Cis-regulatory analysis of Onecut1 expression in fate-restricted retinal progenitor cells

Sruti Patoori[1,2], Nathalie Jean-Charles[2], Ariana Gopal[2], Sacha Sulaiman[2], Sneha Gopal[2,3], Brian Wang[2], Benjamin Souferi[2,4], Mark M. Emerson[1,2,5]

[1] Biology PhD Program, The Graduate Center, The City University of New York, New York, NY, 10016
[2] Department of Biology, The City College of New York, The City University of New York, New York, NY, 10031
[3] Present Address: Doctoral program in Department of Chemical and Biological Engineering, Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy , NY 12180
[4] Present Address: Touro College of Osteopathic Medicine, New York, NY 10027
[5] Biochemistry PhD Program, Graduate Center, City University of New York, NY 10016

*Corresponding author: memerson@ccny.cuny.edu

Email addresses:

Sruti Patoori: spatoori@gradcenter.cuny.edu
Ariana Gopal: ariana.gopal@macaulay.cuny.edu
Sacha Sulaiman: ssulaim000@citymail.cuny.edu
Nathalie Jean-Charles: nathaliejcharles@gmail.com
Sneha Gopal: sneha.gopal94@gmail.com
Brian Wang: brianwangmhc@gmail.com
Benjamin Souferi: benjaminsouferi@gmail.com
Mark Emerson: memerson@ccny.cuny.edu

Running title: Retinal fate-restricted cis-regulatory elements

Keywords: retinal progenitor cells, multipotent, fate-restricted, cone photoreceptors, basic helix-loop-helix, chicken, electroporation

Background: The vertebrate retina consists of six major classes of neuronal cells. During development, these cells are generated from a pool of multipotent retinal progenitor cells (RPCs) that express the gene Vsx2. Fate-restricted RPCs have recently been identified, with limited mitotic potential and cell fate possibilities compared to multipotent RPCs. One population of fate-restricted RPCs, marked by activity of the regulatory element ThrbCRM1, gives rise to both cone photoreceptors and horizontal cells. These cells do not express Vsx2, but co-express the transcription factors (TFs) Onecut1 and Otx2, which bind to ThrbCRM1. The components of the gene regulatory networks that control the transition from multipotent to fate-restricted gene expression are not known. This work aims to identify and evaluate cis-regulatory elements proximal to Onecut1 to identify the gene regulatory networks involved in RPC fate-restriction.

Method: We identified regulatory elements through ATAC-seq and conservation, followed by reporter assays to screen for activity based on temporal and spatial criteria. The regulatory elements of interest were subject to deletion and mutation analysis to identify functional sequences and evaluated by quantitative flow cytometry assays. Finally, we combined the enhancer::reporter assays with candidate TF overexpression to evaluate the relationship between the TFs, the enhancers, and early vertebrate retinal development. Statistical tests included ANOVA, Kruskal-Wallis, or unpaired t-tests.

Results: Two regulatory elements, ECR9 and ECR65, were identified to be active in ThrbCRM1(+) restricted RPCs. Candidate bHLH binding sites were identified as critical sequences in both elements. Overexpression of candidate bHLH TFs revealed specific enhancer-bHLH interactions. Nhlh1 overexpression expanded ECR65 activity into the Vsx2(+) RPC population, and overexpression of NeuroD1/NeuroG2/NeuroD4 had a similar effect on ECR9. Furthermore, bHLHs that were able to activate ectopic ECR9 reporter were able to induce endogenous Otx2 expression.

Conclusions: This work reports a large-scale screen to identify spatiotemporally specific regulatory elements near the Onecut1 locus. These elements were used to identify distinct populations in the developing retina. In addition, fate-restricted regulatory elements responded differentially to bHLH factors, and suggest a role for retinal bHLHs upstream of the Otx2 and Onecut1 genes during the formation of restricted RPCs from multipotent RPCs.

85 **Introduction**:

86 The vertebrate retina is comprised of six main classes of neuronal cells and one class of glial
87 cells, organized into three discrete nuclear layers and two plexiform layers. These
88 morphologically and functionally diverse cells have been characterized in multiple vertebrate
89 species to originate from multipotent retinal progenitor cells (RPCs) (Turner and Cepko, 1987;
90 Fekete et al., 1994). The vertebrate retina is therefore a valuable model to study neuronal cell
91 fate choice. The process of retinal development is highly conserved throughout the vertebrate
92 subphylum, with regards to the birth order of the various cell types (Young, 1985; Wong and
93 Rapaport, 2009) and the developmental regulatory networks involved. However, the gene
94 regulatory networks (GRNs) that mediate the generation of specific restricted RPCs from
95 multipotent RPCs are largely unknown, as are the networks that function in restricted RPCs to
96 define their fate potential.

97

98 One restricted RPC type that has been identified across zebrafish, chick, and mouse models
99 preferentially generates cones and horizontal cells (HCs) and has been identified in zebrafish
100 and chick through regulatory elements associated with the Thrb and Olig2 genes (Emerson et
101 al., 2013; Hafler et al., 2012; Suzuki et al., 2013). For example, analysis in zebrafish and mouse
102 RPCs showed that the same RPC can give rise to both cone and horizontal cell precursor cells
103 (Hafler et al., 2012; Suzuki et al., 2013). Endogenous Thrb expression has been observed in
104 Otx2-expressing early RPCs in the chick, suggesting that reporters driven by regulatory
105 elements correspond to *in vivo* regulatory events in the retina (Trimarchi et al., 2008). While
106 Otx2 expression is involved in multiple cell fates during retinal development, it has been shown

107 that the combination of Otx2 and Onecut1 activates ThrbCRM1, which is a specific Thrb cis-
108 regulatory element (CRE) active in cone/HC restricted RPCs (RPC[CH]) (Emerson et al., 2013).
109 Loss-of-function mutations in Otx2 and Onecut1 affect early cone gene expression, cone
110 number, cone type, and horizontal cell genesis (Nishida et al., 2003; Sapkota et al., 2014; Wu et
111 al., 2013), suggesting that these transcription factors (TFs) are critical in the gene regulatory
112 networks of ThrbCRM1 restricted RPCs.
113
114 The population of restricted RPCs marked by ThrbCRM1 have been shown to be molecularly
115 distinct from multipotent RPCs. ThrbCRM1(+) RPCs downregulate multipotent RPC genes such
116 as Vsx2 while Onecut1 is upregulated (Buenaventura et al., 2018). Onecut1 expression is further
117 upregulated in the HC progeny of these cells but is downregulated in the cone photoreceptor
118 progeny. However, it is not known how Onecut1 expression is activated in the ThrbCRM1 RPC
119 population or what distinguishes it from other Onecut1(+) cell populations.
120
121 The regulatory module that connects Otx2 and Onecut1 to Thrb expression demonstrates the
122 importance of both cis- and trans- regulatory elements in directing retinal cell fate. Cell fate
123 specification and fate restriction require the combinatorial expression of multiple
124 developmental transcription factors. As such, cell-type specific cis-regulatory elements can
125 define the intermediate/restricted RPCs that may be difficult to identify through only the
126 transient expression of developmental transcription factors that are often involved in the
127 specification of multiple retinal cell types. These regulatory elements can be used to facilitate

128 imaging, lineage tracing, and molecular analysis, while also providing insights into the
129 relationships between RPC populations.
130
131 We sought to identify the cis-regulatory elements upstream of Onecut1 expression in cone/HC
132 restricted RPCs and to determine the transcription factors that occupy these elements. To this
133 end, we conducted a multi-step screen to identify Onecut1-associated non-coding DNA
134 elements capable of driving reporter transcription in early retinal RPCs that give rise to cones
135 and horizontal cells in the early embryonic chick retina. The candidate regulatory elements that
136 emerged from the screen were then bioinformatically analyzed for transcription factor binding
137 sites. Mutational analyses facilitated the functional evaluation of these predicted TF binding
138 sites and overexpression experiments were used to determine the relationship between the
139 predicted transcription factors and Onecut1 expression. We identified two regulatory elements,
140 ECR9 and ECR65, upstream of the Onecut1 coding region. These both contain predicted binding
141 sites for bHLH transcription factors, which are known to be functionally important for retinal
142 development. We show that both of these elements require the predicted bHLH binding sites
143 for their activity and that each element responds to distinct bHLH factors, the transcripts of
144 which are enriched or present in the ThrbCRM1 population. Finally, we show that NeuroD1,
145 NeuroG2, and NeuroD4 are sufficient to induce expression of Otx2 and that all four TFs
146 including Nhlh1 are able to induce the activity of their corresponding regulatory elements in
147 Vsx2(+) multipotent RPCs. Ultimately, this work further clarifies components of the gene
148 regulatory network leading to the early retinal cell fates of cone photoreceptors, horizontal
149 cells, and retinal ganglion cells.

150

151 **Results**:

152

153 ***Regulatory Element Identification:***

154

155 Two methods were employed to identify candidate cis-regulatory elements for Onecut1 (Figure
156 1A,B). We first examined the intergenic region 5' of the chicken Onecut1 coding region and 3' of
157 WDR72, as well as a short stretch of the intergenic region 3' of the Onecut1 coding region for
158 evolutionarily conserved regions (ECRs). As not all cis-regulatory sequences are strongly
159 conserved and as ECRBrowser (Ovcharenko et al., 2004) utilizes an older chick genome
160 assembly, we also used chromatin accessibility as a means of candidate enhancer identification.
161 Chick E5 retinae were electroporated with ThrbCRM1::GFP and UbiqC::TdT, cultured for 18-22
162 hours *ex vivo* and sorted into two populations: ThrbCRM1(+) cells, which were marked by both
163 reporters and ThrbCRM1(-), which were marked by only TdT. These cells were then processed
164 for ATAC-seq (Buenrostro et al., 2013), aligned against the galGal5 assembly and the data was
165 visualized in the UCSC Genome Browser (Kent et al., 2002) to identify accessible chromatin
166 regions (ACRs) as potential cis-regulatory elements. In total, we screened 98 ECRs and ACRs
167 near Onecut1 (Additional File 1, Additional File 2) for their ability to drive reporter activity in
168 the developing retina.

169

170 The first criterion of our screen was that the non-coding elements should be capable of driving
171 reporter expression at E5 in the chick retina. To this end, we used a sensitive alkaline

172 phosphatase (AP) reporter assay. Each potential cis-regulatory region was amplified from chick
173 genomic DNA and cloned into Stagia3 (Billings et al., 2010), a dual GFP and AP reporter vector.
174 Retinae at E5 (HH26) were electroporated with the reporter construct along with CAG::mCherry
175 as the co-electroporation control. These retinae were cultured for approximately 18-22h and
176 then fixed before the AP stain was developed.
177
178 Previously identified Thrb reporters served as positive controls as they are known to be active
179 in the E5 chick retina (Emerson et. al., 2013). The empty Stagia3 vector served as the negative
180 control to demonstrate the baseline levels of transcription when no cis-regulatory element is
181 present. At this time point, the majority of the candidate sequences tested did not drive
182 reporter expression above the baseline level defined by the negative control. The active
183 elements that were initially chosen based on evolutionary conservation were largely found to
184 have open chromatin states in the within the ATAC-Seq datasets. All active cis-regulatory
185 elements (Figure 1C, Additional File 3) were categorized as weak, moderate, or strong (Figure
186 1B) and further examined for activity in cone/HC RPCs.
187
188 ***Regulatory Activity within the Cone/HC Restricted RPC population:***
189 The second criterion of the screen was specificity to the population of fate-restricted early
190 retinal RPCs that express Onecut1. The ThrbCRM1 element is active in the cone/HC restricted
191 RPC population that expresses Onecut1 and Otx2 (Emerson et al., 2013). It has been previously
192 reported that at 20 hours post-electroporation, 30% of ThrbCRM1(+) cells are in S-phase or
193 G2/M (Buenaventura et al., 2018). Therefore, to determine which active elements drove

194 transcription in the same restricted RPC population, we co-electroporated the CRE::GFP
195 reporter constructs with a ThrbCRM1::AU1 reporter into the chick retina at E5 and cultured
196 overnight for approximately 20 hours. Retinal sections were stained for AU1 and GFP and
197 qualitatively evaluated with regards to the specificity of each active enhancer to the Onecut1(+)
198 restricted RPC population marked by the AU1 reporter (Figure 2, Additional File 4).
199
200 Many of the active CREs drove reporter expression in both the ThrbCRM1(+) and ThrbCRM1(-)
201 populations, such as ECR42, ECR46 and ACR10-C (Additional File 4). ACR10-B (Additional File 4)
202 activity appears to be distributed throughout the retina with no specific preference for
203 ThrbCRM1(+) cells. Though the populations marked by these CREs overlap with our population
204 of interest, they do not meet the criterion for specificity. Despite driving robust reporter
205 expression in E5 chick retinae, ACR8 and ACR10-A (Figure 2, Additional File 4) appear biased
206 towards ThrbCRM1(-) cells. It is difficult to assess the specificity ECR26, ECR29, and ECR35 to
207 the ThrbCRM1 RPC population due to the weak activity of these enhancers (Additional Files 3
208 and 4).
209
210 ACR2 does not exhibit activity in many cells, but most observed ACR2-positive cells are
211 ThrbCRM1-positive (Figure 2). ECR65 appears to be highly active in ThrbCRM1-positive cells and
212 was qualitatively the most specific enhancer to the ThrbCRM1 population. Likewise, ECR9
213 activity is highly biased to the ThrbCRM1(+) RPC population but drives reporter activity in some
214 ThrbCRM1(-) cells.
215

