Heterochronic transcription factor expression drives cone-dominant retina development in 13-lined ground squirrels.

Kurt Weir[1*], Pin Lyu[2*], Sangeetha Kandoi[1], Roujin An[1], Nicole Pannullo[1], Isabella Palazzo[1], Jared A. Tangeman[1], Jun Shi[6], Steven H. DeVries[6], Dana K. Merriman[7], Jiang Qian[2†], and Seth Blackshaw[1-5†].

Affiliations: [1]Department of Neuroscience, [2]Department of Ophthalmology, [3]Department of Neurology, [4]Institute for Cell Engineering, [5]Kavli Neuroscience Discovery Institute, Johns Hopkins University School of Medicine, Baltimore, MD 21205 USA
[6]Department of Ophthalmology, Northwestern University Feinberg School of Medicine, Chicago, IL, 60611, USA
[7]Department of Biology, University of Wisconsin Oshkosh, Oshkosh, WI, USA.

Abstract:

Evolutionary adaptation to diurnal vision in ground squirrels has led to the development of a cone-dominant retina, in stark contrast to the rod-dominant retinas of most mammals. The molecular mechanisms driving this shift remain largely unexplored. Here, we perform single-cell RNA sequencing (scRNA-Seq) and chromatin accessibility profiling (scATAC-Seq) across developmental retinal neurogenesis in the 13-lined ground squirrel (13LGS) to uncover the regulatory basis of this adaptation. We find that 13LGS cone photoreceptors arise not only from early-stage neurogenic progenitors, as seen in rod-dominant species like mice, but also from late-stage neurogenic progenitors. This extended period of cone generation is driven by a heterochronic shift in transcription factor expression, with cone-promoting factors such as *Onecut2*, *Pou2f1*, and *Zic3* remaining active in late-stage progenitors, and factors that promote cone differentiation such as *Thrb*, *Rxrg*, and *Mef2c* expressed precociously in late-stage neurogenic progenitors. Functional analyses reveal that *Zic3* and *Mef2c* are sufficient to promote cone photoreceptor and repress rod specification, and act through species-specific regulatory elements that drive their expression in late-stage progenitors. These results demonstrate that evolutionary modifications to gene regulatory networks underlie the development of cone-dominant retinas and provide insight into mechanisms of sensory adaptation and potential strategies for cone photoreceptor regeneration in vision disorders.

Introduction:

Evolutionary adaptation to new ecological niches often requires the development of new sensory, motor, and/or behavioral capacities. These adaptations are typically mediated by structural changes to the nervous system which arise from changes in the composition and connectivity of individual neuronal subtypes [1,2]. While the molecular mechanisms underlying these changes are generally poorly understood, the vertebrate retina serves as an ideal system for investigating them. Shifts in temporal niche—such as transitions between predominantly nocturnal and diurnal behavior—have occurred repeatedly across diverse vertebrate lineages [3–5]. These shifts inevitably require adaptive modifications in the rod and cone visual pathways, which detect dim and bright light, respectively.

Mammals represent a notable exception to the general trend in which diurnal vertebrate species possess cone-dominant retinas. This deviation reflects the long-term effects of a selective nocturnal bottleneck that occurred before the Cretaceous-Paleogene mass extinction event, preceding the mammalian radiation [5,6]. Despite considerable variation in photoreceptor subtype composition among mammals [7–9], even strongly diurnal species typically have rod-dominant retinas, with rod-to-cone ratios ranging from approximately 5:1 to 20:1 [8,10]. Strictly nocturnal species, such as mice, exhibit even more pronounced rod dominance of 33:1 [11]. However, diurnal ground squirrels are a major exception to this rule, possessing an inverted rod-to-cone ratio, having only one rod for every 6–7 cones throughout most of the retina [12,13], and as many as 50 cones for every rod in the central visual streak [13]. By examining the development of the cone-dominant ground squirrel retina and comparing it to that of nocturnal rodents like mice, we can identify molecular mechanisms underlying this dramatic and evolutionarily rapid shift in photoreceptor subtype composition.

In rod-dominant mouse and human retinas, cones are generated primarily during early stages of retinal neurogenesis, whereas rods are predominantly produced in later stages [14–17]. The progenitors that selectively give rise to cones and rods are in turn distinguished by the activity of a network of transcription factors that confer early and late-stage temporal identity. The progenitors that give rise to cones and rods are distinguished by the activity of a network of transcription factors that establish early- and late-stage temporal identities. Cones originate from a subpopulation of early-stage *Otx2* and *Neurod1*-positive neurogenic progenitors that express *Onecut1/2* and *Pou2f1/2*, which confer competence to generate both cones and horizontal cells and activate transcription factors that promote cone differentiation while simultaneously repressing rod-specific factor [18,19]. In contrast, late-stage *Otx2*- and *Neurod1*-positive neurogenic progenitors primarily generate postmitotic rod precursors, which express transcription factors such as *Nrl* and *Nr2e3* that suppress cone-specific genes, while also activating rod-specific genes [20–24]. Postmitotic cone precursors, in turn, express transcription factors such as *Thrb*, *Rxrg*, and *Sall3*, which further activate cone-specific genes and inhibit rod-specific gene expression [25,26]. Evolutionary changes in the rod-to-cone ratio may result directly from alterations in the expression or activity of these factors, or from other, as-yet-uncharacterized components of the transcriptional regulatory networks that govern temporal patterning and photoreceptor identity.

To investigate the molecular mechanisms underlying the development of cone-dominant mammalian retinas, we performed unbiased single-cell RNA- and ATAC-Seq analyses, to comprehensively profile gene expression and regulatory changes in the developing retina of the cone-dominant 13-lined ground squirrel (13LGS) across the full course of neurogenesis. We directly compared these data with similar datasets from the developing retina of rod-dominant mice. Our findings reveal that cone generation occurs at high levels even during late neurogenesis in 13LGS, coinciding with two broad heterochronic shifts in transcription factor expression in late-stage retinal progenitors. The first shift involves an expanded expression of a subset of transcription factors that are typically active in early-stage progenitors in rod-dominant retinas—factors known or predicted to promote cone specification. In addition to previously characterized factors such as *Onecut2* and *Pou2f1*, this group includes previously uncharacterized genes, such as *Zic3*, which we show to be both necessary and sufficient for cone specification during early neurogenesis. The second shift involves precocious expression of transcription factors typically found in postmitotic cone precursors, including *Thrb* and *Rxrg*, along with newly identified factors such as *Mef2c*, which cooperatively activate cone-specific genes while repressing rod-specific ones. The expression of these transcription factors in both late-stage progenitors and committed photoreceptor precursors is driven by multiple 13LGS-specific enhancer elements. These enhancers are directly targeted by *Otx2* and *Neurod1*, which broadly regulate photoreceptor specification and differentiation, as well as by heterochronically-expressed transcription factors specific to early-stage retinal progenitors and cone precursors. Our results demonstrate that heterochronic expansion of the expression of transcription factors that promote cone development is a key event in the evolution of the cone-dominant 13LGS retina.

Results:

Single-cell RNA- and ATAC-Seq profiling of developing 13LGS retina

In 13LGS, retinal neurogenesis begins around embryonic day 18 (E18), with birth occurring at E27 (blue-dotted line in Fig. 1A), and concludes by postnatal day 17 (P17)[27]. To comprehensively profile cellular-level changes in gene expression and chromatin accessibility, we conducted scRNA-Seq and scATAC-Seq analyses across ten developmental time points, spanning the full course of neurogenesis and the immediate post-neurogenesis period (E18–P21). We selected these time points to align with stages previously profiled in the developing mouse retina (E11–P14)) (Fig. 1A) [15,17]. High-quality data were obtained for all stages, with scATAC-Seq data showing expected fragment length and distribution patterns (Fig. S1A, B). Using selective markers of cell types previously identified in the developing mouse retina, we successfully annotated cell populations identified through UMAP clustering of aggregated scRNA-Seq and scATAC-Seq data. This approach distinguished all major retinal cell types, including both early- and late-stage primary and neurogenic retinal progenitor cells (RPCs and N. RPCs) (Fig. 1B).

Cell type-specific expression and chromatin accessibility patterns were highly correlated (Fig. S1C), with expected expression and accessibility patterns observed for known cell type-specific marker genes (Fig. 1C). Each annotated progenitor state and cell type (C1–C12)

exhibited significant numbers of differentially expressed genes (DEGs) (Fig. S1D, Table ST1) and differentially accessible chromatin regions (DARs) (Fig. S1E). Evolutionarily conserved patterns of motif accessibility (Fig. S1F, Table ST1) and transcription factor footprinting (Fig. S1G, H) were also observed. Notably, *Nr2f1* showed strong activity in early-stage RPCs, *Pou4f2* in retinal ganglion cells, *Nfix* in late-stage RPCs and Müller glia, and *Crx* and *Neurod1* in photoreceptor precursors (Fig. S1F, Table ST1), consistent with previous studies in developing mouse and human retina [17,28].

Direct comparisons of cell type composition across individual time points in 13LGS and mouse retina indicated a high degree of stage-matching and broad concurrence (Fig. S1I). The transition from early- to late-stage states occurred between E26 and P1 in 13LGS, compared to E16–E18 in mice [15,17]. The expected temporal patterns of neurogenesis were observed in both species: retinal ganglion cells and amacrine cells were generated predominantly by early-stage progenitors, whereas bipolar cells and Müller glia were produced by late-stage progenitors. Photoreceptor precursors were detected from E18 through P12 in 13LGS, with cone cells becoming clearly distinct in the early postnatal period. Notably, a large proportion of cones were generated during the first postnatal week in 13LGS. In contrast to mice, where rods are distinct from E18 onward [15], rods in 13LGS were not detected until the late postnatal period (P12). Both scRNA- and scATAC-Seq data demonstrated that, at all developmental stages, cones were significantly more abundant than rods in 13LGS, comprising approximately 25% of all retinal cells following the completion of neurogenesis at P12. In contrast, in mice, rods progressively increased in proportion, reaching approximately 70% of retinal cells by P8 (Fig. 1D).

Comparative multiomic analysis of photoreceptor differentiation in 13LGS and mouse.

These findings suggest that late-stage RPCs in 13LGS retina are competent to generate cone photoreceptors but not other early-born cell types, with rod specification occurring only at very late postnatal stages (>P12) of neurogenesis. To investigate the mechanisms underlying this heterochronic shift in developmental competence, we integrated scRNA- and scATAC-Seq data from all late-stage time points in 13LGS and mice (Fig. 2A, Table S2).

In the 13LGS scRNA-Seq data, we observed well-defined, continuous differentiation trajectories linking late-stage primary and neurogenic RPCs to cones, bipolar, and amacrine cells. However, no clear trajectory for rod differentiation was detected (Fig. 2A). In contrast, scRNA-Seq data from mice revealed differentiation trajectories connecting neurogenic progenitors to rods, amacrine, and bipolar cells, but not to cones at this stage (indicated by arrows in Fig. 2A). Similarly, scATAC-Seq analysis in 13LGS identified a strong progenitor-to-cone photoreceptor trajectory, but no corresponding progenitor-to-rod trajectory (Fig. 2B).

To further characterize species-specific patterns of gene expression and regulation during photoreceptor development, we analyzed differential gene expression, chromatin accessibility, and motif enrichment across late-stage primary and neurogenic progenitors, immature photoreceptor precursors, rods, and cones. This analysis revealed eight distinct species-specific differential expression patterns (C1–C8) (Fig. 2C, D; Table ST3).

Clusters C1–C3 were selectively expressed in 13LGS. C1 was active exclusively in RPCs, C2 active across all stages of cone specification, and C3 restricted to mature cones. Notably, C2 included several genes previously implicated in photoreceptor development, such as the cone-promoting transcription factors *Onecut1* [18,29] and *Mef2c*, a bHLH transcription factor which is cone-enriched in human retina [30,31] and essential for photoreceptor survival [32]. The smaller cluster C3 contained transcription factors such as *Thrb*, *Rxrg*, and *Sall3*, which are all selectively expressed in cone precursors and essential for cone-specific gene expression [26,33,34]. Additionally, *Zic3*, which is selectively expressed in early-stage RPCs in mice [15,17], was also enriched in C3.

Clusters C4–C8 contained genes that were selectively active in mice relative to 13LGS (Fig. 2C). Clusters C4–C6 showed higher expression levels in mice across different cellular contexts: C4 was specific to progenitors, C5 was active throughout specification, and C6 spanned progenitors and photoreceptor precursors. C4 included the early-stage RPC transcription factor *Nr2f1*, C5 contained *Sox8* and *Sox11*, and C6 featured *Dlx2*. While the direct relevance of most of these differences remains unclear, *Nr2f1* is essential for establishing the dorsal-ventral gradients in M- and S-cone opsin expression that are observed in mouse cones [35], but which do not exist in 13LGS [13].

Clusters C7 and C8, which were selectively active in mouse photoreceptor precursors and rods, contained several transcription factors known to promote rod differentiation and repress cone fate. In C7, we identified *Casz1*, a temporal patterning factor that represses the competence of late-stage progenitors to generate cones, while also promoting rod specification [36]. In C8, as expected, we observed *Nrl*, *Nr2e3*, and *Samd7*, all of which activate rod-specific genes while repressing cone-specific genes [20,22,37,38].

For each of these genes, integrated scRNA- and ATAC-Seq analysis identified corresponding regulatory elements in both 13LGS and mouse (Fig. 2D). Notably, genes in C2 and C3, which were more highly expressed in 13LGS and are strong candidates for promoting cone specification, exhibited significantly higher numbers of predicted functional cis-regulatory elements than their mouse counterparts (Fig. 2E). This species-specific increase in predicted regulatory elements was either absent from, or far less pronounced in, other gene clusters.

Identifying transcription factors that promote cone specification in 13LGS.

To identify transcription factors that strongly contribute to cone specification, we compared differential motif enrichment in accessible chromatin regions across cell types between 13LGS and mice (Fig. 2F). Motifs associated with neurogenic bHLH factors such as Neurog2, Olig2, Ascl1, and Neurod1 were highly enriched in neurogenic RPCs and photoreceptor precursors in both species, as was the motif for Otx2, a key factor required for both rod and cone specification [18,39]. Similarly, motifs for transcription factors strongly expressed in late-stage RPCs, such as Lhx2, Sox8, and Nfia/b, were prominent in progenitors of both species.

Motifs for cone-specific transcription factors, including Thrb, Rxrg, Esrrg, and Isl2, were significantly more prominent in cones of 13LGS. Furthermore, in 13LGS, the Zic3 motif was accessible in both progenitors and mature cones, whereas in mice, its accessibility declined in mature cones (Fig. 2F,G). These data suggest that species-specific gene regulatory networks (GRNs) observed in 13LGS late-stage RPCs are characterized by relatively higher activity levels of transcription factors that are enriched in both early-stage RPCs and postmitotic cone precursors in mice contribute to the development of the cone-dominant retina in 13-LGS.

