Pathogenic variants in CRX have distinct *cis*-regulatory effects on enhancers and silencers in photoreceptors

James L. Shepherdson[1,2,*], Ryan Z. Friedman[1,2,*], Yiqiao Zheng[3], Chi Sun[3], Inez Y. Oh[3], David M. Granas[1,2], Barak A. Cohen[1,2], Shiming Chen[3,4,^], Michael A. White[1,2,^]

[1] Department of Genetics,

[2] Edison Family Center for Genome Sciences & Systems Biology,

[3] Department of Ophthalmology and Visual Sciences,

[4] Department of Developmental Biology, Washington University School of Medicine in St. Louis,

St. Louis, MO 63110

* These authors contributed equally to this work.

^ Correspondence: chenshiming@wustl.edu and mawhite@wustl.edu

Running Title: Effect of CRX variants on enhancers and silencers

ABSTRACT

Dozens of variants in the photoreceptor-specific transcription factor (TF) CRX are linked with different human blinding diseases that vary in their severity and age of onset. How different variants in a single TF cause a range of pathological phenotypes is not understood. We deployed massively parallel reporter assays (MPRAs) to measure changes to CRX *cis*-regulatory function in live mouse retinas carrying knock-ins of two phenotypically distinct human disease-causing *Crx* variants, one in the DNA binding domain (p.R90W) and the other in the transcriptional effector domain (p.E168d2). We found that the effects of CRX variants on global *cis*-regulatory activity patterns correspond with the severity of their phenotypes. The variants affect similar sets of enhancers but to different degrees. A subset of silencers were converted to enhancers in retinas lacking a functional CRX effector domain, but were unaffected by p.R90W. Episomal MPRA activities of CRX-bound sequences showed some correspondence with chromatin environments at their original genomic loci, including an enrichment of silencers and depletion of strong enhancers among distal elements whose accessibility increases later in retinal development. Many distal silencers were de-repressed by p.E168d2, but not by p.R90W, suggesting that loss of developmentally timed silencing caused by p.E168d2 may contribute to phenotypic differences between the two variants. Our findings indicate that phenotypically distinct disease variants in different domains of CRX have partially overlapping effects on its *cis*-regulatory function, leading to mis-regulation of similar sets of enhancers, while having a qualitatively different impact on silencers.

INTRODUCTION

The transcription factor (TF) Cone-Rod Homeobox (CRX) plays a key role in the differentiation and maintenance of photoreceptors (Chen et al. 1997; Furukawa et al. 1997; Freund et al. 1997; Swaroop et al. 2010). Variants in *CRX* have been shown to cause several inherited retinal diseases, including Retinitis Pigmentosa (RP), Cone-Rod Dystrophy (CoRD), and

Leber Congenital Amaurosis (LCA) (Sohocki et al. 1998; Swaroop et al. 1999; Rivolta et al. 2001; Perrault et al. 2003; Nichols et al. 2010; Koyanagi et al. 2019; Fujinami-Yokokawa et al. 2020). These retinopathies present with different phenotypes, varying in their cone vs. rod predominance, severity, and age of onset. Some *CRX* variants have been shown to cause severe dominant disease, while others are pathogenic only in a recessive context (Tran and Chen 2014). *CRX* is the only gene implicated in all three of these diseases.

CRX functions by binding sequence motifs in non-coding *cis*-regulatory elements (CREs), including enhancers and promoters, and recruiting cellular transcriptional machinery (Chen et al. 1997; Peng and Chen 2007). Using Massively Parallel Reporter Assays (MPRAs) conducted in wild-type and *Crx*$^{-/-}$ retinal explants, we previously characterized the transcriptional activity of large libraries of genomic and synthetic DNA sequences that are bound by CRX *in vivo* or contain CRX-binding sequence motifs (White et al. 2013; White et al. 2016; Friedman et al. 2021). These studies revealed some features of the sequence grammar of CRX transcriptional regulation, but how pathogenic variants alter CRX activity across the full range of its target CREs remains unknown. Variants in different protein domains of CRX may differ in their effects on binding affinity, target specificity, and interactions with co-factors, and thus distinct pathogenic variants could alter CRX activity in different ways. Understanding how distinct variants in the same TF change its regulatory function could reveal important differences in the mechanisms by which these variants cause disease. Previous RNA-seq studies of knock-in mouse models carrying the retinopathy-associated variants *p.R90W* or *p.E168d2* suggest that they have graded effects on gene expression that correspond with the severity of their retinal phenotypes (Ruzycki et al. 2015; Tran et al. 2014). Genes affected by p.R90W are also mis-regulated by the more phenotypically severe p.E168d2, which causes additional mis-regulation. However, changes in gene expression are not always a direct measure of changes in the regulatory activity of a specific TF, since RNA levels can be altered by indirect effects, such changes in concentrations of other TFs or by factors that

alter transcript stability. To directly measure the effects of variants on CRX regulatory function, we used MPRAs to assay the activity of 1,723 CRX-bound CREs in retinas from mouse models carrying knock-ins of human disease-causing variants at the native *Crx* locus.

We tested these CREs in retinal explants from mice wild-type for *Crx* or carrying prototypical pathogenic *CRX* variants: the *p.R90W* single residue substitution variant and the *p.E168d2* frameshift variant (Figure 1A). Arginine 90 lies in the DNA-binding homeodomain of CRX, and the R90W variant has been shown to severely weaken binding to target CREs and reported to cause mild late-onset CoRD or LCA depending on zygosity (Swaroop et al. 1999; Tran et al. 2014; Zheng et al. 2023). Glutamic acid 168 is located within the WSP domain of the transcriptional effector region, and frameshift into the +2 reading frame is predicted to lead to a stop codon within 3 residues. Because this stop codon is located within the final exon, the *p.E168d2* transcript is not predicted to undergo nonsense-mediated decay (Karousis and Mühlemann 2019). Although much of the transcriptional effector domain is lost, the p.E168d2 variant retains the ability to bind DNA (Chau et al. 2000; Tran et al. 2014). The p.E168d2 variant has been associated with dominant LCA in patients Tran et al. (2014).

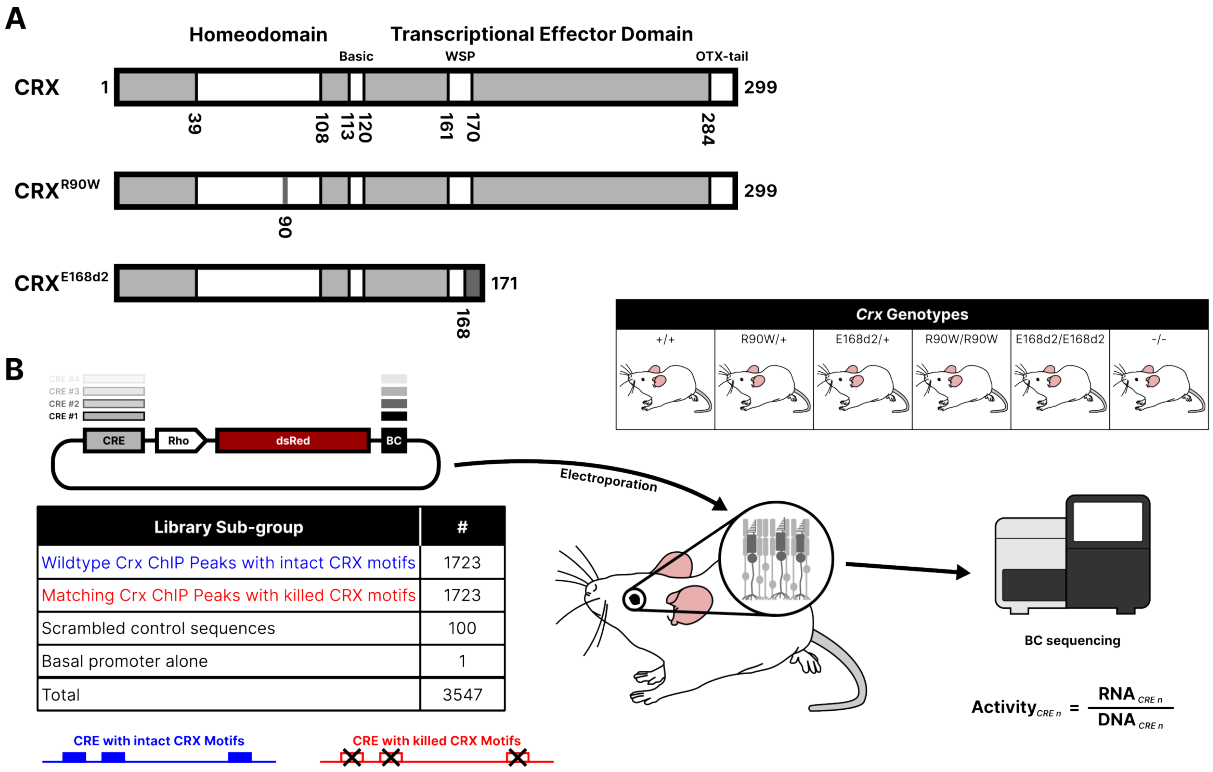
We found that the global changes in measured CRE activity across the *Crx* mutant genotypes broadly correlate with phenotypic severity and the degree of gene expression changes (Ruzycki et al. 2015). We found that the effects of the mutants on enhancers were similar, and differed primarily in their severity. The effects on silencers, however, differed qualitatively between the mutants. We identified potentially bifunctional CREs that behave as transcriptional silencers in wildtype retinas but act as enhancers in the $Crx^{E168d2/E168d2}$ and $Crx^{-/-}$ genotypes. These CREs exhibit both silencer or enhancer activity, depending on *Crx* genotype, and contain sequence features characteristic of both enhancers and silencers. Our results show that the biochemically distinct p.R90W and p.E168d2 variants have largely overlapping effects on the *cis*-regulatory

function of CRX, though the loss of the transcriptional effector domain has specific effects on a subset of CRX-bound silencers.

RESULTS

Biochemically distinct CRX variants differ largely in the severity, but not quality, of their effects

To construct an MPRA library that tests the effects of variants on the *cis*-regulatory function of CRX, we selected 1,723 genomic CRE sequences, each containing at least one CRX binding motif, that are bound by CRX as measured by ChIP-seq in whole adult mouse retina (Ruzycki et al. 2018; Corbo et al. 2010). For each CRE, we included a matching motif mutant, using a mutation known to abrogate TF binding (White et al. 2013; Lee et al. 2010) (Figure 1B). We cloned these 3,446 elements in two libraries, one with each CRE located upstream of a minimal promoter derived from the endogenous *rhodopsin* (*Rho*) locus driving expression of the fluorescent protein dsRed, and one substituted with the mouse *hsp68* minimal promoter. Each CRE could be uniquely identified by a short sequence barcode located in the 3' untranslated region of the dsRed transcript. In each library, we also included 100 scrambled elements as negative controls, with sequence composition matching that of the average CRE but with no detectable CRX binding motifs as called by FIMO using a p-value cutoff of 0.0025 (Grant et al. 2011).