216 This assay indicates that the active enhancers display a wide range in specificity to the
217 ThrbCRM1(+) cell population (Figure 2A, Additional File 4). ECR 65, ECR 9 and ACR 2 are the
218 most promising candidates for a regulatory element that contains the regulatory information
219 sufficient to promote Onecut1 expression in RPC[CH]s. ECR9 and ECR65 were further assessed
220 for specificity to this restricted RPC cell population as they demonstrated more robust reporter
221 expression than ACR2.
222
223 To quantify the specificity of each enhancer to the Onecut1(+) restricted RPCs, E5 chick retinae
224 were electroporated with ECR9:: or ECR65::GFP reporters, ThrbCRM1::AU1, and a co-
225 electroporation control. It was observed that 86.80% of ECR65(+) cells and 70.73% of ECR9(+)
226 cells were also marked by ThrbCRM1 (Figure 2C). However, the ThrbCRM1(+) cell population
227 does not consist of only RPCs. To determine the extent to which each enhancer was active in
228 RPCs, E5 chick retinae electroporated with CRE::GFP constructs were pulsed with EdU for one
229 hour prior to harvest. Approximately 30% of ECR9(+) and ECR65(+) cells are marked by EdU
230 (Figure 2B, 2C), which is comparable to the proportion of EdU(+) cells within the ThrbCRM1 cell
231 population (Buenaventura et al., 2018). Overall, these data provide evidence that the two
232 regulatory elements ECR65 and ECR9 are active in ThrbCRM1 RPCs.
233
234 ***History of ECR 9 and ECR 65 Activity in Early-born Retinal Cell Types***
235 To further test if these regulatory elements label RPCs which produce cells with the same fates
236 that develop from the ThrbCRM1(+) RPC population, we used a PhiC31 lineage tracing system.
237 In this system, a cis-regulatory element is used to drive the expression of PhiC31, which can

238 activate a GFP responder vector through site-specific recombination to label cells with a history
239 of cis-regulatory activity (Schick et al., 2019). We combined this lineage trace system with
240 immunohistochemistry and cell-specific markers to determine which cell types develop from
241 RPCs marked by ECR65 and ECR9. For comparison, and to demonstrate that not every active
242 regulatory element lineage traces to the same populations at this time point, we also lineage
243 traced ACR2 and ECR42.
244
245 It was observed that approximately 40% and 5.5% of ECR9-lineage traced cells were cone
246 photoreceptor (Visinin) and horizontal cell (Lim1) fates, respectively. However, not all ECR9-
247 lineage traced cells correspond to one of these two cells types. Some cells in the ECR9 lineage
248 exhibit long axonal projections which appear to originate from GFP-positive cells in the
249 innermost retina, suggesting that ECR9 is active at some point in the formation of retinal
250 ganglion cells (RGCs). 4.37% of all GFP(+) cells were positive for pan-Brn3, suggesting the
251 presence of RGCs arising from ECR9(+) cells. This may indicate that ECR 9 participates in the
252 gene regulatory network responsible for generating cones, HCs, and RGCs. It is worth nothing
253 that the ThrbCRM1 lineage-trace also includes a similar percentage of pan-Brn3(+) cells as
254 ECR9, despite the lack of inner retinal projections seen with ThrbCRM1 and previous *in ovo*
255 lineage tracing of ThrbCRM1 that detected only a small number of RGCs (Schick et al., 2019) .
256 This is potentially due to the antibody's specificity – it may not mark all Brn3(+) RGC
257 populations or it may be marking cells outside of the target population, including other inner
258 retinal cells such as horizontal or amacrine cells.
259

260 Under the same experimental conditions, ECR65::PhiC31 marked overall fewer cells. However,
261 the cell types with a history of ECR65 appeared biased towards the outer retina (Figure 3A).
262 63.3% of GFP-positive cells were also marked by Visinin (Figure 3B) , confirming that ECR65-
263 positive RPCs are capable of giving rise to cone photoreceptors. Surprisingly, very few ECR65-
264 positive cells are marked by Lim1 despite the clear presence of GFP in inner retinal cells (Figure
265 3A) and reported data that the ThrbCRM1 population gives rise preferentially to Lim1-positive
266 HCs (Schick et al., 2019). To determine whether these cells may be RGCs or Isl1-positive HC, we
267 stained with pan-Brn3 and Isl1. There were few pan-Brn3(+) GFP(+) cells. However, 5.9% of all
268 GFP-positive cells were marked by Isl1. Lineage tracing of ThrbCRM1 yields similar results, with
269 6.16% of GFP(+) cells also positive for Isl1 (Figure 3B).
270
271 ACR 2 does not lineage trace to very many cells, but is nearly exclusive to the outer retinal cells
272 marked by Visinin, indicating that ACR2's role in retinal development is specific to
273 photoreceptors, which are predicted to be cones at this timepoint. The lineage tracing of ECR
274 42 illustrates that this enhancer's activity is not limited to the cone and horizontal cell fate and
275 marks a much broader RPC population, as evidenced by the pan-retinal distribution of cells with
276 a history of this regulatory element (Additional File 5).
277
278 At the conclusion of our screen, eighteen active regulatory elements were identified in the
279 retina, two of which drove GFP reporter expression in a spatial pattern that excluded
280 ThrbCRM1(+) cells. ACR 10.B was found to drive expression non-specifically throughout the
281 retina and three elements, ECR 26; ECR 29; and ECR 35, drove expression too weakly to

282 determine their specificity. ECR 22 was found to mark a population that excludes cone
283 photoreceptors (Gonzalez and Schick et. al., in preparation). ACR2, ECR9, and ECR65 best met
284 the criteria for our screen. These elements are either specific or strongly biased to the
285 ThrbCRM1(+) population, mark early retinal RPC cells expressing Onecut1, and are active in cells
286 that give rise to three early retinal fates: cone photoreceptors; horizontal cells; and RGCs. As
287 the lineage marked by ACR2 is largely specific only to a smaller population of cone
288 photoreceptors, the remainder of this study is focused on evaluation of ECR 9 and ECR 65.
289
290 ***Bioinformatic Analysis of Regulatory Element Sequences:***
291 Though reporter assays indicate that ECR65 and ECR9 drive transcription in the ThrbCRM1(+)
292 population during the cone and HC specification windows in the chick retina, we do not know
293 what role these regulatory elements play in the GRN that gives rise to cone photoreceptors and
294 horizontal cells.
295
296 To determine which transcription factors bind to ECR 65 and ECR 9, we first attempted to
297 identify conserved motifs present within the sequence. We used UCSC Blat (Kent, 2002) to find
298 homologous sequences to the originally identified chick ECR 65 sequence from the golden
299 eagle, barn owl, American alligator, thirteen-lined ground squirrel, northern treeshrew,
300 chimpanzee, and human. ECR 65 is well-conserved among all of the avian species as well as the
301 American alligator and is conserved between avians and mammals.
302

303 When searching for the 515 bp chick sequence in the mouse genome, UCSC BLAT returned a
304 214 bp homologous stretch (Additional File 6A). This 214 bp mouse sequence contained a 58 bp
305 stretch that did not align to the chick sequence. In contrast with chick ECR 65, which is wholly
306 located within accessible chromatin, only 150 bp of the homologous mouse sequence are in
307 accessible chromatin region. The accessible chromatin extends past the homologous sequence.
308 Mouse ECR 65 (mECR65), as defined by chromatin accessibility, is 295 bp long (Additional File
309 6A). In summary, ECR 65 demonstrated both a large sequence divergence and an apparent shift
310 in chromatin accessibility between avian and mammalian species.
311
312 To assay this 295 bp mECR65 element for activity, we cloned it into the Stagia3 vector. When
313 mECR65::GFP was electroporated into the chick retina at E5 along with ThrbCRM1::AU1,
314 mECR65(+) cells were observed in the ThrbCRM1 population, which suggests that despite the
315 sequence divergence, mECR65 has retained the regulatory information for activity in these
316 cells. MEME motif analysis (Bailey and Elkan, 1994) of the various species-specific ECR 65
317 sequences revealed that three motifs appeared conserved between avian and mammalian
318 species (Additional File 6B). It is therefore likely the TF binding sites important for ECR 65
319 activity are within the conserved Motifs 1, 5, and 2 (Figure 4).
320
321 The ECR 9 sequence is within accessible chromatin in both chick and mouse early retinal cells.
322 The ECR 9 sequence used was cloned out of the mouse genome and as seen above, is able to
323 drive transcription of a reporter in chick retinal cells during the peak of cone and horizontal cell
324 development. Though the sequence tested is over 400bp long, only a ~270 bp span is conserved

325 between mouse and chick. Motif analysis of this enhancer through MEME returns seven motifs,
326 of which six are conserved between birds and mammals (Figure 4, Additional Files 6C and 7).
327
328 ***Deletions and Mutations of Regulatory Element Sequences***:
329
330 To take an unbiased approach to determine which regions of ECR 65 are functional, five serial
331 deletions of the chick ECR 65 sequence were tested for their ability to drive reporter activity
332 (Figure 4, Additional File 7). The full-length/wild type version of ECR65 driving TdTomato was
333 compared against the truncated versions driving GFP using flow cytometry. To ensure that any
334 observed effects were not due to changes in the proximity between TF binding sites and the
335 TATA box, all deletions were oriented such that the distance of all remaining sequence to the
336 TATA box was preserved and these deletion constructs were compared to the full-length
337 enhancer in the same orientation. To determine the effect of each deletion, we calculated the
338 percent of GFP(+) cells relative to the amount of TdT(+) cells. WT ECR65::GFP marks about 56-
339 65% as many cells as WT ECR65::TdT (Additional File 8). When these control values were
340 normalized to 100%, ECR 65 Deletions 1-3 respectively marked 33.62%, 42.05% and 8.2% as
341 many cells as the control WT ECR65::GFP (Figure 4A). When deletions were made on the other
342 end of the regulatory element, GFP driven by ECR65 Deletion 4 or Deletion 5 respectively marks
343 10.82% and 5.68% as many cells as WT ECR 65. Region 4 also encompasses MEME-predicted
344 Motif 2. The severe loss of GFP expression upon deleting the 80 bp Region 4 suggests that
345 Region 4 of ECR 65 contributes significantly to the activity of this regulatory element.
346

347 ECR 65 Motif 2 within Region 4 contains a potential binding site for a bHLH transcription factor.
348 bHLH binding sites, known as E-boxes, typically follow the sequence CANNTG. Mutation of this
349 potential E-box sequence resulted in a significant loss of enhancer activity (Figure 4C, Additional
350 Files 7 and 8). This result suggests that the functional sequence within Region 4 may be this 6-
351 bp motif, predicted by TOMTOM (Gupta et al., 2007) to bind the transcription factor
352 NHLH1/NSCL1. Another mutation within Region 4 encompassing a predicted homeobox TF
353 binding site resulted in some loss of GFP reporter activity (Figure 4C, Additional File 8).
354
355 A similar deletion strategy was used to investigate ECR 9 and reporter constructs were tested
356 which removed the ECR9 sequences on either side of the MEME-identified motifs as well as
357 Motif 1, 2, 3, and 7 (Figure 4, Additional File 7). As a control, we calculated the percent of
358 GFP(+) cells relative to the amount of TdT(+) cells marked by full length versions of ECR9
359 (Additional File 8). Once again, we ensured that deletions were orientated away from the TATA
360 box and compared only to full-length ECR9 of the same orientation. Deletions of Region R,
361 Motif 3, and Motif 7 did not result in a significant change to ECR9 activity. However, deletions
362 of Motif 2, Motif1 and Region F all resulted in a decrease of ECR9 activity as compared to the
363 full-length enhancer. Examination of the sequence for putative TF binding sites led to four
364 predicted bHLH sites (Figure 4, Additional File 7). Site 1 and Site 2 are located in Region F and
365 Motif 2, respectively. Sites 3 and 4 are both located within Motif 3. We hypothesized that if any
366 of these sites were important for ECR9 function, mutation of one or more of them directly
367 would result in a change in reporter expression. Mutation of Site 4, located within Motif 3,
368 resulted in a 97.5% loss of ECR9 activity (Figure 4D, Additional Files 7 and 8).

369

370 ***Interactions with bHLH Transcription Factors:***

371 Our screen for cis-regulatory elements that could regulate Onecut1 expression in early
372 restricted RPCs led to ECR9 and ECR65, which appear to be active in overlapping populations of
373 early RPCs that give rise to distinct subsets of retinal cell types. The tested serial deletions and
374 mutations suggest that both of these elements have critical regulatory input from bHLH family
375 transcription factors. In addition to the bioinformatically predicted Nhlh1 binding site in ECR65,
376 bulk RNA-seq indicated that NeuroD1, NeuroD4, NeuroG2, and Atoh7 transcripts were enriched
377 in the ThrbCRM1(+) population (Buenaventura et. al., 2018) and therefore also candidates to
378 interact with these two cis-regulatory elements. To explore these possibilities, each of the five
379 bHLH factors was overexpressed under the control of the ubiquitous CAG promoter in the E5
380 chick retina along with ECR65 and ECR9.

381

382 Under control conditions in which CAG did not drive any open reading frame, we calculated the
383 percent of ECR9(+) or ECR65(+) cells out of the total electroporated population. Overexpression
384 of either NeuroD1, NeuroG2, and NeuroD4 induced an increase in ECR9 reporter output, while
385 Nhlh1 and Atoh7 did not. Confocal microscopy showed that the increase in ECR9(+) cells upon
386 overexpression NeuroD1, NeuroG2, and NeuroD4 was predominantly in the inner retina (Figure
387 5B). However, individual overexpression of these three bHLH factors is not sufficient to increase
388 activity of the ECR9 Site 4 mutation (Figure 5D).