To more systematically identify transcription factors regulating cone specification in 13-LGS, we used scRNA- and ATAC-Seq data from late-stage RPCs and differentiating photoreceptors to reconstruct GRNs predicted to activate cone-specific genes (Fig. 2H). This analysis focused on transcription factors with differential expression in late-stage RPCs between 13LGS and mice. Some transcription factors in this network, such as *Otx2*, *Neurod1*, *Crx*, *Prdm1*, and *Mef2d*, have been shown to promote both rod and cone differentiation [39–43], and their expression patterns did not exhibit strong species-specific differences.

In contrast, transcription factors identified in this network that are known to be required for committed cone precursor differentiation, including *Thrb*, *Rxrg*, *Sall3*, and *Isl2* [25,26,44], consistently showed stronger expression in late-stage RPCs of 13LGS compared to mice. Other genes such as *Mef2c*, which have not been functionally linked to cone specification, are also most prominently expressed in differentiated cones, also showed expression in late-stage RPCs.This suggests a species-specific pattern of precocious expression of transcription factors that promote cone differentiation. Additionally, transcription factors in the network known to promote cone specification in early-stage mouse RPCs, such as *Onecut2* and *Pou2f1*, exhibited enriched expression in late-stage RPCs of 13LGS, implying a heterochronic expansion of cone-promoting factors into later developmental stages. This pattern was also observed for *Zic3*, which emerged as the most likely candidate for driving cone-specific gene expression in 13-LGS. In mice, *Zic3* was selectively enriched in early-stage RPCs [15,17], supporting its potential role in cone specification.

Histological analysis validates cone-enriched expression of candidate transcription factors in 13LGS

To prioritize novel candidate cone-promoting factors for further analysis, we performed immunohistochemistry (IHC) and RNA hybridization chain reaction (HCR) analysis to confirm their expression in late-stage neurogenic RPCs and/or differentiating cones in 13LGS. We analyzed co-expression of the candidate cone-promoting factors *Zic3* and *Mef2c*, as well as *Pou2f1*, a cone-promoting factor active in early-stage RPCs in mouse [18,19]. At P1, P5, P10, and P24, *Zic3* was co-expressed with the neurogenic RPC/photoreceptor precursor marker *Otx2* and the cone precursor marker *Rxrg* in 13LGS (Fig. 3A). *Pou2f1* co-expression with *Otx2* was observed at all ages examined (Fig. 3B). Weak Mef2c protein expression was observed in Otx2-positive neurogenic RPCs at P1 (data not shown), with increased expression at P5, with stronger expression detected in mature photoreceptors at later developmental stages (Fig. S2A).

Zic3 Promotes Cone Photoreceptor Specification and Differentiation

Previous studies have demonstrated that overexpression of either *Onecut1* or *Pou2f1* in late-stage mouse retinal progenitors is sufficient to induce the expression of cone-specific genes while also repressing rod-specific genes [18,19]. We employed a similar approach to determine whether *Zic3*, either alone or in combination with the heterochronically expressed cone-promoting factors *Pou2f1* and *Onecut1*, can likewise induce cone differentiation from late-stage mouse RPCs. To test this, we conducted electroporation of P0 *ex vivo* mouse retinal explants, as previously described [45], and used scRNA-Seq to assess the effects of *Zic3* overexpression alone, or in combination with either *Onecut1* or *Pou2f1* (Fig. 4A). ScRNA-Seq analysis of GFP-positive electroporated cells at five days *in vitro* showed that overexpression of heterochronically expressed transcription factors produced a cone-like photoreceptor precursor population (Fig. S3A, Table ST5). Comparison of gene expression between this photoreceptor cluster and rod photoreceptors revealed that while *Zic3* induced the expression of several cone-specific genes at this stage, it also downregulated the expression of many rod-specific genes (Fig. 4B,C). *Pou2f1* overexpression upregulated an overlapping but distinct set of cone-specific genes relative to *Zic3*, while also downregulating many of the same rod-specific genes, often to a greater extent (Fig. 4C). Co-overexpression of *Zic3* and *Pou2f1* in late-stage RPCs had a synergistic effect, resulting in more differentially expressed genes than did either *Zic3* or *Pou2f1* alone, and greater quantitative changes in cellular levels of expression (Fig. 4B,C, Table ST5).

Onecut1 overexpression similarly produced a population of cone-like precursors (Fig S3B, Table ST5). It upregulated a largely distinct set of cone-specific genes compared to *Pou2f1* alone, but shared some upregulated cone-specific genes with the combined overexpression of *Zic3* and *Pou2f1* (Fig. S3C,D, Table ST5). Furthermore, *Onecut1* overexpression downregulated many of the same rod-specific genes as did the combined overexpression of *Pou2f1* and *Zic3* (Fig. S3C,D). Similarly to *Onecut1* and *Pou2f1* overexpression, *Zic3* overexpression produced horizontal cell-like precursors in addition to cone-like precursors [18,19] (Fig. S3A,B, Table ST5)

Because conventional *Zic3* knockouts show widespread and complex developmental defects and high gestational lethality [46], we generated an early RPC-specific *Zic3* loss-of-function (LOF) mutant by combining the retinal progenitor-specific Chx10-CreGFP transgenic line with a conditional *Zic3* allele [47,48] (Fig. 4D). This resulted in a significant reduction in the number of cone photoreceptors in the mutant retina (Fig. 4E,F), while the relative numbers of rods and horizontal cells remained unaffected (Fig. S4A-D).

ScRNA-Seq analysis (n=4) of P14 *Zic3* conditional mutant retinas identified quantitative changes in gene expression in several terminally differentiated cell types (Fig. S4E-G, Table ST6). Mutant cones showed reduced expression of *Thrb* but increased expression of multiple cone-specific markers, including *Opn1sw*, *Pde6h*, and *Gngt2* (Fig. S4F). Additionally, increased expression of rod-specific genes (*Rho*, *Nrl*, *Pde6g*, *Gngt1*) and pan-photoreceptor genes (*Crx*, *Stx3*, *Rcvrn*) was observed in Müller glia (Fig. S4G). Furthermore, several Müller glia-specific genes were downregulated, including *Clu*, *Aqp4*, and Notch pathway components such as *Hes1* and *Id3*, with the exception of *Hopx*, which was upregulated (Fig. S4G). These findings indicate

that *Zic3* is essential for the generation of the full physiological complement of cones while also playing a role in repressing the expression of rod photoreceptor-specific genes.

***MEF2C* overexpression promotes cone-specific gene expression and inhibits rod photoreceptor maturation.**

We next investigated the effects of both *Mef2c* overexpression in late-stage RPCs and RPC-specific loss of function of *Mef2c* in mice. As described for *Zic3*, we conducted *ex vivo* electroporation of P0 mouse retinal explants with plasmid vectors expressing *Mef2c*, using empty vectors that expressed GFP alone as controls (Fig. 5A). Since *Mef2c* expression is initiated at later stages of cone development relative to *Zic3* (Fig. 3A, S2A), explants were cultured until P8 for both scRNA-Seq and immunohistochemical analysis.

ScRNA-Seq analysis revealed substantial alterations in cell type composition following *MEF2C* overexpression, with a marked increase in both neurogenic RPCs and immature photoreceptor precursors, which suggested a broad delay or arrest in rod differentiation (Fig. 5B,C Table ST7). Additionally, a novel subpopulation of cone-like precursors emerged, accompanied by an overall increase in the number of mature cones (Fig. 5B). Cone-like precursors exhibited upregulation of genes such as *Gnat2*, *Arr3*, and *Pde6c*, while also maintaining expression of genes that are typically enriched in both immature rod and cone photoreceptors, including *Tpi1*, *Rtn4*, and *Tpt1* (Fig. 5C).

Histological analysis indicated that *MEF2C* overexpression consistently led to morphological abnormalities in electroporated regions, including the formation of whorls and rosettes in the outer nuclear layer, a phenotype reminiscent of *Nrl* and *Nr2e3* mutants that exhibit ectopic cone-specific gene expression in rods (Fig. 5D, Fig. S5A,B) [23,49]. A significant increase in the fraction of GFP-positive electroporated cells co-expressing the cone-specific marker Gnat2 was detected (Fig. 5E), though no obvious reduction in the relative number of Nrl-positive cells was observed (Fig. S5A), consistent with scRNA-Seq findings. No changes were observed in the distribution or abundance of other retinal cell types (Fig. S5B-D).

A previous study reported that *Nestin-Cre;Mef2c*$^{lox/lox}$ mice, which are predicted to exhibit loss of function of *Mef2c* throughout the developing central nervous system, showed a rapid degeneration of both rod and cone photoreceptors, that resembled photoreceptor loss observed in fast-acting *rd1/rd1* mutants [32]. To evaluate the effects of retinal progenitor-specific loss of function of *Mef2c*, we generated *Chx10-Cre;Mef2c*$^{lox/lox}$ mice [47,50] (Fig. S6A). No changes in cone numbers were observed at P14 (Fig. S6B,C), nor was there any loss of rod or cone photoreceptors through P180 (data not shown). These findings suggest that *Mef2c* is dispensable for cone development in mice, and that previous reports of rapid photoreceptor degeneration may have resulted from inadvertent contamination with either the *rd1* allele itself or another common retinal dystrophy mutation [51].

Species-specific *cis*-regulatory elements activate expression of cone-promoting genes in late-stage RPCs of 13LGS.

These findings demonstrate that, relative to mice, late-stage RPCs and photoreceptor precursors of 13LGS exhibit significantly elevated expression of multiple transcription factors that promote cone differentiation while inhibiting rod development. To investigate the mechanism underlying this species-specific expansion of expression, we performed bulk CUT&RUN analysis in P5 13LGS retinas and P2 mouse retinas, analyzing stage-matched samples highly enriched for late-stage RPCs and photoreceptor precursors. We profiled histone modifications associated with active promoters and both active and poised enhancers, as well as the photoreceptor-promoting transcription factors Otx2 and Neurod1. This data was integrated with scATAC-Seq-derived chromatin accessibility regions from late-stage RPCs and photoreceptor precursors (Fig. 6A, Table ST8).

We examined the regulatory elements of genes that exhibited specifically elevated expression in late-stage progenitors and photoreceptor precursors of 13LGS, which were identified in clusters C2 and C3 of the analysis described in Figure 2D. Integration of scATAC-Seq and CUT&RUN data showed that 13LGS consistently harbors both greater numbers of accessible chromatin regions associated with markers of both active (H3K27Ac-positive) enhancers and poised (H3K4me1-positive) enhancers. However, most of these elements were not evolutionarily conserved, with intergenic enhancers exhibiting particularly low levels of conservation (Fig. 6B). In contrast, H3K4me3-positive promoter elements were relatively well-conserved and were found in similar numbers in both species (Fig. 6B).

Though conserved and species-specific enhancers showed enrichment for binding sites for the broad photoreceptor precursor transcription factors Otx2 and Neurod1, transcription factors known to regulate cone differentiation— including *Mef2c*, *Rxrg*, *Thrb* and *Zic3* — demonstrated far more active regulatory elements in 13LGS than in mice, though few of these elements exhibited evolutionary conservation (Fig. 6C,D, Table ST9).

Enhancer elements in 13LGS are predicted to be directly targeted by a considerably greater number of transcription factors than in mice. This is particularly evident in enhancers associated with *Thrb* in both species, which are shown to be directly targeted by Otx2 (Fig. 6E,F). While the *Thrb* promoter is both accessible and bound by Otx2 in mouse photoreceptor precursors, it is predicted to be targeted by significantly fewer transcription factors compared to the active elements in 13LGS (Fig. 6E,F). In 13LGS, we identified a pair of species-specific enhancers that are active in late-stage neurogenic RPCs, photoreceptor precursors, and mature cones. These enhancers are predicted to be directly targeted by a combination of factors, including those broadly expressed in developing rods and cones (Otx2, Crx, Neurod1, Prdm1); cone-promoting factors that are restricted to early-stage RPCs in mice (Zic3, Onecut2, Pou2f1); and by transcription factors that selectively promote cone precursor differentiation (Rxrg, Mef2c). The gene regulatory networks (GRNs) driving cone specification in 13LGS become progressively more stable over time, as additional transcription factors are recruited and multiple layers of positive regulation are incorporated (Fig. S7A-D, S8A-C, S9A-C).

We conclude that the development of the cone-dominant retina in 13LGS is driven by the evolution of novel cis-regulatory elements that promote heterochronic expression of cone-

promoting and rod-inhibiting transcription factors that are restricted to either early-stage RPCs or postmitotic cone photoreceptor precursors in mice. These *cis*-regulatory elements are activated by a combination of these factors themselves, in conjunction with transcription factors such as Otx2 and Neurod1, that are expressed in both developing rods and cones. The sustained expression of these factors is maintained by an intricate and progressively expanding network of cross-activating regulatory interactions (Fig. 6G).

Discussion

This study represents the first comprehensive molecular analysis of gene regulatory networks that controls the development of a cone-dominant retina derived from a rod-dominant ancestral state [5,6]. Using integrated scRNA- and scATAC-Seq analyses, we systematically examined developmental changes in gene expression and regulation in the cone-dominant 13-lined ground squirrel (13LGS) retina, and directly compared these findings to stage-matched data from the rod-dominant mouse. Our findings indicate that 13LGS exhibit a selective expansion of the temporal window of cone generation, extending into late stages of neurogenesis. This is mediated by the heterochronic expression of a subset of transcription factors in late-stage progenitors that both promote cone and repress rod specification. In addition to previously characterized cone-promoting and rod-inhibiting factors such as *Onecut2* and *Pou2f1*, we identified *Zic3*, which is selectively expressed in early-stage RPCs in mice, as a key driver of cone specification. Late-stage 13LGS RPCs also exhibit precocious expression of transcription factors such as *Thrb*, *Rxrg*, and *Sall3*, which in mice function post-mitotically to promote cone differentiation, and in many cases also act in parallel to inhibit rod differentiation [25,26].

Transcription of these genes in late-stage RPCs is controlled by 13LGS-specific enhancer elements, which are directly targeted by transcription factors that are broadly expressed in both early- and late-stage RPCs, such as Otx2, as well as by heterochronically and precociously expressed cone-promoting transcription factors through a dynamic process of positive feedback and cross-activation. The gene regulatory network (GRN) governing cone specification becomes progressively more complex, incorporating additional transcription factors over time. These factors include genes expressed in both rod and cone precursors, such as Neurod1 and Prdm1; genes known to promote cone precursor differentiation, such as Sall3; and previously uncharacterized regulators such as Mef2c, which we show to both activate cone-specific genes and repress rod-specific genes.

The evolution of novel morphological features, such as bat wings or the elongated legs of jerboas [52,53], is often driven by the emergence of lineage-specific cis-regulatory elements, which confer novel spatiotemporal patterns of expression to structurally conserved transcription factors or secreted patterning regulators. The evolution of the cone-dominant 13LGS retina follows this paradigm, in which species-specific enhancer elements are used to drive the ectopic heterochronic expression of transcription factors that regulate neural progenitor temporal identity. This in turn promotes cone, and represses rod, photoreceptor specification. In all vertebrates examined to date, including 13LGS, separate and mutually cross-repressive GRNs specify early- and late-stage temporal identity in retinal progenitors [15–17]. In mice, cone

generation occurs only in progenitors in which the early-stage network is active [18,19,54], whereas in 13LGS, multiple transcription factors from the early-stage network persist in late-stage RPCs.

