

Rainbow Enhancers Regulate Restrictive Transcription in Teleost Green, Red, and Blue Cones

Wei Fang,¹ Chuanyu Guo,¹ and Xiangyun Wei^{1,2,3}

Departments of ¹Ophthalmology, ²Microbiology and Molecular Genetics, and ³Developmental Biology, University of Pittsburgh, Pittsburgh, Pennsylvania 15213

Photoreceptor-specific transcription of individual genes collectively constitutes the transcriptional profile that orchestrates the structural and functional characteristics of each photoreceptor type. It is challenging, however, to study the transcriptional specificity of individual photoreceptor genes because each gene's distinct spatiotemporal transcription patterns are determined by the unique interactions between a specific set of transcription factors and the gene's own *cis*-regulatory elements (CREs), which remain unknown for most of the genes. For example, it is unknown what CREs underlie the zebrafish *mpp5b^{ponli}* (*ponli*) and *crumbs2b* (*crb2b*) apical polarity genes' restrictive transcription in the red, green, and blue (RGB) cones in the retina, but not in other retinal cell types. Here we show that the intronic enhancers of both the *ponli* and *crb2b* genes are conserved among teleost species and that they share sequence motifs that are critical for RGB cone-specific transcription. Given their similarities in sequences and functions, we name the *ponli* and *crb2b* enhancers collectively *rainbow* enhancers. *Rainbow* enhancers may represent a *cis*-regulatory mechanism to turn on a group of genes that are commonly and restrictively expressed in RGB cones, which largely define the beginning of the color vision pathway.

Key words: crumbs; nagie oko; photoreceptor; ponli; promoter; *rainbow* enhancer

Significance Statement

Dim-light achromatic vision and bright-light color vision are initiated in rod and several types of cone photoreceptors, respectively; these photoreceptors are structurally distinct from each other. In zebrafish, although quite different from rods and UV cones, RGB cones (red, green, and blue cones) are structurally similar and unite into mirror-symmetric pentamers (G-R-B-R-G) by adhesion. This structural commonality and unity suggest that a set of genes is commonly expressed only in RGB cones but not in other cells. Here, we report that the *rainbow* enhancers activate RGB cone-specific transcription of the *ponli* and *crb2b* genes. This study provides a starting point to study how RGB cone-specific transcription defines RGB cones' distinct functions for color vision.

Introduction

Vertebrate retinal rod photoreceptors sense dim light stimuli for achromatic scotopic vision, whereas various types of cone photoreceptors sense bright light stimuli for color photopic vision. These functional differences result from distinct subcellular structures and molecular compositions of rods and cones, which in turn are determined by cell type-specific transcription profiles (Arshavsky et al., 2002; Pujic et al., 2006; Zhang et al., 2006;

Solomon and Lennie, 2007; Jacobs and Nathans, 2009). Clearly, understanding the regulation of photoreceptors' transcriptional specificity is a fundamental aspect of visual science research.

To study transcriptional specificity in photoreceptors, we focus on apical polarity genes in zebrafish. Many polarity genes are expressed in the retina, and they play essential roles in the development and maintenance of zebrafish's rod photoreceptors and four types of cone photoreceptors (UV, red, green, and blue [RGB]) (Tsujikawa and Malicki, 2004; Gosens et al., 2008; Pocha and Knust, 2013). Moreover, in the retina, these polarity genes are expressed in cell type-specific manners. For example, the zebrafish *mpp5b^{ponli}* apical polarity gene (photoreceptor-layer-nok-like, *ponli* for short) and *crumbs2b* (*crb2b*) are not expressed in rods, UV cones, or any other retinal cells; rather, both *ponli* and *crb2b* are restrictively expressed in RGB cones (Zou et al., 2010, 2012, 2013), which align mirror-symmetrically into the G-R-B-R-G configuration (Robinson et al., 1993; Raymond et al., 1995) and are further adhered together as pentameric cone units (see Fig. 1A) (Zou et al., 2012). Conversely, the closest homologs of

Received Nov. 4, 2016; revised Dec. 31, 2016; accepted Jan. 27, 2017.

Author contributions: W.F. and X.W. designed research; W.F., C.G., and X.W. performed research; W.F., C.G., and X.W. analyzed data; W.F. and X.W. wrote the paper.

This work was supported by the National Institutes of Health Grants P30EY008098, EY016099, EY025638, and R21EY023665, Eye and Ear Foundation of Pittsburgh, and Research to Prevent Blindness. We thank Lynne Sunderman for proofreading the manuscript.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Xiangyun Wei, Department of Ophthalmology, University of Pittsburgh, 3501 Fifth Avenue, Pittsburgh, PA 15213. E-mail: weix@upmc.edu.

DOI:10.1523/JNEUROSCI.3421-16.2017

Copyright © 2017 the authors 0270-6474/17/372834-15\$15.00/0

Table 1. The major constructs that were analyzed in the study^a

Construct no. (Figure)	Candidate <i>cis</i> -regulatory DNAs	Corresponding fish lines
1 (Fig. 2A)	4 kb upstream intergenic DNA, exon 1	NA
2, 4 (Figs. 2A, 4A)	6102 bp <i>ponli</i> DNA (130 bp upstream DNA, exon 1, intron 1, and the noncoding region of exon 2)	Tg(<i>ponli</i> ^{6,102} : HA-mCherry) ^{pt118b}
3 (Fig. 2A)	2 kb LCR ^{RH2} -RH2-1 regulatory DNA (Tsujimura et al., 2007)	Tg(RH2-1:mCherry) ^{pt120}
5–10 (Fig. 4B)	Deletion variants of the 6102 bp DNA, lacking various regions of intron 1	NA
11 (Fig. 4C)	1904 bp <i>ponli</i> DNA (130 bp upstream DNA, exon 1, 818 bp intronic DNA, 5' and 3' intron 1 splicing junction regions, and noncoding region of exon 2)	Tg(<i>ponli</i> ^{1,904} : GFP) ^{pt151-#2, #3} Tg(<i>ponli</i> ^{1,904} : GFP) ^{w1465-#1, #11}
12 (Fig. 5A)	<i>nok</i> core promoter region and <i>ponli</i> 818 bp intronic DNA (<i>nok</i> 's 624 bp upstream DNA, <i>nok</i> 's noncoding exon 1, 1097 bp 5' part of <i>nok</i> intron 1, <i>ponli</i> 's 818 bp intronic DNA, 3' splicing junction region of <i>ponli</i> intron 1, and noncoding region of <i>ponli</i> exon 2)	NA
13 (Fig. 5A)	<i>nok</i> core promoter region (including the same DNA as in construct 12 but lacking <i>ponli</i> 818 bp intronic DNA)	NA
14 (Fig. 6D)	212 bp tilapia <i>ponli</i> enhancer region (the same as construct 11, except the 818 bp zebrafish <i>ponli</i> enhancer region was replaced)	Tg(<i>ponli</i> ^{tilapia} : GFP) ^{pt152-#14, #0}
15 (Fig. 9B)	882 bp medaka <i>crb2b</i> enhancer region (the same as construct 11, except the 818 bp zebrafish <i>ponli</i> enhancer region was replaced)	NA

^aThe structure features, corresponding fish lines, and figures showing the schematic drawings of the constructs are tabulated.

ponli and *crb2b*, the *mpp5a*^{naïve oko} (*nok* for short), and *crb2a* genes are expressed in undifferentiated retinal epithelium cells as well as in all types of photoreceptors and Müller glial cells (see Fig. 1A) (Wei et al., 2002, 2006; Zou et al., 2012). Thus, *ponli* and *crb2b* are ideal subjects to study RGB cone-specific transcription.

Like other RNA polymerase II-driven (Pol II) genes (Warner et al., 2008; Yáñez-Cuna et al., 2013), the transcription of *ponli* and *crb2b* should be regulated by their *cis*-regulatory elements (CREs) and trans-regulatory transcription factors. Because the two genes display identical retinal expression patterns, it is possible that they are regulated by similar CREs and transcription factors. However, virtually nothing is known about the mechanisms underlying their distinct transcription patterns.

Cis-regulation of Pol II-driven tissue-specific gene transcription requires both core promoters and enhancers. Core promoters are required for Pol II docking and transcription initiation but not necessarily for the spatiotemporal specificity of transcription; core promoters normally localize within 100 bp around transcription start sites (Lenhard et al., 2012). By contrast, enhancers regulate the spatiotemporal specificity of transcription, and enhancers can localize fairly distally from transcription start sites, either upstream or downstream (Kulaeva et al., 2012; Rouault et al., 2014). Thus, the unique RGB cone-specific transcription of *ponli* and *crb2b* may be regulated by novel and similar enhancers in conjunction with core promoters.

The purpose of this study is to address two questions: First, what are the CREs of the teleost *ponli* and *crb2b* genes that are critical for RGB cone-specific transcription in the retina? Second, do *ponli* and *crb2b* enhancers bear any similarities that may imply a common *cis*-regulatory mechanism? By assessing the transcriptional activities of various regions of the *ponli* and *crb2b* genes and effects of mutations with transgenic approaches, here we provide affirmative answers to both questions.

Materials and Methods

Zebrafish and Medaka care. Tubingen wild-type zebrafish, Tg(SWS1:GFP) (Takeuchi et al., 2003), Tg(LCR^{RH2}-RH2-2:GFP)^{pt115k} and Tg(RH2-1:GFP)^{pt112} fish lines (Fang et al., 2013), as well as medaka fish (obtained from a local pet store) were maintained on a 14 h light/10 h dark cycle. Zebrafish embryos were raised at 28.5°C. Melanin pigmentation was blocked with 0.003% phenylthiourea. In this study, both male and female fish were used for experiments. For larval fish, their sex was undetermined; for founder fish, they were either males or females; and for F1 stable fish, one female and one male were analyzed. Animal care and handling were in accordance with University of Pittsburgh guidelines.