389

390 We detected no change to the ECR65 reporter output upon misexpression of NeuroD1,
391 NeuroD4, NeuroG2, or Atoh7 (Figure 5A). However, overexpression of Nhlh1 led to a significant
392 increase in the number of GFP(+) cells. Confocal microscopy data showed that the population
393 marked by ECR65 appears to expand towards the inner retina upon Nhlh1 overexpression, with
394 the morphology of some cells resembling multipotent RPCs marked by Vsx2 (Figure 5A, 5B).
395 Furthermore, Nhlh1 overexpression was unable to rescue the loss of GFP activity seen in
396 ECR65bHLH Mut1::GFP, suggesting that the site CATCAG within ECR65 is not only required for
397 regulatory activity, but possibly mediates the interaction between ECR65 and Nhlh1 (Figure 5C).
398
399 Despite their effects on ECR9 and ECR65 activity, none of the four candidate bHLH factors were
400 able to increase GFP reporter output driven by ThrbCRM1 (Figure 5A) or drive any changes in
401 the spatial activity of the enhancer in the chick retina at E5 (Figure 5B). In addition, none of the
402 bHLH genes were sufficient to ectopically induce ThrbCRM1 activity in the P0 mouse retina,
403 suggesting that these genes were not sufficient to induce Onecut1 expression at this time
404 (Additional File 9). These results suggest that either ECR9 and ECR65 activation may occur after
405 ThrbCRM1 activation or that the interactions between the bHLH factors and enhancers are not
406 sufficient to induce Onecut1 expression.
407
408 ***Timeline of Regulatory Element Activity:***
409 ECR65::TdT and ECR9::GFP were electroporated with ThrbCRM1::GFP or ThrbCRM1::TdT
410 respectively into retinas that were then cultured for 8 hours to assess whether the two
411 regulatory elements activated at the same time as ThrbCRM1. ECR65, ECR9, and ThrbCRM1 are

412 all able to drive reporter expression 8 hours after electroporation. At this early time point,
413 similar to the 18-22 hour time point, all ECR65(+) cells are ThrbCRM1(+). ECR9 activity at 8
414 hours is also similar to its activity at 18-22 hours, as there are both ThrbCRM1(+) and
415 ThrbCRM1(-) cells in this population (Figure 6A). Some of the ECR9(+) ThrbCRM1(-) cells in the
416 inner part of the developing retina also appeared to be Vsx2(+) or EdU(+) (Figure 6B, 6C). These
417 results indicate that onset of ECR9 and ECR65 activity is not later than or dependent on
418 ThrbCRM1 activation.
419
420 ***Molecular Events Upstream of ThrbCRM1 Activity:***
421 We then sought to determine whether any of the bHLH factors which impact ECR9 and ECR65
422 activity affected one or more of the factors upstream of ThrbCRM1 such as Onecut1 and Otx2,
423 or affected expression of the multipotent gene Vsx2. Overexpression of NeuroD1, NeuroD4, or
424 NeuroG2 individually resulted in an increase of electroporated Otx2(+) cells (Figure 7A).
425 Together with the increase in ECR9::GFP(+) cells, this would suggest that the newly GFP(+) cells
426 are expressing Otx2 and indeed there is no change in the proportion of GFP(+) Otx2(+) cells
427 upon overexpression of each bHLH factor (Figure 7A). We then examined whether this
428 enhancer-marked population shared a relationship with the Vsx2(+) multipotent RPC
429 population. Vsx2, shown to be largely absent in the ThrbCRM1(+) restricted RPC population,
430 marks 48.8% of all electroporated cells after one day in culture while ECR9(+) Vsx2(+) comprise
431 less than 1% of all electroporated cells. Overexpression of bHLH factors led to an increase in
432 ECR9(+)Vsx2(+) cells. There was a trend for Vsx2(+) cells within the electroporated population
433 to decrease upon bHLH overexpression but this was not statistically significant (Figure 7A).

434

435 Overexpression of Nhlh1 does not result in any change to electroporated Otx2(+) cells nor to
436 those cells positive for both Otx2 and ECR65::GFP. Staining with Vsx2 confirmed that there is an
437 increase in ECR65(+) Vsx2(+) cells upon overexpression of Nhlh1 (Figure 7B), as suggested by
438 the data from Figure 5, though the overall percent of Vsx2(+) cells out of the total
439 electroporated population does not change. This result may indicate that Nhlh1 overexpression
440 induces some ECR65 activity within the Vsx2(+) cell population.

441

442 Onecut1 staining was used to qualitatively assess the relationship between the enhancers
443 ECR65 and ECR9, the four candidate bHLH factors, Vsx2, and Onecut1. ECR9(+) Onecut1(+) cells
444 appear biased towards the outer retina, where ECR9(+) ThrbCRM1(+) cells are also observed
445 (Figure 2A), whereas ECR9(+) Vsx2(+) cells are more common towards the inner retina (Figure
446 7C). This does not change when NeuroD1, NeuroD4, or NeuroG2 are overexpressed. However,
447 it appears that many of the strongly GFP-pos cells that appear in the inner retina in response to
448 bHLH overexpression are neither Onecut1(+) nor Vsx2(+).

449

450 We observed that under control conditions ECR65 activity overlaps greatly with Onecut1
451 expression and there are very few ECR65(+) cells which express both OC1 and Vsx2 (Figure 7D).
452 However, ECR65(+) cells which express Vsx2 but not Onecut1 appeared more frequently upon
453 Nhlh1 overexpression.

454

455 ***Relationship Between bHLH Factors and Early-born Retinal Cell types***

456 We investigated the possibility of these factors affecting retinal cell types developing from E5 to
457 E7. Previously it has been shown that overexpression of Onecut1 in the postnatal mouse retina
458 can induce an increase in cones and horizontal cells while suppressing the rod photoreceptor
459 fate (Emerson et al., 2013). Here, we overexpressed the four candidate bHLH factors, cultured
460 for two days to mirror the lineage tracing experiments, and stained for the same cell-type
461 specific markers (Additional File 10) to determine whether these bHLHs play a role in the
462 development of the cell types marked by the lineage tracing of ECR65 and ECR9.The lineage
463 tracing of ECR9 in conjunction with the data from overexpression of NeuroD1, NeuroD4, and
464 NeuroG2 suggests that these one or more of these transcription factors may play a role in the
465 development of Lim1(+) horizontal cells or RGCs. However, no conclusive increase of these cell
466 types was observed upon bHLH overexpression. Consistent with previously reported data (Li et
467 al., 1999) overexpression of Nhlh1 was not sufficient to induce an increase in Isl1(+) cells. Lastly,
468 none of the four candidate bHLH factors had an effect on photoreceptors marked by Visinin.
469

470 **Discussion:**

471 The work presented here is intended to serve as a step towards understanding the distinction
472 between cell fate multipotency and restriction. The expression of multipotent RPC genes such
473 as Vsx2 and genes such as Onecut1 that mark the ThrbCRM1(+) fate-restricted RPC population
474 appear to be mutually exclusive (Buenaventura et al., 2018). It is not understood what gene
475 regulatory networks are involved in the establishment of these two populations. We therefore
476 sought to identify cis-regulatory elements specific to the RPC[CH] population that function
477 upstream of Onecut1 and may be involved in the restriction process.

478

479 Our large-scale, multi-step screen for regulatory elements resulted in the identification of three
480 regulatory elements that drove a spatial expression pattern biased to the cone/HC restricted
481 RPC population. In addition to ECR9 and ECR65, we found multiple regulatory elements which
482 drove expression in non-specific expression patterns and two which completely excluded the
483 population of interest. These two regulatory elements, ACR8 and ACR10-A, may be involved in
484 generation or maintenance of the multipotent RPC population that gives rise to a broader range
485 of mature cell types as well as the restricted RPC populations. It may also be that the more
486 broadly-acting or ThrbCRM1(-) elements from this screen require further genomic context to
487 act specifically in Onecut1(+) restricted RPCs. Our reporter assay is demonstrably effective in
488 finding minimal elements that can drive specific spatiotemporal expression patterns, but we
489 cannot be certain that we replicate the genomic function of every assayed element given the
490 importance of chromatin state and the surrounding sequences. Our assay does capture some
491 aspects of genomic context as every active element is found within accessible chromatin
492 regardless of sequence conservation. Highly conserved elements located within closed
493 chromatin in E5 chick retinal cells were not found to be active. Despite this, differential
494 chromatin accessibility may not be a strong indicator of enhancer specificity in cell populations
495 at the same developmental time point. ECR65 in particular is accessible in both ThrbCRM1(+)
496 and ThrbCRM1(-) cells, yet under normal conditions is only active in the ThrbCRM1(+)
497 population, demonstrating that appropriate combinations of transcription factors are still
498 required to confer specific activity to regulatory elements.

499

500 The regulatory elements ECR9 and ECR65 mark distinct but overlapping populations of
501 developing chick retinal cells from E5-E6, which includes RPCs (Figure 2). ECR65 has also been
502 identified through DNaseI hypersensitivity in the mouse retina as “OC1 A” (Nadadur et al.,
503 2019) but has not been further characterized. Here, we report that the ECR65(+) population
504 overlaps almost entirely with the ThrbCRM1(+) population of cone/HC RPCs (Figure 2) and
505 excludes the Vsx2(+) multipotent RPC population (Figure 7). By combining bioinformatic
506 analyses of the active enhancer sequences with functional tests and previously published RNA-
507 seq data, we were able to connect the activity of ECR9 and ECR65 to bHLH transcription factors.
508
509 The overexpression assays demonstrate specificity in the interactions between the four
510 candidate bHLH factors and the two candidate enhancers. The involvement of these
511 transcription factors in retinal development and specifically in retinal cell fate choice has also
512 been well-documented (Cepko, 1999; Hatakeyama and Kageyama, 2004; Dennis et al., 2019).
513 For example, Nhlh1 RNA is known to be present in Isl1(+) RGCs (Li et al., 1999) and has been
514 observed in scRNA-seq from chick retinal cells to mark a cone photoreceptor subtype (Ghinia-
515 Tegla et al., 2019). When overexpressed alongside both candidate enhancers, Nhlh1 is only able
516 to affect the activity of ECR65 (Figure 5A, 5B) which lineage traces to an Isl1(+) population
517 (Figure 3). However, further investigation is required to identify and characterize Nhlh1(+)
518 developing cone photoreceptors and determine whether they develop from the ECR65(+)
519 population. Similarly, NeuroG2 has been found to be important for RGC genesis (Hufnagel et al.,
520 2010; Maurer et al., 2018) and involved in HC fate choice (Akagi et al., 2004). This role may

521 underlie its interaction with ECR9 (Figure 5A, 5B), which lineage traces to horizontal cells and
522 morphologically characteristic RGCs (Figure 3).
523
524 The ECR9(+) population also distinguishes itself with a subpopulation that does not overlap with
525 ThrbCRM1 activity. However, this ECR9 (+) ThrbCRM1(-) population that includes RPCs also
526 does not overlap strongly with the multipotent Vsx2(+) population(Figure 6, Figure 7). Though
527 NeuroD1, NeuroG2, and NeuroD4 were considered candidate TFs largely due to their
528 enrichment in the ThrbCRM1(+) population (Buenaventura et. al., 2018), our data did not
529 indicate that overexpression of individual candidate bHLH factors affected ThrbCRM1 activity
530 on the timescale that we examined. Both the quantitative and qualitative assessment along
531 with the visible increase in both Otx2(+) and ECR9(+) cells in the inner retina suggest that
532 NeuroD1, NeuroD4, and/or NeuroG2 mediate ECR9 activity in a population of Otx2(+) cells
533 distinct from the ThrbCRM1(+) restricted RPC population.
534
535 It has been hypothesized that an intermediate restricted RPC, marked by ThrbICR, gives rise to
536 ThrbCRM1(+) restricted RPCs (Schick et al., 2019). Lineage tracing of ThrbICR, which is bound
537 by NeuroD1 (Liu et al., 2008), shows that the cells marked by that element can give rise to RGCs
538 as well as cones and horizontal cells (Schick et al., 2019). In conjunction with the ECR9 lineage
539 tracing data and NeuroD1 overexpression data, ECR9 activity in cells which are neither
540 ThrbCRM1(+) nor Vsx2(+) suggests that ECR9 could also label this intermediate population,
541 referred to as RPCs[CHG] (Figure 8A).
542

543 The regulatory elements uncovered in this and previous screens can therefore be used to mark
544 distinct developing populations in the early vertebrate retina. OC1 ECR65 and ThrbCRM1 are
545 both active in the restricted RPC[CH], while OC1 ECR9 and ThrbICR activity may be marking the
546 hypothesized RPC[CHG] which is distinct from multipotent RPCs characterized by the activity of
547 VSX2 ECR4 (Buenaventura et al., 2018). Other elements found in the screen, such as ACR2
548 (Figure 2, Additional File 5) and ECR22 (Gonzalez and Schick et. al., in preparation), may be
549 more specific to either the cone photoreceptor precursors or horizontal cell precursors,
550 respectively, from the ThrbCRM1 lineage (Figure 8A). Though ECR9(+) and ECR65(+) cell
551 populations overlap with each other and other CREs such as ThrbCRM1 and potentially ThrbICR,
552 the two enhancers characterized here are unique in their responses to the bHLH factors
553 described above. We hypothesize that the interactions between ECR9 and ECR65 and their
554 respective bHLHs play a role in the transition from multipotent RPCs to the fate-restricted
555 RPC[CH] population (Figure 8B). It may be that one or more of the bHLH factors shown to
556 interact with ECR9 may play a role in coordinating Vsx2 downregulation and activation of
557 Onecut1 expression and ThrbCRM1 activity (Figure 8B).
558
559 While we have not determined which factor(s) among NeuroD1, NeuroD4 and NeuroG2
560 interact with ECR9 *in vivo*, the ability of all three factors to increase ECR9 activity is consistent
561 with a previous study demonstrating their functional redundancy during retinogenesis (Akagi et
562 al., 2004). As ECR9 has multiple putative E-box sites, it may be that some combination of two or
563 all three factors is required for the correct regulatory output of this enhancer. Of the ECR9
564 bHLH sites, Sites 2-4 all vary in sequence but all three sites and their flanking sequences are