This raises the question of whether similar heterochronic shifts in key temporal identity regulators may underlie other cases of rapid evolutionary changes in the rod-cone ratio. Examples in mammals include the development of the cone-dominant retina in primate-like tree shrews and the formation of the rod-free foveola in the primate fovea [55,56] . However, rapid changes in the rod-cone ratio are observed across vertebrate classes [8,57]. Significant differences in the relative abundance of other retinal cell types are also observed among species, often correlated with temporal niche adaptation [58]. Moreover, shifts in the relative ratios of neuronal subtypes in the central nervous system can underlie species-specific behavioral, cognitive, and sensory processes between closely-related vertebrate species [59–64]. These variations may also stem from species-specific heterochronic expression of transcription factors that control temporal identity in the developing brain.

By performing an unbiased and comprehensive analysis of gene regulatory networks that are selectively active in late-stage RPCs of 13LGS, we have identified numerous previously uncharacterized factors that are strong candidates for promoting cone specification. Working in mice, we demonstrate that *Zic3* and *Mef2c* are necessary and/or sufficient for both promoting cone and repressing rod photoreceptor differentiation, though many additional candidates remain to be tested. Our analysis suggests that gene regulatory networks controlling cone specification are highly redundant, with transcription factors acting in complex, redundant, and potentially synergistic combinations. This is further supported by our findings on the combined overexpression effects of *Zic3* and *Pou2f1* (Fig. 4) and the relatively mild or undetectable phenotypes following loss of function of *Zic3* and *Mef2c* (Fig. 4, Fig. S6), as well as other cone-promoting factors such as *Onecut1* and *Pou2f1*[18,19].

These findings have potentially significant clinical implications, as the loss of cone photoreceptors is the primary trigger for irreversible blindness in hereditary diseases such as retinitis pigmentosa and age-related macular degeneration. Effective strategies for driving cone specification and differentiation are essential for developing cell-based therapies to replace photoreceptors lost to degenerative disease, and the findings of comparative studies like ours can help identify these mechanisms.

Methods:

Mouse retinal development scRNA-Seq and scATAC-Seq analysis
Raw data and metadata for the mouse scRNA-Seq and scATAC-Seq datasets [15,17] were downloaded from the Gene Expression Omnibus (accession numbers GSE118614, GSE181251) and analyzed following the Seurat SCT and Signac workflows [65–67]. Cell type classifications and ages were taken from the downloaded metadata.

Sample collection
The use of animals for these studies was conducted using protocols approved by the Johns Hopkins Animal Care and Use Committee, the UW Oshkosh Animal Care and Use Committee, and the Northwestern University Animal Care and Use Committee, in compliance with ARRIVE guidelines and the ARVO Statement for the Use of Animals in Ophthalmic and Vision Research, and were performed in accordance with relevant guidelines and regulations.

Timed pregnant CD1 mice were ordered from Charles River Laboratories for the postnatal retinal explant plasmid overexpression analysis. A pair of *Zic3*$^{+/flox}$ [68](JAX stock #023162) female mice and a pair of *Mef2c*$^{+/flox}$ mice [50](JAX stock #025556) were ordered from the Jackson Laboratory and outbred to *Chx10Cre-GFP* CD1 mice [47]. The progeny were bred to produce the *Zic3*$^{lox/lox}$ (or *Zic3*$^{lox/Y}$); *Chx10Cre-GFP*$^{+/-}$, *Zic3*$^{+/+}$ (or *Zic3*$^{+/Y}$); *Chx10Cre-GFP*$^{+/wt}$, *Mef2c*$^{lox/lox}$; *Chx10Cre-GFP*$^{+wt}$, and *Mef2c*$^{lox/lox}$ used in this study. Pups were grown until postnatal day 14 before retinal samples were taken. Mice were euthanized by cervical dislocation.

Embryonic 13-lined ground squirrel (13LGS) retinal samples were collected from timed-pregnant captive dams and embryonic day was confirmed by comparison with mouse Theiler stages. Postnatal 13LGS retinal samples were collected from captive-bred litters with day of birth set as Postnatal Day 0. 13LGS were euthanized by decapitation.

Sample preservation
Mouse retinas and retinal explants used for immunostaining were fixed in freshly prepared 4% paraformaldehyde in PBS at 4°C for 1 hour (ZIC3) or 2 hours (MEF2C), then briefly washed with PBS for 5 mins. After the wash, the explants or retinas were transferred to 30% sucrose in PBS at 4 °C overnight. The dehydrated samples were then embedded in VWR Clear Frozen Section Compound, frozen on dry ice, then stored at -80°C.

For single-cell analysis, 13LGS retinas were dissected and either methanol-fixed or flash-frozen as previously shown [69]. Samples were shipped to Johns Hopkins on dry ice and stored at -80°C.

For IHC and RNA-HCR experiments, postnatal 13LGS whole eye cups were fixed in 4% paraformaldehyde (PFA; 157-8, Electron Microscopy Sciences, USA) made in 1X phosphate buffered saline (PBS) after the removal of lens, cornea and vitreous humor and processed as previously [70,71]. PFA-fixed eye cups were washed three times with 1X PBS at 15 min intervals each, followed by rehydration with 15% and 30% sucrose (S1888-5KG, Sigma Life Science,

USA) solutions (made in 1X PBS) until the eyecups had sunk to the bottom of the tube. The whole cups were then equilibrated, embedded and frozen on dry ice in a sucrose:OCT (4583, Sakura Finetek, Torrance, CA) mixture (2:1) and stored at -80ºC.

For CUT&RUN experiments, retinas were dissected out from postnatal 13LGS after the removal of lens, cornea and vitreous humor. Dissected retinas were immediately flash frozen using liquid nitrogen and shipped to Johns Hopkins University using dry ice and stored at -80°C.

Retinal explant culture

Retinas were dissected from CD1 mice at Postnatal Day 0 in warm PBS, then transferred to a micro electroporation chamber. The retinas were electroporated with plasmid solution at 30 V of 50-ms duration with 950-ms intervals [72] using either a pCAGIG plasmid that expresses IRES-GFP [73,74] or a test plasmid that expresses a gene of interest as a bicistronic transcript together with IRES-GFP [73], then flattened by a series of radial cuts and mounted on 0.2 µm Nucleopore Track-Etch membranes on culture media (2 ml DMEM F12, 10% FBS with 0.1% streptomycin/puromycin) in 12-well plates [75]. In experiments that included a condition where two genes of interest were overexpressed in the same retinal explants, the concentration of plasmids was kept consistent across conditions by supplementing with empty plasmids. For instance, empty condition: 100% control plasmid, condition A: 50% gene A plasmid and 50% control plasmid, condition B: 50% gene A plasmid and 50% control plasmid, combined condition: 50% gene A plasmid and 50% gene B plasmid.

Explants were incubated at 37 °C, 5% CO_2 with the media changed every 72 hours. Two-thirds of the old media (1/2 old media in *MEF2C* overexpression) in each well was replaced by slowly adding new media to the well's wall, avoiding disturbing the membranes. The explants are harvested at P5 (*ZIC3*) or P8 (*MEF2C*) for single-cell RNA sequencing and P8 (*MEF2C*) for immunostaining.

Retinal sectioning

Mouse embedded samples were sectioned with Leica 3050S cryostat at a thickness of 16 µm and collected on clean Superfrost Plus slides (48311-703, VWR, Radnor, PA). Sections were then stored at -20°C. For postnatal 13LGS, 8-10µm thick sections were collected and stored at -80ºC until use for histological assessments.

Fluorescence-Activated Cell Sorting and multiplexing

Dissociated single cells from electroporated retinal explants were sorted based on GFP expression on a Sony SH800S cell sorter. GFP+ cells from each condition were labeled and pooled into a single cell suspension according to the 10x Cell Multiplexing Oligo Labeling for Single Cell RNA Sequencing Protocols with Feature Barcode technology protocol (Rev B) prior to GEM generation and library prep.

Single-cell library preparation

Retina and explants were dissociated, methanol-fixed cells rehydrated, and nuclei isolated from flash-frozen samples as previously described [69]. GEM generation and library preparation for

scRNA-Seq were completed according to the 10x 3' v3.1 user guide (Rev D) using the Chromium Next GEM Single Cell 3' Kit v3.1 and Dual Index Kit TT Set A and for Cellplex samples according to the 10x Next GEM Single Cell 3' Reagent Kits v3.1 (Dual Index) with Feature Barcode technology for Cell Multiplexing (Rev A) using the 3' CellPlex Kit Set A and Dual Index Kit NN Set A. Transposition, GEM generation, and library preparation for scATAC-Seq were completed according to the 10x ATAC v1.1 user guide (Rev D) using the Chromium Next GEM Single Cell ATAC Reagent Kits v1.1 and Single Index Kit N, Set A. Quality control and sequencing were performed by the Johns Hopkins Transcriptomics and Deep Sequencing Core.

ScRNA-Seq and scATAC-Seq analysis

Sequencing results for 13LGS and mouse scRNA-Seq were processed through the 10x CellRanger pipeline (version 3.1.0), using the 13LGS genome GCF_016881025.1_HiC_Itri_2 for 13LGS [76], and the output loaded into Seurat (version 4.0.4). Cells from Cellplex-labeled libraries were deconvoluted using the cellranger multi pipeline (version 6.0). ScRNA-Seq Seurat objects were analyzed following the Seurat SCT workflow [65,66]. Cell type was determined cluster by cluster based on marker gene expression derived from the Seurat function 'FindAllMarkers' using default parameters. Differential gene expression between Cone-like Precursors and Rods was identified using the Seurat function "FindMarkers". ScRNA-Seq samples were aggregated using the CellRanger aggr pipeline (version 3.1.0).

Sequencing results for 13LGS scATAC-Seq data were processed through the 10x CellRanger ATAC pipeline (version 1.2.0) and the output loaded into Signac (version 1.4.0) and analyzed according to the Signac workflow. Cell type was determined cluster by cluster based on marker gene expression derived from the Seurat function 'FindAllMarkers' using default parameters and based on transfer learning from the scRNA-Seq samples following the Signac workflow using default parameters. ScATAC-Seq samples were integrated as described [17].

Immunostaining and visualization

Mouse retinal sections were stained with rabbit anti-Arr3 polyclonal IgG (Millipore, No. AB15282, 1:200) or rabbit anti-Gnat2 polyclonal IgG (Invitrogen, PA5-119558, 1:200); chicken anti-GFP polyclonal IgY (Invitrogen, A10262, 1:400); and mouse anti-Nr2e3 monoclonal IgG (R&D systems, PP-H7223-00, 1:200), goat anti-Nrl polyclonal IgG (R&D Systems, AF2945, 1:200), goat anti-Otx2 polyclonal IgG (R&D Systems, AF1979, 1:400), rabbit anti-Sox9 polyclonal (Millipore, AB5535, 1:400), mouse anti-Tfap2a monoclonal (DSHB, Cat# 3b5, RRID:AB_528084, 1:400), or mouse anti-Lhx1 (DHSB,Cat#4F2, concentrated,) primary antibodies. For the secondary staining, CF633 donkey anti-rabbit IgG polyclonal (Sigma,1:500); Alexa Fluor™488 goat anti-chicken IgY polyclonal (Invitrogen,1:500), and Alexa Fluor™ TM 568 donkey anti-mouse IgG polyclonal (Invitrogen,1:500) secondary antibodies were used. DAPI was used to counterstain nuclei (1:5000). Stained sections were imaged with a Zeiss AxioObserver with LSM700 confocal module at 20X resolution within a fixed scanning area of 319.87 nm^2. Z-stack images were taken for the stained sections and cell counting was performed blinded on Z-projections.

For 13LGS immunohistochemistry experiments, retinal sections were thawed, rehydrated three times with 1X PBS at 15 min intervals and stained *via* IHC for identification of retinal cell types. For IHC, sections were permeabilized for 15 mins at room temperature with 0.1% Triton X-100 (0694-1L, VWR Life Science, USA) made in 10% normal donkey serum in 1X PBS (NDS, S30-100ML, EMD Millipore Corp, USA). The sections were then blocked for 1 hour at room temperature using 10% normal donkey serum and incubated at 4°C overnight with primary antibodies diluted in 10% normal donkey serum. Primary antibodies utilized for staining 13LGS sections included goat anti-Otx2 polyclonal IgG (R&D Systems, BAF1979, 1:400) and rabbit anti-Mef2c polyclonal (Atlas Antibodies, HPA005533, 1:50). Next day, primary antibodies were removed, sections were washed three times in 1X PBS at room temperature with 15 min interval each and fluorescent conjugated secondary antibodies (donkey anti-mouse IgG Alexa Fluor™ Plus 647, ThermoFisher Scientific, A-11055, 1:250; donkey anti-rabbit IgG Alexa Fluor™ 488, ThermoFisher Scientific, A-21206, 1:250, and donkey anti-goat IgG Alexa Fluor™ 555, ThermoFisher Scientific, A-21432, 1:250) diluted in 10% normal donkey serum were applied to sections and incubated in the dark for 1 hour. Cell nuclei were counterstained with DAPI (1 µg/mL, Roche, 10236276001) for 10 min. The slides were then washed three times in 1X PBS at 5 min intervals and coverslipped using Fluoromount G (17984-25, Electron Microscopy Sciences, USA). Images were acquired using ZEN software on LSM700 confocal microscope (Zeiss, Inc) at 63X oil resolution and processed using ImageJ [77].