Library Sub-group	#
Wildtype Crx ChIP Peaks with intact CRX motifs	1723
Matching Crx ChIP Peaks with killed CRX motifs	1723
Scrambled control sequences	100
Basal promoter alone	1
Total	3547

Figure 1: (**A**) Schematics of wildtype CRX protein and the p.R90W and p.E168d2 pathogenic variants, containing a point DNA binding domain mutant or resulting in truncation of the transcriptional effector domain, respectively. (**B**) Outline of the experimental procedure for constructing and testing the CRE libraries. A library of CRE sequences is cloned upstream of a rhodopsin ("Rho", shown) or hsp68 minimal promoter, driving expression of the dsRed fluorescent protein, with each CRE marked by a unique sequence barcode ("BC") in the 3' UTR. This library was electroporated into retinal explants with the six indicated Crx genotypes. RNA was collected from the retinas, and transcript counts for each element were measured and normalized to abundance in the input DNA library to calculate transcriptional activity.

We introduced these CRE libraries via electroporation into P0 retinal explants from wild-type C57BL/6J mice (+/+), mice heterozygous or homozygous for the *p.R90W* (*R90W/+* and *R90W/W*) or *p.E168d2* (*E168d2/+* and *E168d2/d2*) variant, and *Crx* knockout mice (-/-) (all on C57BL/6J background) (Tran et al. 2014). After extracting RNA at P8, we sequenced CRE barcodes and normalized read counts to the abundance of each element in the input plasmid library (retina RNA read count / plasmid DNA read count) to compute each element's transcriptional activity. This method generates highly reproducible measurements of CRE transcriptional activity, and Pearson

correlations between biological replicates are correspondingly high (*Rho* minimal promoter library mean = 0.80; Figure S1).

We classified each element into one of five functional classes based on their transcriptional activity relative to the basal promoter alone: strong enhancer (significant increase in transcriptional activity relative to basal and greater than the 95th percentile of the scrambled sequences), weak enhancer (significant increase in transcriptional activity relative to basal and less than or equal to the 95th percentile of the scrambled sequences), inactive (no significant transcriptional activity difference compared to basal), weak silencer, and strong silencer (using the same criteria and cutoffs as for enhancers, but with significantly less transcriptional activity than basal) (Friedman et al. 2021).

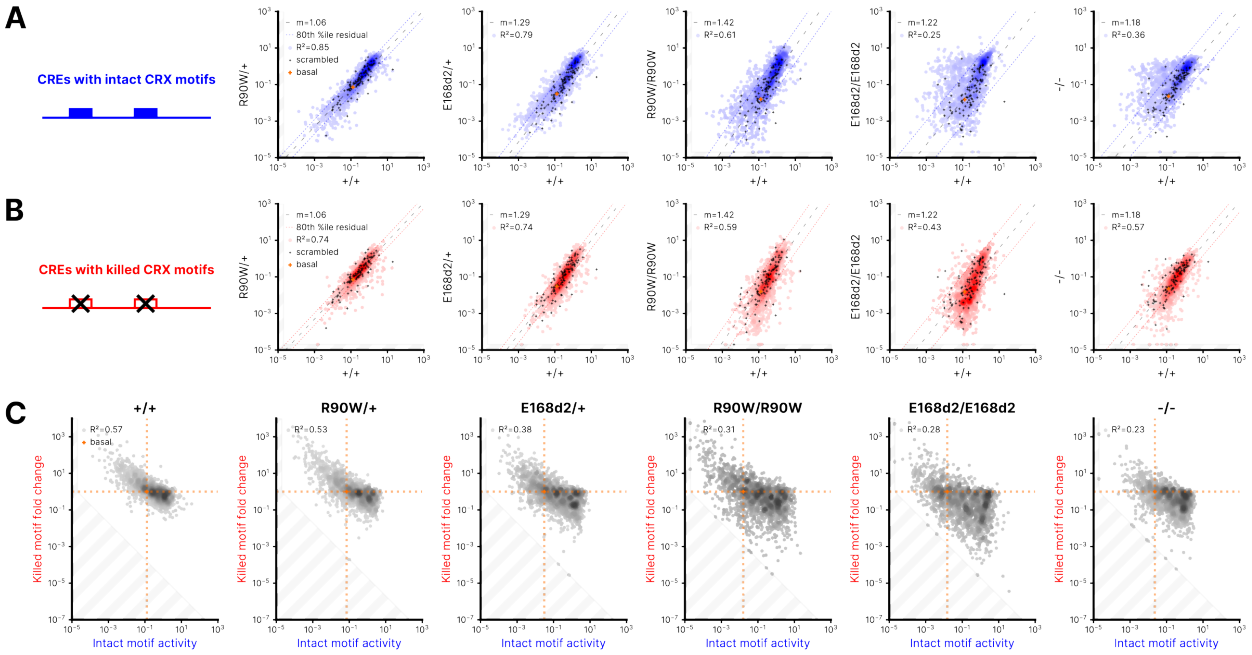


Figure 2: (**A,B**) Correlations of the transcriptional activity of each genomic CRE sequence with intact CRX motifs (**A**) or CRE sequence with killed CRX motifs (**B**) in the indicated genotypes. Superimposed black cross symbols indicate the transcriptional activity of the 100 scrambled control sequences, and the orange cross shows the activity of the basal rhodopsin promoter alone. Using the scrambled control sequences to establish an empirical null distribution of CRE activity, we performed robust regression using a trimmed mean M-estimator (fit plotted as a

dashed line) for each genotype comparison. The outer dotted lines indicate the 80th percentile of the residuals of the library CREs relative to the regression of the scrambled sequences. (**C**) Activity of each CRE with intact CRX motifs versus the fold change in activity for each matching CRE with killed motifs, in each of the indicated genotypes. The orange dotted lines indicate the activity of the basal promoter alone, and a fold change of 1 (no difference in activity between CRE with intact or killed CRX motifs). In all panels, transcriptional activity was adjusted for visualization purposes using a fixed pseudocount of 2e-5; the out-of-bound region is indicated by the shading in light grey.

To first evaluate the global *cis*-regulatory effects of the tested CRX variants on transcriptional activity with the *Rho* minimal promoter, we compared the expression of each CRE in retinas from +/+ mice to that in retinas from each of the CRX variant mouse lines (Figure 2A). We observed an increasingly pronounced global dysregulation of CRE activity that correlated with increasing phenotypic severity. Prior characterization of these mouse models found that phenotypic severity corresponded with the identity and zygosity of the mutants (Tran et al. 2014). Little retinal impairment was observed in *R90W/+* mice, while progressive rod dystrophy occurred in *E168d2/+* mice. Unlike the heterozygotes, all homozygous mutant mice are blind, with *E168d2/d2* mice exhibiting the most severe photoreceptor degeneration among the homozygous mutant mice. Consistent with the relative phenotypic severities of the genotypes, we found that CRE activity in retinas from mice carrying the least-severe *R90W/+* genotype was tightly correlated with activity in corresponding +/+ retinas (R^2 = 0.87). In contrast, CRE activity in retinas with the severe *E168d2/d2* genotype was considerably deviated from matched +/+ measurements (R^2 = 0.26). The most striking difference observed between the mutants was the change in silencer activities in the homozygous mutant genotypes. Moderate de-repression of some silencers was observed in *R90W/W* retinas, while a much stronger loss of repression was observed in *E168d2/d2* and -/- retinas, with some CREs that were silencers in wild-type retinas gaining significant activity in these genotypes. This suggests that the transcriptional effector domain of CRX is critical for its function at silencers. The greater loss of repression observed in *E168d2/d2* retinas relative to *R90W/W* retinas is consistent with a prior RNA-seq analysis, which found that 37% (248 of 673) of

differentially expressed genes in *E168d2/d2* retinas were upregulated, while only 26% (70 of 265) of differentially expressed genes in *R90W/R90W* retinas were upregulated ($p = 0.003$, Fisher's exact test) (Ruzycki et al. 2015). These upregulated genes, many of which are cone photoreceptor-specific, are generally not bound by CRX within 1 kb of the TSS and thus their increase in expression may be an indirect effect of the mutants (Ruzycki et al. 2015). However, the disproportionate degree of derepression in *E168d2/d2* retinas observed in both RNA-seq and MPRA datasets suggests that, relative to the p.R90W variant, the p.E168d2 variant may have a disproportionate impact on CRX function at distal silencers.

The *Rho* minimal promoter common to all library members contains CRX binding motifs, and thus its intrinsic transcriptional activity is also altered by pathogenic CRX variants. This effect can be observed in the regression slope calculated from the 100 scrambled control sequences, which deviates in the more severe genotypes. Although global transcriptional perturbation in the severe genotypes contributes to a greater divergence of transcriptional activity even in the scrambled control sequences, a general weakening of *Rho* minimal promoter activity can be observed (Figure 2A).

Measurements of the *hsp68* minimal promoter library demonstrate broadly similar CRE transcriptional activity effects in the pathogenic genotypes (Figure S2, Figure S3A). However, the low basal transcriptional activity of the *hsp68* minimal promoter itself resulted in low transcriptional activity of all tested CREs, contributing to measurement noise and dropout of some library members (particularly pronounced in the lowest-activity library members). Despite this measurement noise, the effects of the different genotypes on basal promoter activity observed for the *Rho* minimal promoter are not reproduced for the *hsp68* minimal promoter, consistent with its lack of CRX binding motifs.

Examining the activity of the mutant CREs lacking CRX binding motifs, we found that much of the dysregulated activity in the homozygous genotypes depends on CRX motifs (Figure 2B).

This indicates that these are direct effects due to altered CRX *cis*-regulatory function at these CREs. Across all genotypes, most CREs exhibit CRX motif-dependent activity (Figure 2C). In the most-severe *E168d2/d2* and *-/-* retinas, the increased activity of de-repressed silencers is lost in the motif mutants, suggesting that an additional TF, likely the CRX ortholog OTX2 (Koike et al. 2007), acts at these sites in *-/-* and possibly *E168d2/d2* retinas. Together, these results show that silencers depend on both CRX motifs and on the CRX effector domain, indicating that CRX acts as a direct repressor at these CREs.

A subset of CREs convert from silencers to enhancers in *E168d2/d2* and *-/-* retinas.