565 highly conserved between mouse, chick and human (Additional File 7). It is also worth noting
566 that while the ECR65 bHLH site shares the same sequence as ECR 9 Site 3, the two enhancers
567 still differ in their ability to respond to Nhlh1 overexpression. This may be due to the
568 differences in their flanking sequences. Previous work has shown that some bHLH factors are
569 able to utilize each other's binding sites (Mao et al., 2013) and that the sequence flanking the
570 core E-box may also be important for binding affinity (Gordân et al., 2013).
571
572 Our results do not suggest that any of these bHLH factors individually are capable of inducing or
573 suppressing any of the early-born cell fates in the chick retina. Previous studies using Xenopus
574 were able to yield an increase in photoreceptors upon overexpression of NeuroD1 and
575 NeuroD4 (Wang and Harris, 2005). In chick, NeuroD1 has been reported to induce more
576 photoreceptors when virally overexpressed in the chick retina at E2 and cultured until E8.5/E9
577 (Yan and Wang, 1998). Our lack of a similar result may be due to differences in the
578 experimental timepoints, as our experimental conditions included a maximum culture time of
579 48 hours. Though we were also not able to induce horizontal cell or RGC fates, it may be
580 because the overexpressed factors require the co-expression of other TFs in order to specify or
581 induce particular cell fates. For instance, the prediction of a homeobox TF binding site so close
582 to the bHLH binding site in ECR65 may suggest that both Nhlh1 and a homeobox factor are
583 required to induce the cell fates observed from lineage tracing ECR65. Future studies are
584 required in which bHLH factors are overexpressed in combination with each other and with
585 homeobox factors to determine which factors are sufficient to drive early retinal cell fates in
586 chick.

587

588 **Conclusions:** This study examined the upstream regulatory events of the Onecut1 gene that
589 occur in chick RPCs. Guided by both sequence conservation and chromatin accessibility, we
590 identified two regulatory elements near the Onecut1 gene, ECR9 and ECR65, that are
591 preferentially active in Onecut1-expressing ThrbCRM1(+) RPCs. We find that both of these
592 elements are predicted to contain bHLH transcription factor binding sites, which are required
593 for activity of these elements. Overexpression of specific bHLH members leads to ectopic
594 activity of ECR9 and ECR65 in Vsx2(+) RPCs. These bHLH factors are able to upregulate
595 endogenous Otx2 expression, a protein normally expressed in fate-restricted RPCs. Taken
596 together, these results suggest a role for bHLH factors in promoting the formation of fate-
597 restricted RPCs from multipotent RPCs through the activation of Onecut1 and Otx2.

598

599 **Abbreviations:**

600 GRN – gene regulatory network

601 RPC – retinal progenitor cell

602 HC – horizontal cell

603 RPC[CH] – restricted retinal progenitor cell that gives rise to cones, horizontal cells

604 RPC[CHG] – restricted retinal progenitor cell that gives rise to cones, horizontal cells, retinal
605 ganglion cells

606 CRE – cis-regulatory element

607 TF – transcription factor

608 GFP – green fluorescent protein

609 TdT – TdTomato

610

611 **Declarations:**

612 **Acknowledgements:**

613 Diego Buenaventura provided guidance with the bioinformatic analyses and Sean McCaffery
614 generated the lineage tracing plasmids for ECR9 and ECR42. Miruna Ghinia-Tegla, Estie Schick,
615 Xueqing Chen and the rest of the Emerson Lab provided valuable feedback. The Lim1 and Isl1
616 antibodies developed by TM Jessell and S. Brenner-Morton and the Visinin antibody developed
617 by C. Cepko and S. Bruhn were obtained from the Developmental Studies Hybridoma Bank,
618 created by the NICHD of the NIH and maintained at The University of Iowa, Department of
619 Biology, Iowa City, IA 52242. Jeffrey Walker and Jorge Morales provided assistance with flow
620 cytometry and confocal microscopy, respectively.

621

622 **Funding:**

623 Support was provided to ME by a National Science Foundation grant 1453044 and a Sloan
624 Foundation Junior Faculty Research Award in Science and Engineering. BS was supported by
625 NSF DBI 1156512 to CCNY and core facility usage by 3G12MD007603-30S2 to CCNY.

626

627 **Availability of Data and Materials:**

628 ATAC-seq data generated from chick retinal cells has been submitted to GEO as data set
629 XXXXXX and will be available upon publication. ATAC-seq data generated from mouse retinal

630 cells is not currently available as it is part of another study but can be made available upon
631 request.
632
633 **Authors Contributions:**
634 ME and SP generated the ATAC-seq libraries, planned and conducted experiments and wrote
635 the manuscript. NJC led the initial screen shown in Figure 1 and Additional Files 1 and 2. AG, SG,
636 BW, and BS generated plasmids and processed tissue for Figures 1 and 2 and Additional Files 1
637 and 2. BS also generated mutagenized plasmids for Figure 4. SS generated plasmids and
638 processed tissue for Figures 4 and 5.
639
640 **Ethics approval and consent to participate**
641 The City College of New York Institutional Animal Care and Use Committee approved all animal
642 procedures under protocol 932.
643
644 **Competing interests**
645 The authors declare that they have no competing interests
646
647 **Consent for publication**
648 Not applicable
649
650 **Methods:**
651 **Animals**

652 Fertilized chick eggs were acquired from Charles River and stored at 16 °C room for a maximum
653 of ten days. Embryonic days were counted from E0 when eggs were moved to a 38 °C
654 humidified incubator for five days.
655
656 **ATAC-seq**
657 ATAC-seq libraries were collected and amplified as outlined by Buenrostro et al., 2013. Mouse
658 retinas were collected from embryonic day 12.5 (plug morning equal to time 0.5) embryos, and
659 dissociated using manual douncing. Libraries were analyzed for quality control on Bioanalyzer
660 and Qubit and then sequenced at a depth of 37.5 million reads per sample. Sequenced libraries
661 were prepared for analysis using FASTQGroomer (Blankenberg et al., 2010) on default settings
662 and then analyzed using Bowtie for Illumina (Langmead et al., 2009) with default settings
663 except -X 2000 and – m 1, through the usegalaxy.org (Afgan et al., 2018) web platform.
664 Resulting SAM files were converted to the BAM format (Li et al., 2009) and the BigWig (Kent et
665 al., 2010) format.
666
667 **Electroporation**
668 Retinae were electroporated ex vivo as previously described (Schick et al., 2019). CAG::reporter
669 plasmids were used at a concentration of 0.1ug/uL with the exception of CAG::mCherry, which
670 was used at a concentration of 0.04 ug/uL. Reporter plasmids under the control of a tissue-
671 specific enhancer (ex: ThrbCRM1, ECR9, ACR2, etc) were used at a concentration of 0.16 ug/uL
672 For lineage tracing experiments, all plasmids (recombinase, responder plasmid, co-
673 electroporation control) were used at a concentration of 0.1 ug/uL. Recombinase and

674 responder plasmids are described in Schick et. al., 2019. Plasmids used in initial identification
675 screen for cis-regulatory elements used miniprep scale DNA purification (Zymo Research, D4020
676 or 5 prime/Eppendorf, FastPlasmid kit) and all subsequent experiments used midiprep scale
677 DNA purification (Qiagen, 12145 or 12243).
678
679 **Alkaline Phosphatase Staining**
680 Retinas were harvested from culture and fixed in 4% paraformaldehyde (PFA), washed 3X in
681 PBS, and incubated in 1 mL NTM (pH 9.5) buffer for 15 minutes while shaking at a low speed
682 before addition of 1mL NTM with NBT(0.25 mg/mL) and BCIP (0.125 mg/mL). Retinae were
683 incubated with the AP substrates in the dark for 2-3 hours until the positive control was well-
684 stained.
685
686 **Immunohistochemistry**
687 Retinas were harvested from culture, prepared for cryosectioning, and 20 micron vertical
688 sections were collected as outlined in Schick et al., 2019. Sections were incubated for 10
689 minutes in 0.1% Tween (VWR, 97062-332) in PBS (PBT) and then blocked for 1 hour in 5% serum
690 (Jackson ImmunoResearch, Donkey - 017-000121, Goat - 005-000-121) in PBT at room
691 temperature prior to incubation with primary antibodies overnight at 4 C. Sections were
692 washed 3X with PBT, prior to blocking at room temperature for 30 min and then incubated with
693 secondary antibodies overnight at 4 C. DAPI was added at 1 ug/uL in PBT while washing off the
694 secondary antibodies. Sections were then mounted using Flouromount-G (Southern Biotech,
695 0100-01) and coverslips (VWR, 48393-106). Primary antibodies are listed in the table below. All

696 secondary antibodies were obtained from Jackson ImmunoResearch and suitable for multiple
697 labelling. All Alexa-conjugated secondary antibodies were used at dilution of 1:400 and Cy3-
698 conjugated secondary antibodies were used at a dilution of 1:250 from secondary antibody
699 stocks in 50% glycerol.

Antibody; dilution	Vendor	Catalog number
Chick anti-GFP 1:2000	Abcam	ab13970
Rabbit anti-GFP 1:1000	Invitrogen	A6455
Mouse anti-AU1 1:2000	Enzo	ENZ-ABS135-0200
Chick anti-Bgal 1:1000	Abcam	ab9361
Rabbit anti-RFP 1:250	Rockland	600-401-379
Mouse anti-Onecut1 1:200	Santa Cruz	Sc-376308
Goat anti-Otx2 1:500	R&D	AF1979
Sheep anti-Vsx2 1:200	Ex Alpha	x1180p
Mouse anti-Visinin 1:250	DHSB	764-s
Mouse anti-Lim1 1:15	DHSB	4F2-C
Mouse anti-Isl1 1:50	DHSB	39.3F7
Mouse anti-panBrn3 1:100	Santa Cruz	Sc-390781

700

701 **Microscopy and Cell Counting**

702 Images of whole AP-stained retinas were acquired using a Zeiss Axiozoom V16 microscope with
703 a 1X objective and Zen 2 Blue 2011 software. All confocal microscopy images were acquired
704 using a Zeiss LSM 710 inverted confocal microscope with a 40x oil immersion objective, 488nm
705 laser, 561 nm laser, 633 nm laser, 405 nm laser, and Zen Black 2015 21 SP2 software at a
706 resolution of 1024 x 1024, acquisition speed of 6 and averaging number of 2. For all confocal
707 images shown, Z-stacks consisting of 8-12 Z planes were collected and are shown as maximum
708 intensity projections. Cells were counted in the Fiji (Schindelin et al., 2012) distribution of
709 ImageJ, using the Cell Counter plug-in developed by Kurt De Vos. In the lineage-tracing
710 experiments and for the EdU assay, all GFP(+) cells which co-localized with DAPI were counted

711 first and cells positive for EdU or for cell-specific markers such as Visinin were counted among
712 this population. In both instances, multiple images per retina were analyzed if necessary to
713 reach at least 45-50 GFP(+) cells. Brightness and contrast were adjusted uniformly across each
714 image in Affinity Designer vector editor (Serif [Europe] Ltd).
715
716 **EdU Pulse and Detection**
717 5 uL of 10mM EdU solution was added to 1 mL of culture media during the last hour of
718 incubation before retinas are harvested as described above. EdU was detected using the Click-iT
719 Plus EdU Kit for Imaging (Invitrogen, C10640). The tissue was first incubated for 15 minutes in
720 0.1% Tween in PBS at room temperature before incubating with the EdU Reaction Cocktail for
721 30 min in the dark at room temperature. The EdU Reaction Cocktail was then removed with 3
722 washes of PBT prior to antibody staining.
723
724 **Deletions and Mutagenesis**
725 Deletions or truncated versions of regulatory element sequences were generated through the
726 use of PCR primers that began internally within the sequence and excluded the portions of the
727 sequence to be deleted. Mutant regulatory element sequences were generated using overlap
728 extension PCR, in which primers included short sequence mismatches at potential TF binding
729 sites. A second set of primers encompassed the ends of the regulatory element sequence and
730 the restriction sites within the plasmid template for easy cloning of the mutant sequence into
731 the Stagia3 reporter vector. Site 3 and Site 4 within ECR9 were mutated to the sequences
732 “TTAGAC” and “AGGCCA”. The bHLH site within ECR65 Region 4 was also mutated twice, to the

733 sequence “CTGATGAATGGCG” to include the E-box site, 4 bp upstream and 3 bp downstream
734 and to the sequence “TTTCCCAAAG” to include the E-box site and 4 bp upstream. (Figure 4,
735 Additional File 8).
736
737 **MEME-suite**
738 To identify conserved motifs within ECR65, we used MEME with settings to find a maximum of
739 12 motifs, with a width of 6-50 bp each and 2-9 sites per motif. For ECR9, we used settings to
740 find a maximum of 7 motifs, with a width of 6-50 bp each and 2-8 sites per motif. The output
741 from MEME was then used as input for TOMTOM under default settings.
742
743 **Multiple Sequence Alignments**
744 The alignments between the chick, mouse, and human sequences for ECR9 and ECR65
745 (Additional File 7) were produced in Clustal Omega (Madeira et al., 2019) version 1.2.4 with
746 default settings.
747
748 **Plasmids**
749 Plasmids containing coding sequences of candidate TF genes were obtained from Transomics.
750 Mouse Nhlh1 (Clone ID: BC051018) and mouse NeuroD4 (Clone ID: BC054391) were cloned
751 using EcoR1 to insert into a modified pCAG vector that allows for EcoR1 flanked insert cloning
752 (Emerson et al., 2013) while mouse NeuroD1 (Clone ID: BC018241), human NeuroG2 (Clone ID:
753 BC036847) and human Atoh7 (Clone ID: BC032621) were cloned using a combination of EcoR1
754 and Not1 (NeuroD1, NeuroG2, Atoh7) into pCAG::EGFP (Matsuda and Cepko, 2004) such that

