

Identification of a modular super-enhancer in murine retinal development

Victoria Honnell ^{1,2}, Jackie L. Norrie ¹, Anand G. Patel ¹, Cody Ramirez ¹, Jiakun Zhang ¹, Yu-Hsuan Lai¹, Shibiao Wan ³ & Michael A. Dyer ¹ 

Super-enhancers are expansive regions of genomic DNA comprised of multiple putative enhancers that contribute to the dynamic gene expression patterns during development. This is particularly important in neurogenesis because many essential transcription factors have complex developmental stage- and cell-type specific expression patterns across the central nervous system. In the developing retina, *Vsx2* is expressed in retinal progenitor cells and is maintained in differentiated bipolar neurons and Müller glia. A single super-enhancer controls this complex and dynamic pattern of expression. Here we show that deletion of one region disrupts retinal progenitor cell proliferation but does not affect cell fate specification. The deletion of another region has no effect on retinal progenitor cell proliferation but instead leads to a complete loss of bipolar neurons. This prototypical super-enhancer may serve as a model for dissecting the complex gene expression patterns for neurogenic transcription factors during development. Moreover, it provides a unique opportunity to alter expression of individual transcription factors in particular cell types at specific stages of development. This provides a deeper understanding of function that cannot be achieved with traditional knockout mouse approaches.

¹Department of Developmental Neurobiology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA. ²Graduate School of Biomedical Sciences, St. Jude Children's Research Hospital, Memphis, TN 38105, USA. ³Center for Applied Bioinformatics, St. Jude Children's Research Hospital, Memphis, TN 38105, USA. ✉email: michael.dyer@stjude.org

The developing retina must express thousands of genes in a precise spatiotemporal order for proper cell-fate determination¹. While some genes, such as those regulated by Rb1, have simple transcriptional regulation at the transcription start site, others rely on enhancers that may be hundreds of kilobases away^{2,3}. Looping out of the intervening chromatin and higher-order chromatin structures are believed to be important in developmental stage-specific promoter-enhancer interactions^{4,5}. Indeed, in the case of *Sox2* the enhancer is not only physically separated from the promoter when the gene is inactive but it is sequestered in a different nuclear compartment^{6,7}. Advances in high-throughput sequencing assays such as chromatin immunoprecipitation sequencing (ChIP-seq) has allowed for tens of thousands of putative enhancers to be identified in the developing murine and human retina^{8–10}. However, few of these enhancers have been functionally analyzed in vivo¹¹. One challenge is determining which enhancers are essential in retinogenesis and which genes they regulate. Another challenge is determining how genes with complex expression patterns in multiple cell types at different stages of development are regulated.

Super-enhancers (SEs) are large regulatory elements containing clusters of smaller transcriptional enhancers believed to define cell identity and cell fate specification across development^{11–15}. Originally, SEs were identified in ESCs and were differentiated from regular enhancers through enrichment of Med1, OSN, H3K27ac, and H3K4me1, and DNaseI hypersensitivity regions¹⁴. Efforts to dissect SEs by genetic manipulation suggest that SEs are comprised of both necessary and dispensable elements^{16–21}. Whether constituent enhancers comprising SEs are responsible for gene expression, or whether SEs act as a single functional regulatory unit is still debated in the field.

Many central nervous system-specific transcription factors have complex expression patterns since they are expressed in multiple cell types across multiple stages of development^{22–24}. This makes uncoupling early-stage effects from late-stage effects difficult in gene knockout mice. For example, *Pax6* expression is required for both brain and eye formation early in development, is downregulated as neurons differentiate, and is maintained in a subset of differentiated cell types in the adult^{25,26}. Since a *Pax6* knockout animal model is embryonic lethal, studying its late-stage effects in postmitotic neurons is difficult even with conditional knockout approaches. The same is true for many of the transcription factors that are essential for neurogenesis.

The *Vsx2* gene encodes a homeodomain containing transcription factor that is expressed in several cell types across multiple stages of eye development^{27,28}. *Vsx2* is found in the majority of proliferating retinal progenitor cells (RPCs) during development as well as differentiated bipolar cells, and Müller glia in the adult retina^{27–30}. Mutations in human and mouse *VSX2*/*Vsx2* lead to microphthalmia because it is required for retinal progenitor cell proliferation^{31–34}. Bipolar cells have not been detected in the *Vsx2* mutant retina, but it is not clear if this is due to a direct role of *Vsx2* in bipolar cell fate specification or a secondary consequence of the defects in RPC proliferation^{31,33}.

We previously identified a super-enhancer upstream of the *Vsx2* gene³⁵ and computational methods³⁶ suggested it was a core regulatory circuit SE (CRC-SE). That is, the *Vsx2* SE has a *Vsx2* binding site and may regulate its own expression. Deletion of a portion of the *Vsx2* SE including the *Vsx2* binding site led to a complete absence of bipolar neurons³⁵. Importantly, the region we deleted contained a much smaller element that had been previously shown to confer bipolar-specific expression by square wave electroporation in vivo³⁷. In a more recent study, that bipolar-specific element was deleted using CRISPR/Cas9 by in vivo square wave electroporation leading to a reduction in bipolar neurons³⁸. *Vsx2* expression was maintained in Müller glia

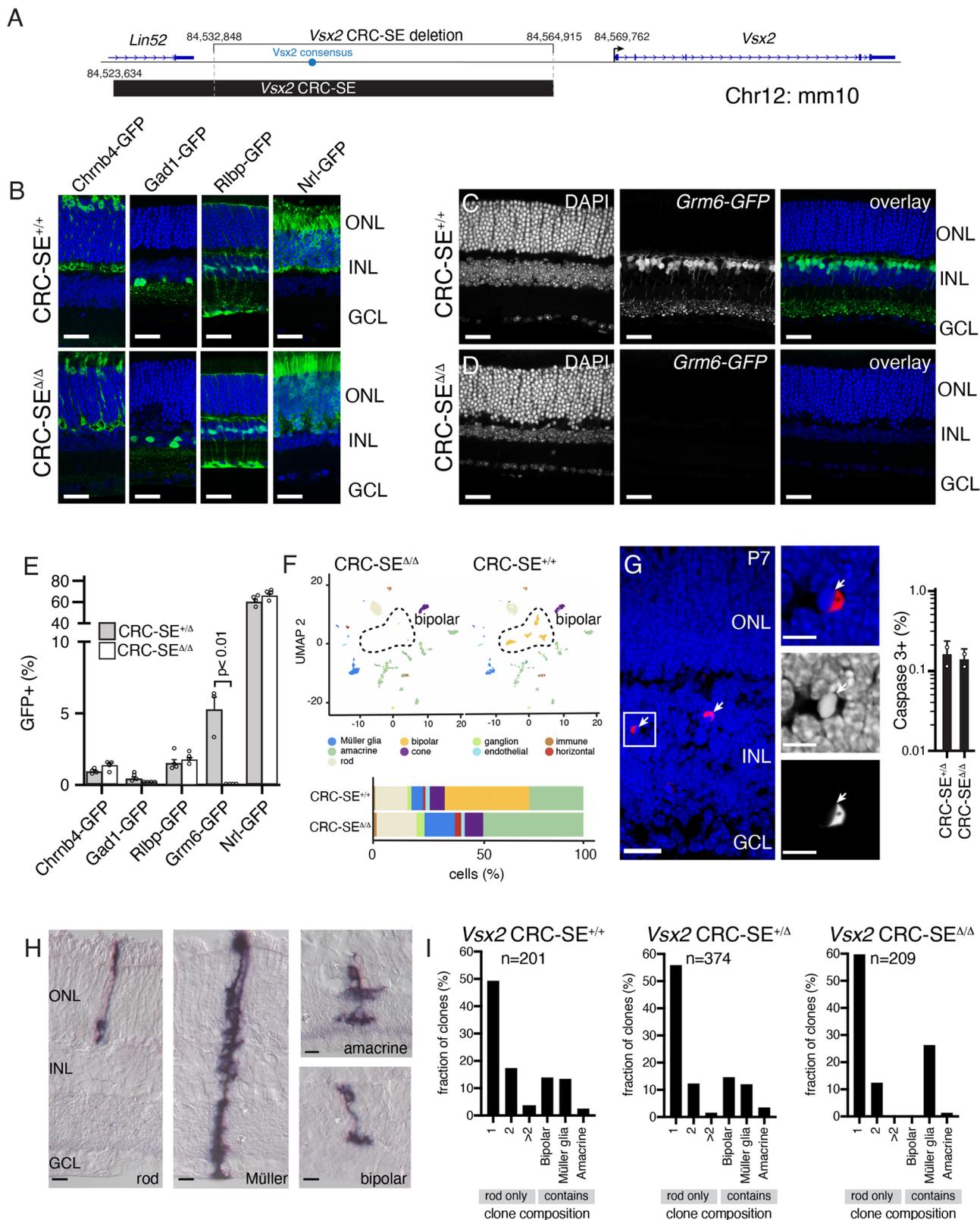
of *Vsx2*-SE mice and there was no evidence of microphthalmia suggesting that the retinal progenitor expression of *Vsx2* was not altered.

