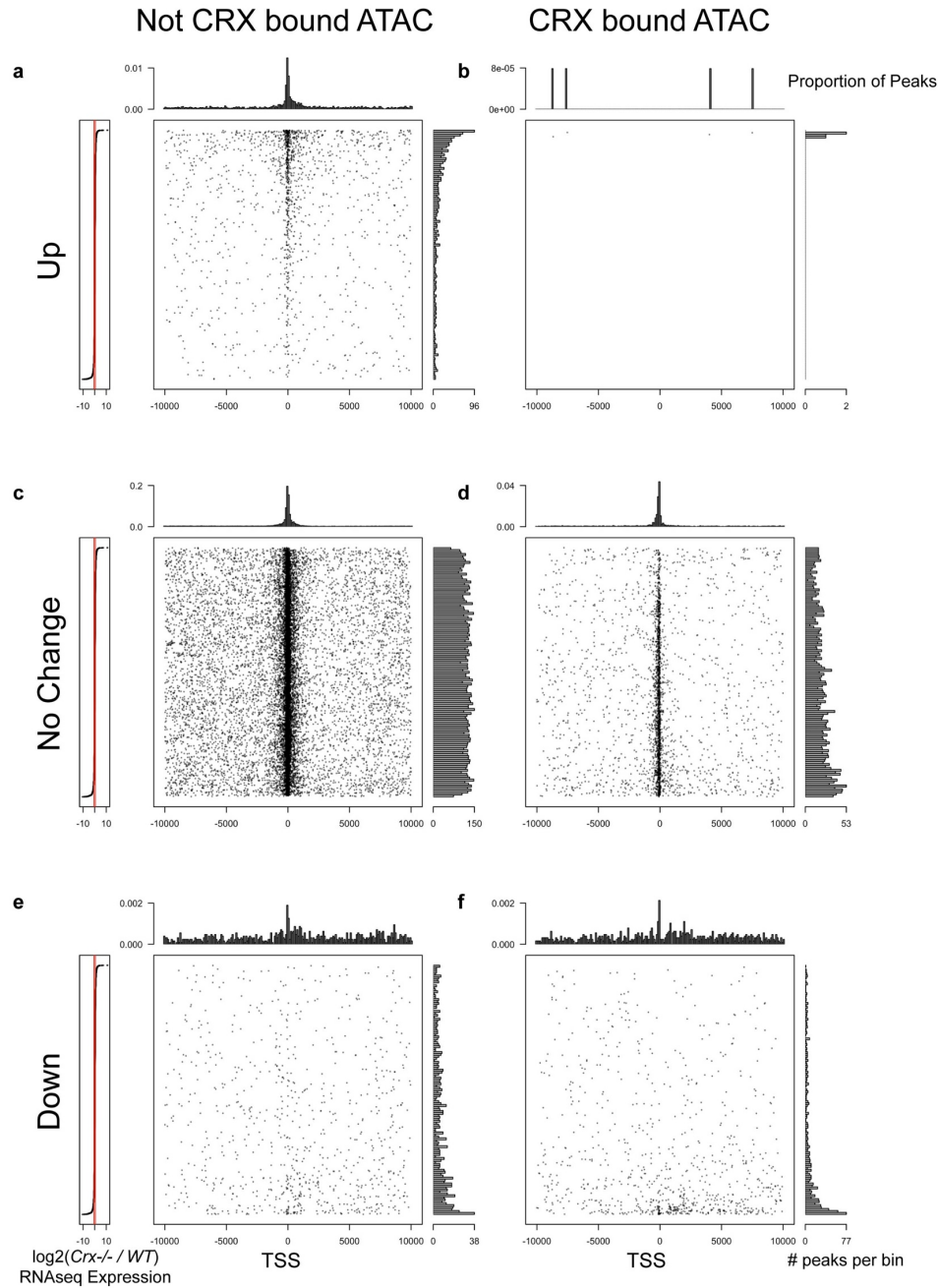
**Figure 4.5: CRX activates a subset of distal enhancers over development.** (a) Plots display read density of DNase1 experiments centered on CRX binding sites of Dependent and Independent Group C sites. (b) Analysis of RNAseq of nearest gene to each peak, displayed as boxplot of normalized RPKM values at P2 and P21 in *WT* and P21 in *Crx*<sup>-/-</sup>. (Wilcoxon Rank Sum Test, Paired; \*  $p < 0.05$ , \*\*  $p < 2.2 \times 10^{-5}$ , \*\*\*  $p < 2.2 \times 10^{-10}$ , \*\*\*\*  $p < 2.2 \times 10^{-16}$ )



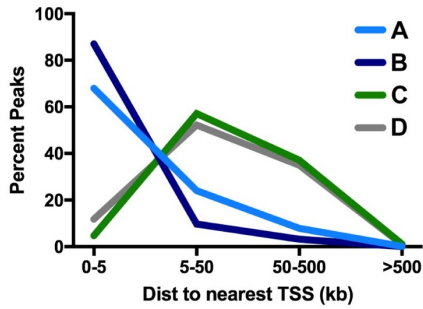
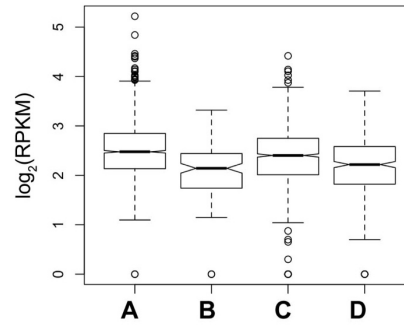
**Figure 4.6: Sequence analyses differentiate Dependent and Independent CRX sites.** (a) Nucleotide density and (b) conservation scores display differences between Groups A, C and D, and between Dependent and Independent classes of CRX sites in each group. Both are calculated in 20bp windows +/- 1kb from center of CRX peak. (Wilcoxon Rank Sum Test, \* p < 0.05, \*\* p < 2.2x10<sup>-5</sup>, \*\*\* p < 2.2x10<sup>-10</sup>, \*\*\*\* p < 2.2x10<sup>-16</sup>). (c, d) CRX occupancy scores, as defined by BEEML algorithm are displayed for two DNA:protein ratios ( $\mu$  values).



**Figure 4.7: Model of CRX mechanism of action.** Model synthesizes new insights into difference in CRX activity and mechanism of action at CRX Dependent vs Independent proximal and distal regulatory sites. The model describes differences in activity state over time (left), and distils motif analysis and base conservation data.

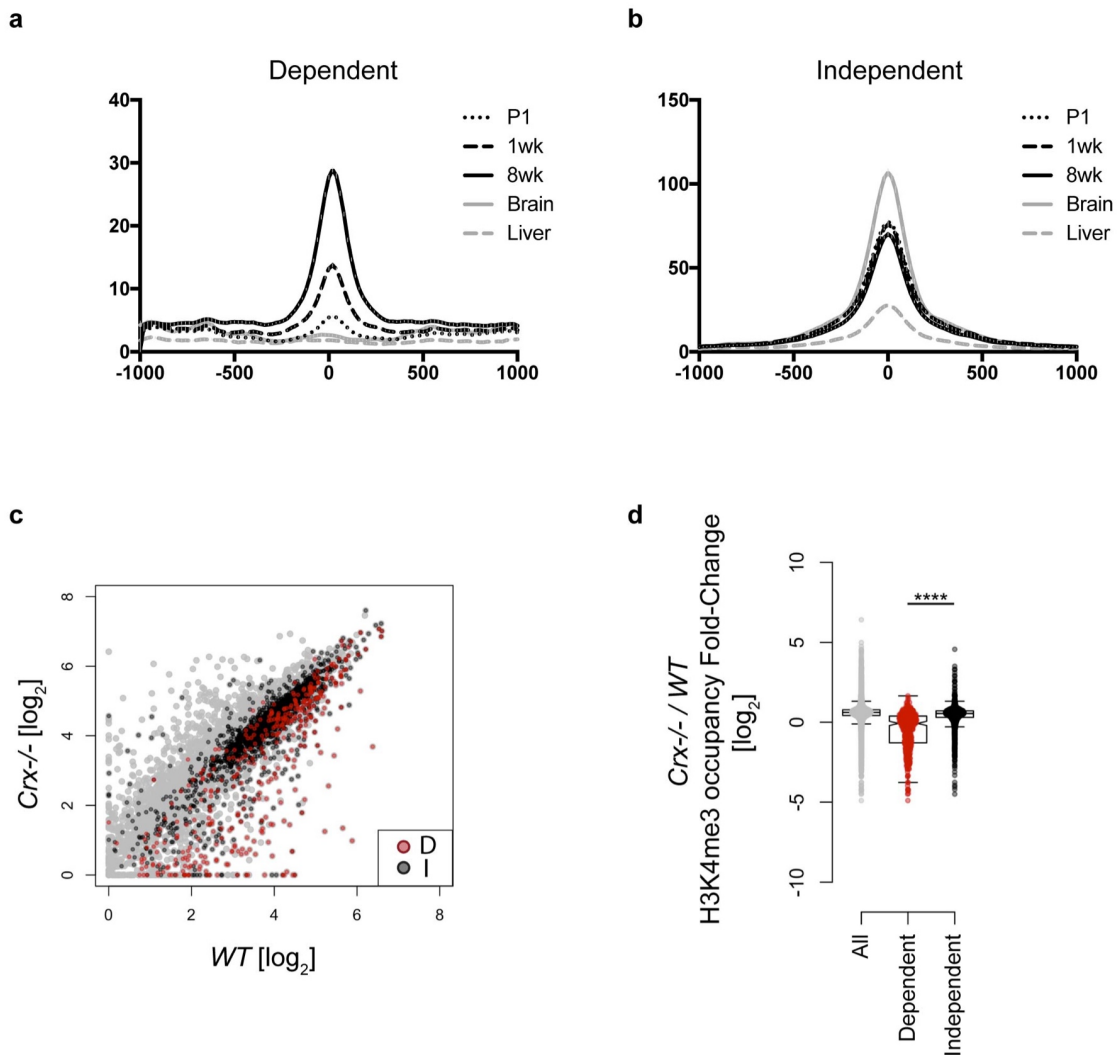


**Supplemental Figure 4.1: CRX bound ATAC peak signal changes correlate with expression changes of nearby genes.** (a, c, e - left panel) Meta-gene plots of all genes expressed in P21 *WT* and *Crx*<sup>-/-</sup> retinas, ordered by [ $\log_2$ ] fold-change (as depicted in plot on left). Black dots represent the center of ATAC regulatory site relative to TSS of all ATAC peaks not bound by CRX (a, c, e - right panel) and of only the subset bound by CRX (b, d, f). Peaks are divided by their changes in ATAC signal- sites increased in *Crx*<sup>-/-</sup> relate to *WT* (a, b), those that are not changed (c, d), and those that decrease in *Crx*<sup>-/-</sup> (e, f). Histograms of X and Y axes display density and distribution of the data.

**a****b**

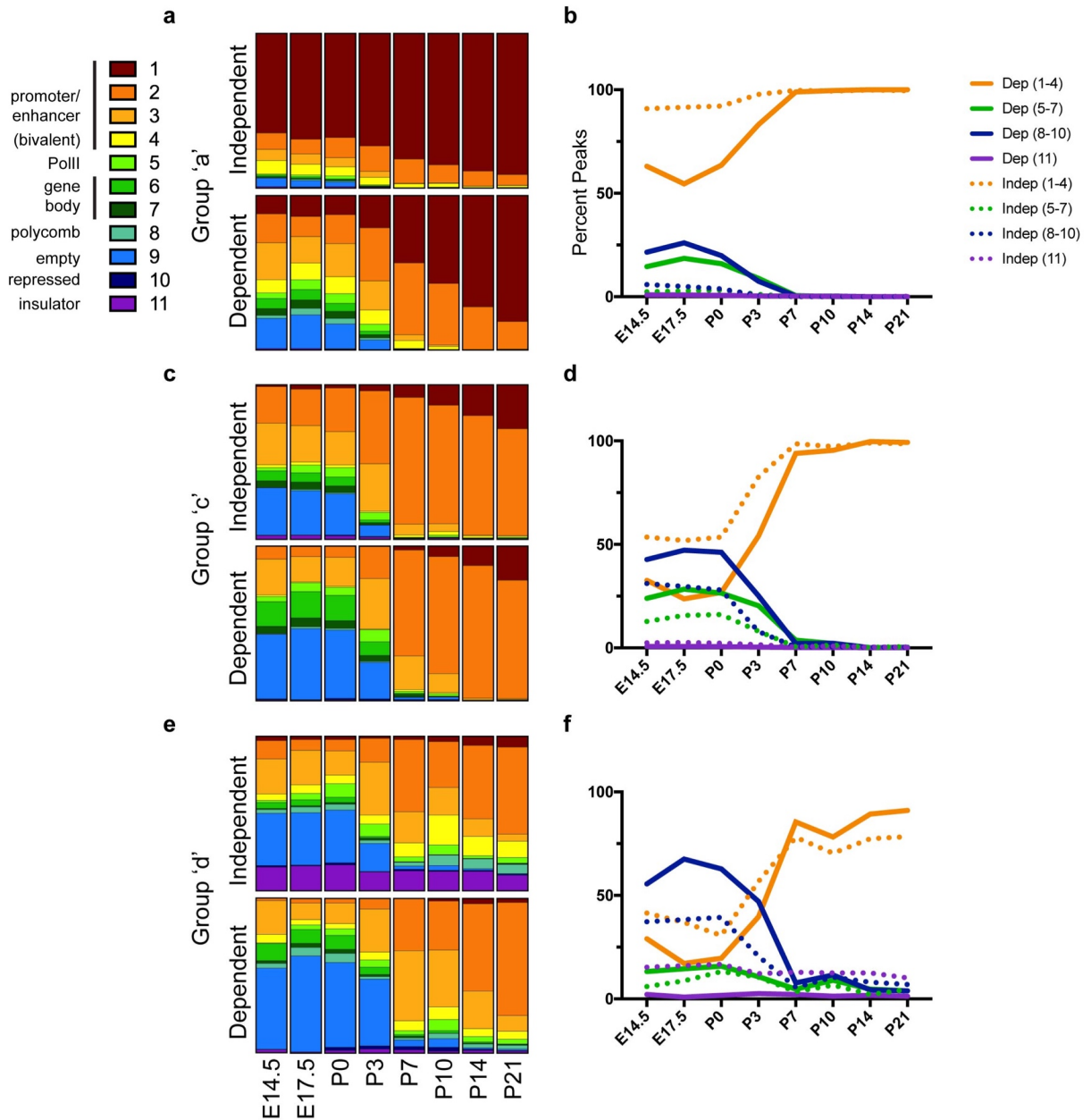
**Supplemental Figure 4.2: CRX binds in distinct regulatory environments. (a)** Quantification of the distribution of peaks within Groups A-D and their distance to the nearest TSS. **(b)** Boxplots represent normalized expression (RPKM) of the nearest gene to each peak within Groups A-D.



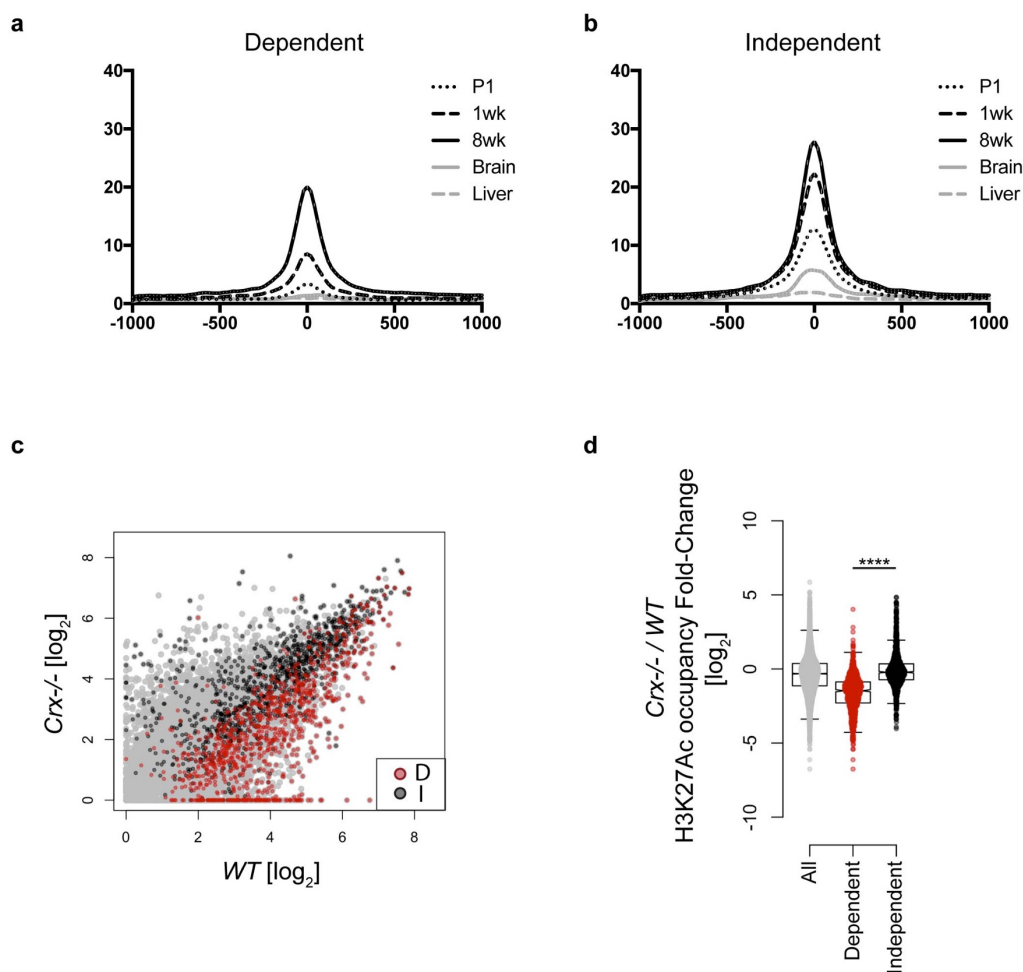


**Supplemental Figure 4.3: CRX is required to activate Dependent local regulatory elements and to remodel chromatin.** Plots display mean and SEM of DNase1 data presented in Figure 4.4a for Dependent (**a**) and Independent (**b**) CRX sites. (**c**) Scatterplot displays changes in H3K4me3 deposition at all peaks (light gray). Black and red denote the subset of H3K4me3 peaks that contain Independent and Dependent CRX Group A sites. (**d**) Quantification of the fold change of H3K4me3 deposition in *WT* vs *Crx*<sup>-/-</sup>. (Wilcoxon Rank Sum Test; \*\*\*\*  $p < 2.2 \times 10^{-16}$ )

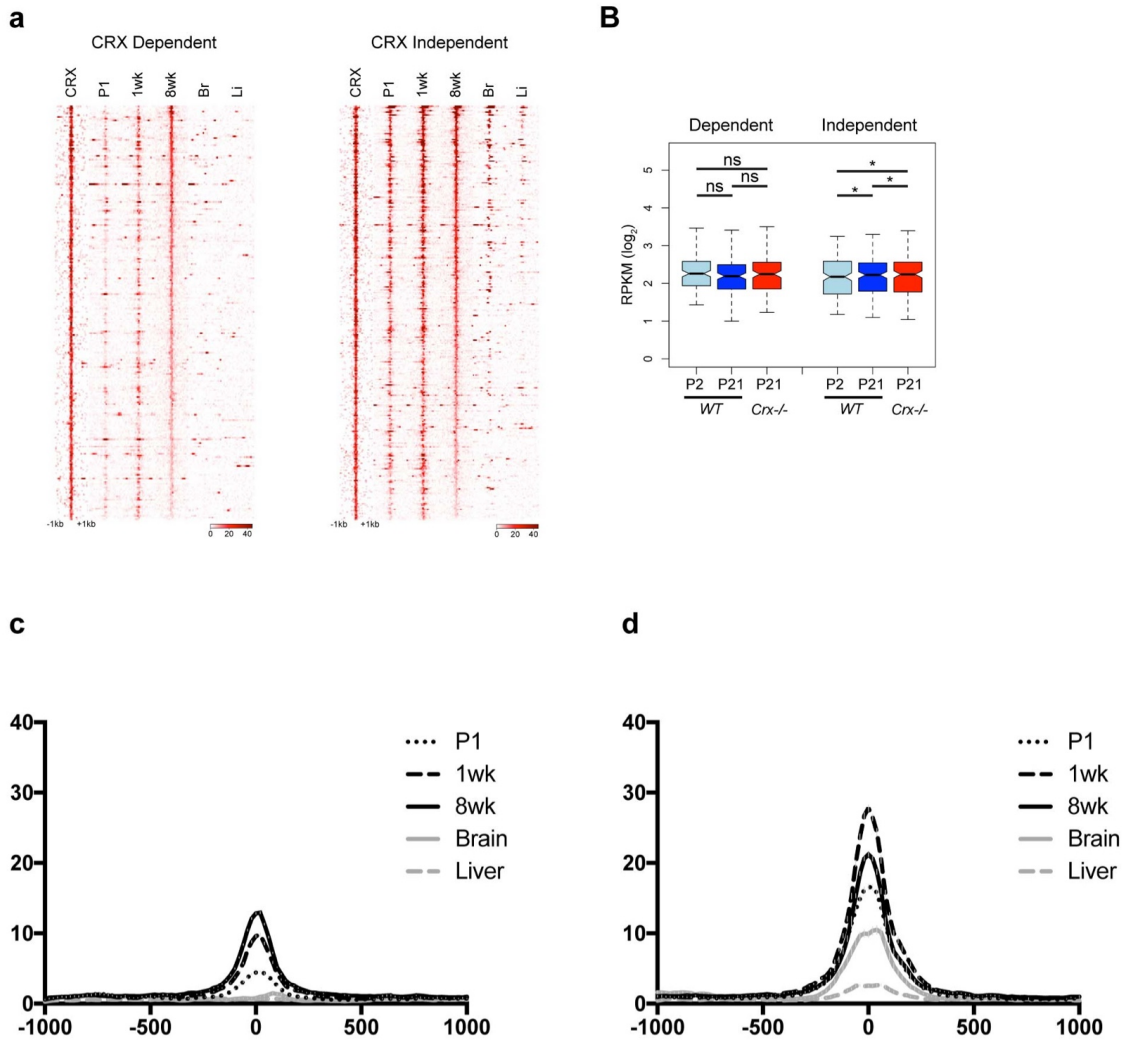




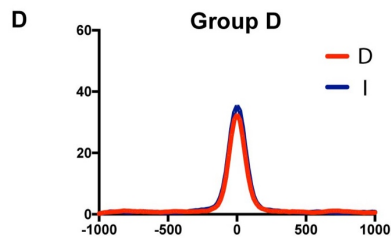
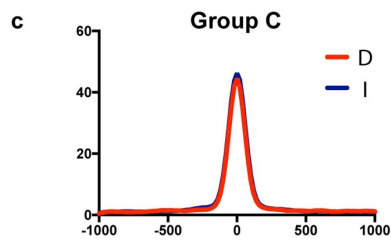
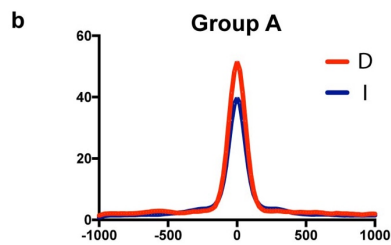
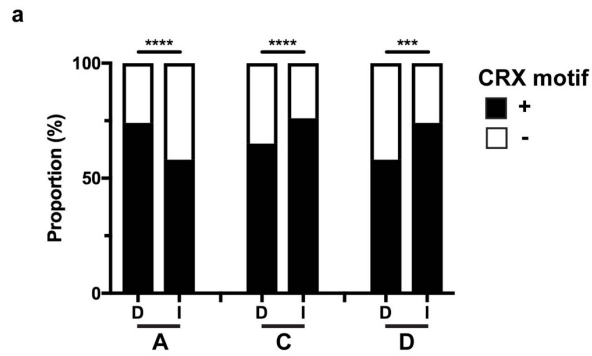
**Supplemental Figure 4.4: Dependent and Independent sites show different chromatin state dynamics. (a, c, e)** Stacked bargraphs represent proportion of Group A-B sites that are contained within each HMM defined chromatin state at 8 developmental ages. **(b, d, f)** Quantification of data binned into HMM classes 1-4, 5-7, 8-10, and 11, for Dependent and Independent sites show different dynamics of reorganization over development. Legend defines basic classification of HMM classes.[Aldiri et al., 2017]