RNA-Hybridization Chain Reaction (RNA-HCR)

Prepare the samples by thawing and dehydrating the 13LGS retinal sections by immersing them into a series of ethanol (EtOH) solution - 50% EtOH, 70% EtOH, 100% EtOH for 5 mins each. Immerse the slide into another fresh 100% EtOH for 5 mins and allow the slides to air dry for 5 mins at room temperature. Apply 10 ug/uL of Proteinase K solution (3001-2-60, Zymo Research) on the sections and incubate the slides in a humidified chamber for 10 mins at 37°C. Wash the slides by immersing in fresh 1X PBS twice and proceed to the detection stage. Dry the slides by blotting the edges on a kimwipe. Pre-hybrize the sections in a humidified chamber pre-warmed to 37°C, 10 mins by adding 200 uL of probe hybridization buffer (240904, HCR™ Buffer sets, Molecular Instruments). Buffer was removed and 50 - 100 uL of DNA probe solutions (1.6 pmol of each probe diluted in probe hybridization buffer) were added to the sections. DNA probes were custom designed using in situ probe generator v0.3.2 at the visual studio code platform for the 13-LGS gene targets as shown in Figure 3 (60 oligos/target at 50 pmol per oligos), obtained from IDT, and reconstituted as 1 uM stock in nuclease free water. Sections with probe solution were covered with coverslip and incubated overnight (~18-24 h) at 37°C in a humidified chamber. Next day, immerse the slide in the probe wash buffer at 37°C to float off the coverslip. Excess probes were removed by incubating the slides at 37°C with pre-heated probe wash buffer (240904, HCR™ Buffer sets, Molecular Instruments) and 5X SSCT solutions (S6639-1L, Sigma-Aldrich): 75% probe wash buffer/25% 5X SSCT for 15 mins; 50% probe wash buffer/50% 5X SSCT for 15 mins; 25% probe wash buffer/75% 5X SSCT for 15 mins and 100% 5X SSCT for 15 mins. Immerse the slide in 5X SSCT for 5 mins at room temperature, and then proceed to the amplification stage. Dry the slides by blotting the edges on a kimwipe. Pre-amplify the HCR by adding 200 uL of amplification buffer (240904, HCR™ Buffer sets, Molecular Instruments) on top of the sections and incubating the slides for 30 mins

at room temperature. Remove the solution and add 50 - 100 uL of hairpin solution and incubate overnight (12-18 h) in a dark humidified chamber at room temperature. Hairpin solution (Fluorophore labeled for 488/546/647) was prepared by heating (95°C for 90 seconds) and snap-cooling (in a dark drawer at room temperature for 30 mins) 2 uL of 3 uM stocks of h1 and h2 hairpins (6 pmol each) to 100 uL of amplification buffer at room temperature. Next day, immerse the slide in a 5X SSCT buffer at room temperature to float off the coverslip. Counterstain with DAPI (1 µg/mL, Roche, 10236276001) for 10 mins, at dark in room temperature. Excessive hair pins were removed by incubating the slides in a 5X SSCT buffer at room temperature thrice: 2X for 30 mins and 1X for 5 mins. The slides were coverslipped using Fluoromount G (17984-25, Electron Microscopy Sciences, USA). Images were acquired using ZEN software on LSM700 confocal microscope (Zeiss, Inc) at 63X oil resolution and processed using ImageJ (NIH, USA).

CUT&RUN library preparation

Nuclei were isolated from snap-frozen P5 13LGS and P2 mouse retinas for use in CUT&RUN analysis. Two biological replicates were assayed for each species, with each 13LGS replicate consisting of a pool of retinas from two individuals, and each mouse replicate consisting of a pool of retinas from 3 individuals. Nuclei were isolated according to the described protocol [78], using a lysis buffer modified to contain 0.01% Tween-20 and 0.01% Nonidet P40, as well as 1X protease inhibitor cocktail (Roche, 11873580001) in all solutions. Samples were lysed for 5 minutes in the lysis buffer with gentle trituration throughout, as described. CUT&RUN was performed using the CUTANA™ ChIC/CUT&RUN Kit (Epicypher, 14-1048). For each reaction, either 500,000 13LGS nuclei or 385,000 mouse nuclei were used. The CUTANA protocol v4.0 protocol was followed with minor modifications. In brief, isolated nuclei were resuspended in CUT&RUN wash buffer and added directly to Concanavalin A beads. Then 0.5 µg of the following antibodies were added to each respective reaction: IgG control (EpiCypher, 13-0042), H3K4me1 (EpiCypher, 13-0057), H3K4me3 (EpiCypher, 13-0041), H3K27ac (EpiCypher, 13-0059), H3K27me3 (EpiCypher, 13-0055), NeuroD1 (Cell Signaling, 62953), Otx2 (R&D Systems, BAF1979), or Otx2 (Atlas Antibodies, HPA000633). H3K9me3 (Abcam, ab8898) was used at 0.41 µg of antibody per reaction. The total CUT&RUN DNA yield, up to 5 ng of total purified DNA, was used for library preparation, using reagents from CUTANA™ CUT&RUN Library Prep Kit (EpiCypher, 14-1001) or NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (NEB, E7645S). The EpiCypher CUTANA Library Prep Kit User Manual v1.3 was followed and indexing performed using NEBNext® Multiplex Oligos for Illumina® (NEB, E7395S) or Dual Index Kit TT Set A (10X Genomics, PN-1000215). Final libraries were sequenced at the Novogene sequencing core (Sacramento, CA, USA) on the NovaSeq X Plus (Illumina) to generate paired-end 150 base pair reads.

CUT&RUN analysis

Sequence reads were trimmed of low quality bases and adapters using cutadapt v3.5 [79] and Trim Galore v0.6.4_dev [80], using arguments *---paired --length 20 -e 0.1 --illumina -q 20 --stringency 3*. Trimmed reads were aligned to mouse genome GRCm38 and 13LGS genome GCF_016881025.1_HiC_Itri_2 using Bowtie2 [81] with arguments *-X 1000 --very-sensitive-local --no-mixed --no-discordant --no-dovetail*. Unmapped reads with a MAPQ score of 30 or below

were filtered from the analysis with SAMtools v1.12 [82]. Aligned reads were deduplicated using UMI-tools [83] dedup with parameters -chimeric-pairs="discard" --unpaired-reads="discard" --paired --ignore-umi. Technical replicates were collapsed into single biological replicates using samtools merge. For visualization, normalization was applied using deepTools v3.5.1 [84] bamCoverage using arguments *--binSize 1 --normalizeUsing RPGC --effectiveGenomeSize 2311060300* (13LGS) or *--effectiveGenomeSize 2651675564* (mouse). For each protein target, peaks were called jointly across replicates against IgG control using macs2 [85] using arguments *-f BAMPE --keep-dup all -q 0.05 -g 2311060300* (13LGS) or *-g 2651675564* (mouse). For mouse samples, peaks on mitochondrial reads and blacklisted ENCODE regions [86] were excluded from analysis.

Integration of 13LGS and mouse late-stage single-cell datasets

To compare cone development between 13LGS and mouse during the late developmental stages, scRNA-Seq datasets spanning different periods (E21–P8 for 13LGS and E18–P5 for mouse) were integrated. First, the raw gene expression matrices from both species were filtered to retain only conserved genes. Next, 13LGS gene identifiers were mapped to their corresponding mouse gene identifiers as shown in Table S2, and when multiple 13LGS genes corresponded to a single mouse gene, their expression values were summed. The resulting 13LGS and mouse cell-by-gene matrices were directly merged, and the combined raw count matrix was normalized using SCTransform (SCT) before being integrated with Harmony in Seurat. Dimensionality reduction and clustering analyses were then performed on the Harmony dimensions to identify combined cell types (RPCs, N. RPCs, BC/Photoreceptor Precursors, Cones, and Rods). Finally, cell annotations from the scRNA-Seq data were transferred to the scATAC-Seq dataset using the addGeneIntegrationMatrix function in ArchR (version) [87]. Based on the scATAC-Seq clusters and transferred annotations, the final cell types were manually annotated. For scATAC-Seq datasets, peaks were re-called based on the combined cell types using the addReproduciblePeakSet function in ArchR.

Identification of differential photoreceptor development genes and their regulatory elements between 13LGS and mouse

To identify differentially expressed genes (DEGs) involved in photoreceptor development between the 13LGS and mouse datasets, we utilized the integrated cell-by-gene matrix from the previous integration step. Differential comparisons were performed using the “FindMarkers” function from the Seurat package with the criteria of p-value < 1e-6 and average $\log_2$ fold change > 0.35. Specifically, the following comparisons were made: 13LGS_RPCs vs. Mouse_RPCs, 13LGS_N. RPCs vs. Mouse_N. RPCs, 13LGS_BC/Photoreceptor Precursors vs. Mouse_BC/Photoreceptor Precursors, 13LGS_Cone vs. Mouse_Cone, and 13LGS_Cone vs. Mouse_Rod. All DEGs from these comparisons were then merged and clustered according to their expression levels across RPCs, N. RPCs, BC/Photoreceptor precursors, cones, and rods in both species using a k-means algorithm (k = 8). Genes in clusters 2 and 3 were subsequently identified as 13LGS-specific cone-promoting genes.

To identify regulatory elements associated with differentially expressed genes (DEGs), scRNA-Seq and scATAC-Seq datasets were first integrated separately for 13LGS and mouse. Next,

using the integrated multi-omics data, cells from RPC, NG, BC/Photoreceptor Precursor, Cone, and rod populations were analyzed to calculate Peak-to-Gene (PtoG) correlations via the "addPeak2GeneLinks" function in ArchR. For each gene, peaks located within 500 kb that showed a PtoG correlation greater than 0.25 and an FDR below 0.01 were defined as its regulatory elements. Subsequently, the number of regulatory elements per cluster was compared between 13LGS and mouse at the level of conserved gene pairs. For each species, the gene-associated regulatory element counts were normalized by the average number of regulatory elements per gene. Finally, pairwise t-tests were performed to assess the differences in the number of cis-regulatory elements across clusters.

Construction of gene regulatory networks

Late-stage cell type-specific gene regulatory networks (GRNs) were constructed independently for 13LGS and mouse. To enable a direct comparison between the two networks, we first filtered the cells and genes prior to GRN analysis. For the 13LGS dataset, retinal progenitor cells (RPCs), neurogenic retinal progenitor cells (N. RPCs), bipolar cell/photoreceptor precursors (BC/Photoreceptor precursors), and cones were selected. In contrast, the mouse dataset includes RPCs, N. RPCs, BC/photoreceptor precursors, cones, and rods. Additionally, only the conserved genes present in both datasets were retained for GRN construction.

1.Identifying cis-regulatory elements

For each target gene, three categories of peaks were identified as potential regulatory elements based on the following criteria: 1）TSS Peaks: Peaks located within a 1 kb region surrounding the target gene's transcription start site (TSS). 2）Gene Body Peaks: Peaks situated within the target gene body. Only peaks with a peak-to-gene (PtoG) correlation greater than 0.25 and statistically significant (FDR < 0.05) were retained. 3）Intergenic Peaks: Peaks found within 500 kb of the target gene's TSS, but not overlapping with the TSS or gene body regions of other target genes. These peaks were also required to have a PtoG correlation greater than 0.25 and to be statistically significant (FDR < 0.05).

2. Predicting TF binding sites

Cell type-specific TF binding sites were identified based on TF expression, motif matching, and footprint scores. First, TFs with an average expression level below 0.1 in each cell type were excluded. Next, binding sites for the remaining expressed TFs were determined by matching their motifs—sourced from the TRANSFAC2018 database—to the peak sequences using the matchMotifs function from the motifmatchr package (p-value threshold = 5e-5). Second, footprint scores were calculated using merged scATAC-Seq signals for each cell type. The Tn5 insertion sites were extracted using the "getFragmentsFromProject" function from the ArchR package, converted to BED files, and normalized with TOBIAS software[ref]. Finally, for each motif's binding region, the average normalized Tn5 insertion signals were calculated by its center (S_center) and flanking regions (S_left and S_right). The flanking region is triple the size of the center region. Motifs were retained if they met the following criteria: S_left - S_center > 0.05 and S_right - S_center > 0.05.

3. TF-target Correlation

TF-target regulatory scores were computed using the GBM-based GRN inference algorithm *via* the arboreto package's grnboost2 function [ref]. Pearson correlations for each TF-target pair were also calculated. Regulatory interactions were classified as positive if the importance score > 0.1 and Pearson correlation > 0.05, and as negative if the importance score > 0.1 and Pearson correlation < -0.05.

4. Construction of GRNs
Cell type-specific regulons were constructed by integrating data from Steps 1–3. The TF–peak–target regulons were classified as activating or repressive based on their TF–target correlations. Duplicated TF–peak–target regulons were removed prior to all downstream analyses. Additionally, TF–target regulons were derived from the TF–peak–target regulons and used to identify cone-promoting transcription factors.

5. Identification of Cone-promoting key activator TFs in 13LGS.
To identify key transcription factors (TFs) promoting cone specification in 13LGS, cone-specific genes were first identified by comparing with other late-stage neurons using the FindMarkers function (cutoffs: $\log_2$FC > 0.3 and adjusted p-value < 0.01). Next, for each TF, the enrichment of its regulatory targets within the cone-specific genes was assessed using a hypergeometric test via R's phyper function, where the "population" was defined as the total number of target genes in the positive GRNs, the "sample" as the number of positive targets for that TF, and the "successes" as the number of positive targets present in the cone-specific gene list. Additionally, the coverage of each TF was calculated as the proportion of its positive targets that are cone-specific. Finally, TFs with a hypergeometric test p-value < 1e-6 and coverage > 0.2 were identified as key activators. The Top30 TFs were ranked by the number of their positive regulated cone-specific genes.

Identification of TSS, aEnhancers and pEnhancers by CUT & RUN
TSS, bTSS (bivalent TSS), aEnhancer (active enhancer), and pEnhancer (poised enhancer) elements were identified based on the H3K4me3, H3K4me1, and H3K27ac CUT&RUN datasets in conjunction with the candidate cis-regulatory elements defined in the previous step using scATAC-Seq analysis. To determine the presence of a histone modification within each candidate cis-regulatory element, the corresponding histone peaks were overlapped with these elements, and the average signal intensity for each histone mark was calculated within their regions. A cis-regulatory element was considered to exhibit a histone modification if it both overlapped with a histone peak and showed an average signal greater than 2.5. Finally, cis-regulatory elements exhibiting H3K4me3+ were classified as TSS, those with H3K4me3– and H3K27ac+ as aEnhancer, and those with H3K4me1+ as pEnhancer.

Identification of conserved regulatory elements
Conserved regulatory elements between 13LGS and mouse were identified by comparing ATAC-Seq peaks (all potential cis-regulatory elements), as well as putative regulatory elements associated with promoters, active, and poised enhancers that were associated with conserved gene pairs. For each conserved gene pair, candidate regulatory peaks were collected from both species and paired one-to-one, with each pair comprising one peak from 13LGS and one peak

from mouse. Subsequently, the sequence of each 13LGS peak and the forward and reverse complement sequences for each mouse peak were extracted using the Biostrings package in R. Similarity scores for each candidate peak pair were computed with the “pairwiseAlignment” function using local alignment parameters (type = "local", gapOpening = -2, gapExtension = -1). To determine the cutoff for conserved peak pairs, a normal model was fitted to all similarity scores using the MASS package’s fitdistr function, and a threshold corresponding to a p-value < 0.05 under the fitted normal distribution was defined. Peak pairs with scores exceeding this threshold were considered to be conserved.

ChromVAR analysis

Global transcription factor (TF) activity at the single-cell level was evaluated using ChromVAR (version) [88]. The analysis began with the raw cell-by-peak matrix, which was processed with the addGCBias function to adjust for GC content biases. Next, motif information from the TransFac2018 database was incorporated to generate a TF z-score matrix via the matchmotifs and computeDeviations functions. Finally, these z-scores were visualized as heatmaps to illustrate TF activity patterns across the cell population.

To identify cell type-specific motifs in 13LGS, the z-score matrix generated from ChromVAR was converted into a Seurat object. The FindAllMarkers function from the Signac package was then applied with the following parameters: only.pos = TRUE, mean.fxn = rowMeans, and fc.name = "avg_diff". Finally, motifs with an adjusted p-value < 0.01 and an average difference (avg_diff) > 2 were identified as cell type-specific.