In the present study, we further refine the domains of the *Vsx2* CRC-SE and determine if *Vsx2* binding is required for its function by generating eight new mouse strains lacking different regions of the *Vsx2* SE. We show that the previously identified evolutionarily conserved *Vsx2* binding sites are not required for function and we identify distinct subdomains or modules of the larger *Vsx2* SE that are required for retinal progenitor cell expression and bipolar neuron expression. Single-cell RNA-seq combined with single-cell ATAC-seq was particularly useful for identifying these distinct domains, but in vivo validation was essential for establishing their function. It is possible that SEs and/or CRC-SEs associated with neurogenic transcription factors that have complex spatial and temporal patterns of expression may be modular with respect to the patterns of expression. This is in contrast to SEs and/or CRC-SEs where the individual elements are simply additive, but do not confer temporal or spatial specificity¹⁵.

Results

Bipolar cells switch fate in the *Vsx2*-SE^{Δ/Δ} retina. We have previously identified a 41 kb CRC-SE upstream of the *Vsx2* gene (Fig. 1A)¹⁰. Deletion of 32 kb of that SE (Fig. 1A) in vivo using CRISPR/Cas9 led to a complete loss of bipolar neurons based on immunostaining and RNA-seq³⁵. To further elucidate the cell-type specific effects of the *Vsx2*-SE deletion, we crossed the *Vsx2*-SE^{Δ/Δ} mice to those with cell-type specific GFP reporter transgenes (Fig. 1B–D). The pattern of GFP expression in cones (*Chrn4-GFP*), horizontal cells, and a subset of amacrine cells (*Gad1-GFP*), Müller glia (*Rbp-GFP*), and rods (*Nrl-GFP*) was normal other than the laminar disruption that occurs due to the loss of bipolar cells (Fig. 1B). However, ON bipolar cells were completely absent in the *Vsx2*-SE^{Δ/Δ}; *Grm6-GFP* mice^{39–42} (Fig. 1C, D, F). Quantitative scoring of the proportion of GFP+ cells in each line by flow cytometry showed that there were no statistically significant changes in retinal cell types other than bipolar neurons (Fig. 1E). However, subtle differences may be obscured by the large proportion (~80%) of rods in the murine retina.

The reduction in bipolar neurons in the *Vsx2*-SE^{Δ/Δ} retinae raises the possibility that the cells fated to be bipolar neurons die during development or become other cell types born at the same stage of development (rods or Müller glia). To determine if bipolar neurons died during development, we performed immunostaining for activated Caspase 3 during the peak period of bipolar genesis (P3, P7) in wild type and *Vsx2*-SE^{Δ/Δ} retinae and scored the proportion of immunopositive cells. There was no observable difference in the proportion of apoptotic cells between wild type and *Vsx2*-SE^{Δ/Δ} retinae (Fig. 1G). Next, to determine if there is an increase in rods and/or Müller glia, we performed retroviral lineage analysis using an alkaline phosphatase replication-incompetent retrovirus (LIA-E, Supplemental Information). The retrovirus was injected into the subretinal space in P0 pups from a *Vsx2*-SE^{+Δ} intracross. After 3 weeks, the retinae were harvested, stained for alkaline phosphatase, and serially cryosectioned. Clones of cells from individual retinal progenitor cells were reconstructed from serial sections and scored for cell number and cell type by morphology and laminar position (Fig. 1H). After scoring, the mice were genotyped to determine if they were *Vsx2*-SE^{Δ/Δ}, *Vsx2*-SE^{+Δ}, or *Vsx2*-SE^{+/+}. We found significant differences in cell type distribution according to genotype ($p = 0.0053$). Notably, bipolar cell containing clones were absent in the *Vsx2*-SE^{Δ/Δ} retinae while they were present in the *Vsx2*-SE^{+Δ} and *Vsx2*-SE^{+/+} retina. The absence of bipolar



containing clones was largely filled by an increased relative prevalence of Müller glia or rod containing clones in the *Vsx2-SE Δ/Δ* relative to the wild type and heterozygous littermates (Fig. 1I and Supplemental Information). This is consistent with the birthdates of late born cell types (rods, bipolar neurons, and Müller glia). Taken together, our data suggest that the bipolar neurons become rods and Müller glia in the *Vsx2-SE Δ/Δ* retina.

Retinal progenitor cell proliferation is normal in the *Vsx2-SE Δ/Δ* retina. To determine if there is any perturbation in retinal progenitor cell proliferation as a result of the super-enhancer deletion, we performed EdU labeling. *Vsx2-SE Δ/Δ* mice were intercrossed to generate *Vsx2-SE Δ/Δ* , *Vsx2-SE $+/Δ$* , and *Vsx2-SE $+/+$* littermates. Animals were injected with EdU and 1 h later the retinæ were harvested, stained for EdU and scored (Supplemental Information).

Fig. 1 Cell-type specification in the *Vsx2* CRC-SE deletion retina. **A** Drawing of the original *Vsx2* CRC-SE identified by H3K27Ac ChIP-seq (black bar) and the original *Vsx2* CRC-SE deletion with coordinates in mm10. **B** The *Vsx2* CRC-SE deletion strain was crossed to GFP reporter strains for different retinal cell types as indicated. Micrographs of representative sections for wild type (upper) and knockout (lower panel) are shown. **C, D** Micrograph of *Grm6*-GFP expression in ON bipolar cells of wild type (**C**) and *Vsx2* SE knockout. DAPI staining is blue and GFP expression is green. **E** Bar plot showing mean with SEM for FACS analysis of the proportion of GFP+ cells for each strain shown in **B–D** ($n = 4$ retinæ from two mice, unpaired two-sided t -test, $p = 0.0016$). **F** tSNE plot of scRNA-seq of rod depleted retinæ showing each individual cell type for each genotype. The bipolar cluster is outlined. **G** Micrograph of DAPI (blue) and activated caspase three (red) immunostaining of retinal sections at P7 for the *Vsx2* SE knockout retina. A bar plot showing the scoring from two independent experiments with mean and standard deviation adjacent to the micrographs. **H** Micrographs of alkaline phosphatase stained clones of cells showing rods, Müller, amacrine, and bipolar neurons based on laminar position and morphology. **I** Bar plot showing the scoring of clone composition for each genotype and the total number of clones. There is a loss of bipolar containing clones and a corresponding increase in rod and Müller glia containing clones (chi-squared test, $p = 0.0053$). Abbreviations: ONL, outer nuclear layer; INL, inner nuclear layer; GCL, ganglion cell layer; CRC, core regulatory circuit. Scale bars: **B–D**, 25 mm, **G, H**, 10 mm. Source data are provided as a Source Data file.

