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]Solomon H. Snyder 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.

* Co-first authors (contributed equally).
† Corresponding authors.

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 and repress rod photoreceptor-specific gene expression, and act through species-specific regulatory elements that drive their expression in late-stage progenitors. These results demonstrate that 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 (Arendt, 2008; Katz, 2011). 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 (Le Duc and Schöneberg, 2016; Slavenko et al., 2022; Walls, 1942). 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 (Gerkema et al., 2013; Walls, 1942). Despite considerable variation in photoreceptor subtype composition among mammals (Ahnelt and Kolb, 2000; Kaskan et al., 2005; Peichl, 2005), even strongly diurnal species typically have rod-dominant retinas, with rod-to-cone ratios ranging from approximately 5:1 to 20:1 (Mustafi et al., 2009; Peichl, 2005). Strictly nocturnal species, such as mice, exhibit even more pronounced rod dominance of 33:1 (Jeon et al., 1998). 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 (Kryger et al., 1998; Verra et al., 2020), and as many as 50 cones for every rod in the central visual streak (Kryger et al., 1998). 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 (Cepko, 2014; Clark et al., 2019; Lyu et al., 2021; Zhang et al., 2023). 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 (Emerson et al., 2013; Javed et al., 2020). 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 (Cheng et al., 2004; Chen et al., 2005; Daniele et al., 2005; Mears et al., 2001; Milam et al., 2002). 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 (de Melo et al., 2011; Swaroop et al., 2010). 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 development 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)(Mossman and Weisfeldt, 1939). 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) (Clark et al., 2019; Lyu et al., 2021). 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 showed greatest correlation in the same and closely-related cell types across modalities (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). We investigated enrichment of monomeric transcription factor motifs within DARs. Conserved patterns of motif accessibility, identified using ChromVAR and the TRANSFAC2018 database, (Fig. S1F, Table ST1) and transcription factor footprinting (Fig. S1G, H) were observed. Notably, motifs annotated as belonging to *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). Motifs assigned to one transcription factor can have more complicated binding situations *in vivo*, but these enrichments are markedly consistent with previous studies in developing mouse and human retina (Lyu et al., 2021; Thomas et al., 2022).

Direct comparisons of cell type composition across individual time points in the 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 (Clark et al., 2019; Lyu et al., 2021). The expected temporal patterns of neurogenesis were observed in both species: retinal ganglion cells and horizontal cells were generated predominantly in the early stage, amacrine cells in the late embryonic and early postnatal period, and bipolar cells and Müller glia were produced in the late stage (Table ST1). 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. As expected, identification of mature cones in the 13LGS scATAC data preceded the scRNA data, as chromatin opening precedes gene expression ("Chromatin Potential Identified by Shared Single-Cell Profiling of RNA and Chromatin," 2020). In contrast to mice, where rods are distinct from E18 onward (Clark et al., 2019), 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 (Table ST1). By subclustering cones, we can separate distinct clusters of S-cones from the far more numerous M-cones at each postnatal time point, except P21 where there are too few S-cones present to form a separate cluster, and identify genes enriched in each cone subtype (Table ST2) (Kryger et al., 1998). In contrast, in mice, rods progressively increased in proportion, reaching approximately 70% of retinal cells by P8 (Fig. 1D). These proportions are roughly in agreement with what has been found *in vivo* by prior studies using orthogonal approaches (Jeon et al., 1998; Kryger et al., 1998).

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 used Harmony on conserved genes to integrate scRNA- and scATAC-Seq data from all late-stage time points in 13LGS and mice (Fig. 2A, Table ST3).

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, likely due to the very low number of rod cells detected prior to P17 (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). Interestingly, both species exhibited a bipolar cell/photoreceptor precursor population (Fig. 2A,B). This may suggest that late-stage 13LGS retina contain bipolar cell/cone bipotent progenitors similar to the bipolar cell/rod bipotent progenitors of late-stage mouse retina (Clark et al., 2019; Lyu et al., 2021), though further investigation would be required to confirm this. Nonetheless, this highlights the specificity of the effects of heterochronic expansion of cone-promoting factors.

To further characterize species-specific patterns of gene expression and regulation during postnatal 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. Due to their very low number before time point P17, we were unable to include 13LGS rods in the analysis. This analysis revealed eight distinct species-specific differential expression patterns identified by k-means clustering (C1–C8) (Fig. 2C, D; Table ST4).

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 factor *Onecut1* (Emerson et al., 2013; Sapkota et al., 2014) and *Mef2c*, a MADS-box transcription factor which is cone-enriched in human retina (Kallman et al., 2020; Wolf et al., 2017) and reported to be essential for photoreceptor survival (Nagar et al., 2017). 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 (de Melo et al., 2011; Ng et al., 2001; Roberts et al., 2005). Additionally, *Zic3*, which is selectively expressed in early-stage RPCs in mice (Clark et al., 2019; Lyu et al., 2021), was also enriched in C3.

212 Clusters C4–C8 contained genes that were selectively active in mice relative to 13LGS
213 (Fig. 2C). Clusters C4–C6 showed higher expression levels in mice across different cellular
214 contexts: C4 was specific to progenitors, C5 was active throughout specification, and C6
215 spanned progenitors and photoreceptor precursors. C4 included the early-stage RPC
216 transcription factor *Nr2f1*, C5 contained *Sox8* and *Sox11*, and C6 featured *Dlx2*. While the direct
217 relevance of most of these differences remains unclear, *Nr2f1* is essential for establishing the
218 dorsal-ventral gradients in M- and S-cone opsin expression that are observed in mouse cones
219 (Satoh et al., 2009), but which do not exist in 13LGS (Kryger et al., 1998).
220
221 Clusters C7 and C8, which were selectively active in mouse photoreceptor precursors
222 and rods, contained several transcription factors known to promote rod differentiation and
223 repress cone fate. In C7, we identified *Casz1*, a temporal patterning factor that represses the
224 competence of late-stage progenitors to generate cones, while also promoting rod specification
225 (Mattar et al., 2015). In C8, as expected, we observed *Nrl*, *Nr2e3*, and *Samd7*, all of which
226 activate rod-specific genes while repressing cone-specific genes (Cheng et al., 2004; Chen et
227 al., 2005; Hao et al., 2012; Omori et al., 2017).
228
229 Integrated scRNA- and ATAC-Seq analysis using ArchR identified regulatory elements
230 predicted to control each of these genes in both 13LGS and mouse (Fig. 2D). Notably, genes in
231 C2 and C3, which were more highly expressed in 13LGS and are strong candidates for
232 promoting cone specification, exhibited significantly higher numbers of predicted functional cis-
233 regulatory elements than their mouse counterparts (Fig. 2E). This species-specific increase in
234 predicted regulatory elements was either absent from, or far less pronounced in, other gene
235 clusters.
236
237 **Identifying transcription factors that promote cone specification in 13LGS.**
238 To identify transcription factors that strongly contribute to cone specification, we
239 compared differential motif enrichment in accessible chromatin regions across cell types
240 between 13LGS and mice (Fig. 2F, Table ST4). Motifs associated with neurogenic bHLH factors
241 such as Neurog2, Ascl1, Olig2, and Neurod1 were highly enriched in neurogenic RPCs and
242 photoreceptor precursors in both species, as was the motif for Otx2, a key factor required for
243 both rod and cone specification (Emerson et al., 2013; Nishida et al., 2003). Similarly, motifs for
244 transcription factors strongly expressed in late-stage RPCs, such as Lhx2, and Sox8, were
245 prominent in progenitors of both species.
246
247 Motifs for cone-specific transcription factors, including Thrb, Rxrg, and Esrrg were
248 significantly more prominent in cones of 13LGS. Furthermore, in 13LGS, the Zic3 motif was
249 accessible in both progenitors and mature cones, whereas in mice, its accessibility declined in
250 mature cones (Fig. 2F,G). These data suggest that species-specific gene regulatory networks
251 (GRNs) observed in 13LGS late-stage RPCs are characterized by relatively higher activity
252 levels of transcription factors that are enriched in both early-stage RPCs and postmitotic cone
253 precursors in mice contribute to the development of the cone-dominant retina in 13-LGS.
254

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, Table ST5). 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 (Andzelm et al., 2015; Furukawa et al., 1999; Katoh et al., 2010; Nishida et al., 2003; Pennesi et al., 2003), and their expression patterns did not exhibit strong species-specific differences.

In contrast, two broad patterns of differential expression of cone-promoting transcription factors were observed between mouse and 13LGS. First, transcription factors identified in this network that are known to be required for committed cone precursor differentiation, including *Thrb*, *Rxrg*, and *Sall3* (de Melo et al., 2011; Lu et al., 2020; Swaroop et al., 2010), consistently showed stronger expression in both late-stage RPCs and early-stage primary and/or neurogenic RPCs of 13LGS compared to mice. Other genes such as *Mef2c*, which, while not 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 cone differentiation-promoting transcription factors in the course of cone differentiation. Second, transcription factors in the network known to promote cone specification in early-stage mouse RPCs, such as *Onecut2* and *Pou2f1*, exhibited enriched expression in early and late-stage primary and/or neurogenic 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 (Clark et al., 2019; Lyu et al., 2021), supporting its potential role in cone specification. In contrast, genes such as *Casz1*, which act in late neurogenic RPCs to promote rod specification (Mattar et al., 2015), are downregulated in 13LGS late neurogenic RPCs relative to mice.

To confirm heterochronic expression of candidate cone-promoting factors in late-stage developing 13LGS retina, we performed immunohistochemistry (IHC) and RNA hybridization chain reaction (HCR) analysis. 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(Emerson et al., 2013; Javed et al., 2020), with markers of neurogenic RPCs and/or differentiating cones. *Pou2f1* co-expression with *Otx2* was observed at P1, P5, P10, and P24 in 13LGS (Fig. S2A). At all ages examined, *Zic3* was co-expressed with the neurogenic RPC/photoreceptor precursor marker *Otx2* and the cone precursor marker *Rxrg* (Fig. S2B-E). 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. S2F,G).

***Zic3* promotes cone-specific gene expression and is necessary for generating the full complement of cone photoreceptors.**

298 Previous studies have demonstrated that overexpression of either *Onecut1* or *Pou2f1* in
299 late-stage mouse retinal progenitors is sufficient to induce the expression of cone-specific genes
300 while also repressing rod-specific genes (Emerson et al., 2013; Javed et al., 2020). We
301 employed a similar approach to determine whether *ZIC3*, either alone or in combination with the
302 heterochronically expressed cone-promoting factor *POU2F1*, can likewise induce cone
303 differentiation from late-stage mouse RPCs. To test this, we conducted electroporation of P0 *ex*
304 *vivo* mouse retinal explants, as previously described (de Melo and Blackshaw, 2018), and used
305 scRNA-Seq, running different conditions together using CellPlex, to assess the effects of *ZIC3*
306 overexpression alone, or in combination with *POU2F1* (Fig. 3A). ScRNA-Seq analysis of GFP-
307 positive electroporated cells at five days *in vitro* showed that overexpression of
308 heterochronically expressed transcription factors produced a cone-like photoreceptor precursor
309 population (Fig. S3A, Table ST6). Comparison of gene expression between this photoreceptor
310 cluster and rod photoreceptors revealed that empty plasmids induced one differentially-
311 expressed gene: *GFP* (Fig. 3B,C, Table ST6). *ZIC3* not only induced the expression of several
312 cone photoreceptor-specific genes (Uhlén et al., 2015) at this stage, it also downregulated the
313 expression of many rod-specific genes (Fig. 3B,C). *POU2F1* overexpression upregulated an
314 overlapping but distinct, and larger, set of cone-specific genes relative to *ZIC3*, while also
315 downregulating many of the same rod-specific genes, often to a greater extent (Fig. 3C). Co-
316 overexpression of *ZIC3* and *POU2F1* in late-stage RPCs had a synergistic effect, resulting in
317 more differentially expressed genes than did either *ZIC3* or *POU2F1* alone, and greater
318 quantitative changes in cellular levels of expression (Fig. 3B,C, Table ST6).
319
320 *Onecut1* overexpression similarly produced a population of cone-like precursors (Fig
321 S3B, Table ST6). It upregulated a largely distinct set of cone-specific genes compared to
322 *POU2F1* alone but shared some upregulated cone-specific genes with the combined
323 overexpression of *ZIC3* and *POU2F1* (Fig. S3C,D, Table ST6). Furthermore, *Onecut 1*
324 overexpression downregulated many of the same rod-specific genes as did the combined
325 overexpression of *POU2F1* and *ZIC3* (Fig. S3C,D). Similarly to *Onecut 1* and *POU2F1*
326 overexpression, *ZIC3* overexpression produced horizontal cell-like precursors in addition to
327 cone-like precursors (Emerson et al., 2013; Javed et al., 2020) (Fig. S3A,B, Table ST6).
328
329 Because conventional *Zic3* knockouts show widespread and complex developmental
330 defects and high gestational lethality (Purandare et al., 2002), we generated an early RPC-
331 specific *Zic3* loss-of-function (LOF) mutant by combining the retinal progenitor-specific Chx10-
332 CreGFP transgenic line with a conditional *Zic3* allele (Rowan and Cepko, 2004; Sutherland et
333 al., 2013) (Fig. 3D). This resulted in a statistically significant ~20% reduction in the density of
334 cone photoreceptors in the mutant retina (Fig. 3E,F), while the relative numbers of rods and
335 horizontal cells remained unaffected (Fig. S4A-D).
336
337 ScRNA-Seq analysis (n=4) of P14 *Zic3* conditional mutant retinas identified quantitative
338 changes in gene expression in several terminally differentiated cell types (Fig. S4E-G, Table
339 S7). The remaining mutant cones showed reduced expression of *Thrb* but increased expression
340 of multiple cone-specific markers, including *Opn1sw*, *Pde6h*, and *Gngt2* (Fig. S4F). Despite the
341 increased *Opn1sw* expression, there does not seem to be a differential impact on genes