To identify differential motifs associated with photoreceptor development between 13LGS and mouse, the ChromVAR-generated z-score matrix was utilized. The following pairwise comparisons were performed: 13LGS_RPCs vs. Mouse_RPCs, 13LGS_N. RPCs vs. Mouse_N. RPCs, 13LGS_BC/Photoreceptor Precursors vs. Mouse_BC/Photoreceptor Precursors, 13LGS_Cone vs. Mouse_Cone, and 13LGS_Cone vs. Mouse_Rod. Motifs exhibiting an adjusted p-value < 0.01 and an average difference (avg_diff) > 2, with their corresponding genes showing consistent changes, were designated as differential motifs. Subsequently, all differential motifs were merged and clustered via hierarchical clustering based on their average ChromVAR z-score levels across RPCs, N. RPCs, BC/Photoreceptor precursors, cone, and rod cell types in both species.

Data availability: All scRNA-Seq, scATAC-Seq, and CUT&RUN data are available as GEO GSE295358. The custom scripts used in this paper: https://github.com/lp871/13LGS-Development.

Acknowledgements

We thank J. Nathans, A. Kolodkin, R. Johnston Jr, D.W. Kim, and W. Yap for comments on the manuscript and V. Busa for feedback on figure construction. This work was supported by NIH grants R21EY32281 and R01EY036173 to S.B, F31EY031942 to K.W, K00EY036684 to J.T, R01EY012141 to S.de V., U24EY029891 to D.K.M., an unrestricted grant from Research to Prevent Blindness to S. de V. a Bright Focus Macular Degeneration Postdoctoral Fellowship

(M2023005F) to S.K, and a Pediatric Ophthalmology Career-Starter Research Grant from Knights Templar Eye Foundation, Inc to S.K.

Figure legends:

Figure 1 : Overview of the study. (**A**) Schematic summary of the study. scRNA-Seq and scATAC-Seq of whole 13LGS retinas were performed at 10 different time points. Major retinal cell types are identified by clustering in both scRNA-Seq and scATAC-Seq datasets (**B**) UMAP projections of all 13LGS retinal cells from scRNA-Seq and scATAC-Seq. Each point represents a single cell, with cell type and time point indicated by shading. (**C**) Cell-specific marker gene expression and accessibility are shown across 11 major retinal cell types. (**D**) The relative abundances of rods, rod precursors, cones and cone precursors are different between developing 13LGS and mouse retina.

Figure 2. Identification of 13LGS–specific regulatory mechanisms in cone development by integrating late-stage retinal scRNA-Seq and scATAC-Seq data with mouse.
(**A**) UMAP projections of the integrated scRNA-Seq data from 13LGS and mouse. (**B**) UMAP projections of the scATAC-Seq data, showing cell types re-annotated based on the integrated scRNA-Seq analysis. (**C**) Heatmaps displaying eight clusters of genes that are differentially expressed between 13LGS and mouse. Column labels correspond to the five major cell types relevant to photoreceptor development: RPCs (late-stage primary retinal progenitor cells), N. RPCs (late-stage neurogenic progenitor cells), BC/photoreceptor precursors (bipolar cell, cone, and rod precursor cells), cone photoreceptors and rod photoreceptors. (**D**) Heatmaps showing the chromatin accessibility of positively-correlated regulatory elements for each of the eight DEG clusters. (**E**) Dot plots depicting the relative number of positively correlated regulatory elements for each conserved gene pair between 13LGS and mouse. Each dot represents a conserved gene pair, with the x-axis indicating the number of regulatory elements in mouse and the y-axis showing the number in 13LGS. Pairwise t-tests were used to compare the number of elements in each species for each gene pair, and the median values are indicated below each plot. (**F**) Differential motif analysis in five cone development–associated cell types between 13LGS and mouse. (**G**) *Zic3* expression motif accessibility is higher in 13LGS Photoreceptor Precursors and mature Cones than in mouse. (**H**) Heatmap showing the expression levels of the top 30 Cone-promoting genes identified by GRN analysis in 13LGS across retinal progenitor cells (RPCs), neurogenic progenitor cells (N. RPCs), and photoreceptor precursors in both 13LGS and mouse.

Figure 3: Transcription factors that show heterochronic gene expression display co-expression with Otx2 in 13LGS retinas.(**A**) RNA hybridization chain reaction (HCR) showing *Zic3*, *Rxrg*, and *Otx2* expression in P1, P5, P10, and P24. (**B**) HCR showing *Pou2f1* and *Otx2* expression in P1, P5, P10, and P24. Scale bars, 10 μm. ONL, outer nuclear layer; INL, inner nuclear layer; DAPI, 4′,6-diamidino-2-phenylindole. Arrowheads indicate cells co-expressing indicated markers.

Figure 4: Zic3 is necessary for normal mouse cone development and sufficient to promote cone gene expression. (**A**) Diagram of overexpression strategy used to test effects of *ZIC3* overexpression. (**B**) Heatmap of log-2 fold change in expression between cone-like precursors and Rods separated by condition for select genes, scaled by gene. (**C**) Upset plot of significant

differentially expressed genes shared between conditions. (**D**) Diagram of retinal progenitor cell-specific conditional loss of function analysis of *Zic3*. (**E**) Immunohistochemistry showing GFP and Arr3 expression in P14 wild type (WT) and *Zic3*$^{lox/lox}$ (*Zic3* cKO) mouse retinas. Scale bars, 20 µm. (**F**) Box plot of the number of cones per micron along the retinal slice (n = 11 WT, 10 *Zic3* cKO). P-values calculated by Wilcoxon rank-sum test. P0, postnatal day 0; P5, postnatal day 5; P14, postnatal day 14; FACS, fluorescence-activated cell sorting; scRNA-Seq, single-cell RNA sequencing; ONL, outer nuclear layer; INL, inner nuclear layer; DAPI, 4′,6-diamidino-2-phenylindole; GFP, green fluorescent protein; OE, overexpression; log2FC, log2 fold change; DEG, differentially expressed gene; WT, wild type; cKO, conditional knockout.

Figure 5: Mef2c is sufficient to promote cone-specific gene expression and repress rod-specific gene expression in mouse. (**A**) Diagram of overexpression strategy used to test effects of *MEF2C* overexpression. (**B**) UMAP representation of P8 mouse retinal explants electroporated with a plasmid expressing GFP alone (Empty) or GFP in a bicistronic transcript with human MEF2C (*MEF2C*), (n = 7445 cell Empty, 11949 cells *MEF2C*). Each point represents a single cell and is colored by cell type as determined by clustering and marker gene expression. (**C**) Heatmap of expression for select genes for cones, cone-like photoreceptor precursors, rod photoreceptor precursors, and rods in cells overexpressing *MEF2C*, scaled by gene. (**D**) Immunohistochemistry showing GFP and GNAT2 expression in P8 mouse retinas from Empty and *MEF2C* conditions. Scale bars, 50 µm. (**E**) Box plot of the number of Gnat2+, GFP+ cells divided by the total number of GFP+ cells (n = 8 for both conditions). P-values calculated by Wilcoxon rank-sum test. P0, postnatal day 0; P8, postnatal day 8; FACS, fluorescence-activated cell sorting; scRNA-Seq, single-cell RNA sequencing; ONL, outer nuclear layer; INL, inner nuclear layer; DAPI, 4′,6-diamidino-2-phenylindole; GFP, green fluorescent protein; OE, overexpression.

Figure 6: Conserved TFs bind to species-specific enhancers to promote Cone specification in 13LGS. (**A**) Schematic illustrating annotation of cis-regulatory elements in RPCs and Photoreceptor Precursors by integration of scATAC-Seq and CUT&RUN 13LGS and mouse datasets. (**B**) Heatmaps show annotated accessible regulatory elements in both 13LGS and mouse. Promoters, activated enhancers (AEs), and poised enhancers (PEs) which are associated with histone markers associated with genes in clusters C2 and C3, which are selectively active in 13LGS RPCs and/or photoreceptor precursors. Shading indicates CUT&TAG signal for the corresponding histone modification within 2kb of the scATAC-Seq peak center. Bar plots displaying the number of each category of regulatory element in each species that are conserved or species-specific. (**C**) Dot plots showing the enrichment of binding sites for Otx2 and Neurod1, TFs which are broadly expressed in both neurogenic RPC and rod and cone precursors, which are enriched in both evolutionarily-conserved cis-regulatory elements in both species. (**D**) Bar plots showing the number of evolutionarily-conversed and species-specific enhancers per TSS in four cone-promoting genes between 13LGS and mouse. (E) The gene regulatory networks regulating *Thrb* expression in 13LGS and mouse late N. RPCs. (**F**) An example of a Thrb-related regulon and its corresponding scATACseq and CUT&RUN tracks. The arrow indicates the consistent regulatory relationships between GRNs prediction and experimental validations. (**G**) The epigenetic model of cone specification in 13LGS and mouse.

Figure S1 : Quality control of scATAC-Seq and scRNA-Seq data. (**A**) Transcriptional start site enrichment profiles of scATAC-Seq datasets. Lines are colored by time points. (**B**) Fragment size distribution of scATAC-Seq datasets. Individual time points are indicated by colored lines. (**C**) Heatmap showing the Pearson correlation between gene expression and gene accessibility for each retina cell type. (**D**) Heatmap of cell type-specific genes. (**E**) Heatmap of cell type-specific peaks. (**F**) Heatmap of cell type-specific motifs. (**G**) Examples of TF footprint profiles for Pou4f2, Crx, Nfix and Onecut1 in indicated scATAC-Seq cell types. (**H**) Examples of chromVAR score are shown for Otx2, Pou2f2, Nfix and Neurod1 using scATAC-Seq datasets.
(**I**) The relative abundance of retinal cell types in scRNA-Seq and scATAC-Seq is different between developing 13LGS and mouse retina.

Figure S2: Transcription factors that show heterochronic gene expression and co-expression with Otx2 show sustained expression in 13LGS retinas. Immunohistochemistry showing co-expression of Mef2c and Otx2 in P5, P10 and P24. Scale bars, 10 µm. DAPI, 4′,6-diamidino-2-phenylindole; NbONL, neuroblastic outer nuclear layer; ONL, outer nuclear layer; INL, inner nuclear layer. Arrowheads indicate cells co-expressing indicated markers.

Figure S3: *ZIC3* overexpression promotes the formation of cone-like photoreceptor precursor cells. (**A**) UMAP representation of P5 mouse retinal explants electroporated with a plasmid expressing GFP alone (Empty), GFP in a bicistronic transcript with ZIC3 (*ZIC3, ZIC3+POU2F1*) or POU2F1 (*POU2F1, ZIC3+POU2F1*) (n = 1901 cells Empty, 1578 cells *ZIC3*, 904 cells *POU2F1*, 2111 cells *ZIC3+POU2F1*), split by condition. Each point represents a single cell and is colored by cell type as determined by clustering and marker gene expression. (**B**) UMAP representation of P5 mouse retinal explants electroporated with a plasmid expressing GFP alone (Empty), GFP in a bicistronic transcript with ZIC3 (*ZIC3+POU2F1*), POU2F1 (*POU2F1, ZIC3+POU2F1*), or ONECUT1 (*ONECUT1*) (n = 4313 cells Empty, 6287 cells *POU2F1*, 2476 cells *ZIC3+POU2F1*, 4455 cells *ONECUT1*), split by condition. Each point represents a single cell and is colored by cell type as determined by clustering and marker gene expression. (**C**) Heatmap of log-2 fold change in expression between cone-like precursors and rods separated by condition for select genes, scaled by gene. (**D**) Upset plot of significant differentially expressed genes shared between conditions. RPC, retinal progenitor cell; N. RPC, neurogenic retinal progenitor cells; HC, horizontal cell, BC, bipolar cell; AC, amacrine cell; log2FC, log2 fold change; DEG, differentially expressed gene.

Figure S4: *Zic3* is required for normal gene expression in cones and Müller glia. (**A**) Immunohistochemistry showing GFP and Nr2e3 expression in P14 wild type (WT) and *Zic3*$^{lox/lox}$ (*Zic3* cKO) mouse retinas. (**B**) Box plot of the number of rods per square micron along the retinal slice (n = 12 WT, 10 *Zic3* cKO). (**C**) Immunohistochemistry showing GFP and LHX1 expression in P14 WT and *Zic3* cKO mouse retinas. (**D**) Box plot of the number of horizontal cells per micron along the retinal slice (n = 3 WT, 2 *Zic3* cKO). (**E**) UMAP representation of P14 WT and *Zic3* cKO mouse retinas (n = 6174 cells (N=4) WT, 7562 cells (N=4) *Zic3*), split by genotype. Each point represents a single cell and is colored by cell type as determined by clustering and marker gene expression. Heatmaps of expression in (**F**) Cones and (**G**) Müller

glia for select genes for WT and *Zic3* cKO retinas, scaled by gene. Scale bars, 10 μm. P-value calculated by Wilcoxon rank-sum test. DAPI, 4′,6-diamidino-2-phenylindole; GFP, green fluorescent protein; ONL, outer nuclear layer; INL, inner nuclear layer; HC, horizontal cell; AC/HC, amacrine cells/horizontal cell; RGC, retinal ganglion cell; BC, bipolar cell; MG, Müller glia.

Figure S5: *Mef2c* overexpression does not obviously impact non-Cone cell type proportions. Immunohistochemistry showing GFP and (**A**) Nrl, (**B**) Otx2, (**C**) Sox9, or (**D**) Tfap2a expression in P8 mouse retinal explants electroporated with a plasmid expressing GFP alone (Empty) or GFP in a bicistronic transcript with MEF2C (*MEF2C*). Scale bars, 50 μm. DAPI, 4′,6-diamidino-2-phenylindole; GFP, green fluorescent protein; ONL, outer nuclear layer; INL, inner nuclear layer.

Figure S6: *Mef2c* loss of function does not affect Cone cell production or survival. (**A**) Diagram of conditional knockout strategy used to test for the necessity of *Mef2c* for normal cone development. (**B**) Immunohistochemistry showing GFP and Arr3 expression in P14 wild type (WT) and *Chx10-Cre;Mef2c*$^{lox/lox}$ (*Mef2c* cKO) mouse retinas. Scale bars, 50 μm. (**C**) Box plot of the number of ARR3+, GFP+ cells divided by the total number of GFP+ cells (n = 4). P-value calculated by Wilcoxon rank-sum test. WT, wildtype; cKO, conditional knockout; DAPI, 4′,6-diamidino-2-phenylindole; GFP, green fluorescent protein; ONL, outer nuclear layer; INL, inner nuclear layer, GCL, ganglion cell layer.

Figure S7: Transcriptional regulatory networks controlling *Thrb* expression are more complex in 13LGS than mouse at all developmental stages. is regulated by more extensive gene regulatory networks in 13LGS than mouse. Gene regulatory networks directing *Thrb* expression in (**A**) 13LGS photoreceptor precursors, (**B**) mouse photoreceptor precursors, (**C**) 13LGS cones, and (**D**) mouse cones.