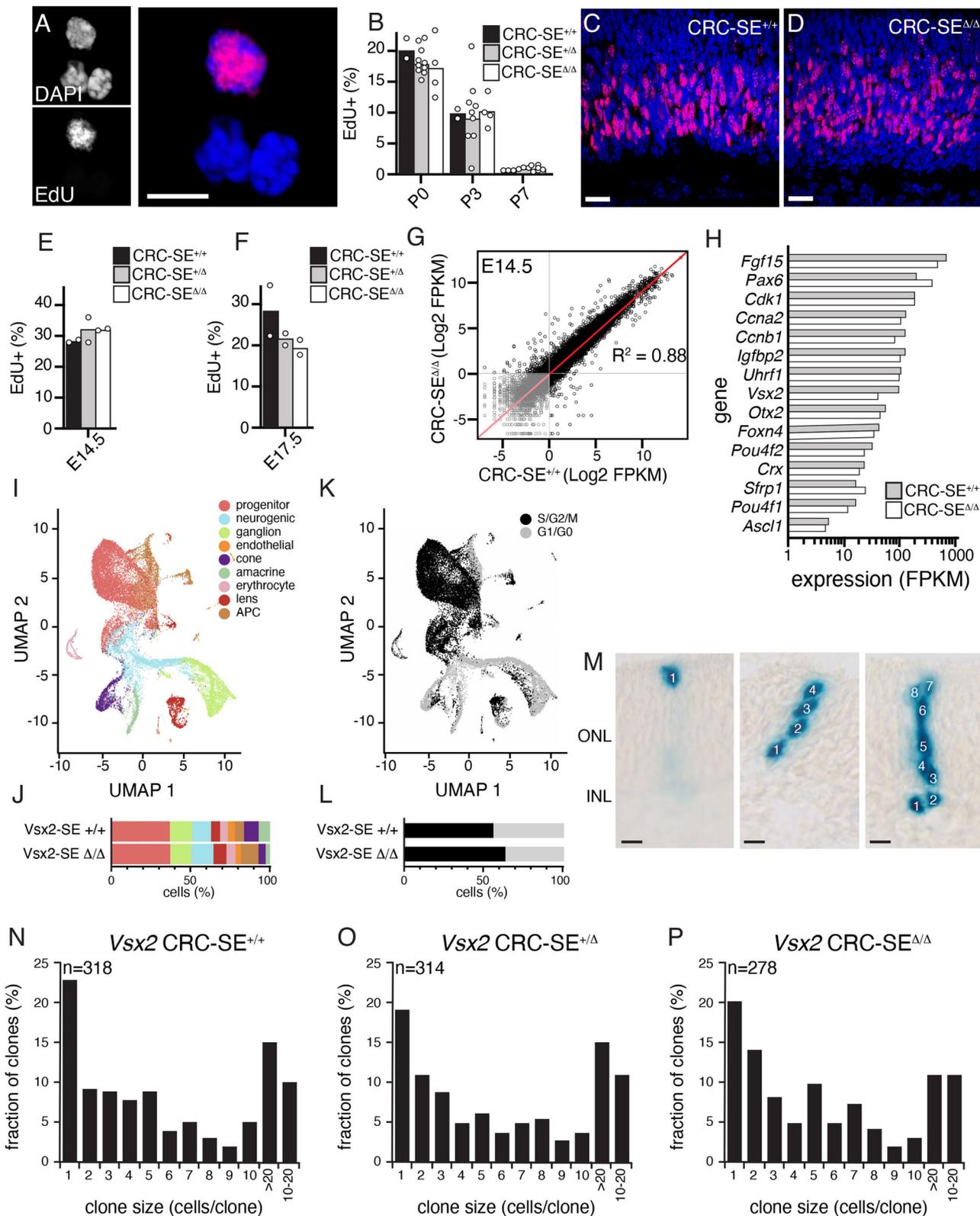
There was no observed difference in the proportion of EdU+ cells at E14.5, E17.5, P0, P3, or P7 (Fig. 2A–F). RNA-seq on E14.5 *Vsx2-SE Δ/Δ* and *Vsx2-SE $^{+/+}$* retinæ showed no observable difference in expression of cell cycle, retinal progenitor, or retinal development genes (Fig. 2G, H). We confirmed those results at single-cell resolution by performing scRNA-seq on E14.5 *Vsx2-SE Δ/Δ* and *Vsx2-SE $^{+/+}$* retinæ using the analysis method described above for Fig. 1 (Fig. 2I–L). Finally, to confirm that the proliferation of individual retinal progenitor cells was similar in *Vsx2-SE Δ/Δ* , *Vsx2-SE $^{+/+}$* , and *Vsx2-SE $^{+/+}$* retinæ, we performed clonal analysis using a replication incompetent retrovirus that expresses nuclear β -galactosidase (NIN-E, Supplemental Information). Briefly, retinæ from E14.5 embryos were harvested from a *Vsx2-SE $^{+/+}$* intercross and cultured as explants. They were infected with the NIN-E retrovirus, cultured for 14 days, genotyped and stained for β -galactosidase expression (Fig. 2M). Individual clones of cells derived from single infected retinal progenitor cells were reconstructed from serial sections and scored for cell number. There are no statistically significant differences in the number of clones ($p = 0.784$) or median number of cells per clone ($p = 0.663$) according to genotype (Fig. 2N, P). Taken together, these data suggest that deletion of the 3' 32 kb portion of the 41 kb *Vsx2-SE* affects the expression of *Vsx2* in bipolar neurons but not Müller glia or retinal progenitor cells. There is no defect in retinal progenitor cell proliferation and the cells that would normally become bipolar neurons now develop as rod photoreceptors and Müller glia.

The *Vsx2*-SE contains multiple evolutionarily conserved regions. As a first step to identify the module(s) of the *Vsx2* SE required for *Vsx2* expression in bipolar neurons, we aligned the evolutionarily conserved regions with our previous ChIP-seq and ATAC-seq data from the developing mouse retina³⁵. We identified three evolutionarily conserved regions (R1-28, R2-22, and R3-17) that are 28, 22, and 17 kb upstream of the *Vsx2* transcription start site and contained within the original 32 kb *Vsx2-SE* deletion (Fig. 3A). An additional evolutionarily conserved region (R0-37) was identified 37 kb upstream of the *Vsx2* transcription start site. R0-37 was not part of the original 32 kb deletion but is contained in the 41 kb *Vsx2* SE computationally defined from our original H3K27Ac ChIP-seq data (Fig. 3A). This region overlaps with an evolutionarily conserved sequence from the human genome that was previously shown to be sufficient for driving reporter gene expression in retinal progenitor cells in chick and mouse retina³⁰. R0-37 and R1-28 have ATAC-seq, Hi-C, and ChIP-seq profiles consistent with retinal progenitor cell expression while R3-17 is consistent with bipolar and/or Müller glial cell expression in adult retina (Fig. 3B, C)^{10,35}. Importantly, R3-17 overlaps with a genomic segment that has been previously shown by the Cepko research group to be sufficient for bipolar expression of a transgene in vivo^{28,37,43}. The

Hi-C data suggests that there is a looping interaction between the conserved elements and the *Vsx2* promoter in E14.5, P0, and adult retina (Fig. 3B).

The evolutionarily conserved regions have cell-type specific accessibility and activity. To further refine the cell-type specificity of evolutionarily conserved regions with open chromatin, we performed scATAC-seq on E14.5 and adult retinæ. Cell types were identified based on the open chromatin at promoters for cell-type specific genes (Supplemental Information). Consistent with the data in Fig. 3, chromatin in regions 0 and 1 were most accessible in retinal progenitor cells (Fig. 4A). Interestingly, those regions were also in an open chromatin state in Müller glia of the adult retina (Fig. 4B). This is in contrast to R3-17 which has scATAC-seq peaks specifically in bipolar neurons of the adult retina (Fig. 4B). Next, we identified putative transcription factors that may contribute to *Vsx2* gene expression by identifying consensus binding sites. There are 1,748 predicted motifs for 63 different transcription factors (p -value < 0.0001) and 1,453 motifs (62 transcription factors) have a positional weight matrix (PWM) > 10.0 (Fig. 4C). Only eight of the 62 transcription factors (*Dbp*, *Hlf*, *Otx2*, *Zfhx3*, *Prox1*, *Vsx2*, *Isl1*, and *Lhx4*) are expressed in bipolar neurons (102 motifs) and four (*Isl1*, *Lhx4*, *Prox1*, and *Vsx2*) have greater than 2-fold reduction in expression in the *Vsx2-SE Δ/Δ* retinæ (Fig. 4C–I). There are 40 motifs for *Isl1*, *Lhx4*, *Vsx2*, and *Prox1* across the entire *Vsx2-SE* but only 18 overlap with the scATAC-seq peaks in the evolutionarily conserved regions (Fig. 4B and Dataset S1). *Vsx2* is expressed in bipolar neurons and Müller glia, *Isl1*, and *Lhx4* are expressed primarily in bipolar neurons and *Prox1* is expressed in bipolar neurons and a subset of amacrine cells (Fig. 4E–I). We performed ChIP-seq for *Isl1*, *Prox1*, and *Lhx4* and compared those data to our previous results for *Vsx2*. *Lhx4* ChIP-seq failed quality control and *Vsx2*, *Isl1*, and *Prox1* all had ChIP-seq peaks at the consensus binding sites predicted from our analysis (Fig. 4D and Fig. S1A–C). While there was a one-to-one correspondence for ChIP-seq peaks and the predicted binding sites for *Vsx2* and *Prox1*, there were several predicted *Isl1* binding sites that did not have corresponding ChIP-seq peaks. This highlights the importance of validating predicted binding sites with ChIP-seq.