Most assayed CREs exhibit a graded response across the CRX mutants, with the effects differing primarily in their severity. However, a subset of silencers exhibit not only loss of repression, but an increase in activity in *E168d2/d2* and *-/-* retinas (Figure 3A). Examining these CREs, we found 235 sequences classified as weak or strong silencers in wild-type retinas that become weak or strong enhancers in retinas lacking the CRX transcriptional effector domain (Figure 3B, Figure S4). Comparing CREs that convert from silencers to enhancers and their matched motif mutants across all genotypes, we observed an inversion of the effect of motif mutations. (Figure 3C). Specifically, mutating the CRX motifs caused activity to increase in retinas expressing *Crx* with an intact transcriptional effector domain, while the same motif mutations caused activity to decrease in the *E168d2/d2* and *-/-* retinas. These results demonstrate that these CREs act as motif-dependent silencers in the presence of the CRX transcriptional effector domain, and as motif-dependent enhancers upon the loss of this domain.

To further characterize these CREs, we examined two sequence features that distinguish CRX-bound enhancers from CRX-bound silencers, as reported in our prior work (Friedman et al. 2021). Compared to enhancers, CRX-bound silencers tend to contain more copies of the CRX motif, which we quantify using a threshold-free scan called CRX predicted occupancy (White et al. 2013; White et al. 2016; Friedman et al. 2021). Enhancers are characterized by the presence

of a more diverse collection of motifs for eight other lineage-specific TFs. We quantify motif diversity with a score called information content, based on Boltzmann entropy, that considers the number and diversity of motifs in a CRE. Information content considers motifs or motif families from CRX/OTX2, NRL (leucine zipper), NEUROD1 (E-box binding factor), RORB (nuclear receptor), MAZ/SP4 (C2H2 zinc finger), MEF2D (MADS-box), RAX (Q50 homeodomain) and GFI1 (CH2H2 zinc finger repressor). We previously reported that CRX-bound enhancers have higher information content than silencers (Friedman et al. 2021). Thus, enhancers are characterized by high information content and modest CRX predicted occupancy, while silencers are characterized by low information content and high CRX predicted occupancy.

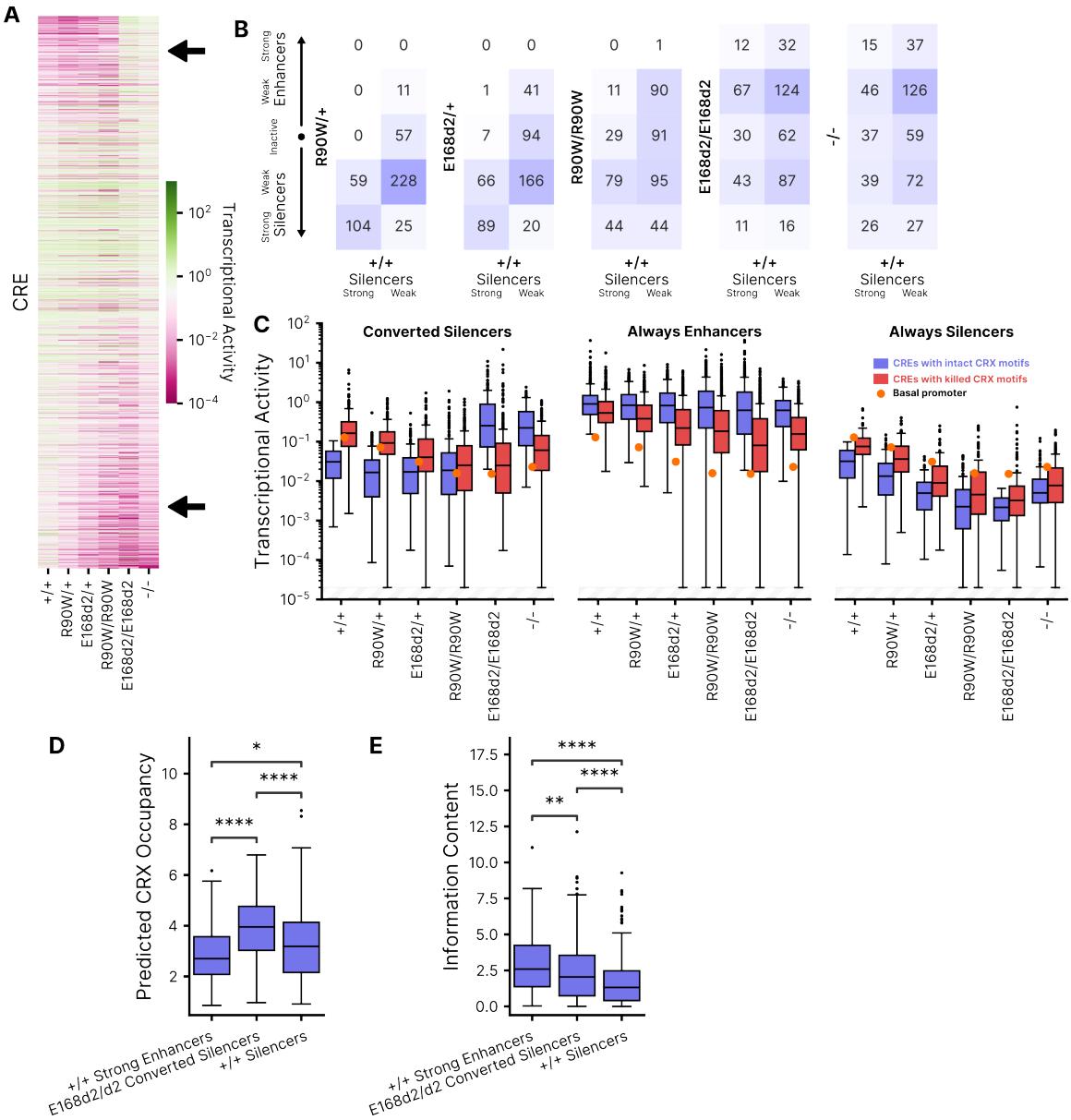


Figure 3: (**A**) Heatmap comparing the transcriptional activity of each CRE across the six CRX genotypes. In the top right (upper arrow), note a group of CREs with generally low activity in all but the E168d2/d2 and -/- genotypes. Along the bottom (lower arrow) note the stepwise reduction in activity with increasing phenotypic severity. CREs are sorted by the ratio of transcriptional activity in +/+ and -/- genotypes. (**B**) Classifications across genotypes for CREs classified as strong silencers (left column of each heatmap) or weak silencers (right column) in wildtype retinas. (**C**) Transcriptional activity of wildtype or motif mutant CREs across genotypes for 235 silencers that convert to enhancers in E168d2/d2. The orange dot indicates the transcriptional activity of the basal promoter alone in each genotype. (**D**) Predicted CRX occupancy of silencers that convert to enhancers, versus CREs that act as strong enhancers in +/+ and CREs that act as silencers in +/+ that are not derepressed in E168d2/168d2. (**E**) Information content of the same subgroups as panel A. *: $p < 0.05$, **: $p < 0.01$, ****: $p < 0.0001$ via two-sided Mann-Whitney-Wilcoxon test. In panel C, transcriptional activity was adjusted for visualization purposes using a fixed pseudocount of 2e-5; the out-of-bound region is indicated by the shading in light grey.

We determined the predicted CRX occupancy and information content of CREs that convert from silencers to enhancers in *E168d2/d2* and -/- retinas, and found that these CREs have features of both silencers and enhancers. Compared with CREs marked as strong enhancers in wild-type retina, silencers that convert to enhancers have high predicted CRX occupancy (Figure 3D). Notably, the CRX predicted occupancy of these potentially bifunctional CREs is even higher than the elevated predicted occupancy of the other silencers (median 3.95 vs. 3.19). Unlike other silencers, CREs that convert to enhancers also have high information content, nearly as high as the information content of strong enhancers (median 2.05 vs. 2.59) (Figure 3E). The mixed sequence characteristics of these CREs may explain why they convert from silencers to enhancers in *E168d2/d2* and -/- retinas. These CREs might function as silencers in wild-type retina due to their high CRX occupancy. Upon loss of the CRX transcriptional effector domain, these CREs cease to function as silencers, possibly due to the loss of interactions with a CRX-interacting repressor complex. However, the presence of other TF motifs (high information content) enables these CREs to function as enhancers in *E168d2/d2* and -/- retinas, in contrast to low-information content silencers that simply lose their repressive activity. The additional TFs recruited by these diverse binding motifs may interact with OTX2 or some other TF that at the CRX motifs in the *E168d2/d2* and -/- genotypes.

We used Simple Enrichment Analysis (SEA) to search for additional enriched motifs in these potentially bifunctional CREs (Bailey and Grant 2021). Relative to enhancers and silencers in wild-type retina, as well as CREs that remain enhancers and silencers in *E168d2/d2* retina, we found that CREs that convert from silencers to enhancers are enriched in motifs for a variety of paired-type homeodomain TFs, as well as homeodomain dimer motifs. CRX and its ortholog OTX2 are K50 paired-type homeodomains that contain a lysine at residue 50 of the DNA binding domain which determines the specificity of these TFs for a TAATCC consensus sequence. While CRX and OTX2 play critical roles specifically in photoreceptors and bipolar cells, other non-K50

paired-type homeodomain TFs are active in multiple retinal cell types (Hennig et al. 2008; Zagozewski et al. 2014). Silencers that convert to enhancers are enriched in motifs for TFs active in several retinal cell types: LHX-family TFs (horizontal and Mueller glial cells, $p = 2.02x10^{-8}$); VSX1 and VSX2 (cone bipolar cells and retinal ganglion cells, $p = 1.23x10^{-5}$); ISL2 (retinal ganglion cells, $p = 2.87x10^{-6}$); VAX1 and other NK-family TFs (role in ventral retina dorsalization, $p = 1.78x10^{-5}$); and RAX (photoreceptors and bipolar cells, $p = 2.13 \times 10^{-4}$). These CREs are also strongly enriched in homeodomain motif dimers. Compared CREs that remain enhancers and silencers in wild-type and *E168d2/d2* retina, CREs that convert from silencers to enhancers exhibit a 1.5-fold enrichment of homeodomain motif dimers (71.2% of bifunctional CREs contain at least 1 dimer versus 47.4% among silencers and enhancers, $p = 1.78x1^{-10}$). Paired-type homeodomain TFs bind as both homo- and hetero-dimers, and the presence of these dimers in potentially bifunctional CREs may enable these elements to play regulatory roles in multiple cell types (Wilson et al. 1993; Rister et al. 2015; Hughes et al. 2018).

Finally, we asked whether potentially bifunctional CREs are co-bound by the co-repressor BCOR, a Polycomb repressive complex 1 (PRC1) factor that co-binds with CRX and is linked with early-onset retinal degeneration (Langouët et al. 2022). Only a minority of all silencers in wild-type retina overlapped a BCOR binding site (11.8%). However, silencers that convert to enhancers in *E168d2/d2* and *-/-* retinas were significantly less likely to be bound at P0 by BCOR (4.7%, chi square test $p = 7.2x10^{-4}$), suggesting that these potentially bifunctional elements are distinct from other silencers, and that their silencing may not depend on PRC1.