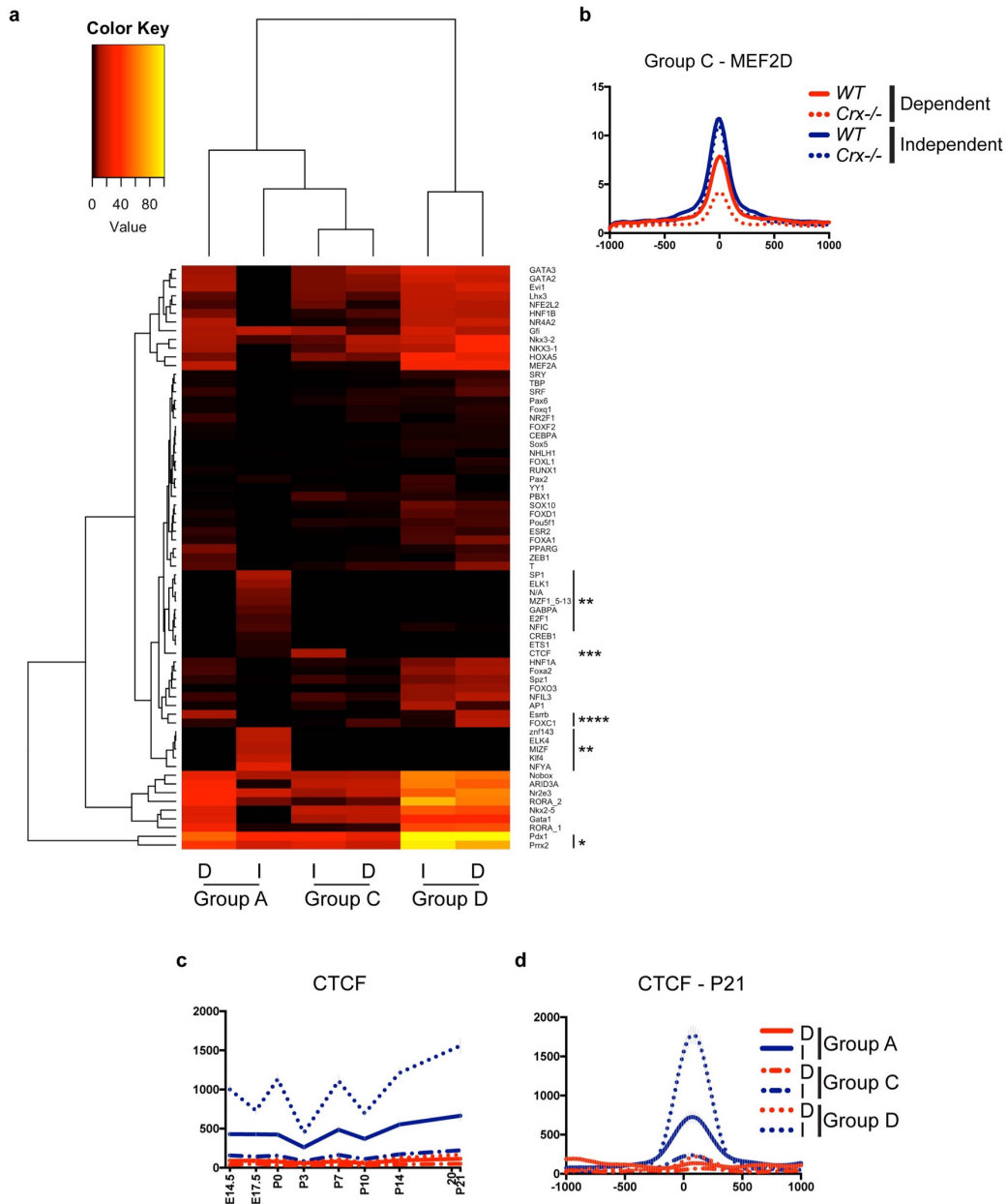
**Supplemental Figure 4.5: CRX is required to activate Dependent enhancer elements and remodel chromatin.** Plots display mean and SEM of DNase1 data presented in Fig.5a for Dependent (**a**) and Independent (**b**) CRX sites. (**c**) Scatterplot displays changes in H3K27Ac deposition at all peaks (light gray). Black and red denote the subset of H3K27Ac peaks that contain Independent and Dependent CRX Group C sites. (**d**) Quantification of the fold change of H3K27Ac deposition in *WT* vs *Crx*<sup>-/-</sup>. (Wilcoxon Rank Sum Test; \*\*\*\*  $p < 2.2 \times 10^{-16}$ )



**Supplemental Figure 4.6: CRX is required to activate Dependent distal regulatory elements.** (a) Plots display read density of DNase1 experiments centered on CRX binding site of Dependent and Independent Group D sites. (b) Analysis of RNAseq of nearest gene to each peak, displayed as boxplot of normalized RPKM values at P2 and P21 in *WT* and P21 in *Crx*<sup>-/-</sup>. Plots display mean and SEM of DNase1 data above for Dependent (d) and Independent (e) CRX sites. (Wilcoxon Rank Sum Test, Paired; \*  $p < 0.05$ , \*\*  $p < 2.2 \times 10^{-5}$ , \*\*\*  $p < 2.2 \times 10^{-10}$ , \*\*\*\*  $p < 2.2 \times 10^{-16}$ )



**Supplemental Figure 4.7: Sites display different presence of CRX motif and binding. (a)** Proportion of sites within each of the noted groups that contains the published CRX motif. **(b-d)** Quantification of read depth of CRX ChIP-seq at Dependent (red) and Independent (blue) sites within the specified groups. (Fisher's exact test; \*\*\*  $p = 0.0001$ , \*\*\*\*  $p < 0.0001$ )



**Supplemental Figure 4.8: TF motifs explain nature of Independent site activation in absence of CRX.** (a) Heatmap shows unsupervised clustering of  $[-\log_{10}]$  transformed p-values representing the significance of representation of the noted TF motif within the set of sites. (\* - \*\*\*\*) represent TFs referenced in details in the text. (b) Quantification of MEF2D ChIP-seq data at Group C enhancer sites displays loss of signal specifically at Dependent sites (red) in the *Crx*<sup>-/-</sup>. Independent sites (blue) show no change in signal. (c) Quantification of CTCF binding over time (maximum peak intensity over development) and (d) relative to CRX bindings site at P21 display consistent signal at Group D Independent sites.

<b>Increased in Crx<sup>-/-</sup> (Up)</b>	
# Term Name	Binom FDR Q-Val
detection of visible light	2.66E-08
Notch signaling pathway	1.59E-07
oligodendrocyte cell fate specification	8.29E-07
somatic motor neuron differentiation	1.20717E-06
peripheral nervous system development	1.78783E-06
neuromuscular process controlling balance	3.88239E-06
smoothed signaling pathway involved in dorsal/ventral neural tube patter	1.07769E-05
body fluid secretion	1.91156E-05
regulation of timing of cell differentiation	2.15899E-05
regulation of development, heterochronic	2.56657E-05

<b>Group A Independent</b>	
# Term Name	Binom FDR Q-Val
sensory perception of light stimulus	4.09E-18
protein transport	4.95E-18
retina homeostasis	1.91E-17
visual perception	2.86E-17
establishment of protein localization	7.57E-17
photoreceptor cell maintenance	3.29E-15
RNA processing	1.98E-13
peptidyl-lysine modification	6.54E-11
photoreceptor cell differentiation	1.40E-10
detection of light stimulus	8.22E-10

<b>Group C Independent</b>	
# Term Name	Binom FDR Q-Val
actin-mediated cell contraction	0.0001399599
aspartate transport	0.011581789
C4-dicarboxylate transport	0.01317136
cell-cell junction organization	0.01471496
glycoprotein biosynthetic process	0.031408027
cell junction organization	0.038060625

<b>Group D Independent</b>	
# Term Name	Binom FDR Q-Val
no terms significantly associated	

<b>Reduced in Crx<sup>-/-</sup> (Down)</b>	
# Term Name	Binom FDR Q-Val
phototransduction	3.89E-22
detection of light stimulus	1.07E-20
sensory perception of light stimulus	3.87E-20
visual perception	9.66E-20
microtubule-based process	9.16E-16
regulation of nuclear-transcribed mRNA catabolic process, deadenylation-dependent decay	2.82E-14
positive regulation of nuclear-transcribed mRNA catabolic process, deadenylation-dependent decay	1.27E-13
positive regulation of mRNA catabolic process	5.86E-13
retina development in camera-type eye	3.01E-12
regulation of Rab GTPase activity	4.46E-10

<b>Group A Dependent</b>	
# Term Name	Binom FDR Q-Val
visual perception	5.91E-23
sensory perception of light stimulus	6.61E-23
phototransduction	1.44E-23
detection of light stimulus	4.51E-23
retina development in camera-type eye	1.74E-20
response to light stimulus	8.43E-13
camera-type eye development	1.61E-12
eye development	2.87E-11
detection of abiotic stimulus	7.26E-10
photoreceptor cell development	9.01E-10

<b>Group C Dependent</b>	
# Term Name	Binom FDR Q-Val
cilium organization	0.000735145
cilium assembly	0.001590798
cilium morphogenesis	0.002355205
microtubule-based process	0.002403853
peptidyl-serine phosphorylation	0.003606846
peptidyl-serine modification	0.007377643
regulation of myeloid cell differentiation	0.015567533
negative regulation of insulin receptor signaling pathway	0.016723291
establishment of organelle localization	0.023717345
positive regulation of myeloid cell differentiation	0.027914745

<b>Group D Dependent</b>	
# Term Name	Binom FDR Q-Val
no terms significantly associated	

**Supplemental Table 4.1: GREAT analysis of CRX bound regulatory sites.** Each table lists the top 10 enriched GO categories for the specified group of sites.








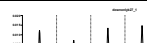







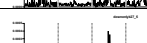







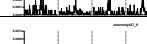
### Group A Dependent

Rank	Motif	Best Match/Details		P-value	log P-pvalue	% of Target	% of Backg	STD(Bg STD)
1		Crx/MA0467.1/Jaspar(0.914)		1.00E-244	-5.63E+02	66.19%	12.18%	36.2bp (74.4bp)
2		PITX3/MA0714.1/Jaspar(0.769)		1.00E-23	-3.37E+01	6.42%	0.90%	40.1bp (76.7bp)
3		RORgt(NR)/EL4-RORgt-Flag-ChIP-Seq(GSE36019)/Homer(0.895)		1.00E-18	-4.31E+01	9.56%	2.59%	50.5bp (62.3bp)
4		PHI130.1/Orc2/Jaspar(0.609)		1.00E-16	-3.78E+01	1.00%	0.00%	42.3bp (36.2bp)
5		LIN54/MA0619.1/Jaspar(0.761)		1.00E-13	-3.01E+01	6.56%	1.77%	48.6bp (75.1bp)
6		SMAD2:SMAD3:SMAD4/MA0513.1/Jaspar(0.565)		1.00E-12	-2.90E+01	1.28%	0.03%	60.7bp (54.4bp)
7		POU6F2/MA0793.1/Jaspar(0.788)		1.00E-12	-2.81E+01	1.00%	0.01%	40.7bp (12.5bp)








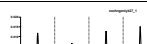








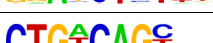
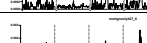







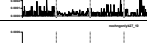

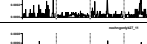


### Group A Independent

Rank	Motif	Best Match/Details		P-value	log P-pvalue	% of Target	% of Backg	STD(Bg STD)
1		OTX1/MA0711.1/Jaspar(0.963)		1e-452	-1.04E+03	46.35%	9.95%	36.9bp (75.6bp)
2		POL003.1 GC-box/Jaspar(0.918)		1.00E-152	-3.52E+02	31.89%	11.54%	51.3bp (69.7bp)
3		NFY(CCAAT)/Promoter/Homer(0.911)		1.00E-144	-3.33E+02	25.36%	7.94%	51.4bp (74.1bp)
4		GFY(?)/Promoter/Homer(0.985)		1.00E-118	-2.74E+02	7.25%	0.63%	53.2bp (74.8bp)
5		Crem/MA0609.1/Jaspar(0.733)		1.00E-103	-2.38E+02	65.97%	43.96%	55.0bp (75.5bp)
6		NRF1(NRF)/MCF7-NRF1-ChIP-Seq(Unpublished)/Homer(0.989)		1.00E-46	-1.08E+02	9.64%	3.22%	55.7bp (71.1bp)
7		RORgt(NR)/EL4-RORgt-Flag-ChIP-Seq(GSE36019)/Homer(0.927)		1.00E-32	-7.51E+01	3.31%	0.59%	48.0bp (60.8bp)
8		BHLHE41/MA0636.1/Jaspar(0.901)		1.00E-28	-6.59E+01	18.73%	10.95%	53.6bp (68.7bp)
9		ZBED1/MA0749.1/Jaspar(0.833)		1.00E-24	-5.69E+01	12.87%	6.89%	53.3bp (75.5bp)
10		Atf1(bZIP)/K562-ATF1-ChIP-Seq(GSE31477)/Homer(0.959)		1.00E-22	-5.26E+01	6.71%	2.76%	45.9bp (63.9bp)
11		YY1(Zf)/Promoter/Homer(0.849)		1.00E-20	-4.76E+01	3.02%	0.78%	56.4bp (69.0bp)
12		Pbx3(Homeobox)/GM12878-PBX3-ChIP-Seq(GSE32465)/Homer(0.955)		1.00E-20	-4.71E+01	2.72%	0.65%	51.0bp (58.5bp)
13		STAT3/MA0144.2/Jaspar(0.849)		1.00E-18	-4.37E+01	9.18%	4.77%	53.1bp (65.2bp)
14		FOXJ1/MA0042.2/Jaspar(0.684)		1.00E-17	-4.08E+01	5.95%	2.63%	49.3bp (81.6bp)
15		ETV2/MA0762.1/Jaspar(0.671)		1.00E-16	-3.74E+01	5.74%	2.61%	56.8bp (70.0bp)
16		PAX3-FKHR-fusion(Paired.Homeobox)/Rb4-PAX3-FKHR-ChIP-Seq(GSE19063)/Homer(0.630)		1.00E-16	-3.70E+01	0.42%	0.01%	44.4bp (35.9bp)
17		Atf1/MA0604.1/Jaspar(0.710)		1.00E-15	-3.63E+01	2.81%	0.85%	52.9bp (76.1bp)
18		PB0077.1 Spdef.1/Jaspar(0.664)		1.00E-15	-3.52E+01	2.05%	0.50%	43.6bp (72.5bp)
19		NFY(CCAAT)/Promoter/Homer(0.668)		1.00E-12	-2.83E+01	3.39%	1.37%	61.8bp (63.6bp)

### Group C Dependent




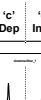
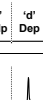
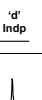
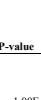

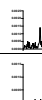

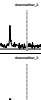
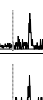

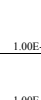









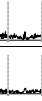



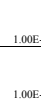







Rank	Motif	Best Match/Details	 'a' Dep	 'a' Indp	 'c' Dep	 'c' Indp	 'g' Dep	 'g' Indp	P-value	log P-pvalue	% of Target	% of Backg	STD(Bg STD)
1		OTX1/MA0711.1/Jaspar(0.928)							1.00E-244	-5.63E+02	68.30%	15.88%	37.0bp (73.4bp)
2		PB0048.1 Nkx3-1.1/Jaspar(0.782)							1.00E-75	-1.73E+02	24.94%	5.38%	42.9bp (72.0bp)
3		PH0045.1 Hoxa1/Jaspar(0.716)							1.00E-31	-7.37E+01	22.48%	8.69%	50.4bp (73.7bp)
4		MEF2B/MA0660.1/Jaspar(0.937)							1.00E-30	-6.99E+01	9.95%	2.03%	42.1bp (74.3bp)
5		E2A(bHLH)/near_PU.1/Bcell-PU.1-ChIP-Seq(GSE21512)/Homer(0.753)							1.00E-19	-4.49E+01	3.93%	0.44%	45.4bp (60.4bp)
6		PH0130.1 Otx2/Jaspar(0.622)							1.00E-19	-4.43E+01	12.29%	4.38%	48.0bp (71.2bp)
7		Smad3(MAD)/NPC-Smad3-ChIP-Seq(GSE36673)/Homer(0.753)							1.00E-15	-3.67E+01	38.57%	25.42%	56.1bp (66.9bp)
8		RORgt(NR)/EL4-RORgt-Flag-ChIP-Seq(GSE56019)/Homer(0.961)							1.00E-14	-3.30E+01	2.95%	0.35%	51.7bp (61.9bp)
9		TFCP2/MA0145.3/Jaspar(0.708)							1.00E-13	-3.06E+01	6.27%	1.80%	51.8bp (65.3bp)

### Group C Independent









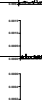

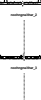


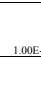








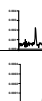

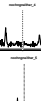
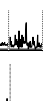

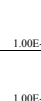


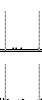
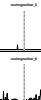
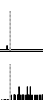
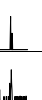
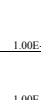

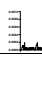



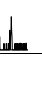
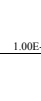
Rank	Motif	Best Match/Details	 'a' Dep	 'a' Indp	 'c' Dep	 'c' Indp	 'g' Dep	 'g' Indp	P-value	log P-pvalue	% of Target	% of Backg	STD(Bg STD)
1		OTX1/MA0711.1/Jaspar(0.944)							1e-437	-1.01E+03	77.17%	14.24%	36.1bp (72.7bp)
2		NeuroD1(bHLH)/Islet-NeuroD1-ChIP-Seq(GSE30298)/Homer(0.961)							1.00E-37	-8.57E+01	13.66%	3.76%	52.2bp (67.2bp)
3		RORgt(NR)/EL4-RORgt-Flag-ChIP-Seq(GSE56019)/Homer(0.947)							1.00E-34	-7.93E+01	9.07%	1.82%	45.5bp (64.7bp)
4		Me2d(MADS)/Retina-Me2d-ChIP-Seq(GSE61391)/Homer(0.961)							1.00E-31	-7.32E+01	4.49%	0.38%	45.6bp (64.3bp)
5		MyoG(bHLH)/C2C12-MyoG-ChIP-Seq(GSE36024)/Homer(0.653)							1.00E-19	-4.52E+01	6.58%	1.67%	40.9bp (73.8bp)
6		Meis1(Homeobox)/MastCells-Meis1-ChIP-Seq(GSE48085)/Homer(0.894)							1.00E-17	-4.00E+01	19.24%	10.14%	54.1bp (68.3bp)
7		USF1(bHLH)/GM12878-Ust1-ChIP-Seq(GSE32465)/Homer(0.959)							1.00E-15	-3.65E+01	4.89%	1.16%	52.2bp (63.4bp)
8		RHOXF1/MA0719.1/Jaspar(0.622)							1.00E-14	-3.24E+01	10.57%	4.64%	52.8bp (65.6bp)
9		GATA3(ZF)/Treg-Gata3-ChIP-Seq(GSE20898)/Homer(0.699)							1.00E-13	-3.06E+01	2.79%	0.45%	52.7bp (68.7bp)
10		Nr2e3/MA0164.1/Jaspar(0.699)							1.00E-12	-2.88E+01	3.29%	0.68%	44.3bp (65.7bp)
11		PH0040.1 Hmbox1/Jaspar(0.727)							1.00E-12	-2.86E+01	6.98%	2.61%	52.2bp (69.9bp)
12		GMEB2/MA0862.1/Jaspar(0.726)							1.00E-12	-2.85E+01	14.56%	7.81%	55.8bp (67.2bp)



### Group D Dependent

Rank	Motif	Best Match/Details	'a' Dep	'a' Indp	'c' Dep	'c' Indp	'd' Dep	'd' Indp	P-value	log P-value	% of Target	% of Backg	STD(Bg STD)
1		Crx/MA0467.1/Jaspar(0.886)							1.00E-64	-1.49E+02	75.30%	15.58%	37.0bp (76.3bp)
2		Crx/MA0467.1/Jaspar(0.690)							1.00E-18	-4.17E+01	24.10%	4.37%	45.2bp (72.7bp)
3		SCL(bHLH)/HPC7-Scf-ChIP-Seq(GSE13511)/Homer(0.564)							1.00E-14	-3.33E+01	8.43%	0.37%	35.3bp (71.3bp)
4		Pitx1(Homeobox)/Chicken-Pitx 1-ChIP-Seq(GSE38910)/Homer(0.743)							1.00E-13	-3.12E+01	13.86%	1.74%	42.9bp (73.1bp)
5		ISL2/MA0914.1/Jaspar(0.680)							1.00E-12	-2.97E+01	13.86%	1.88%	57.7bp (74.2bp)

### Group D Independent

Rank	Motif	Best Match/Details	'a' Dep	'a' Indp	'c' Dep	'c' Indp	'd' Dep	'd' Indp	P-value	log P-value	% of Target	% of Backg	STD(Bg STD)
1		CRX(Homeobox)/Retina-Crx-ChIP-Seq(GSE20012)/Homer(0.995)							1.00E-107	-2.48E+02	77.08%	14.85%	38.5bp (77.1bp)
2		CTCF(Zf)/CD4+CTCF-ChIP-Seq(Barski et al.)/Homer(0.908)							1.00E-32	-7.46E+01	13.44%	0.67%	44.2bp (69.2bp)
3		YY2/MA0748.1/Jaspar(0.646)							1.00E-14	-3.45E+01	4.74%	0.13%	48.8bp (59.8bp)
4		PB0048.1 Nkx3-1.1/Jaspar(0.754)							1.00E-14	-3.23E+01	25.30%	8.82%	51.3bp (70.8bp)
5		E2F8/MA0865.1/Jaspar(0.589)							1.00E-13	-3.08E+01	1.98%	0.00%	41.1bp (0.0bp)
6		PHOX2A/MA0713.1/Jaspar(0.835)							1.00E-12	-2.90E+01	6.32%	0.48%	45.1bp (78.1bp)

**Supplemental Table 4.2: De novo motif analysis identifies sequences enriched in different sets of CRX bound regions.** Data presented in columns A-C & E-I are generated by HOMER (see Appendix 2 methods). Column D represents the scanning of the identified sequence back across the 6 classes of CRX sites (+/- 1kb from peak center). The plot in line 1 depicts the layout of the plot, and the y axis quantifies ‘motifs per bp.’