In order to determine if regions 1–3 were sufficient to drive expression in bipolar neurons and/or Müller glia, we subcloned them into a GFP reporter plasmid with a minimal promoter that has been described previously³⁷ (Fig. 4J). The individual R1-28, R2-22 and R3-17 reporter plasmids were separately co-electroporated in vivo at P0 with a constitutive Scarlet reporter to label all electroporated cells⁴⁴. We also included a reporter plasmid with the *Vsx2* consensus binding sites within R1-28. As a positive control for cell-type specific expression we used a previously identified bipolar-specific regulatory element³⁷. We scored proportion of GFP+, Scarlet+ /Scarlet+ cells for each cell



type for each reporter plasmid at P21 as done previously^{45,46}. R2-22 had little activity in this assay and the *Vsx2* binding sites conferred Müller glial cell expression (Fig. 4K, L). R1-28 had significantly more bipolar cell expression than the negative control ($p = 0.0018$) but R3-17 had significantly more bipolar cell activity than R1-28 ($p = 0.0028$). As further confirmation, we performed in vivo square wave electroporation of a plasmid that

confers bipolar neuron specific expression of CRISPR and gRNAs for the *Vsx2* gene or R3-17 enhancer element (see Supplemental Materials and Methods). For both sets of gRNAs, there was a reduction in bipolar neurons in the adult retina (Fig. S1E). Taken together, the scATAC-seq data and in vivo square wave electroporation data are consistent with the major bipolar-specific regulatory element residing in R3-17.

Fig. 2 Proliferation is normal in the *Vsx2* CRC-SE deletion retina. **A** Micrograph of dissociated P0 retinal cells showing EdU labeling (red) and DAPI (blue). **B** Bar plot showing the mean of retinæ from littermates for dissociated cell scoring. At P0 $n = 8$, P3 $n = 7$, P7 $n = 7$. There is no observable difference across genotypes or stages. **C, D** Representative micrograph of sections of P0 retina showing DAPI (blue) and EdU (red). **E, F** Scoring of EdU+ cells from micrographs of sections. Bar plots display the mean of biological duplicates. There was no observable difference across genotypes. **G** Scatterplot of each gene in RNA-seq analysis in the CRC-SE deletion strain and wild-type mice at E14.5 with a correlation coefficient of 0.88. **H** Bar plot of expression in FPKM for representative retinal progenitor, cell cycle, and differentiation genes. **I, J** tSNE of scRNA-seq for wild type and *Vsx2* CRC-SE deletion retina at E14.5 and corresponding stack bar plot. **K, L** Similar data as **I, J** but showing cell cycle distribution across retinal cell populations as indicated. **M** Representative micrograph of clones of cells from E14.5 retinal explants cultured for 2 weeks to score clone size. **N–P** Scoring of clone size distribution across the genotypes from littermates of a heterozygous intercross (Kruskal-Wallis test, number of clones $p = 0.784$, median number of cells per clone $p = 0.6634$). Abbreviations: FPKM, fragments per kilobase per million reads. Scale bars: **A**, 5 mm; **C, D**, and **M**, 10 mm. Source data are provided as a Source Data file.

To test the activity of regions 1–3 in proliferating retinal progenitor cells, we performed ex vivo square wave electroporation on E14.5 retinæ and in vivo square wave electroporation on P0 retinæ as described above. We also included R0-37 because of the strong scATAC-seq peak in retinal progenitor cells. Forty-eight hours after electroporation, EdU was added to the explant culture medium (E14.5 electroporation) or administered to the pup (P0 electroporation) by intraperitoneal injection and retinæ were harvested 1 h later. We scored the proportion of EdU+, GFP+, Scarlet+/Scarlet+ for each plasmid in biologic triplicate. R0-37 showed the highest level of expression in proliferating retinal progenitor cells (Fig. S1F, G).

The *Vsx2* SE contains distinct developmental—and cell type—specific modules. The developmental—and cell type—specific chromatin profiling combined with the in vivo electroporation data suggest that the 41 kb *Vsx2* SE is bifunctional with R0-37 and R1-28 being active in retinal progenitor cells and Müller glia and R3-17 being active in bipolar neurons. We used CRISPR-Cas9 to delete each conserved region and the *Vsx2* binding sites in mice (Fig. 5A). There are two adjacent *Vsx2* consensus binding sites in R1-28 that were deleted individually and in combination (Fig. 5A). We also deleted the combination of R0-37 through R1-28 to determine if they function together in retinal progenitor cells. In total, eight new mouse lines were analyzed. Multiple independent founder lines were generated for each deletion and they were outcrossed to wild type (C57BL/6J) mice before intracrossing individual sublines for each deletion. We compared these *Vsx2* SE deletion strains to the mouse strain (*orf*) that has a spontaneous mutation in the *Vsx2* gene and microphthalmia^{31,34}. The R0-37 mice had microphthalmia (Fig. 5B, C) and the retinæ were hypocellular (Fig. 5D). The phenotype was not as severe as that observed in the *orf* mice (Fig. 5B–D). The R1-28 mice had a normal size eye but the retina was approximately half the thickness of a wild-type retina (Fig. 5B–D). The combined deletion of R0-37 through R1-28 led to microphthalmia and the eyes were smaller than either single deletion alone (Fig. 5B–D). Indeed, the eyes were so small that we were unable to isolate sufficient retinæ from the eyes for molecular profiling or histology (Fig. 5D). Visual acuity measurements under scotopic conditions were consistent with the microphthalmia (R0-37 and R0-37–R1-28), reduction in retinal thickness (R1-28) and loss of bipolar neurons (R3-17) (Fig. 5E). The mutations in the *Vsx2* consensus binding sites and deletion of R2-22 had no effect on visual acuity or retinal development (Fig. 5A–E).

To elucidate any changes in cell-type distribution across these eight deletion strains, we performed bulk RNA-seq, scRNA-seq, immunostaining and qRT-PCR for all retinal cell types in adult retina. R2-22 deletion and the three consensus binding site mutant strains had no discernable retinal defect even though ChIP-seq confirmed reduced binding in the double consensus mutant (Fig. 6A, Fig. S1, and Datasets S1–S3). R3-17 deletion

fully recapitulated the loss of bipolar neurons in the original 32 kb *Vsx2*-SE deletion by RNA-seq, qRT-PCR, immunostaining, and morphologic analysis (Fig. 6A–C and Datasets S1–S3). Deletion of R1-28 led to a hypocellular retina that had no reduction in the proportion of EdU+ cells at E14.5, E17.5 or P0 and normal distribution of bipolar and amacrine neurons, cone and rod photoreceptors and Müller glia by RNA-seq, qRT-PCR, and immunostaining (Fig. 6A–E and Datasets S2–S4). There was a reduction in expression of some ganglion cell genes in the R1-28 deletion suggesting that there may be a defect in early cell cycle exit (Fig. 6B, C, E). However, bipolar cells and Müller glia were present in the R1-28 deletion strain (Fig. 6B, C, E, F). The R0-37 deletion mouse strain had microphthalmia but bipolar neurons and Müller glia were normal (Datasets S1–S3, S5). Therefore, there is a functional and physical separation of the region of the *Vsx2* SE that is required for retinal progenitor cell proliferation (R0-37 and R1-28) and for bipolar neuron formation (R3-17).