342 enriched in S- or M-cones. For instance, *Pde6h,* which we find to show reduced expression in
343 S-cones relative to M-cones in 13-LGS, is increased in mutant cones and *Pex5l*, which is higher
344 in S-cones, is reduced in mutant cones (Fig. S4F, Table ST7). In addition, we observe a broad
345 decrease in expression of genes expressed at high levels in both cones and rods (*Rpgrip1,*
346 *Drd4*) and rod-specific genes (*Rho*, *Cnga1*, *Pde6b*) in mutant cones (Fig. S4F). Since rods are
347 fragile cells that are located immediately adjacent to cones, some level of contamination of rod-
348 specific genes is inevitable in single-cell analysis of dissociated cones (Clark et al., 2019; Lyu et
349 al., 2021), and this reduced level of rod contamination could result from altered adhesion
350 between mutant rods and cones. In contrast, increased expression of rod-specific genes (*Rho*,
351 *Nrl*, *Pde6g*, *Gngt1*) and pan-photoreceptor genes (*Crx*, *Stx3*, *Rcvrn*) was observed in Müller glia
352 (Fig. S4G), which may likewise result from altered adhesion between Muller glia and rods.
353 Finally, several Müller glia-specific genes were downregulated, including *Clu*, *Aqp4*, and Notch
354 pathway components such as *Hes1* and *Id3*, with the exception of *Hopx*, which was upregulated
355 (Fig. S4G). This likely reflects the indirect effects of *Zic3* loss of function in retinal progenitors.
356 These findings indicate that *Zic3* is essential for the proper expression of photoreceptor genes
357 in cones while also playing a role in regulating expression of Müller glia-specific genes.
358
359 ***MEF2C* overexpression promotes cone-specific gene expression and inhibits rod**
360 **photoreceptor maturation.**
361 We next investigated the effects of both *MEF2C* overexpression in late-stage RPCs and
362 RPC-specific loss of function of *Mef2c* in mice. As described for *ZIC3*, we conducted *ex vivo*
363 electroporation of P0 mouse retinal explants with plasmid vectors expressing *MEF2C*, using
364 empty vectors that expressed GFP alone as controls (Fig. 4A). Since *Mef2c* expression is
365 initiated at later stages of cone development relative to *Zic3* (Fig. 3A, S2A), explants were
366 cultured until P8 for both scRNA-Seq and immunohistochemical analysis.
367
368 ScRNA-Seq analysis revealed substantial alterations in cell type composition following
369 *MEF2C* overexpression, with a marked increase in both neurogenic RPCs and immature
370 photoreceptor precursors, which suggested a broad delay or arrest in rod differentiation (Fig.
371 4B,C Table ST8). Additionally, a novel subpopulation of cone-like precursors emerged (Fig. 4B).
372 Cone-like precursors exhibited upregulation of genes such as *Gnat2*, *Arr3*, and *Pde6c*, while
373 also maintaining expression of genes that are typically enriched in both immature rod and cone
374 photoreceptors, including *Tpi1*, *Rtn4*, and *Tpt1* (Fig. 4C).
375
376 Histological analysis indicated that *MEF2C* overexpression consistently led to
377 morphological abnormalities in electroporated regions, including the formation of whorls and
378 rosettes in the outer nuclear layer, a phenotype reminiscent of *Nrl* and *Nr2e3* mutants that
379 exhibit ectopic cone-specific gene expression in rods (Fig. 4D, Fig. S5A,B) (Akhmedov et al.,
380 2000; Mears et al., 2001). Although a few GFP-positive electroporated cells co-expressing the
381 cone-specific marker *Gnat2* were detected in control (likely due to the electroporation of cone
382 precursors, which we have previously observed in P0 retinal explants (Clark et al., 2019; Leavey
383 et al., 2025; Lyu et al., 2021; Onishi et al., 2010)), there was a significant increase in double-
384 positive cells in the test condition, matching the novel cone-like precursor population found in
385 the scRNA-Seq dataset (Fig. 4E). No obvious reduction in the relative number of Nrl-positive

386 cells was observed (Fig. S5A). No changes were observed in the distribution or abundance of
387 other retinal cell types (Fig. S5B-D).
388
389 A previous study reported that *Nestin-Cre;Mef2c*$^{lox/lox}$ mice, which are predicted to exhibit
390 loss of function of *Mef2c* throughout the developing central nervous system, showed a rapid
391 degeneration of both rod and cone photoreceptors, that resembled photoreceptor loss observed
392 in fast-acting *rd1/rd1* mutants (Nagar et al., 2017). To evaluate the effects of retinal progenitor-
393 specific loss of function of *Mef2c*, we generated *Chx10-Cre;Mef2c*$^{lox/lox}$ mice (Rowan and Cepko,
394 2004; Vong et al., 2005) (Fig. S6A). No changes in cone numbers were observed at P14 (Fig.
395 S6B,C), nor was there any loss of rod or cone photoreceptors through P180 (data not shown).
396 These findings suggest that *Mef2c* is dispensable for cone development in mice, and that
397 previous reports of rapid photoreceptor degeneration may have resulted from inadvertent
398 contamination with either the *rd1* allele itself or another common retinal dystrophy mutation
399 (Chang et al., 2002).
400
401 **Species-specific *cis*-regulatory elements activate expression of cone-promoting**
402 **genes in late-stage RPCs of 13LGS.**
403 These findings demonstrate that, relative to mice, late-stage RPCs and photoreceptor
404 precursors of 13LGS exhibit significantly elevated expression of multiple transcription factors
405 that promote cone differentiation while inhibiting rod development. To investigate the
406 mechanism underlying this species-specific expansion of expression, we performed bulk
407 CUT&RUN analysis in P5 13LGS retinas and P2 mouse retinas, analyzing stage-matched
408 samples highly enriched for late-stage RPCs and photoreceptor precursors. We profiled histone
409 modifications associated with active promoters and both active and poised enhancers, as well
410 as the photoreceptor-promoting transcription factors Otx2 and Neurod1. This data was
411 integrated with scATAC-Seq-derived chromatin accessibility regions from late-stage RPCs and
412 photoreceptor precursors (Fig. 5A, Table ST9).
413
414 We examined the regulatory elements of genes that exhibited specifically elevated
415 expression in late-stage progenitors and photoreceptor precursors of 13LGS, which were
416 identified in clusters C2 and C3 of the analysis described in Figure 2D. Integration of scATAC-
417 Seq and CUT&RUN data showed that 13LGS consistently harbors both greater numbers of
418 accessible chromatin regions associated with markers of both active (H3K27Ac-positive)
419 enhancers and poised (H3K4me1-positive) enhancers for these genes. However, most of these
420 elements mapped to sequences that were not shared between 13LGS and mouse, with
421 intergenic enhancers exhibiting particularly low levels of conservation (Fig. 5B). In contrast,
422 H3K4me3-positive promoter elements were relatively well-conserved and were found in similar
423 numbers in both species (Fig. 5B).
424
425 Though conserved and species-specific enhancers showed enrichment for binding sites
426 for the broad photoreceptor precursor transcription factors Otx2 and Neurod1 profiled by
427 CUT&RUN, transcription factors known to regulate cone differentiation— including *Mef2c*, *Rxrg*,
428 *Thrb* and *Zic3* —demonstrated far greater motif enrichment in active regulatory elements in

429 13LGS than in mice, though few of these elements mapped to sequences that were shared
430 between 13LGS and mouse (Fig. 5C,D, Table ST10).
431
432 Enhancer elements associated with cone-specific genes in 13LGS are predicted to be
433 directly targeted by a considerably greater number of transcription factors in late-stage
434 neurogenic RPCs than in mice, as might be expected, given the higher expression levels of
435 these genes. This is particularly evident in enhancers associated with *Thrb* in both species,
436 which are shown to be directly targeted by Otx2 (Fig. 5E,F). While the *Thrb* promoter is both
437 accessible and bound by Otx2 in mouse photoreceptor precursors, it is predicted to be targeted
438 by significantly fewer transcription factors compared to the active elements in 13LGS (Fig.
439 5E,F). In 13LGS, we identified a pair of species-specific enhancers that are active in late-stage
440 neurogenic RPCs, photoreceptor precursors, and mature cones. These enhancers are predicted
441 to be directly targeted by a combination of factors, including those broadly expressed in
442 developing rods and cones (Otx2, Crx, Neurod1, Prdm1); cone-promoting factors that are
443 restricted to early-stage RPCs in mice (Zic3, Onecut2, Pou2f1); and by transcription factors that
444 selectively promote cone precursor differentiation (Rxrg, Mef2c). The gene regulatory networks
445 (GRNs) driving cone specification in 13LGS become progressively more stable over time, as
446 additional transcription factors are recruited and multiple layers of positive regulation are
447 incorporated (Fig. S7A-D, S8A-C, S9A-C).
448
449 We conclude that the development of the cone-dominant retina in 13LGS is driven by
450 novel cis-regulatory elements that promote heterochronic expression of cone-promoting and
451 rod-inhibiting transcription factors that are restricted to either early-stage RPCs or postmitotic
452 cone photoreceptor precursors in mice. These *cis*-regulatory elements are activated by a
453 combination of these factors themselves, in conjunction with transcription factors such as Otx2
454 and Neurod1, that are expressed in both developing rods and cones. The sustained expression
455 of these factors is maintained by an intricate and progressively expanding network of cross-
456 activating regulatory interactions (Fig. 5G).
457
458 **Discussion**
459 This study represents the first comprehensive molecular analysis of gene regulatory
460 networks that controls the development of a cone-dominant retina derived from a rod-dominant
461 ancestral state (Gerkema et al., 2013; Walls, 1942). Using integrated scRNA- and scATAC-Seq
462 analyses, we systematically examined developmental changes in gene expression and
463 regulation in the cone-dominant 13-lined ground squirrel (13LGS) retina, and directly compared
464 these findings to stage-matched data from the rod-dominant mouse. Though BrdU/EdU labeling
465 would be required to unambiguously demonstrate species-specific differences in birthdating, our
466 findings strongly indicate that 13LGS exhibit a selective expansion of the temporal window of
467 cone generation, extending into late stages of neurogenesis. This is mediated by the
468 heterochronic expression of a subset of transcription factors in late-stage progenitors that both
469 promote cone and repress rod specification. In addition to previously characterized cone-
470 promoting and rod-inhibiting factors such as *Onecut2* and *Pou2f1*, we identified *Zic3*, which is
471 selectively expressed in early-stage RPCs in mice, as also contributing to cone specification.
472 Late-stage 13LGS RPCs also exhibit precocious expression of transcription factors such as

473 *Thrb*, *Rxrg*, and *Sall3*, which in mice function post-mitotically to promote cone differentiation,
474 and in many cases also act in parallel to inhibit rod differentiation (de Melo et al., 2011; Swaroop
475 et al., 2010).
476
477 Transcription of these genes in late-stage RPCs is controlled by 13LGS-specific
478 enhancer elements, which are directly targeted by transcription factors that are broadly
479 expressed in both early- and late-stage RPCs, such as Otx2, as well as by heterochronically
480 and precociously expressed cone-promoting transcription factors through a dynamic process of
481 positive feedback and cross-activation. The gene regulatory network (GRN) governing cone
482 specification becomes progressively more complex, incorporating additional transcription factors
483 over time. These factors include genes expressed in both rod and cone precursors, such as
484 Neurod1 and Prdm1; genes known to promote cone precursor differentiation, such as Sall3; and
485 previously uncharacterized regulators such as Mef2c, which we show can both activate cone-
486 specific genes and repress rod-specific genes.
487
488 The evolution of novel morphological features, such as bat wings or the elongated legs
489 of jerboas (Cretekos et al., 2008; Saxena et al., 2022), is often driven by the emergence of
490 lineage-specific cis-regulatory elements, which confer novel spatiotemporal patterns of
491 expression to structurally conserved transcription factors or secreted patterning regulators.
492 Based on our bioinformatic analysis, the cone-dominant 13LGS retina follows this paradigm, in
493 which species-specific enhancer elements, identified by the overlap of species-specific
494 H3K27Ac peaks and cell type-specific chromatin accessibility, are used to drive both the ectopic
495 heterochronic expression of transcription factors that regulate neural progenitor temporal
496 identity and precocious expression in retinal progenitors of transcription factors that normally act
497 to promote cone differentiation in postmitotic precursor cells. These in turn promote cone, and
498 repress rod, photoreceptor specification. The highly interconnected nature of the GRN
499 controlling 13LGS cone specification should make it robust, limiting the impact of the loss of any
500 single enhancer element on cone production. Similarly, the GRN's highly derived nature, with
501 TF binding combinations not seen in rod-dominant retinas, means that enhancers that are
502 functional in the 13LGS cannot necessarily be expected to behave similarly in mice. In all
503 vertebrates examined to date, including 13LGS, separate and mutually cross-repressive GRNs
504 specify early- and late-stage temporal identity in retinal progenitors (Cepko, 2014; Clark et al.,
505 2019; Lyu et al., 2021). In mice, cone generation occurs only in progenitors in which the early-
506 stage network is active (Emerson et al., 2013; Javed et al., 2020; Lonfat et al., 2021), whereas
507 in 13LGS, multiple transcription factors from the early-stage network persist in late-stage RPCs.
508
509 This raises the question of whether similar heterochronic shifts in key temporal identity
510 regulators may underlie other cases of rapid evolutionary changes in the rod-cone ratio.
511 Examples in mammals include the development of the cone-dominant retina in primate-like tree
512 shrews and the formation of the rod-free foveola in the primate fovea (Müller and Peichl, 1989;
513 Yuodelis and Hendrickson, 1986). However, rapid changes in the rod-cone ratio are observed
514 across vertebrate classes (Lamb, 2013; Peichl, 2005). Significant differences in the relative
515 abundance of other retinal cell types are also observed among species, often correlated with
516 temporal niche adaptation (Hahn et al., 2023). Moreover, shifts in the relative ratios of neuronal