## 4.6 References

- Aldiri, I., Xu, B., Wang, L., Chen, X., Hiler, D., Griffiths, L., Valentine, M., et al. (2017). The Dynamic Epigenetic Landscape of the Retina During Development, Reprogramming, and Tumorigenesis. *Neuron*, *94*, 550–568.
- Andzelm, M. M., Cherry, T. J., Harmin, D. A., Boeke, A. C., Lee, C., Hemberg, M., Pawlyk, B., et al. (2015). MEF2D drives photoreceptor development through a genome-wide competition for tissue-specific enhancers. *Neuron*, *86*, 247–263.
- Baden, T., Berens, P., Franke, K., Roman-Roson, M., Bethge, M., & Euler. (2016). The functional diversity of mouse retinal ganglion cells. *Nature*, *529*, 345–350.
- Benayoun, B. A., Pollina, E. A., Ucar, D., Mahmoudi, S., Karra, K., Wong, E. D., Devarajan, K., et al. (2014). H3K4me3 Breadth Is Linked to Cell Identity and Transcriptional Consistency. *Cell*, *158*, 673–688.
- Brooks, M. J., Rajasimha, H. K., Roger, J. E., & Swaroop, A. (2011). Next-generation sequencing facilitates quantitative analysis of wild-type and *Nrl*(<sup>-/-</sup>) retinal transcriptomes. *Molecular Vision*, *17*, 3034–54.
- Buenrostro, J. D., Wu, B., Chang, H. Y., & Greenleaf, W. J. (2015). ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Current Protocols In Molecular Biology*, *2015*, 21.29.1-21.29.9.
- Chen, J., Rattner, A., & Nathans, J. (2005). The rod photoreceptor-specific nuclear receptor Nr2e3 represses transcription of multiple cone-specific genes. *The Journal Of Neuroscience*, *25*, 118–29.
- Chen, S., Wang, Q. L., Nie, Z., Sun, H., Lennon, G., Copeland, N. G., Gilbert, D. J., et al. (1997). Crx, a novel Otx-like paired-homeodomain protein, binds to and transactivates photoreceptor cell-specific genes. *Neuron*, *19*, 1017–30.
- Chen, S., Wang, Q., Xu, S., Liu, I., Li, L. Y., Wang, Y., & Zack, D. J. (2002). Functional analysis of cone – rod homeobox ( CRX ) mutations associated with retinal dystrophy. *Human Molecular Genetics*, *11*, 873–884.
- Corbo, J. C., Lawrence, K. a, Karlstetter, M., Myers, C. a, Abdelaziz, M., Dirkes, W., Weigelt, K., et al. (2010). CRX ChIP-seq reveals the cis-regulatory architecture of mouse photoreceptors. *Genome Research*, *20*, 1512–25.
- Furukawa, T., Morrow, E. M., & Cepko, C. L. (1997). Crx, a novel otx-like homeobox gene, shows photoreceptor-specific expression and regulates photoreceptor differentiation. *Cell*, *91*, 531–41.
- Furukawa, T., Morrow, E. M., Li, T., Davis, F. C., & Cepko, C. L. (1999). Retinopathy and attenuated circadian entrainment in Crx-deficient mice. *Nature Genetics*, *23*, 466–70.
- Grossman, S. R., Zhang, X., Wang, L., Engreitz, J., Melnikov, A., Rogov, P., Tewhey, R., et al. (2017). Systematic dissection of genomic features determining transcription factor binding and enhancer function. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, *114*, E1291–E1300.
- Hennig, A. K., Peng, G.-H., & Chen, S. (2013). Transcription Coactivators p300 and CBP Are Necessary for Photoreceptor-Specific Chromatin Organization and Gene Expression. *PloS One*, *8*, e69721.
- Hsiao, T. H.-C., Diaconu, C., Myers, C. A., Lee, J., Cepko, C. L., & Corbo, J. C. (2007). The cis-

- regulatory logic of the mammalian photoreceptor transcriptional network. *PloS One*, 2, e643.
- Hughes, A. E. O., Enright, J. M., Myers, C. A., Shen, S. Q., & Corbo, J. C. (2017). Cell Type-Specific Epigenomic Analysis Reveals a Uniquely Closed Chromatin Architecture in Mouse Rod Photoreceptors. *Scientific Reports*, 7, 43184.
- Jeon, C. J., Strettoi, E., & Masland, R. H. (1998). The major cell populations of the mouse retina. *The Journal Of Neuroscience*, 18, 8936–46.
- Kim, J.-W., Yang, H.-J., Oel, A. P., Brooks, M. J., Jia, L., Plachetzki, D. C., Li, W., et al. (2016). Recruitment of Rod Photoreceptors from Short-Wavelength-Sensitive Cones during the Evolution of Nocturnal Vision in Mammals. *Developmental Cell*, 37, 520–532.
- Landolin, J. M., Johnson, D. S., Trinklein, N. D., Landolin, J. M., Johnson, D. S., Trinklein, N. D., Aldred, S. F., et al. (2010). Sequence features that drive human promoter function and tissue specificity. *Genome Research*, 20, 890–898.
- Lem, J., Krasnoperova, N. V., Calvert, P. D., Kosaras, B., Cameron, D. a, Nicolò, M., Makino, C. L., et al. (1999). Morphological, physiological, and biochemical changes in rhodopsin knockout mice. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 96, 736–41.
- Macosko, E. Z., Basu, A., Satija, R., Nemes, J., Shekhar, K., Goldman, M., Tirosh, I., et al. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*, 161, 1202–1214.
- McLean, C. Y., Bristor, D., Hiller, M., Clarke, S. L., Schaar, B. T., Lowe, C. B., Wenger, A. M., et al. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnology*, 28, 495–501.
- Mears, A., Kondo, M., Swain, P. K., Takada, Y., Bush, R. A., Saunders, T. L., Sieving, P. A., et al. (2001). Nrl is required for rod photoreceptor development. *Nature Genetics*, 29, 447–52.
- Mo, A., Luo, C., Davis, F. P., Mukamel, E. A., Henry, G. L., Nery, J. R., Urich, M. A., et al. (2016). Epigenomic landscapes of retinal rods and cones. *ELife*, 5, 1–29.
- Montana, C. L., Lawrence, K. a, Williams, N. L., Tran, N. M., Peng, G.-H., Chen, S., & Corbo, J. C. (2011). Transcriptional regulation of neural retina leucine zipper (Nrl), a photoreceptor cell fate determinant. *The Journal Of Biological Chemistry*, 286, 36921–31.
- Nasonkin, I. O., Merbs, S. L., Lazo, K., Oliver, V. F., Brooks, M., Patel, K., Enke, R. A., et al. (2013). Conditional knockdown of DNA methyltransferase 1 reveals a key role of retinal pigment epithelium integrity in photoreceptor outer segment morphogenesis. *Development*, 140, 1330–1341.
- Palhan, V. B., Chen, S., Peng, G.-H., Tjernberg, A., Gamper, A. M., Fan, Y., Chait, B. T., et al. (2005). Polyglutamine-expanded ataxin-7 inhibits STAGA histone acetyltransferase activity to produce retinal degeneration. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 102, 8472–7.
- Peng, G.-H., Ahmad, O., Ahmad, F., Liu, J., & Chen, S. (2005). The photoreceptor-specific nuclear receptor Nr2e3 interacts with Crx and exerts opposing effects on the transcription of rod versus cone genes. *Human Molecular Genetics*, 14, 747–64.
- Peng, G.-H., & Chen, S. (2005). Chromatin immunoprecipitation identifies photoreceptor transcription factor targets in mouse models of retinal degeneration: new findings and challenges. *Visual Neuroscience*, 22, 575–86.
- Roger, J. E., Hiriyanna, A., Gotoh, N., Hao, H., Cheng, D. F., Ratnapriya, R., Kautzmann, M. I., et al.

- (2014). OTX2 loss causes rod differentiation defect in CRX-associated congenital blindness. *The Journal Of Clinical Investigation*, *124*, 631–643.
- Ruzycki, P. A., Linne, C. D., Hennig, A. K., & Chen, S. (2017). CRX-L253X Mutation Produces Dominant Photoreceptor Defects in TVRM65 Mice. *Investigative Ophthalmology & Visual Science*, *58*, 4644–4653.
- Ruzycki, P. A., Tran, N. M., Kefalov, V. J., Kolesnikov, A. V., & Chen, S. (2015). Graded gene expression changes determine phenotype severity in mouse models of CRX-associated retinopathies. *Genome Biology*, *16*, 171.
- Shekhar, K., Lapan, S. W., Whitney, I. E., Tran, N. M., Macosko, E. Z., Kowalczyk, M., Adiconis, X., et al. (2016). Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. *Cell*, *166*, 1308–1323.
- Siebert, S., Cabuy, E., Scherf, B. G., Kohler, H., Panda, S., Le, Y.-Z., Fehling, H. J., et al. (2012). Transcriptional code and disease map for adult retinal cell types. *Nature Neuroscience*, *15*, 487–95.
- La Spada, A. R., Fu, Y., Sopher, B. L., Libby, R. T., Wang, X., Li, L. Y., Einum, D. D., et al. (2001). Polyglutamine-Expanded Ataxin-7 Antagonizes CRX Function and Induces Cone-Rod Dystrophy in a Mouse Model of SCA7. *Neuron*, *31*, 913–927.
- Swaroop, A., Kim, D., & Forrester, D. (2010). Transcriptional regulation of photoreceptor development and homeostasis in the mammalian retina. *Nature Reviews. Neuroscience*, *11*, 563–576.
- Tran, N. M., & Chen, S. (2014). Mechanisms of blindness: animal models provide insight into distinct CRX-associated retinopathies. *Developmental Dynamics*, *243*, 1153–66.
- Tran, N. M., Zhang, A., Zhang, X., Huecker, J. B., Hennig, A. K., & Chen, S. (2014). Mechanistically Distinct Mouse Models for CRX-Associated Retinopathy. *PLoS Genetics*, *10*, e1004111.
- White, M. A., Kwasnieski, J. C., Myers, C. A., Shen, S. Q., Corbo, J. C., & Cohen, B. A. (2016). A Simple Grammar Defines Activating and Repressing cis-Regulatory Elements in Photoreceptors. *Cell Reports*, *17*, 1247–1254.
- White, M. A., Myers, C. A., Corbo, J. C., & Cohen, B. A. (2013). Massively parallel in vivo enhancer assay reveals that highly local features determine the cis-regulatory function of ChIP-seq peaks. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, *110*, 11952–11957.
- Whyte, W. A., Orlando, D. A., Hnisz, D., Abraham, B. J., Lin, C. Y., Kagey, M. H., Rahl, P. B., et al. (2013). Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. *Cell*, *153*, 307–319.
- Young, R. W. (1985). Cell differentiation in the retina of the mouse. *The Anatomical Record*, *212*, 199–205.
- Yue, F., Cheng, Y., Breschi, A., Vierstra, J., Wu, W., Ryba, T., Sandstrom, R., et al. (2014). A comparative encyclopedia of DNA elements in the mouse genome. *Nature*, *515*, 355–364.
- Zhao, Y., Granas, D., & Stormo, G. D. (2009). Inferring binding energies from selected binding sites. *PLoS Computational Biology*, *5*, e1000590.

# **Chapter 5**

## **Organization of the photoreceptor genome**

## 5.1 Author Contributions

This project was conceived of by Shiming Chen and Philip Ruzycki to build upon past work by former lab member Guang Hua Peng [Peng & Chen, 2011, 2012] and to address several questions in the growing field of genomic organization as we think the rod photoreceptor is a useful model system for this purpose.

I optimized Circularized Chromosome Conformation Capture (4C) and Fluorescence in situ hybridization (FISH) protocols for retinal tissue. Irina Solovei, LMU München, was very generous in sharing her protocols for FISH probe generation and tissue preparation. I performed all 4C molecular biology, RNA-seq, and qRT-PCR and related data analyses. Morphometry and FISH data collection and analyses were performed by myself and Courtney Linne. This project is a work in progress, and the contents of the chapter will be submitted for publication soon.

## 5.2 Abstract

Purpose: The packaging of DNA is a highly regulated and critical process of cellular differentiation. Euchromatin and heterochromatin territories must be determined to assure both proper gene expression and repression. In mouse rod photoreceptors, this packaging is unique as nuclei are inverted; euchromatin occupies only the outer ring of the nucleus while heterochromatin is packaged densely into the nuclear core. While studies have generally described this packaging in terms of gene-rich vs gene-poor, very little is known about the organization at a molecular scale.

Methods: To interrogate fragments of DNA that interact with *Rhodopsin (Rho)*, the highest expressed gene in the cell, we have performed Circularized Chromosome Conformation Capture (4C) and confirmed our analysis with DNA Fluorescence *in situ* Hybridization (FISH). We generated a mouse model where *Rho* has been inactivated and assessed the effects on *Rho*-contacting loci using FISH and RNA-seq.

Results: 4C data suggested a majority of *Rho* interactions were very local, and restricted to a stretch of multiple topologically associated domains (TADs) that are all segregated to the A compartment (active territory). Distant *trans* interactions that were confirmed using FISH were also within A compartment DNA and were highly enriched for active chromatin and known photoreceptor regulatory elements. Inactivation of the *Rho* promoter resulted in a loss of distal contacts, and an increase in gene expression within the interacting *cis* TADs.

Conclusion: This data shows that the rod photoreceptor genome, although very distinct at microscopic resolution, is packaged into a very similar A/B compartment structure as other cell types. This data also suggests that the active A compartment can be subdivided into distinct

units, perhaps based on chromatin state. The inclusion of *Rho* into this sub compartment was dependent upon the activity state of *Rho*. RNA-seq further suggested that within a compartment, machinery is a resource that when otherwise not in use, can affect the expression of other genes.



## 5.3 Introduction

At the molecular level, each chromosome is packaged into the nucleus to ensure proper expression and repression of genes. Previous work has focused on the role of epigenomic gene regulation on the relatively local state of the region surrounding a particular gene. New technologies now allow for the interrogation of the compaction of DNA on a local scale (promoter-enhancer interactions), and beyond to understand how 2 meters of DNA is folded to fit into each nucleus.

The chromosome is first segregated into discrete topologically associated domains (TADs), which serve as ‘genomic neighborhoods’ (~1Mb in size) thought to isolate and ensure proper connections of enhancers and promoters.[Reviewed in [Bonev & Cavalli, 2016]] Indeed, the disruption of a single TAD boundary has been shown to result in the improper spread of repressive or active chromatin while their global depletion results in a gross change in genomic organization; both experiments cause changes to gene expression. [Narendra et al., 2015, 2016; Rao et al., 2017; Schwarzer et al., 2017] Many of these transcriptional changes were spurious novel transcripts from enhancers, thought to result from the lack of a directed connection with its proper promoter.[Schwarzer et al., 2017]

Both studies that depleted TADs genome wide by interfering with the boundary cohesin protein found that the loss of TADs resulted in the strengthening of the next level of “compartment domain” organization.[Rao et al., 2017; Schwarzer et al., 2017] Each chromosome can be subdivided into 5-10Mb alternating fragments that segregate into either the A (active) or B (inactive) compartments, synonymous with eu- and hetero- chromatin.[Lieberman-Aiden et al., 2009] The segregation of a fragment into A or B is surprisingly well conserved across cell types,

with only minimal switching of individual segments between compartments associated with cell type specific gene expression.[Dixon et al., 2015; Schmitt et al., 2016] Within either compartment, there is limited evidence of further sub-compartments (A1/2 & B1/2/3) that are loosely defined by the epigenetic state of the contained DNA.[Rao et al., 2014]

The role of TADs and A/B compartments are still under debate as there is only limited direct evidence for the local regulatory potential of a TAD and the mechanism of inclusion of a fragment into a particular compartment or sub-compartment has not yet been elucidated. However, the fact that interactions defined by epigenetic state can supersede compartment domain structure upon the loss of TADs, suggests that these are opposing forces constantly in play within the nucleus. Previous studies have been limited in their ability to test the inclusion/exclusion of individual loci in A/B compartments, the effect of the exclusion of any individual gene, and to this point there is little evidence of further specialization (finer organization) within these compartments.

The rod photoreceptor presents an excellent system in which to test these questions. The rod photoreceptor is a highly specialized neuron responsible for the initial conversion of a photon of light into an electrical signal. Rods comprise ~80% of the cells in the mouse retina, [Jeon et al., 1998; Macosko et al., 2015] making them an accessible system for cell fate, epigenetic, and gene expression studies. Rod birth peaks at postnatal day 0 (P0), and the cells are specified by P14 and fully differentiated by P21.[Young, 1985] During this prolonged developmental window, studies have defined dramatic changes to the rod epigenome, gene expression, and even to microscopically visible genomic organization. [Aldiri et al., 2017; Hughes et al., 2017; Kim et al., 2016; Mo et al., 2016; Solovei et al., 2009]

Mouse rods have an inverted nuclear architecture that is characteristic of many nocturnal mammals,[Solovei et al., 2009] and the result of the lack of *Lbr* and *Lmna* expression, normally required to tether heterochromatin to the nuclear lamina.[Hughes et al., 2017; Solovei et al., 2013] This organization is perturbed in many models of photoreceptor disease. Several models never assume the proper organization,[Corbo & Cepko, 2005; Hennig et al., 2013; Tran et al., 2014] while others show misorganization that manifests with disease progression.[Helmlinger et al., 2006] It is currently unknown why diseased rods would show clear changes in genomic organization, but could be related to the many gene expression changes observed in retinas of these mouse models.[Kim et al., 2016; Roger et al., 2014; Ruzycski et al., 2015]

To function properly, rods pack a modified cilia structure called the outer segment (OS) with Rhodopsin, the photopigment essential for capturing light.(Reviewed in [Arshavsky et al., 2002; Kefalov, 2012]) The rod sheds a portion of this structure every day, and thus must maintain an extraordinary high rate of expression of this transcript. (~2% of the total RNA within the cell- [Blackshaw et al., 2001]) The *Rho* locus thus must be highly epigenetically active, as evidenced by strong ATAC, histone, and cell type specific transcription factor signals.[Aldiri et al., 2017; Corbo et al., 2010; Hao et al., 2012; Mo et al., 2016]

In this work, we use the mouse photoreceptor to explore several principles of genomic organization. We show that *Rho* is located within the A compartment, and it co-localizes with other very epigenetically active areas across the chromosome. We also thoroughly characterize a new mouse model where we have inactivated *Rho*, and determine the effects on its localization and on expression of other genes within the A compartment.

## 5.4 Results

### 5.4.1 *Rho* displays strong interactions with local stretch of ‘A’ compartment DNA

We performed Circularized Chromosome Conformation Capture (4C) on P14 retinas of *WT* and *Nrl*<sup>-/-</sup> mice. NRL is a TF essential for rod development and gene expression, that when removed (*Nrl*<sup>-/-</sup>), the cells instead develop into functionally mature S-cones.[Mears et al., 2001] We selected a bait region just upstream of *Rho* in the middle of a number of known regulatory elements and within the larger H3K27Ac marked domain (Figure 5.1a). 4C raw read counts showed the expected strong enrichment for fragments of DNA in very close proximity to *Rho* (Supplemental Figure 5.1a). Raw data suggested that compared to *Nrl*<sup>-/-</sup>, *WT* retinas show more interaction with an actively marked and CTCF bound enhancer 30kb upstream (marked by \*\*). When *Rho* is inactive in *Nrl*<sup>-/-</sup> mice, there was increased interaction with a regulatory site further upstream also bound by CTCF as well as a broad relatively inactive region downstream of *Rho* (marked by \* and \*\*\* respectively).

We analyzed the 4C data with a novel scoring algorithm that ranks based on the consistency of local coverage compared to a background average (see Appendix 2 Methods section for more details). This analysis clearly identified a strong region of interacting DNA within 1Mb in both libraries that corresponded well with the local TAD boundaries of the closest related tissue analyzed by HiC, the mouse cortex (Figure 5.1b).[Dixon et al., 2012] Both libraries also displayed significant interactions with a larger local ~5Mb region of Chr6 surrounding *Rho* that extended beyond the boundaries of the local TAD. We wondered whether this larger region represented another level of genomic organization. By computing the Pearson Correlation of the HiC data, we determined that this extended region identified by our analysis included more than

50% of a linear stretch of DNA segregated to the A compartment (termed *cis*), but the interactions abruptly stopped before the flanking B compartment DNA segments (Figure 5.1c-d and Supplemental Figure 5.1b).

Our analysis also noted positive associations between *Rho* and other fragments of DNA beyond these boundaries. These *trans* interacting fragments were also predominantly located within A compartment chromatin (Figure 5.1c and Supplemental Figure 5.1b), but represented only a small subset of the rest of A compartment DNA (Figure 5.1e). Together, A Compartment DNA represented a significant majority of the DNA identified as interacting with *Rho* (Figure 5.1f).

#### **5.4.2 Distal interactions are dependent on *Rho* activation**

To test the effect of local epigenetic state on this genomic organization, we created two new mouse lines that lack either ~200bp of the *Rho* promoter (*PPR*<sup>-/-</sup>) or the Rhodopsin Enhancer Region (*RER*<sup>-/-</sup>) that lies 2kb upstream (Figure 5.2a). As expected, deletion of *PPR* completely abrogated expression of *Rho*, and largely inactivated the locus as measured by ATAC-seq (Figure 5.2). Surprisingly, even though previous data has suggested the *RER* is a very strong enhancer,[Corbo et al., 2010; Nie et al., 1996] its deletion had no effect on local epigenetic state or gene expression (Figure 5.2). Morphometric analysis showed that the *PPR*<sup>-/-</sup> retina displays dramatic shortening of the OS as early as P14, (Supplemental Figure 5.2c) but cells do not die until after P21,(Supplemental Figure 5.2e vs h) consistent with the previously published *Rho*<sup>-/-</sup> that removed exon 2 [Humphries et al., 1997; Toda et al., 1999]. This analysis also confirmed that loss of the *RER* has no effect on photoreceptor morphology (Supplemental Figure 5.2).

We chose four sites spanning Chromosome 6 to test and verify our 4C analysis. Regions 1 and 3 represent positive interactions as determined by 4C that lie 86 and ~9 Mb away respectively (Figure 5.3a). Both 4C peaks included a number of genes active in photoreceptors, but we chose to generate FISH probes using BACs that spanned the most epigenetically active. Site 1 includes a cluster of very highly expressed *miRNAs* (*mir182/183/96*) and the lncRNA *Rncr4*, the loss of which cause severe retinal defects [Krol et al., 2015] (Supplemental Figure 5.3b). Site 3 includes many regulatory elements and active genes, but we chose to focus on *Ptms* which has a very strong enrichment of H3K27Ac (Supplemental Figure 5.3d). Sites 2 and 4 both represent control fragments that were not identified by 4C analysis. Site 2 includes the *Tcrb* locus that lies within a distal B compartment stretch of Chromosome 6, and is not near any retinal active regulatory sites (Supplemental Figure 5.3c). Finally, site 4 includes the cone photoreceptor expressed *Pde6h* gene. This gene is classified within the A compartment, but is inactive in rods (Supplemental Figure 5.3e).

We performed FISH and measured the distance between pairs of probes in P14 retinal sections of a number of mouse models. In *WT* and *RER*<sup>-/-</sup> retinas, the *Rho* locus is epigenetically and functionally normal. These were compared to mutants in which *Rho* is inactive and not expressed including the *PPR*<sup>-/-</sup> and models lacking transcription factors important in rod differentiation and gene expression (*Nrl*<sup>-/-</sup> and *Crx*<sup>-/-</sup>).

FISH results first and foremost confirmed our novel 4C analysis (Figure 5.3b-e). We noted strong co-localization between *Rho* and regions 1 (*mir182*) and 3 (*Ptms*), but not with regions 2 (*Tcrb*) and 4 (*Pde6h*) in *WT* rods. This pattern was the same when analyzed in the phenotypically normal *RER*<sup>-/-</sup>, but positive interactions were lost in *PPR*<sup>-/-</sup>, *Nrl*<sup>-/-</sup> and *Crx*<sup>-/-</sup> retinas.

We wondered whether regions 1 and 3 that showed strong interactions with *Rho* in *WT* and *RER*<sup>-/-</sup> rods would also show interactions independently with each other. Region 1 displayed a higher rate of association with region 3 than with inactive region 2, even though 2 is much closer on the linear genome. Importantly, this interaction was unaffected in *PPR*<sup>-/-</sup> or either TF mutant (Figure 5.3f-g). Together, these data support our novel 4C data analysis, and suggest that the long-range interactions of *Rho* are dependent upon its activity or chromatin state, but that *Rho* is not responsible for ‘recruiting’ or establishing the state of other genes.