Figure S8: Zic3 is regulated by more extensive gene regulatory networks in 13LGS than mouse at all developmental stages. (**A**) Gene regulatory networks controlling *Zic3* expression in 13LGS late-stage neurogenic RPCs, photoreceptor precursors, and cones. (**B**) Example of Zic3-related regulons and their corresponding scATAC-Seq and CUT&RUN tracks in 13LGS. (**C**) Example of Zic3-related regulons and their corresponding scATAC-Seq and CUT&RUN tracks in mouse.

Figure S9: Transcriptional regulatory networks controlling cone-specific gene expression in 13LGS increase in connectivity and complexity as differentiation proceeds. Cell type-specific 13LGS gene regulatory networks with cone-promoting TFs in (**A**) late-stage neurogenic retinal progenitor cells (late N. RPCs), (**B**) photoreceptor precursors, and (**C**) cones.

Table ST1: List of cell type-specific patterns of differential gene expression, chromatin accessibility, and motif enrichment in accessible chromatin for 13LGS.

Table ST2: Gene ortholog pairs for 13LGS and mouse used in this analysis.

Table ST3: Genes and accessible transcription factor motifs showing differential activity between 13LGS and mouse in late-stage progenitors and differentiating photoreceptors.

Table ST4: Top cone-promoting transcription factors active in 13LGS.

Table ST5: List of genes differentially expressed following overexpression of *ZIC3*, *ONECUT1*, *POU2F1*, and *ZIC3*+*POU2F1* as detected by scRNA-Seq.

Table ST6: List of genes differentially expressed in *Chx10-Cre;Zic3*$^{lox/lox}$ retina.

Table ST7: List of genes differentially expressed following overexpression of *MEF2C* as detected by scRNA-Seq.

Table ST8: Cis-regulatory elements inferred by scATAC-Seq and CUT&TAG for genes in Clusters 2 and 3 in 13LGS and mouse.

Table ST9: Active and poised enhancers associated with *Mef2c*, *Rxrg*, *Thrb* and *Zic3* in 13LGS and mouse.

References:

1. Arendt, D. (2008). The evolution of cell types in animals: emerging principles from molecular studies. Nat Rev Genet *9*, 868–882.

2. Katz, P.S. (2011). Neural mechanisms underlying the evolvability of behaviour. Philos Trans R Soc Lond B Biol Sci *366*, 2086–2099.

3. Le Duc, D., and Schöneberg, T. (2016). Adaptation to nocturnality - learning from avian genomes. Bioessays *38*, 694–703.

4. Slavenko, A., Dror, L., Camaiti, M., Farquhar, J.E., Shea, G.M., Chapple, D.G., and Meiri, S. (2022). Evolution of diel activity patterns in skinks (Squamata: Scincidae), the world's second-largest family of terrestrial vertebrates. Evolution *76*, 1195–1208.

5. Walls, G.L. (1942). The Vertebrate Eye and Its Adaptive Radiation.

6. Gerkema, M.P., Davies, W.I.L., Foster, R.G., Menaker, M., and Hut, R.A. (2013). The nocturnal bottleneck and the evolution of activity patterns in mammals. Proc Biol Sci *280*, 20130508.

7. Ahnelt, P.K., and Kolb, H. (2000). The mammalian photoreceptor mosaic-adaptive design. Prog Retin Eye Res *19*, 711–777.

8. Peichl, L. (2005). Diversity of mammalian photoreceptor properties: adaptations to habitat and lifestyle? Anat Rec A Discov Mol Cell Evol Biol *287*, 1001–1012.

9. Kaskan, P.M., Franco, E.C.S., Yamada, E.S., Silveira, L.C. de L., Darlington, R.B., and Finlay, B.L. (2005). Peripheral variability and central constancy in mammalian visual system evolution. Proc Biol Sci *272*, 91–100.

10. Mustafi, D., Engel, A.H., and Palczewski, K. (2009). Structure of cone photoreceptors. Prog Retin Eye Res *28*, 289–302.

11. Jeon, C.J., Strettoi, E., and Masland, R.H. (1998). The major cell populations of the mouse retina. J Neurosci *18*, 8936–8946.

12. Verra, D.M., Sajdak, B.S., Merriman, D.K., and Hicks, D. (2020). Diurnal rodents as pertinent animal models of human retinal physiology and pathology. Prog Retin Eye Res *74*, 100776.

13. Kryger, Z., Galli-Resta, L., Jacobs, G.H., and Reese, B.E. (1998). The topography of rod and cone photoreceptors in the retina of the ground squirrel. Vis Neurosci *15*, 685–691.

14. Zhang, X., Leavey, P., Appel, H., Makrides, N., and Blackshaw, S. (2023). Molecular mechanisms controlling vertebrate retinal patterning, neurogenesis, and cell fate specification. Trends Genet *39*, 736–757.

15. Clark, B.S., Stein-O'Brien, G.L., Shiau, F., Cannon, G.H., Davis-Marcisak, E., Sherman, T., Santiago, C.P., Hoang, T.V., Rajaii, F., James-Esposito, R.E., et al. (2019). Single-Cell RNA-Seq Analysis of Retinal Development Identifies NFI Factors as Regulating Mitotic Exit and Late-Born Cell Specification. Neuron *102*, 1111–1126.e5.

16. Cepko, C. (2014). Intrinsically different retinal progenitor cells produce specific types of progeny. Nat Rev Neurosci *15*, 615–627.

17. Lyu, P., Hoang, T., Santiago, C.P., Thomas, E.D., Timms, A.E., Appel, H., Gimmen, M., Le, N., Jiang, L., Kim, D.W., et al. (2021). Gene regulatory networks controlling temporal patterning, neurogenesis, and cell-fate specification in mammalian retina. Cell Rep *37*, 109994.

18. Emerson, M.M., Surzenko, N., Goetz, J.J., Trimarchi, J., and Cepko, C.L. (2013). Otx2 and Onecut1 promote the fates of cone photoreceptors and horizontal cells and repress rod photoreceptors. Dev Cell *26*, 59–72.

19. Javed, A., Mattar, P., Lu, S., Kruczek, K., Kloc, M., Gonzalez-Cordero, A., Bremner, R., Ali, R.R., and Cayouette, M. (2020). Pou2f1 and Pou2f2 cooperate to control the timing of cone photoreceptor production in the developing mouse retina. Development *147*. https://doi.org/10.1242/dev.188730.

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

21. Milam, A.H., Rose, L., Cideciyan, A.V., Barakat, M.R., Tang, W.-X., Gupta, N., Aleman, T.S., Wright, A.F., Stone, E.M., Sheffield, V.C., et al. (2002). The nuclear receptor NR2E3 plays a role in human retinal photoreceptor differentiation and degeneration. Proc Natl Acad Sci U S A *99*, 473–478.

22. Cheng, H., Khanna, H., Oh, E.C.T., Hicks, D., Mitton, K.P., and Swaroop, A. (2004). Photoreceptor-specific nuclear receptor NR2E3 functions as a transcriptional activator in rod photoreceptors. Hum Mol Genet *13*, 1563–1575.

23. Mears, A.J., Kondo, M., Swain, P.K., Takada, Y., Bush, R.A., Saunders, T.L., Sieving, P.A., and Swaroop, A. (2001). Nrl is required for rod photoreceptor development. Nat Genet *29*, 447–452.

24. Daniele, L.L., Lillo, C., Lyubarsky, A.L., Nikonov, S.S., Philp, N., Mears, A.J., Swaroop, A., Williams, D.S., and Pugh, E.N., Jr (2005). Cone-like morphological, molecular, and electrophysiological features of the photoreceptors of the Nrl knockout mouse. Invest Ophthalmol Vis Sci *46*, 2156–2167.

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

26. de Melo, J., Peng, G.-H., Chen, S., and Blackshaw, S. (2011). The Spalt family transcription factor Sall3 regulates the development of cone photoreceptors and retinal horizontal interneurons. Development *138*, 2325–2336.

27. Mossman, H.W., and Weisfeldt, L.A. (1939). The fetal membranes of a primitive rodent, the thirteen-striped ground squirrel. Am. J. Anat. *64*, 59–109.

28. Thomas, E.D., Timms, A.E., Giles, S., Harkins-Perry, S., Lyu, P., Hoang, T., Qian, J., Jackson, V.E., Bahlo, M., Blackshaw, S., et al. (2022). Cell-specific cis-regulatory elements and mechanisms of non-coding genetic disease in human retina and retinal organoids. Dev

Cell *57*, 820–836.e6.

29. Sapkota, D., Chintala, H., Wu, F., Fliesler, S.J., Hu, Z., and Mu, X. (2014). Onecut1 and Onecut2 redundantly regulate early retinal cell fates during development. Proc. Natl. Acad. Sci. U. S. A. *111*, E4086–E4095.

30. Wolf, A., Aslanidis, A., and Langmann, T. (2017). Retinal expression and localization of Mef2c support its important role in photoreceptor gene expression. Biochem Biophys Res Commun *483*, 346–351.

31. Kallman, A., Capowski, E.E., Wang, J., Kaushik, A.M., Jansen, A.D., Edwards, K.L., Chen, L., Berlinicke, C.A., Joseph Phillips, M., Pierce, E.A., et al. (2020). Investigating cone photoreceptor development using patient-derived NRL null retinal organoids. Commun Biol *3*, 82.

32. Nagar, S., Trudler, D., McKercher, S.R., Piña-Crespo, J., Nakanishi, N., Okamoto, S.-I., and Lipton, S.A. (2017). Molecular Pathway to Protection From Age-Dependent Photoreceptor Degeneration in Mef2 Deficiency. Invest Ophthalmol Vis Sci *58*, 3741–3749.

33. Roberts, M.R., Hendrickson, A., McGuire, C.R., and Reh, T.A. (2005). Retinoid X receptor (gamma) is necessary to establish the S-opsin gradient in cone photoreceptors of the developing mouse retina. Invest Ophthalmol Vis Sci *46*, 2897–2904.

34. Ng, L., Hurley, J.B., Dierks, B., Srinivas, M., Saltó, C., Vennström, B., Reh, T.A., and Forrest, D. (2001). A thyroid hormone receptor that is required for the development of green cone photoreceptors. Nat Genet *27*, 94–98.

35. Satoh, S., Tang, K., Iida, A., Inoue, M., Kodama, T., Tsai, S.Y., Tsai, M.-J., Furuta, Y., and Watanabe, S. (2009). The spatial patterning of mouse cone opsin expression is regulated by bone morphogenetic protein signaling through downstream effector COUP-TF nuclear receptors. J Neurosci *29*, 12401–12411.

36. Mattar, P., Ericson, J., Blackshaw, S., and Cayouette, M. (2015). A conserved regulatory logic controls temporal identity in mouse neural progenitors. Neuron *85*, 497–504.

37. Hao, H., Kim, D.S., Klocke, B., Johnson, K.R., Cui, K., Gotoh, N., Zang, C., Gregorski, J., Gieser, L., Peng, W., et al. (2012). Transcriptional regulation of rod photoreceptor homeostasis revealed by in vivo NRL targetome analysis. PLoS Genet *8*, e1002649.

38. Omori, Y., Kubo, S., Kon, T., Furuhashi, M., Narita, H., Kominami, T., Ueno, A., Tsutsumi, R., Chaya, T., Yamamoto, H., et al. (2017). Samd7 is a cell type-specific PRC1 component essential for establishing retinal rod photoreceptor identity. Proc Natl Acad Sci U S A *114*, E8264–E8273.

39. Nishida, A., Furukawa, A., Koike, C., Tano, Y., Aizawa, S., Matsuo, I., and Furukawa, T. (2003). Otx2 homeobox gene controls retinal photoreceptor cell fate and pineal gland development. Nat Neurosci *6*, 1255–1263.

40. Furukawa, T., Morrow, E.M., Li, T., Davis, F.C., and Cepko, C.L. (1999). Retinopathy and attenuated circadian entrainment in Crx-deficient mice. Nat Genet *23*, 466–470.

41. Pennesi, M.E., Cho, J.-H., Yang, Z., Wu, S.H., Zhang, J., Wu, S.M., and Tsai, M.-J. (2003).

BETA2/NeuroD1 null mice: a new model for transcription factor-dependent photoreceptor degeneration. J Neurosci *23*, 453–461.

42. Katoh, K., Omori, Y., Onishi, A., Sato, S., Kondo, M., and Furukawa, T. (2010). Blimp1 suppresses Chx10 expression in differentiating retinal photoreceptor precursors to ensure proper photoreceptor development. J Neurosci *30*, 6515–6526.

43. Andzelm, M.M., Cherry, T.J., Harmin, D.A., Boeke, A.C., Lee, C., Hemberg, M., Pawlyk, B., Malik, A.N., Flavell, S.W., Sandberg, M.A., et al. (2015). MEF2D drives photoreceptor development through a genome-wide competition for tissue-specific enhancers. Neuron *86*, 247–263.

44. Lu, Y., Shiau, F., Yi, W., Lu, S., Wu, Q., Pearson, J.D., Kallman, A., Zhong, S., Hoang, T., Zuo, Z., et al. (2020). Single-Cell Analysis of Human Retina Identifies Evolutionarily Conserved and Species-Specific Mechanisms Controlling Development. Dev Cell *53*, 473–491.e9.

45. de Melo, J., and Blackshaw, S. (2018). In Vivo Electroporation of Developing Mouse Retina. Methods Mol Biol *1715*, 101–111.

46. Purandare, S.M., Ware, S.M., Kwan, K.M., Gebbia, M., Bassi, M.T., Deng, J.M., Vogel, H., Behringer, R.R., Belmont, J.W., and Casey, B. (2002). A complex syndrome of left-right axis, central nervous system and axial skeleton defects in Zic3 mutant mice. Development *129*, 2293–2302.

47. Rowan, S., and Cepko, C.L. (2004). Genetic analysis of the homeodomain transcription factor Chx10 in the retina using a novel multifunctional BAC transgenic mouse reporter. Dev Biol *271*, 388–402.

48. Sutherland, M.J., Wang, S., Quinn, M.E., Haaning, A., and Ware, S.M. (2013). Zic3 is required in the migrating primitive streak for node morphogenesis and left-right patterning. Hum Mol Genet *22*, 1913–1923.

49. Akhmedov, N.B., Piriev, N.I., Chang, B., Rapoport, A.L., Hawes, N.L., Nishina, P.M., Nusinowitz, S., Heckenlively, J.R., Roderick, T.H., Kozak, C.A., et al. (2000). A deletion in a photoreceptor-specific nuclear receptor mRNA causes retinal degeneration in the rd7 mouse. Proc Natl Acad Sci U S A *97*, 5551–5556.

50. Vong, L.H., Ragusa, M.J., and Schwarz, J.J. (2005). Generation of conditional Mef2cloxP/loxP mice for temporal- and tissue-specific analyses. Genesis *43*, 43–48.

51. Chang, B., Hawes, N.L., Hurd, R.E., Davisson, M.T., Nusinowitz, S., and Heckenlively, J.R. (2002). Retinal degeneration mutants in the mouse. Vision Res *42*, 517–525.