517 subtypes in the central nervous system can underlie species-specific behavioral, cognitive, and
518 sensory processes between closely-related vertebrate species (Chen et al., 2025; Fernald,
519 2012; Jacobs and Spencer, 1994; Wada et al., 2006; Yang et al., 1996; Young and Wang,
520 2004). These variations may also stem from species-specific heterochronic expression of
521 transcription factors that control temporal identity in the developing brain.
522
523 By performing an unbiased and comprehensive analysis of gene regulatory networks
524 that are selectively active in late-stage RPCs of 13LGS, we have identified numerous previously
525 uncharacterized factors that are strong candidates for promoting cone specification. In line with
526 previous studies of *Onecut1* and *Pou2f1* (Emerson et al., 2013; Javed et al., 2020), we find that
527 these factors each activate expression of a subset of cone-specific genes while broadly and
528 potently repress expression of rod-specific genes. Working in mice, we demonstrate that *Zic3*
529 and *Mef2c* behave similarly, activating expression of partially overlapping sets of cone-specific
530 genes and strongly repressing rod differentiation. Indeed, overexpression of *Mef2c* increased
531 the number of both neurogenic RPCs and immature photoreceptor precursors, suggesting that
532 rod differentiation was delayed. Many additional candidates, such as *Hivep2* or *Isl2*, remain to
533 be tested and are likely to serve similar functions. Our analysis suggests that gene regulatory
534 networks controlling cone specification are highly redundant, with transcription factors acting in
535 complex, redundant, and potentially synergistic combinations. This is further supported by our
536 findings on the synergistic effects of combined overexpression of *ZIC3* and *POU21* increasing
537 both the number of differentially expressed genes and their level of change in expression
538 relative to the modest changes seen with overexpression of either gene alone (Fig. 3) and the
539 relatively mild or undetectable phenotypes observed following loss of function of *Zic3* and *Mef2c*
540 (Fig. 3, Fig. S6), as well as other cone-promoting factors such as *Onecut1* and *Pou2f1*
541 (Emerson et al., 2013; Javed et al., 2020).
542
543 These findings have potentially significant clinical implications, as the loss of cone
544 photoreceptors is the primary trigger for irreversible blindness in hereditary diseases such as
545 retinitis pigmentosa and age-related macular degeneration. Effective strategies for driving cone
546 specification and differentiation are essential for developing cell-based therapies to replace
547 photoreceptors lost to degenerative disease, and the findings of comparative studies like ours
548 can help identify these mechanisms.
549

550
551 **Methods:**
552
553 **Mouse retinal development scRNA-Seq and scATAC-Seq analysis**
554 Raw data and metadata for the mouse scRNA-Seq and scATAC-Seq datasets (Clark et al.,
555 2019; Lyu et al., 2021) were downloaded from the Gene Expression Omnibus (accession
556 numbers GSE118614, GSE181251) and analyzed following the Seurat SCT and Signac
557 workflows (Hafemeister and Satija, 2019; Hao et al., 2021; Stuart et al., 2021). Cell type
558 classifications and ages were taken from the downloaded metadata.
559
560 **Sample collection**
561 The use of animals for these studies was conducted using protocols approved by the Johns
562 Hopkins Animal Care and Use Committee, the UW Oshkosh Animal Care and Use Committee,
563 and the Northwestern University Animal Care and Use Committee, in compliance with ARRIVE
564 guidelines and the ARVO Statement for the Use of Animals in Ophthalmic and Vision Research,
565 and were performed in accordance with relevant guidelines and regulations.
566
567 Timed pregnant CD1 mice were ordered from Charles River Laboratories for the postnatal
568 retinal explant plasmid overexpression analysis. A pair of *Zic3*$^{+/lox}$ (Jiang et al., 2013)(JAX stock
569 #023162) female mice and a pair of *Mef2c*$^{+/lox}$ mice (Vong et al., 2005)(JAX stock #025556)
570 were ordered from the Jackson Laboratory and outbred to *Chx10Cre-GFP* CD1 mice (Rowan
571 and Cepko, 2004). The progeny were bred to produce the *Zic3*$^{lox/lox}$ (or *Zic3*$^{lox/Y}$); *Chx10Cre-*
572 *GFP*$^{+/-}$, *Zic3*$^{+/+}$ (or *Zic3*$^{+/Y}$); *Chx10Cre-GFP*$^{+/wt}$, *Mef2c*$^{lox/lox}$; *Chx10Cre-GFP*$^{+wt}$, and *Mef2c*$^{lox/lox}$
573 used in this study. Pups were grown until postnatal day 14 before retinal samples were taken.
574 Mice were euthanized by cervical dislocation.
575
576 Embryonic 13-lined ground squirrel (13LGS) retinal samples were collected from timed-pregnant
577 captive dams and embryonic day was confirmed by comparison with mouse Theiler stages.
578 Postnatal 13LGS retinal samples were collected from captive-bred litters with day of birth set as
579 Postnatal Day 0. 13LGS were euthanized by decapitation.
580
581 **Sample preservation**
582 Mouse retinas and retinal explants used for immunostaining were fixed in freshly prepared 4%
583 paraformaldehyde in PBS at 4°C for 1 hour (ZIC3) or 2 hours (MEF2C), then briefly washed with
584 PBS for 5 mins. After the wash, the explants or retinas were transferred to 30% sucrose in PBS
585 at 4 °C overnight. The dehydrated samples were then embedded in VWR Clear Frozen Section
586 Compound, frozen on dry ice, then stored at -80°C.
587
588 For single-cell analysis, 13LGS retinas were dissected and either methanol-fixed or flash-frozen
589 as previously shown (Weir et al., 2021). Samples were shipped to Johns Hopkins on dry ice and
590 stored at -80°C.
591
592 For IHC and RNA-HCR experiments, postnatal 13LGS whole eye cups were fixed in 4%
593 paraformaldehyde (PFA; 157-8, Electron Microscopy Sciences, USA) made in 1X phosphate

594 buffered saline (PBS) after the removal of lens, cornea and vitreous humor and processed as
595 previously (Kandoi et al., 2022; Yu et al., 2024). PFA-fixed eye cups were washed three times
596 with 1X PBS at 15 min intervals each, followed by rehydration with 15% and 30% sucrose
597 (S1888-5KG, Sigma Life Science, USA) solutions (made in 1X PBS) until the eyecups had sunk
598 to the bottom of the tube. The whole cups were then equilibrated, embedded and frozen on dry
599 ice in a sucrose:OCT (4583, Sakura Finetek, Torrance, CA) mixture (2:1) and stored at -80ºC.
600
601 For CUT&RUN experiments, retinas were dissected out from postnatal 13LGS after the removal
602 of lens, cornea and vitreous humor. Dissected retinas were immediately flash frozen using liquid
603 nitrogen and shipped to Johns Hopkins University using dry ice and stored at -80°C.
604
605 **Retinal explant culture**
606 Retinas were dissected from CD1 mice at Postnatal Day 0 in warm PBS, then transferred to a
607 micro electroporation chamber. The retinas were electroporated with plasmid solution at 30 V of
608 50-ms duration with 950-ms intervals (Matsuda and Cepko, 2008) using either a pCAGIG
609 plasmid that expresses IRES-GFP (Matsuda and Cepko, 2004; Onishi et al., 2009) or a test
610 plasmid that expresses the gene of interest as a bicistronic transcript together with IRES-GFP
611 (Matsuda and Cepko, 2004). The coding sequences for the transcription factors overexpressed
612 in these IRES-GFP constructs correspond to the following accession numbers: human *ZIC3*,
613 NM_003413; human *MEF2C*, XM_011509654; mouse *Onecut1*, NM_008262; human *POU2F1*,
614 NM_002397.
615
616 Following harvesting, retinas were flattened by a series of radial cuts and mounted on 0.2 µm
617 Nucleopore Track-Etch membranes on culture media (2 ml DMEM F12, 10% FBS with 0.1%
618 streptomycin/puromycin) in 12-well plates (Tabata et al., 2004). In experiments that included a
619 condition where two genes of interest were overexpressed in the same retinal explants, the
620 concentration of plasmids was kept consistent across conditions by supplementing with empty
621 plasmids. For instance, empty condition: 100% control plasmid, condition A: 50% gene A
622 plasmid and 50% control plasmid, condition B: 50% gene A plasmid and 50% control plasmid,
623 combined condition: 50% gene A plasmid and 50% gene B plasmid.
624
625 Explants were incubated at 37 °C, 5% CO_2 with the media changed every 72 hours. Two-thirds
626 of the old media (1/2 old media in *MEF2C* overexpression) in each well was replaced by slowly
627 adding new media to the well's wall, avoiding disturbing the membranes. The explants are
628 harvested at P5 (*ZIC3*) or P8 (*MEF2C*) for single-cell RNA sequencing and P8 (*MEF2C*) for
629 immunostaining.
630
631 **Retinal sectioning**
632 Mouse embedded samples were sectioned with Leica 3050S cryostat at a thickness of 16 µm
633 and collected on clean Superfrost Plus slides (48311-703, VWR, Radnor, PA). Sections were
634 then stored at -20°C. For postnatal 13LGS, 8-10µm thick sections were collected and stored at -
635 80ºC until use for histological assessments.
636
637 **Fluorescence-Activated Cell Sorting and multiplexing**

638 Dissociated single cells from electroporated retinal explants were sorted based on GFP
639 expression on a Sony SH800S cell sorter. GFP+ cells from each condition were labeled and
640 pooled into a single cell suspension according to the 10x Cell Multiplexing Oligo Labeling for
641 Single Cell RNA Sequencing Protocols with Feature Barcode technology protocol (Rev B) prior
642 to GEM generation and library prep.
643
644 **Single-cell library preparation**
645 Retina and explants were dissociated, methanol-fixed cells rehydrated, and nuclei isolated from
646 flash-frozen samples as previously described (Weir et al., 2021). GEM generation and library
647 preparation for scRNA-Seq were completed according to the 10x 3' v3.1 user guide (Rev D)
648 using the Chromium Next GEM Single Cell 3' Kit v3.1 and Dual Index Kit TT Set A and for
649 Cellplex samples according to the 10x Next GEM Single Cell 3' Reagent Kits v3.1 (Dual Index)
650 with Feature Barcode technology for Cell Multiplexing (Rev A) using the 3' CellPlex Kit Set A
651 and Dual Index Kit NN Set A. Transposition, GEM generation, and library preparation for
652 scATAC-Seq were completed according to the 10x ATAC v1.1 user guide (Rev D) using the
653 Chromium Next GEM Single Cell ATAC Reagent Kits v1.1 and Single Index Kit N, Set A. Quality
654 control and sequencing were performed by the Johns Hopkins Transcriptomics and Deep
655 Sequencing Core.
656
657 **ScRNA-Seq and scATAC-Seq analysis**
658 Sequencing results for 13LGS and mouse scRNA-Seq were processed through the 10x
659 CellRanger pipeline (version 3.1.0), using the 13LGS genome GCF_016881025.1_HiC_Itri_2
660 (Fu et al., 2020), and the output loaded into Seurat (version 4.0.4). Cells from Cellplex-labeled
661 libraries were deconvoluted using the CellRanger multi pipeline (version 6.0). ScRNA-Seq
662 Seurat objects were analyzed following the Seurat SCT workflow (Hafemeister and Satija, 2019;
663 Hao et al., 2021). Cell type was determined cluster by cluster based on marker gene expression
664 derived from the Seurat function 'FindAllMarkers' using default parameters. Differential gene
665 expression between Cone-like Precursors and Rods was identified using the Seurat function
666 "FindMarkers". ScRNA-Seq samples were aggregated using the CellRanger aggr pipeline
667 (version 3.1.0).
668
669 Sequencing results for 13LGS scATAC-Seq data were processed through the 10x CellRanger
670 ATAC pipeline (version 1.2.0) and the output loaded into Signac (version 1.4.0) and analyzed
671 according to the Signac workflow. Cell type was determined cluster by cluster based on marker
672 gene expression derived from the Seurat function 'FindAllMarkers' using default parameters and
673 based on transfer learning from the scRNA-Seq samples following the Signac workflow using
674 default parameters. ScATAC-Seq samples were integrated as described (Lyu et al., 2021).
675
676 **Immunostaining and visualization**
677 Mouse retinal sections were stained with rabbit anti-Arr3 polyclonal IgG (Millipore, No.
678 AB15282, 1:200) or rabbit anti-Gnat2 polyclonal IgG (Invitrogen, PA5-119558, 1:200); chicken
679 anti-GFP polyclonal IgY (Invitrogen, A10262, 1:400); and mouse anti-Nr2e3 monoclonal IgG
680 (R&D systems, PP-H7223-00, 1:200), goat anti-Nrl polyclonal IgG (R&D Systems, AF2945,
681 1:200), goat anti-Otx2 polyclonal IgG (R&D Systems, AF1979, 1:400), rabbit anti-Sox9

682 polyclonal (Millipore, AB5535, 1:400), mouse anti-Tfap2a monoclonal (DSHB, Cat# 3b5,
683 RRID:AB_528084, 1:400), or mouse anti-Lhx1 (DHSB,Cat#4F2, concentrated,) primary
684 antibodies. For the secondary staining, CF633 donkey anti-rabbit IgG polyclonal (Sigma,1:500);
685 Alexa Fluor™488 goat anti-chicken IgY polyclonal (Invitrogen,1:500), and Alexa Fluor™ TM 568
686 donkey anti-mouse IgG polyclonal (Invitrogen,1:500) secondary antibodies were used. DAPI
687 was used to counterstain nuclei (1:5000). Stained sections were imaged with a Zeiss
688 AxioObserver with LSM700 confocal module at 20X resolution within a fixed scanning area of
689 319.87 nm^2. Z-stack images were taken for the stained sections and cell counting was
690 performed blinded on Z-projections.
691
692 For 13LGS immunohistochemistry experiments, retinal sections were thawed, rehydrated three
693 times with 1X PBS at 15 min intervals and stained *via* IHC for identification of retinal cell types.
694 For IHC, sections were permeabilized for 15 mins at room temperature with 0.1% Triton X-100
695 (0694-1L, VWR Life Science, USA) made in 10% normal donkey serum in 1X PBS (NDS, S30-
696 100ML, EMD Millipore Corp, USA). The sections were then blocked for 1 hour at room
697 temperature using 10% normal donkey serum and incubated at 4°C overnight with primary
698 antibodies diluted in 10% normal donkey serum. Primary antibodies utilized for staining 13LGS
699 sections included goat anti-Otx2 polyclonal IgG (R&D Systems, BAF1979, 1:400) and rabbit
700 anti-Mef2c polyclonal (Atlas Antibodies, HPA005533, 1:50). Next day, primary antibodies were
701 removed, sections were washed three times in 1X PBS at room temperature with 15 min interval
702 each and fluorescent conjugated secondary antibodies (donkey anti-mouse IgG Alexa Fluor™
703 Plus 647, ThermoFisher Scientific, A-11055, 1:250; donkey anti-rabbit IgG Alexa Fluor™ 488,
704 ThermoFisher Scientific, A-21206, 1:250, and donkey anti-goat IgG Alexa Fluor™ 555,
705 ThermoFisher Scientific, A-21432, 1:250) diluted in 10% normal donkey serum were applied to
706 sections and incubated in the dark for 1 hour. Cell nuclei were counterstained with DAPI (1
707 µg/mL, Roche, 10236276001) for 10 min. The slides were then washed three times in 1X PBS
708 at 5 min intervals and coverslipped using Fluoromount G (17984-25, Electron Microscopy
709 Sciences, USA). Images were acquired using ZEN software on LSM700 confocal microscope
710 (Zeiss, Inc) at 63X oil resolution and processed using ImageJ (Schindelin et al., 2012).
711
712 **RNA-Hybridization Chain Reaction (RNA-HCR)**
713 The samples were prepared by thawing and dehydrating the 13LGS retinal sections by
714 immersion into a series of ethanol (EtOH) solution - 50% EtOH, 70% EtOH, 100% EtOH for 5
715 mins each. The slides were immersed into another fresh 100% EtOH for 5 mins and allowed to
716 air dry for 5 mins at room temperature. 10 µg/µL of Proteinase K solution (3001-2-60, Zymo
717 Research) were applied on the sections and the slides incubated in a humidified chamber for 10
718 mins at 37°C. The slides were washed by immersion in fresh 1X PBS twice. The slides were
719 dried by blotting the edges on a Kimwipe. The sections were pre-hybridized in a humidified
720 chamber pre-warmed to 37°C for 10 mins after adding 200 µL of probe hybridization buffer
721 (240904, HCR™ Buffer sets, Molecular Instruments). Buffer was removed and 50 - 100 µL of
722 DNA probe solutions (1.6 pmol of each probe diluted in probe hybridization buffer) were added
723 to the sections. DNA probes were custom designed using in situ probe generator v0.3.2 at the
724 visual studio code platform for the 13-LGS gene targets as shown in Figure 3 (60 oligos/target
725 at 50 pmol per oligos), obtained from IDT, and reconstituted as 1 µM stock in nuclease free