755 each coding sequence is under the control of the CAG promoter. The following plasmids were
756 previously reported: CAG::OC1, ThrbCRM1::GFP, ThrbCRM1::AU1 plasmids (Emerson et al.,
757 2013); CAG::iRFP (Buenaventura et al., 2018); Bp::PhiC31 lineage tracing and CAaNa::GFP
758 responder plasmids (Schick et al., 2019); UbiC::TdTomato (Rompani and Cepko, 2008); and
759 TdTomato reporter plasmid (Jean-Charles et al., 2018). The CAG::mCherry and CAG::nucBgal
760 plasmids were constructed by Takahiko Matsuda and reported in (Wang et al., 2014) and
761 obtained from the Cepko lab, respectively. The ThrbCRM1::TdTomato plasmid was made by
762 ligating a Not1/EcoR1 fragment from ThrbCRM1::GFP into the TdTomato reporter plasmid
763 (Jean-Charles et al., 2018). Candidate Onecut1 cis-regulatory elements were amplified from
764 chick or mouse genomic DNA with Herculase II polymerase (Agilent, 600677-51), treated for 10-
765 30 minutes with Taq polymerase (Qiagen, 201203) to generate Adenine overhangs, and ligated
766 into PGemTeasy (Promega, A1360). Inserts were sequence verified by Sanger sequencing
767 (Genewiz) and moved into Stagia3 after EcoR1 digestion. In cases where elements contained
768 EcoR1 sites, the original sequences used to amplify candidate elements were used to generate
769 modified oligos with Xho1, Sal1, or Mfe1 restriction sites to allow for PCR-amplification of
770 elements from the verified PGemTeasy clones and subsequent insertion into an appropriately
771 digested Stagia3 plasmid. As candidate elements could be inserted into Stagia3 in two
772 orientations, for some elements both possible orientations were tested.
773
774 **Dissociation and Flow Cytometry**
775 Upon harvest, retinae were dissociated into single cells as described in Schick et al, 2019 using
776 papain (Worthington, LS003126) and an activation solution of L-cysteine (VWR, 97063-478) and

777 10mM EDTA at 37 C. 10% FBS (ThermoFisher, A3160602) solution in DMEM (Life Technologies,
778 11995-073) was used to stop the dissociation .Cells were further digested with DNaseI (Sigma,
779 4536282001) and subsequently washed in DMEM prior to fixing in 4% PFA. Dissociated cells
780 were then analyzed on a BD LSRII machine using the 488nm, 561nm, and 633nm lasers. The
781 collected data was analyzed using FlowJo version 10.4.2.
782

783 **Statistical tests**

784 Statistical tests were conducted using GraphPad Prism. Data sets were tested for normality
785 (Shapiro-Wilk) prior to ANOVA, Kruskal-Wallis, or t-tests. Significant Kruskal-Wallis or ANOVA p-
786 values were followed up with Dunn's or Dunnett's post hoc test, respectively.

Figure	Exp. Condition	Dunnett p-value	Dunn p-value	Unpaired t-test
Figure 4a	ECR65 Deletion 3		0.0011	
	ECR65 Deletion 4	< 0.0001		
	ECR65 Deletion 5	< 0.0001		
Figure 4c	ECR9 Deletion F			0.0016
	ECR9 Deletion 1			0.006
	ECR9 Deletion 2			0.0025
Figure 4b	ECR65 bHLH Mut A	< 0.0001		
	ECR65 hb Mut	0.0242		
Figure 4d	ECR9 site 3 Mut	0.0007		
	ECR9 site 4 Mut	0.001		
Figure 5a	ECR65 + Nhlh1	0.046		
	ECR9 + NeuroD1	0.0019		
	ECR9 + NeuroG2	0.0085		
	ECR9 + NeuroD4	0.021		
Figure 5c	ECR65 + Nhlh1	0.046		
Figure 5d	ECR9 + NeuroD1	0.0137		
	ECR9 + NeuroD4	0.01		
	ECR9 + NeuroG2	0.0164		
Figure 7a	Nhlh1 Vsx2 + GFP			0.035
	ECR9 + NeuroD1 Otx2	0.0245		
	NeuroD1 Vsx2 + GFP	0.0287		
	ECR9 + NeuroG2 Otx2	0.0098		
	ECR9 + NeuroD4 Otx2	0.0069		
	NeuroD4 Vsx2 + GFP	0.037		

787

788 **References**

789 Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Cech, M., Chilton, J., Clements, D.,
790 Coraor, N., Grüning, B.A., et al. (2018). The Galaxy platform for accessible, reproducible and
791 collaborative biomedical analyses: 2018 update. Nucleic Acids Res. *46*, W537–W544.

792 Akagi, T., Inoue, T., Miyoshi, G., Bessho, Y., Takahashi, M., Lee, J.E., Guillemot, F., and
793 Kageyama, R. (2004). Requirement of Multiple Basic Helix-Loop-Helix Genes for Retinal
794 Neuronal Subtype Specification. J. Biol. Chem. *279*, 28492–28498.

795 Bailey, T.L., and Elkan, C. (1994). Fitting a Mixture Model By Expectation Maximization To
796 Discover Motifs In Biopolymer. Proc. Int. Conf. Intell. Syst. Mol. Biol. *2*, 28–36.

797 Billings, N.A., Emerson, M.M., and Cepko, C.L. (2010). Analysis of thyroid response element
798 activity during retinal development. PloS One *5*, e13739.

799 Blankenberg, D., Gordon, A., Von Kuster, G., Coraor, N., Taylor, J., and Nekrutenko, A. (2010).
800 Manipulation of FASTQ data with Galaxy. Bioinformatics *26*, 1783–1785.

801 Buenaventura, D.F., Ghinia-Tegla, M.G., and Emerson, M.M. (2018). Fate-restricted retinal
802 progenitor cells adopt a molecular profile and spatial position distinct from multipotent
803 progenitor cells. Dev. Biol. *443*, 35–49.

804 Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition
805 of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding
806 proteins and nucleosome position. Nat. Methods *10*, 1213–1218.

807 Cepko, C.L. (1999). The roles of intrinsic and extrinsic cues and bHLH genes in the determination
808 of retinal cell fates. Curr. Opin. Neurobiol. *9*, 37–46.

809 Dennis, D.J., Han, S., and Schuurmans, C. (2019). bHLH transcription factors in neural
810 development, disease, and reprogramming. Brain Res. *1705*, 48–65.

811 Emerson, M.M., Surzenko, N., Goetz, J.J., Trimarchi, J., and Cepko, C.L. (2013). Otx2 and
812 Onecut1 Promote the Fates of Cone Photoreceptors and Horizontal Cells and Repress Rod
813 Photoreceptors. Dev. Cell *26*, 59–72.

814 Fekete, D.M., Perez-Miguelsanz, J., Ryder, E.F., and Cepko, C.L. (1994). Clonal Analysis in the
815 Chicken Retina Reveals Tangential Dispersion of Clonally Related Cells. Dev. Biol. *166*, 666–682.

816 Ghinia-Tegla, M.G., Buenaventura, D.F., Kim, D.Y., Thakurdin, C., Gonzalez, K.C., and Emerson,
817 M.M. (2019). Single cell profiling of CRISPR/Cas9-induced OTX2 deficient retinas reveals fate
818 switch from restricted progenitors. BioRxiv 538710.

819 Gordân, R., Shen, N., Dror, I., Zhou, T., Horton, J., Rohs, R., and Bulyk, M.L. (2013). Genomic
820 regions flanking E-box binding sites influence DNA binding specificity of bHLH transcription
821 factors through DNA shape. Cell Rep. *3*, 1093–1104.

822 Gupta, S., Stamatoyannopoulos, J.A., Bailey, T.L., and Noble, W. (2007). Quantifying similarity
823 between motifs. Genome Biol. *8*, R24.

824 Hafler, B.P., Surzenko, N., Beier, K.T., Punzo, C., Trimarchi, J.M., Kong, J.H., and Cepko, C.L.
825 (2012). Transcription factor Olig2 defines subpopulations of retinal progenitor cells biased
826 toward specific cell fates. Proc. Natl. Acad. Sci. *109*, 7882–7887.

827 Hatakeyama, J., and Kageyama, R. (2004). Retinal cell fate determination and bHLH factors.
828 Semin. Cell Dev. Biol. *15*, 83–89.

829 Hufnagel, R.B., Le, T.T., Riesenberg, A.L., and Brown, N.L. (2010). Neurog2 controls the leading
830 edge of neurogenesis in the mammalian retina. Dev. Biol. *340*, 490–503.

831 Jean-Charles, N., Buenaventura, D.F., and Emerson, M.M. (2018). Identification and
832 characterization of early photoreceptor cis-regulatory elements and their relation to Onecut1.
833 Neural Develop. *13*, 26.

834 Kent, W.J. (2002). BLAT—The BLAST-Like Alignment Tool. Genome Res. *12*, 656–664.

835 Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, and
836 D. (2002). The Human Genome Browser at UCSC. Genome Res. *12*, 996–1006.

837 Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S., and Karolchik, D. (2010). BigWig and BigBed:
838 enabling browsing of large distributed datasets. Bioinformatics *26*, 2204–2207.

839 Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient
840 alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

841 Li, C.-M., Yan, R.-T., and Wang, S.-Z. (1999). Misexpression of cNSCL1 Disrupts Retinal
842 Development. Mol. Cell. Neurosci. *14*, 17–27.

843 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and
844 Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*,
845 2078–2079.

846 Liu, H., Etter, P., Hayes, S., Jones, I., Nelson, B., Hartman, B., Forrest, D., and Reh, T.A. (2008).
847 NeuroD1 Regulates Expression of Thyroid Hormone Receptor β2 and Cone Opsins in the
848 Developing Mouse Retina. J. Neurosci. *28*, 749–756.

849 Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey,
850 A.R.N., Potter, S.C., Finn, R.D., et al. (2019). The EMBL-EBI search and sequence analysis tools
851 APIs in 2019. Nucleic Acids Res. *47*, W636–W641.

852 Mao, C.-A., Cho, J.-H., Wang, J., Gao, Z., Pan, P., Tsai, W.-W., Frishman, L.J., and Klein, W.H.
853 (2013). Reprogramming amacrine and photoreceptor progenitors into retinal ganglion cells by
854 replacing Neurod1 with Atoh7. Development *140*, 541–551.

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

857 Maurer, K.A., Kowalchuk, A., Shoja-Taheri, F., and Brown, N.L. (2018). Integral bHLH factor
858 regulation of cell cycle exit and RGC differentiation. Dev. Dyn. *247*, 965–975.

859 Nadadur, R.D., Perez-Cervantes, C., Lonfat, N., Smith, L.A., Hughes, A.E.O., Wang, S., Corbo, J.C.,
860 Cepko, C., and Moskowitz, I.P. (2019). Enhancer transcription identifies cis-regulatory elements
861 for photoreceptor cell types. BioRxiv 513598.

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

865 Ovcharenko, I., Nobrega, M.A., Loots, G.G., and Stubbs, L. (2004). ECR Browser: a tool for
866 visualizing and accessing data from comparisons of multiple vertebrate genomes. Nucleic Acids
867 Res. *32*, W280-286.

868 Rompani, S.B., and Cepko, C.L. (2008). Retinal progenitor cells can produce restricted subsets of
869 horizontal cells. Proc. Natl. Acad. Sci. U. S. A. *105*, 192–197.

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

873 Schick, E., McCaffery, S.D., Keblish, E.E., Thakurdin, C., and Emerson, M.M. (2019). Lineage
874 tracing analysis of cone photoreceptor associated cis-regulatory elements in the developing
875 chicken retina. Sci. Rep. *9*, 1–14.

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

879 Suzuki, S.C., Bleckert, A., Williams, P.R., Takechi, M., Kawamura, S., and Wong, R.O.L. (2013).
880 Cone photoreceptor types in zebrafish are generated by symmetric terminal divisions of
881 dedicated precursors. Proc. Natl. Acad. Sci. U. S. A. *110*, 15109–15114.

882 Trimarchi, J.M., Harpavat, S., Billings, N.A., and Cepko, C.L. (2008). Thyroid hormone
883 components are expressed in three sequential waves during development of the chick retina.
884 BMC Dev. Biol. *8*, 101.

885 Turner, D.L., and Cepko, C.L. (1987). A common progenitor for neurons and glia persists in rat
886 retina late in development. Nature *328*, 131–136.

887 Wang, J.C.-C., and Harris, W.A. (2005). The role of combinational coding by homeodomain and
888 bHLH transcription factors in retinal cell fate specification. Dev. Biol. *285*, 101–115.

889 Wang, S., Sengel, C., Emerson, M.M., and Cepko, C.L. (2014). A gene regulatory network
890 controls the binary fate decision of rod and bipolar cells in the vertebrate retina. Dev. Cell *30*,
891 513–527.