Discussion

Over the past 2 decades, several of the major neurogenic transcription factors have been knocked out in mice resulting in dramatic defects in central nervous system development. However, a major challenge in the field is discerning the primary effect from secondary non-specific effects. For example, if neural progenitor proliferation is altered by knocking out a transcription factor in mice, there may be secondary non-specific effects on cell fate specification due to the evolutionarily conserved birth order of neurons and/or glia. Conditional knockout approaches can help discern between primary and secondary effects but it may be difficult to obtain *Cre* transgenic lines with the level of cell and/or developmental stage-specific control that is required. Our data presented here on the *Vsx2* super-enhancer provides an alternative strategy for elucidating the function of neurogenic transcription factors in development. A single super-enhancer contains all the modules required for expression of *Vsx2* in retinal progenitor cells, Müller glia and bipolar neurons. By deleting those individual modules in mice using CRISPR-Cas9, we were able to specifically study the function of *Vsx2* in the different cell types at distinct stages of development. Super-enhancers have been identified for many of the neurogenic transcription factors and using existing scATAC-seq and scRNA-seq along with enhancer profiling it is now possible to identify similar modules other super-enhancers that can be tested and validated in vivo.

A modular super-enhancer. In our original study³⁵, we did not delete the entire super-enhancer that was identified by H3K27Ac ChIP-seq because of the close proximity of the 5' end of the *Vsx2* SE to the 3' end of the *Lin52* gene. Previous studies had shown that a region in that deletion contained a module that was important for bipolar expression using in vivo square wave

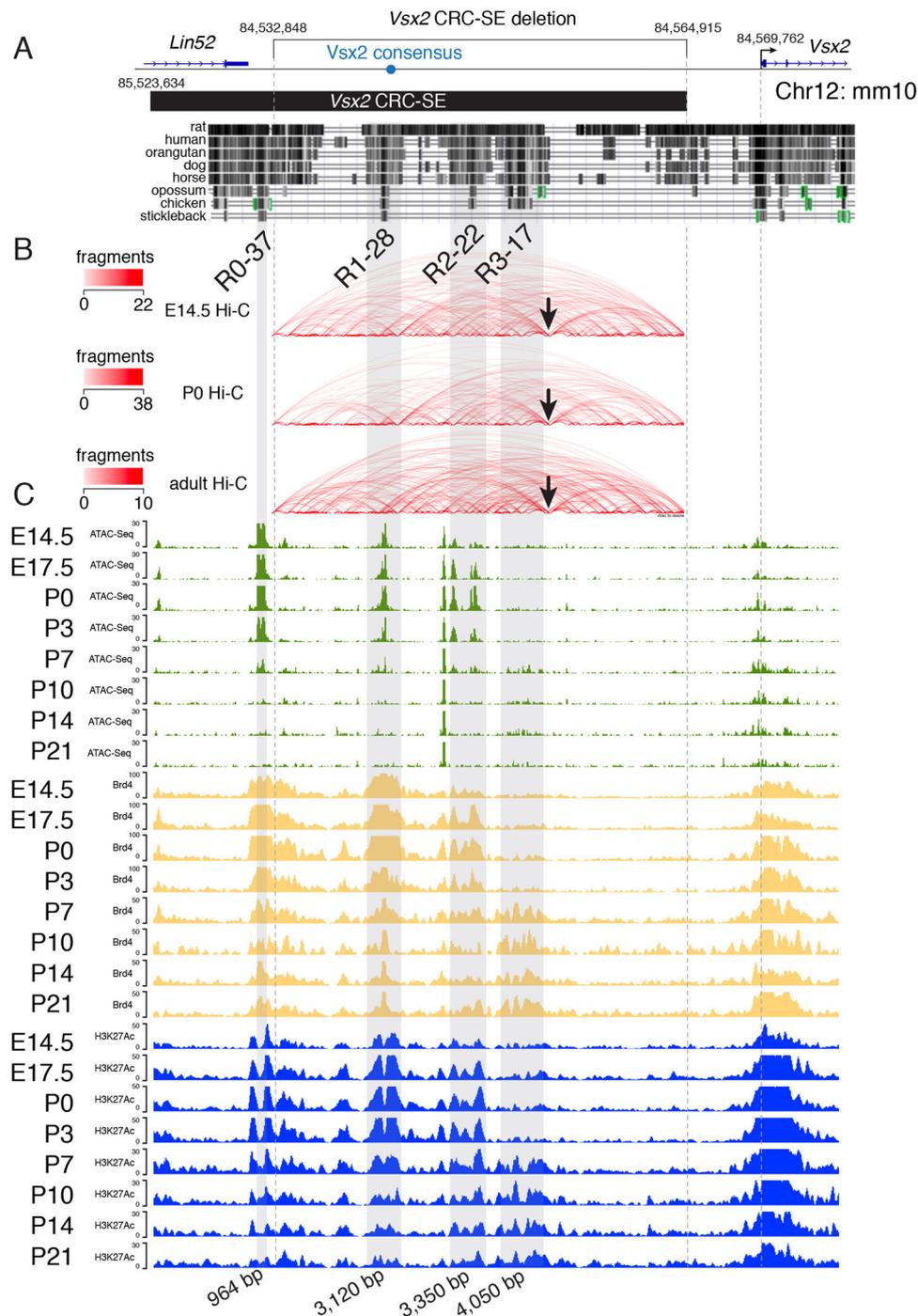
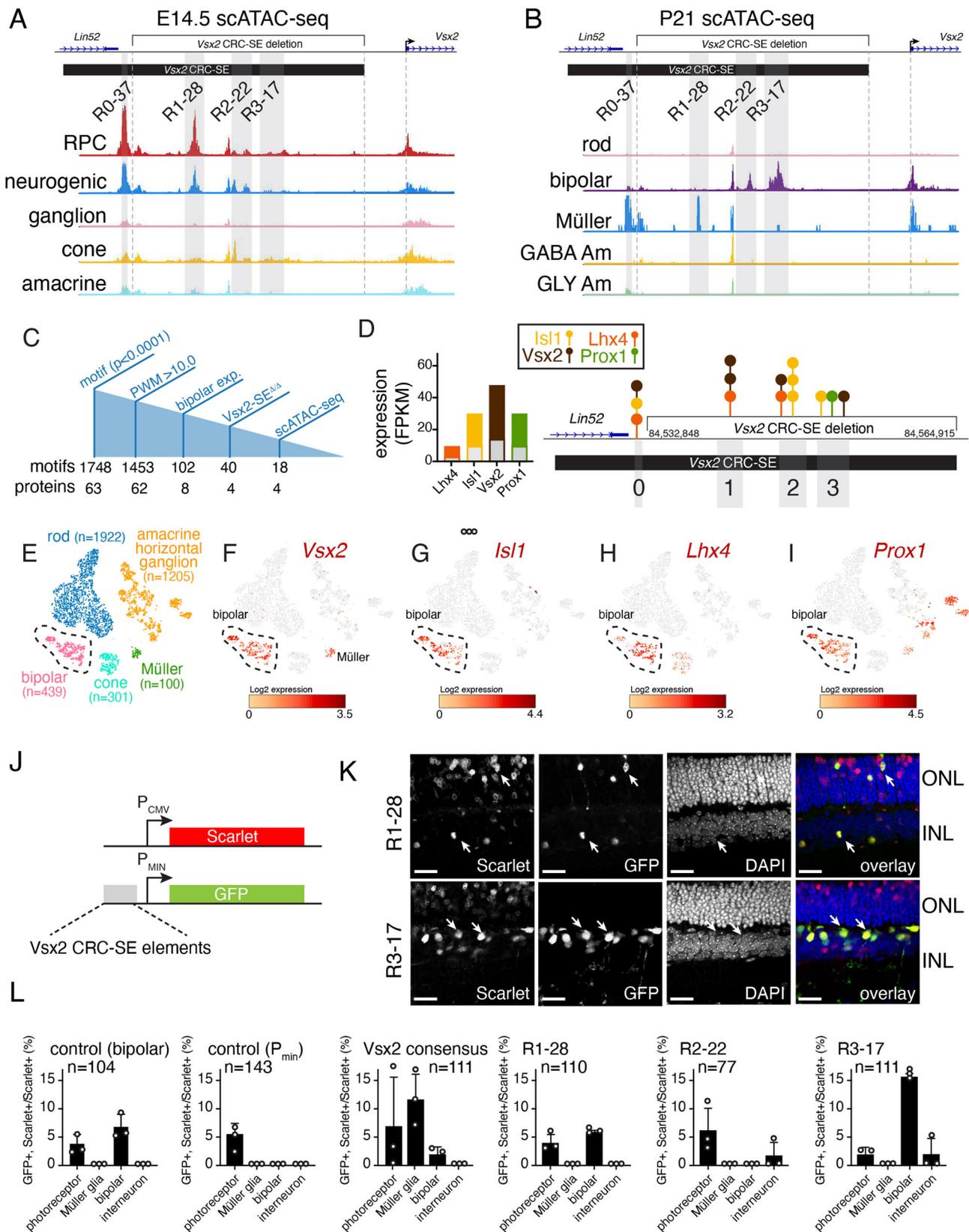


Fig. 3 Chromatin landscape of Vsx2 CRC-SE. **A** Evolutionary conservation of Vsx2 CRC-SE regions. **B** Hi-C interaction map at E14.5, P0 and adult showing a separation just downstream of the evolutionarily conserved regions (arrow). **C** Representative ATAC-seq and ChIP-seq for Brd4 and H3K27Ac across developmental stages for each region.