### 5.4.3 Distal interactions enriched for highly active chromatin

To better understand the nature of the regions identified as interacting with *Rho*, we analyzed the ChromHMM state within P14 retina.[Aldiri et al., 2017] As expected, the cortex defined A Compartment was enriched for active chromatin and facultative heterochromatin. The B compartment was specifically enriched for Class 9 (empty) which likely reflects the constitutively repressed nature of such DNA (Figure 5.4a). This stark distinction suggests that mouse cortex and rod photoreceptors have a very similar compartmental organization even though their nuclei are organized very differently at the microscopic scale.

We also compared the regions defined by 4C data as interacting with *Rho*. These regions showed even stronger enrichment for active HMM states (Figure 5.4a). Further subdivision of these interactions as *cis* or *trans* displayed a stronger enrichment for highly active promoter and enhancer chromatin within *trans* loci (Supplemental Figure 5.4a).

We chose several aspects of this data to explore further. As *trans* interacting regions displayed stronger enrichment for the most active HMM states, we wondered whether a new class of very active regions of the genome would be enriched within the interacting sites. These

‘Super Enhancers’ (SEs) [Hnisz et al., 2013] were indeed highly enriched within *trans* contacted regions (Figure 5.4b). In fact, *Rho*, *Ptms*, and *mir182/183/96* are all specifically contained within SEs. We also found significant enrichment at both *cis* and *trans* interacting regions for binding of the photoreceptor master TF CRX and regions termed insulators (CTCF bound), but no enrichment for repressed chromatin (Figure 5.4c-e). We wondered whether there was any specificity for Dependent or Independent CRX sites (see Chapter 4 of this dissertation). *Trans* interacting fragments showed very significant enrichment for CRX Group A (promoter proximal) Dependent and Independent sites, but interestingly only Independent Group C (active enhancer) sites (Supplemental Figure 5.3b-e). Together, these comparisons suggest that *Rho* interacts with regions with dense CRX and CTCF binding sites and fragments that are also highly epigenetically active.

#### **5.4.4 Inactivation of *Rho* affects expression of other genes**

As *Rho* is the most transcriptionally active gene in the rod photoreceptor, we wondered what effect its inactivation would have on other genes. RNA-seq data profiling P14 *WT* and *PPR*<sup>-/-</sup> retinas showed very similar gene expression between the two models, with few genes other than *Rho* affected more than 2-fold (Figure 5.5a, *Rho* marked in red). We compared the expression of annotated transcripts identified as interacting with *Rho* with all transcripts on Chromosome 6 and found a modest but significant increase in gene expression. We binned the genes into those within the *cis* A Compartment vs *trans* interactions, and found that this signal largely originated from genes within the *cis* Compartment (Figure 5.5b). While the A Compartment as a whole showed no significant difference from the B Compartment (Figure 5.5c), we further binned this data to the 11 linear segments of the A Compartment (A1-A11).



Contacted genes within A2 and A5 showed significant changes from other genes within that segment as well as the chromosome on a whole, but the number of genes within each was very small (Supplemental Figure 5.5).

## 5.5 Discussion

### 5.5.1 General principles of genomic organization

By analyzing the organization of the highly active *Rho* gene within the rod photoreceptor nucleus we have uncovered several novel insights into genomic organization. First, although the rod nucleus is inverted and microscopically differently from virtually every other cell type in the body, its genome still organized at the molecular scale into a very similar TAD and A/B Compartment structures. This is both surprising and reassuring as even though it varies so dramatically, the forces working to maintain these structures are still in play and likely important.

We have shown through comparison of the interactions of *Rho* with other active regions of the genome in *WT* and mutant retinas, that this organization depends on the activity of the *Rho* locus. Even though we predict in the *PPR*<sup>-/-</sup> retina that CRX and NRL could still bind enhancer elements and other nearby regulatory sites, the activity of the locus is necessary for the co-localization. We showed the clear loss of this activity by ATAC-seq, but future ChIP experiments for H3K27Ac would be necessary to prove that the broad ‘Super Enhancer’ domain is in fact disrupted.

The discovery that the loss of the *RER* had no effect on *Rho* expression was very surprising. This enhancer has been studied extensively previously and it has proven to be very active at increasing transcription both *in vitro* and *in vivo*. [Corbo et al., 2010; Nie et al., 1996] However, as browser shots display, (Supplemental Figure 5.1 and Figure 5.1) there are many other regulatory elements that may compensate. Future work will be necessary to test the role of these elements individually and in combination (see preliminary results in Appendix 1).

### **5.5.2 *Rho* interacts with other active regions of its chromosome**

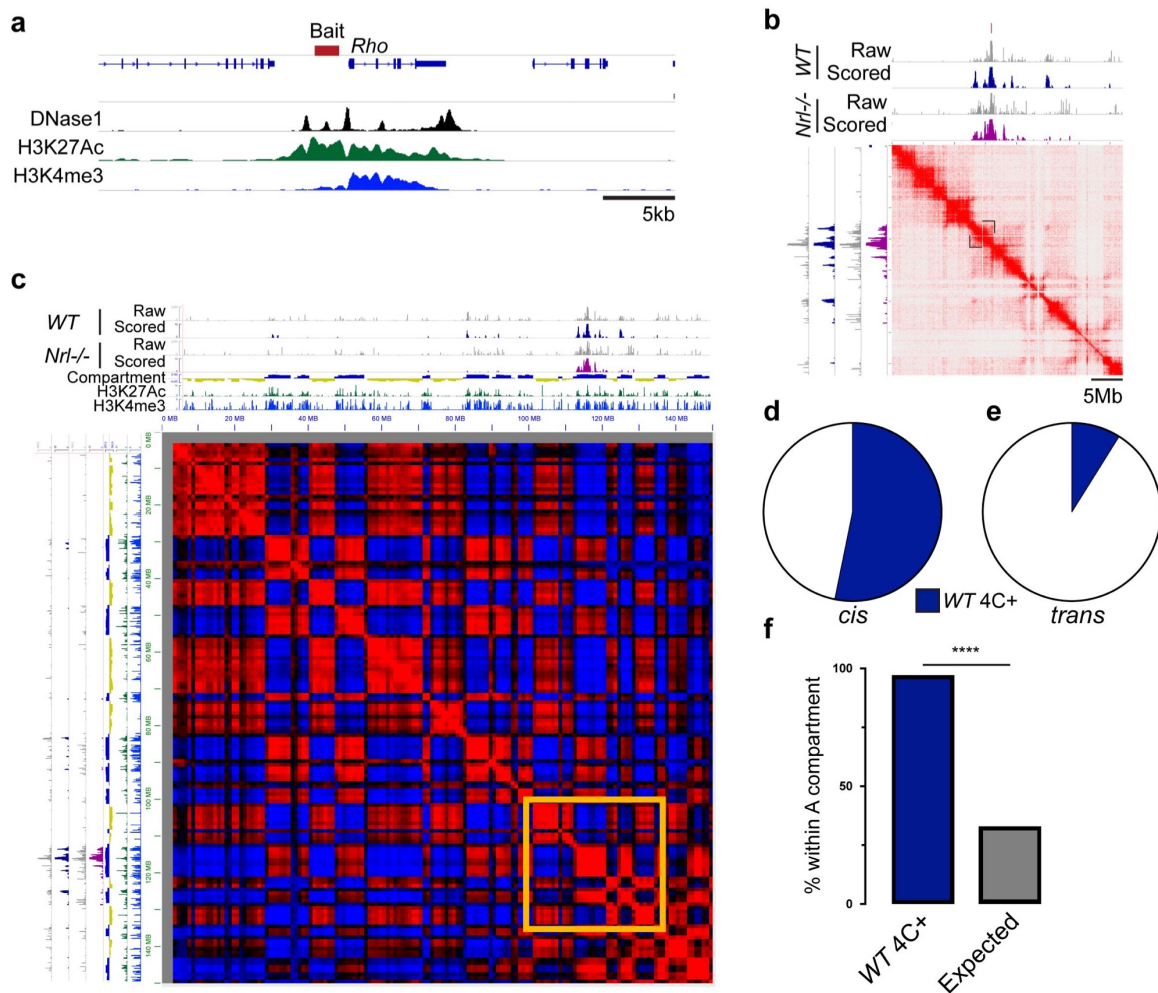
By analyzing the nature of the regions of Chr6 identified as interacting with *Rho* in *trans*, we determined that *Rho* is in contact with regions of DNA that are also very active. As depicted in the model (Figure 5.6a), these regions are highly enriched for Super Enhancers, CTCF binding sites, and regulatory elements bound by photoreceptor master TF CRX. This analysis of the 4C data used a relatively broad window to score, further analyses comparing smaller windows to determine if more discrete signals are more relevant or selective for these important genomic regulatory elements.

### **5.5.3 Promoters do not act independently**

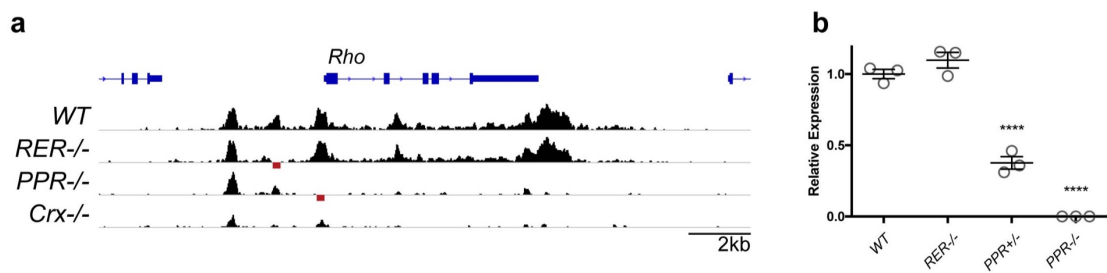
Finally, by comparing the transcriptome of the *PPR*<sup>-/-</sup> mutant with *WT* litter mate controls, we have determined that the loss of such a strong promoter as that of *Rho* can affect the expression of other genes. These effects were extremely modest, but the significant increase in expression of genes within the local A compartment suggests that the machinery that would have been active at the *Rho* promoter was instead available to be used at other promoters in proximity to *Rho* (Figure 5.6b). This finding is not surprising, as past experiments have previously shown the association of certain active genes with nuclear speckles.[Brown et al., 2008; Khanna et al., 2014; Lamond & Spector, 2003]

Together, our data supports the model put forward in Figure 5.6 where genes, like *Rho*, are within physical proximity with other regions of the chromosome that are most similar epigenetically. Especially on a local scale, the activity of a promoter is actually in constant balance with every other promoter. When a strong promoter such as *Rho* is removed, the machinery can be mis-used by others. Future work will be necessary to test this in other cell

types that display such strong expression of individual genes and to determine any functional consequences of these small changes to gene expression.

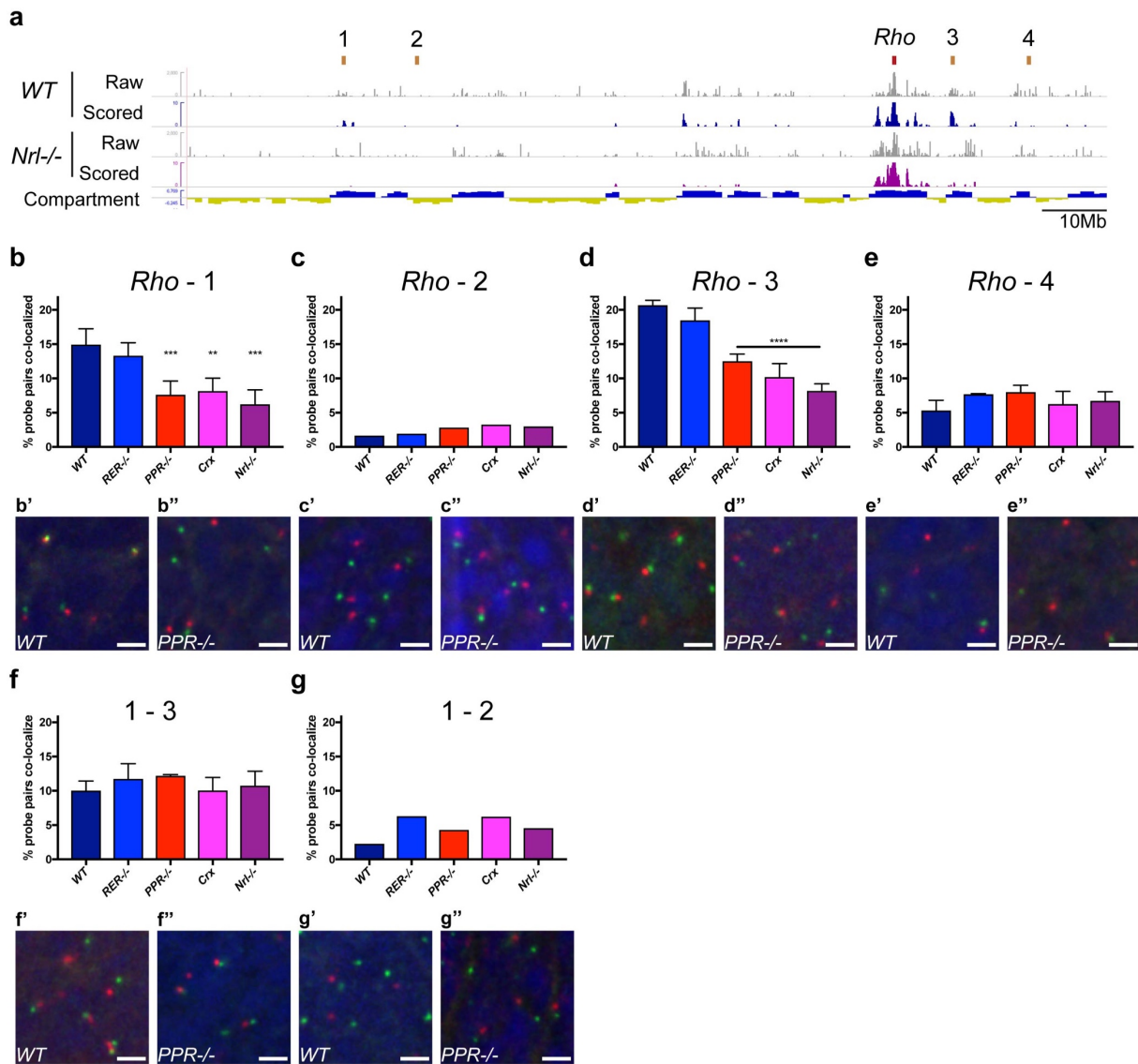


**Figure 5.1: *Rho* displays strong interactions with local DNA in both *WT* and *Nrl*<sup>-/-</sup> retinas** (a) Browser tracks display relevant epigenetic marks in the P14 *WT* retina and the fragment of DNA used at the 4C bait (red). Raw (gray) and analyzed (blue and purple) 4C data display strong association of *Rho* with local TAD structure as defined in cortex (bracketed in matrix in black) (b), and with larger A compartment DNA when data is analyzed as Pearson Correlation matrix (c) (area enlarged in Supplemental Figure 5.1b highlighted in orange). Proportion of *cis* (d) and *trans* (e) linear segments of A Compartment that are identified as interacting with *Rho*. (f) Proportion of interacting regions within A compartment compared to permutation of fragments (see Appendix 2 Methods) (\*\*\*\*  $p < 0.0001$ ; Fisher's Exact Test).

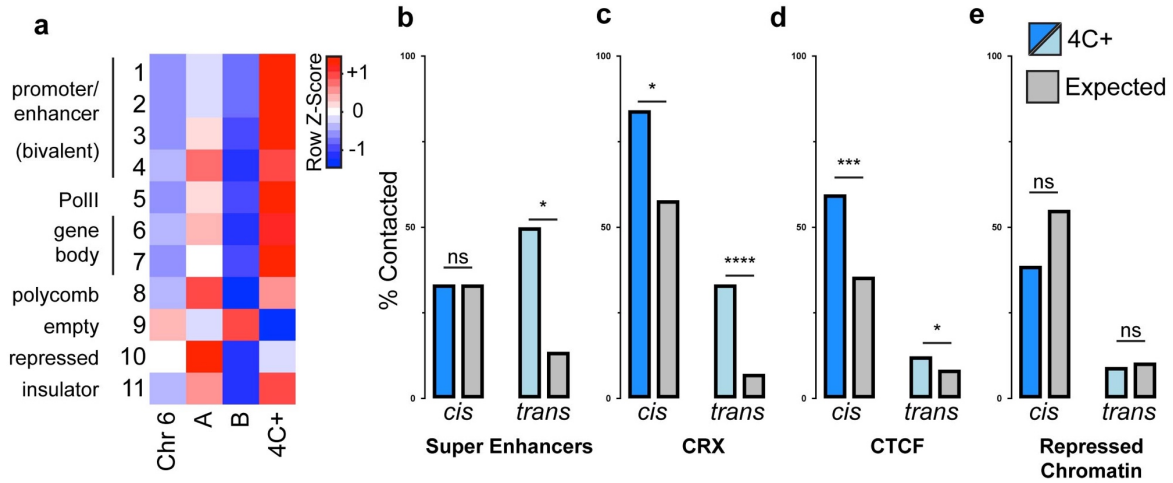


**Figure 5.2: Disruption of the *Rho* promoter affects local epigenetic landscape**

(a) ATAC-seq of P14 retinas displays epigenetic accessibility changes in mutants. Red marks deleted sequences in the *RER*<sup>-/-</sup> and *PPR*<sup>-/-</sup> strains. (b) ddPCR analysis of *Rho* expression (*Actb* as stable reference) in P14 retinas. (\*\*\*\*  $p < 0.0001$ ; ANOVA with Tukey's multiple comparison test)



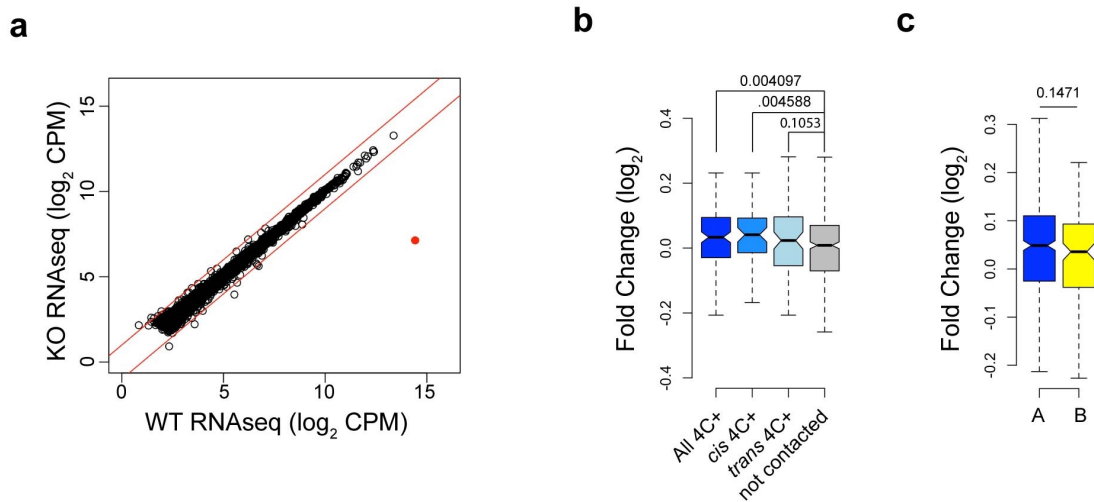
**Figure 5.3: FISH confirms *Rho* interactions with distal fragments of its chromosome**  
**(a)** Browser track displays raw and scored 4C data from *WT* and *Nrl*<sup>-/-</sup> retinas for all of Chromosome 6. A/B Compartments are labelled by analysis of cortex HiC data (blue and yellow respectively). Regions selected for further analysis are marked '1-4' above track. **(b-g)** Analysis of FISH data quantified as the proportion of probe-pairs co-localized (centers of mass within 0.5um). Representative images of *WT* (**b'-g'**) and *PPR*<sup>-/-</sup> (**b''-g''**) are displayed below associated data. (mean + SEM, \*\*p<0.01, \*\*\*\*p<0.0001; ANOVA with Tukey's multiple comparison test)



**Figure 5.4: *Rho* trans interactions are enriched at active segments of the genome**

(a) Heatmap displays relative enrichment of P14 retina ChromHMM state [Aldiri et al., 2017] in all of Chromosome 6 compared to Compartments A/B and those regions identified by 4C as interacting with *Rho*. (b-e) Bargraphs show quantification of the proportion of Super Enhancers, CRX binding sites, CTCF binding sites (ChromHMM state 11), and repressed chromatin (ChromHMM state 10) covered by *cis* and *trans* 4C positive regions. Expected values represent permutation of the marks within that *cis* or *trans* genomic region (see Appendix 2 Methods). (\* $p < 0.05$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ ; Fisher's Exact Test)

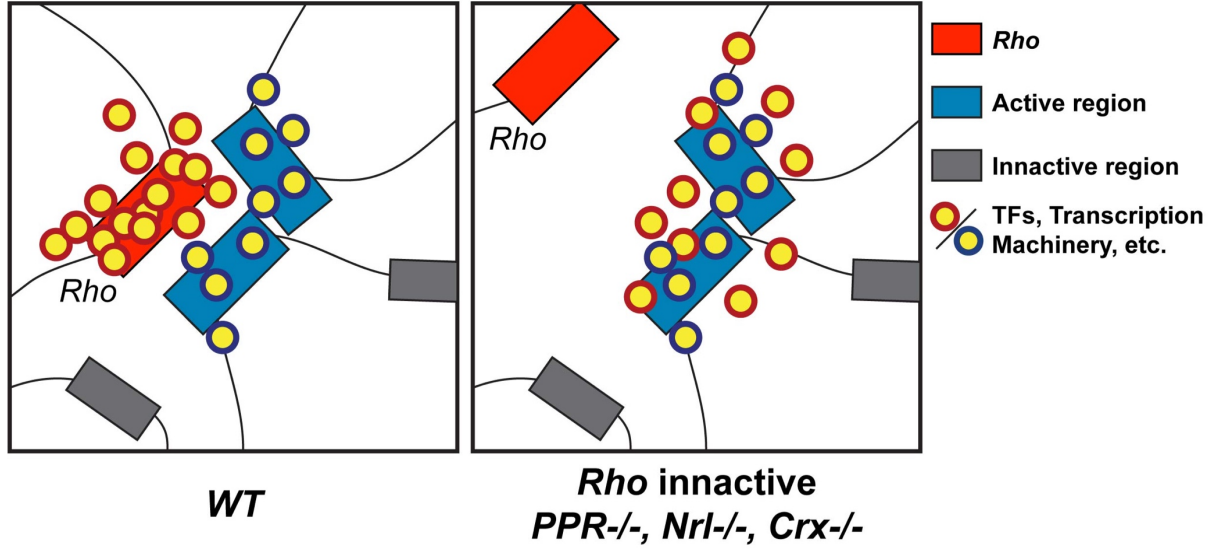




**Figure 5.5: Loss of *Rho* expression affects other genes that interact with *Rho***

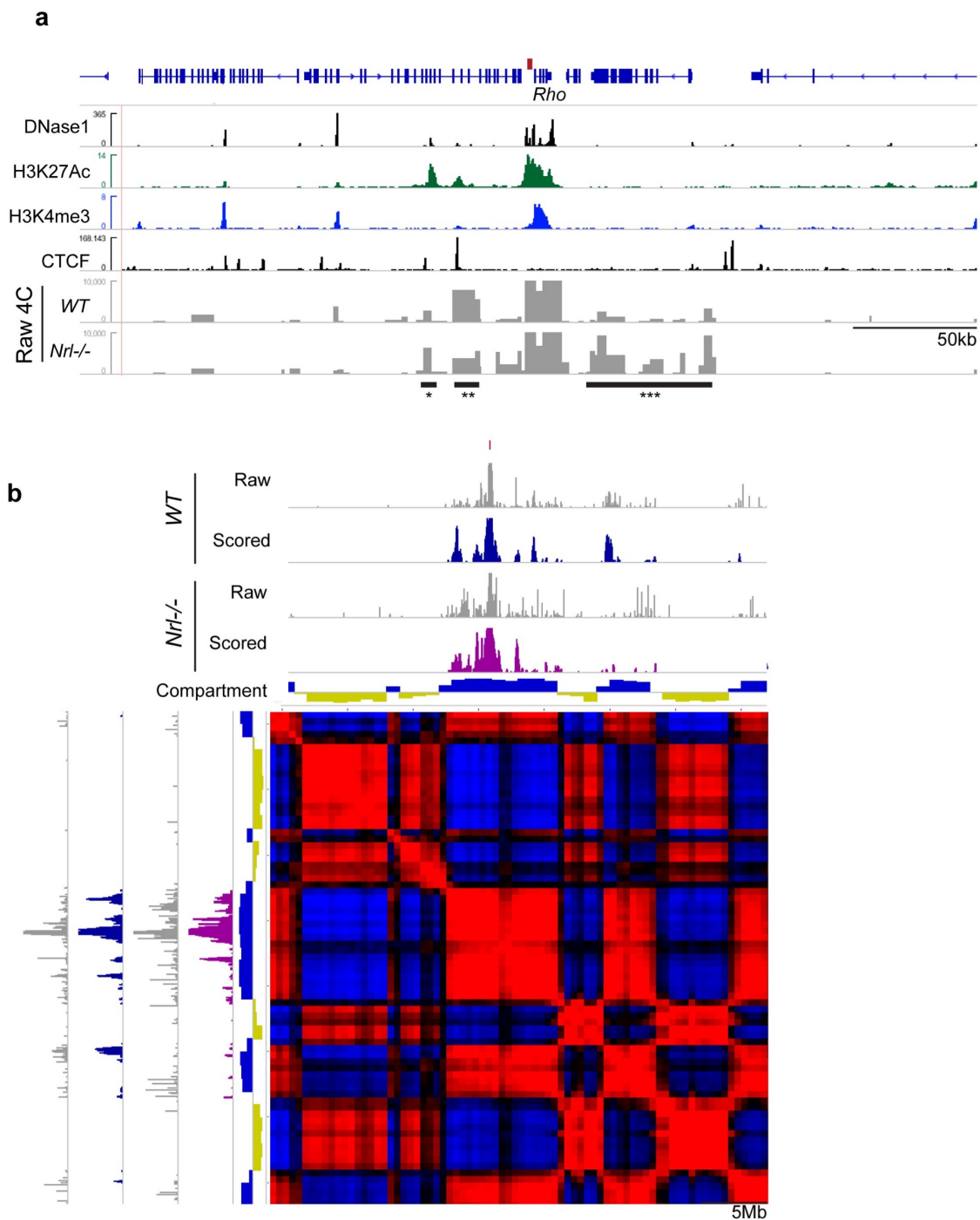
(a) Scatterplot displays very consistent expression of all genes except *Rho* (highlighted in red) in *WT* and *PPR*<sup>-/-</sup> retinas at P14 as measured by RNA-seq. Red lines denote 2-fold change.