892 Wong, L.L., and Rapaport, D.H. (2009). Defining retinal progenitor cell competence in Xenopus
893 laevis by clonal analysis. Development *136*, 1707–1715.

894 Wu, F., Li, R., Umino, Y., Kaczynski, T.J., Sapkota, D., Li, S., Xiang, M., Fliesler, S.J., Sherry, D.M.,
895 Gannon, M., et al. (2013). Onecut1 Is Essential for Horizontal Cell Genesis and Retinal Integrity.
896 J. Neurosci. *33*, 13053–13065.

897 Yan, R.-T., and Wang, S.-Z. (1998). neuroD induces photoreceptor cell overproduction in vivo
898 and de novo generation in vitro. J. Neurobiol. *36*, 485–496.

899 Young, R.W. (1985). Cell differentiation in the retina of the mouse. Anat. Rec. *212*, 199–205.

900

901 **Figure 1. Identification and initial screening for regulatory elements active in E5 chick retinae.**

902 (A) Generation of ATAC-seq libraries of ThrbCRM1-positive and ThrbCRM1-negative cell

903 populations. Chick retinae at embryonic day 5 (E5) were electroporated ex vivo with both

904 ThrbCRM1::GFP and UbiqC::TdT plasmids and incubated in culture for 18-22 hours prior to

905 dissociation. Dissociated cells were sorted via FACS into GFP and TdT double-positive cells, and

906 GFP-negative, TdT-positive cells. Each population was processed for ATAC-seq. (B) Visualization

907 of aligned ATAC-seq reads to the galGal5 genome in UCSC Genome Browser in intergenic region

908 between Onecut1 and WDR72 (labelled). Active regulatory elements represented by colored,

909 labelled lines based on activity level from alkaline phosphatase assay (C) Alkaline phosphatase

910 reporter assay to screen for regulatory activity. E5 chick retinae were electroporated with

911 CRE::AP plasmids and CAG::mCherry plasmids. Empty Stagia3 vector (No Enhancer) represents

912 the negative control. Scale bar in the first panel represents 500 µm and applies to all.

913

914 **Figure 2 - Candidate enhancers ECR9 and ECR65 demonstrate specificity to ThrbCRM1(+)**
915 **population and overlap with mitotic progenitors.** (A) Overlap between CRE::GFP expression
916 (cyan) and ThrbCRM1::AU1 expression (magenta) 18-22 hours after electroporation. DAPI is
917 shown in the last column. Electroporated retina is found above the dotted line (B) Overlap of 1
918 hous EdU-pulsed cells with ECR9::GFP and ECR65::GFP (cyan) 18-22 hours after electroporation.
919 (C) Quantification of EdU(+) and ThrbCRM1 reporter(+) cells in ECR9::GFP and ECR65::GFP cells.
920 Percentages of EdU(+) cells were calculated from confocal images by determining the number
921 of EdU and enhancer::GFP double-positive cells out of all enhancer::GFP-positive cells. Each
922 point represents a biological replicate with data collected from two images. Percentages of
923 ThrbCRM1(+) cells were calculated from a flow cytometry assay in which retinae were
924 electroporated with ECR9::GFP or ECR65::GFP and ThrbCRM1::TdT and the number of GFP/ TdT
925 double-positive cells of the total GFP(+) population was calculated. Each data point represents a
926 biological replicate. Error bars represent 95% confidence interval.
927
928 **Figure 3. ECR65(+) and ECR9(+) cells lineage trace to similar cell fates as ThrbCRM1(+)**
929 **population.** (A) E5 chick retinas were electroporated with enhancer::PhiC31 constructs and
930 CAG::Bgal as an electroporation control before two days of tissue culture followed by harvest
931 and immunohistochemistry. (A) Retinal sections were stained with GFP, Bgal, and DAPI. (B)
932 Markers of early retinal cell types within lineage traced populations. Sections were stained with
933 the markers Isl1, Visinin, Lim1, and pan-Brn3. Percentages of cells marked by these factors were
934 calculated out of the total number of electroporated Bgal(+) cells per retinal section. Each point
935 represents a biological replicate. Error bars represent 95% confidence interval.

936

937 **Figure 4. Deletions and mutations reveal sites important for regulatory activity** E5 chick
938 retinae were electroporated with full length (A) ECR65 or (B) ECR9 driving TdT along truncated
939 or mutated versions of the enhancers and CAG::IRFP as a co-electroporation control. Retinae
940 were cultured for 18-22 hours before dissociation and analysis by flow cytometry. Labelled
941 blocks in (A) and (B) represent the enhancer constructs labelled along the Y-axis. Grey blocks in
942 (C) and (D) denote putative TF binding sites. Deleted versions of enhancers were oriented in the
943 expression vector such that the truncated end is farther from the TATA box. Percent of
944 enhancer activity was calculated as a ratio between the total GFP(+) cells and the total TdT(+)
945 cells and scaled to the activity of the full-length enhancer. Error bars represent 95% confidence
946 interval. See Methods section for statistical tests.

947

948 **Figure 5. Effect of overexpression of bHLH factors on regulatory activity of ECR9, ECR65, and**
949 **ThrbCRM1. (A)** E5 retinae were electroporated with Nhlh1, NeuroD1, NeuroD4, NeuroG2 or
950 Atoh7 under the control of the CAG promoter in combination with ECR65, ECR9 or ThrbCRM1
951 driving either TdT or GFP. Percentages of ECR9(+), ECR65(+) and ThrbCRM1(+) cells were
952 calculated out of the total cells marked by co-electroporation control CAG::IRFP. **(B)** Retinae
953 were electroporated with ECR9::GFP or ECR65::GFP in combination with ThrbCRM1::AU1 and a
954 CAG::bHLH plasmid. **(C,D)** Comparison of the effect of bHLH overexpression on mutant and WT
955 versions of ECR9 and ECR65. Percentages were calculated out of total number of CAG::IRFP(+)
956 cells. Percentages of WT enhancer(+) cells in the control condition were scaled to 100% for C, D.
957 Percent of enhancer(+) cells in the empty CAG vector control condition were normalized to

958 100% for A. Error bars represent 95% confidence interval. In all graphs, each datapoint
959 represents a biological replicate. See Methods for statistical tests
960
961 **Figure 6. Onset of ECR9 and ECR65 activity compared to ThrbCRM1** (A,B) E5 retinae were
962 electroporated with ThrbCRM1::TdT and either ThrbCRM1::GFP, ECR65::GFP or ECR9::GFP and
963 cultured for 8 hours before harvest. Immunohistochemistry was used to amplify GFP and TdT
964 signal. Yellow arrow in (B) indicates a cell that is ThrbCRM1::AU1(-), ECR9::GFP(+), Vsx2(+). (C)
965 E5 retinae electroporated with ECR9::GFP and ThrbCRM1::TdT were cultured for 8 hours and
966 pulsed with EdU from hour 7-8. Yellow arrow indicates a cell that is ECR9::GFP(+)
967 ThrbCRM1::AU1(-) EdU(+). White arrow indicates a cell that is ECR9::GFP(+) ThrbCRM1::AU1(-)
968 EdU(-). Scale bars in A, B, and C represent 50 µm and apply to all panels. Scale bars in B and C
969 insets represent 20 µm and apply to all panels.
970
971 **Figure 7. Co-localization of bHLH-induced activity of ECR9 and ECR65 with early retinal**
972 **development markers**. E5 retinas were electroporated with ECR9::GFP or ECR65::GFP and their
973 corresponding bHLH factors and cultured for 18-22 hours before harvest and
974 immunohistochemistry. (A) Cell quantitation derived from confocal images of retinas stained
975 for ECR9::GFP (enhancer), co-electroporation marker Bgal, and either Vsx2 or Otx2. Percentages
976 of Otx2(+) or Vsx2(+) cells as well as GFP(+)Otx2(+) or GFP(+)Vsx2(+) cells were calculated out of
977 the total number of Bgal(+) cells. Error bars represent 95% CI (B) Same as for (A) but for
978 ECR65::GFP. (C) Confocal images of retinas were stained for GFP (enhancer), Onecut1 and Vsx2.
979 Scale bar represents 50 µm.

Figure 1

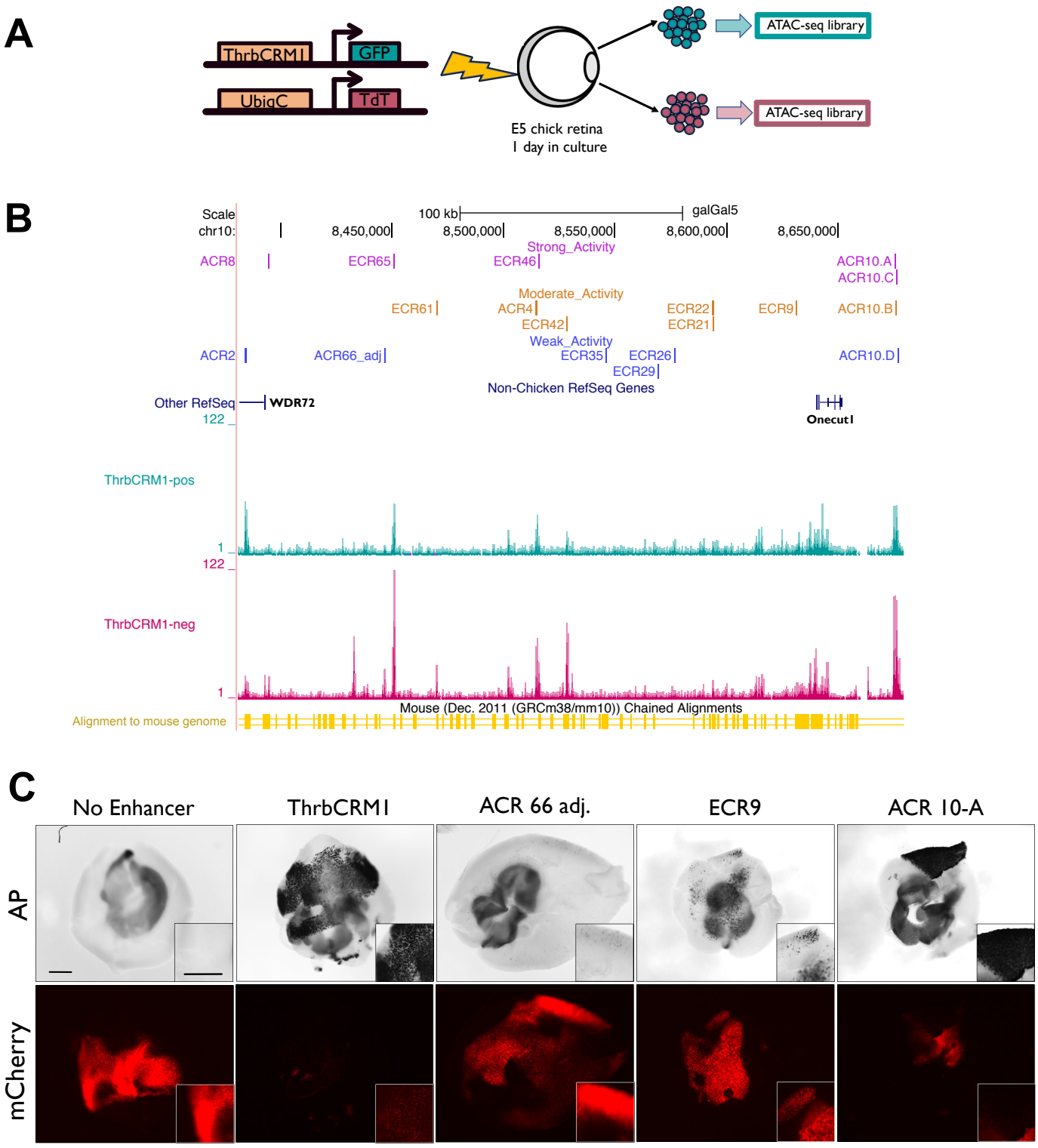

A
ThrbCRM1
GFP
UbiqC
TdT
E5 chick retina
1 day in culture
ATAC-seq library
ATAC-seq library
B
Scale
chr10:
100 kb
galGal5
8,450,000
8,500,000
8,550,000
8,600,000
8,650,000
Strong_Activity
ACR8
ECR65
ECR46
ACR10.A
ACR10.C
Moderate_Activity
ECR61
ACR4
ECR22
ECR9
ACR10.B
ECR42
ECR21
Weak_Activity
ACR2
ACR66_adj
ECR35
ECR26
ACR10.D
ECR29
Non-Chicken RefSeq Genes
Other RefSeq
WDR72
Onecut1
ThrbCRM1-pos
ThrbCRM1-neg
Mouse (Dec. 2011 (GRCm38/mm10)) Chained Alignments
Alignment to mouse genome
C
No Enhancer
ThrbCRM1
ACR 66 adj.
ECR9
ACR 10-A
AP
mCherry