Differences in episomal MPRA activity correspond with differences in chromatin environments of CRX-bound sequences

CREs in the MPRA library were selected from different chromatin environments as measured in wild-type retina. All library CREs were selected from sites of accessible chromatin (measured by ATAC-seq at P14) (Ruzycki et al. 2018) that overlapped CRX ChIP-seq peaks (measured in adult retina) (Corbo et al. 2010), but these sites varied in other chromatin annotations as

determined in a previous study (Ruzycki et al. 2018). These annotations include 1) the presence or absence of histone marks H3K4me3 and H3K27ac in P14 wild-type retina, and 2) the loss or retention of chromatin accessibility in P14 *Crx*$^{-/-}$ retina (*i.e.*, whether accessibility is CRX dependent or independent). Based on these annotations, CRX-bound accessible sites were previously classified into four groups (Table 1) (Ruzycki et al. 2018). Group A sequences are proximal regulatory elements that are positive for H3K4me3 and H3K27ac and typically lie within 5 kb of the nearest TSS. Groups C and D are distal regulatory elements, though at P14 group D sites lack both H3K4me3 and H3K27ac. (Group B sites are rare and were not included in the MPRA library.) Within each group, CRX-bound sites exhibit either CRX-dependent or independent chromatin accessibility. A previous analysis found that these chromatin annotations from P14 retina strongly correlated with the cell-type specificity and developmental timing of the accessibility of these sites (Ruzycki et al. 2018). Group A independent sequences are accessible in multiple non-retinal cell types and are open in the retina throughout postnatal development. Group A dependent sites are retina-specific promoters that are largely inaccessible at P1 but increase their accessibility over 8 weeks of postnatal development. Group C and D sites are retina-specific distal regulatory elements whose accessibility increases over development, with independent sites showing earlier activation than dependent sites.

Group	Histone Marks		Typical Distance to TSS	CRX-dependent Accessibility	Timing and specificity of accessibility	# CREs in MPRA library
	H3K4me3	H3K27ac				
A	+	+	0-5 kb	Dependent	Retina-specific, late accessibility	249
				Independent	Open in other cell types, always accessible	632
C	-	+	5-500 kb	Dependent	Retina-specific, late accessibility	311
				Independent	Retina-specific, early accessibility	366
D	-	-	5-500 kb	Dependent	Retina-specific, late accessibility	76
				Independent	Retina-specific, early accessibility	89

Table 1: Features of chromatin environments of MPRA library sequences.

We found that CREs taken from different chromatin environments exhibited systematic differences in episomal MPRA activity in P8 wild-type retina, and that these differences correspond with the developmental timing and CRX-dependence of accessibility at their original genomic loci (Figure 4A). Group A CRX-independent CREs exhibit on average the highest MPRA activity, and contain the highest proportion of enhancers and the lowest proportion of silencers of all subgroups tested. This is consistent with their strong chromatin accessibility throughout retinal development. Compared to the CRX-independent CREs, Group A CRX-dependent CREs contain fewer enhancers and more silencers, and exhibit significantly lower average activity (mean log2 RNA/DNA = -1.083 vs. -0.477 for dependent vs. independent Group A, $p = 5.0 \times 10^{-4}$, independent two-sample t-test). The weaker activity of Group A dependent CREs is consistent with their later activation in development, reaching a peak weeks after P8 when MPRA activity was measured. A similar trend was observed for the distal regulatory elements of Group C, which also show increased activation over development. Group C dependent sequences also contain proportionally more silencers and fewer enhancers, and their average MPRA activity was similar

to that of Group A dependent CREs (mean log2 RNA/DNA = -1.101 vs -1.083 for Group C dependent vs Group A dependent). Group C independent CREs, which show earlier developmental activation than dependent CREs, exhibited intermediate MPRA activity (mean log2 RNA/DNA = -0.672) and intermediate proportions of enhancers and silencers. Finally, more than half of Group D CREs are silencers, consistent with their lack of active chromatin marks at P14 and their later developmental activation. The correspondence between episomal MPRA activity and features of the genomic chromatin environment suggests that the sequence-encoded features responsible for MPRA activity also drive key chromatin changes at native genomic loci over the course of retinal development. Furthermore, the increased fraction of silencers among regulatory elements that show later chromatin activation suggests that these CREs may be actively silenced early in development, preventing premature expression of later developmental genes.

CREs from different chromatin groups also exhibit different responses to CRX mutants, consistent with differences in their proportion of silencers and enhancers. The proximal Group A CREs show the strongest loss of average activity in *R90W/W*, *E168d2/d2* and *-/-* retinas (Figure 4B). Distal Group C CREs show a weaker average loss of in these genotypes, likely because these elements have not yet been strongly induced at P8 and because this group includes a larger fraction of de-repressed silencers. Notably, Group C independent CREs exhibit the least change in the mutant retina. This suggests that other TFs contribute more than CRX to the activity of these CREs, and these other TFs may be responsible for continued accessibility of these sites in *-/-* retina. Group D sites show significantly higher activity in the mutant retina, which reflects the fact that this group contains a high fraction of distal silencers that are derepressed in the mutant lines. We also found that 69% of CREs that convert from silencers to enhancers in *E168d2/d2* and *-/-* retinas are distal elements (Groups C and D, Figure 4C), which is a significant enrichment compared to all CREs in the library (49% distal, chi square test $p = 1.81 \times 10^{-9}$). Thus,

changes in episomal *cis*-regulatory activity in different *Crx* genotypes correspond with differences in the original genomic chromatin environments of assayed CREs. Taken together, our findings suggest that a significant fraction of photoreceptor CREs act as distal silencers during early postnatal development, and that these silencers are most strongly disrupted by the loss of the CRX effector domain in *E168d2/d2* and *-/-* genotypes.

CRX mutants have no systematic effects on cooperativity between CRX binding sites

Included in the MPRA library are 100 CREs with exactly two CRX binding motifs. To test for effects of the p.R90W and p.E168d2 CRX mutants on homotypic cooperativity, we designed single and double CRX motif mutants for each of these CREs. If the two CRX motifs in a CRE act cooperatively, the loss of activity caused by the single motif mutants, added together, will exceed the effect of the double motif mutant. To assess cooperativity we therefore computed the sum of the effects of the single motif mutants on MPRA activity, and divided that by the effect of the double motif mutant, for each CRE. A ratio greater than one indicates that the two motifs act cooperatively, while a ratio near one indicates no cooperativity. A ratio less than one could indicate either redundancy or anti-cooperativity between the two CRX motifs. We found that very few of the 100 tested CREs showed evidence of cooperativity between CRX sites, and none of the mutant *Crx* genotypes systematically increased or decreased cooperativity (Figure S5). While the p.R90W and p.E168d2 CRX variants may affect cooperative interactions between CRX and other TFs, or may affect CRX-CRX cooperativity in a very specific contexts, they have no pervasive effects on homotypic cooperative interactions between CRX molecules.

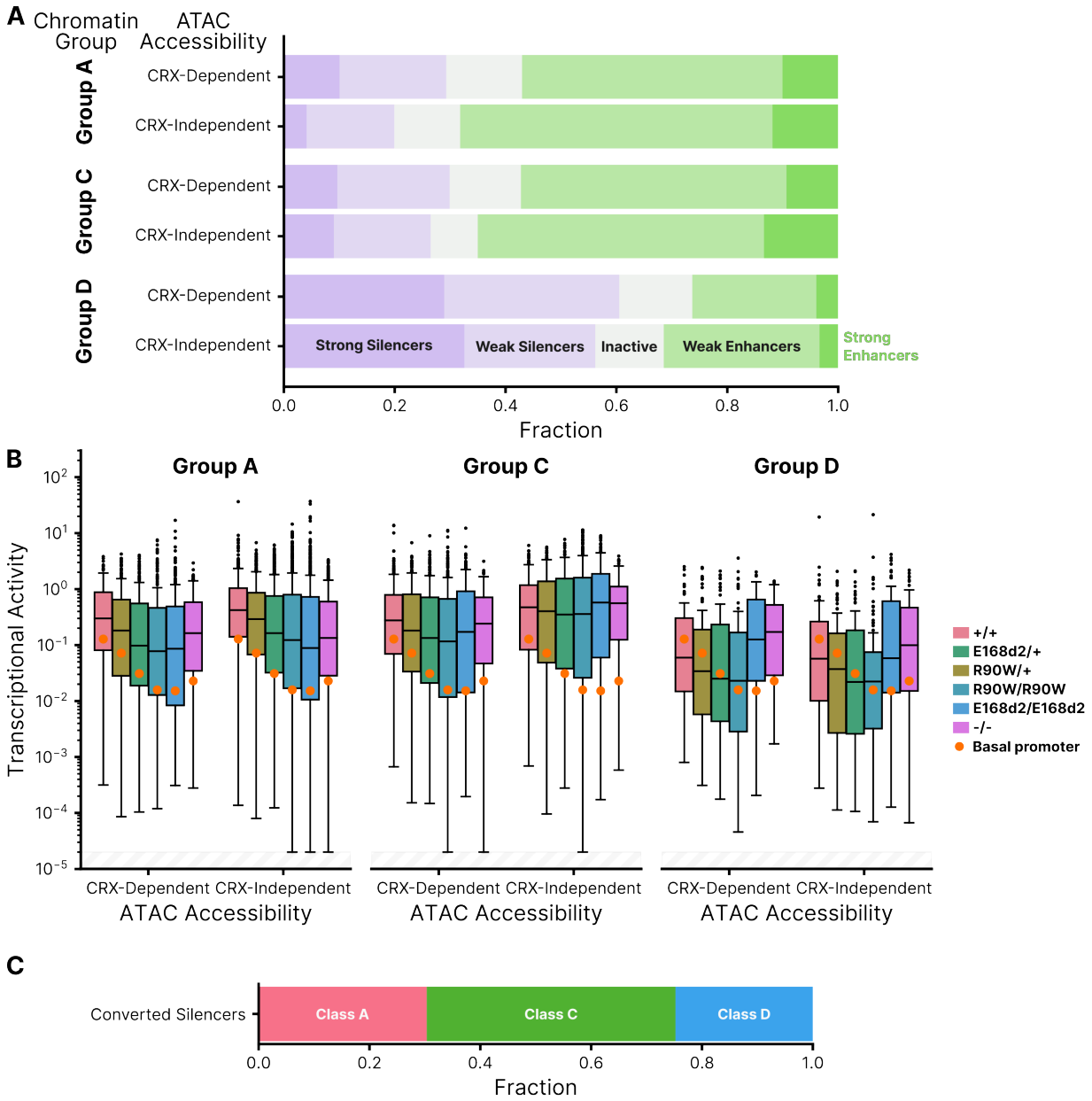


Figure 4: (**A**) Distribution of CRE activity classes in +/+ retina separated into groups based on chromatin annotations of the site of genomic origin and CRX-dependent ATAC accessibility (Ruzycki et al. 2018). Group A: H3K4me3+ H3K27Ac+, Group C: H3K4me3- H3K27Ac+, Group D: H3K4me3- H3K27Ac-. (**B**) Transcriptional activity of all CREs, separated by chromatin and ATAC group, across all genotypes. (**C**) Chromatin classification breakdown of E168d2/d2 converted silencers (elements classified as weak or strong silencers in +/+ that become weak or strong enhancers in E168d2/d2). In panel B, transcriptional activity was adjusted for visualization purposes using a fixed pseudocount of 2e-5; the out-of-bound region is indicated by the shading in light grey.