726 water. Sections with probe solution were covered with coverslip and incubated overnight (~18-
727 24 h) at 37°C in a humidified chamber. Next day, the slides were immersed in the probe wash
728 buffer at 37°C to float off the coverslip. Excess probes were removed by incubating the slides at
729 37°C with pre-heated probe wash buffer (240904, HCR™ Buffer sets, Molecular Instruments)
730 and 5X SSCT solutions (S6639-1L, Sigma-Aldrich): 75% probe wash buffer/25% 5X SSCT for
731 15 mins; 50% probe wash buffer/50% 5X SSCT for 15 mins; 25% probe wash buffer/75% 5X
732 SSCT for 15 mins and 100% 5X SSCT for 15 mins. The slides were immersed in 5X SSCT for 5
733 mins at room temperature. The slides were dried by blotting the edges on a Kimwipe. The HCR
734 was pre-amplified by adding 200 μL of amplification buffer (240904, HCR™ Buffer sets,
735 Molecular Instruments) on top of the sections and incubating the slides for 30 mins at room
736 temperature. The solution was removed and 50 - 100 μl of hairpin solution added before
737 incubating overnight (12-18 h) in a dark humidified chamber at room temperature. Hairpin
738 solution (Fluorophore labeled for 488/546/647) was prepared by heating (95°C for 90 seconds)
739 and snap-cooling (in a dark drawer at room temperature for 30 mins) 2 μL of 3 μM stocks of h1
740 and h2 hairpins (6 pmol each) to 100 uL of amplification buffer at room temperature. Next day,
741 the slides were immersed in a 5X SSCT buffer at room temperature to float off the coverslip and
742 counterstained with DAPI (1 µg/mL, Roche, 10236276001) for 10 mins, at dark in room
743 temperature. Excessive hair pins were removed by incubating the slides in a 5X SSCT buffer at
744 room temperature thrice: 2X for 30 mins and 1X for 5 mins. The slides were coverslipped using
745 Fluoromount G (17984-25, Electron Microscopy Sciences, USA). Images were acquired using
746 ZEN software on an LSM700 confocal microscope (Zeiss, Inc) at 63X oil resolution and
747 processed using ImageJ (NIH, USA).
748
749 **CUT&RUN library preparation**
750 Nuclei were isolated from snap-frozen P5 13LGS and P2 mouse retinas for use in CUT&RUN
751 analysis. Two biological replicates were assayed for each species, with each 13LGS replicate
752 consisting of a pool of retinas from two individuals, and each mouse replicate consisting of a
753 pool of retinas from 3 individuals. Nuclei were isolated according to the described protocol
754 (Tangeman et al., 2025), using a lysis buffer modified to contain 0.01% Tween-20 and 0.01%
755 Nonidet P40, as well as 1X protease inhibitor cocktail (Roche, 11873580001) in all solutions.
756 Samples were lysed for 5 minutes in the lysis buffer with gentle trituration throughout, as
757 described. CUT&RUN was performed using the CUTANA™ ChIC/CUT&RUN Kit (Epicypher,
758 14-1048). For each reaction, either 500,000 13LGS nuclei or 385,000 mouse nuclei were used.
759 The CUTANA protocol v4.0 protocol was followed with minor modifications. In brief, isolated
760 nuclei were resuspended in CUT&RUN wash buffer and added directly to Concanavalin A
761 beads. Then 0.5 µg of the following antibodies were added to each respective reaction: IgG
762 control (EpiCypher, 13-0042), H3K4me1 (EpiCypher, 13-0057), H3K4me3 (EpiCypher, 13-
763 0041), H3K27ac (EpiCypher, 13-0059), H3K27me3 (EpiCypher, 13-0055), NeuroD1 (Cell
764 Signaling, 62953), Otx2 (R&D Systems, BAF1979), or Otx2 (Atlas Antibodies, HPA000633).
765 H3K9me3 (Abcam, ab8898) was used at 0.41 µg of antibody per reaction. The total CUT&RUN
766 DNA yield, up to 5 ng of total purified DNA, was used for library preparation, using reagents
767 from CUTANA™ CUT&RUN Library Prep Kit (EpiCypher, 14-1001) or NEBNext® Ultra™ II DNA
768 Library Prep Kit for Illumina® (NEB, E7645S). The EpiCypher CUTANA Library Prep Kit User
769 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 (Martin, 2011) and Trim Galore v0.6.4_dev (Krueger et al., 2023), 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 (Langmead and Salzberg, 2012) 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 (Danecek et al., 2021). Aligned reads were deduplicated using UMI-tools (Smith et al., 2017) 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 (Ramírez et al., 2016) 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 (Zhang et al., 2008) 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 (Amemiya et al., 2019) 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) (Granja et al., 2021). 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 (Alicia Schep [aut, cre], Stanford University [cph], 2017) (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 ("Website," n.d.)). 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 (doi: 10.1038/nmeth.4463). 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 Top 30 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

902 elements, and the average signal intensity for each histone mark was calculated within their
903 regions. A cis-regulatory element was considered to exhibit a histone modification if it both
904 overlapped with a histone peak and showed an average signal greater than 2.5. Finally, cis-
905 regulatory elements exhibiting H3K4me3+ were classified as TSS, those with H3K4me3– and
906 H3K27ac+ as aEnhancer, and those with H3K4me1+ as pEnhancer.
907
908 **Identification of conserved regulatory elements**
909 Conserved regulatory elements between 13LGS and mouse were identified by comparing
910 ATAC-Seq peaks (all potential cis-regulatory elements), as well as putative regulatory elements
911 associated with promoters, active, and poised enhancers that were associated with conserved
912 gene pairs. For each conserved gene pair, candidate regulatory peaks were collected from both
913 species and paired one-to-one, with each pair comprising one peak from 13LGS and one peak
914 from mouse. Subsequently, the sequence of each 13LGS peak and the forward and reverse
915 complement sequences for each mouse peak were extracted using the Biostrings package in R.
916 Similarity scores for each candidate peak pair were computed with the “pairwiseAlignment”
917 function using local alignment parameters (type = "local", gapOpening = -2, gapExtension = -1).
918 To determine the cutoff for conserved peak pairs, a normal model was fitted to all similarity
919 scores using the MASS package’s fitdistr function, and a threshold corresponding to a p-value <
920 0.05 under the fitted normal distribution was defined. Peak pairs with scores exceeding this
921 threshold were considered to be conserved.
922
923 **ChromVAR analysis**
924 Global transcription factor (TF) activity at the single-cell level was evaluated using ChromVAR
925 (Schep et al., 2017). The analysis began with the raw cell-by-peak matrix, which was processed
926 with the addGCBias function to adjust for GC content biases. Next, motif information from the
927 TransFac2018 database was incorporated to generate a TF z-score matrix via the matchmotifs
928 and computeDeviations functions. Finally, these z-scores were visualized as heatmaps to
929 illustrate TF activity patterns across the cell population.
930
931 To identify cell type-specific motifs in 13LGS, the z-score matrix generated from ChromVAR
932 was converted into a Seurat object. The FindAllMarkers function from the Signac package was
933 then applied with the following parameters: only.pos = TRUE, mean.fxn = rowMeans, and
934 fc.name = "avg_diff". Finally, motifs with an adjusted p-value < 0.01 and an average difference
935 (avg_diff) > 2 were identified as cell type-specific.
936
937 To identify differential motifs associated with photoreceptor development between 13LGS and
938 mouse, the ChromVAR-generated z-score matrix was utilized. The following pairwise
939 comparisons were performed: 13LGS_RPCs vs. Mouse_RPCs, 13LGS_N. RPCs vs. Mouse_N.
940 RPCs, 13LGS_BC/Photoreceptor Precursors vs. Mouse_BC/Photoreceptor Precursors,
941 13LGS_Cone vs. Mouse_Cone, and 13LGS_Cone vs. Mouse_Rod. Motifs exhibiting an
942 adjusted p-value < 0.01 and an average difference (avg_diff) > 2, with their corresponding
943 genes showing consistent changes, were designated as differential motifs. Subsequently, all
944 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.

Key resources table:

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
rabbit anti-Arr3 polyclonal	Millipore	No. AB15282
rabbit anti-Gnat2 polyclonal	Invitrogen	PA5-119558
chicken anti-GFP polyclonal	Invitrogen	A10262
mouse anti-Nr2e3 monoclonal	R&D systems	PP-H7223-00
goat anti-Nrl polyclonal	R&D Systems	AF2945
goat anti-Otx2 polyclonal	R&D Systems	AF1979
rabbit anti-Sox9 polyclonal	Millipore	AB5535
mouse anti-Tfap2a monoclonal	DSHB	Cat# 3b5, RRID:AB_528084
mouse anti-Lhx1 monoclonal	DHSB	Cat#4F2
goat anti-Otx2 polyclonal	R&D Systems	BAF1979
rabbit anti-Mef2c polyclonal	Atlas Antibodies	HPA005533
donkey anti-mouse Alexa Fluor™ Plus 647 polyclonal	ThermoFisher Scientific	A-11055
donkey anti-rabbit Alexa Fluor™ 488 polyclonal	ThermoFisher Scientific	A-21206
donkey anti-goat IgG Alexa Fluor™ 555 polyclonal	ThermoFisher Scientific	A-21432
CF633 donkey anti-rabbit polyclonal	Sigma	SAB4600132-50UL
goat anti-chicken Alexa Fluor™ 488 polyclonal	Invitrogen	A-11039
donkey anti-mouse Alexa Fluor™ 568 polyclonal	Invitrogen	A10037
CUTANA™ IgG Negative Control Antibody for CUT&RUN and CUT&Tag	EpiCypher	13-0042
H3K4me1 Antibody, SNAP-Certified™ for CUT&RUN and CUT&Tag	EpiCypher	13-0057
H3K4me3 Antibody, SNAP-Certified™ for CUT&RUN	EpiCypher	13-0041
H3K27ac Antibody, SNAP-Certified™ for CUT&RUN and CUT&Tag	EpiCypher	13-0059

H3K27me3 Antibody, SNAP-Certified™ for CUT&RUN and CUT&Tag	EpiCypher	13-0055
NeuroD1 (E3E4F) Rabbit mAb	Cell Signaling	62953
rabbit Anti-OTX2 Antibody polyclonal	Atlas Antibodies	HPA000633
Anti-Histone H3 (tri methyl K9) antibody - ChIP Grade	Abcam	ab8898
Biological Samples		
13LGS retinal samples	This paper	N/A
Critical commercial assays		
Chromium Next GEM Single Cell 3' Kit v3.1	10X Genomics	PN-1000268
Dual Index Kit TT Set A	10X Genomics	PN-1000215
Chromium Next GEM Single Cell ATAC Reagent Kits v1.1	10X Genomics	PN-1000175
Single Index Kit N, Set A	10X Genomics	PN-1000212
3' CellPlex Kit Set A	10X Genomics	PN-1000261
Dual Index Kit NN Set A	10X Genomics	PN-1000243
CUTANA™ ChIC/CUT&RUN Kit	Epicypher	14-1048
CUTANA™ CUT&RUN Library Prep Kit	EpiCypher	14-1001
NEBNext® Ultra™ II DNA Library Prep Kit for Illumina®	NEB	E7645S
NEBNext® Multiplex Oligos for Illumina®	NEB	E7395S
Deposited data		
Mouse retinal development scRNA-seq data	Clark et al.[15]	GEO: GSE118614
Mouse retinal development scATAC-seq data	Lyu et al.[17]	GEO: GSE181251
13LGS retinal development scRNA-seq, scATAC-seq data 13LGS and Mouse Cut&Run data Mouse *Zic3*, *Pou2f1*, *Onecut1*, *Mef2c* overexpression and *Zic3* cKO scRNA-seq data	This paper	GEO: GSE295358
13LGS genome GCF_016881025.1_HiC_ltri_2	Fu et al.[76]	https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/016/881/025/GCF_016881025.1_HiC_ltri_2/
Experimental models: Organisms/strains		
Mus Musculus: *Zic3*$^{+/lox}$: B6;129-Zic3tm2.1Jwb/J	The Jackson Lab Jiang et al.[68]	RRID:IMSR_JAX:023162