52. Saxena, A., Sharma, V., Muthuirulan, P., Neufeld, S.J., Tran, M.P., Gutierrez, H.L., Chen, K.D., Erberich, J.M., Birmingham, A., Capellini, T.D., et al. (2022). Interspecies transcriptomics identify genes that underlie disproportionate foot growth in jerboas. Curr Biol *32*, 289–303.e6.

53. Cretekos, C.J., Wang, Y., Green, E.D., Martin, J.F., Rasweiler, J.J., 4th, and Behringer, R.R. (2008). Regulatory divergence modifies limb length between mammals. Genes Dev *22*, 141–151.

54. Lonfat, N., Wang, S., Lee, C., Garcia, M., Choi, J., Park, P.J., and Cepko, C. (2021). Cis-regulatory dissection of cone development reveals a broad role for Otx2 and Oc transcription factors. Development *148*. https://doi.org/10.1242/dev.198549.

55. Müller, B., and Peichl, L. (1989). Topography of cones and rods in the tree shrew retina. J Comp Neurol *282*, 581–594.

56. Yuodelis, C., and Hendrickson, A. (1986). A qualitative and quantitative analysis of the human fovea during development. Vision Res *26*, 847–855.

57. Lamb, T.D. (2013). Evolution of phototransduction, vertebrate photoreceptors and retina. Prog Retin Eye Res *36*, 52–119.

58. Hahn, J., Monavarfeshani, A., Qiao, M., Kao, A.H., Kölsch, Y., Kumar, A., Kunze, V.P., Rasys, A.M., Richardson, R., Wekselblatt, J.B., et al. (2023). Evolution of neuronal cell classes and types in the vertebrate retina. Nature *624*, 415–424.

59. Fernald, R.D. (2012). Social control of the brain. Annu Rev Neurosci *35*, 133–151.

60. Yang, L., Liu, Q., and Pollak, G.D. (1996). Afferent connections to the dorsal nucleus of the lateral lemniscus of the mustache bat: evidence for two functional subdivisions. J Comp Neurol *373*, 575–592.

61. Jacobs, L.F., and Spencer, W.D. (1994). Natural space-use patterns and hippocampal size in kangaroo rats. Brain Behav Evol *44*, 125–132.

62. Young, L.J., and Wang, Z. (2004). The neurobiology of pair bonding. Nat Neurosci *7*, 1048–1054.

63. Wada, K., Howard, J.T., McConnell, P., Whitney, O., Lints, T., Rivas, M.V., Horita, H., Patterson, M.A., White, S.A., Scharff, C., et al. (2006). A molecular neuroethological approach for identifying and characterizing a cascade of behaviorally regulated genes. Proc Natl Acad Sci U S A *103*, 15212–15217.

64. Chen, J., Richardson, P.R., Kirby, C., Eddy, S.R., and Hoekstra, H.E. (2025). Cellular evolution of the hypothalamic preoptic area of behaviorally divergent deer mice. bioRxiv. https://doi.org/10.1101/2024.08.22.608850.

65. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., 3rd, Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29.

66. Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol *20*, 296.

67. Stuart, T., Srivastava, A., Madad, S., Lareau, C.A., and Satija, R. (2021). Single-cell chromatin state analysis with Signac. Nat Methods *18*, 1333–1341.

68. Jiang, Z., Zhu, L., Hu, L., Slesnick, T.C., Pautler, R.G., Justice, M.J., and Belmont, J.W. (2013). Zic3 is required in the extra-cardiac perinodal region of the lateral plate mesoderm for left-right patterning and heart development. Hum Mol Genet *22*, 879–889.

69. Weir, K., Leavey, P., Santiago, C., and Blackshaw, S. (2021). Multiplexed Analysis of

Retinal Gene Expression and Chromatin Accessibility using scRNA-Seq and scATAC-Seq. J Vis Exp. https://doi.org/10.3791/62239.

70. Kandoi, S., Martinez, C., Merriman, D.K., and Lamba, D.A. (2022). Characterization of Retinal Development in 13-Lined Ground Squirrels. Transl Vis Sci Technol *11*, 17.

71. Yu, C.T., Kandoi, S., Periasamy, R., Reddy, L.V.K., Follett, H.M., Summerfelt, P., Martinez, C., Guillaume, C., Bowie, O., Connor, T.B., et al. (2024). Human iPSC-derived photoreceptor transplantation in the cone dominant 13-lined ground squirrel. Stem Cell Reports *19*, 331–342.

72. Matsuda, T., and Cepko, C.L. (2008). Analysis of gene function in the retina. Methods Mol Biol *423*, 259–278.

73. Matsuda, T., and Cepko, C.L. (2004). Electroporation and RNA interference in the rodent retina in vivo and in vitro. Proc Natl Acad Sci U S A *101*, 16–22.

74. Onishi, A., Peng, G.-H., Hsu, C., Alexis, U., Chen, S., and Blackshaw, S. (2009). Pias3-dependent SUMOylation directs rod photoreceptor development. Neuron *61*, 234–246.

75. Tabata, Y., Ouchi, Y., Kamiya, H., Manabe, T., Arai, K.-I., and Watanabe, S. (2004). Specification of the retinal fate of mouse embryonic stem cells by ectopic expression of Rx/rax, a homeobox gene. Mol Cell Biol *24*, 4513–4521.

76. Fu, R., Gillen, A.E., Grabek, K.R., Riemondy, K.A., Epperson, L.E., Bustamante, C.D., Hesselberth, J.R., and Martin, S.L. (2020). Dynamic RNA Regulation in the Brain Underlies Physiological Plasticity in a Hibernating Mammal. Front Physiol *11*, 624677.

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

78. Tangeman, J.A., Charris Dominguez, C.M., Bendezu-Sayas, S., and Del Rio-Tsonis, K. (2025). Nuclei Isolation from Ocular Tissues of the Embryonic Chicken for Single-Nucleus Profiling. Methods Mol Biol *2848*, 105–116.

79. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. *17*, 10.

80. Krueger, F., James, F., Ewels, P., Afyounian, E., Weinstein, M., Schuster-Boeckler, B., Hulselmans, G., and sclamons (2023). FelixKrueger/TrimGalore: v0.6.10 - add default decompression path (Zenodo) https://doi.org/10.5281/ZENODO.7598955.

81. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods *9*, 357–359.

82. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., et al. (2021). Twelve years of SAMtools and BCFtools. Gigascience *10*. https://doi.org/10.1093/gigascience/giab008.

83. Smith, T., Heger, A., and Sudbery, I. (2017). UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. Genome Res *27*, 491–499.

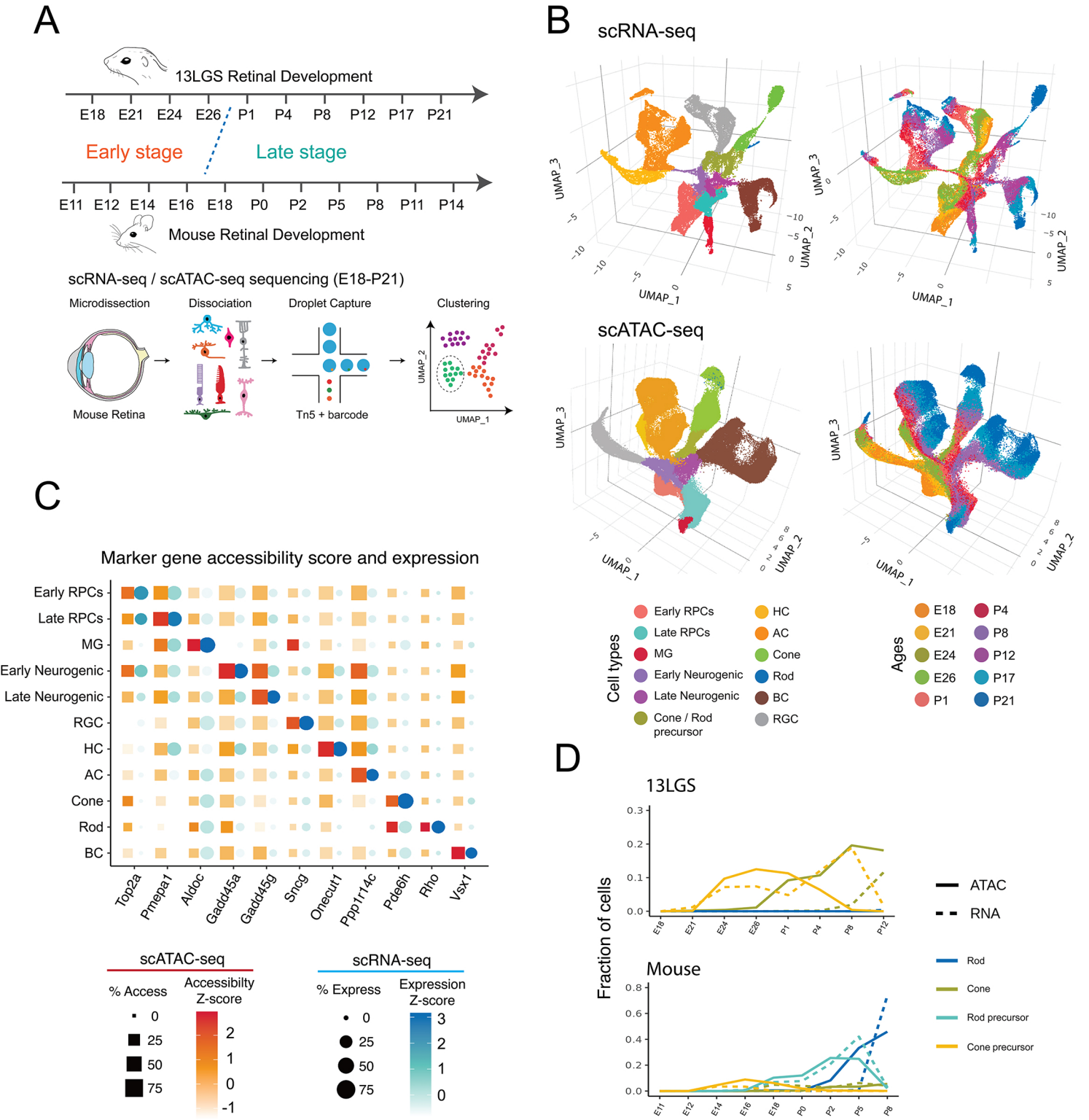
84. Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res *44*, W160–W165.

85. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

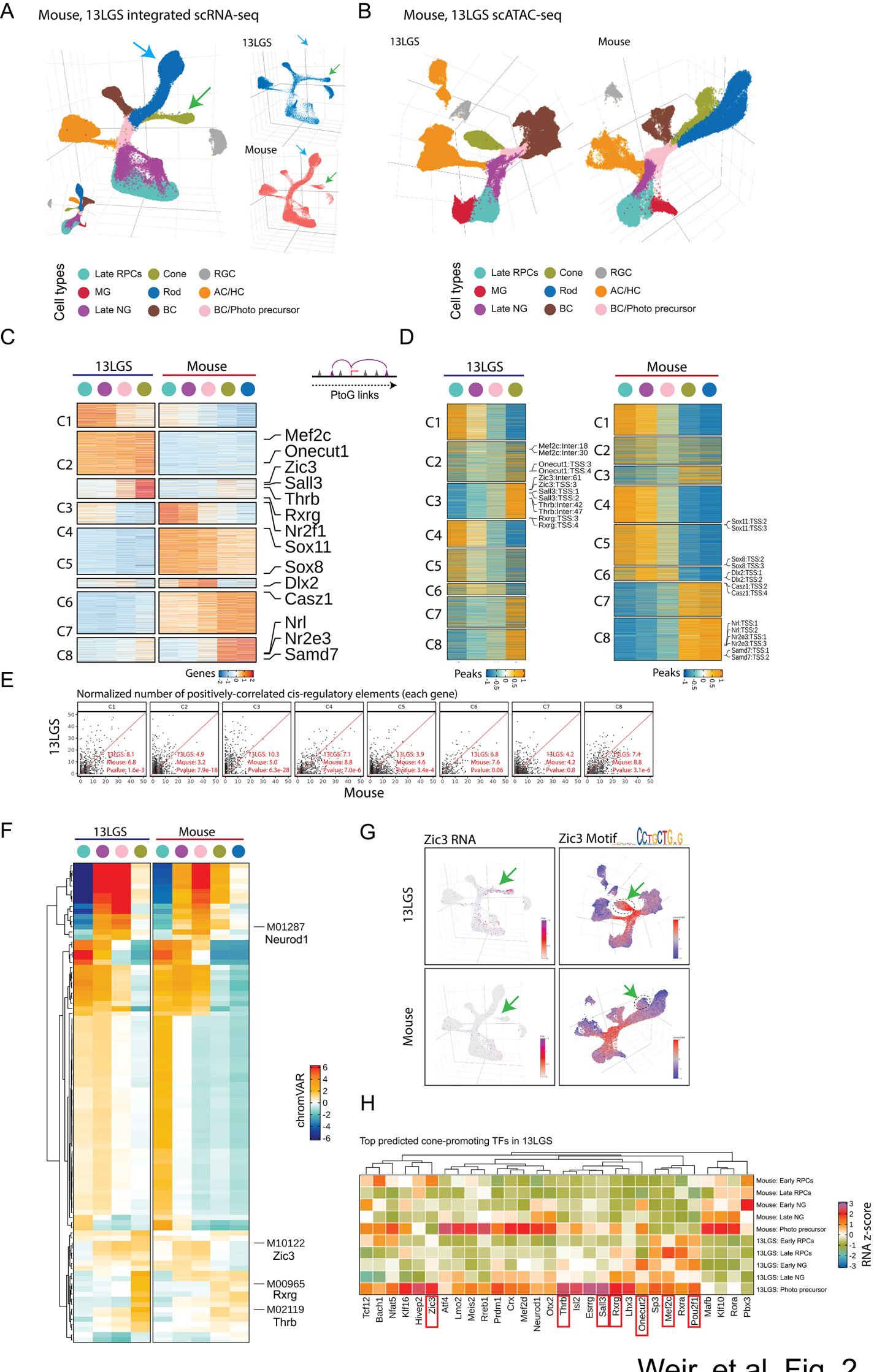
86. Amemiya, H.M., Kundaje, A., and Boyle, A.P. (2019). The ENCODE Blacklist: Identification of Problematic Regions of the Genome. Sci Rep *9*, 9354.

87. Granja, J.M., Corces, M.R., Pierce, S.E., Bagdatli, S.T., Choudhry, H., Chang, H.Y., and Greenleaf, W.J. (2021). ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. Nat Genet *53*, 403–411.