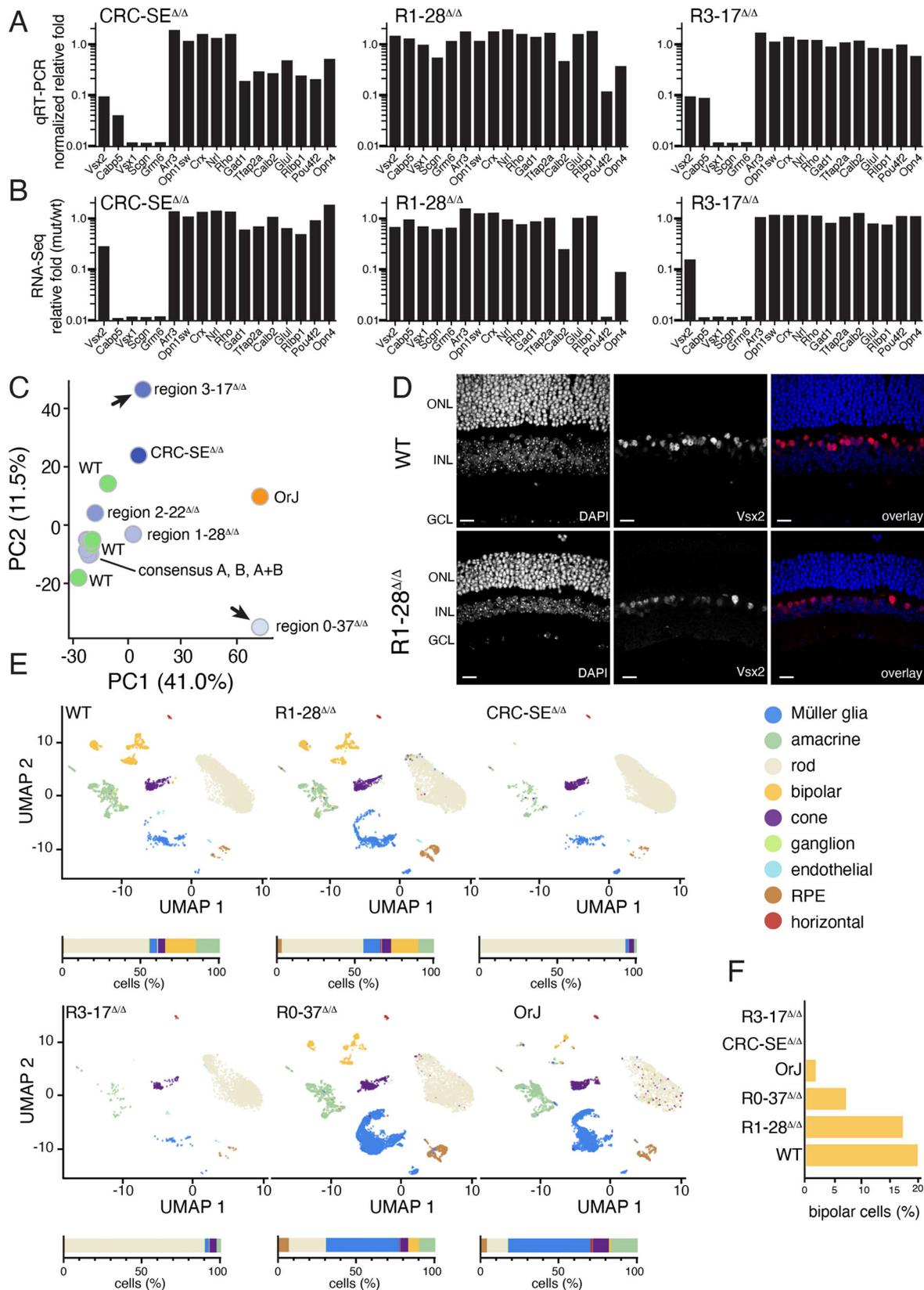
electroporation^{37,38}. Our analysis of R3-17 of the Vsx2 SE in genetically engineered mice in vivo agreed with their original experimental results and highlights the power of combining in vivo square wave electroporation with genetic engineering of mice using CRISPR/Cas9. The retinal progenitor cell regulatory elements were not identified in those prior studies and here we show that it resides adjacent to the *Lin52* gene but within the original Vsx2-SE defined computationally. Our in vivo studies allowed us to separate the bipolar gene expression regulatory module from the retinal progenitor cell module. This is in contrast to the *Vsx2 orf* mutant mouse strain that has

microphthalmia and lacks bipolar neurons. Previous analyses of SE elements in vivo have shown that multiple regions work cooperatively or in an additive manner to regulate gene expression^{16,17,19,21}. Our identification of two modules (R0-37 and R1-28) important for retinal progenitor cell expression agree with the additive model and deletion of R0-37 through R1-28 led to a further reduction in eye/retina size. Previous studies suggest that there are proliferating retinal progenitor cells that do not express Vsx2^{29,30} so it is possible that any residual retinal progenitor cells in the R0-37/R1-28 deletion mouse retina would be those Vsx2⁻ retinal progenitor cells.



Bipolar neurons become rods and Müller glia in the *Vsx2* SE deletion. Individual retinal progenitor cells can produce combinations of rods, bipolar neurons, and Müller glia⁴⁷⁻⁵¹. Using complementary approaches, we showed that retinal progenitor cells produce more rods and Müller glia in the *Vsx2* SE deletion mouse

strain. The bipolar element has now been localized to R3-17 which overlaps with the region previously shown to control bipolar cell expression^{37,38}. We cannot rule out the possibility that there are a small number of bipolar neurons born that undergo apoptosis in the *Vsx2* SE deletion strain but if that is the case, it would be a



expression. The *Vsx2* CRC-SE is an important case in point. It meets the computational criteria for a CRC-SE¹⁰ because there are two *Vsx2* consensus binding sites in R1-28 as well as other sites in the *Vsx2* CRC-SE. We showed by ChIP-seq that *Vsx2* binds the R1-28 sites and mutating the sites in mice in vivo led to reduced *Vsx2* binding. However, there was no discernable

phenotype in the retina from mice lacking those sites. It is possible that the *Vsx2* SE is a CRC-SE because the ChIP-seq showed other *Vsx2* binding sites that overlapped with *Isl1* and *Prox1*. It is possible that these transcription factors work together at R3-17 to regulate bipolar cell expression of *Vsx2*. Interestingly, the consensus binding sites for *Vsx2* in R1-28 were sufficient to promote

Fig. 6 Separation of functional regions of the *Vsx2* SE. A, B qRT-PCR and RNA-seq of genes for retinal cell types for each of the indicated strains. Bipolar cells are missing from the original SE deletion strain and the R3-17 strain while there is a defect in retinal ganglion cell formation in the R1-28 deletion. **C** Principal component analysis (PCA) of bulk RNA-seq for each deletion mouse strain. The three consensus binding mutants are stacked on top of each other and indistinguishable from WT. **D** Micrographs showing the expression of *Vsx2* in bipolar cells and Müller glia in the adult wild type and R1-28 deletion strain. The retinal thickness is about half that of the wild type but most cell types are present. **E** UMAP plots for WT, R1-28, CRC-SE, R3-17, R0-37, and *orl* retinae. Stack bar plots show the percentage of each cell type below each UMAP. **F** Bar plot of the percentage of bipolar neurons in each of the retinae shown in **(E)**. Abbreviations: ONL, outer nuclear layer; INL, inner nuclear layer; GCL, ganglion cell layer; PC1, PC2, principal components 1 and 2. Scale bars: 25 μ m.

expression of a heterologous reporter gene in Müller glia even though there was no defect in Müller glial cell fate specification in the double consensus knockout. Thus, the *Vsx2* binding sites in R1-28 appear to be sufficient for expression but not necessary.