Analysis of fold-change of genes (*PPR*<sup>-/-</sup> vs *WT*) that interact with *Rho* by 4C (b) and all genes within Chromosome 6 A and B Compartments (c). (Wilcoxon Rank Sum test)



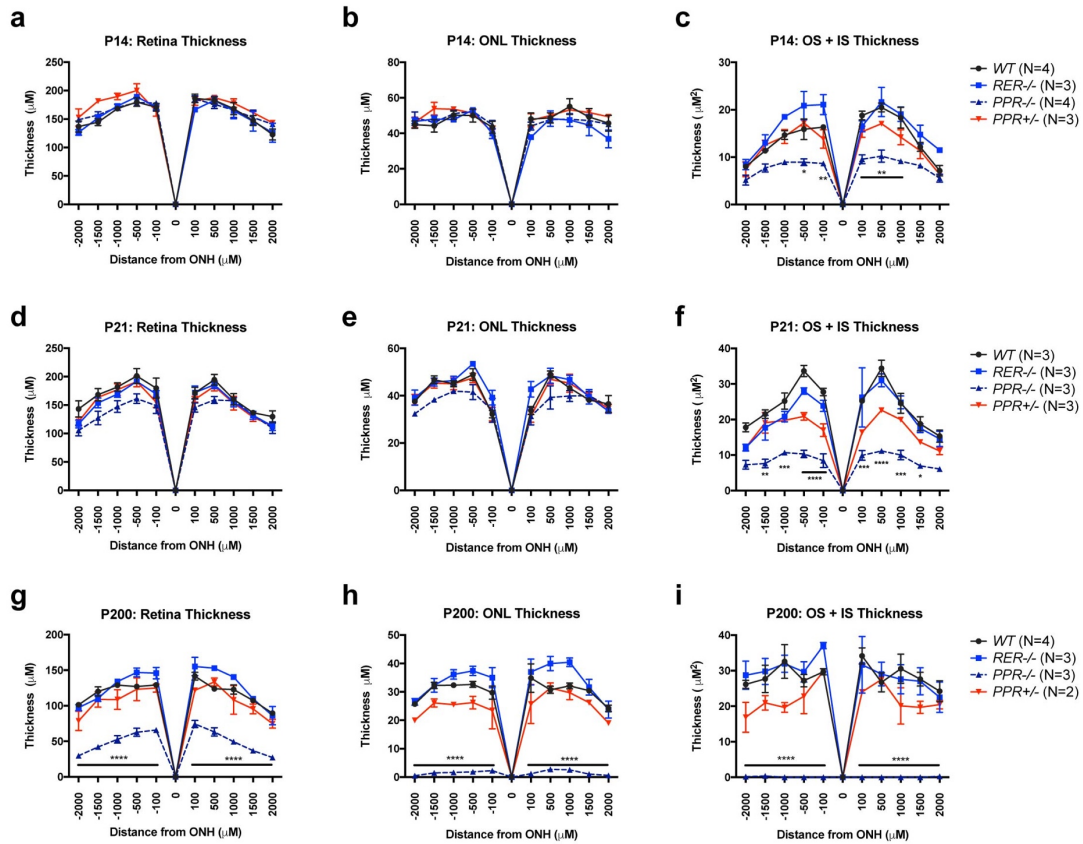
**Figure 5.6: Model of *Rho* genomic interactions and effects of loss**

Model describes local interactions of *Rho* in *WT* retina that are lost when *Rho* is inactivated and the competition between these genes for transcription and activity associated proteins.



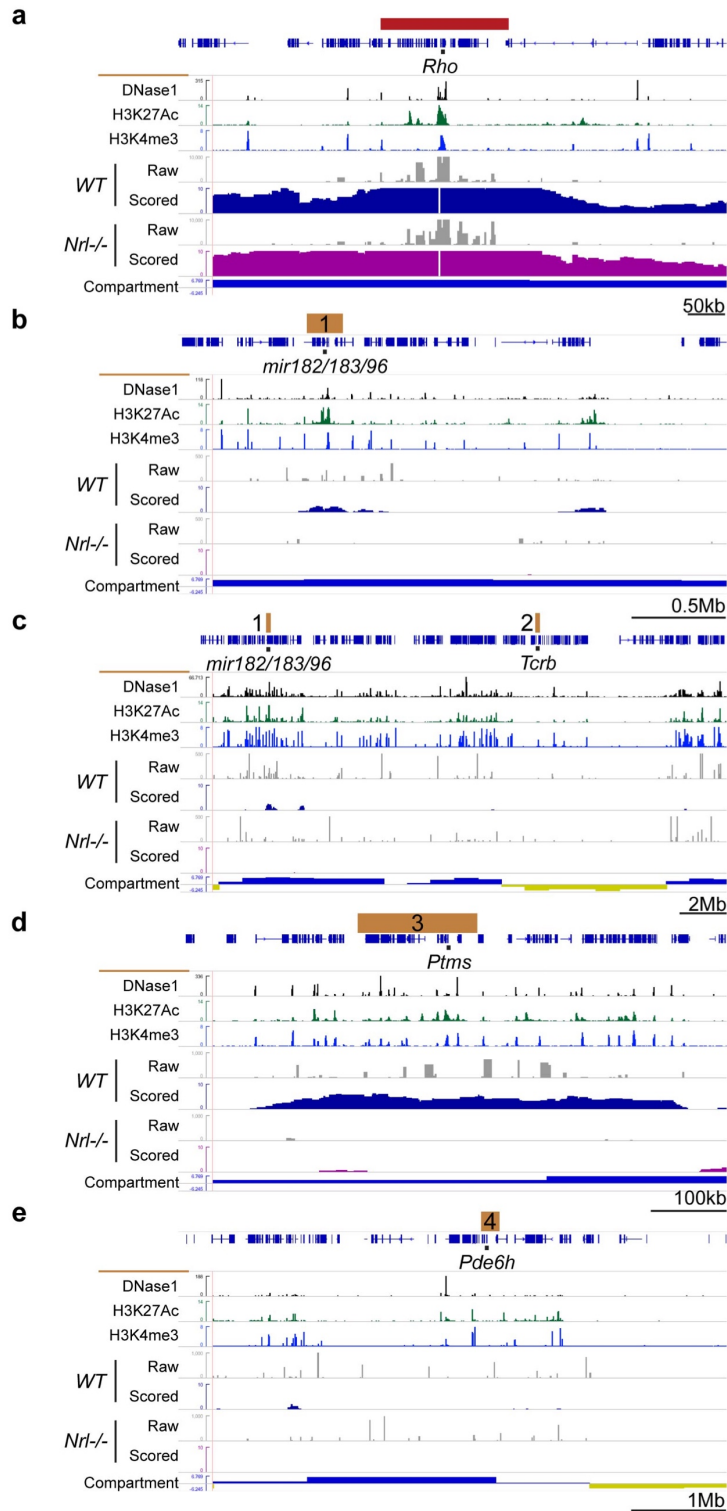
**Supplemental Figure 5.1: 4C detects interactions of *Rho* within its TAD and local region of A compartment**

(a) Browser tracks display relevant epigenetic marks in the P14 *WT* retina and the raw 4C data from *WT* and *Nrl*<sup>-/-</sup> retina. (b) Raw and scored 4C data from *WT* and *Nrl*<sup>-/-</sup> retina with A/B compartment designation and Pearson Correlation matrix of cortex HiC.



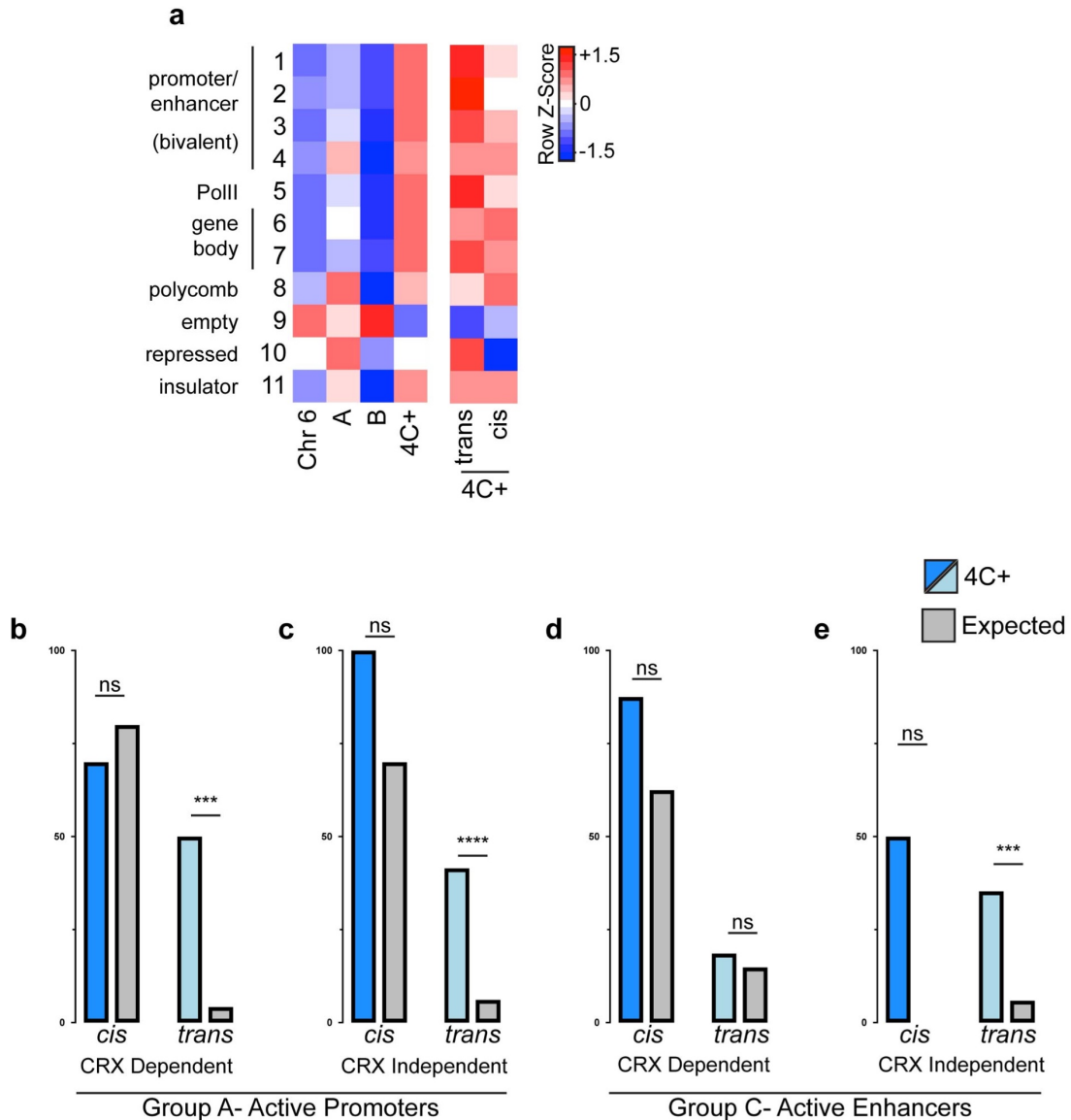
**Supplemental Figure 5.2: Morphometric analysis of *PPR*<sup>-/-</sup> and *RER*<sup>-/-</sup> retinas**

Measurement of thickness of indicated layers at P14 (a-c), P21 (d-f), and P200 (g-i) in *WT*, *RER*<sup>-/-</sup>, *PPR*<sup>-/-</sup>, and *PPR*<sup>+/-</sup> retinas. (mean  $\pm$  SEM,  $n \geq 3$ ; \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ ; 2-Way Anova with Tukey's Multiple Comparisons Test).



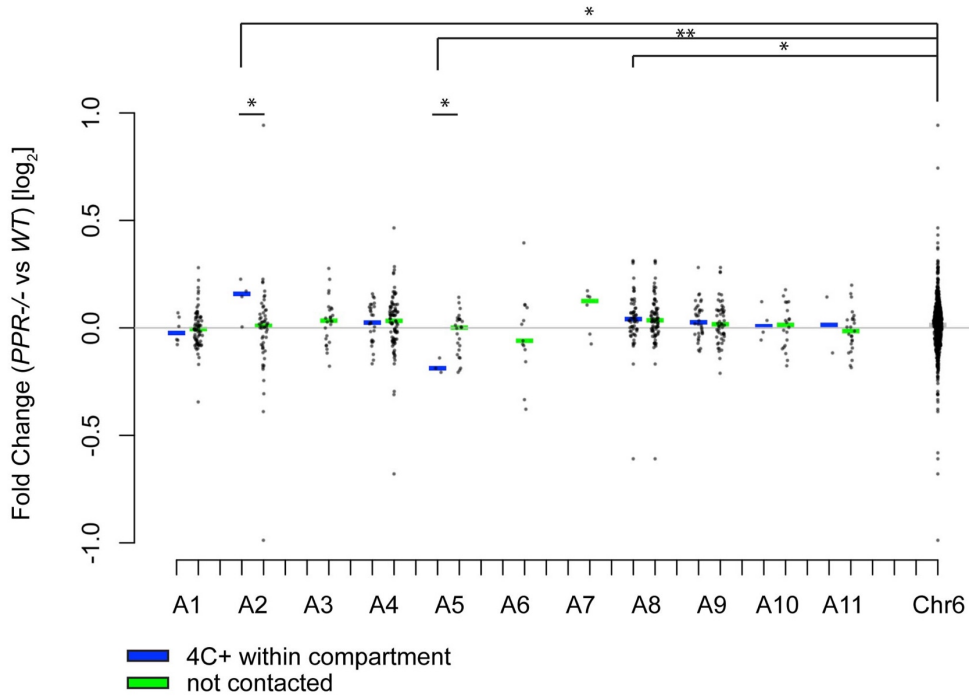
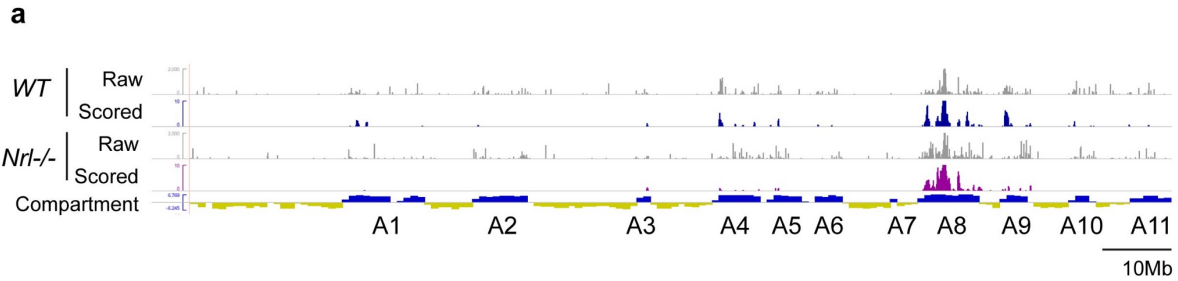
**Supplemental Figure 5.3: BAC selection for FISH analysis**

Browser shots display relevant epigenetic data and raw and scored 4C data surrounding *Rho* (a) and regions 1-4 (b-e) selected for FISH experiments. Location of BAC is indicated by red / orange bar above gene track.



**Supplemental Figure 5.4: *Trans* interacting regions show very strong enrichment for active epigenetic state**

(a) Heatmap displays relative enrichment of P14 retina ChromHMM state [Aldiri et al., 2017] in all of Chromosome 6 compared to Compartments A/B and *cis* and *trans* regions identified by 4C as interacting with *Rho*. (b-e) Bargraphs show quantification of the proportion of types of CRX sites (see Chapter 4 of this dissertation) covered by *cis* and *trans* interactions. Expected values represent permutation of the marks within that *cis* or *trans* genomic region (see Appendix 2 Methods). (\*\*\*)  $p < 0.001$ , (\*\*\*\*)  $p < 0.0001$ ; Fisher's Exact Test)



**Supplemental Figure 5.5: A Compartment segments show differential effects upon loss of *Rho* activity**

(a) Browser track displaying the locations of the 11 linear segments of Chromosome 6 that are designated as A compartment DNA (blue vs yellow compartment track). (b) Plot displays the Fold Change ( $\log_2$ ) between *WT* and *PPR*<sup>-/-</sup> retina of individual genes and the median of 4C+ genes (blue) and those not contacted (green) located within each of the segments. Each 4C+ set is also compared to the changes in expression of all Chromosome 6 genes. (black dots represent individual genes; blue, green and gray bars represent median; \*  $p < 0.05$ , \*\*  $p < 0.01$ ; Wilcoxon Rank Sum Test)

## 5.6 References

- Aldiri, I., Xu, B., Wang, L., Chen, X., Hiler, D., Griffiths, L., Valentine, M., et al. (2017). The Dynamic Epigenetic Landscape of the Retina During Development, Reprogramming, and Tumorigenesis. *Neuron*, *94*, 550–568.
- Arshavsky, V. Y., Lamb, T. D., & Pugh, E. N. (2002). G Proteins and Phototransduction. *Annual Review Of Physiology*, *64*, 153–187.
- Blackshaw, S., Fraioli, R. E., Furukawa, T., & Cepko, C. L. (2001). Comprehensive analysis of photoreceptor gene expression and the identification of candidate retinal disease genes. *Cell*, *107*, 579–589.
- Bonev, B., & Cavalli, G. (2016). Organization and function of the 3D genome. *Nature Reviews Genetics*, *17*, 772–772.
- Brown, J. M., Green, J., das Neves, R. P., Wallace, H. a C., Smith, A. J. H., Hughes, J., Gray, N., et al. (2008). Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. *The Journal Of Cell Biology*, *182*, 1083–97.
- Corbo, J. C., & Cepko, C. L. (2005). A hybrid photoreceptor expressing both rod and cone genes in a mouse model of enhanced S-cone syndrome. *PLoS Genetics*, *1*, e11.
- Corbo, J. C., Lawrence, K. a, Karlstetter, M., Myers, C. a, Abdelaziz, M., Dirkes, W., Weigelt, K., et al. (2010). CRX ChIP-seq reveals the cis-regulatory architecture of mouse photoreceptors. *Genome Research*, *20*, 1512–25.
- Dixon, J. R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-bourget, J. E., Lee, A. Y., Ye, Z., et al. (2015). Chromatin architecture reorganization during stem cell differentiation. *Nature*, *518*, 331–336.
- Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., et al. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, *485*, 376–380.
- Hao, H., Kim, D. S., Klocke, B., Johnson, K. R., Cui, K., Gotoh, N., Zang, C., et al. (2012). Transcriptional Regulation of Rod Photoreceptor Homeostasis Revealed by In Vivo NRL Targetome Analysis. *PLoS Genetics*, *8*, e1002649.
- Helmlinger, D., Hardy, S., Abou-Sleymane, G., Eberlin, A., Bowman, A. B., Gansmüller, A., Picaud, S., et al. (2006). Glutamine-expanded ataxin-7 alters TFIIIC/STAGA recruitment and chromatin structure leading to photoreceptor dysfunction. *PLoS Biology*, *4*, e67.
- Hennig, A. K., Peng, G.-H., & Chen, S. (2013). Transcription Coactivators p300 and CBP Are Necessary for Photoreceptor-Specific Chromatin Organization and Gene Expression. *PloS One*, *8*, e69721.
- Hnisz, D., Abraham, B. J., Lee, T. I., Lau, A., Saint-André, V., Sigova, A. a, Hoke, H. a, et al. (2013). Super-Enhancers in the Control of Cell Identity and Disease. *Cell*, *155*, 934–947.
- Hughes, A. E. O., Enright, J. M., Myers, C. A., Shen, S. Q., & Corbo, J. C. (2017). Cell Type-Specific Epigenomic Analysis Reveals a Uniquely Closed Chromatin Architecture in Mouse Rod Photoreceptors. *Scientific Reports*, *7*, 43184.
- Humphries, M., Rancourt, D., Farrar, G. J., Kenna, P., Hazel, M., Bush, R., Sieving, P. A., et al. (1997). Retinopathy induced in mice by targeted disruption of the rhodopsin gene. *Nature Genetics*, *15*, 216–219.
- Jeon, C. J., Strettoi, E., & Masland, R. H. (1998). The major cell populations of the mouse retina. *The*



- Journal Of Neuroscience*, 18, 8936–46.
- Kefalov, V. J. (2012). Rod and cone visual pigments and phototransduction through pharmacological, genetic, and physiological approaches. *Journal Of Biological Chemistry*, 287, 1635–1641.
- Khanna, N., Hu, Y., & Belmont, A. S. (2014). Hsp70 transgene directed motion to nuclear speckles facilitates heat shock activation. *Current Biology*, 24, 1138–1144.
- Kim, J.-W., Yang, H.-J., Brooks, M. J., Zelinger, L., Karakulah, G., Gotoh, N., Boleda, A., et al. (2016). NRL-Regulated Transcriptome Dynamics of Developing Rod Photoreceptors. *Cell Reports*, 17, 2460–2473.
- Krol, J., Krol, I., Alvarez, C. P. P., Fiscella, M., Hierlemann, A., Roska, B., & Filipowicz, W. (2015). A network comprising short and long noncoding RNAs and RNA helicase controls mouse retina architecture. *Nature Communications*, 6, 7305.
- Lamond, A. I., & Spector, D. L. (2003). Nuclear speckles: a model for nuclear organelles. *Nature Reviews. Molecular Cell Biology*, 4, 605–12.
- Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, 326, 289–93.
- Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., et al. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*, 161, 1202–1214.
- Mears, A., Kondo, M., Swain, P. K., Takada, Y., Bush, R. A., Saunders, T. L., Sieving, P. A., et al. (2001). Nrl is required for rod photoreceptor development. *Nature Genetics*, 29, 447–52.
- Mo, A., Luo, C., Davis, F. P., Mukamel, E. A., Henry, G. L., Nery, J. R., Urich, M. A., et al. (2016). Epigenomic landscapes of retinal rods and cones. *ELife*, 5, 1–29.
- Narendra, V., Bulaji, M., Dekker, J., Mazzoni, E. O., & Reinberg, D. (2016). CTCF-mediated topological boundaries during development foster appropriate gene regulation. *Genes And Development*, 30, 2657–2662.
- Narendra, V., Rocha, P. P., An, D., Raviram, R., Skok, J. A., Mazzoni, E. O., & Reinberg, D. (2015). CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science*, 347, 1017–1021.
- Nie, Z., Chen, S., Kumar, R., & Zack, D. J. (1996). RER, an evolutionarily conserved sequence upstream of the rhodopsin gene, has enhancer activity. *Journal Of Biological Chemistry*, 271, 2667–2675.
- Peng, G., & Chen, S. (2011). Active opsin loci adopt intrachromosomal loops that depend on the photoreceptor transcription factor network. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, 108, 17821–6.
- Peng, G., & Chen, S. (2012). Revealing Looping Organization of Mammalian Photoreceptor Genes Using Chromosome Capture (3C) Assays. *Retinal Development: Methods And Protocols*, 884, 305–318.
- Rao, S. S. P., Huang, S.-C., Hilaire, B. G. S., Engreitz, J. M., Perez, E. M., Kieffer-Kwon, K.-R., Sanborn, A. L., et al. (2017). Cohesin Loss Eliminates All Loop Domains. *Cell*, 171, 305–320.
- Rao, S. S. P., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., Sanborn, A. L., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell*, 159, 1–16.