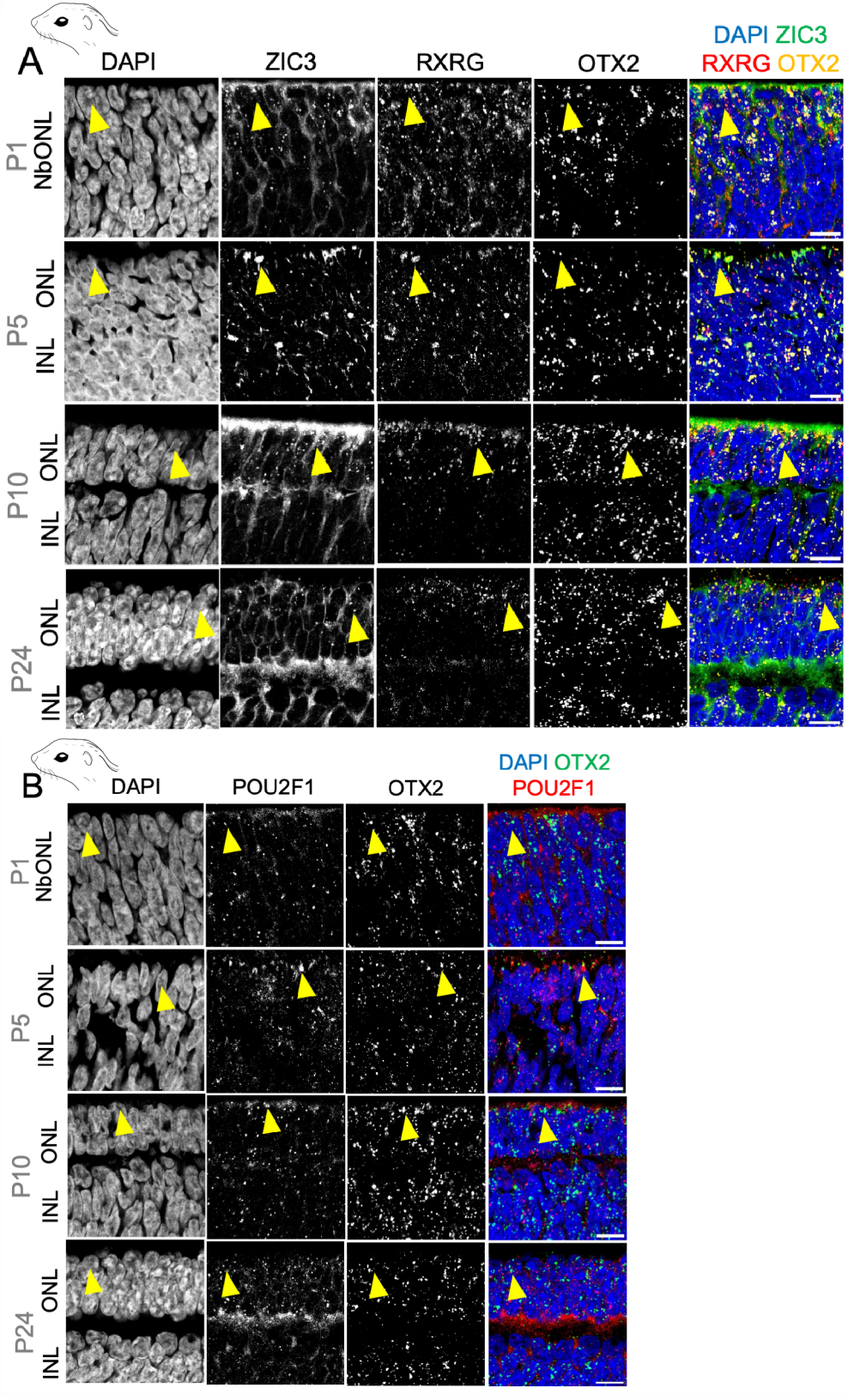
88. Schep, A.N., Wu, B., Buenrostro, J.D., and Greenleaf, W.J. (2017). chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. Nat Methods *14*, 975–978.



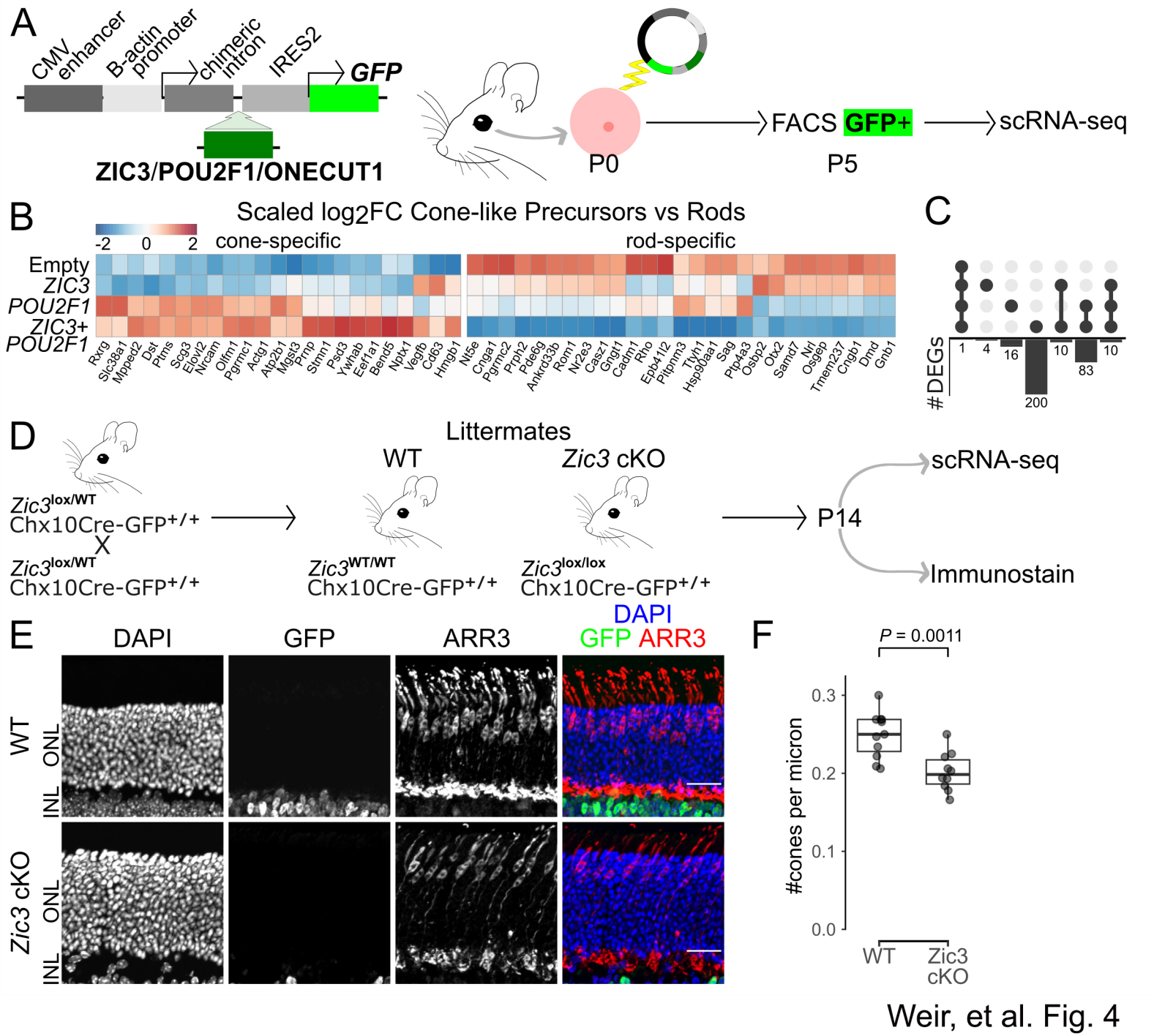
Weir, et al. Fig. 1



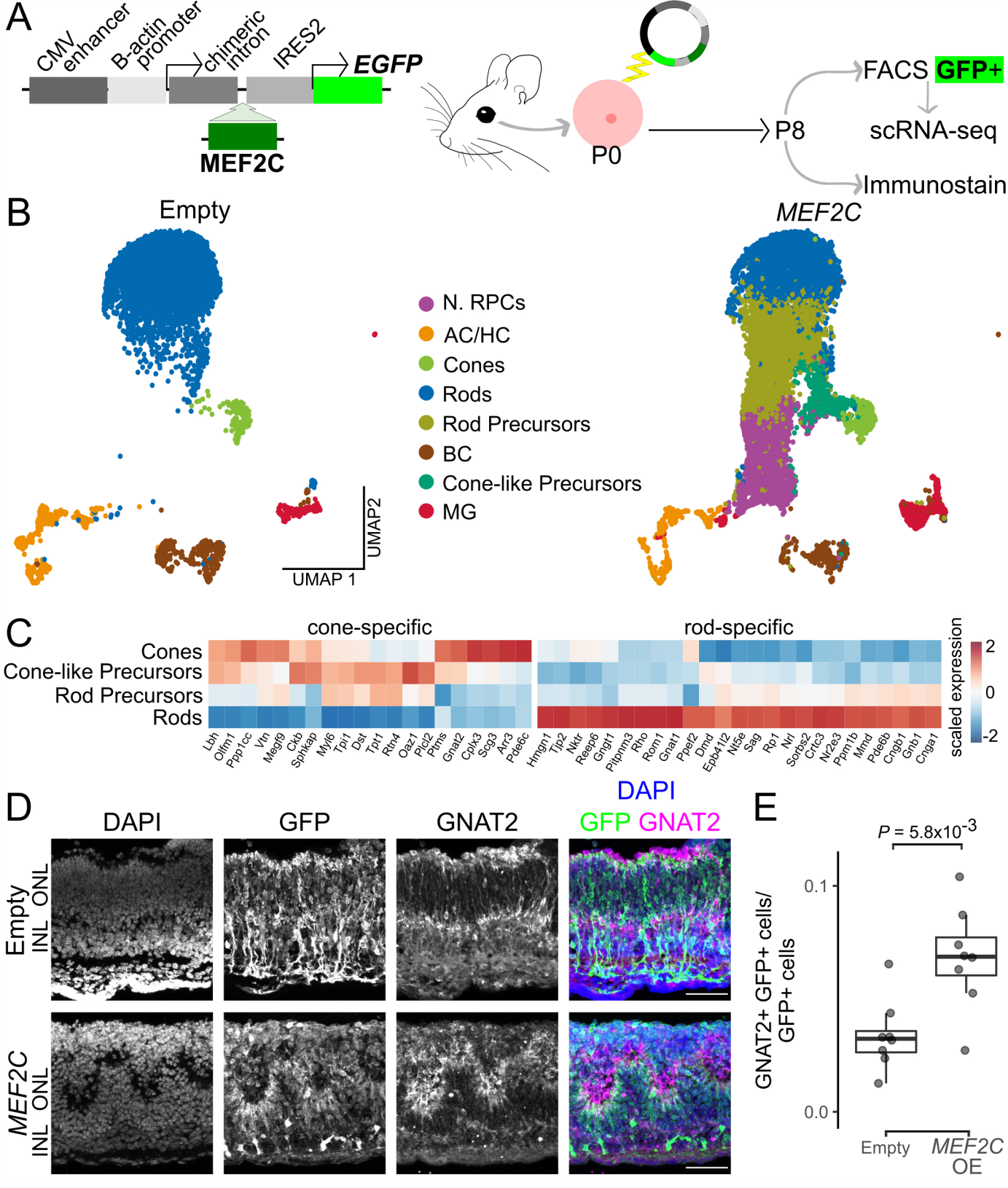
Weir, et al. Fig. 2



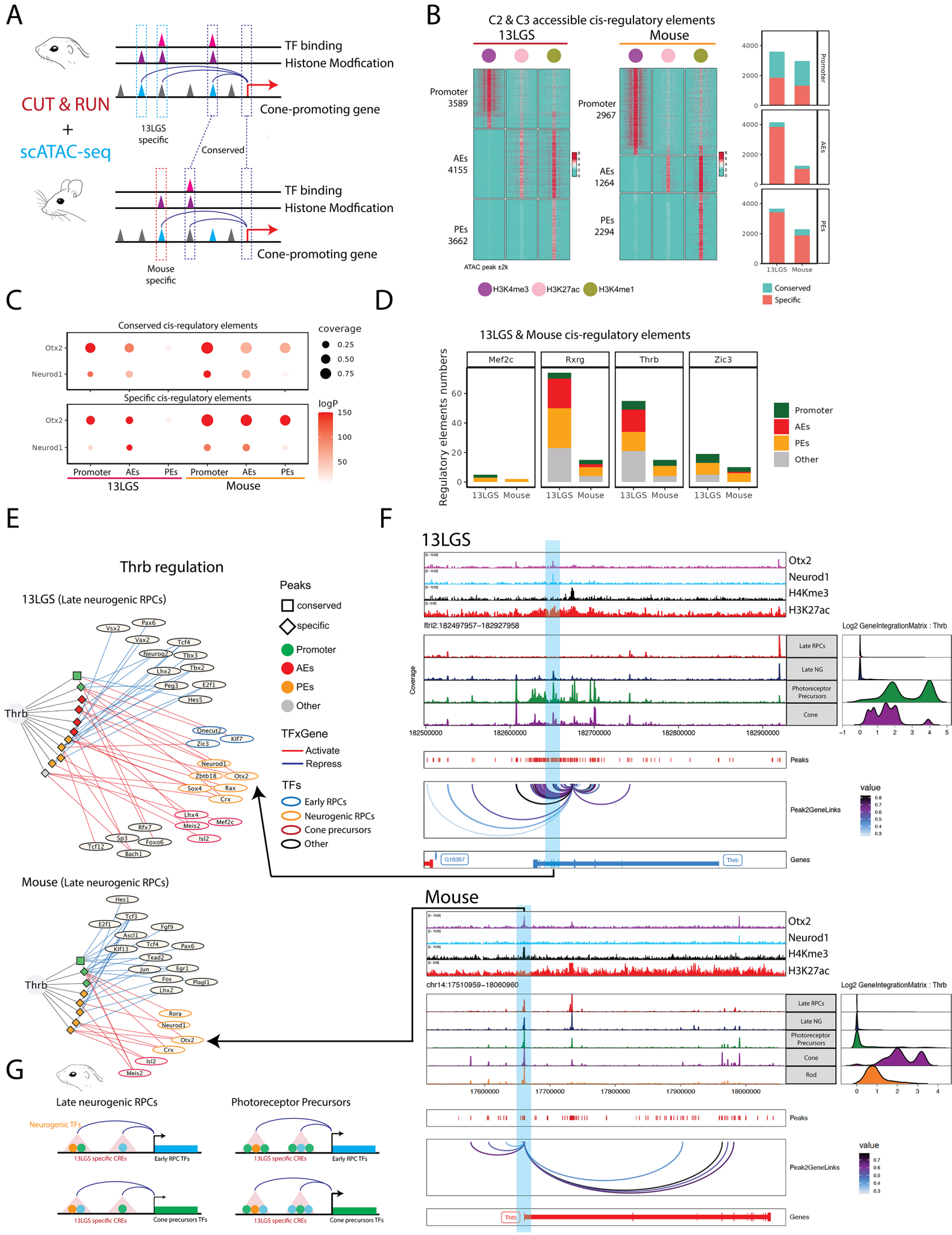
Weir, et al. Fig. 3



Weir, et al. Fig. 4



Weir, et al. Fig. 5



Weir, et al. Fig. 6