DISCUSSION

CRX plays a critical regulatory role in multiple photoreceptor cell types, and it interacts with both co-activators and co-repressors to maintain cell-type specific gene expression in rod and cone photoreceptors (Yanagi et al. 2000; Chen et al. 2005; Peng et al. 2005; Peng and Chen 2007; Mitton et al. 2003; Hlawatsch et al. 2013; Hennig et al. 2013). Because CRX controls both activation and repression of gene expression, different disease-causing variants in CRX may have distinct effects on its *cis*-regulatory function, resulting in the range of phenotypes observed in CRX-linked retinopathies (Tran and Chen 2014; Sun and Chen 2023). While genome-wide transcriptional profiling can reveal important differences in cis-regulatory function between mutants (Ruzycki et al. 2015), some gene expression changes may be due to downstream, indirect effects of TF variants or to differences in transcript stability. We thus used MPRAs in p.R90W and p.E168d2 knock-in mouse models of human disease as a more direct measure of the changes to CRX *cis*-regulatory function caused by two biochemically distinct pathogenic variants.

We observed graded degrees of global misregulation in CRE activity that correspond with the severity of the retinal phenotypes of the mouse models. This finding is consistent with a prior RNA-seq analysis in the same mouse models (Ruzycki et al. 2015) and suggests that the phenotypic differences between p.R90W and p.E168d2 are in part due to the fact that p.E168d2 causes more extreme mis-regulation at more CREs. The p.R90W and p.E168d2 variants affect largely overlapping sets of wild-type enhancers Figure 3A, lower arrow), suggesting that a significant reduction in DNA binding or loss of the transcriptional effector domain cause similar deficits at these sites. Because CRX-bound enhancers are enriched in additional motifs for other cell type-specific TFs, the loss of CRX activity at enhancers may be partially mitigated by the activity of other TFs.

In contrast to the similar effects on enhancers, the p.R90W and p.E168d2 mutants differed most in their effects on silencers. Relatively few silencers showed loss of repressive activity in *R90W/W* retinas, while much stronger de-repression occurred in *E168d2/d2* and *-/-* retinas Figure 2A). We previously observed loss of silencing in *-/-* retinas, showing that CRX is required for the repressive activity of silencers (White et al. 2016). Our results here now show that silencing specifically depends on the CRX transcriptional effector domain, which has been classically considered an activation domain (Chen et al. 1997; Peng and Chen 2007). However, a recent study found that the transcriptional effector domain of human CRX is capable of both activation and repression, and that homeodomain TFs are enriched in such bifunctional effector domains (DelRosso et al. 2023). The ability of CRX to act as both a repressor and an activator is consistent with its role as a regulator of distinct gene expression programs in multiple photoreceptor cell types Swaroop et al. (2010). Results of multiple MPRA studies show that whether CRX acts as an activator or repressor depends on *cis*-regulatory context, though the sequence features that determine that context remain poorly understood (White et al. 2013; White et al. 2016; Hughes et al. 2018; Friedman et al. 2021; Loell et al. 2023).

The greater loss of silencing measured by MPRA in *E168d2/d2* retina is consistent with the finding that proportionally more genes are up-regulated in *E168d2/d2* retina relative to *R90W/W* retina (Ruzycki et al. 2015). Many of these up-regulated transcripts are from cone photoreceptor-specific genes that are being mis-expressed in rod photoreceptors. Notably, the up-regulated genes tend to lack proximal CRX binding sites, in contrast to down-regulated genes, which frequently have a CRX-bound site within 5 kb of the TSS (Ruzycki et al. 2015). This indicates that up-regulated genes are either indirect targets of CRX, or else that they are direct targets regulated by distal CRX-bound silencers. Our results here support the latter hypothesis. We found that distal CRX-bound CREs (>5 kb from the TSS) are more enriched in silencers compared to proximal CREs, which are enriched in strong enhancers. Thus, an important effect of the p.E168d2 mutant

may be the loss of distal silencing, causing genes specific to other cell types to be mis-expressed in rod photoreceptors.

Among the mis-regulated silencers in *E168d2/d2* and *-/-* retinas are potentially bifunctional CREs that may act as silencers in rod photoreceptors and as enhancers in another retinal cell type. These silencers convert to enhancers of the *Rho* basal promoter in retinal genotypes lacking a CRX effector domain, and they contain sequence features characteristic of both silencers and enhancers. Our results show that the bifunctional activity of these CREs depends on their CRX binding sites. When these binding sites are abolished, silencing activity is lost in wild-type retina, while enhancer activity is lost in retinas missing the CRX effector domain. This suggests that these binding sites are re-used by another TF in a cellular context lacking functional CRX, and that this other TF activates at sites where CRX represses (White et al. 2016). A probable candidate is the CRX ortholog OTX2, which plays a major role in retinal development and whose binding specificity is nearly identical to that of CRX (Furukawa et al. 2002; Koike et al. 2007). While we observed the conversion of potential bifunctional CREs only in mutant genotypes, in wild-type retina these CREs may act as CRX-bound silencers in rod photoreceptors and as enhancers in cell types with high OTX2 and low CRX expression, such as bipolar cells. CRX and OTX2, despite close homology in both their DNA binding and effector domains, have non-overlapping functions (Yamamoto et al. 2020; Terrell et al. 2012), which would allow them to play different regulatory roles at the same CREs. Currently there are a limited number of examples of CREs that act as silencers in one cell type and enhancers in another (Gisselbrecht et al. 2020). However, such bifunctional elements could be part of an effective regulatory strategy during the development of a population of multipotent progenitor cells into closely related cell types with overlapping regulators, as occurs in the retina.

CREs bound by CRX differ in their chromatin environments, notably the cell-type specificity and developmental timing of chromatin accessibility (Ruzycki et al. 2018). We found that the

activities of these CREs measured episomally by MPRA at P8 showed some correspondence with the chromatin environment at their native genomic loci. Distal CREs whose chromatin accessibility does not peak until 8 weeks after birth were depleted of strong enhancers and enriched for silencers, while proximal CREs that are already accessible by P0 contained the highest fraction of strong enhancers of any chromatin group. The enrichment of silencers among late-acting CRX-bound CREs suggests that these elements may have sequence-encoded, early acting repressive features that prevent premature activation. Our results indicate that episomal MPRAs capture some sequence-encoded features that determine differences in chromatin environment in the genome.

Many inherited human diseases are linked with variants in TFs (Tremblay et al. 2018; Herman et al. 2021; Sun and Chen 2023). To understand how different variants in the same TF gene produce different disease phenotypes, it is critical to understand the distinct effects of these variants on *cis*-regulatory function. Our work suggests that using MPRAs to directly measure changes in *cis*-regulatory function can yield insights into variant effects that complement the information learned from transcriptional and chromatin profiling assays.

METHODS

Library Design and Cloning

The MPRA library contained 1,723 genomic sequences centered on CRX ChIP-seq peaks that overlapped ATAC-seq peaks from P14 retina (Ruzycki et al. 2018; Corbo et al. 2010). Sequences were 164 bp in length and represented in the library by four unique barcodes. As negative controls we included 100 scrambled sequences. Basal promoters were represented with 16 unique barcodes to ensure precise measurement of basal levels.

ChIP-seq peaks were sampled based on two sets of annotations described by Ruzycki et al. (Ruzycki et al. 2018) (Table 1). These annotations include 1) the presence or absence of histone

marks H3K4me3 and H3K27ac in P14 wild-type retina, and 2) the loss or retention of chromatin accessibility in *Crx*$^{-/-}$ retina (*i.e.*, whether accessibility is CRX dependent or independent). Generally, H3K4me3^{+}/H3K27ac^{+} sites include proximal regulatory elements typically within 5 kb of the nearest TSS (Class A), H3K4me3^{-}/H3K27ac^{+} sites include distal regulatory elements typically 5-500 kb from the nearest TSS (Class C), and H3K4me3^{-}/H3K27ac^{-} sites are also distal sequences typically 5-500 kb from the nearest TSS (Class D) that are accessible and bound by CRX but lack other evidence of activity. For each of the wild-type ChIP peaks in the library, we created a mutant version in which all occurrences of the CRX motif were abolished by a point mutation, following the procedure described in (White et al. 2013). To test the hypothesis that CRX mutants might alter cooperative activity at CREs, we selected 100 library sequences with only two copies of the CRX motif, and created single and double mutants.

The MPRA library was cloned following our previously described protocol (White et al. 2013; White et al. 2016; Friedman et al. 2021). Library sequences were synthesized by Agilent as barcoded oligonucleotides with the following structure: 5' priming sequence (GTAGCGTCTGTCCGT)/EcoRI site/164 bp Library sequence/SpeI site/C/SphI site/Barcode sequence/NotI site/3' priming sequence (CAACTACTACTACAG). The oligo library was amplified by four cycles of PCR using New England Biolabs (NEB) Phusion High-fidelity Master Mix. Amplicons were purified on a 2% agarose gel, digested with EcoRI-HF and NotI-HF (NEB), then cloned into the EagI and EcoRI sites of the plasmid pJK03 with T4 ligase (NEB). The libraries were transformed into 5-alpha electrocompetent cells (NEB). The plasmid library was prepared by maxi-prep (Sigma), digested with SphI-HF and SpeI-HF (NEB) and then treated with Antarctic phosphatase (NEB). Basal promoters with *DsRed* were amplified from pJK01 (*RHO*) (Kwasnieski et al. 2012) or pGL-HSP68 (*HSP68*) (Kwasnieski et al. 2014) and purified from a 1% agarose gel. Basal promoter amplicons were digested with NheI-HF and SphI-HF, then cloned into the libraries

with T4 ligase (NEB). The libraries were transformed into 5-alpha electrocompetent cells (NEB), grown in liquid culture, then prepared by maxi-prep (Sigma).