Analysis of transcriptional activities of candidate CREs by transgenesis and mutagenesis. The RH2-1, *ponli*, and *crb2b* *cis*-regulatory genomic DNAs were obtained from the following sources: The 2 kb RH2-1 green opsin gene regulatory DNA (LCR^{RH2}-RH2-1) includes a local control region (LCR), RH2-1 promoter, and 5'-UTR region of zebrafish RH2-1 gene (a gift from Dr. Shoji Kawamura) (Tsujimura et al., 2007); various regions of zebrafish *ponli* gene, including the 818 bp *ponli* enhancer (KY379333), were amplified from BAC clone CH211-105D18 (BACPAC Resources Center); the 212 bp tilapia *ponli* enhancer (KY379335) was amplified from tilapia genomic DNA, which was isolated from a frozen tilapia fish sold in a local supermarket; and the 208 bp medaka *ponli* (KY379334) and 882 bp *crb2b* (KY399459) enhancers were amplified from medaka genomic DNA, which was isolated from a whole medaka fish.

To test the transcriptional activities of various *cis*-regulatory DNAs, we used the I-Sce I meganuclease-based transgenesis system (Thermes et al., 2002) (vector provided by Dr. Michael Tsang). Specifically, each *cis*-regulatory DNA was inserted between a KpnI and an NheI restriction site and was followed sequentially by a Kozak sequence (GCCACC), the GFP or HA-tagged mCherry reporter ORFs, and an *sv40* polyadenylation sequence-containing 3'-UTR (see Fig. 2A, constructs 1–3; Fig. 4A–C, constructs 4–11; Fig. 5A, constructs 12 and 13; Fig. 6D, construct 14; Fig. 9B, construct 15; Table 1). All constructs were verified by restriction digestion and sequencing. After the constructs were made, 20 pg of each construct was coinjected with 0.01 U of I-Sce I enzyme into embryos at 1-cell stage. Reporter expression was examined at larval or adult stages either in founder fish or in F1 stable transgenic fish.

For the deletion analysis of the zebrafish *ponli* intron 1, various internal regions were deleted from construct 4 using the PCR-based Q5 Site-Directed Mutagenesis Kit (NEB, #E0554S) (see Fig. 4B, constructs 5–10), with inverse PCR primers matching to the sequences that flank the deletion regions. Construct 11 was generated by inserting the 531 bp core promoter region, the 818 bp *ponli* enhancer, and the 3' end of intron 1, noncoding region of exon 1 between the KpnI and NheI sites.

Constructs 14 and 15 were generated by replacing the 818 bp zebrafish *ponli* enhancer between the ApaI and AsiI sites of construct 11 with the 212 bp tilapia *ponli* enhancer and the 882 bp medaka *crb2b* enhancer, respectively.

For substitution mutation analysis of enhancer motifs, the candidate enhancer motifs in constructs 11 and 14 were substituted with unrelated sequences of the same length (8–18 bp; see Fig. 8), with the center sequence as an AscI site (GGCGCGCC) for construct selection by restriction digestion. Again, these sequence substitutions were performed with the Q5 Site-Directed Mutagenesis Kit.

Real time RT-PCR analysis. To compare the mRNA levels of the endogenous *ponli* and RH2 green opsin genes, we performed real time RT-PCR analysis with the 2^{−(ΔΔCt)} method (Livak and Schmittgen, 2001), and we used an iQ5 real-time PCR detection system and an iQ SYBR Green Super-

mix kit (Bio-Rad). The experiments were performed with triplicate samples and repeated three times and normalized against the mRNA of the *elf2a* housekeeping gene. The following PCR primer pairs were used: *ponli* forward (ACATCGCTGGTTCTACACC) and *ponli* reverse (CCAACACGG TACACGTTTCAG); *RH2-1* forward (GCCGCTCAACTACATTTTGG) and *RH2-1* reverse (TGTAGCCCAAGTAGAGCAG); *Elf2a* forward (TT GAGAAGAAAATCGGTGGTGCTG) and *Elf2a* reverse (GGAACGGTGT GATTGAGG GAAATTC).

Bioinformatics. Candidate CREs of the *ponli* and *crb2b* genes were identified by searching homologous sequences among teleost fish genomes with the BLAT searching utility of the UCSC database (<http://genome.ucsc.edu/>). At the medaka (*Oryzias latipes*) genome browser, enter *ponli* or *crb2b* to locate the genes. The search also automatically align homologous genes of the following species: zebrafish (*Danio rerio*), stickleback (*Gasterosteus aculeatus*), fugu (*Takifugu rubripes*), tetraodon (*Tetraodon nigroviridis*), and Nile tilapia (*Oreochromis niloticus*). The sequence similarities were calculated with Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/>, Research resource identifier [RRID]: SCR_001591).

Confirming the transcription and splicing of the *ponli*⁶⁰⁹²:HA-mCherry gene by RT-PCR. We performed RT-PCR analysis to confirm that the reporter gene was properly transcribed and spliced in *Tg(ponli^{6,092}:HA-mCherry)^{pt118b}* (pt118b for short). Briefly, we isolated the total RNA from 5 dpf eyes of wild-type and pt118b fish with Trizol (Invitrogen) and then performed reverse transcription using SuperScript III kit (Invitrogen). We then PCR amplified the cDNAs with the following primer pairs to confirm transcription and splicing (see Fig. 2A,B): F1 (GGACCTGATTGGAAACATCG, in exon 1) and R1 (GGCAAAAAGCCAACACTGTACA, in exon 2), which can detect both the endogenous *ponli* gene and the reporter gene but cannot distinguish between the two; F1 and R2 (GGATGTGCGCGGGGT GCTTCACGTA, in the mCherry coding regions), which can only detect the reporter gene (see Fig. 2A,B).

Comparison of reporter gene expression by Western blotting. We performed Western blotting to compare the protein expression levels of reporter genes. Briefly, fish eyes were removed and homogenized on ice in a lysis buffer (1 × PBS buffer with 0.5% Triton X-100, 1 mM PMSF, 1 mM dithiothreitol, and 1 × Roach proteinase inhibitor mixture). The lysates were mixed with 5 × SDS-PAGE loading buffer and heated at 100°C for 5 min. The supernatants of the protein samples were separated in 12% SDS-PAGE gels and transferred to PVDF membranes for Western blotting with anti-mCherry antibody (1:2000; BioVision, catalog #5993-100; RRID:AB_1975001), anti-GFP antibody (1:2000; Sigma-Aldrich, catalog #G6539; RRID:AB_25 9941), and anti-γ-tubulin antibody (1:3000; Sigma-Aldrich, catalog #T3320; RRID:AB_261655).

Immunohistochemical analysis of reporter expression patterns in the retina. We performed immunohistochemical confocal microscopy to examine the cellular expression patterns of reporter proteins in the retina. Briefly, both adult fish and larval eyes were fixed in 4% PFA in 1 × PBS buffer at room temperature for 2 h, infiltrated with 40% sucrose in 1 × PBS overnight, and embedded in tissue freezing media (Tissue-Tek; Sakura Finetek). Eye samples were cryosectioned at 30 μm thickness and immunostained as described previously (Wei et al., 2006). We used the following primary and secondary antibodies: rabbit polyclonal mCherry antibodies (1:300; BioVision, catalog #5993-100; RRID:AB_1975001), mouse monoclonal Zpr1 antibodies (1:300; Zebrafish International Resource Center, catalog #ab174435; RRID: AB_10013803), Cy3-conjugated goat anti-mouse IgG (1:300; Jackson ImmunoResearch Laboratories, catalog #115-165-166; RRID:AB_23 38692), goat anti-rabbit IgG (1:300; Jackson ImmunoResearch Laboratories, catalog #111-165-144; RRID:AB_2338006), and Cy5-conjugated donkey anti-rabbit IgG (1:300; Jackson ImmunoResearch Laboratories, catalog #711-175-152; RRID:AB_2340607). AlexaFluor-647 phalloidin (1:300; Thermo Fisher Scientific, catalog #A22287; RRID:AB_2620155) was used for F-actin staining. Images were taken with a Fluoview FV1000 confocal microscope (Olympus).

Results

We took a strategy to first identify *ponli* CREs, particularly its enhancer; we then used the characteristics of *ponli* CREs to guide

our search for the enhancer of the *crb2b* gene. Because enhancers can localize distally, either upstream or downstream of transcription sites, and because the transcribed sequence of the *ponli* gene itself already extends over 75 kb, it was a challenging task to identify the *ponli*'s enhancer region in this broad region and beyond. Therefore, for the sake of practicality and because many enhancers often reside upstream of the transcription start site or within the first intron, we limited our search for *ponli* enhancer in the 4 kb upstream intergenic region and in the 5.7 kb downstream intron 1 (Fig. 1B).