Mus Musculus: *Mef2c*$^{+/lox}$: Mef2ctm1Jjs/J	The Jackson Lab Vong et al.[50]	RRID:IMSR_JAX:025556
Mus Musculus: *Chx10Cre-GFP*:	This lab Rowan et al.[47]	N/A
Oligonucleotides		
ZIC3 Gateway Ultimate™ ORF Clone	ThermoFisher Scientific	NM_003413 (coding sequence only)
POU2F1 Gateway Ultimate™ ORF Clone	ThermoFisher Scientific	XM_011509654 (coding sequence only)
Onecut1 Gateway Ultimate™ ORF Clone	ThermoFisher Scientific	NM_008262 (coding sequence only)
MEF2C Gateway Ultimate™ ORF Clone	ThermoFisher Scientific	NM_002397 (coding sequence only)
Recombinant DNA		
Plasmid: pCAGIG IRES-GFP	This lab Matsuda et al.[73]	N/A
Software and algorithms		
CellRanger Pipeline (version 3.1.0)	10X Genomics	https://www.10xgenomics.com/support/software
CellRanger Multi Pipeline (version 6.0)	10X Genomics	https://www.10xgenomics.com/support/software
CellRanger Aggr Pipeline (version 3.1.0)	10X Genomics	https://www.10xgenomics.com/support/software
CellRanger ATAC pipeline (version 1.2.0)	10X Genomics	https://www.10xgenomics.com/support/software
Seurat (version 4.0.4)	Hao, Hao et al.[65]	https://satijalab.org/seurat/articles/install_v5.html
Signac (version 1.4.0)	Stuart et al.[67]	https://satijalab.org/signac
cutadapt (version 3.5)	Martin[79]	https://cutadapt.readthedocs.io/en/stable/
Trim Galore (version 0.6.4_dev)	Krueger et al.[80]	https://github.com/FelixKrueger/TrimGalore
Bowtie2	Langmead et al.[81]	https://github.com/BenLangmead/bowtie2
SAMtools (version 1.12)	Danacek et al.[82]	https://www.htslib.org
UMI-tools	Smith et al.[83]	https://umi-tools.readthedocs.io/en/latest/index.html
deepTools (version 3.5.1)	Ramirez et al.[84]	https://deeptools.readthedocs.io/en/latest/
macs2	Zhang et al.[85]	https://pypi.org/project/MACS2/

ArchR	Granja et al.[87]	https://www.archrproject.com
ChromVAR	Schep et al.[88]	https://greenleaflab.github.io/chromVAR/
ImageJ	Schindelin et al.[77]	https://imagej.net/ij/

953

954 **Figure legends:**
955

956 **Figure 1 :** Overview of the study. (**A**) Schematic summary of the study. scRNA-Seq and
957 scATAC-Seq of whole 13LGS retinas were performed at 10 different time points. Major retinal
958 cell types are identified by clustering in both scRNA-Seq and scATAC-Seq datasets (**B**) UMAP
959 projections of all 13LGS retinal cells from scRNA-Seq and scATAC-Seq. Each point represents
960 a single cell, with cell type and time point indicated by shading. (**C**) Cell-specific marker gene
961 expression and accessibility are shown across 11 major retinal cell types. (**D**) The relative
962 abundances of rods, rod precursors, cones and cone precursors are different between
963 developing 13LGS and mouse retina.
964

965 **Figure 2.** Identification of 13LGS–specific regulatory mechanisms in cone development by
966 integrating late-stage retinal scRNA-Seq and scATAC-Seq data with mouse.
967 (**A**) UMAP projections of the integrated scRNA-Seq data from 13LGS and mouse. (**B**) UMAP
968 projections of the scATAC-Seq data, showing cell types re-annotated based on the integrated
969 scRNA-Seq analysis. (**C**) Heatmaps displaying eight clusters of genes that are differentially
970 expressed between 13LGS and mouse. Column labels correspond to the five major cell types
971 relevant to photoreceptor development: Late RPCs (late-stage primary retinal progenitor cells),
972 Late NG (late-stage neurogenic progenitor cells), BC/Pho pre(bipolar cell, cone, and rod
973 precursor cells), cone photoreceptors and rod photoreceptors. (**D**) Heatmaps showing the
974 chromatin accessibility of positively-correlated regulatory elements for each of the eight DEG
975 clusters. (**E**) Dot plots depicting the relative number of positively correlated regulatory elements
976 for each conserved gene pair between 13LGS and mouse. Each dot represents a conserved
977 gene pair, with the x-axis indicating the number of regulatory elements in mouse and the y-axis
978 showing the number in 13LGS. Pairwise t-tests were used to compare the number of elements
979 in each species for each gene pair, and the median values are indicated below each plot. (**F**)
980 Differential motif analysis in five cone development–associated cell types between 13LGS and
981 mouse. (**G**) *Zic3* expression motif accessibility is higher in 13LGS Photoreceptor Precursors and
982 mature Cones than in mouse. (**H**) Heatmap showing the expression levels of the top 30 Cone-
983 promoting genes identified by GRN analysis in 13LGS across retinal progenitor cells (RPCs),
984 neurogenic progenitor cells (N. RPCs), and photoreceptor precursors in both 13LGS and
985 mouse.
986

987 **Figure 3:** Zic3 is necessary for normal mouse cone development and sufficient to promote cone
988 gene expression. (**A**) Diagram of overexpression strategy used to test effects of *ZIC3*
989 overexpression. (**B**) Heatmap of log2-fold change in expression between cone-like precursors

990 and Rods separated by condition for select genes, scaled by gene. (**C**) Upset plot of significant
991 differentially expressed genes shared between conditions. (**D**) Diagram of retinal progenitor cell-
992 specific conditional loss of function analysis of *Zic3*. (**E**) Immunohistochemistry showing GFP
993 and Arr3 expression in P14 wild type (WT) and *Zic3*$^{lox/lox}$ (*Zic3* cKO) mouse retinas. Scale bars,
994 20 µm. (**F**) Box plot of the number of cones per micron along the retinal slice (n = 11 WT, 10
995 *Zic3* cKO). P-values calculated by Wilcoxon rank-sum test. P0, postnatal day 0; P5, postnatal
996 day 5; P14, postnatal day 14; FACS, fluorescence-activated cell sorting; scRNA-Seq, single-cell
997 RNA sequencing; ONL, outer nuclear layer; INL, inner nuclear layer; DAPI, 4′,6-diamidino-2-
998 phenylindole; GFP, green fluorescent protein; OE, overexpression; log2FC, log2 fold change;
999 DEG, differentially expressed gene; WT, wild type; cKO, conditional knockout.
1000
1001 **Figure 4**: Mef2c is sufficient to promote cone-specific gene expression and repress rod-specific
1002 gene expression in mouse. (**A**) Diagram of overexpression strategy used to test effects of
1003 *MEF2C* overexpression. (**B**) UMAP representation of P8 mouse retinal explants electroporated
1004 with a plasmid expressing GFP alone (Empty) or GFP in a bicistronic transcript with human
1005 MEF2C (*MEF2C*), (n = 7445 cell Empty, 11949 cells *MEF2C*). Each point represents a single
1006 cell and is colored by cell type as determined by clustering and marker gene expression. (**C**)
1007 Heatmap of expression for select genes for cones, cone-like photoreceptor precursors, rod
1008 photoreceptor precursors, and rods in cells overexpressing *MEF2C*, scaled by gene. (**D**)
1009 Immunohistochemistry showing GFP and GNAT2 expression in P8 mouse retinas from Empty
1010 and *MEF2C* conditions. Scale bars, 50 µm. (**E**) Box plot of the number of Gnat2+, GFP+ cells
1011 divided by the total number of GFP+ cells (n = 8 for both conditions). P-values calculated by
1012 Wilcoxon rank-sum test. P0, postnatal day 0; P8, postnatal day 8; FACS, fluorescence-activated
1013 cell sorting; scRNA-Seq, single-cell RNA sequencing; ONL, outer nuclear layer; INL, inner
1014 nuclear layer; DAPI, 4′,6-diamidino-2-phenylindole; GFP, green fluorescent protein; OE,
1015 overexpression.
1016
1017 **Figure 5:** Conserved TFs bind to species-specific enhancers to promote Cone specification in
1018 13LGS. (**A**) Schematic illustrating annotation of cis-regulatory elements in RPCs and
1019 Photoreceptor Precursors by integration of scATAC-Seq and CUT&RUN 13LGS and mouse
1020 datasets. (**B**) Heatmaps show annotated accessible regulatory elements in both 13LGS and
1021 mouse. Promoters, activated enhancers (AEs), and poised enhancers (PEs) which are
1022 associated with histone markers associated with genes in clusters C2 and C3, which are
1023 selectively active in 13LGS RPCs and/or photoreceptor precursors. Shading indicates
1024 CUT&TAG signal for the corresponding histone modification within 2kb of the scATAC-Seq peak
1025 center. Bar plots displaying the number of each category of regulatory element in each species
1026 that are conserved or species-specific. (**C**) Dot plots showing the enrichment of binding sites for
1027 Otx2 and Neurod1, TFs which are broadly expressed in both neurogenic RPC and
1028 photoreceptor precursors, which are enriched in both conserved cis-regulatory elements in both
1029 species. (**D**) Bar plots showing the number of conversed and species-specific enhancers per
1030 TSS in four cone-promoting genes between 13LGS and mouse. (**E**) The gene regulatory
1031 networks regulating *Thrb* expression in 13LGS and mouse late N. RPCs. (**F**) An example of a
1032 Thrb-related regulon and its corresponding scATACseq and CUT&RUN tracks. The arrow

1033 indicates the consistent regulatory relationships between GRNs prediction and experimental
1034 validations. (**G**) The epigenetic model of cone specification in 13LGS and mouse.
1035

1036 **Figure S1 (related to Figure 1) :** Quality control of scATAC-Seq and scRNA-Seq data. (**A**)
1037 Transcriptional start site enrichment profiles of scATAC-Seq datasets. Lines are colored by time
1038 points. (**B**) Fragment size distribution of scATAC-Seq datasets. Individual time points are
1039 indicated by colored lines. (**C**) Heatmap showing the Pearson correlation between gene
1040 expression and gene accessibility for each retina cell type. (**D**) Heatmap of cell type-specific
1041 genes. (**E**) Heatmap of cell type-specific peaks. (**F**) Heatmap of cell type-specific motifs. (**G**)
1042 Examples of TF footprint profiles for Pou4f2, Crx, Nfix and Onecut1 in indicated scATAC-Seq
1043 cell types. (**H**) Examples of chromVAR score are shown for Otx2, Pou2f2, Nfix and Neurod1
1044 using scATAC-Seq datasets.
1045 (**I**) The relative abundance of retinal cell types in scRNA-Seq and scATAC-Seq is different
1046 between developing 13LGS and mouse retina.
1047

1048 **Figure S2 (related to Figure 2):** Transcription factors that show heterochronic gene expression
1049 and co-expression with Otx2 show sustained expression in 13LGS retinas. Fluorescent *in situ*
1050 hybridization showing co-expression of (**A**) *Pou2f1* and *Otx2* or (**B**) *Zic3*, *Rxrg*, and *Otx2* in P1,
1051 P5, P10, and P24 retinas. Insets show higher power images of highlighted areas. (**C**) *Zic3,*
1052 *Rxrg,* and *Otx2* fluorescent *in situ* hybridization from P24 with matched (**C**') negative controls.
1053 (**D**) *Pou2f1* and *Otx2* fluorescent *in situ* hybridization from P24 with matched (**D**') negative
1054 controls. (**E**) Quantification of the fraction of *Otx2*-positive cells in the outer neuroblastic layer
1055 (P1, P5) and ONL (P10, P24) that also express *Zic3*. (**F**) Immunohistochemical analysis Mef2c
1056 and Otx2 expression in P1, P5, P10, and P24 retinas. (**G**) Mef2c and Otx2
1057 immunohistochemistry from P24 with matched (**G**') negative controls. Negative controls for
1058 fluorescent *in situ* hybridization omit the probe and for immunohistochemistry omit primary
1059 antibodies. Scale bars, 10 μm (S2A-F), 50 μm (S2G) and 5 μm (inset). Cell counts in E were
1060 analyzed using one-way ANOVA analysis with Sidak multiple comparisons test and 95%
1061 confidence interval. ** = $p < 0.01$, **** = $p < 0.0001$, and ns = non-significant. N=3 independent
1062 experiments. DAPI, 4′,6-diamidino-2-phenylindole; NbONL, neuroblastic outer nuclear layer;
1063 ONL, outer nuclear layer; INL, inner nuclear layer. Arrowheads indicate cells co-expressing
1064 indicated markers.
1065

1066 **Figure S3 (related to Figure 3)**: *ZIC3* overexpression promotes the formation of cone-like
1067 photoreceptor precursor cells. (**A**) UMAP representation of P5 mouse retinal explants from four
1068 conditions: electroporated with a plasmid expressing GFP alone (Empty), GFP in a bicistronic
1069 transcript with ZIC3 (*ZIC3*) or POU2F1 (*POU2F1*), or both the ZIC3 and POU2F1 plasmids
1070 (*ZIC3*+*POU2F1*) (n = 1901 cells Empty, 1578 cells *ZIC3*, 904 cells *POU2F1*, 2111 cells
1071 *ZIC3*+*POU2F1*), split by condition. Each point represents a single cell and is colored by cell type
1072 as determined by clustering and marker gene expression. (**B**) UMAP representation of P5
1073 mouse retinal explants from four conditions: electroporated with a plasmid expressing GFP
1074 alone (Empty), GFP in a bicistronic transcript with POU2F1 (*POU2F1*) or Onecut1 (*Onecut1*), or
1075 both the ZIC3 and POU2F1 plasmids (*ZIC3*+*POU2F1*) (n = 4313 cells Empty, 6287 cells
1076 *POU2F1*, 4455 cells *Onecut1,* 2476 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. (**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 (related to Figure 3): *Zic3* is required for normal patterns of 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. Orange arrows indicate horizontal cell nuclei and yellow arrows indicate mouse-on-mouse vascular staining. (**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 (related to Figure 4): *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 (related to Figure 4): *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 (related to Figure 5): 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.