Retinal Explant Electroporation and Library Sequencing

All mouse lines were on the C57BL/6J background and free of rd1 and rd8 mutations. The $Crx^{R90W/W}$, $Crx^{R90W/+}$, $Crx^{E168d2/d2}$and $Crx^{E168d2/+}$ lines and their genotyping procedures were described by Tran et al. (Tran et al. 2014). All procedures involving mice were approved by the Animal Studies Committee of Washington University in St. Louis and performed under Animal Welfare Assurance # A-3381-01 and Protocols # 21-0414 (to SC). 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 Sciences. Every effort was made to minimize the animals' suffering, anxiety and discomfort.

Retinal explant electroporations were conducted as previously described (Kwasnieski et al. 2012). Retinas were isolated from P0 mice and electroporated with 30 µg MPRA library and 30 µg pCAG-GFP as a transfection control. After culturing for 8 days retinas were placed in TRIzol (ThermoFisher) and RNA was extracted. Three replicate electroporations (3 retinas each) were performed.

cDNA was prepared using Superscript Reverse Transcriptase III (ThermoFisher) with oligo dT primers following the manufacturer's instructions. Barcodes were amplified from cDNA or input plasmid DNA using 24 cycles of PCR with Phusion High Fidelity Master Mix (NEB). Amplicons were purified with the Monarch PCR kit (NEB), digested with MfeI-HF and SphI-HF (NEB) and ligated to custom sequencing adapters with PE2 indexing barcodes and phased P1 barcodes. Following 20 cycles of enrichment PCR, libraries were sequenced on the Illumina NextSeq platform to greater than 1000x coverage per sample.

MPRA Quantification and Analysis

All sequencing reads were processed regardless of quality score. Sequenced barcodes were matched to expected library members, and only reads with a sequenced barcode that perfectly matched an expected library barcode were counted. As described previously (Kwasnieski et al. 2012; White et al. 2013), the transcriptional activity of each barcoded reporter gene was determined by the number of RNA reads of the corresponding barcode. To account for differences in barcode representation in the pooled library, RNA reads were normalized to input DNA reads. RNA/DNA ratios were averaged over all barcodes for each element, and across biological replicates. CREs were classified into one of five activity groups based on their activity relative to the basal promoter alone. CREs with a transcriptional activity that was not significantly different from basal were deemed “inactive”. Remaining CREs with an activity significantly greater than basal were deemed “strong enhancers” if their activity was greater than the 95th percentile of the 100 scrambled control sequences, and “weak enhancers” otherwise. “Strong silencers” and “weak silencers” were classified by having activity significantly lower than the basal promoter, with strong silencers less than the 5th percentile of the 100 scrambled controls.

Each CRE was annotated with the characteristics of the genomic sequence of origin using the ChIP-seq data from which the CREs were derived (Ruzycki et al. 2018). The information content and predicted CRX occupancy of each CRE sequence was computed as previously described (Friedman et al. 2021). To analyze motif enrichment across CRE subgroups, SEA (“Simple Enrichment Analysis”, MEME Suite v5.5.2) was used with default settings and the HOCOMOCO Mouse v11 CORE motif database (Bailey and Grant 2021; Kulakovskiy et al. 2018). Analysis scripts and Python notebooks for the described analyses are available at https://github.com/barakcohenlab/retinopathy-manuscript, including the source code necessary to generate all data figure panels. The completed MPRA analysis with activity and classifications of each CRE can be found in Supplemental Table S1.

DATA ACCESS

All raw and processed sequencing data generated in this study have been deposited at the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE230090.

COMPETING INTEREST STATEMENT

B.A.C is on the scientific advisory board of Patch Biosciences. The authors declare no other competing interests.

ACKNOWLEDGEMENTS

This work was supported by NIH grants R01 EY027784 (to BC and SC), R01 EY012543 (to SC), R01 GM121755 (to MAW), F30 EY033640 (to JLS), F31 HG011431 (to RZF), and Research to Prevent Blindness (to SC and DOVS).

REFERENCES

Bailey TL, Grant CE. 2021. *SEA: Simple Enrichment Analysis of motifs*. bioRxiv doi: 10.1101/2021.08.23.457422.

Chau K-Y, Chen S, Zack DJ, Ono SJ. 2000. *Functional Domains of the Cone-Rod Homeobox (CRX) Transcription Factor*. *Journal of Biological Chemistry* **275**: 37264–37270.

Chen J, Rattner A, Nathans J. 2005. *The rod photoreceptor-specific nuclear receptor Nr2e3 represses transcription of multiple cone-specific genes*. *J Neurosci* **25**: 118–129.

Chen S, Wang Q-L, Nie Z, Sun H, Lennon G, Copeland NG, Gilbert DJ, Jenkins NA, Zack DJ. 1997. *Crx, a Novel Otx-like Paired-Homeodomain Protein, Binds to and Transactivates Photoreceptor Cell-Specific Genes*. *Neuron* **19**: 1017–1030.

Corbo JC, Lawrence KA, Karlstetter M, Myers CA, Abdelaziz M, Dirkes W, Weigelt K, Seifert M, Benes V, Fritsche LG, et al. 2010. *CRX ChIP-seq reveals the cis-regulatory architecture of mouse photoreceptors*. *Genome Research* **20**: 1512–1525.

DelRosso N, Tycko J, Suzuki P, Andrews C, Aradhana null, Mukund A, Liongson I, Ludwig C, Spees K, Fordyce P, et al. 2023. *Large-scale mapping and mutagenesis of human transcriptional effector domains*. *Nature* **616**: 365–372.

Freund CL, Gregory-Evans CY, Furukawa T, Papaioannou M, Looser J, Ploder L, Bellingham J, Ng D, Herbrick J-AS, Duncan A, 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**: 543–553.

Friedman RZ, Granas DM, Myers CA, Corbo JC, Cohen BA, White MA. 2021. *Information content differentiates enhancers from silencers in mouse photoreceptors*. *eLife* **10**: e67403.

Fujinami-Yokokawa Y, Fujinami K, Kuniyoshi K, Hayashi T, Ueno S, Mizota A, Shinoda K, Arno G, Pontikos N, Yang L, et al. 2020. *Clinical and Genetic Characteristics of 18 Patients from 13 Japanese Families with CRX-associated retinal disorder: Identification of Genotype-phenotype Association*. *Sci Rep* **10**: 9531.

Furukawa A, Koike C, Lippincott P, Cepko CL, Furukawa T. 2002. *The mouse Crx 5'-upstream transgene sequence directs cell-specific and developmentally regulated expression in retinal photoreceptor cells*. *J Neurosci* **22**: 1640–1647.

Furukawa T, Morrow EM, Cepko CL. 1997. *Crx, a Novel otx-like Homeobox Gene, Shows Photoreceptor-Specific Expression and Regulates Photoreceptor Differentiation*. *Cell* **91**: 531–541.

Gisselbrecht SS, Palagi A, Kurland JV, Rogers JM, Ozadam H, Zhan Y, Dekker J, Bulyk ML. 2020. *Transcriptional Silencers in Drosophila Serve a Dual Role as Transcriptional Enhancers in Alternate Cellular Contexts*. *Mol Cell* **77**: 324–337.e8.

Grant CE, Bailey TL, Noble WS. 2011. *FIMO: Scanning for occurrences of a given motif*. *Bioinformatics* **27**: 1017–1018.

Hennig AK, Peng G-H, Chen S. 2008. *Regulation of photoreceptor gene expression by Crx-associated transcription factor network*. *Brain Research* **1192**: 114–133.

Hennig AK, 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.

Herman L, Todeschini A-L, Veitia RA. 2021. *Forkhead Transcription Factors in Health and Disease*. *Trends Genet* **37**: 460–475.

Hlawatsch J, Karlstetter M, Aslanidis A, Lückoff A, Walczak Y, Plank M, Böck J, Langmann T. 2013. *Sterile alpha motif containing 7 (Samd7) is a novel crx-regulated transcriptional repressor in the retina*. *PLoS One* **8**: e60633.

Hughes AEO, Myers CA, Corbo JC. 2018. *A massively parallel reporter assay reveals context-dependent activity of homeodomain binding sites in vivo*. *Genome Research* **28**: 1520–1531.

Jacobson SG, Cideciyan AV, Huang Y, Hanna DB, Freund CL, Affatigato LM, Carr RE, Zack DJ, Stone EM, McInnes RR. 1998. Retinal degenerations with truncation mutations in the cone-rod homeobox (CRX) gene. *Invest Ophthalmol Vis Sci* **39**: 2417–2426.

Karousis ED, Mühlemann O. 2019. *Nonsense-Mediated mRNA Decay Begins Where Translation Ends*. *Cold Spring Harb Perspect Biol* **11**: a032862.

Koike C, Nishida A, Ueno S, Saito H, Sanuki R, Sato S, Furukawa A, Aizawa S, Matsuo I, Suzuki N, et al. 2007. *Functional Roles of Otx2 Transcription Factor in Postnatal Mouse Retinal Development*. *Molecular and Cellular Biology* **27**: 8318–8329.

Koyanagi Y, Akiyama M, Nishiguchi KM, Momozawa Y, Kamatani Y, Takata S, Inai C, Iwasaki Y, Kumano M, Murakami Y, et al. 2019. *Genetic characteristics of retinitis pigmentosa in 1204 Japanese patients*. *J Med Genet* **56**: 662–670.

Kulakovskiy IV, Vorontsov IE, Yevshin IS, Sharipov RN, Fedorova AD, Rumynskiy EI, Medvedeva YA, Magana-Mora A, Bajic VB, Papatsenko DA, et al. 2018. *HOCOMOCO: Towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis*. *Nucleic Acids Research* **46**: D252–D259.

Kwasnieski JC, Fiore C, Chaudhari HG, Cohen BA. 2014. *High-throughput functional testing of ENCODE segmentation predictions*. *Genome Res* **24**: 1595–1602.

Kwasnieski JC, Mogno I, Myers CA, Corbo JC, Cohen BA. 2012. *Complex effects of nucleotide variants in a mammalian cis-regulatory element*. *Proceedings of the National Academy of Sciences* **109**: 19498–19503.

Langouët M, Jolicoeur C, Javed A, Mattar P, Gearhart MD, Daiger SP, Bertelsen M, Tranebjærg L, Rendtorff ND, Grønskov K, et al. 2022. *Mutations in BCOR , a co-repressor of CRX/OTX2 , are associated with early-onset retinal degeneration*. *Sci Adv* **8**: eabh2868.

Lee J, Myers CA, Williams N, Abdelaziz M, Corbo JC. 2010. *Quantitative fine-tuning of photoreceptor cis-regulatory elements through affinity modulation of transcription factor binding sites*. *Gene Ther* **17**: 1390–1399.

Loell KJ, Friedman RZ, Myers CA, Corbo JC, Cohen BA, White MA. 2023. *Transcription factor interactions explain the context-dependent activity of CRX binding sites*. bioRxiv doi: 10.1101/2023.03.05.531194.