The 4 kb upstream intergenic DNA cannot drive RGB cone-specific transcription

To assess the transcriptional activity of the 4 kb upstream intergenic DNA, we made transgenic construct 1, which contains the 4 kb intergenic DNA, *ponli*'s 110 bp exon 1, a Kozak sequence, and an mCherry-coding sequence (Fig. 2A, construct 1). We included *ponli* exon 1 in the construct to best ensure the intactness of the core promoter of the *ponli* gene. The construct was then injected into embryos at 1-cell stage and monitored for retinal expression at 4 dpf (days post fertilization); we found that none of the 316 injected fish showed mCherry signals in the retina (data not shown). Thus, the 4 kb upstream intergenic DNA does not contain *ponli*'s enhancer and the core promoter region alone is not sufficient to drive retinal transcription.

ponli's first intron for RGB cone-specific transcription

We next assessed the transcriptional activity of intron 1. To do so, we made transgenic construct 2 by fusing an mCherry reporter gene downstream of a 6102 bp DNA, which covers a 140 bp upstream region, exon 1, intron 1, and the 67 bp noncoding region of exon 2 (Fig. 2A, construct 2). We injected this construct into zebrafish embryos and found that 19 of 208 injected embryos indeed expressed mCherry in the retina at 4 dpf. These mCherry-expressing fish were raised to adulthood, and from their outcross progeny, we identified a stable fish line, which was named *Tg(ponli^{6,102}:HA-mCherry)^{pt118b}* (pt118b for short). Confocal immunohistochemistry of adult pt118b retina revealed that mCherry is restrictively expressed in RGB cones just as endogenous Ponli protein (Fig. 2C–E) (Zou et al., 2010). In addition, RT-PCR results indicate that the intron 1 DNA of the transgene could be transcribed and subsequently removed from mature mRNA by splicing as the endogenous *ponli* transcript is processed, although the RT-PCR analysis could not evaluate the efficiency of splicing (Fig. 2B). Interestingly, the mCherry signals in pt118b are ~10 times weaker than those driven by the *RH2-1green opsin cis-regulatory DNA* in *Tg(Rh2-1:HA-mCherry)^{pt120}* (Fig. 2A, construct 3; Figs. 2F, 3A–C); this weaker transgenic expression echoes the fact that *ponli* endogenous mRNA is much less abundant than green *opsin* mRNA (Fig. 3D).

These results indicate that the 6102 bp DNA harbors sufficient CREs for *ponli*'s RGB cone-specific transcription. Considering that construct 2 shares *ponli*'s core promoter region with construct 1, which lacks any transcriptional activity (Fig. 2A), we conclude that the RGB cone-specific transcriptional activity of the 6102 bp DNA must be determined by an enhancer located in *ponli*'s 5785 bp intron 1.

An 818 bp mid-region of *ponli* intron 1 mediates RGB cone-specific transcription

We next sought to narrow down the intronic DNA region that drives RGB cone-specific transcription. To do so, we generated deletion constructs 4–11 and analyzed the transcriptional activ-

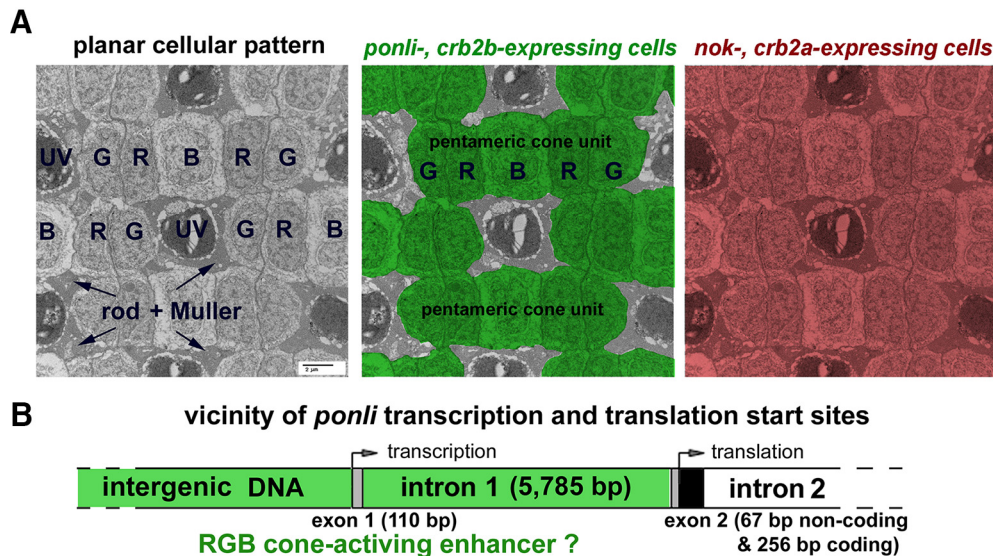


Figure 1. The *ponli*, *crb2b*, *nok*, and *crb2a* apicobasal polarity genes are differentially expressed in the zebrafish retina. **A**, The regular photoreceptor mosaic patterns in adult zebrafish retina are revealed by a transmission electron micrograph of a transverse section of the photoreceptor layer crossing the inner segments of RGB cones (left). RGB cones adhere to each other and align mirror-symmetrically into pentameric units: G-R-B-R-G (Zou et al., 2012). Arrows indicate the electron-dense clusters of Müller cell apical processes and rod inner segments. RGB cones that specifically express *ponli* and *crb2b* are illustrated in green (middle) (Zou et al., 2010, 2012); by contrast, all photoreceptors and Müller cells that express *nok* and *crb2a* are highlighted in red (right) (Wei et al., 2006; Zou et al., 2012). UV, UV cones; G, green cones; R, red cones; B, blue cones. **B**, Search of *ponli*'s enhancer was focused in the intergenic DNA upstream of the transcription start site and in intron 1 (both highlighted in green). Arrows indicate the transcription start site in exon 1 and the translation start site in exon 2. Gray rectangles represent the untranslated exonic regions. Black rectangle represents the protein-coding region of exon 2.

ities of the intronic DNA sequences (Fig. 4A–C). These constructs all contain *ponli*'s core promoter region (including the 140 bp upstream sequence and 110 bp exon 1), but each includes different subregions of intron 1 (Fig. 4A–C). We injected these constructs into embryos at 1-cell stage and monitored their retinal GFP expression at 4 dpf. We found that, in conjunction with the core promoter region, an 818 bp mid-intronic DNA, but not its flanking 4232 bp intronic DNAs, is sufficient to drive retinal expression (Fig. 4B,C). We believe that this 818 bp intronic DNA regulates retinal expression by promoting transcription but not by ensuring the proper splicing between exon 1 and exon 2 through a splicing branch site. This is because, when exon 1 and exon 2 were fused together by deleting the entire intron 1, retinal expression was inhibited (Fig. 4B, construct 9).

To confirm that the 818 bp DNA drives GFP expression specifically in RGB cones, we next raised two stable fish lines, *Tg(ponli^{1,904}:GFP)^{pt151-#2}* (pt151-#2 and pt151-#3 for short; Fig. 4C–G), to reveal the cellular expression patterns of GFP in developed retina. Indeed, in *Tg(ponli^{1,904}:GFP)^{pt151-#3}*, GFP is restrictively expressed in RGB cones (Fig. 4D–F); however, in *Tg(ponli^{1,904}:GFP)^{pt151-#2}*, GFP is expressed not only in RGB cones but also in <20% of rods and very sparsely in bipolar cells (Fig. 4G). We also used the same 1904 bp *ponli* cis-regulatory DNA to drive RGB cone-specific gene expression in 13 other unrelated stable transgenic fish lines; nine of these lines display RGB cone-specific expression, but the remaining four fish lines also display some ectopic GFP expression in 10%–80% of rods and sparsely in bipolar cells, as in *Tg(ponli^{1,904}:GFP)^{pt151-#2}* (data not shown). We believe that such ectopic expression in rods and bipolar cells is due to position effects because the ectopic expression does not always happen in every stable fish line. Nevertheless, when ectopic expression does occur, it occurs more often in rods than in other non-RGB-cone cells, suggesting that some rod-specific transcription factors may activate the 1904 bp *ponli* cis-regulatory DNA under certain position effects.

In addition to the high reproducibility of RGB cone-specific transgenic expression, the *ponli^{1,904}* cis-regulatory DNA also characteristically express GFP at modest levels: Western blotting showed that the protein expression levels in *Tg(ponli^{1,904}:GFP)^{pt151-#2}*, #3; w1465-#1, w1465-#11 were 50–100 times lower than those driven by *opsin* regulatory DNAs in *Tg(LCR^{RH2}-RH2-2:GFP)^{pt115-k}* and *Tg(SWS1:GFP)* (Takechi et al., 2003; Fang et al., 2013) (Fig. 4H). We also observed modest variations in GFP expression among *Tg(ponli^{1,904}:GFP)* fish (Fig. 4H).

The 818 bp intronic DNA contains the enhancer of the *ponli* gene

The high reproducibility of RGB cone-specific transcription in these transgenic fish suggests that 1904 bp DNA contains sufficient basic CREs of the *ponli* gene for its RGB cone-specific transcription: namely, *ponli*'s core promoter and the enhancer (or at least one of its enhancers). We suspected that the 818 bp intronic DNA harbored the *ponli* enhancer, which directs *ponli*'s transcriptional specificity; in addition, the 140 bp upstream sequence and exon 1 contain *ponli*'s core promoter, which may be replaced with other generic core promoters without compromising *ponli*'s transcriptional specificity.

To test this idea, we next engineered a chimeric cis-regulatory DNA by combining the 818 bp *ponli* intronic DNA with the core promoter region of the *nok* gene, which is expressed in all photoreceptors and Müller cells in developed retina (Wei et al., 2006) (Fig. 1A), and then evaluated its transcriptional activity with a transgenic reporter construct (Fig. 5A, construct 12). Embryonic injection of this construct led to preferred retinal GFP expression at 4 dpf (Fig. 5B,C). By contrast, injection of another construct that only contained *nok*'s core promoter regions produced no retinal expression, even though many muscle cells displayed transient GFP expression (Fig. 5A–C, construct 13).