- Roger, J. E., Hiriyanna, A., Gotoh, N., Hao, H., Cheng, D. F., Ratnapriya, R., Kautzmann, M. I., et al. (2014). OTX2 loss causes rod differentiation defect in CRX-associated congenital blindness. *The Journal Of Clinical Investigation*, *124*, 631–643.
- Ruzycki, P. A., Tran, N. M., Kefalov, V. J., Kolesnikov, A. V., & Chen, S. (2015). Graded gene expression changes determine phenotype severity in mouse models of CRX-associated retinopathies. *Genome Biology*, *16*, 171.
- Schmitt, A. D., Hu, M., Jung, I., Xu, Z., Qiu, Y., Tan, C. L., Li, Y., et al. (2016). A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. *Cell Rep*, *17*, 2042–2059.
- Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N. A., et al. (2017). Two independent modes of chromatin organization revealed by cohesin removal. *Nature*, *551*, 51–56.
- Solovei, I., Kreysing, M., Lanctôt, C., Kösem, S., Peichl, L., Cremer, T., Guck, J., et al. (2009). Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. *Cell*, *137*, 356–68.
- Solovei, I., Wang, A. S., Thanisch, K., Schmidt, C. S., Krebs, S., Zwerger, M., Cohen, T. V., et al. (2013). LBR and Lamin A/C Sequentially Tether Peripheral Heterochromatin and Inversely Regulate Differentiation. *Cell*, *152*, 584–598.
- Toda, K., Bush, R. a, Humphries, P., & Sieving, P. a. (1999). The electroretinogram of the rhodopsin knockout mouse. *Visual Neuroscience*, *16*, 391–8.
- Tran, N. M., Zhang, A., Zhang, X., Huecker, J. B., Hennig, A. K., & Chen, S. (2014). Mechanistically Distinct Mouse Models for CRX-Associated Retinopathy. *PLoS Genetics*, *10*, e1004111.
- Young, R. W. (1985). Cell differentiation in the retina of the mouse. *The Anatomical Record*, *212*, 199–205.

# **Chapter 6**

## **Conclusions and Future Directions**

In this dissertation I have thoroughly investigated and demonstrated novel insights into the following: The disease inheritance and mechanism of a new model of CRX-associated disease (Chapter 2), the principles of gene mis-regulation across several classes of CRX-associated retinal diseases (Chapter 3), the role of CRX in establishing the epigenome of a mature photoreceptor and its specific function at individual regulatory elements (Chapter 4), and the organization of the genome within the rod photoreceptor nucleus and the regulating principles of this structure (Chapter 5).

## **6.1 A cohesive understanding of CRX related diseases**

In Chapters 2 and 3, I expanded our knowledge of the diversity of and shared mechanism between CRX disease models.[Ruzycki et al., 2017, 2015] Mutations to CRX result in a broad range of phenotypic severity in human patients, the diversity of which has been carefully modelled and characterized in several knock-in or naturally occurring mutant mice.[Roger et al., 2014; Tran et al., 2014] By analyzing and comparing the transcriptomes of several *Crx* mutant mouse models, I determined that the severity of the phenotype directly corresponds with the effects on transcription of a core set of rod photoreceptor genes. The classification of the *L253X* mutant as a dominant Class III mutant shed new light on this phenotypic continuum, and further supports the understanding that very small changes to the transcriptome can have clinically relevant effects on photoreceptor development and maintenance.

This research serves as the basis for several new lines of investigation. First, the classification of *L253X* as a Class III mutant expands our understanding of dominant CRX disease. Class III mutants are categorized by a premature termination codon that causes hyperstability of the mutant transcript. We discovered that the late *L253X* mutation does cause an

increase in transcript stability, but to a lesser degree than that observed in the more severe *E168d2* model. The only difference between these mutant transcripts is the length of the apparent UTR, suggesting that distinct elements additively contribute to hyperstability. High throughput analysis of deletion libraries could uncover the mechanism of this hyperstability. Experiments testing the effects of variants of the *Crx* UTR on the stability of a transcript should elucidate the RNA encoded stabilizing elements or the structural form of the folded RNA structure that makes this RNA resistant to degradation.

Second, this data suggests that even though patients may have very different mutations and phenotypes, a potential therapeutic ‘target’ or strategy can be similar for all *CRX* related disease. Gene therapy now offers a viable option to deliver a copy of normal *CRX* to mutant photoreceptors *in vivo*. Our findings demonstrate that very small perturbations to the target gene set can have dramatic effects on phenotype which suggests that even inefficient gene therapy treatments could benefit Class I and less dramatic Class III mutants like the *L253X* model. Future efforts to treat severe Class III mutations may require a combined augmentation based gene-therapy with concurrent knockdown of the hyperstable mutant transcript.

## **6.2 A comprehensive genomic map of CRX regulatory activity**

Previous efforts to understand the *in vivo* molecular mechanism of CRX have focused primarily on the transcriptional changes in mutant animals. Studies have shown that genes that lose expression are strongly enriched for those that are near CRX bound regulatory elements. However, no study had asked the reciprocal question and discovered that there are actually more CRX bound regulatory elements near genes that show no expression changes. In Chapter 4, I

discovered that CRX binds within a variety of chromatin environments, and is only responsible for the activity of a small subset (~1/3) of these sites. These ‘CRX Dependent’ sites represent those that are near genes that lose expression in the *Crx* mutant retinas. By comparing these with the ‘CRX Independent’ sites that show no change in activity upon the loss of CRX, I discovered several key factors including developmental activation, conservation, and motif enrichment that distinguish a CRX Dependent vs Independent site.

These findings influence how we understand the global role of a transcription factor (TF). In this case, we are studying CRX, but the fact that a TF does not necessarily have a functionally relevant role at every ChIP-seq defined binding site is important for the study of all TFs regardless of the cell type. The work presented here shows that, especially for developmentally important TFs, ATAC-seq is an important and useful tool to elucidate this functional role. ATAC-seq is relatively inexpensive, highly reproducible, and the data corresponded well with downstream read-outs of activation state such as histone ChIP-seq data. Future experiments should be designed to understand the effects of other retinal TF mutant models. In particular, a thorough combined analysis with data from *Nrl*<sup>-/-</sup>, *Nr2e3*<sup>rd7/rd7</sup>, among others would allow for a better understanding of the relationship between these key factors and could explain the shared or unique activity of each at the many co-bound regulatory sites. Experiments that profile at higher developmental resolution (daily from P3-P10) may determine a finer resolution to the temporal activation state of CRX Dependent vs Independent sites than that described in my work. This high-resolution approach may also be mandatory to understand the role of NRL as *Nrl*<sup>-/-</sup> results in a different transcriptional program to activate in the absence of this positive rod signals prior to P14.

The TF network of the rod photoreceptor that includes CRX, NRL, and NR2E3 is largely based on developmental and cell fate decision analyses. The findings presented in Chapter 4 show that CRX is required developmentally to activate certain sites, but raise the question whether CRX is also required to maintain that activity in the mature photoreceptor. By generating a conditional *Crx* allele and utilizing a tamoxifen responsive *Cre* line, future experiments could allow the rod to develop normally but deplete CRX only after epigenetic and functional maturity. A time course of ATAC-seq over the subsequent weeks would define the role of CRX in maintaining a mature epigenome. A similar experiment was performed testing the effect of adult loss of *Nrl*. Authors determined that while expression of many genes was lost, DNA methylation remained unchanged.[Montana et al., 2012] This result suggests that the epigenome was left unperturbed, perhaps due to the sustained expression of CRX. Our conditional deletion of *Crx* will test this question and determine whether NRL and/or NR2E3 can compensate and continue to promote gene expression alone upon loss of CRX.

Finally, these data highlight a core set of CRX regulatory sites. We predict that the mutation of these sites would have unique consequences on rod photoreceptor development but largely spare other neuronal or unrelated tissues. We predict that this finding is directly relevant to the understanding of non-coding variants in human disease. As more families with retinal specific disorders undergo whole genome sequencing, we propose that these sites represent a list of high priority sites to analyze for variants, especially within the CRX homeodomain motif. Additional experiments could investigate this hypothesis by employing a CRISPR screening technique in mice. By constructing a dual guide RNA library to remove the regulatory elements, mass electroporation of this library into the retina of a *Rho-GFP* mouse would test for the functional relevance of each regulatory site in the development of a mature rod photoreceptor (as

marked by GFP expression). This experiment could further define the most relevant binding sites that globally affect photoreceptor development and maintenance.

### **6.3 The rod photoreceptor as a model for genomic organization**

Previous work in the retina had thoroughly described the unique inverted nuclear organization of the mouse rod photoreceptor nucleus.[Solovei et al., 2009, 2013] Meanwhile, studies have made tremendous strides in the past several years to understand the organization of DNA within the nucleus primarily using cell culture systems. In Chapter 5, I described ongoing work where I use the rod photoreceptor as a model to understand organizing principles of the genome *in vivo*. My data suggests a complex relationship between expressed genes and has uncovered basic principles governing the inclusion of genomic regions within a highly active sub-compartment of the nucleus.

This study suggests a number of future experiments to further understand genomic organization in the rod photoreceptor. First, although the analysis presented here suggested the relevance of HiC data from a related neuronal tissue (cortex), such a whole genome ‘all vs all’ approach comparing rods and cones (*WT* vs *Nrl*<sup>-/-</sup>) would be useful. This data could highlight key differences between the two cell types and display the general nature of compartmental organization upon inversion of the rod nucleus. Future experiments could attempt to ‘recover’ a conventional organization by expressing LBR or LMNA/C under a rod photoreceptor specific promoter. Full morphological and functional characterization would determine whether the inversion of the nucleus is required for proper gene expression, while molecular assessment by



4C, HiC or FISH would report any effects on TAD/Compartmental organization or establishment.

Our 4C and FISH data suggested a strong association of *Rho* with other active regions of the chromosome that were lost when *Rho* was inactivated (in the *PPR*<sup>-/-</sup> retina). However, our data that assessed the relative position of *Rho* and other genes did not have the resolution to determine what interactions *Rho* gained. Future 4C or HiC experiments comparing *PPR*<sup>-/-</sup> and *WT* photoreceptors would first ensure that the rest of Chromosome 6 maintains the same conformation, and determine whether *Rho* has been relegated to the B compartment or just out of a highly active A sub-compartment. Our data only showed a strong correlation between chromatin state and association with *Rho*, but did not test any active mechanism for this organization. Previous efforts to perform dual IHC/FISH experiments failed in retinal tissue sections. However, new FISH variations have not yet been tested. These experiments could be useful to determine how close *Rho* is to the nuclear lamina, or its proximity to nuclear speckles or densities of other active chromatin associated proteins in *WT* vs *PPR*<sup>-/-</sup> cells.

Our data suggested that, when *Rho* was inactivated, genes within close proximity showed excess transcriptional activation. This is evidence that local genes compete for transcriptional machinery. *Rho* is a relatively unique gene in that it is so highly expressed relative to all other transcripts in the cell. However, epigenetically it looks quite similar to other highly active chromatin regions (super enhancers in particular). Future experiments could test whether similar effects are observed when other highly expressed genes or ‘super enhancer’ regions are inactivated in rods or even in other cell types.

Finally, on a ‘local’ chromatin scale, this study has challenged the conventional understanding of *Rho* expression regulation and the function of upstream enhancers. We

generated the *RER*<sup>-/-</sup> mutant mouse and discovered that the *RER* has very little if any contribution to bulk *Rho* expression. Gene expression was only assessed at P14, so additional experiments will be necessary to test whether the *RER* is necessary for proper expression earlier in development or in adult photoreceptors. Alternatively, the more distal *CBRI* could be compensating for the loss of the *RER*. Ongoing experiments will test the combined and unique activities of the *CBRI* and *RER* regulatory elements (see preliminary data in Appendix 1).

## 6.4 Conclusions

The work presented in this dissertation has investigated the mechanisms that establish and the importance of the rod photoreceptor epigenome. My results have specific implications for the understanding of retinal development and disease pathogenesis as well as broadly applicable principles important for the understanding of gene regulation and genomic organization. In the retina, I have elucidated the role of the TF CRX, implicated in a number of severe blinding disorders. These experiments provide a better understanding of the disease mechanisms of distinct classes of CRX mutations and identify specific non-coding elements in the genome that are likely critical for proper CRX directed photoreceptor gene expression. More broadly, this work suggests that all TFs involved in cellular development likely have complex and diverse roles, and their binding at a particular site cannot be interpreted as important or indicative of activity. Finally, my work presented here regarding genomic organization offers further support that we cannot think of each gene as being controlled independently; expression is influenced by that of nearby genes, and this organization within the nucleus is highly regulated and likely a fundamental and important part of cellular development and homeostasis.

## 6.5 References

- Montana, C. L., Kolesnikov, A. V., Shen, S. Q., Myers, C. A., Kefalov, V. J., & Corbo, J. C. (2012). Reprogramming of adult rod photoreceptors prevents retinal degeneration. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, *110*, 1732–1737.
- Roger, J. E., Hiriya, A., Gotoh, N., Hao, H., Cheng, D. F., Ratnapriya, R., Kautzmann, M. I., et al. (2014). OTX2 loss causes rod differentiation defect in CRX-associated congenital blindness. *The Journal Of Clinical Investigation*, *124*, 631–643.
- Ruzycki, P. A., Linne, C. D., Hennig, A. K., & Chen, S. (2017). Crx-L253X Mutation Produces Dominant Photoreceptor Defects in TVRM65 Mice. *Investigative Ophthalmology & Visual Science*, *58*, 4644–4653.
- Ruzycki, P. A., Tran, N. M., Kefalov, V. J., Kolesnikov, A. V., & Chen, S. (2015). Graded gene expression changes determine phenotype severity in mouse models of CRX-associated retinopathies. *Genome Biology*, *16*, 171.
- Solovei, I., Kreysing, M., Lanctôt, C., Kösem, S., Peichl, L., Cremer, T., Guck, J., et al. (2009). Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. *Cell*, *137*, 356–68.
- Solovei, I., Wang, A. S., Thanisch, K., Schmidt, C. S., Krebs, S., Zwerger, M., Cohen, T. V., et al. (2013). LBR and Lamin A/C Sequentially Tether Peripheral Heterochromatin and Inversely Regulate Differentiation. *Cell*, *152*, 584–598.
- Tran, N. M., Zhang, A., Zhang, X., Huecker, J. B., Hennig, A. K., & Chen, S. (2014). Mechanistically Distinct Mouse Models for CRX-Associated Retinopathy. *PLoS Genetics*, *10*, e1004111.

# **Appendix 1**

## **The role of the *Rhodopsin* enhancers**

Early efforts to understand the very high rod photoreceptor specific expression of *Rhodopsin* focused on the discovery and classification of the promoter and nearby enhancers.[Chen & Zack, 1996; Corbo et al., 2010; Nie et al., 1996] These experiments discovered three highly conserved ~200bp elements termed the proximal promoter region (*PPR*), *Rhodopsin* enhancer region (*RER*; 2kb upstream of *PPR*), and the CRX bound region 1 (*CBRI*; 4kb upstream of *PPR*). All elements drove strong retinal expression by plasmid reporter assays, and the *PPR* and *RER* showed similar specificity in transgenic animals.[Jimeno et al., 2006; Nie et al., 1996] All elements have since shown strong *in vivo* binding of both important transcription factors CRX and NRL (Figure A1.1).[Corbo et al., 2010; Hao et al., 2012]

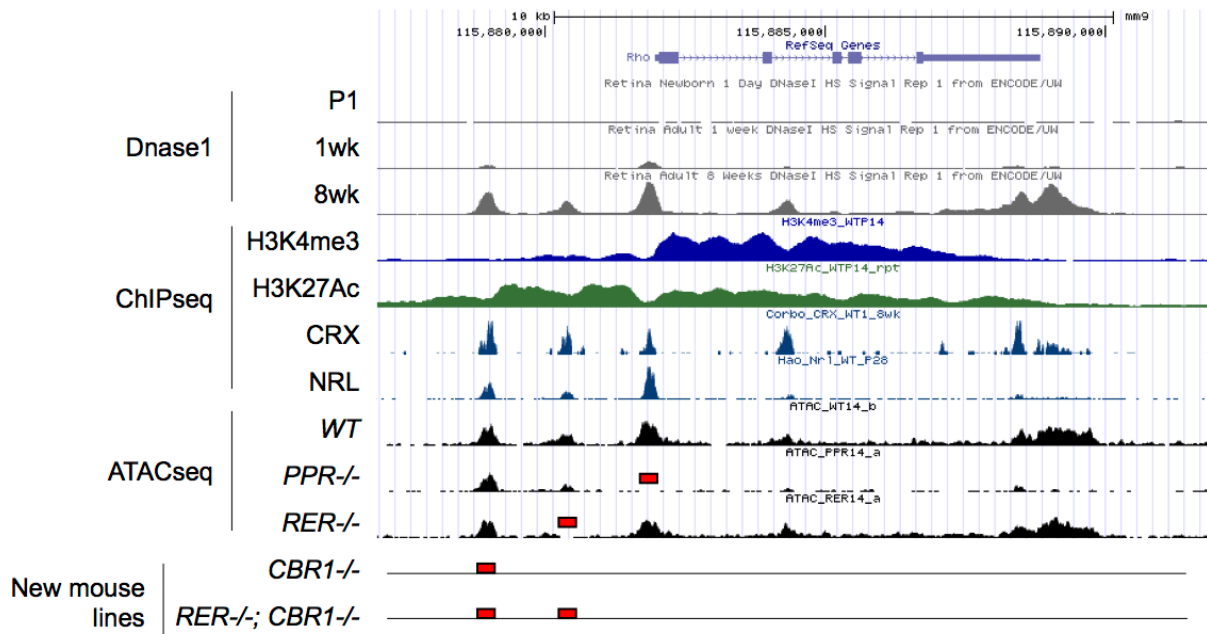
In Chapter 5 of this dissertation, I presented an experiment where we used CRISPR-Cas9 to create germline knockout mouse lines lacking either the *PPR* or the *RER* (*PPR*<sup>-/-</sup> and *RER*<sup>-/-</sup> respectively). We analyzed both the epigenome and transcriptome of these mice by ATAC-seq and ddPCR and were surprised to find that there was no detectable defect in the *RER*<sup>-/-</sup> retina. While *PPR*<sup>-/-</sup> mice displayed dramatic loss of *Rho* expression, thinning of the outer nuclear layer (ONL), and loss of local ATAC-seq signal at nearby regulatory elements, *RER*<sup>-/-</sup> mice showed no detectable differences from *WT* (Chapter 5: Figure 5.2 and Supplemental Figure 5.2).

We noticed that the *PPR*<sup>-/-</sup> retina still displayed strong ATAC-seq sensitivity at the *CBRI*, but this signal was largely lost at the *RER* (Figure A1.1). This data suggests that, at P14, the activity of the *RER* is dependent upon promoter activation, but the state of the *CBRI* is independent. Interestingly, in the *RER*<sup>-/-</sup>, the *CBRI* and *PPR* were both fully activated, (consistent with normal *Rho* expression) suggesting that the *RER* alone was not essential for the activation of the *CBRI* or *PPR*.

Together, this data suggest that the *CBRI* may compensate for the loss of the *RER*. To test this question, we again used CRISPR-Cas9 to generate two new mouse lines; the first lacks only the *CBRI* (*CBRI*<sup>-/-</sup>) and the second lacks both the *RER* and the *CBRI* (*RER*<sup>-/-</sup>;*CBRI*<sup>-/-</sup>) (Figure A1.1).

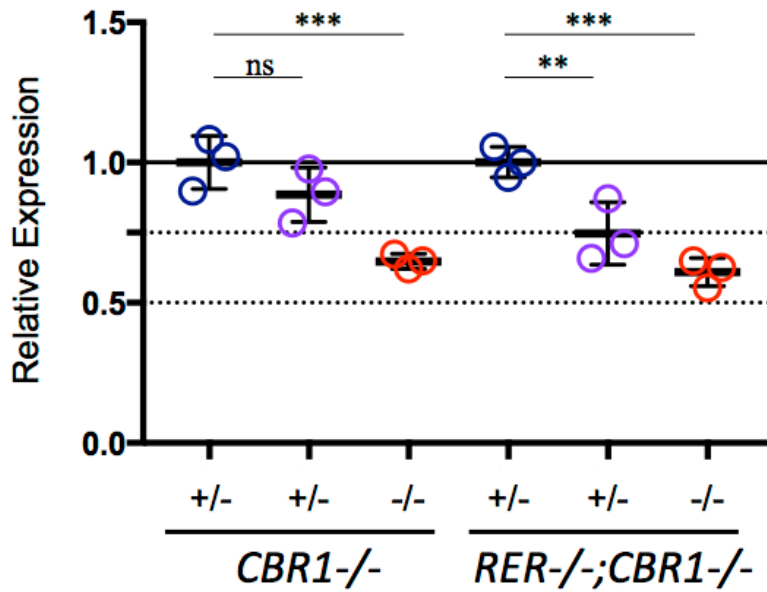
Experiments are ongoing to determine the epigenomic consequences of the loss of the *CBRI* alone or in conjunction with the *RER*. However, initial results do strongly support a role of the *CBRI* in *Rhodopsin* gene expression regulation. To test *Rhodopsin* expression, I collected retinas at P14 from all mouse lines and performed ddPCR. Results show that the loss of the *CBRI* alone leads to a ~40% reduction in *Rho* expression (Figure A1.2). Surprisingly the combined loss in the *RER*<sup>-/-</sup>;*CBRI*<sup>-/-</sup> retina showed no additional deficit in expression.