Mitton KP, Swain PK, Khanna H, Dowd M, Apel IJ, Swaroop A. 2003. *Interaction of retinal bZIP transcription factor NRL with Flt3-interacting zinc-finger protein Fiz1: Possible role of Fiz1 as a transcriptional repressor*. *Hum Mol Genet* **12**: 365–373.

Nichols LL, Alur RP, Boobalan E, Sergeev YV, Caruso RC, Stone EM, Swaroop A, Johnson MA, Brooks BP. 2010. *Two novel CRX mutant proteins causing autosomal dominant Leber congenital amaurosis interact differently with NRL*. *Hum Mutat* **31**.

Oh ECT, Cheng H, Hao H, Jia L, Khan NW, Swaroop A. 2008. *Rod differentiation factor NRL activates the expression of nuclear receptor NR2E3 to suppress the development of cone photoreceptors*. *Brain Res* **1236**: 16–29.

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**: 747–764.

Peng G-H, Chen S. 2007. *Crx activates opsin transcription by recruiting HAT-containing co-activators and promoting histone acetylation*. *Human Molecular Genetics* **16**: 2433–2452.

Perrault I, Hanein S, Gerber S, Barbet F, Dufier J-L, Munnich A, Rozet J-M, Kaplan J. 2003. *Evidence of autosomal dominant Leber congenital amaurosis (LCA) underlain by a CRX heterozygous null allele*. *J Med Genet* **40**: e90.

Rister J, Razzaq A, Boodram P, Desai N, Tsanis C, Chen H, Jukam D, Desplan C. 2015. *Single–base pair differences in a shared motif determine differential Rhodopsin expression*. *Science* **350**: 1258–1261.

Rivolta C, Berson EL, Dryja TP. 2001. *Dominant Leber congenital amaurosis, cone-rod degeneration, and retinitis pigmentosa caused by mutant versions of the transcription factor CRX*. *Hum Mutat* **18**: 488–498.

Ruzycki PA, Tran NM, Kolesnikov AV, Kefalov VJ, Chen S. 2015. *Graded gene expression changes determine phenotype severity in mouse models of CRX-associated retinopathies*. *Genome Biology* **16**.

Ruzycki PA, Zhang X, Chen S. 2018. *CRX directs photoreceptor differentiation by accelerating chromatin remodeling at specific target sites*. *Epigenetics & Chromatin* **11**.

Sohocki MM, Sullivan LS, Mintz-Hittner HA, Birch D, Heckenlively JR, Freund CL, McInnes RR, Daiger SP. 1998. *A Range of Clinical Phenotypes Associated with Mutations in CRX, a Photoreceptor Transcription-Factor Gene*. *The American Journal of Human Genetics* **63**: 1307–1315.

Sun C, Chen S. 2023. *Disease-causing mutations in genes encoding transcription factors critical for photoreceptor development*. *Front Mol Neurosci* **16**: 1134839.

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

Swaroop A, Wang Q-L, Wu W, Cook J, Coats C, Xu S, Chen S, Zack DJ, Sieving PA. 1999. *Leber Congenital Amaurosis Caused by a Homozygous Mutation (R90W) in the Homeodomain of the Retinal Transcription Factor CRX: Direct Evidence for the Involvement of CRX in the Development of Photoreceptor Function*. *Human Molecular Genetics* **8**: 299–305.

Terrell D, Xie B, Workman M, Mahato S, Zelhof A, Gebelein B, Cook T. 2012. *OTX2 and CRX rescue overlapping and photoreceptor-specific functions in the Drosophila eye*. *Dev Dyn* **241**: 215–228.

Tran NM, Chen S. 2014. *Mechanisms of blindness: Animal models provide insight into distinct CRX-associated retinopathies*. *Developmental Dynamics* **243**: 1153–1166.

Tran NM, Zhang A, Zhang X, Huecker JB, Hennig AK, Chen S. 2014. *Mechanistically Distinct Mouse Models for CRX-Associated Retinopathy* ed. G.S. Barsh. *PLoS Genetics* **10**: e1004111.

Tremblay M, Sanchez-Ferras O, Bouchard M. 2018. *GATA transcription factors in development and disease*. *Development* **145**: dev164384.

White MA, Kwasnieski JC, Myers CA, Shen SQ, Corbo JC, Cohen BA. 2016. *A Simple Grammar Defines Activating and Repressing cis-Regulatory Elements in Photoreceptors*. *Cell Reports* **17**: 1247–1254.

White MA, Myers CA, Corbo JC, Cohen BA. 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* **110**: 11952–11957.

Wilson D, Sheng G, Lecuit T, Dostatni N, Desplan C. 1993. *Cooperative dimerization of paired class homeo domains on DNA. Genes Dev* **7**: 2120–2134.

Yamamoto H, Kon T, Omori Y, Furukawa T. 2020. *Functional and Evolutionary Diversification of Otx2 and Crx in Vertebrate Retinal Photoreceptor and Bipolar Cell Development*. *Cell Reports* **30**: 658–671.e5.

Yanagi Y, Masuhiro Y, Mori M, Yanagisawa J, Kato S. 2000. *P300/CBP acts as a coactivator of the cone-rod homeobox transcription factor*. *Biochem Biophys Res Commun* **269**: 410–414.

Zagozewski JL, Zhang Q, Pinto VI, Wigle JT, Eisenstat DD. 2014. *The role of homeobox genes in retinal development and disease*. *Developmental Biology* **393**: 195–208.

Zheng Y, Sun C, Zhang X, Ruzycki P, Chen S. 2023. *Missense mutations in CRX homeodomain cause dominant retinopathies through two distinct mechanisms*. bioRxiv doi: 10.1101/2023.02.01.526652.

SUPPLEMENTARY MATERIAL

Supplementary Table 1 Table of transcriptional activities and classifications for all CREs in all genotypes in Apache Parquet format.

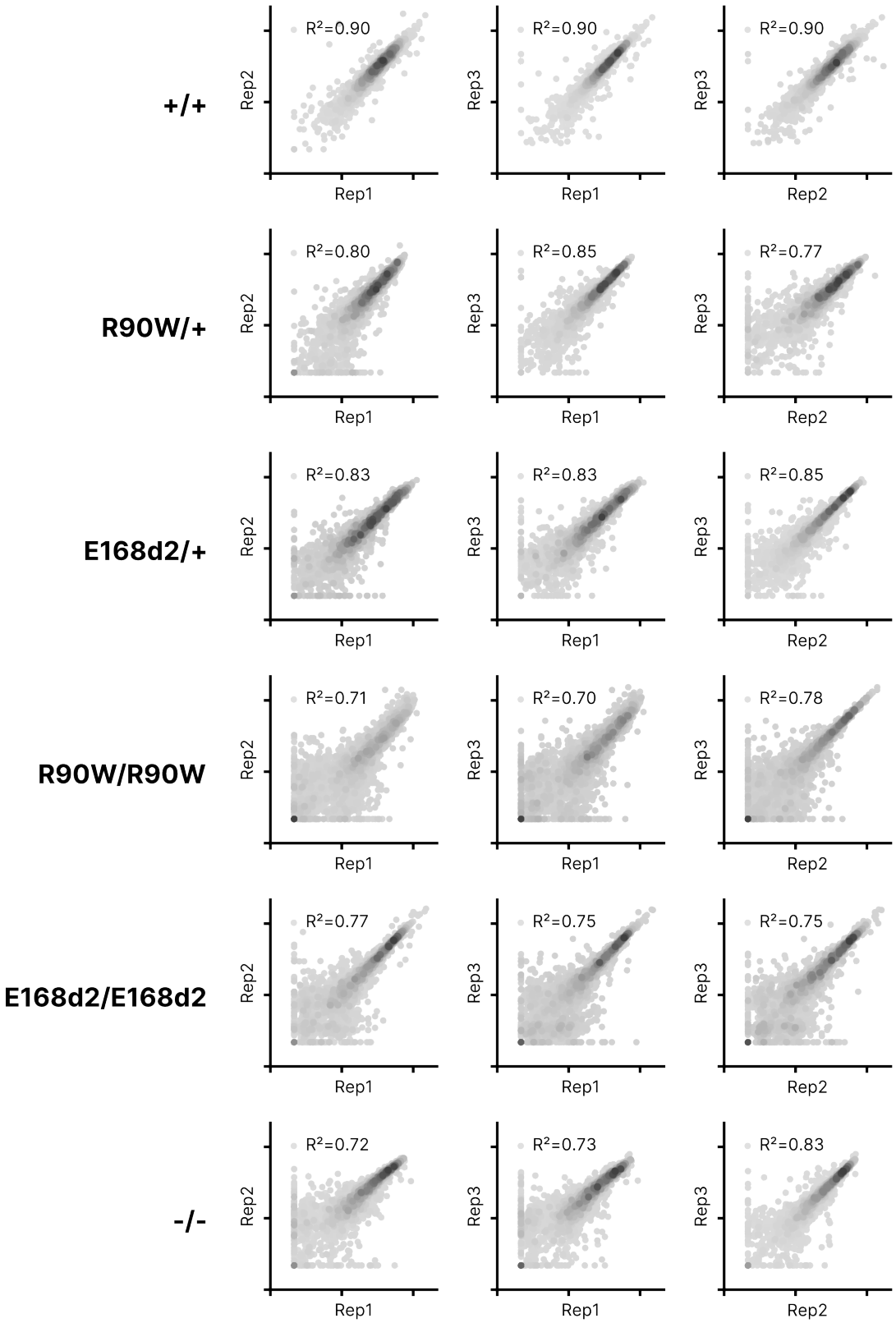


Figure S1: CRE-level correlation across replicates for the Rho promoter library for the indicated genotypes.