To evaluate the transcriptional specificity of the GFP reporter in the construct-12-injected fish, we raised these fish to adulthood and

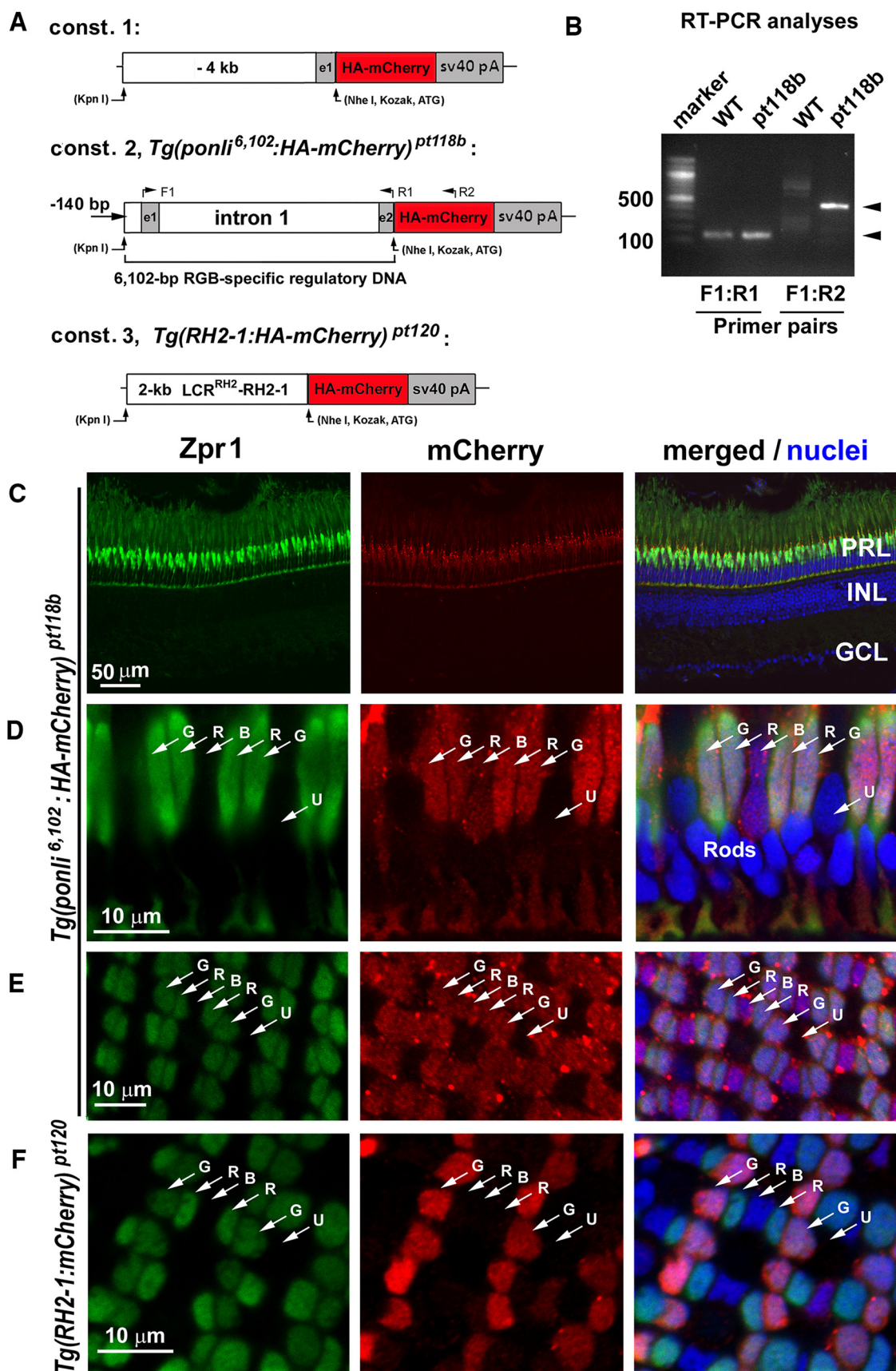


Figure 2. A *ponli* enhancer resides in its first intron. **A**, Transgenic constructs 1–3 were used to assess the transcriptional activities of the 4 kb upstream intergenic DNA (construct 1), *ponli*'s core promoter region and intron 1 (construct 2, to generate *Tg(ponli^{6,102}:HA-mCherry)*^{pt118b}), and green opsin's cis-regulatory DNA LCR^{RH2}-RH2-1 (the locus control region of RH2 genes, followed by RH2-1's proximal core promoter region, including the 5'UTR of the RH2-1 gene; Tsujimura et al., 2007) as a green cone-expressing control (construct 3, to generate *Tg(RH2-1:mCherry)*^{pt120}). (Figure legend continues.)

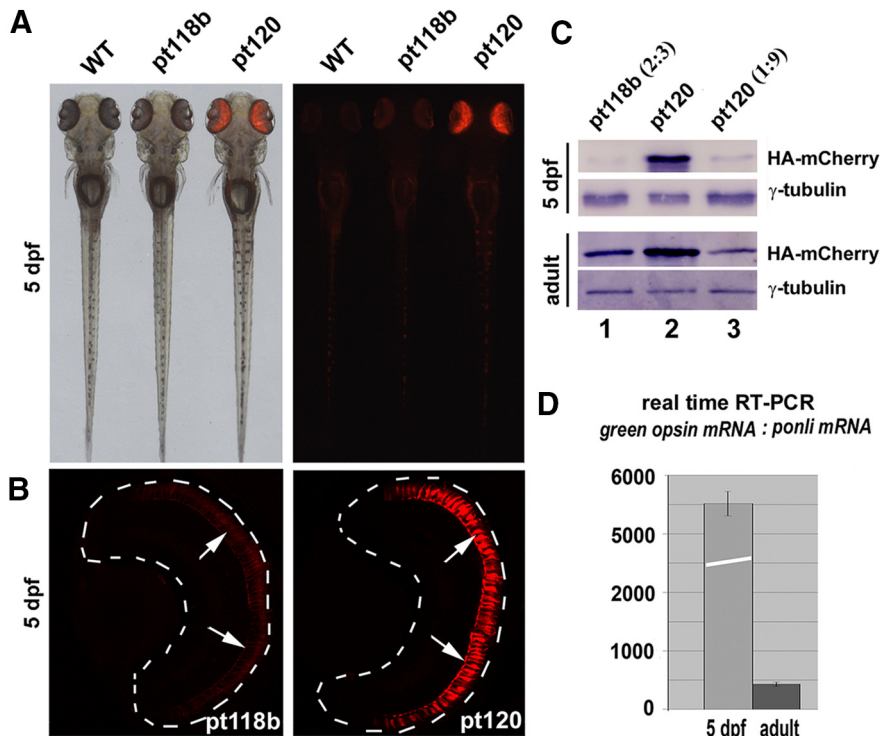


Figure 3. The 6102 bp *ponli* cis-regulatory DNA drives modest transgenic expression. **A**, As revealed by stereo-fluorescence microscopy, the retinal mCherry signals in 5 dpf *Tg(ponli^{6,102}:HA-mCherry)^{pt118}* (pt118b for short) were much weaker than that in *Tg(RH2-1:HA-mCherry)^{pt120}* (pt120 for short). **B**, At 5 dpf, the weak mCherry signals in pt118b and strong mCherry signals in pt120 all localized to the photoreceptor layer (arrows), as shown by confocal immunohistochemistry. White dashed lines encircle retinal profiles. **C**, mCherry protein levels in pt118b were ~1 order of magnitude less than in pt120, as suggested by Western blotting of ten 5 dpf eyes or one-tenth of adult eyes. To ensure that the same amount of mCherry-expressing cells was compared between pt118b and pt120 (lanes 1 and 2), the pt118b samples were diluted with the wild-type retinal lysates by a ratio of 2:3 (lane 1); this dilution was chosen because mCherry is expressed in all five members of pentameric cone units in pt118b but only in two green cone members of cone pentamers in pt120. In lane 3, pt120 lysates were diluted with wild-type lysates at a ratio of 1:9 to facilitate signal comparison. The γ -tubulin blotting served as loading controls. **D**, Real time RT-PCR analysis suggested that the endogenous *ponli* gene is >5000 and 400 times less active than the *green opsin* genes in 5 dpf and adult fish, respectively.

then performed confocal microscopy to examine the expression patterns of transgenic clones. We found that GFP was only expressed in the RGB cones in 49 of the 60 transgenic cell clones examined in adult retina (Fig. 5D,E); in the remaining 11 clones, in addition to GFP expression in RGB cones, we also observed ectopic GFP expression in <20% of rods (Fig. 5F,G). These expression patterns are similar to those observed in *Tg(ponli^{1,904}:GFP)^{pt151-#3}* and *Tg(ponli^{1,904}:GFP)^{pt151-#2}* (Fig. 4E–G), suggesting that the 818 bp intronic DNA underlies the transgenic expression pattern that is characteristic of the *ponli* gene but not that of the *nok* gene. Although we cannot exclude the possibility that there is more than one *ponli* enhancer, our results suggest that the 818 bp intronic DNA contains

a *ponli* enhancer, which when supplied with a basic core promoter (not necessarily *ponli*'s own core promoter), is sufficient to confer modest and RGB cone-specific transcription.

ponli cis-regulatory elements are conserved in teleost fish

To identify the *ponli*'s CREs in its core promoter and enhancer regions, we next aligned the DNA sequences of the 5' ends of *ponli* genes of six teleost fish (medaka, tilapia, fugu, tetraodon, stickleback, and zebrafish). The alignment reveals high sequence conservation near the transcription start site, the 5' splicing site of intron 1, the middle region of intron 1, and exon 2 (Fig. 6A). In the core promoter regions, we found three conserved core promoter motifs (CPM1–3) in all six fish species. CPM1s locate <20 bp upstream of the transcription start sites; CPM2s cover the transcription start sites (Fig. 6B, arrows); and CPM3 locates 30–70 bp downstream of transcription start sites. The sequence of CPM2s (CCACTTG in zebrafish and CCACGGA in four other fish) is similar to the consensus sequence of human transcription initiators (YYANWYY) (Juven-Gershon et al., 2010), suggesting that CPM2s are the initiator motifs of the teleost *ponli* genes (Fig. 6B). We could not identify a TATA box, which normally resides between positions –33 and –28 bp upstream of the transcription start site (Carninci et al., 2006; Cooper et al., 2006). Because core promoters normally reside within 50 bp 5' and 3' of the transcription

start sites (Juven-Gershon et al., 2010; Yáñez-Cuna et al., 2013), these conserved sequence motifs likely constitute the *ponli* core promoters, which belong to the TATA-less core promoters category (Carninci et al., 2006).