A systematic approach to dissection transcription factor SEs in neurogenesis. Our original deletion of the *Vsx2* SE was imprecise because it was based only on identification of the super-enhancer domain and proximity to adjacent genes. In this study, we combined those previous ChIP-seq results with scRNA-seq and scATAC-seq to refine the developmental stage—and cell type—specific regions of the *Vsx2* SE. Using this integrated approach, we were able to identify each module within the *Vsx2* SE and discern their developmental stage—and cell type—specific activity *in vivo*. We propose that this approach using bulk RNA-seq and bulk ChIP-seq combined with scRNA-seq and scATAC-seq can be used to produce an atlas of multi-functional super-enhancers for neurogenic transcription factors during development. This is important because it would allow researchers to discern specific functions of transcription factors in particular cell types or at particular stages of development. Moreover, by comparing those essential domains across species, one can begin to elucidate the role of such super-enhancer sub regions in evolution of brain expression patterns and brain development. This provides an advantage to traditional gene knockout or conditional knockout approaches because it can be difficult to discern primary from secondary effects. With advances in human neural organoid development from embryonic stem cells, it is now possible to perform cross-species comparisons in mice and humans.

Methods

Mouse strains. All animal procedures and protocols were approved by the St. Jude Laboratory Animal Care and Use Committee under protocol number 393-100500. All studies conform to federal and local regulatory standards. Mice were housed on ventilated racks on a standard 12 h light–dark cycle. Wild-type C57BL/6 J mice were purchased from the Jackson Laboratory (Bar Harbor, ME, #000664). For timed pregnancy, individual male mice were housed with three females in a single cage. Plugged/pregnant females (identified by visual examination), were isolated and embryos or pups were harvested at the appropriate time. Both males and females were combined for this study. Conserved regions within the *Vsx2* SE were identified by examination of evolutionary conservation in the UCSC mm9 genome build. *Vsx2* modified mouse models were created using CRISPR-Cas9 technology (Supplemental Information).

Imaging. Images were taken with the Zeiss LSM 700 confocal microscope using the 40X lens. Brightness and contrast were modified for images presented in the figures for the IF studies. Raw original data are available for all datasets and probes.

RNA isolation. RNA was extracted from individual TRIzol (Life Technologies) preparations via a phenol-chloroform extraction. Samples were first dissociated by pipetting retina and TRIzol vigorously. A 1:4 volume of chloroform (Sigma) was then added to each sample and incubated at room temperature for 3 min followed centrifugation at 12,000 3 g at 4 C for 15 min. The aqueous layer was then transferred to a siliconized Eppendorf tube followed by the addition of 2.0 mL glycogen (Roche) and 500 mL isopropanol (Fisher Scientific). Samples were incubated at room temperature for 10 min followed by centrifugation at 12,000 3 g at 4 C for

15 min. Samples were then washed twice with ice-cold 80% EtOH (Fisher) to remove salts, resuspended in DEPC H₂O, and the concentration was determined with a NanoDrop (ThermoScientific).

Libraries were prepared from 500 ng total RNA with the TruSeq Stranded Total RNA Library Prep Kit according to the manufacturer's directions (Illumina). Paired-end 100-cycle sequencing was performed on HiSeq 2000 or HiSeq 2500 sequencers according to the manufacturer's directions (Illumina).

qRT-PCR. cDNA was made from 200 ng of RNA from *Vsx2*-Consensus, *Vsx2*SE-DoubleConsensus, *Vsx2*SE-Region1, *Vsx2*SE-Region2, *Vsx2*SE-Region3, and *Vsx2*SE knockout retinae, as well as for their wild type and heterozygous littermates (Applied Biosystems 4387406). cDNA was loaded onto a Custom TaqMan Array Card (Applied Biosystems 4342249) run on a QuantStudio 7 Flex (ThermoScientific) system. “Undetermined” values were set to a Ct of 40 as the limit of detection of the assay.

RNA-seq. RNA was quantified using the Quant-iT RiboGreen assay (Life Technologies) and quality checked by 2100 Bioanalyzer RNA 6000 Nano assay (Agilent,) 4200 TapeStation High Sensitivity RNA ScreenTape assay (Agilent,) or LabChip RNA Pico Sensitivity assay (PerkinElmer) prior to library generation. Libraries were prepared from total RNA with the TruSeq Stranded Total RNA Library Prep Kit according to the manufacturer's instructions (Illumina, PN 20020599). Libraries were analyzed for insert size distribution on a 2100 BioAnalyzer High Sensitivity kit (Agilent Technologies,) 4200 TapeStation D1000 ScreenTape assay (Agilent Technologies,) or Caliper LabChip GX DNA High Sensitivity Reagent Kit (PerkinElmer.) Libraries were quantified using the Quant-iT PicoGreen ds DNA assay (Life Technologies) or low pass sequencing with a MiSeq Nano kit (Illumina) Paired-end 100-cycle sequencing was performed on a NovaSeq 6000 (Illumina). For PCA analysis, only protein-coding genes with the annotation level (https://www.encodegenes.org/pages/data_format.html) in 1 (verified loci) and 2 (manually annotated loci) were included in the analysis. With input of read counts for all samples, counts per million mapped reads (CPMs) were obtained by using the function *cpm* in the edgeR package. Genes were removed when the corresponding CPMs for all samples were smaller than the CPM whose corresponding raw read count is 10. Then, the top 3000 most variable genes were selected by ranking the mean absolute deviation (MAD) of the log₂-transformed CPMs in descending order. Based on these most variable genes, the PCA analysis was performed on the log₂-transformed CPMs by using the *prcomp* function available in the standard R language. The top two principal components (PCs) were used to draw the PCA figures.

Statistics and reproducibility. Littermates of both sexes were randomly selected for analyses. Investigators were blinded when scoring images. No statistical method was used to predetermine sample size. There were no instances in which repeat experiments yielded conflicting results, suggesting reproducibility of our experiments. GraphPad Prism 8 software was used to calculate statistical measures. No data were excluded from the analyses.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The sequencing data generated in this study are publicly available in the GEO database under accession code “GSE169262”. The reference genome is mm10 and can be accessed by downloading the mouse Gencode release M22 at the following “FTP link [ftp://ftp.ebi.ac.uk/pub/databases/gencode/Gencode_mouse/release_M22/]”. All other relevant data supporting the key findings of this study are available within the article and its Supplementary Information files or from the corresponding author upon reasonable request. Source data are provided with this paper.

Code availability

Code used to process all scRNA-seq data can be found on “Github [https://github.com/CodyRamirez/StJude_Scripts/tree/main/Vsx2_SE]” (<https://doi.org/10.5281/zenodo.5777750>) under the MIT license.

Received: 2 February 2021; Accepted: 20 December 2021;