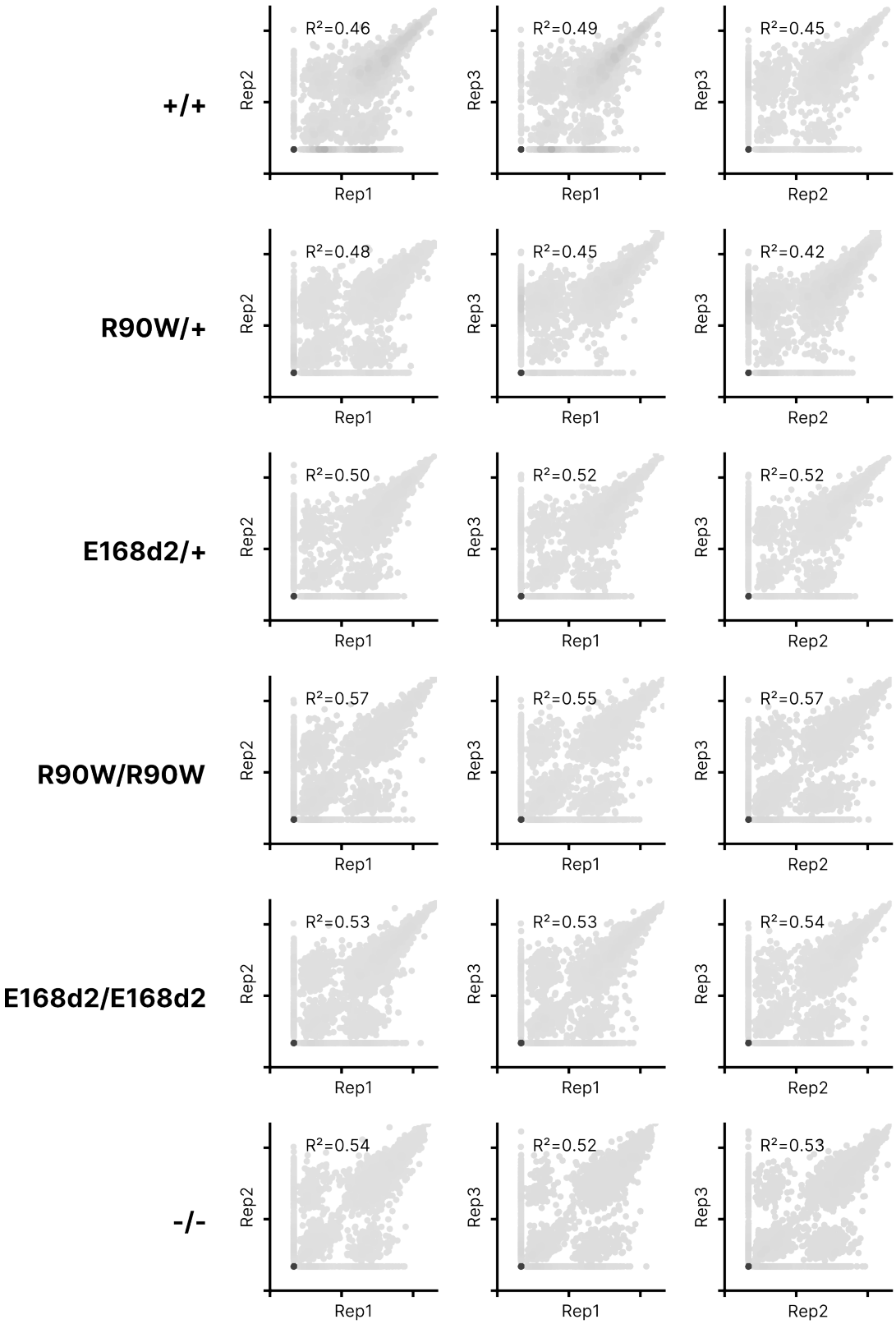


Figure S2: CRE-level correlation across replicates for the hsp68 promoter library for the indicated genotypes.

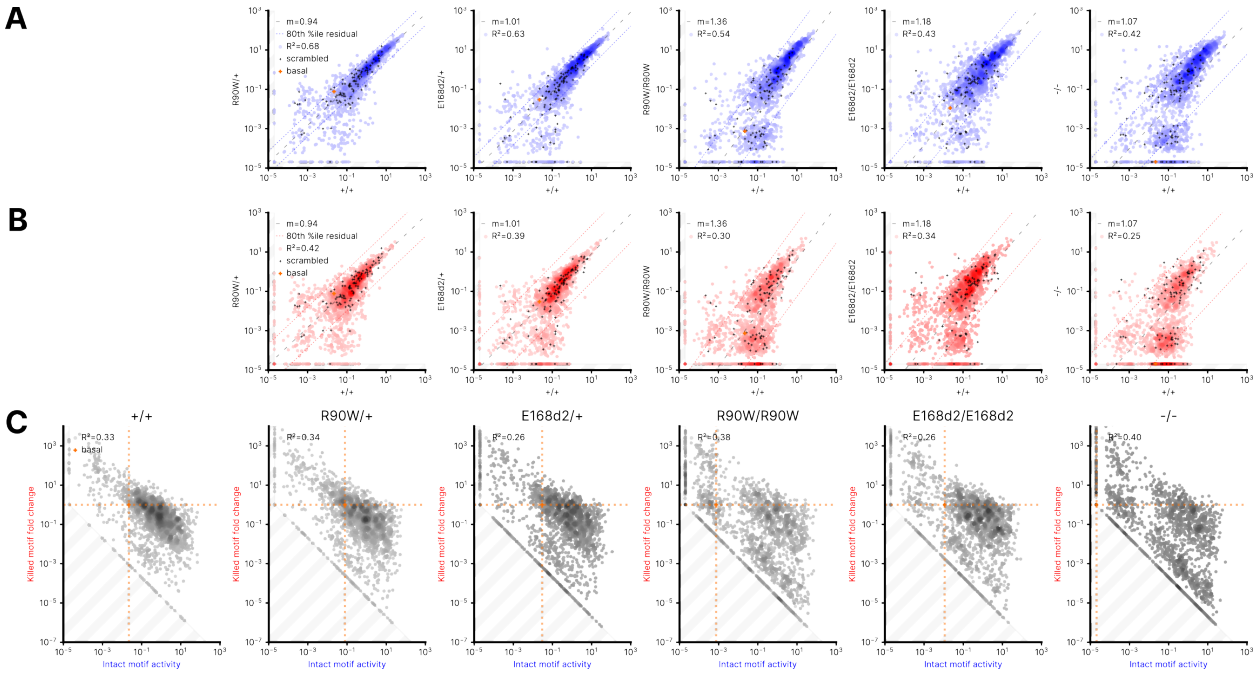


Figure S3: (**A,B**) Correlations of the transcriptional activity of each genomic CRE sequence with intact CRX motifs (**A**) or CRE sequence with killed CRX motifs (**B**) in the indicated genotypes, for the hsp68 promoter library. Superimposed black cross symbols indicate the transcriptional activity of the 100 scrambled control sequences, and the orange cross shows the activity of the basal rhodopsin promoter alone. Using the scrambled control sequences to establish an empirical null distribution of CRE activity, we performed robust regression using a trimmed mean M-estimator (fit plotted as a dashed line) for each genotype comparison. The outer dotted lines indicate the 80th percentile of the residuals of the library CREs relative to the regression of the scrambled sequences. (**C**) Activity of each CRE with intact CRX motifs versus the fold change in activity for each matching CRE with killed motifs, in each of the indicated genotypes. The orange dotted lines indicate the activity of the basal promoter alone, and a fold change of 1 (no difference in activity between CRE with intact or killed CRX motifs). In all panels, transcriptional activity was adjusted for visualization purposes using a fixed pseudocount of 2e-5; the out-of-bound region is indicated by the shading in light grey.

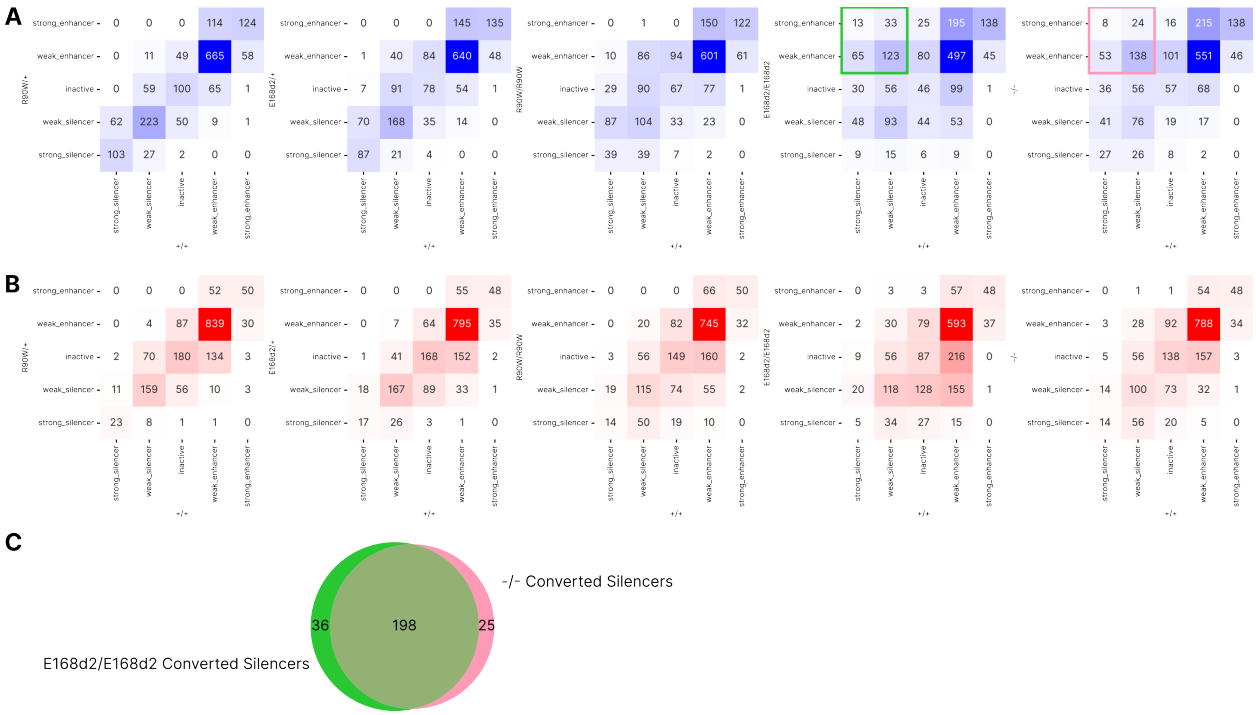


Figure S4: (**A**) Classifications of all CREs with intact CRX motifs. (**B**) Classifications of all CREs with killed CRX motifs. (**C**) Overlap of CREs classified as silencers in +/+ and enhancers in the indicated genotypes; colors match the boxes in panel A.

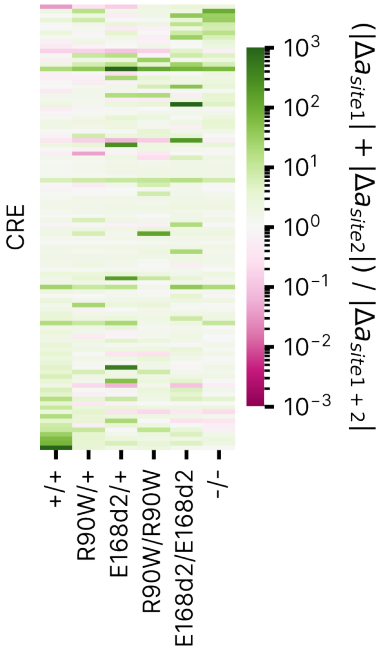


Figure S5: Heatmap of the ratio of the sum of single mutant transcriptional activity delta to the double mutant transcriptional activity delta; each row is a different CRE. CREs are sorted by the ratio of the cooperativity score in +/+ and -/- genotypes.