The conserved mid-intronic region overlaps with the 3' half of the 818 bp zebrafish enhancer-containing intronic DNA (Fig. 6A). Sequence comparison revealed that these mid intronic regions contain five highly conserved DNA sequence motifs, named, enhancer motifs 1–5 (EM1–5). EM1–5s are aligned sequentially and compactly in all teleost fish, except that in zebrafish, they are aligned in the order of EM1, first EM2, second EM2, EM4, EM3, and EM5. Interestingly, zebrafish EM5 and the second EM2 are in reverse orientation compared with EM2 and EM5 in other teleost fish (Fig. 6A,C). Finally, zebrafish EM5 is more separated from the rest of the EMs, being 50 bp downstream of the nearest EM3; by contrast, in other teleost fish, EM5 is only 3 bp downstream of the nearest EM4.

To determine whether the conserved intronic regions of other teleost fish also drive RGB cone-specific transcription, we replaced the zebrafish 818 bp enhancer region in construct 11 with the 212 bp conserved tilapia intronic DNA and analyzed, in zebrafish, the GFP expression patterns of the resulting chimeric transgenic construct 14 (Fig. 6D). As expected, the 212 bp tilapia intronic DNAs indeed drove GFP expression in zebrafish RGB cones in two stable fish lines *Tg(ponli^{1,904}:GFP)^{pt152-#14, #D}* (Fig. 6E). These results suggest that the mid intronic regions of the

←

(Figure legend continued.) HA-mCherry)^{pt120}. The locations and directions of RT-PCR primers (**B**) for construct 2 are indicated with arrows. e1, Exon 1; e2, exon 2. **B**, The endogenous *ponli* gene's transcription and subsequent splicing between exon 1 and exon 2 was verified by RT-PCR using primer pair F1 and R1 (bottom arrowhead), and mCherry's transcription and subsequent splicing was confirmed by RT-PCR with primer pair F1 and R2 (top arrowhead). **C–E**, In adult *Tg(ponli^{6,102}:HA-mCherry)^{pt118b}* retina, mCherry is expressed in the photoreceptor layer (PRL) but not in the inner nuclear layer (INL) or ganglion cell layer (GCL) (**C**, lower magnification); in the PRL, mCherry is expressed in green (arrows G), red (arrows R), and blue (arrows B) cones but not in UV cones (arrows U) or rods (**D**, vertical sections; **E**, transverse sections). **F**, By contrast, mCherry is almost exclusively expressed in green cones in *Tg(RH2-1:HA-mCherry)^{pt120}*. Green/red double cones were visualized by Zpr1 antibody. The cell nuclei were stained with DAPI (blue).

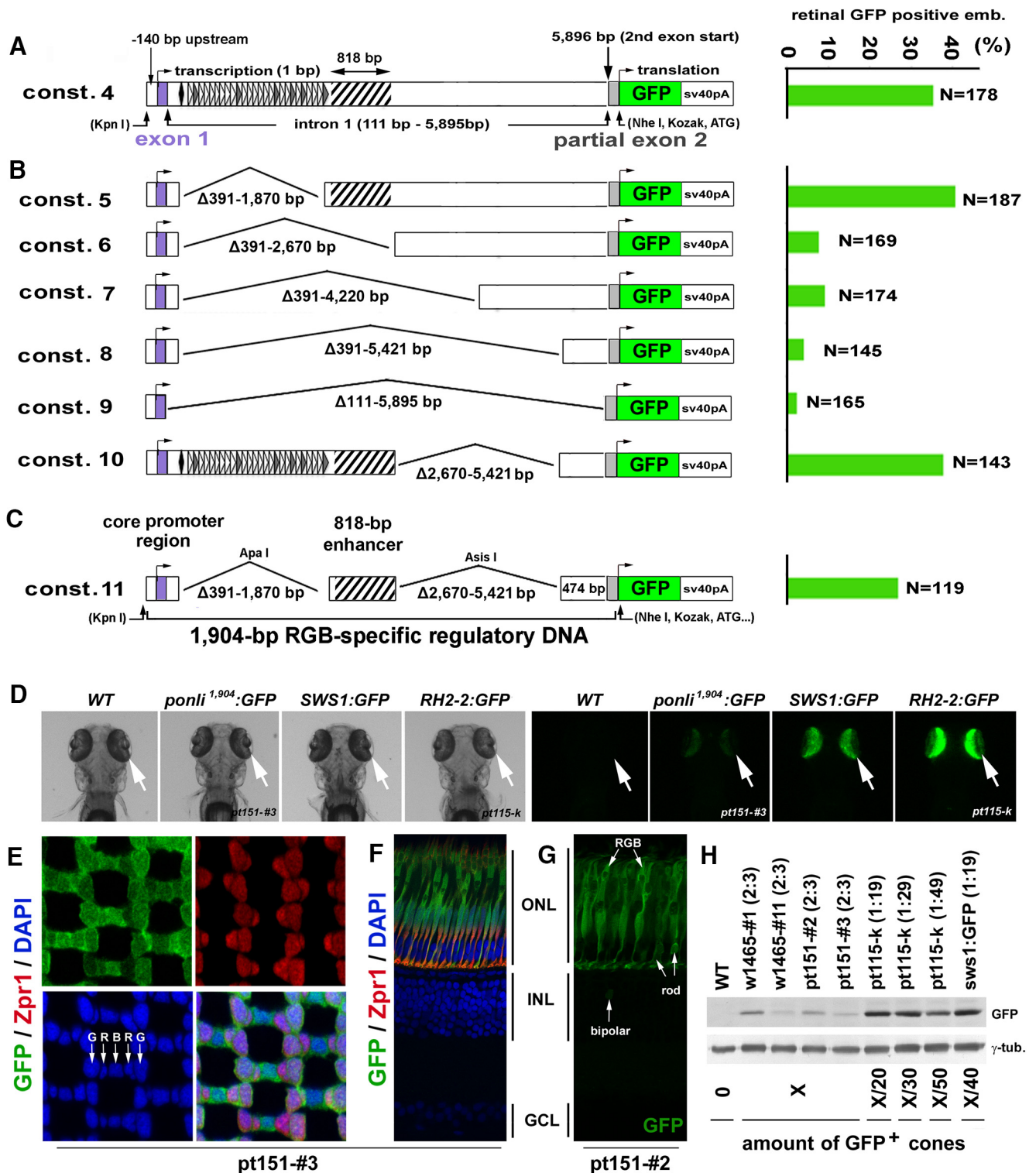


Figure 4. The 818 bp *ponli* intronic DNA is sufficient to mediate RGB cone-specific transcription. **A**, Construct 4 contains a GFP reporter that is driven by *ponli*'s 6102 bp *cis*-regulatory DNA (between –140 bp upstream of the transcription start site and the translation start site of exon 2). Open and gray triangles represent two types of tandem short repeats in intron 1. **B**, Constructs 5–10 contain regional deletion of *ponli* intron 1. **C**, The engineered 1904 bp *ponli* *cis*-regulatory DNA, which contains the 818 bp intronic region (hatched), drove retinal expression in ~30% of injected embryos. **D**, Retinal GFP expression, revealed by stereo-immunofluorescence microscopy, is weaker in *Tg(ponli^{1,904}:GFP)^{pt151-#3}* (pt151-#3 for short) than in *Tg(LCR^{RH2}-RH2-2:GFP)^{pt115-k}* (pt115-k for short) (Fang et al., 2013) and *Tg(SWS1:GFP)* (Takechi et al., 2003). **E**, **F**, In *Tg(ponli^{1,904}:GFP)^{pt151-#3}*, GFP is restrictively expressed in RGB cones, as revealed by confocal immunohistology of photoreceptors at transverse sections (**E**) and vertical sections (**F**). **G**, In *Tg(ponli^{1,904}:GFP)^{pt151-#2}* (pt151-#2 for short), in addition to primary expression in RGB cones, GFP is also sparsely expressed in bipolar cells and rods. **H**, Western blotting reveals that GFP levels in *Tg(ponli^{1,904}:GFP)^{pt151-#2}*, *Tg(ponli^{1,904}:GFP)^{pt151-#3}*, and two other unrelated stable transgenic lines that carry a *ponli^{1,904}:GFP* transgenic reporter gene (w1465-#1 and w1465-#11 for short) are 50–100 times less than those in *Tg(LCR^{RH2}-RH2-2:GFP)^{pt115-k}* and *Tg(SWS1:GFP)*. To make Western blotting signals visually comparable, samples with strong signals were diluted with wild-type eye lysates. The dilution ratios with wild-type retinal lysates are marked in parentheses (transgenic eyes: wild-type eyes). X, Unidentified number of GFP positive cones from w1465-#1, w1465-#11, pt151-#2, or pt151-#3; X/20, X/30, and X/50, 1/20th, 1/30th, and 1/50th of X amount of GFP positive cones, respectively. γ -Tubulin blotting serves as loading controls; each lane was loaded with retinal lysates from ten 7 dpf eyes.