Figure 2

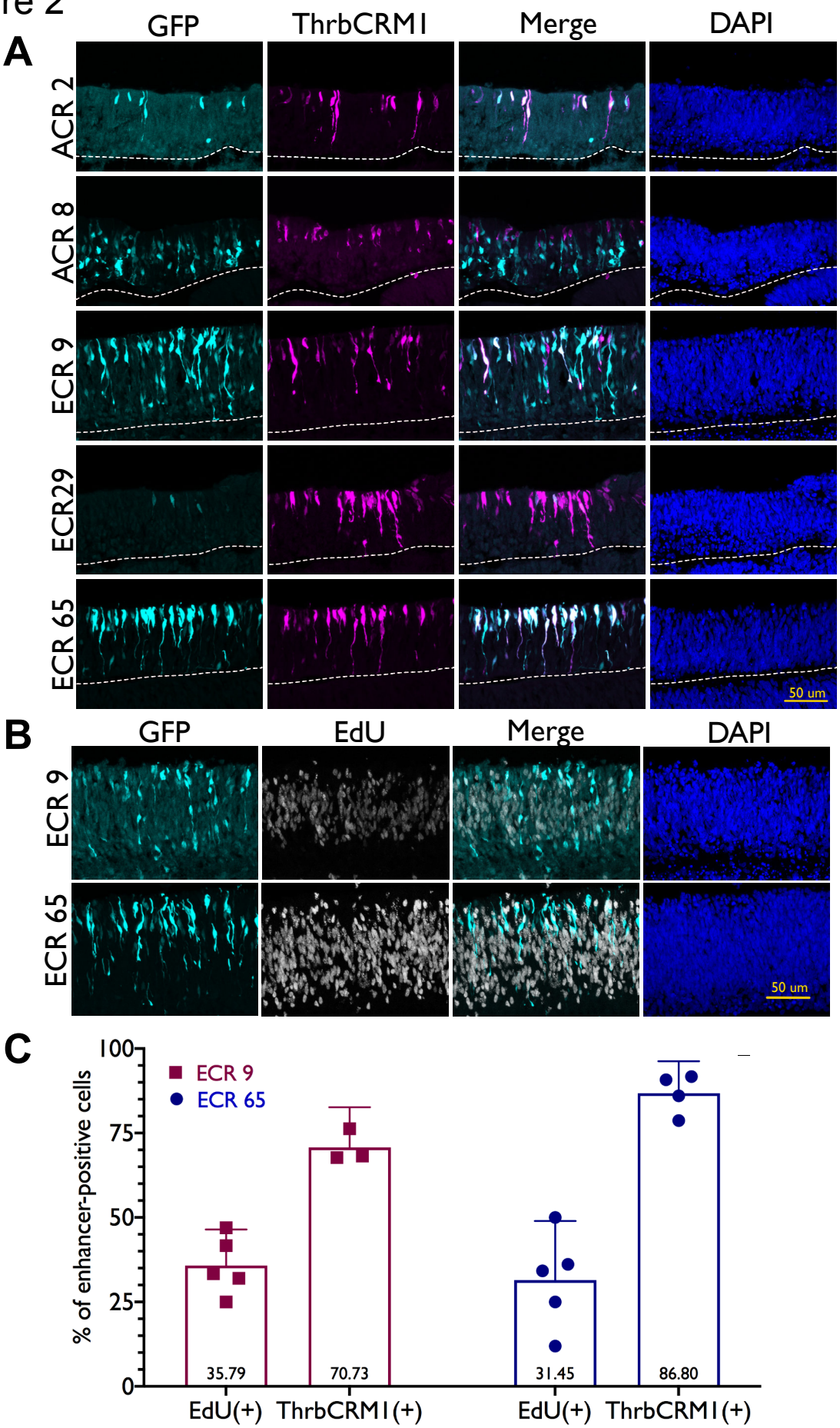

A
GFP
ThrbCRM1
Merge
DAPI
ACR 2
ACR 8
ECR 9
ECR29
ECR 65
50 um
B
GFP
EdU
Merge
DAPI
ECR 9
ECR 65
50 um
C
ECR 9
ECR 65
% of enhancer-positive cells
100
75
50
25
0
35.79
70.73
31.45
86.80
EdU(+)
ThrbCRM1(+)
EdU(+)
ThrbCRM1(+)

Figure 3

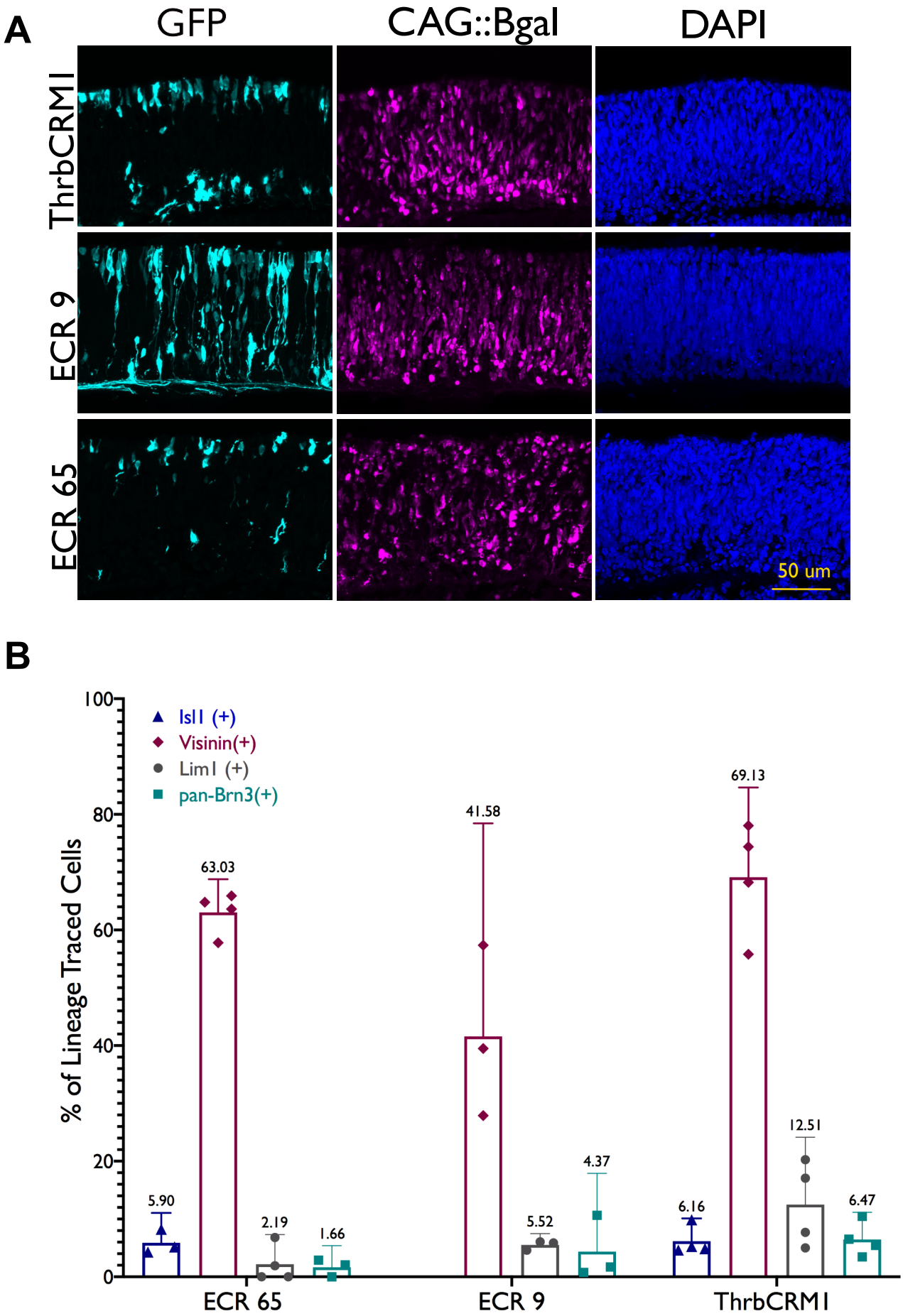

A
GFP
CAG::Bgal
DAPI
ThrbCRM1
ECR 9
ECR 65
50 um
B
Isl1 (+)
Visinin(+)
Lim1 (+)
pan-Brn3(+)
% of Lineage Traced Cells
100
80
60
40
20
0
5.90
63.03
2.19
1.66
41.58
5.52
4.37
6.16
69.13
12.51
6.47
ECR 65
ECR 9
ThrbCRM1

Figure 4

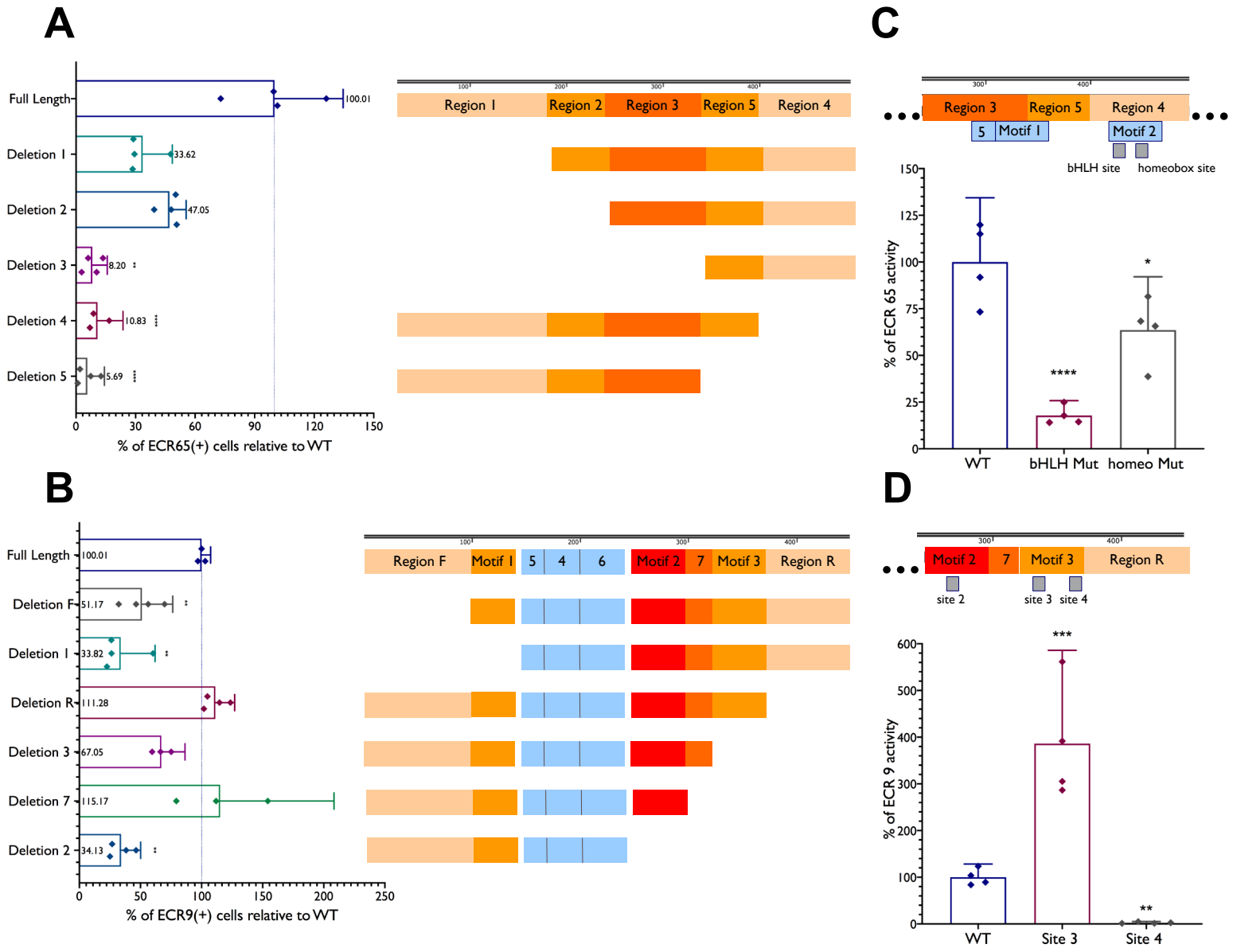
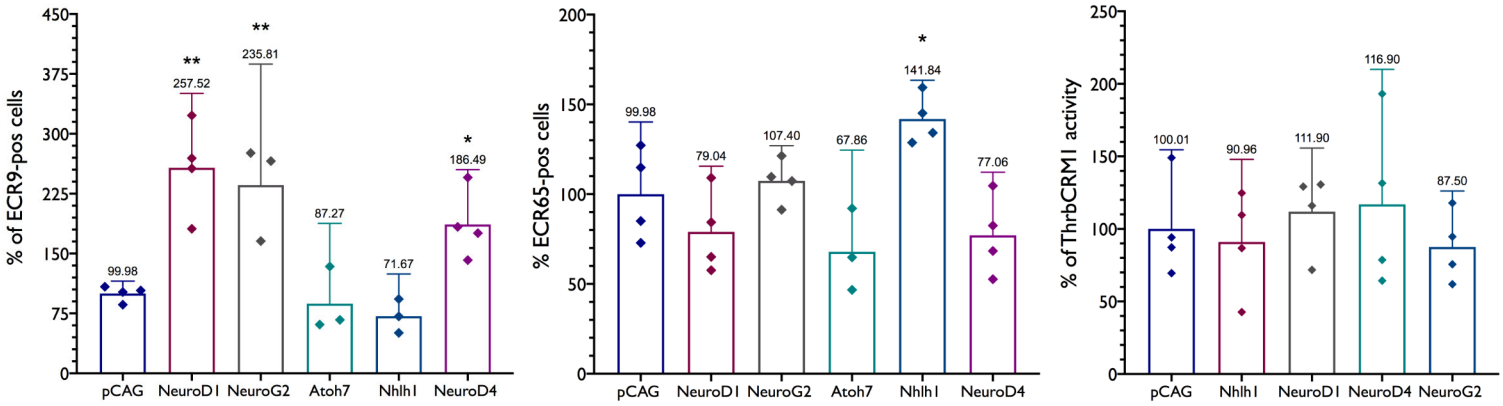
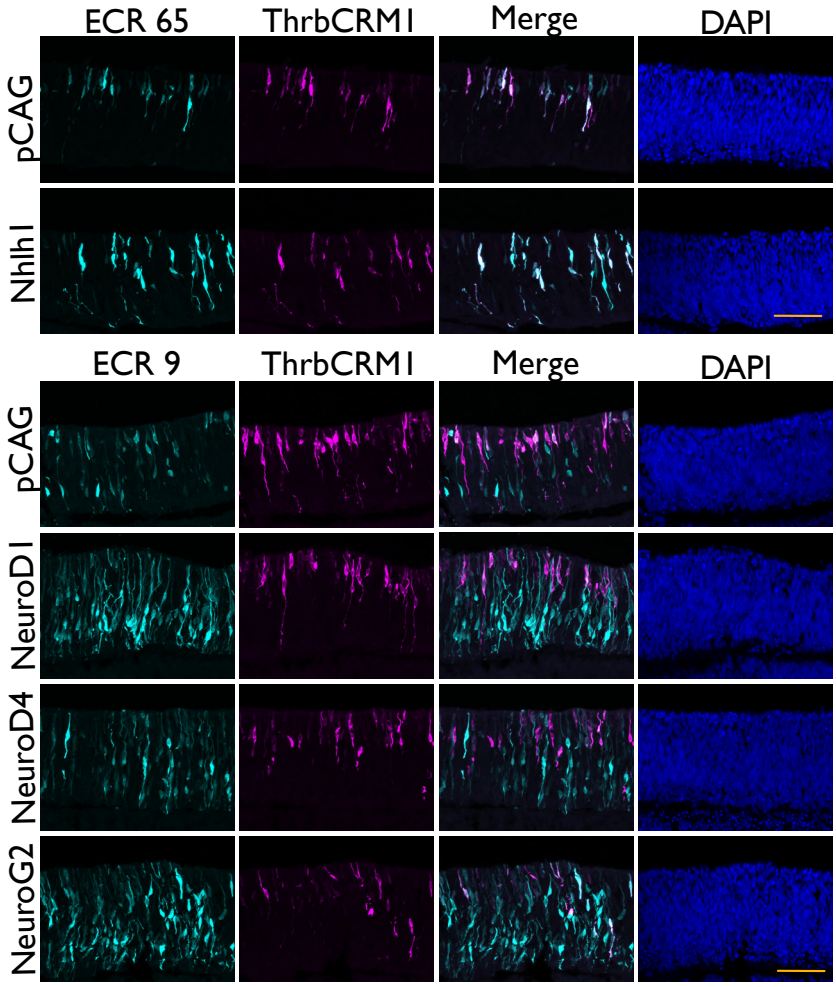