In conclusion, these results suggest that the *CBRI* is critical for proper *Rho* expression, but leave the role of the *RER* unanswered. Ongoing experiments will test the epigenomic landscape of *CBRI*<sup>-/-</sup> and *RER*<sup>-/-</sup>;*CBRI*<sup>-/-</sup> retinas at P14, and expression of *Rhodopsin* at earlier and later time points in development.



**Figure A1.1: Epigenetic landscape of *Rhodopsin***

Browser tracks display relevant epigenomic data showing temporal activation of *PPR*, *RER*, and *CBR1* (by DNase1) and steady state mature profiles of other data. ATAC-seq tracks show changes in ATAC-seq sensitivity as a result of perturbations to *Rho* regulatory elements. Areas deleted in various lines are marked with red box.



**Figure A1.2: Disruption of *CBR1* affects *Rho* expression**

ddPCR analysis of *Rho* expression (*Actb* as control) in designated mouse lines at P14. All samples are littermate controls (n=3; mean  $\pm$  SEM; ; \*\*\* p<0.01, \*\*\* p<0.001; One-way ANOVA).



## Appendix 1 References

- Chen, S., & Zack, D. J. (1996). Ret 4, a positive acting rhodopsin regulatory element identified using a bovine retina in vitro transcription system. *The Journal Of Biological Chemistry*, 271, 28549–57.
- Corbo, J. C., Lawrence, K. a, Karlstetter, M., Myers, C. a, Abdelaziz, M., Dirkes, W., Weigelt, K., et al. (2010). CRX CHIP-seq reveals the cis-regulatory architecture of mouse photoreceptors. *Genome Research*, 20, 1512–25.
- Hao, H., Kim, D. S., Klocke, B., Johnson, K. R., Cui, K., Gotoh, N., Zang, C., et al. (2012). Transcriptional Regulation of Rod Photoreceptor Homeostasis Revealed by In Vivo NRL Targetome Analysis. *PLoS Genetics*, 8, e1002649.
- Jimeno, D., Feiner, L., Lillo, C., Teofilo, K., Goldstein, L. S. B., Pierce, E. A., & Williams, D. S. (2006). Analysis of kinesin-2 function in photoreceptor cells using synchronous Cre-loxP knockout of Kif3a with RHO-Cre. *Investigative Ophthalmology And Visual Science*, 47, 5039–5046.
- Nie, Z., Chen, S., Kumar, R., & Zack, D. J. (1996). RER, an evolutionarily conserved sequence upstream of the rhodopsin gene, has enhancer activity. *Journal Of Biological Chemistry*, 271, 2667–2675.

# **Appendix 2**

## **Methods**

## Chapter 2 Methods

### Nomenclature for the *TVRM65* mutation

There are several annotated transcripts for the murine *Crx* locus. The previously published amino acid change in the *TVRM65* mouse was based on isoform 2 (NM001113330.1) that encodes a CRX protein with 323 amino acids.[Chau et al., 2000; Won et al., 2011] However, the major transcript in the retina is isoform 1 (NM\_007770.4), which encodes a CRX protein homologous to human CRX, with 299 amino acids. Thus, here we refer to the *TVRM65* mutation as *p.L253X*, based on isoform 1 numbering.

### Mice

*Crx-L253X (TVRM65)* mice [Won et al., 2011] were obtained from Jackson Laboratory (Bar Harbor, ME, MGI ID: 4867395). Mice were mated and maintained with *C57BL/6J* (JAX Stock number 000664), and confirmed free of *rd1* and *rd8* mutations by PCR genotyping. *Crx* null mice (*Crx*<sup>-/-</sup>) were provided by Dr. Constance Cepko, Harvard University (Boston, MA) [Furukawa et al., 1999] and back-crossed onto *C57BL/6J* for more than 10 generations. All mice were housed in a barrier facility in the Division of Comparative Medicine of Washington University School of Medicine. All procedures involving the use of mice were approved by Washington University's Animal Care and Use Committee and followed [ARVO guidelines for the Use of Animals in Ophthalmic and Vision Research](#).

### PCR Genotyping, RNA Purification and Reverse Transcription, and qRT-PCR

Genotyping, RNA purification and reverse transcription, and qRT-PCR were performed as previously described.[Tran et al., 2014a] Primer sets for genotyping and RT-PCR are listed in Supplemental Table 1. For qRT-PCR, all primers were tested for proper amplification efficiency prior to use. Relative gene expression was normalized to the retinal constitutively expressed genes *Ubb*, *Tuba1B*, and *Gapdh*. Data for 3 biological replicates were then analyzed using the Delta Cq method in QBase software (Biogazelle, Ghent, Belgium). The results are presented by the heatmap.2 function of the R gplots package (v3.0.1).

## Droplet digital PCR (ddPCR)

Confirmatory genotyping and allelic expression data were generated using the droplet digital PCR system (BioRad, Hercules, CA). A 20ul ddPCR reaction mixture containing custom-designed normal and mutant primer/probe sets (see Supplemental Table 2) was prepared according to manufacturer's directions, droplets generated, and nano-reactions cycled on a C1000 Touch Thermal Cycler (BioRad). The average number of allelic transcripts per biological replicate (n=3) was then determined with the QX 200 Droplet Reader and QuantaSoft Analysis Pro (BioRad).

## Transient Transfection Luciferase Reporter Assays

HEK293T cells (ATCC CRL-11286) were cultured in Dulbecco's Modified Eagle Medium with 10% fetal bovine serum and Penicillin/Streptomycin. Cells were transfected using conventional CaCl and Boric Acid Buffered Saline Method in 6-well plates as previously described.[Chen et al., 2002; Tran et al., 2014a] Experimental plasmids and transfection amounts included *BR130-luc* [2 ug] (*a Rho* promoter-luciferase reporter)[Chen et al., 2002], *pRL-CMV* [1 ng] (transfection normalization control) and protein expression vectors *PED-NRL* [100 ng], *pcDNA3.1-hCrx*, *Crx1-254*, and *Crx-E168d2*. Cells were harvested 48 hours post transfection and assays performed using the Dual-Luciferase Reporter Assay System (Promega, Madison, WI).

## Electroretinogram (ERG) and Statistical Analyses

ERGs and statistical analyses were performed as previously described in [Tran et al., 2014a]. Briefly, tests were performed on a UTAS-E3000 Visual Electrodiagnostic System running EM for Windows (LKC Technologies, Inc., Gaithersburg, MD) while mouse body temperature was maintained at 37±0.5°C with a heating pad controlled by a rectal temperature probe (FHC Inc., Bowdoin, ME). Pupils were dilated with 1.0% atropine sulfate (Bausch & Lomb, Tampa, FL) and dilation and corneal hydration maintained during testing by positioning the platinum wire loop recording electrodes in a mixture of atropine and 1.25% hydroxypropyl methylcellulose (GONAK; Akorn Inc., Buffalo Grove, IL). Mice were tested without knowledge of genotype. Bilateral flash ERG responses were obtained; the set of recordings giving the higher amplitudes was correlated with genotype info for statistical analyses including two-way ANOVA and post-hoc multiple comparison tests were performed using Prism (GraphPad, La Jolla, CA).

## **Quantitative Western blot analysis**

Western Blots were performed as previously described.[Tran et al., 2014a] Experiments were performed using three biological replicates (15ug total protein extract from homogenates of two retinas). Membranes were probed with mouse monoclonal anti-CRX antibody M02 (Abnova, 1:200) and anti-GAPDH antibody (Sigma G9545, 1:10,000), visualized with donkey anti-rabbit 680nm and donkey anti-mouse 800nm IRDyes (LI-COR, Lincoln, NE) (1:10,000), then imaged and quantified using the Odyssey Infrared Imager and ImageStudio 6 software (LI-COR).

## **Histology, Morphometry, Immunohistochemistry, and Statistical Analyses**

Histology, morphometry, immunohistochemistry and statistical analyses were performed as previously described.[Tran et al., 2014a] 4um sagittal sections of paraffin-embedded eyecups, cut through the optic nerve head, were either stained with hematoxylin and eosin (H&E) or prepared for immunofluorescence staining. Slides containing sections were incubated overnight with primary antibody at 4°C (1:400 anti-Rhodopsin monoclonal RetP1, Sigma), washed, and incubated with secondary antibody (1:500 Alexa 568-goat-anti-mouse, Invitrogen) for 2 hours at room temperature. Coverslips were mounted with Vectashield Hard Set mounting medium with DAPI (Vector Labs, Burlingame, CA). All imaging was performed using a Leica DM5500B microscope and Leica DFC365FX camera (Leica Microsystems, Buffalo Grove, IL).

## Chapter 3 Methods

### Mice

Mice were housed in a barrier facility operated and maintained by the Division of Comparative Medicine of Washington University School of Medicine. All mice used in this study were backcrossed to *C57BL/6J* mice obtained from Jackson Laboratories (Bar Harbor, ME, Stock number 000664) for at least 10 generations. Each line was genotyped for common *rd1*, *rd8*, and *Rpe65* variants that affect retinal structure and function; all lines were negative for *rd1* and *rd8* and contain the *C57BL/6J* RPE65<sup>450M</sup> isoform. The *Crx* mutant lines *E168d2*, *E168d2neo* and *R90W* and their genotyping procedures were published previously [Tran et al., 2014b].

All procedures involving mice were approved by the Animal Studies Committee of Washington University in St. Louis (IACUC 20120246, expiration date 01/18/2016). Experiments were carried out in strict accordance with recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health (Bethesda, MD), the Washington University Policy on the Use of Animals in Research; and the Guidelines for the Use of Animals in Visual Research of the Association for Research in Ophthalmology and Visual Science (<http://www.arvo.org/animals/>).

### RNA collection & library preparation

For each genotype and time point: three biological replicates were analyzed, and each replicate consists of four retinas from one male and one female mouse. P21 *WT* samples represent a randomized sampling of *WT* littermates of *Crx* mutants. Retinas were immediately processed for RNA using the PerfectPure RNA tissue kit (5 Prime). The quantity and quality of the RNA was assayed using a Bioanalyzer (Agilent). Samples with a minimum RIN score of 8.0 were then selected for library construction. Library construction was performed by the Genome Technology Access Center (GTAC) (Washington University in St. Louis). mRNA was isolated by poly-A selection using Oligo-dT beads (mRNA Direct Kit, Ambion). mRNA was then fragmented by incubation at 94C for 2.5min in an alkaline buffer (40mM Tris Acetate pH 8.2, 100mM Potassium Acetate, 30mM Magnesium Acetate) and reverse transcribed using Super Script III enzyme (Invitrogen) and random hexamers to yield double stranded cDNA. The cDNA was blunt ended using Klenow DNA polymerase, followed by addition of an A base and ligation of Illumina sequencing adapters to the 3' ends. All enzymes including DNA polymerase, Klenow DNA polymerase, Klenow exo-, RnaseH, T4 DNA polymerase, and T4 Polynucleotide Kinase were purchased from New England BioLabs (NEB). Ligated fragments

were amplified for 12 cycles using primers incorporating the unique index tags before sequenced on the Illumina Hi-Seq 2000 using single reads extending 42 bases.

## **RNA sequencing data analysis**

1x42bp reads were aligned to the mouse genome (version mm9) with the sequence aligner TopHat2 (Version v2.0.5) using the following parameters: -a 5 -m 1 -i 10 -I 500000 -r 100 -p 4 --microexon-search --no-coverage-search -x 20 --segment-length 25. Dependencies included Bowtie (v0.12.8) and Samtools (v0.1.18). Bedgraph files were generated using BEDTools (v2.23.0) and visualized using IGV (Broad Institute). The HTSeq package (Version 0.6.1p1) was used to assign aligned reads to the gene annotation reference track (UCSC Genes Track, UCSC Table Browser, NCBI37/mm9, accessed July 16, 2014). This generated a raw read count per gene which was used in EdgeR [Nikolayeva & Robinson, 2014] for detecting differentially expressed genes. For each of the genotype comparisons, genes that did not pass the filter criteria of counts per million (CPM)  $\geq 5$  in all replicates of at least one comparison group were removed prior to the analysis. Filtered count data was normalized by the EdgeR default normalization method, TMM, and differential expression analysis for each of the comparison groups were performed by the exact test. P-values were subjected to Bonferroni and Hochberg multiple testing correction to include false discovery rate (FDR). Downstream analysis was performed using custom Perl and R scripts. Principal Component Analysis was performed using the princomp function in the stats R package (v3.1.1). Gene expression heatmaps were generated using the heatmap.2 function of the gplots R package (Version 2.16.0). As several heatmaps have included too many genes to properly label each row, we have included a supplemental file that lists the genes included in each heatmap in the order in which they were presented (Table S4).

Raw RNA-seq data was also obtained from previously published work: Gene Expression Omnibus (GEO) accession: GSE52006 [Roger et al., 2014]. This data was aligned also using TopHat and analyzed as described above using HTSeq and EdgeR.

## **Epigenetic data analysis**

Data was obtained from GEO under accession numbers GSE20012 [Corbo et al., 2010], GSE38500 [Popova et al., 2012], and GSM1014198/GSM1014175/GSM1014188 [Vierstra et al., 2014]. NRL ChIP seq [Hao et al., 2012] was downloaded from the NEI Data Share website, accessible through the following link (<http://datashare.nei.nih.gov/dataShareMain.jsp>). Data were analyzed based on mouse reference genome (NCBI37/mm9) and UCSC known gene table. Heatmaps and line graphs depicting epigenetic data were generated using the UCSC defined TSS for each gene and the UNIX software package HOMER (Version 4.7) [Heinz et al., 2010].

Heatmaps were generated by importing the HOMER generated counts back into R, ordering by adult DNase1 level at the TSS, and visualized by the heatmap.2 function in the gplots R package.

## Gene Ontology Analysis

Gene ontology analysis (GO Analysis) [Ashburner et al., 2000] was performed using the online g:Profiler tool [Reimand et al., 2011]. Terms were filtered to only include the best parent term.

## Light damage (LD) and 13-*cis*-RA treatment

For light damage (LD) analysis, mutant mice and littermate controls were dark adapted overnight. Eyes were dilated with 1% Cyclogyl and 2.5% phenylephrine hydrochloride. Mice were then placed in darkness for 30 min before LD. LD was performed in a temperature controlled rat cage with the top removed and lined with reflective material to provide even illumination. Exposure was performed using white fluorescent light and the intensity was measured using a lux meter. Light intensity was 11-13.5 KLUX in all experiments. Mice were placed into cages containing moistened food pellets, and each animal received damaging light for 8 hrs. Mouse pupils were re-dilated and the cage position was rotated every 2 hours.

For 13-*cis*-RA treatments, 13-*cis*-retinoic acid (RA) (Isotretinoin, Accutane) (Sigma-Aldrich, R3255) was reconstituted in DMSO to a concentration of 0.013 mg/μl. 40 mg/kg 13-*cis*-RA or an equal volume of DMSO vehicle control was injected into the mouse's interperitoneal cavity 12 hrs before LD and mice were dark-adapted overnight. Mice were re-dosed with 40 mg/kg 13-*cis*-RA 30 min before LD.

Following all LD experiments, mice were returned to a 12 hr light-dark cycle under normal ambient light conditions (~100 LUX) for 7 days. After 7 days, retinal function was tested by electroretinography (ERG). The peak amplitude responses for rod-driven dark-adapted a-waves, rod ON bipolar cell-driven dark-adapted b-waves and cone bipolar cell-driven light-adapted b-waves were measured for a series of increasing light intensities as described below. Mice were then sacrificed and eyes were collected for retinal histology assessment (see below). Images were taken in the central superior retina (~500 μm from the optic nerve (ON)), which is the area of the retina that is most strongly affected by LD [38].

## Histology, immunohistochemistry and microscopy

Histology analyses of hemotoxylin and eosin-stained (H&E) paraffin-embedded retinal sections were performed as described previously [Tran et al., 2014b]. For ONL morphometry, 20X composites of whole retinal sagittal sections stained with H&E were analyzed using ImageJ



software [Schneider et al., 2012]. The distance from the ON was determined by drawing a curved line along the outer limiting membrane. The ONL thickness was measured at 100  $\mu\text{m}$ , 500  $\mu\text{m}$ , 1000  $\mu\text{m}$ , and 1500  $\mu\text{m}$  from the ON and 200  $\mu\text{m}$  from the peripheral edge on both the superior and inferior retina. Results are presented by a 'spider graph'. The between-group differences in ONL thickness were determined by testing overall genotype\*distance interactions ( $p < 0.05$ ,  $n \geq 3$ ) of each treatment. Statistical comparison of data at each distance was performed using two-way ANOVA for repeated measurement data, followed by a post-hoc test to adjust p-value for multiple comparisons between each genotype and the *WT* control group SAS 9.3 (SAS Institutes, Cary, NC).

Fluorescence immunohistochemistry was performed on frozen retinal sections. Eyes were enucleated and fixed for 30 minutes in 4% paraformaldehyde (PFA) on ice. Eyes were washed in PBS and left overnight in 30% sucrose in PBS at 4 °C. 50% volume of OCT was added and rocked at room temperature for 1 hr. Eyes were then transferred to 100% OCT for 1 hr before frozen in OCT on dry ice. Blocks were cut onto polylysine (Thermo Scientific) slides and stored at -80 °C. For staining, slides were removed and allowed to dry for 30 min before 10 min 4% PFA fixation. Slides were washed (2 x 5 min) with PBS and blocked at room temperature with 5% nonfat dry milk/2% bovine serum albumin (BSA)/10% normal goat serum (NGS) for 2 hrs. Slides were washed (3 x 5 min) with PBS before overnight incubation with primary antibody at 4 °C. Antibodies were diluted as below in PBS with 1% BSA and 0.5% Triton X-100. Slides were washed (3 x 5 min) in PBS. Secondary antibodies (and PNA antibody) were diluted in the same buffer as primaries, applied to the slides, and incubated for 1 hr at RT, and washed again (3 x 5 min) with PBS before the application of coverslip.

Primary antibodies and dilutions used are as follows: rabbit anti-RXR $\gamma$  (1:500, Santa Cruz Biotechnology, sc-555), rabbit anti-CNGB3 (1:500, from Xi-Qin Ding, University of Oklahoma), peanut agglutinin (PNA) conjugated to Rhodamine (1:500, Vector Labs). Secondary antibodies (1:400): goat anti-rabbit or mouse IgG antibodies coupled to Alexa Fluor A488 and Rhodamine 568 (Invitrogen). All slides were counterstained with hard-set mounting medium with DAPI (Vectashield).

Brightfield and fluorescent imaging was performed using either an Olympus BX51 microscope and Spot RT3 Cooled Color Digital camera (Diagnostic instruments Inc.) or a Leica DM5500B microscope and Leica DFC365FX camera (Leica Microsystems, Buffalo Grove, IL)

## **Electroretinography (ERG)**

Mice were dark-adapted overnight and anesthetized with intraperitoneal injection of a mixture of ketamine (80 mg/kg) and xylazine (15 mg/kg). Pupils were dilated with 1% atropine sulfate. A passive-heating pad maintained animal's body temperature at 37 °C. Scotopic ERG responses were measured from both eyes using corneal platinum-ring electrodes held in place by a drop of Gonak solution. Full-field ERGs were recorded with the UTAS-E 3000 system (LKC

Technologies, Inc.), using Ganzfeld-derived test stimuli of calibrated white light intensity. The amplitude of ERG a-wave was determined from the baseline to the primary negative peak of the photoresponse. The amplitude of ERG b-wave was measured from the a-wave peak to the maximum of the secondary positive peak. Photopic ERG recordings were performed under room illumination ( $\sim 25 \text{ cd}\cdot\text{s m}^{-2}$ ) after 5–10 min of light adaptation.

The rate of rod dark adaptation was determined from the recovery of maximal ERG a-wave amplitude after 30-second exposure to light bleaching of  $> 90\%$  of rhodopsin (delivered by 520 nm LEDs focused at the surface of mouse eye cornea and producing  $\sim 2.5 \times 10^8 \text{ photons } \mu\text{m}^{-2}\text{s}^{-1}$ ), using test flashes of  $\sim 470 \text{ cd}\cdot\text{s m}^{-2}$  at indicated times after the bleach. In these experiments, mice were re-anesthetized subcutaneously every 30–40 min with a smaller dose of ketamine ( $\sim 50\%$  of the initial dose). In addition, at the same times a drop of 1% atropine sulfate was added to the eye surface to keep pupils dilated, and a 1:1 mixture of PBS and Gonak solutions was applied to the eyes to protect them from drying and maintain electrode contacts during extended recording sessions.

## qRT-PCR

RNA used for assays included RNA used for Illumina RNAseq library preparation and additional biological replicates. For each genotype,  $n \geq 3$  biological replicates, each replicate consisting of four retinas from one male and one female mouse, were collected. RNA isolation, cDNA preparation, and qRT-PCR reactions were performed as previously described [21]. Briefly, retina tissue was immediately processed for RNA using the PerfectPure RNA tissue kit (5 Prime) and quantified. cDNA was synthesized from 1  $\mu\text{g}$  of RNA using the Transcriptor First Strand cDNA Synthesis kit (Roche Applied Science). A 10  $\mu\text{l}$  QRT-PCR reaction mixture containing  $1 \times$  EvaGreen with Low Rox reaction mix (BioRad), 1  $\mu\text{M}$  primer mix, and diluted cDNA was prepared and run on a BioRad CFX thermocycler in triplicate. Data was analyzed using QBase software (Biogazelle). Relative gene expression was normalized to *Ubb* and *Tuba1b* and fold-change from *WT* was determined. Kruskal-Wallis and Dunn's Multiple Comparison tests were used to determine significant fold change differences from *WT* ( $P < 0.05$ ). Primer sets were designed using MacVector software and synthesized by IDT DNA technologies (Table S5).

## Chapter 4 Methods

### ATAC-seq library prep and sequencing

ATAC-seq was performed as published in Buenrostro, et al.[Buenrostro et al., 2015] Briefly, retinas were dissected from P14 *WT* and *Crx*<sup>-/-</sup> mice and washed in PBS. Tissue was dissociated at 37C using 2% Collagenase in TESCA buffer for 13 min, and the reaction stopped by the addition of 2X volume of DMEM + 10% FBS. DNase1 (0.5 Units) was added for the final 3 minutes to minimize clumping of cells. Cells were counted with hemocytometer, and 50,000 re-suspended in TD buffer for a 1 hour incubation with TDE1 at 37C. Remaining library prep was performed as published. Libraries were pooled and sequenced using the Illumina 2500.

### Chromatin Immunoprecipitation

Chromatin immunoprecipitation (ChIP) assay was performed as previously described.[Tran et al., 2014b] In summary, 6 pooled P14 C57BL/6J Wild type, or *Crx*<sup>-/-</sup> mouse retinas per sample were dissected and chromatin was cross-linked with 1% formaldehyde in PBS for 10 minute at room temperature. Crosslinked cells were lysed and fragmented by sonication. Chromatin fragments were immunoprecipitated with the Histone H3 (acetyl K27 - ab4729) (Abcam, Cambridge, UK) and trimethyl-Histone H3 (Lys4) (Millipore, 07-473) antibodies, or normal rabbit / mouse IgG (Santa Cruz Biotechnology, Dallas, TX) bound to Protein A beads (Millipore, 16-125), or A/G beads (Santa Cruz Biotechnology). After extensive washing, the immunoprecipitated chromatin was eluted, heated to 67°C to reverse the cross-links, and the DNA purified by ethanol precipitation. Libraries were prepared using the DNA SMART ChIP-Seq Kit (Clontech, Mountain View, CA). 10ng of ChIP DNA was used as input for each sample.