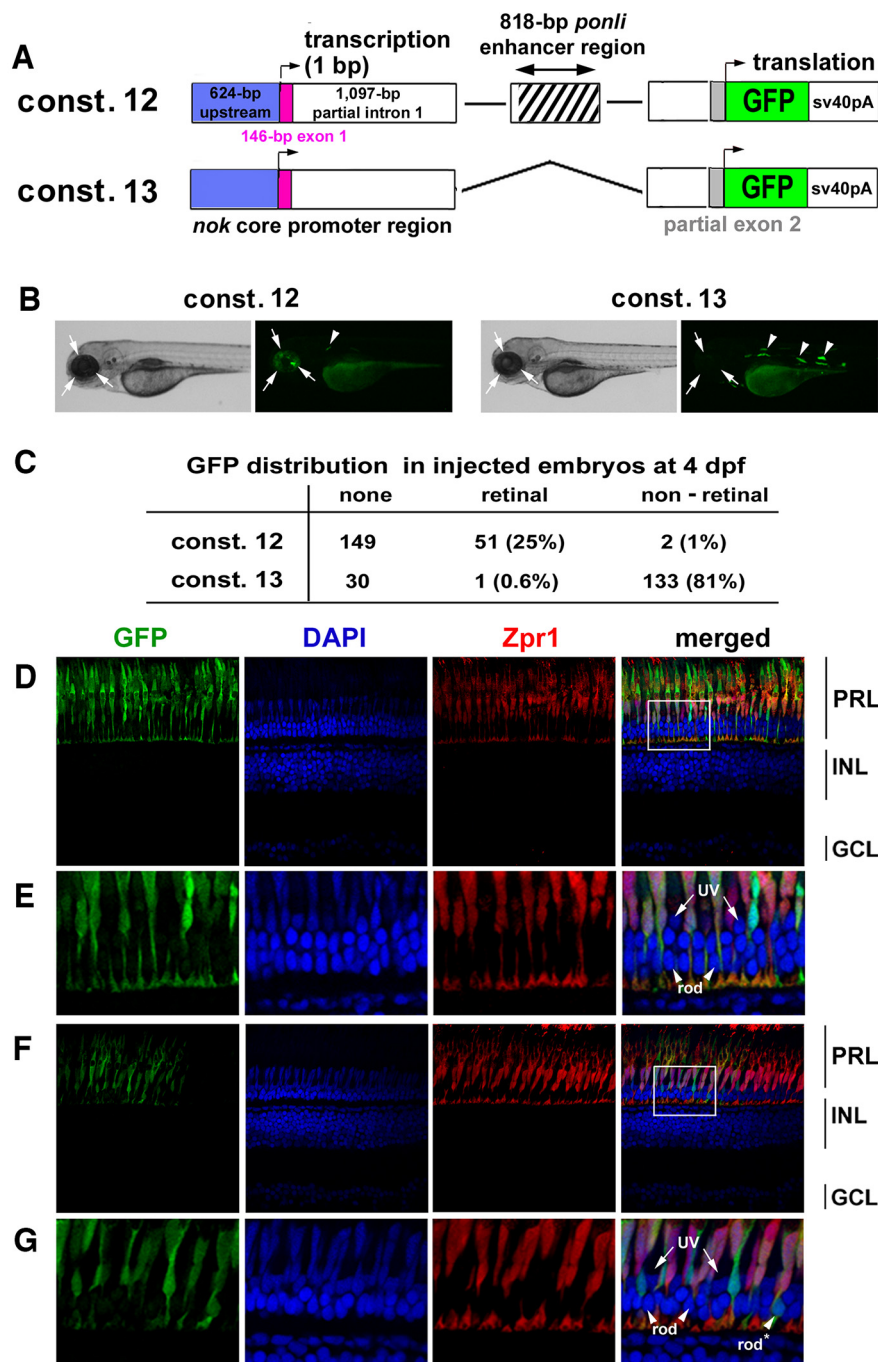


Figure 5. A *ponli* enhancer resides in the 818 bp intronic DNA. **A**, Chimeric construct 12 contains *nok*'s core promoter region, the 818 bp *ponli* intronic DNA, 3' portion of *ponli*'s intron 1, noncoding region of *ponli*'s exon 2, and a GFP reporter. Construct 13 is the same as construct 12 but lacks the 818 bp *ponli* intronic DNA. Pink boxes represent the first noncoding exon of the *nok* gene. Blue boxes represent the 624 bp region upstream of *nok*'s transcription start site. **B**, **C**, Statistics show that construct 12 drove retinal expression in 25% of injected embryos, whereas construct 13 drove nonretinal expression, mainly ectopically in muscle cells (arrowheads). **D–G**, Restrictive GFP expression in RGB cones was observed in 49 of 60 transgenic cell clones that carried construct 12 (**D**, **E**), whereas in the remaining 11 clones, a small number of rods express GFP (**F**, **G**; asterisk). **E**, **G**, Enlarged areas that are boxed in **D** and **F**, respectively. PRL, Photoreceptor layer; INL, inner nuclear layer; GCL, ganglion cell layer.

teleost *ponli* genes are conserved not only in sequences but also in RGB cone-specific transcriptional activity.

Roles of *ponli* EM1–5 and nearby sequences in RGB cone-specific transcription

To determine whether or not *ponli* EM1–5 and their nearby sequences are required for RGB cone-specific transcription, we mu-

tated these motifs individually and then examined the resulting effects on retinal expression patterns. Because such analysis was laborious and because it is impractical to generate and analyze a stable fish line for each mutation, we devised a larval immunohistological microscopic method to evaluate the effects directly in injected fish. The essence of this method is to inject the mutation constructs in *Tg(Rh2-1:HA-mCherry)^{pt120}* embryos at the 1-cell stage and then at 9 dpf, to determine the types of GFP-expressing cells according to the photoreceptors' morphologies, Zpr1 immunoreactivity, and mCherry signals (Fig. 7).

The substitution mutations were generated for both tilapia and zebrafish *ponli* enhancer regions by replacing these short sequences individually with unrelated sequences of the same lengths (~10 bp): a total of 14 substitution mutations for tilapia *ponli* enhancer and 18 substitution mutations for zebrafish *ponli* enhancer (Fig. 8A–C).

These substitution mutation analyses revealed that mutation of EM2 (GAACA GATGG), which is conserved in all six fish species, abolished retinal expression, suggesting that EM2 plays an activation role in RGB cone-specific transcription (Fig. 8B,C). In addition, the sequences around EM2 also appear to be essential, even though there is no clear similarity conservation between zebrafish and the other five fish species in these nearby sequences (Fig. 6C). However, mutation of the second zebrafish EM2, which is in reverse orientation and 10 bp downstream of the first EM2, did not affect retinal expression (Fig. 8C).

Surprisingly, unlike EM2, loss of EM1, EM3, EM4, and EM5 did not appear to affect retinal expression, except when EM1 was next to EM2 in tilapia (Fig. 8B,C). Nevertheless, the cell type-specific transcription can be slightly compromised. For example, mutation of tilapia motif 13 resulted in more ectopic expression in non-*RGB*-cone cells that form vertical clones, suggesting that ectopic expression might be activated in retinal progenitors; this ectopic expression might be augmented by position effect because different transgenic clones in the same retina displayed a different level of specificity (Fig. 8D). Interestingly, even though these mutant constructs resulted in ectopic expression in rods, bipolar cells, amacrine cells, and ganglion cells, we very rarely observed any expression in UV cones. And the ratio of GFP expression in RGB cones remains at an ~2:2:1 ratio (Fig. 8B,C). Together, the mutation analyses suggest that EM2 and its close neighboring sequences are essential for retinal expression in an all-or-none fashion and

interestingly, even though these mutant constructs resulted in ectopic expression in rods, bipolar cells, amacrine cells, and ganglion cells, we very rarely observed any expression in UV cones. And the ratio of GFP expression in RGB cones remains at an ~2:2:1 ratio (Fig. 8B,C). Together, the mutation analyses suggest that EM2 and its close neighboring sequences are essential for retinal expression in an all-or-none fashion and



that other conserved motifs may collectively influence cell type specificity.

crb2b enhancers share *ponli* enhancer motifs

The RGB cones are very similar in their morphologies and functions; such structural and functional similarities must be based on the similarity in their gene expression profiling. This made us wonder whether similar enhancers are used to govern other RGB cone-specific genes, such as *crumbs2b* (*crb2b*), which is expressed with the same spatiotemporal patterns as the *ponli* gene in the retina, and whose protein product colocalizes with Ponli (Zou et al., 2010, 2012).

To explore this possibility, we next compared the first intron of the *crb2b* genes among the five fish species (medaka, fugu, tetraodon, stickleback, and zebrafish) and found four stretches of sequence that are conserved in the first introns (C1–C4; Fig. 9A). Strikingly, the conserved stretch C3 also contains three motifs that are identical or very similar to the EM2, EM3, and EM4 of the medaka *ponli* enhancer (Fig. 9A), implying that this stretch of DNA might function as a *crb2b* enhancer. To test this hypothesis, we replaced the *ponli* enhancer of the transgenic reporter assay construct 11 with an 882 bp C3-containing DNA fragment (Fig. 9B); the resulting construct 15 was injected into zebrafish embryos, and the transcriptional activity in larvae was examined. We found that this 882 bp DNA indeed induced specific RGB cone expression (Fig. 9C,E). As for *ponli* enhancers, loss of EM2 completely abolished the 882 bp DNA's retinal transcriptional activity (Fig. 9D). However, mutations of EM3 and EM4 did not abolish retinal expression; rather, they slightly increased nonphotoreceptor expression and also slightly altered RGB cone ratios (Fig. 9D,E). These results suggest that the 882 bp DNA fragment of *crb2b* intron 1 contains a *crb2b* enhancer that is sufficient to drive RGB cone-specific transcription in conjunction with a core promoter.