A
Full Length
Deletion 1
Deletion 2
Deletion 3
Deletion 4
Deletion 5
100.01
33.62
47.05
8.20
10.83
5.69
0
30
60
90
120
150
% of ECR65(+) cells relative to WT
Region 1
Region 2
Region 3
Region 5
Region 4
B
Full Length
Deletion F
Deletion 1
Deletion R
Deletion 3
Deletion 7
Deletion 2
100.01
51.17
33.82
111.28
67.05
115.17
34.13
0
50
100
150
200
250
% of ECR9(+) cells relative to WT
Region F
Motif 1
5
4
6
Motif 2
7
Motif 3
Region R
C
Region 3
Region 5
Region 4
5
Motif 1
Motif 2
bHLH site
homeobox site
% of ECR 65 activity
150
125
100
75
50
25
0
****
*
WT
bHLH Mut
homeo Mut
D
Motif 2
7
Motif 3
Region R
site 2
site 3
site 4
% of ECR 9 activity
600
500
400
300
200
100
0
***
**
WT
Site 3
Site 4

Figure 5

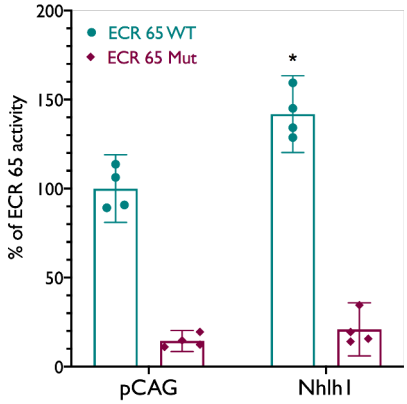
A



B



C



D

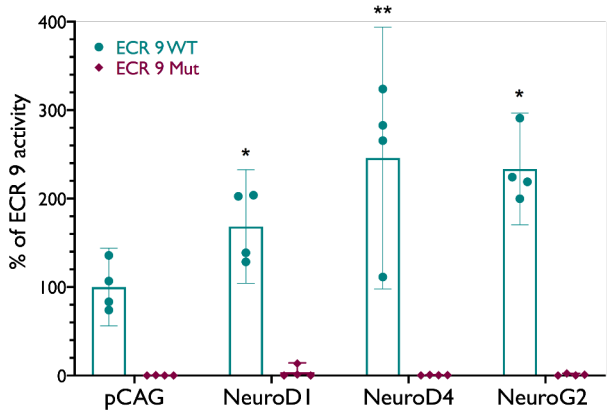


Figure 6

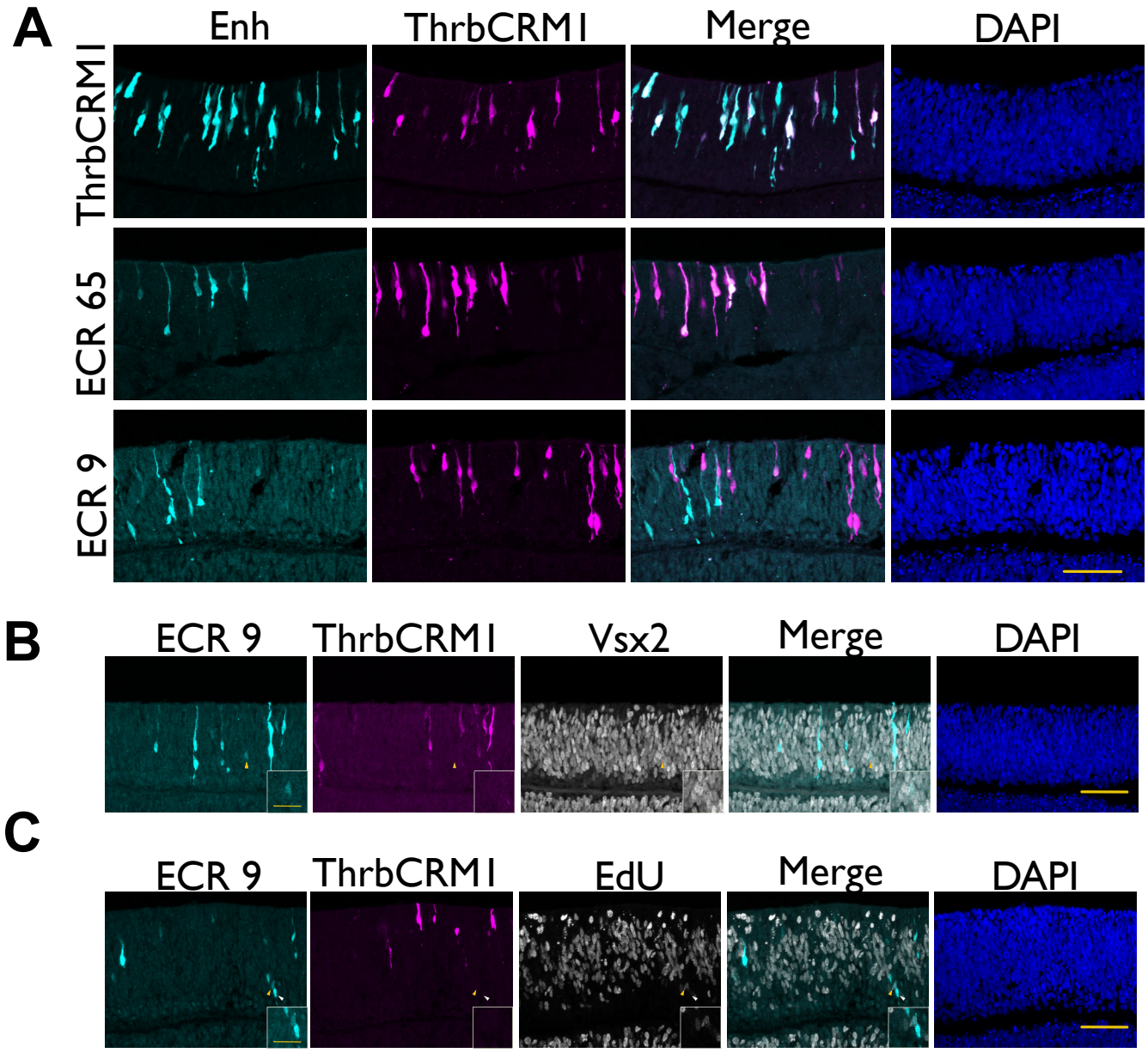
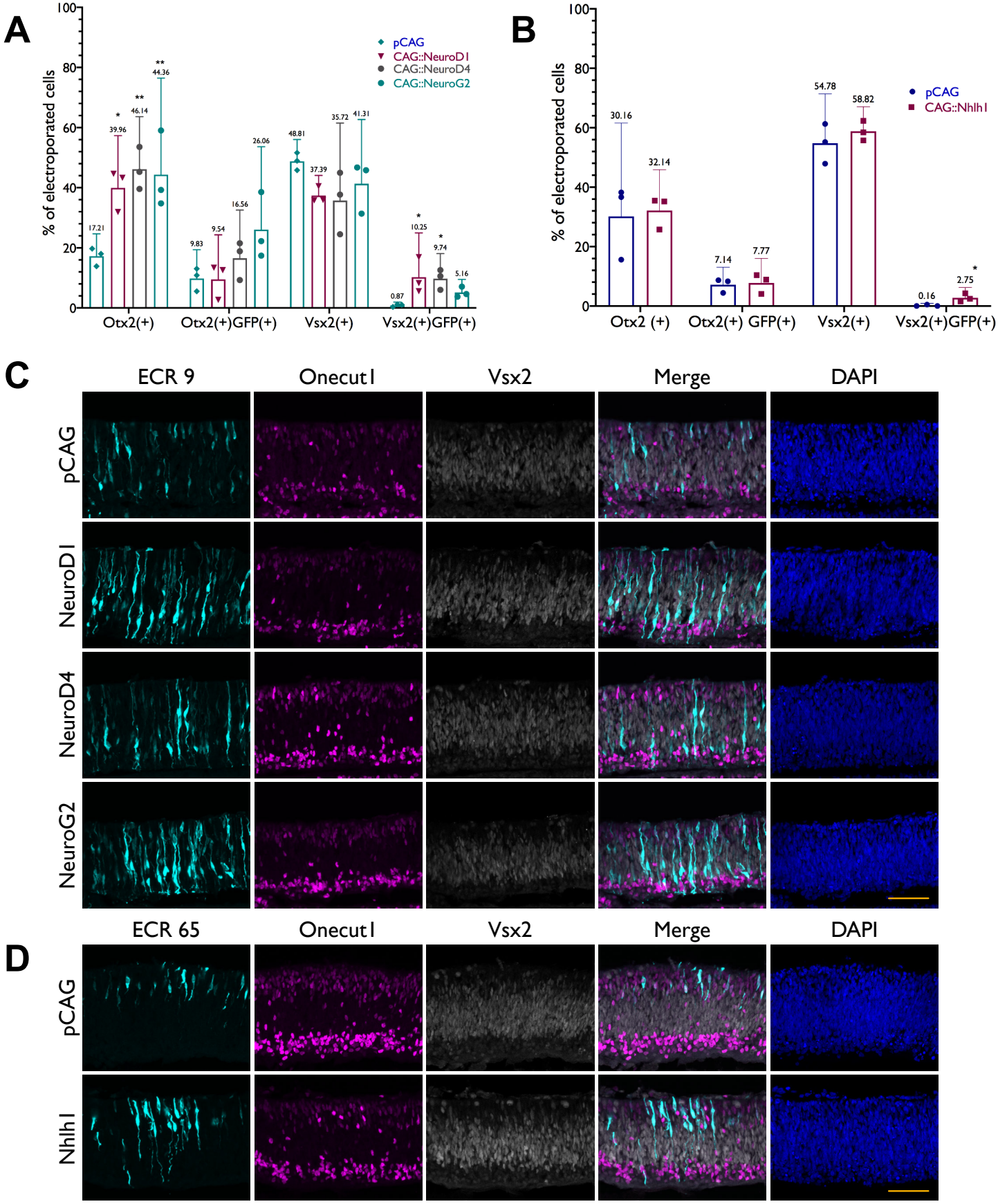

A
Enh
ThrbCRM1
Merge
DAPI
ThrbCRM1
ECR 65
ECR 9
B
ECR 9
ThrbCRM1
Vsx2
Merge
DAPI
C
ECR 9
ThrbCRM1
EdU
Merge
DAPI

Figure 7


A
% of electroporated cells
pCAG
CAG::NeuroD1
CAG::NeuroD4
CAG::NeuroG2
17.21
39.96
46.14
44.36
9.83
9.54
16.56
26.06
48.81
37.39
35.72
41.31
0.87
10.25
9.74
5.16
Otx2(+)
Otx2(+)GFP(+)
Vsx2(+)
Vsx2(+)GFP(+)
B
% of electroporated cells
pCAG
CAG::Nhlh1
30.16
32.14
7.14
7.77
54.78
58.82
0.16
2.75
Otx2 (+)
Otx2(+) GFP(+)
Vsx2(+)
Vsx2(+)GFP(+)
C
ECR 9
Onecut1
Vsx2
Merge
DAPI
pCAG
NeuroD1
NeuroD4
NeuroG2
D
ECR 65
Onecut1
Vsx2
Merge
DAPI
pCAG
Nhlh1