1121
1122 **Figure S8 (related to Figure 5):** Zic3 is regulated by more extensive gene regulatory networks
1123 in 13LGS than mouse at all developmental stages. (**A**) Gene regulatory networks controlling
1124 *Zic3* expression in 13LGS late-stage neurogenic RPCs, photoreceptor precursors, and cones.
1125 (**B**) Example of Zic3-related regulons and their corresponding scATAC-Seq and CUT&RUN
1126 tracks in 13LGS. (**C**) Example of Zic3-related regulons and their corresponding scATAC-Seq
1127 and CUT&RUN tracks in mouse.
1128
1129 **Figure S9 (related to Figure 5):** Transcriptional regulatory networks controlling cone-specific
1130 gene expression in 13LGS increase in connectivity and complexity as differentiation proceeds.
1131 Cell type-specific 13LGS gene regulatory networks with cone-promoting TFs in (**A**) late-stage
1132 neurogenic retinal progenitor cells (late N. RPCs), (**B**) photoreceptor precursors, and (**C**)
1133 cones.
1134
1135 **Table ST1 (related to Figure 1)**: List of cell type-specific patterns of differential gene
1136 expression, chromatin accessibility, and motif enrichment in accessible chromatin for 13LGS.
1137 Counts for each cell type in each scRNA- and scATAC-Seq library are also shown.
1138
1139 **Table ST2 (related to Figure 1)**: Genes differentially expressed in S- and M-cones in 13LGS.
1140
1141 **Table ST3 (related to Figure 2)**: Gene ortholog pairs for 13LGS and mouse used in this
1142 analysis.
1143
1144 **Table ST4 (related to Figure 2)**: Genes and accessible transcription factor motifs showing
1145 differential activity between 13LGS and mouse in late-stage progenitors and differentiating
1146 photoreceptors.
1147
1148 **Table ST5 (related to Figure 2)**: Top cone-promoting transcription factors active in 13LGS.
1149
1150 **Table ST6 (related to Figure 2)**: List of genes differentially expressed following overexpression
1151 of *ZIC3*, *Onecut1*, *POU2F1*, and *ZIC3+POU2F1* as detected by scRNA-Seq and proportions of
1152 major cell types observed in each sample.
1153
1154 **Table ST7 (related to Figure 3)**: List of genes differentially expressed in *Chx10-Cre;Zic3*$^{lox/lox}$
1155 retina and proportions of major cell types observed in each sample.
1156
1157 **Table ST8 (related to Figure 4)**: List of genes differentially expressed following overexpression
1158 of *MEF2C* and proportions of major cell types observed in each sample.
1159
1160 **Table ST9 (related to Figure 5)**: Cis-regulatory elements inferred by scATAC-Seq and
1161 CUT&TAG for genes in Clusters 2 and 3 in 13LGS and mouse.
1162
1163 **Table ST10 (related to Figure 6)**: Active and poised enhancers associated with *Mef2c*, *Rxrg*,
1164 *Thrb* and *Zic3* in 13LGS and mouse.

Acknowledgements

We thank J. Nathans, A. Kolodkin, R. Johnston Jr, J. Tollkuhn, M.L. Wilson, 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 to S.B, F31EY031942 to K.W, K00EY036684 to J.T, R01EY012141 to S.H.D., grants from Research to Prevent Blindness to S.H.D. 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, National Eye Institute (U24EY029891) to D.K.M, and a Foundation Fighting Blindness to D.K.M.

1176 **References:**

1177 Ahnelt PK, Kolb H. 2000. The mammalian photoreceptor mosaic-adaptive design. *Prog Retin*
1178 *Eye Res* **19**:711–777.
1179 Akhmedov NB, Piriev NI, Chang B, Rapoport AL, Hawes NL, Nishina PM, Nusinowitz S,
1180 Heckenlively JR, Roderick TH, Kozak CA, Danciger M, Davisson MT, Farber DB. 2000. A
1181 deletion in a photoreceptor-specific nuclear receptor mRNA causes retinal degeneration in the
1182 rd7 mouse. *Proc Natl Acad Sci U S A* **97**:5551–5556.
1183 Alicia Schep [aut, cre], Stanford University [cph]. 2017. motifmatchr. Bioconductor.
1184 doi:10.18129/B9.BIOC.MOTIFMATCHR
1185 Amemiya HM, Kundaje A, Boyle AP. 2019. The ENCODE Blacklist: Identification of Problematic
1186 Regions of the Genome. *Sci Rep* **9**:9354.
1187 Andzelm MM, Cherry TJ, Harmin DA, Boeke AC, Lee C, Hemberg M, Pawlyk B, Malik AN,
1188 Flavell SW, Sandberg MA, Raviola E, Greenberg ME. 2015. MEF2D drives photoreceptor
1189 development through a genome-wide competition for tissue-specific enhancers. *Neuron*
1190 **86**:247–263.
1191 Arendt D. 2008. The evolution of cell types in animals: emerging principles from molecular
1192 studies. *Nat Rev Genet* **9**:868–882.
1193 Cepko C. 2014. Intrinsically different retinal progenitor cells produce specific types of progeny.
1194 *Nat Rev Neurosci* **15**:615–627.
1195 Chang B, Hawes NL, Hurd RE, Davisson MT, Nusinowitz S, Heckenlively JR. 2002. Retinal
1196 degeneration mutants in the mouse. *Vision Res* **42**:517–525.
1197 Cheng H, Khanna H, Oh ECT, Hicks D, Mitton KP, Swaroop A. 2004. Photoreceptor-specific
1198 nuclear receptor NR2E3 functions as a transcriptional activator in rod photoreceptors. *Hum Mol*
1199 *Genet* **13**:1563–1575.
1200 Chen J, Rattner A, Nathans J. 2005. The rod photoreceptor-specific nuclear receptor Nr2e3
1201 represses transcription of multiple cone-specific genes. *J Neurosci* **25**:118–129.
1202 Chen J, Richardson PR, Kirby C, Eddy SR, Hoekstra HE. 2025. Cellular evolution of the
1203 hypothalamic preoptic area of behaviorally divergent deer mice. *bioRxiv*.
1204 doi:10.1101/2024.08.22.608850
1205 Chromatin Potential Identified by Shared Single-Cell Profiling of RNA and Chromatin. 2020. .
1206 *Cell* **183**:1103–1116.e20.
1207 Clark BS, Stein-O'Brien GL, Shiau F, Cannon GH, Davis-Marcisak E, Sherman T, Santiago CP,
1208 Hoang TV, Rajaii F, James-Esposito RE, Gronostajski RM, Fertig EJ, Goff LA, Blackshaw S.
1209 2019. Single-Cell RNA-Seq Analysis of Retinal Development Identifies NFI Factors as
1210 Regulating Mitotic Exit and Late-Born Cell Specification. *Neuron* **102**:1111–1126.e5.
1211 Cretekos CJ, Wang Y, Green ED, Martin JF, Rasweiler JJ 4th, Behringer RR. 2008. Regulatory
1212 divergence modifies limb length between mammals. *Genes Dev* **22**:141–151.
1213 Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T,
1214 McCarthy SA, Davies RM, Li H. 2021. Twelve years of SAMtools and BCFtools. *Gigascience*
1215 **10**. doi:10.1093/gigascience/giab008
1216 Daniele LL, Lillo C, Lyubarsky AL, Nikonov SS, Philp N, Mears AJ, Swaroop A, Williams DS,
1217 Pugh EN Jr. 2005. Cone-like morphological, molecular, and electrophysiological features of the
1218 photoreceptors of the Nrl knockout mouse. *Invest Ophthalmol Vis Sci* **46**:2156–2167.
1219 de Melo J, Blackshaw S. 2018. In Vivo Electroporation of Developing Mouse Retina. *Methods*
1220 *Mol Biol* **1715**:101–111.
1221 de Melo J, Peng G-H, Chen S, Blackshaw S. 2011. The Spalt family transcription factor Sall3
1222 regulates the development of cone photoreceptors and retinal horizontal interneurons.
1223 *Development* **138**:2325–2336.
1224 Emerson MM, Surzenko N, Goetz JJ, Trimarchi J, Cepko CL. 2013. Otx2 and Onecut1 promote
1225 the fates of cone photoreceptors and horizontal cells and repress rod photoreceptors. *Dev Cell*

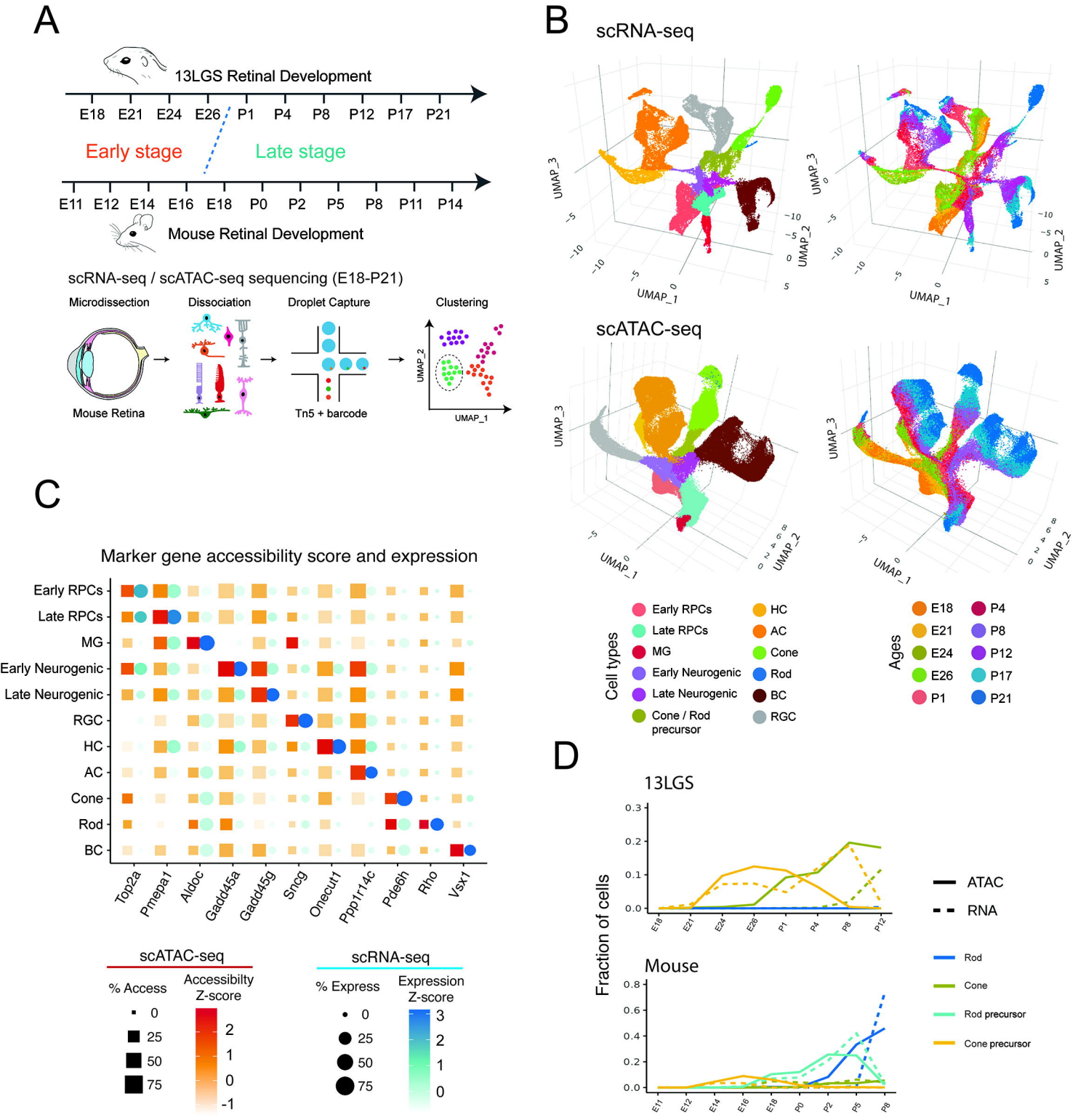
26:59–72.
Fernald RD. 2012. Social control of the brain. *Annu Rev Neurosci* **35**:133–151.
Fu R, Gillen AE, Grabek KR, Riemondy KA, Epperson LE, Bustamante CD, Hesselberth JR, Martin SL. 2020. Dynamic RNA Regulation in the Brain Underlies Physiological Plasticity in a Hibernating Mammal. *Front Physiol* **11**:624677.
Furukawa T, Morrow EM, Li T, Davis FC, Cepko CL. 1999. Retinopathy and attenuated circadian entrainment in Crx-deficient mice. *Nat Genet* **23**:466–470.
Gerkema MP, Davies WIL, Foster RG, Menaker M, Hut RA. 2013. The nocturnal bottleneck and the evolution of activity patterns in mammals. *Proc Biol Sci* **280**:20130508.
Granja JM, Corces MR, Pierce SE, Bagdatli ST, Choudhry H, Chang HY, Greenleaf WJ. 2021. ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat Genet* **53**:403–411.
Hafemeister C, Satija R. 2019. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* **20**:296.
Hahn J, Monavarfeshani A, Qiao M, Kao AH, Kölsch Y, Kumar A, Kunze VP, Rasys AM, Richardson R, Wekselblatt JB, Baier H, Lucas RJ, Li W, Meister M, Trachtenberg JT, Yan W, Peng Y-R, Sanes JR, Shekhar K. 2023. Evolution of neuronal cell classes and types in the vertebrate retina. *Nature* **624**:415–424.
Hao H, Kim DS, Klocke B, Johnson KR, Cui K, Gotoh N, Zang C, Gregorski J, Gieser L, Peng W, Fann Y, Seifert M, Zhao K, Swaroop A. 2012. Transcriptional regulation of rod photoreceptor homeostasis revealed by in vivo NRL targetome analysis. *PLoS Genet* **8**:e1002649.
Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, Hoffman P, Stoeckius M, Papalexi E, Mimitou EP, Jain J, Srivastava A, Stuart T, Fleming LM, Yeung B, Rogers AJ, McElrath JM, Blish CA, Gottardo R, Smibert P, Satija R. 2021. Integrated analysis of multimodal single-cell data. *Cell* **184**:3573–3587.e29.
Jacobs LF, Spencer WD. 1994. Natural space-use patterns and hippocampal size in kangaroo rats. *Brain Behav Evol* **44**:125–132.
Javed A, Mattar P, Lu S, Kruczek K, Kloc M, Gonzalez-Cordero A, Bremner R, Ali RR, Cayouette M. 2020. Pou2f1 and Pou2f2 cooperate to control the timing of cone photoreceptor production in the developing mouse retina. *Development* **147**. doi:10.1242/dev.188730
Jeon CJ, Strettoi E, Masland RH. 1998. The major cell populations of the mouse retina. *J Neurosci* **18**:8936–8946.
Jiang Z, Zhu L, Hu L, Slesnick TC, Pautler RG, Justice MJ, Belmont JW. 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.
Kallman A, Capowski EE, Wang J, Kaushik AM, Jansen AD, Edwards KL, Chen L, Berlinicke CA, Joseph Phillips M, Pierce EA, Qian J, Wang T-H, Gamm DM, Zack DJ. 2020. Investigating cone photoreceptor development using patient-derived NRL null retinal organoids. *Commun Biol* **3**:82.
Kandoi S, Martinez C, Merriman DK, Lamba DA. 2022. Characterization of Retinal Development in 13-Lined Ground Squirrels. *Transl Vis Sci Technol* **11**:17.
Kaskan PM, Franco ECS, Yamada ES, Silveira LC de L, Darlington RB, Finlay BL. 2005. Peripheral variability and central constancy in mammalian visual system evolution. *Proc Biol Sci* **272**:91–100.
Katoh K, Omori Y, Onishi A, Sato S, Kondo M, Furukawa T. 2010. Blimp1 suppresses Chx10 expression in differentiating retinal photoreceptor precursors to ensure proper photoreceptor development. *J Neurosci* **30**:6515–6526.
Katz PS. 2011. Neural mechanisms underlying the evolvability of behaviour. *Philos Trans R Soc Lond B Biol Sci* **366**:2086–2099.
Krueger F, James F, Ewels P, Afyounian E, Weinstein M, Schuster-Boeckler B, Hulselmans G, sclamons. 2023. FelixKrueger/TrimGalore: v0.6.10 - add default decompression path. Zenodo.