Published online: 11 January 2022

References

- Livesey, F. J. & Cepko, C. L. Vertebrate neural cell-fate determination: lessons from the retina. *Nat. Rev. Neurosci.* **2**, 109–118 (2001).
- Amano, T. et al. Chromosomal dynamics at the Shh locus: limb bud-specific differential regulation of competence and active transcription. *Dev. Cell* **16**, 47–57 (2009).
- Kleinjan, D. A. & van Heyningen, V. Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am. J. Hum. Genet.* **76**, 8–32 (2005).
- Mifsud, B. et al. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat. Genet.* **47**, 598–606 (2015).
- Sanyal, A., Lajoie, B. R., Jain, G. & Dekker, J. The long-range interaction landscape of gene promoters. *Nature* **489**, 109–113 (2012).
- Li, Y. et al. CRISPR reveals a distal super-enhancer required for Sox2 expression in mouse embryonic stem cells. *PLoS One* **9**, e114485 (2014).
- Zhou, H. Y. et al. A Sox2 distal enhancer cluster regulates embryonic stem cell differentiation potential. *Genes Dev.* **28**, 2699–2711 (2014).
- Visel, A. et al. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**, 854–858 (2009).
- Shen, Y. et al. A map of the cis-regulatory sequences in the mouse genome. *Nature* **488**, 116–120 (2012).
- Aldiri, I. et al. The dynamic epigenetic landscape of the retina during development, reprogramming, and tumorigenesis. *Neuron* **94**, 550–568 e510 (2017).
- Pott, S. & Lieb, J. D. What are super-enhancers? *Nat. Genet.* **47**, 8–12 (2015).
- Hnisz, D. et al. Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).
- Loven, J. et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320–334 (2013).
- Whyte, W. A. et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
- Blobel, G. A., Higgs, D. R., Mitchell, J. A., Notani, D. & Young, R. A. Testing the super-enhancer concept. *Nat. Rev. Genet.* <https://doi.org/10.1038/s41576-021-00398-w> (2021).
- Hay, D. et al. Genetic dissection of the α -globin super-enhancer in vivo. *Nat. Genet.* **48**, 895–903 (2016).
- Hnisz, D. et al. Convergence of developmental and oncogenic signaling pathways at transcriptional super-enhancers. *Mol. Cell* **58**, 362–370 (2015).
- Huang, J. et al. Dissecting super-enhancer hierarchy based on chromatin interactions. *Nat. Commun.* **9**, 943 (2018).
- Huang, J. et al. Dynamic control of enhancer repertoires drives lineage and stage-specific transcription during hematopoiesis. *Dev. Cell* **36**, 9–23 (2016).
- Parker, S. C. et al. Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. *Proc. Natl. Acad. Sci. USA* **110**, 17921–17926 (2013).
- Shin, H. Y. et al. Hierarchy within the mammary STAT5-driven Wap super-enhancer. *Nat. Genet.* **48**, 904–911 (2016).
- Haubst, N. et al. Molecular dissection of Pax6 function: the specific roles of the paired domain and homeodomain in brain development. *Development* **131**, 6131–6140 (2004).
- Castro, D. S. et al. A novel function of the proneural factor Ascl1 in progenitor proliferation identified by genome-wide characterization of its targets. *Genes Dev.* **25**, 930–945 (2011).
- Nishida, A. et al. Otx2 homeobox gene controls retinal photoreceptor cell fate and pineal gland development. *Nat. Neurosci.* **6**, 1255–1263 (2003).
- Collinson, J. M., Quinn, J. C., Hill, R. E. & West, J. D. The roles of Pax6 in the cornea, retina, and olfactory epithelium of the developing mouse embryo. *Dev. Biol.* **255**, 303–312 (2003).
- Marquardt, T. et al. Pax6 is required for the multipotent state of retinal progenitor cells. *Cell* **105**, 43–55 (2001).
- Liu, I. S. et al. Developmental expression of a novel murine homeobox gene (Chx10): evidence for roles in determination of the neuroretina and inner nuclear layer. *Neuron* **13**, 377–393 (1994).
- Rowan, S. & Cepko, C. L. Genetic analysis of the homeodomain transcription factor Chx10 in the retina using a novel multifunctional BAC transgenic mouse reporter. *Dev. Biol.* **271**, 388–402 (2004).
- Vitorino, M. et al. Vsx2 in the zebrafish retina: restricted lineages through derepression. *Neural Dev.* **4**, 14 (2009).
- Buenaventura, D. F., Ghinia-Tegla, M. G. & Emerson, M. M. Fate-restricted retinal progenitor cells adopt a molecular profile and spatial position distinct from multipotent progenitor cells. *Dev. Biol.* **443**, 35–49 (2018).
- Burmeister, M. et al. Ocular retardation mouse caused by Chx10 homeobox null allele: impaired retinal progenitor proliferation and bipolar cell differentiation. *Nat. Genet.* **12**, 376–384 (1996).
- Ferda Percin, E. et al. Human microphthalmia associated with mutations in the retinal homeobox gene CHX10. *Nat. Genet.* **25**, 397–401 (2000).
- Livne-Bar, I. et al. Chx10 is required to block photoreceptor differentiation but is dispensable for progenitor proliferation in the postnatal retina. *Proc. Natl. Acad. Sci. USA* **103**, 4988–4993 (2006).
- Truslove, G. M. A gene causing ocular retardation in the mouse. *J. Embryol. Exp. Morphol.* **10**, 652–660 (1962).
- Norrie, J. L. et al. Nucleome dynamics during retinal development. *Neuron* **104**, 512–528 e511 (2019).
- Saint-Andre, V. et al. Models of human core transcriptional regulatory circuitries. *Genome Res.* **26**, 385–396 (2016).
- Kim, D. S., Matsuda, T. & Cepko, C. L. A core paired-type and POU homeodomain-containing transcription factor program drives retinal bipolar cell gene expression. *J. Neurosci.* **28**, 7748–7764 (2008).
- Goodson, N. B., Kaufman, M. A., Park, K. U. & Brzezinski, J. A. T. Simultaneous deletion of Prdm1 and Vsx2 enhancers in the retina alters photoreceptor and bipolar cell fate specification, yet differs from deleting both genes. *Development* <https://doi.org/10.1242/dev.190272> (2020).
- Dhingra, A. et al. Probing neurochemical structure and function of retinal ON bipolar cells with a transgenic mouse. *J. Comp. Neurol.* **510**, 484–496 (2008).
- Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 e3529 (2021).
- Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).
- Welch, J. D. et al. Single-cell multi-omic integration compares and contrasts features of brain cell identity. *Cell* **177**, 1873–1887 e1817 (2019).
- Rowan, S. & Cepko, C. L. A POU factor binding site upstream of the Chx10 homeobox gene is required for Chx10 expression in subsets of retinal progenitor cells and bipolar cells. *Dev. Biol.* **281**, 240–255 (2005).
- Matsuda, T. & Cepko, C. L. Electroporation and RNA interference in the rodent retina in vivo and in vitro. *Proc. Natl. Acad. Sci. USA* **101**, 16–22 (2004).
- Emerson, M. M. & Cepko, C. L. Identification of a retina-specific Otx2 enhancer element active in immature developing photoreceptors. *Dev. Biol.* **360**, 241–255 (2011).
- Jean-Charles, N., Buenaventura, D. F. & Emerson, M. M. Identification and characterization of early photoreceptor cis-regulatory elements and their relation to Onecut1. *Neural Dev.* **13**, 26 (2018).
- Wang, S., Sengel, C., Emerson, M. M. & Cepko, C. L. A gene regulatory network controls the binary fate decision of rod and bipolar cells in the vertebrate retina. *Dev. Cell* **30**, 513–527 (2014).
- Turner, D. L. & Cepko, C. L. A common progenitor for neurons and glia persists in rat retina late in development. *Nature* **328**, 131–136 (1987).
- Wang, S. & Cepko, C. L. Photoreceptor fate determination in the vertebrate retina. *Investig. Ophthalmol. Vis. Sci.* **57**, ORSFe1–ORSFe6 (2016).
- Hafler, B. P. et al. Transcription factor Olig2 defines subpopulations of retinal progenitor cells biased toward specific cell fates. *Proc. Natl. Acad. Sci. USA* **109**, 7882–7887 (2012).
- Honnell, V. et al. Identification of a modular super-enhancer in murine retinal development. Github <https://doi.org/10.5281/zenodo.5777750> (2021).

Acknowledgements

This work was supported by Cancer Center Support (CA21765) and grants to M.A.D. from the National Institutes of Health (EY014867, EY018599, and CA168875) and 1F31EY031577 to V.H. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The research was also supported by American Lebanese Syrian Associated Charities. M.A.D. was also supported by Alex's Lemonade Stand, and Tully Family and Peterson Foundations. We thank Shondra Miller in the Center for Advanced Genome Engineering for designing the gRNAs. We thank Valerie Stewart in the Neuroembryology Laboratory for producing transgenic mice by micro-injection. We thank Tom Reynolds and Everest Ouyang for technical assistance. We thank the Center for Applied Bioinformatics for computational assistance.

Author contributions

V.H. and M.A.D. conceived and designed the study, V.H., J.L.N., J.Z., and M.A.D. collected the data, V.H., J.L.N., A.G.P., C.R., Y.-H.L., and S.W. performed computational analysis, V.H., J.L.N., A.G.P., and M.A.D., analyzed and interpreted the data, V.H. and M.A.D. drafted the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-27924-y>.

Correspondence and requests for materials should be addressed to Michael A. Dyer.

Peer review information *Nature Communications* Jian Xu, Judith Zaugg, and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022