### Mapping of ATAC-seq and ChIP-seq data

Libraries were de-multiplexed according to barcodes inserted in the P7 adaptor, and mapped to mm9 using Novoalign (V3.04.06). Alignments were cleaned using Samtools (V1.3.1); duplicate and reads mapping to mitochondrial genome were removed. All other processing for visualization in IGV was performed using samtools and bedtools (V2.24.0). For visualization, bedtools slop function was used to extend reads 300 bp.

## **Peak Calling and Genotype Comparison**

Peak calling was performed using MACS2 (V2.1.0.20140616). Peak calling was performed on replicate samples independently. Only peaks that replicated in both samples were kept, by comparing peak files using the bedtools intersect function. Intersecting peaks were merged using bedtools merge function. Peaks were analyzed for number of reads within each library using bedtools coverage, and statistical comparison for ATAC-seq data was performed using EdgeR (V3.18.1). Histone read count for FC analysis was computed by normalizing all libraries to RPM before subtraction of read count of appropriate input sample within same region..

## **Detection of overlapping ChIP, ATAC, and ChromHMM location**

All co-localization detection of genome-wide datasets was performed using the bedtools intersect function. ChromHMM bed files were graciously provided by Dr. Issam Aldiri, St. Jude's Children's Research Hospital.

## **Read coverage epigenetic analysis**

Heatmaps and line graphs depicting epigenetic data were generated using the UNIX software package HOMER (V4.7).[Heinz et al., 2010] Heatmaps were generated by importing the HOMER generated counts back into R, and visualized by the heatmap.2 function in the gplots (V3.0.1) R package.

## **Motif calling**

Known motif analysis was performed using Transcription Factor Affinity Prediction (TRAP) web Tools.[Manke et al., 2008] Analysis was performed on Jaspar vertebrate matrix file, with mouse promoter background model, and multiple test correction was Benjamini-Hochberg. Heatmap visualizes the  $-\log(10)$  converted corrected p-value of each TF motif. De novo motif analysis was performed using the HOMER findmotifsgenome tool (V4.7).

## **BEEML calculation of affinity**

Analyses were performed as in White et al (2016).[White et al., 2016] Custom scripts to calculate TF occupancy were graciously provided by authors.

## **Nucleotide density and motif density**

All analyses were centered on the CRX ChIP peak and utilized the HOMER annotatepeaks function. Matrix outputs from de novo motif analysis were tested for motif density using the HOMER annotatepeaks function in 20bp windows.

## **Gene Ontology (GO) analysis**

Gene Ontology (GO) analyses were performed using the GREAT tool (Ver 3.0).[McLean et al., 2010] Peak files used were centered on the CRX binding site, and associations were with the settings ‘single nearest gene- no limit on distance.’ These same associations were used as the basis of regulatory site- gene determination for RNA-seq comparisons.

## **RNA-seq**

RNA-seq data was obtained from GSE52006. Data was analyzed as described previously in [Ruzycki et al., 2015]. Briefly, 1x42bp reads were aligned to the mouse genome (version mm9) with the sequence aligner TopHat2 (Version v2.0.5) using the following parameters: -a 5 -m 1 -i 10 -I 500000 -r 100 -p 4 --microexon-search --no-coverage-search -x 20 --segment-length 25. Dependencies included Bowtie (v0.12.8) and Samtools (v0.1.18). Bedgraph files were generated using BEDTools (v2.23.0) and visualized using IGV (Broad Institute). The HTSeq package (Version 0.6.1p1) was used to assign aligned reads to the gene annotation reference track (UCSC Genes Track, UCSC Table Browser, NCBI37/mm9, accessed July 16, 2014). This generated a raw read count per gene which was used in EdgeR [Nikolayeva & Robinson, 2014] for detecting differentially expressed genes. For each of the genotype comparisons, genes that did not pass the filter criteria of counts per million (CPM)  $\geq 5$  in all replicates of at least one comparison group were removed prior to the analysis. Filtered count data was normalized by the EdgeR default normalization method, TMM, and differential expression analysis for each of the comparison groups were performed by the exact test. P-values were subjected to Bonferroni and Hochberg multiple testing correction to include false discovery rate (FDR). Downstream analysis was performed using custom Perl and R scripts.

## **External Datasets**

Other data used in this study were retrieved from the following accession locations on NCBI:

CRX ChIP-seq: GSE20012

CTCF ChIP-seq: GSE87064

DNase1: Retina P1 GSM1014188, Retina P7 GSM1014198, Retina 8wk GSM1014175;

Liver GSM1014195, Brain GSM1014151  
RNA-seq: GSE52006

## Chapter 5 Methods

### 4C Cell Collection

P14 retinas were dissected and dissociated at 37C using 2% Collagenase (Sigma Aldrich, St. Louis, MO) in TESCA buffer (50mM TES pH7.4, 0.36mM CaCl<sub>2</sub>) for 15 minutes. DNase1 (0.5 U) was added for an additional 3 minutes to avoid clumping of cells. Reaction was stopped by addition of 2X volume 37C DMEM+10% FBS and incubation for 5min at RT. Cells were strained through a 40um filter, before fixation in 2% Formaldehyde at room temperature (RT) for 10 minutes. Reaction was stopped by addition of cold glycine to a final concentration of 0.125M on ice. Cells were pelleted, re-suspended in PBS+1%FBS+1%EDTA, and counted using a hemocytometer. Aliquots of cells were pelleted and stored at -80C.

### 4C Template Preparation

4C libraries were prepped essentially as described in [van de Werken et al., 2012]. 10<sup>7</sup> cells were thawed and resuspended in 500ul Cutsmart buffer (NEB, Ipswich, MA). Cells were permeabilized using SDS and Triton-X100 before addition of 400U NcoI-HF restriction enzyme (NEB). Digestion was left overnight at 37C. After cutting efficiency was established, DNA was ligated with 100U T4 DNA Ligase (NEB) for 1 hour at 37C. After ligation, cells were de-crosslinked using Proteinase K (Qiagen, Hilden, Germany) and heated to 65C overnight. DNA was recovered by phenol-chloroform extraction. The DNA was again digested by addition of 100U MspI (NEB) overnight at 37C. Reaction was cleaned by phenol-chloroform precipitation, and again ligated with 100U T4 DNA Ligase (NEB) overnight at 16C. DNA was retrieved by ethanol precipitation and cleaned using Qiaquick PCR purification columns (Qiagen).

### 4C Inverse PCR

4C Inverse PCR was performed essentially as described in [van de Werken et al., 2012] with slight modifications. To increase the specificity of the inverse PCR, a nested approach was employed. Initial PCR was performed with the following primers: [F- GCATCACAACCTGTCCCTGC, R- GTTCCCAAACAAGACACCTGC]. 200ng 4C template was amplified using using KlenTaq enzyme (DNA Polymerase Technology, St. Louis, MO) for 25 cycles divided over 8x50ul reactions. The products were cleaned using PCR purification columns (Qiagen). Second PCR to add Illumina tails was performed using the following primers:  
F-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAGTGCAATCGAGGCTCCATG,  
R- CAAGCAGAAGACGGCATACGAGATGCTAAGCAAACATCTAGGAATCCCG

Elution from first PCR were again split into 8x50ul reactions and amplified using using KlenTaq enzyme (DNA Polymerase Technology) for 10 cycles. Products were cleaned using PCR purification columns (Qiagen). Finally, libraries were size-selected by gel electrophoresis (200-1600bp) and cleaned using the QiaexII Gel Extraction Kit (Qiagen). Libraries were pooled and sequenced (1X50bp) by the Genome Technology Access Center (GTAC) at Washington University, St. Louis using the Illumina HiSeq 2000 (Illumina, San Diego, CA).

## 4C data analysis

Reads were demultiplexed and mapped to the mouse genome (mm9) using bowtie (V1.1.1) with options (-S -v 2 -y -k 1 -m 1 -p 3 -trim5 [length of inver PCR primer sequence]). Resultant bam files were converted to bed format by Samtools (V1.3.1) and processed to count reads per known NcoI fragment using custom perl scripts. Raw data was then visualized using the IGV Browser [Robinson et al., 2011]. Scoring of distant interactions was performed using custom perl scripts. Current scoring methods focused on either binary or quantitative methods of analyzing 4C data. We chose to integrate these two methods together as follows: A running window test/background approach was employed to score the enrichment of our designated score within the test window compared to the much larger background area. All data presented in this Chapter used a 50/500 analysis of restriction enzyme fragments. This represents an enrichment area of approximately 100kb/1Mb and was determined empirically. We tested a range of windows and all highlighted similar areas of interaction. Smaller windows displayed more noise, and thus for this analysis we focused on a wider window.

$$\text{Score} = 1 / (\text{standard deviation}(\text{read coverage of contained fragments}) / (\text{total reads within window} / \text{number of fragments tested})) * (\# \text{ positive fragments})$$

For each window, the score of the test window was divided by that of the background. All scores for the chromosome were shuffled 10K times to determine FDR 0.01 threshold.

## HiC data usage/programs

HiC data was visualized (observed and observed pearson) using Juicebox [Durand et al., 2016] with data generated by Dixon et al [Dixon et al., 2012]. Compartment analysis was performed as described previously [Dixon et al., 2015]. Briefly, observed/expected matrix at 1Mb resolution for Chromosome 6 was retrieved from Juicebox. Data was processed in R (V3.4.1), where Pearson Correlation matrix and principal component analyses were performed using the ‘cov’ and ‘eigen’ functions.



## ATAC-seq library prep and sequencing

ATAC-seq was performed as published in Buenrostro, et al.[Buenrostro et al., 2015] Briefly, retinas were dissected from P14 *WT* and *Crx*<sup>-/-</sup> mice and washed in PBS. Tissue was dissociated at 37C using 2% Collagenase in TESCA buffer for 13 min, and the reaction stopped by the addition of 2X volume of DMEM + 10% FBS. DNase1 (0.5 Units) was added for the final 3 minutes to minimize clumping of cells. Cells were counted with hemocytometer, and 50,000 re-suspended in TD buffer for a 1 hour incubation with TDE1 at 37C. Remaining library prep was performed as published. Libraries were pooled and sequenced using the Illumina 2500.

## RNaseq library prep mapping and genotype comparison

Retinas were dissected from P14 mice and immediately frozen on dry ice. Samples were stored at -80C. When all samples had been collected, retinas from 2 mice (1M, 1F) were pooled and RNA was collected using the 5Prime PerfectPure RNA Tissue Kit (Fisher Scientific). Libraries were prepared, sequenced and data analyzed as described in Chapter 3 Methods. All data represent comparisons between 3 *WT* and 4 *PPR*<sup>-/-</sup> biological replicate samples.

## 4C Interaction enrichment and RNA analyses

All overlap of 4C data with previously published datasets were performed using bedtools (V2.24.0) intersect function. Expected values were computed by merging adjacent 4C positive regions, and permuting these either over the whole chromosome, within only their same compartment, or only within the *cis/trans* areas of the A compartment using the bedtools shuffle function. RNaseq analysis was performed by determining the TSS of each gene, which was used to overlap with 4C positive and AB compartment regions using bedtools intersect.

External datasets used in these experiments include the following

- CRX ChIP-seq: GSE20012
- CTCF ChIP-seq: GSE87064
- ChromHMM classifications generated in [Aldiri et al., 2017]; bed files generously provided by the authors
- DNase1Retina 8wk: GSM1014175;
- WT P14 H3K27Ac and H3K4me3 ChIP-seq were generated and discussed within Chapter 4 of this dissertation

## Droplet digital PCR (ddPCR)

Gene expression analysis was performed using the digital PCR system (BioRad, Hercules, CA). A 20ul ddPCR reaction mixture with ddPCR Supermix for Probes (no dUTP; (BioRad) containing Rho and Actb probe sets (Mm01184405\_m1 and Mm02619580\_g1 respectively; Thermo Fisher Scientific, Waltham, MA) was prepared according to manufacturer's directions, droplets generated, and nano-reactions cycled on a C1000 Touch Thermal Cycler (BioRad). The average number of allelic transcripts per biological replicate (n=3) was then determined with the QX 200 Droplet Reader and QuantaSoft Analysis Pro (BioRad).

## FISH Probe Generation

BAC cultures were obtained from the CHORI BACPAC repository (Oakland, CA). Clones used are listed below. E. coli were grown under normal conditions and BAC DNA harvested using the Nucleobond Xtra BAC Purification Kit (Machery-Nagel, Duren, Germany). 5ug of BAC DNA was used for probe generation using the Nick Translation Kit with either Digoxigenin-11-dUTP or Biotin-16-dUTP (Roche, Basel, Switzerland). Probe was cleaned using PCR purification columns (Qiaquick, Qiagen, Hilden, Germany) and resuspended in 100ul EB and frozen.

<i>Rho</i>	RP24-288F10
<i>mir182</i>	RP24-314K9 / RP23-385J22
<i>Ptms</i>	RP24-330J12
<i>Pde6h</i>	RP23-79O6
<i>Trcb</i>	RP23-421M9

## FISH Tissue Preparation and *in situ* hybridization

Retinas were dissected into 4% PFA and fixed overnight at 4C. Tissue was washed in PBS and dehydrated using a sucrose (in PBS) gradient at 4C as follows: 10% 1hr, 20% 1hr, 30% overnight. An equal volume OCT (Tissue-Tek, optimum cutting temperature formulation; Sakura, Netherlands) was added and incubated at 4C for 8 hours before tissue was transferred to 100% OCT for overnight incubation at 4C. Tissue was frozen with dry ice in OCT and stored at -80C for later use. Tissue sections (20um) were cut by cryostat and stored on slides also at -80C.

Slides were removed from -80C and dried on a hot plate (37C) for 1 hour. OCT was removed by 5min incubation with PBS, and tissue was briefly fixed to the slide to avoid release (5min 4% PFA), before slides were once again washed with PBS (5min). Slides were equilibrated in 10mM Sodium Citrate (pH 6) for 5 minutes before transferred into the same buffer at 95C and incubated for 5min. The coplin jar was removed from the double boiler and

allowed to cool at room temperature for 1 hour. Slides were transferred to 2XSSC (5min) and equilibrated in 2XSSC+50% Formamide (Millipore-Sigma, St. Louis, MO) for 5min. Slides were incubated on hot plate with 2XSSC+50% formamide for 3 minutes before being submersed into ice cold (-20C) 70% EtOH for 2 min. Slides were then further dehydrated in 90% EtOH (2min) and 100% EtOH (5min) before being allowed to dry at room temperature for 15min. Probe mixture for each slide was prepared as follows: 2ul DIG labelled probe, 2ul Biotin labelled probe, 2ul salmon sperm DNA (Invitrogen, Carlsbad, CA), 11ul formamide, and 6.6ul 3.3X FISH buffer (6.6XSSC, 33% Dextran Sulfate (Fisher Scientific, Hampton, NH)). This mixture was pre-incubate for 30min at 37C. Hydrophobic barriers were drawn on slides, allowed to dry, and slides were pre-warmed on 37C hotplate. Probe was on a coverslip to the tissue, and coverslip was affixed to the slide with rubber cement to prevent dessication. Slides were incubated at 37C in humidified chamber for 72 hours. Slides were removed from incubator, rubber cement removed and washed in 2XSSC at 37C for 5 min. Slides were transferred to 60C Wash 1 (0.4XSSC, 0.3% NP40) for 5 minutes, followed by room temperature Wash 2 (2XSSC, 0.1% NP40) for 5 minutes. Slides were again washed in 2XSSC for 5 min at room temperature before application of Rhodamine labelled anti-Digoxigenin (Roche) and Alexa Fluor 488 conjugated Streptavidin (Life Technologies, Carlsbad, CA) in blocking buffer (2XSSC, 0.2% Tween20, 1% BSA) and incubated at 37C for 1hr. Slides were washed 3x10min in wash buffer (4XSSC, 0.2% Tween20), followed by 2x5min in 2XSSC. Coverslips were applied using Vectashield hard-mount media with DAPI (Vector Labs, Burlingame, CA).

## **FISH imaging and analysis**

Slides were imaged at 100X magnification on a Leica DM5500B microscope with Leica DFC365FX camera (Leica Microsystems, Buffalo Grove, IL) with at least 40 optical sections (z-step 0.194um). Stacks of images were first processed using the Leica software deconvolution algorithm before export. 21 sections from the center of the stack were imported to the ImageJ image analysis program (all analysis was performed using Fiji [Schindelin et al., 2012]). Images were converted to a stack and a large region of interest was cropped that included only the outer nuclear layer (mainly rod photoreceptors). Channels were split, and filtered using the Mean 3D and 3D Fast Filters (TopHat) functions of ImageJ. Positive probe locations and centers of mass were then determined in red and green channels using the 3D Objects Counter. Center of mass information was saved in text files, and custom perl scripts were used to determine and report distance between nearest independent pairs of probes in 3-dimensional space.

**Table A2.1: Number of FISH experiments completed for each probe set**

	<i>WT</i>	<i>RER</i> <sup>-/-</sup>	<i>PPR</i> <sup>-/-</sup>	<i>Crx</i> <sup>-/-</sup>	<i>Nrl</i> <sup>-/-</sup>
<i>Rho - mir182</i>	4	3	6	3	3
<i>Rho - Tcrb</i>	1	1	1	1	1
<i>Rho - Ptms</i>	4	4	4	4	4
<i>Rho - Pde6h</i>	3	2	3	3	2
<i>mir182 - Ptms</i>	2	2	2	2	2
<i>Tcrb - Ptms</i>	1	1	1	1	1

## Appendix 2 References

- Aldiri, I., Xu, B., Wang, L., Chen, X., Hiler, D., Griffiths, L., Valentine, M., et al. (2017). The Dynamic Epigenetic Landscape of the Retina During Development, Reprogramming, and Tumorigenesis. *Neuron*, *94*, 550–568.
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., et al. (2000). Gene Ontology : tool for the unification of biology. *Nature Genetics*, *25*, 25–29.
- Buenrostro, J. D., Wu, B., Chang, H. Y., & Greenleaf, W. J. (2015). ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Current Protocols In Molecular Biology*, *2015*, 21.29.1-21.29.9.
- Chau, K. Y., Chen, S., Zack, D. J., & Ono, S. J. (2000). Functional domains of the cone-rod homeobox (CRX) transcription factor. *Journal Of Biological Chemistry*, *275*, 37264–37270.
- Chen, S., Wang, Q.-L., Xu, S., Liu, I., Li, L. Y., Wang, Y., & Zack, D. J. (2002). Functional analysis of cone-rod homeobox (CRX) mutations associated with retinal dystrophy. *Human Molecular Genetics*, *11*, 873–84.
- Corbo, J. C., Lawrence, K. a, Karlstetter, M., Myers, C. a, Abdelaziz, M., Dirkes, W., Weigelt, K., et al. (2010). CRX ChIP-seq reveals the cis-regulatory architecture of mouse photoreceptors. *Genome Research*, *20*, 1512–25.
- Dixon, J. R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-bourget, J. E., Lee, A. Y., Ye, Z., et al. (2015). Chromatin architecture reorganization during stem cell differentiation. *Nature*, *518*, 331–336.
- Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., et al. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, *485*, 376–380.
- Durand, N. C., Robinson, J. T., Shamim, M. S., Machol, I., Mesirov, J. P., Lander, E. S., & Aiden, E. L. (2016). Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Systems*, *3*, 99–101.
- Furukawa, T., Morrow, E. M., Li, T., Davis, F. C., & Cepko, C. L. (1999). Retinopathy and attenuated circadian entrainment in Crx-deficient mice. *Nature Genetics*, *23*, 466–470.
- Hao, H., Kim, D. S., Klocke, B., Johnson, K. R., Cui, K., Gotoh, N., Zang, C., et al. (2012). Transcriptional Regulation of Rod Photoreceptor Homeostasis Revealed by In Vivo NRL Targetome Analysis. *PLoS Genetics*, *8*, e1002649.
- Heinz, S., Benner, C., Spann, N., & Bertolino, E. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Molecular Cell*, *38*, 576–589.
- Manke, T., Roider, H. G., & Vingron, M. (2008). Statistical modeling of transcription factor binding affinities predicts regulatory interactions. *PLoS Computational Biology*, *4*, e1000039.
- McLean, C. Y., Bristor, D., Hiller, M., Clarke, S. L., Schaar, B. T., Lowe, C. B., Wenger, A. M., et al. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnology*, *28*, 495–501.
- Nikolayeva, O., & Robinson, M. D. (2014). EdgeR for Differential RNA-seq and ChIP-seq Analysis: An Application to Stem Cell Biology. *Stem Cell Transcriptional Networks Methods In Molecular Biology*, *1150*, 45–79.

- Popova, E. Y., Xu, X., Dewan, A. T., Salzberg, A. C., Berg, A., Hoh, J., Zhang, S. S., et al. (2012). Stage and Gene Specific Signatures Defined by Histones H3K4me2 and H3K27me3 Accompany Mammalian Retina Maturation In Vivo. *PLoS One*, *7*, e46867.
- Reimand, J., Arak, T., & Vilo, J. (2011). G:Profiler--a web server for functional interpretation of gene lists (2011 update). *Nucleic Acids Research*, *39*, W307-15.
- Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature Biotechnology*, *29*, 24–26.
- Roger, J. E., Hiriyanna, A., Gotoh, N., Hao, H., Cheng, D. F., Ratnapriya, R., Kautzmann, M. I., et al. (2014). OTX2 loss causes rod differentiation defect in CRX-associated congenital blindness. *The Journal Of Clinical Investigation*, *124*, 631–643.
- Ruzycki, P. A., Tran, N. M., Kefalov, V. J., Kolesnikov, A. V., & Chen, S. (2015). Graded gene expression changes determine phenotype severity in mouse models of CRX-associated retinopathies. *Genome Biology*, *16*, 171.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*, *9*, 676–682.
- Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*, 671–675.
- Tran, N. M., Zhang, A., Zhang, X., Huecker, J. B., Hennig, A. K., & Chen, S. (2014a). Mechanistically Distinct Mouse Models for CRX-Associated Retinopathy. *PLoS Genetics*, *10*.
- Tran, N. M., Zhang, A., Zhang, X., Huecker, J. B., Hennig, A. K., & Chen, S. (2014b). Mechanistically Distinct Mouse Models for CRX-Associated Retinopathy. *PLoS Genetics*, *10*, e1004111.
- Vierstra, J., Rynes, E., Sandstrom, R., Zhang, M., Canfield, T., Hansen, R. S., Stehling-Sun, S., et al. (2014). Mouse regulatory DNA landscapes reveal global principles of cis-regulatory evolution. *Science*, *346*, 1007–1012.
- van de Werken, H. J. G., de Vree, P. J. P., Splinter, E., Holwerda, S. J. B., Klous, P., de Wit, E., & de Laat, W. (2012). *4C Technology: Protocols and Data Analysis. Methods In Enzymology* (1st ed., Vol. 513). Elsevier Inc.
- White, M. A., Kwasnieski, J. C., Myers, C. A., Shen, S. Q., Corbo, J. C., & Cohen, B. A. (2016). A Simple Grammar Defines Activating and Repressing cis-Regulatory Elements in Photoreceptors. *Cell Reports*, *17*, 1247–1254.
- Won, J., Shi, L. Y., Hicks, W., Wang, J., Hurd, R., Naggert, J. K., Chang, B., et al. (2011). Mouse Model Resources for Vision Research. *Journal Of Ophthalmology*, *2011*, 1–12.