1277 doi:10.5281/ZENODO.7598955
1278 Kryger Z, Galli-Resta L, Jacobs GH, Reese BE. 1998. The topography of rod and cone
1279 photoreceptors in the retina of the ground squirrel. *Vis Neurosci* **15**:685–691.
1280 Lamb TD. 2013. Evolution of phototransduction, vertebrate photoreceptors and retina. *Prog*
1281 *Retin Eye Res* **36**:52–119.
1282 Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods*
1283 **9**:357–359.
1284 Leavey P, Jiang L, Pannullo N, Santiago C, Blackshaw S. 2025. Overexpression of Meis factors
1285 in late-stage retinal progenitors yields complex effects on temporal patterning and
1286 neurogenesis. *bioRxiv*. doi:10.1101/2025.05.01.651492
1287 Le Duc D, Schöneberg T. 2016. Adaptation to nocturnality - learning from avian genomes.
1288 *Bioessays* **38**:694–703.
1289 Lonfat N, Wang S, Lee C, Garcia M, Choi J, Park PJ, Cepko C. 2021. Cis-regulatory dissection
1290 of cone development reveals a broad role for Otx2 and Oc transcription factors. *Development*
1291 **148**. doi:10.1242/dev.198549
1292 Lu Y, Shiau F, Yi W, Lu S, Wu Q, Pearson JD, Kallman A, Zhong S, Hoang T, Zuo Z, Zhao F,
1293 Zhang M, Tsai N, Zhuo Y, He S, Zhang J, Stein-O'Brien GL, Sherman TD, Duan X, Fertig EJ,
1294 Goff LA, Zack DJ, Handa JT, Xue T, Bremner R, Blackshaw S, Wang X, Clark BS. 2020. Single-
1295 Cell Analysis of Human Retina Identifies Evolutionarily Conserved and Species-Specific
1296 Mechanisms Controlling Development. *Dev Cell* **53**:473–491.e9.
1297 Lyu P, Hoang T, Santiago CP, Thomas ED, Timms AE, Appel H, Gimmen M, Le N, Jiang L, Kim
1298 DW, Chen S, Espinoza DF, Telger AE, Weir K, Clark BS, Cherry TJ, Qian J, Blackshaw S.
1299 2021. Gene regulatory networks controlling temporal patterning, neurogenesis, and cell-fate
1300 specification in mammalian retina. *Cell Rep* **37**:109994.
1301 Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads.
1302 *EMBnet J* **17**:10.
1303 Matsuda T, Cepko CL. 2008. Analysis of gene function in the retina. *Methods Mol Biol* **423**:259–
1304 278.
1305 Matsuda T, Cepko CL. 2004. Electroporation and RNA interference in the rodent retina in vivo
1306 and in vitro. *Proc Natl Acad Sci U S A* **101**:16–22.
1307 Mattar P, Ericson J, Blackshaw S, Cayouette M. 2015. A conserved regulatory logic controls
1308 temporal identity in mouse neural progenitors. *Neuron* **85**:497–504.
1309 Mears AJ, Kondo M, Swain PK, Takada Y, Bush RA, Saunders TL, Sieving PA, Swaroop A.
1310 2001. Nrl is required for rod photoreceptor development. *Nat Genet* **29**:447–452.
1311 Milam AH, Rose L, Cideciyan AV, Barakat MR, Tang W-X, Gupta N, Aleman TS, Wright AF,
1312 Stone EM, Sheffield VC, Jacobson SG. 2002. The nuclear receptor NR2E3 plays a role in
1313 human retinal photoreceptor differentiation and degeneration. *Proc Natl Acad Sci U S A* **99**:473–
1314 478.
1315 Mossman HW, Weisfeldt LA. 1939. The fetal membranes of a primitive rodent, the thirteen-
1316 striped ground squirrel. *Am J Anat* **64**:59–109.
1317 Müller B, Peichl L. 1989. Topography of cones and rods in the tree shrew retina. *J Comp Neurol*
1318 **282**:581–594.
1319 Mustafi D, Engel AH, Palczewski K. 2009. Structure of cone photoreceptors. *Prog Retin Eye*
1320 *Res* **28**:289–302.
1321 Nagar S, Trudler D, McKercher SR, Piña-Crespo J, Nakanishi N, Okamoto S-I, Lipton SA. 2017.
1322 Molecular Pathway to Protection From Age-Dependent Photoreceptor Degeneration in Mef2
1323 Deficiency. *Invest Ophthalmol Vis Sci* **58**:3741–3749.
1324 Ng L, Hurley JB, Dierks B, Srinivas M, Saltó C, Vennström B, Reh TA, Forrest D. 2001. A
1325 thyroid hormone receptor that is required for the development of green cone photoreceptors.
1326 *Nat Genet* **27**:94–98.
1327 Nishida A, Furukawa A, Koike C, Tano Y, Aizawa S, Matsuo I, Furukawa T. 2003. Otx2

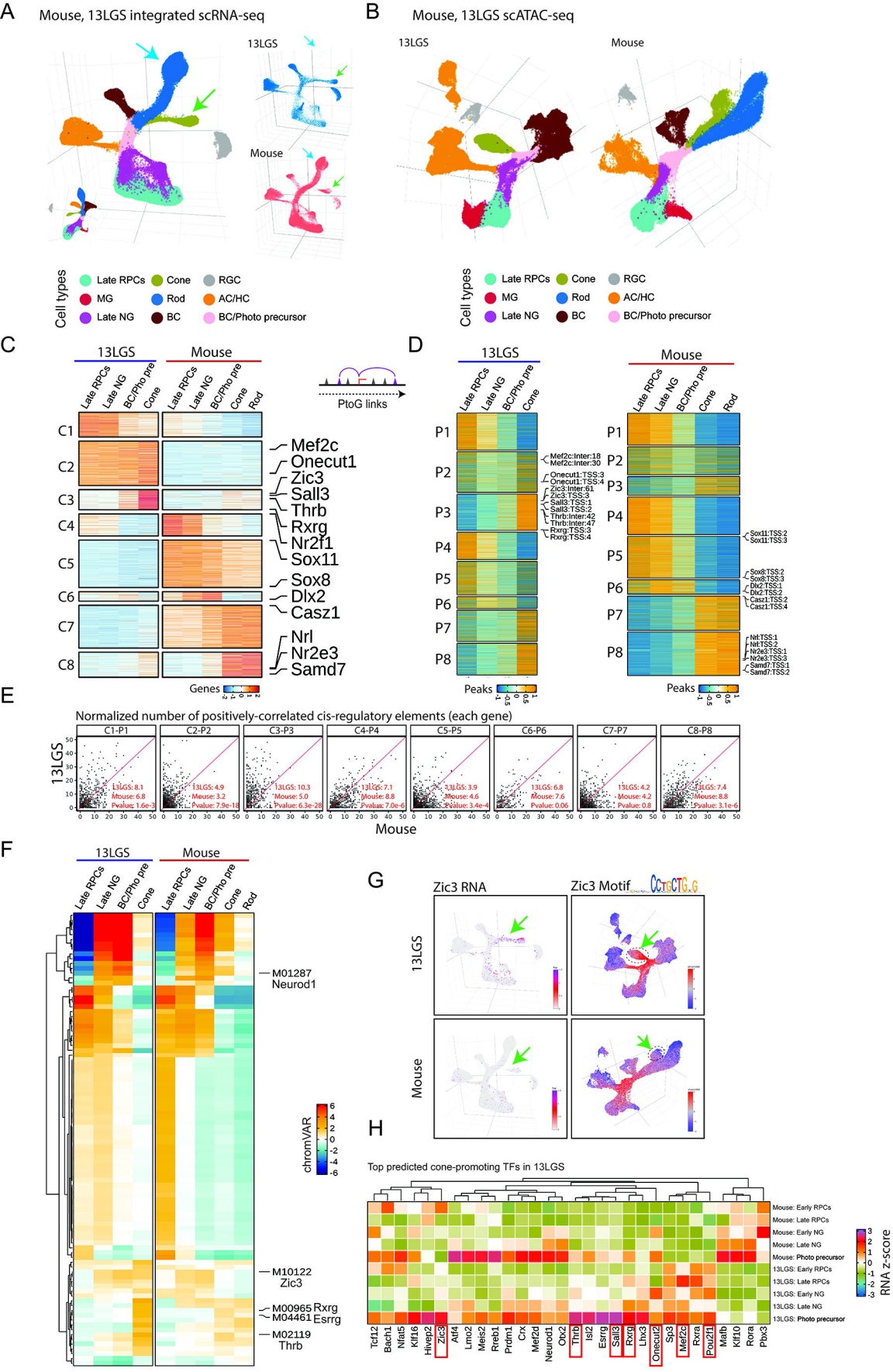
homeobox gene controls retinal photoreceptor cell fate and pineal gland development. *Nat Neurosci* **6**:1255–1263.
Omori Y, Kubo S, Kon T, Furuhashi M, Narita H, Kominami T, Ueno A, Tsutsumi R, Chaya T, Yamamoto H, Suetake I, Ueno S, Koseki H, Nakagawa A, Furukawa T. 2017. Samd7 is a cell type-specific PRC1 component essential for establishing retinal rod photoreceptor identity. *Proc Natl Acad Sci U S A* **114**:E8264–E8273.
Onishi A, Peng G-H, Chen S, Blackshaw S. 2010. Pias3-dependent SUMOylation controls mammalian cone photoreceptor differentiation. *Nat Neurosci* **13**:1059–1065.
Onishi A, Peng G-H, Hsu C, Alexis U, Chen S, Blackshaw S. 2009. Pias3-dependent SUMOylation directs rod photoreceptor development. *Neuron* **61**:234–246.
Peichl L. 2005. Diversity of mammalian photoreceptor properties: adaptations to habitat and lifestyle? *Anat Rec A Discov Mol Cell Evol Biol* **287**:1001–1012.
Pennesi ME, Cho J-H, Yang Z, Wu SH, Zhang J, Wu SM, Tsai M-J. 2003. BETA2/NeuroD1 null mice: a new model for transcription factor-dependent photoreceptor degeneration. *J Neurosci* **23**:453–461.
Purandare SM, Ware SM, Kwan KM, Gebbia M, Bassi MT, Deng JM, Vogel H, Behringer RR, Belmont JW, 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.
Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**:W160–5.
Roberts MR, Hendrickson A, McGuire CR, Reh TA. 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.
Rowan S, Cepko CL. 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.
Sapkota D, Chintala H, Wu F, Fliesler SJ, Hu Z, Mu X. 2014. Onecut1 and Onecut2 redundantly regulate early retinal cell fates during development. *Proc Natl Acad Sci U S A* **111**:E4086–95.
Satoh S, Tang K, Iida A, Inoue M, Kodama T, Tsai SY, Tsai M-J, Furuta Y, 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.
Saxena A, Sharma V, Muthuirulan P, Neufeld SJ, Tran MP, Gutierrez HL, Chen KD, Erberich JM, Birmingham A, Capellini TD, Cobb J, Hiller M, Cooper KL. 2022. Interspecies transcriptomics identify genes that underlie disproportionate foot growth in jerboas. *Curr Biol* **32**:289–303.e6.
Schep AN, Wu B, Buenrostro JD, Greenleaf WJ. 2017. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods* **14**:975–978.
Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**:676–682.
Slavenko A, Dror L, Camaiti M, Farquhar JE, Shea GM, Chapple DG, 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.
Smith T, Heger A, Sudbery I. 2017. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res* **27**:491–499.
Stuart T, Srivastava A, Madad S, Lareau CA, Satija R. 2021. Single-cell chromatin state analysis with Signac. *Nat Methods* **18**:1333–1341.
Sutherland MJ, Wang S, Quinn ME, Haaning A, Ware SM. 2013. Zic3 is required in the migrating primitive streak for node morphogenesis and left-right patterning. *Hum Mol Genet*