The functional and sequence similarities between *ponli* and *crb2b* enhancers suggest that they constitute a conserved *cis*-regulatory mechanism that governs RGB cone-specific transcription of *ponli* and *crb2b*. We thus name these enhancers collectively *rainbow* enhancers to honor their transcriptional activity in RGB cones, which largely constitute the beginning of the color vision pathway.

Discussion

In this study, we set out to identify CREs of the *ponli* and *crb2b* genes that regulate RGB cone-specific transcription and to determine whether or not their *cis*-regulatory mechanisms are conserved. By assessing the transcriptional activities of various DNA

regions of teleost *ponli* and *crb2b* genes with homology comparison and transgenic reporter assays, we found that the *ponli* core promoter is TATA-less and contains three conserved core promoter motifs (CPM1–3); in addition, the intronic *ponli* enhancers, containing five conserved enhancer motifs (EM1–5), determine RGB cone-specific transcription. We further showed that *crb2b* enhancers are also intronic and that they share EMs with *ponli* enhancers, suggesting that *crb2b* and *ponli* enhancers, collectively called *rainbow* enhancers, underlie a conserved *cis*-regulatory mechanism.

Organization of *rainbow* enhancer motifs

The sequence and functional conservation among *ponli* and *crb2b* enhancers suggest that *rainbow* enhancers are constructed according to a common “code.” Indeed, the concept of “enhancer *cis*-regulatory codes” is gaining support from growing evidence. According to this concept, a specific code governs how and what enhancer motifs are organized in an enhancer to confer its tissue-specific transcriptional activity (Yañez-Cuna et al., 2013). However, it remains challenging to interpret *cis*-regulatory codes because enhancer motifs are often short and degenerate; in addition, enhancers are often studied individually, making it more difficult to reveal the underlying general principles. Thus, the identification of *rainbow* enhancers for two different genes offers a good opportunity to use homology and functional conservation to crack the *cis*-regulatory code that governs RGB cone-specific transcription. Although a complete understanding of *rainbow* enhancer code has yet to be reached, our current study revealed two prominent features of the *rainbow* enhancers.

First, the transcriptional activity of *rainbow* enhancers depends on EM2 in an all-or-none fashion because mutation of the EM2s in the zebrafish *ponli* enhancer, tilapia *ponli* enhancer, and medaka *crb2b* enhancer completely abolished their transcriptional activities (Figs. 8B,C, 9D,E). The consensus sequence of EM2 (GAACAGATGG) and its surrounding sequences do not resemble the consensus sequences of the U2 and U12 splicing branch sites TTCTCATTC and TTTTCCTTAACCTTTT, which localize most commonly 20–40 bp upstream of the 3' splicing sites (Yeo et al., 2004; Turunen et al., 2013), suggesting that loss of EM2 is unlikely to block splicing; rather, the transcriptional dependency on EM2 would suggest that EM2 plays an essential role in promoting transcription, possibly by recruiting a transcription factor that has yet to be identified. Supporting this notion, eliminating the need for splicing by fusing exon 1 and exon 2 blocked retinal expression completely (Fig. 4B, construct 9). The EM2-binding transcription factor may itself be expressed in an RGB cone-specific manner. In such a case, the enhancer of an EM2-binding transcription factor gene may also be a *rainbow* enhancer, and it may augment its own transcription in a positive-feedback loop to increase the robustness of RGB cone-specific transcription. Alternatively, the EM2-binding transcription factor may not be restrictively expressed in RGB cones; instead, it may promote nonspecific transcription. In such a case, other enhancer motifs may restrict *rainbow* enhancers' transcriptional activity to RGB cones. Interestingly, the sequences immediately next to EM2 are also important for transcriptional activity, even though they do not have to be as highly conserved as EM2 itself (Figs. 6, Fig. 8). Such sequence divergence of essential motifs is not unique; another example is the critical *cis*-regulatory elements of the fly *sparkling* enhancer, which has rapidly changed over a short period of time during evolution (Swanson et al., 2011).

(Figure legend continued.) presumably constitute the *ponli* minimal core promoter motifs: CPM1, CPM2, and CPM3 (also marked with consensus sequences). **C**, The candidate enhancer motifs (EM1–5) and nearby conserved sequences in the mid-intron 1 region of the six teleost species are labeled with numbers and alphabets (also see Fig. 8). **B, C**, Red represents sequences that display the highest conservations in all six teleost species. Black and gray represent complete and partial nucleotide conservations among five nonzebrafish species, respectively. Hyphens indicate sequence gaps. Numbers and green lines indicate the motifs in the tilapia *ponli* enhancer. Letters and blue lines indicate motifs in the zebrafish *ponli* enhancer. The motifs that are conserved among all six teleost fish are marked with numbers (4, 5, 7, 9, and 10) (also marked with EM1–5 and consensus sequences), and superscripts *t* and *z* are added to indicate their being tilapia motifs and zebrafish motifs, respectively. Black lines indicate the candidate CRX binding sites within EM1s. **D**, To assess the transcriptional activity of tilapia *ponli* intronic DNA, the 818 bp zebrafish mid-intronic DNA in construct 11 was replaced with the 212 bp tilapia *ponli* intronic DNA (construct 14). **E**, Tilapia 212 bp intronic DNA drove RGB cone-specific GFP expression in *Tg(ponli^{tilapia};GFP)^{pt152-#14}*.

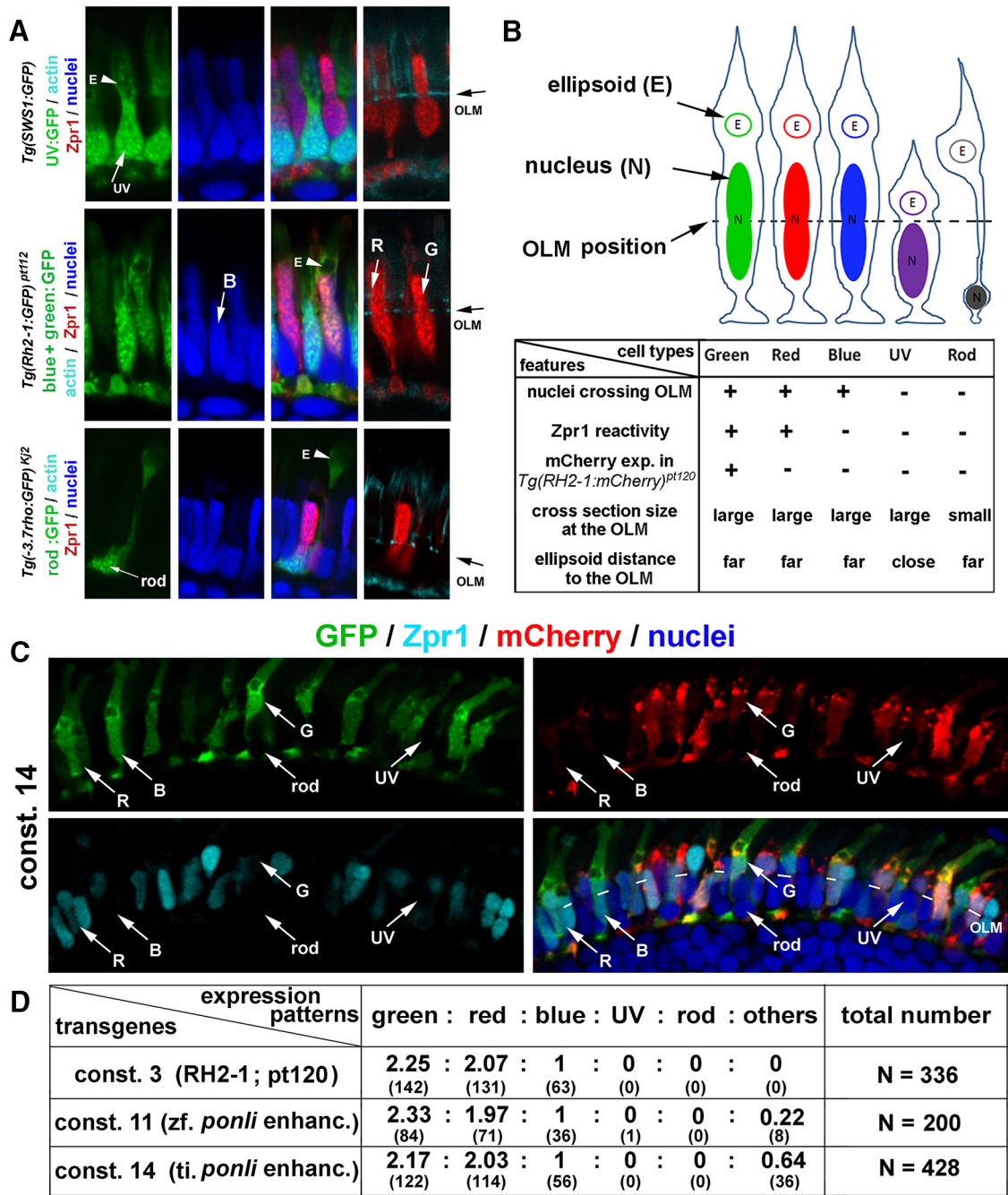


Figure 7. A larval immunohistological microscopic method was used to assess the transcriptional activity of *cis*-regulatory DNAs in 9 dpf *Tg(RH2-1:mCherry)^{pt120}* larvae. **A, B**, By analyzing photoreceptor characteristics in *Tg(-3.7rho:eGFP)^{K12}*, *Tg(SWS1:GFP)*, and *Tg(RH2-1:GFP)^{pt112}* fish lines (**A**), a set of criteria were established to distinguish each of five photoreceptor types in the *Tg(RH2-1:mCherry)^{pt120}* background (**B**) at 9 dpf, when stereotypical mosaic alignment of cones has not yet developed in larvae: immunoreactivity to the Zpr1 antibody, nuclear and ellipsoid positions relative to the outer limiting membrane (OLM), cross section cell diameter at the OLM, and mCherry expression in green cones. **C, D**, In agreement with the expression pattern in stable fish lines *Tg(ponli^{1,904}:GFP)^{pt151-#3}* and *Tg(ponli^{stlapia}:GFP)^{pt152-#14}*, the larval immunohistological microscopic method confirmed that zebrafish or tilapia *ponli* enhancer drove specific GFP expression in RGB cones in transient analysis (**C**, example images; quantifications are summarized in the table in **D**).