22:1913–1923.
Swaroop A, Kim D, Forrest D. 2010. Transcriptional regulation of photoreceptor development and homeostasis in the mammalian retina. *Nat Rev Neurosci* **11**:563–576.
Tabata Y, Ouchi Y, Kamiya H, Manabe T, Arai K-I, 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.
Tangeman JA, Charris Dominguez CM, Bendezu-Sayas S, Del Rio-Tsonis K. 2025. Nuclei Isolation from Ocular Tissues of the Embryonic Chicken for Single-Nucleus Profiling. *Methods Mol Biol* **2848**:105–116.
Thomas ED, Timms AE, Giles S, Harkins-Perry S, Lyu P, Hoang T, Qian J, Jackson VE, Bahlo M, Blackshaw S, Friedlander M, Eade K, Cherry TJ. 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.
Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CA-K, Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist P-H, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, von Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, von Heijne G, Nielsen J, Pontén F. 2015. Proteomics. Tissue-based map of the human proteome. *Science* **347**:1260419.
Verra DM, Sajdak BS, Merriman DK, Hicks D. 2020. Diurnal rodents as pertinent animal models of human retinal physiology and pathology. *Prog Retin Eye Res* **74**:100776.
Vong LH, Ragusa MJ, Schwarz JJ. 2005. Generation of conditional Mef2cloxP/loxP mice for temporal- and tissue-specific analyses. *Genesis* **43**:43–48.
Wada K, Howard JT, McConnell P, Whitney O, Lints T, Rivas MV, Horita H, Patterson MA, White SA, Scharff C, Haesler S, Zhao S, Sakaguchi H, Hagiwara M, Shiraki T, Hirozane-Kishikawa T, Skene P, Hayashizaki Y, Carninci P, Jarvis ED. 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.
Walls GL. 1942. The Vertebrate Eye and Its Adaptive Radiation.
Website. n.d. DOI: https://doi.org/10.1038/s41467-020-18035-1)
Weir K, Leavey P, Santiago C, Blackshaw S. 2021. Multiplexed Analysis of Retinal Gene Expression and Chromatin Accessibility using scRNA-Seq and scATAC-Seq. *J Vis Exp*. doi:10.3791/62239
Wolf A, Aslanidis A, Langmann T. 2017. Retinal expression and localization of Mef2c support its important role in photoreceptor gene expression. *Biochem Biophys Res Commun* **483**:346–351.
Yang L, Liu Q, Pollak GD. 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.
Young LJ, Wang Z. 2004. The neurobiology of pair bonding. *Nat Neurosci* **7**:1048–1054.
Yu CT, Kandoi S, Periasamy R, Reddy LVK, Follett HM, Summerfelt P, Martinez C, Guillaume C, Bowie O, Connor TB, Lipinski DM, Allen KP, Merriman DK, Carroll J, Lamba DA. 2024. Human iPSC-derived photoreceptor transplantation in the cone dominant 13-lined ground squirrel. *Stem Cell Reports* **19**:331–342.
Yuodelis C, Hendrickson A. 1986. A qualitative and quantitative analysis of the human fovea during development. *Vision Res* **26**:847–855.
Zhang X, Leavey P, Appel H, Makrides N, Blackshaw S. 2023. Molecular mechanisms controlling vertebrate retinal patterning, neurogenesis, and cell fate specification. *Trends Genet* **39**:736–757.
Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**:R137.

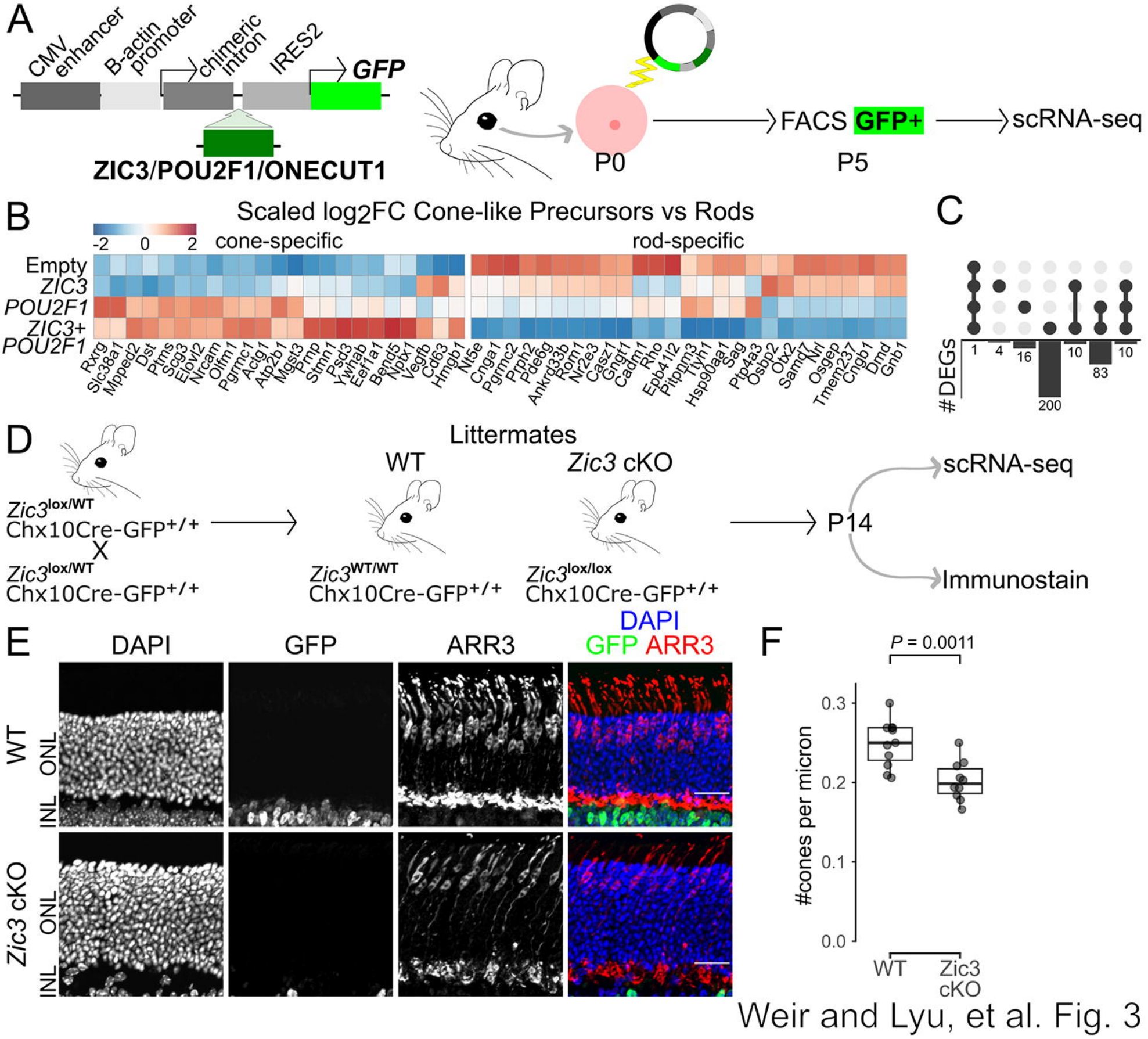
1430



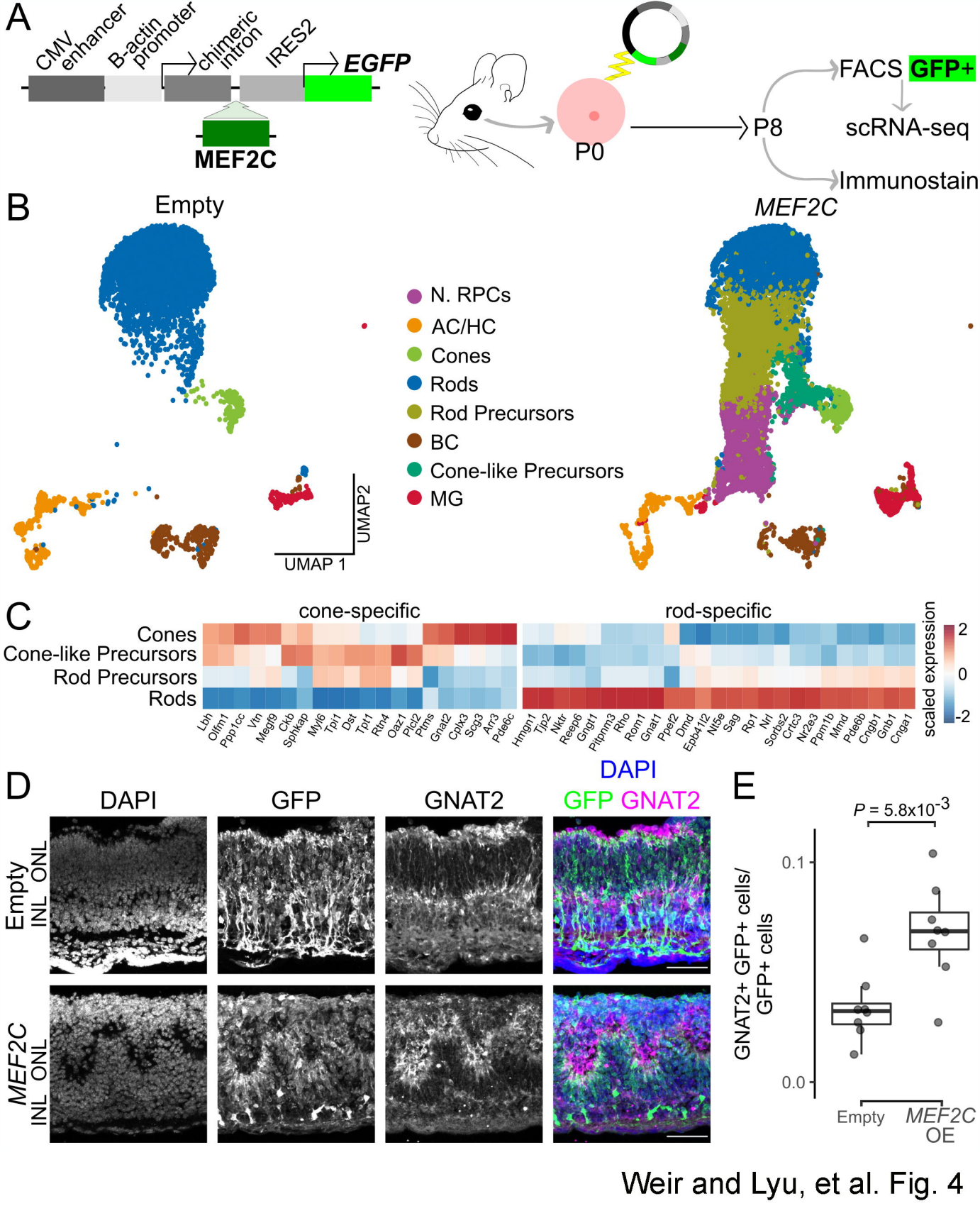
Weir and Lyu, et al. Fig. 1



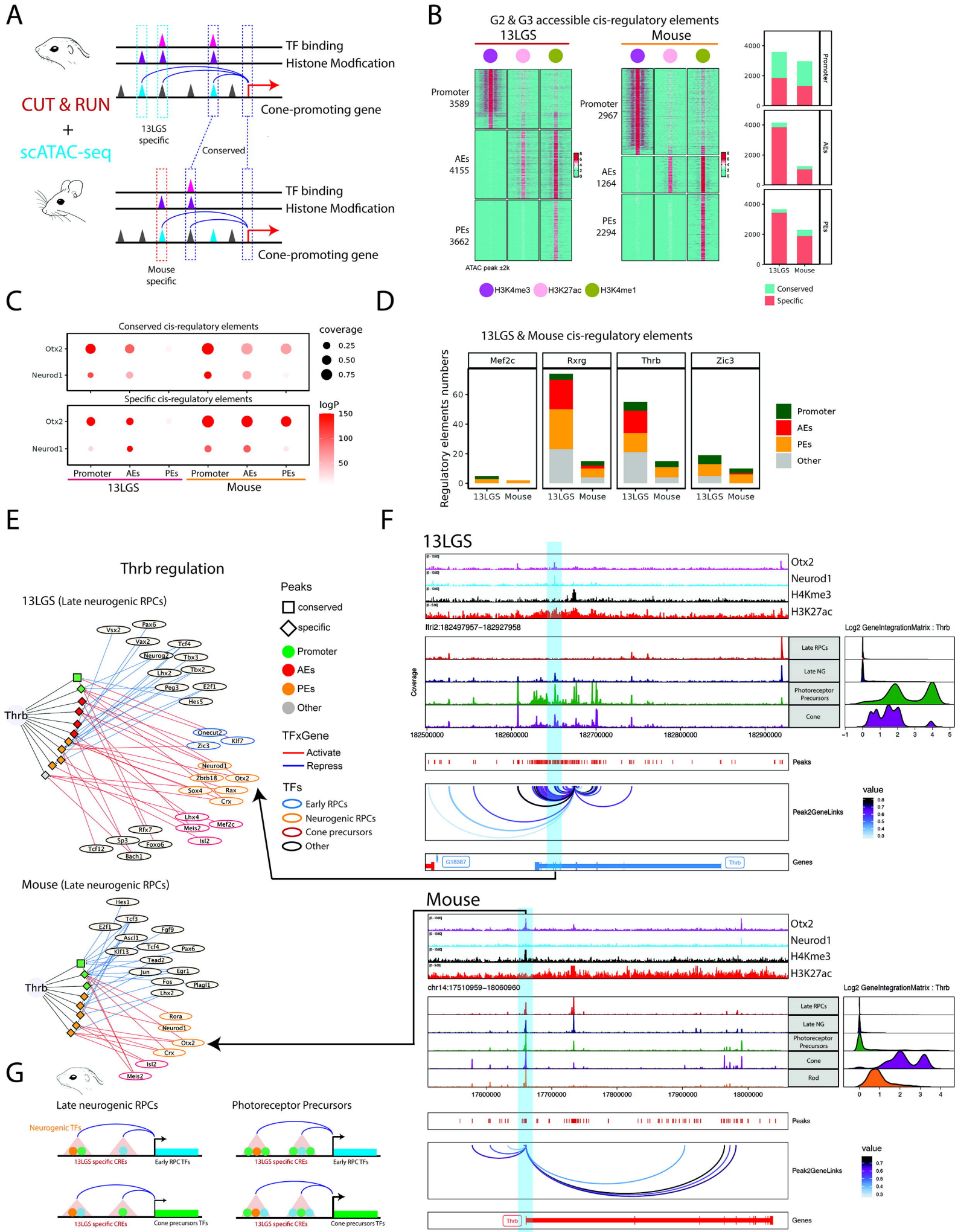
Weir and Lyu, et al. Fig. 2



Weir and Lyu, et al. Fig. 3



Weir and Lyu, et al. Fig. 4



Weir and Lyu, et al. Fig. 5