Second, in addition to EM2, *rainbow* enhancers also contain several other highly conserved motifs, which are concentrated around EM2 in a range of ~200 bp (Figs. 6C, 9A), a typical size for many known enhancers (Yáñez-Cuna et al., 2013). However, these conserved motifs are not absolutely required because some of them, such as EM5, are even absent or highly degenerated and could not be identified in medaka *crb2b* enhancer (Fig. 9A); in addition, mutations of these motifs did not abolish retinal expression as EM2 mutations, even though these mutations slightly compromised cell type specificity of transgenic expression. In addition to sequence

variations, the distance between highly conserved motifs can vary between species and between genes (Fig. 8B,C and Fig. 9A, respectively). Moreover, the orientation of zebrafish *ponli* enhancer's EM5 and the second duplicated EM2 are even in a reverse orientation compared with the same motifs in tilapia *ponli* enhancer (Fig. 6C). Despite these variations among *rainbow* enhancers, both tilapia *ponli* enhancer and medaka *crb2b* enhancer can still activate RGB cone-specific transcription in zebrafish, suggesting that the transcription factors are not required to bind to *rainbow* enhancers in a strict spatial order in order for them to restrict transcriptional

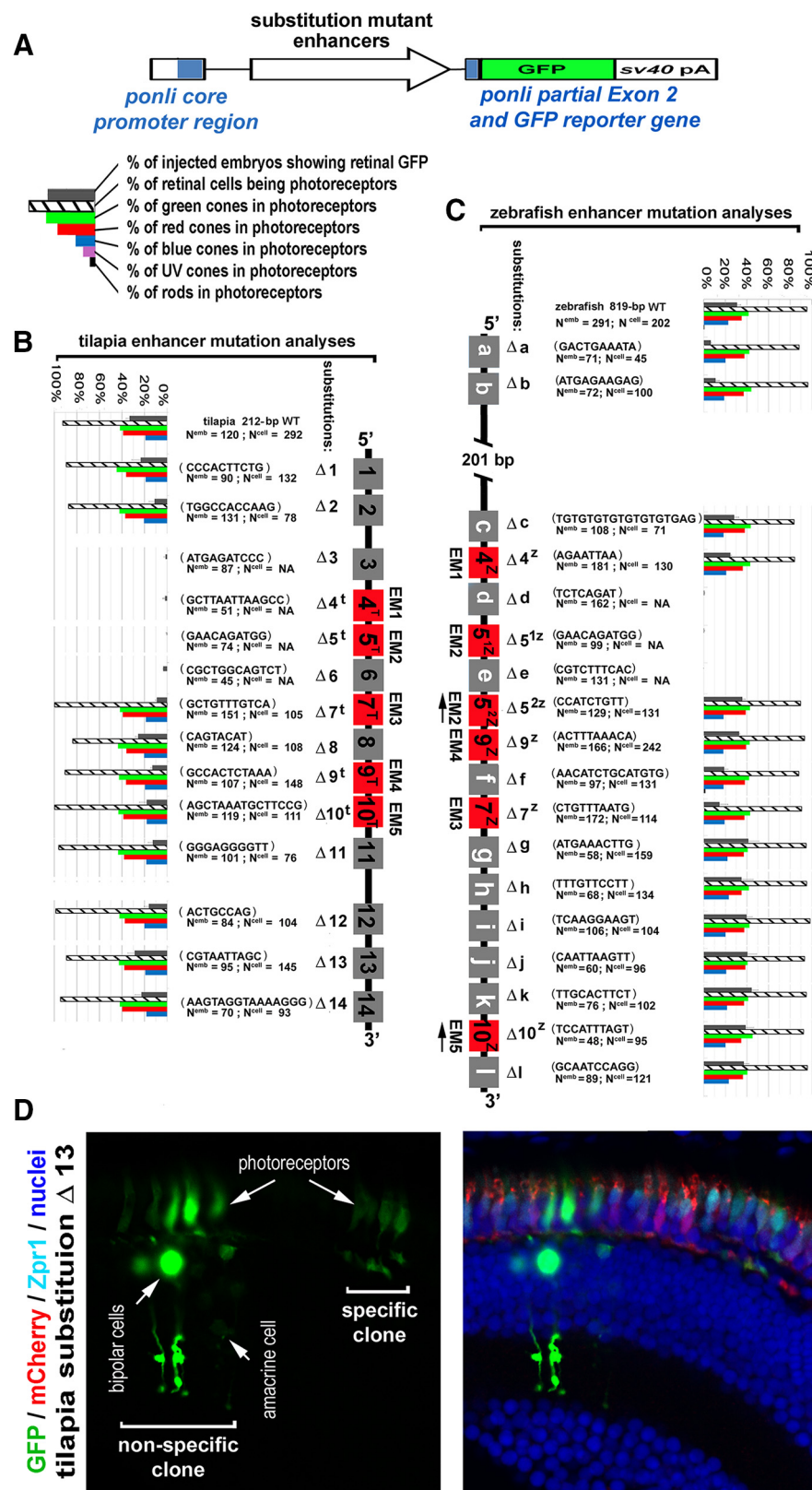


Figure 8. Transgenic reporter analyses revealed essential motifs of zebrafish and tilapia *ponli* enhancers. **A**, A scheme illustrates the transgenic reporting assay for analyzing the functions of individual motifs of *ponli* enhancers. The example histograms illustrate the quantification strategy. Gray bars represent the percentages of surviving embryos that displayed retinal GFP expression. Hatched bars represent the percentages of photoreceptors among total GFP-positive retinal cells. Green, red, blue, magenta, and black bars represent percentages of green cones, red cones, blue cones, UV cones, and rods among total GFP-positive photoreceptors, respectively. **B**, **C**, Substitution mutations of tilapia motifs 3, 4^t (EM1), 5^t (EM2), and 6 as well as zebrafish motifs d, 5^{1z} (EM2), and completely abolished retinal transgenic expression, whereas substitutive mutations of some other motifs resulted in ectopic

activation in RGB cones. This functional conservation is surprising because it is thought that the composition, orientation, spacing, and alignment of enhancer motifs play critical roles in transcriptional specificity by creating unique and continuous recognition surfaces for transcription factor binding and activation in a distinct combinatorial fashion, as exemplified by the paradigmatic interferon- β enhanceosome (Panne et al., 2007; Panne, 2008; Levine, 2010; Swanson et al., 2010). The coexistence of both organizational variations of *rainbow* enhancer motifs and functional conservation of *rainbow* enhancers raises an interesting question: How can enhancers be structurally varied on the one hand and still maintain their strict transcriptional specificity on the other hand? Perhaps loose *cis*-regulatory codes confer the freedom that transcription factors need to assemble in distinct configurations and to paradoxically enforce strict transcriptional specificity.

Similarities and differences between *rainbow* enhancers and other cone-specific CREs

Compared with other cone-specific CREs, *rainbow* enhancers display three prominent features. First, *rainbow* enhancers differ from others in the cellular patterns of transcriptional activities. To date, a limited number of cone-specific CREs have been identified; they are either active in one distinct type of cone photoreceptor or active in all types of cone photoreceptor. Examples of the former are the CREs of various cone opsin genes (Takechi et al., 2003, 2008; Luo et al., 2004; Tsujimura et al., 2010, 2015); examples of the latter are the CREs of cone arrestin and cone transducin α subunit (Pickrell et al., 2004; Smyth et al., 2008). Unlike these two classes of cone-specific CREs, *rainbow* enhancers' unique RGB cone-specific expression implies that the underlying mechanism of *rainbow* enhancers is distinct from others. Moreover, because the

expression of GFP in nonphotoreceptor retinal cells. Motif sequences that were substituted are shown in parentheses. Red boxes represent the highly conserved enhancer motifs (EM1–5). Gray boxes represent less-conserved sequences. N^{emb}, Total number of embryos that survived at 9 dpf after injection; N^{cell}, total number of GFP-positive retinal cells that were identified across all retinal GFP-positive embryos. The motifs are labeled with numbers and letters in the same way as in Figure 6. **D**, In an embryo injected with the tilapia substitution $\Delta 13$ construct, two clones of transgenic cells displayed different expression patterns: one showed GFP expression in photoreceptors as well as in bipolar and amacrine cells; the other showed exclusive photoreceptor expression.

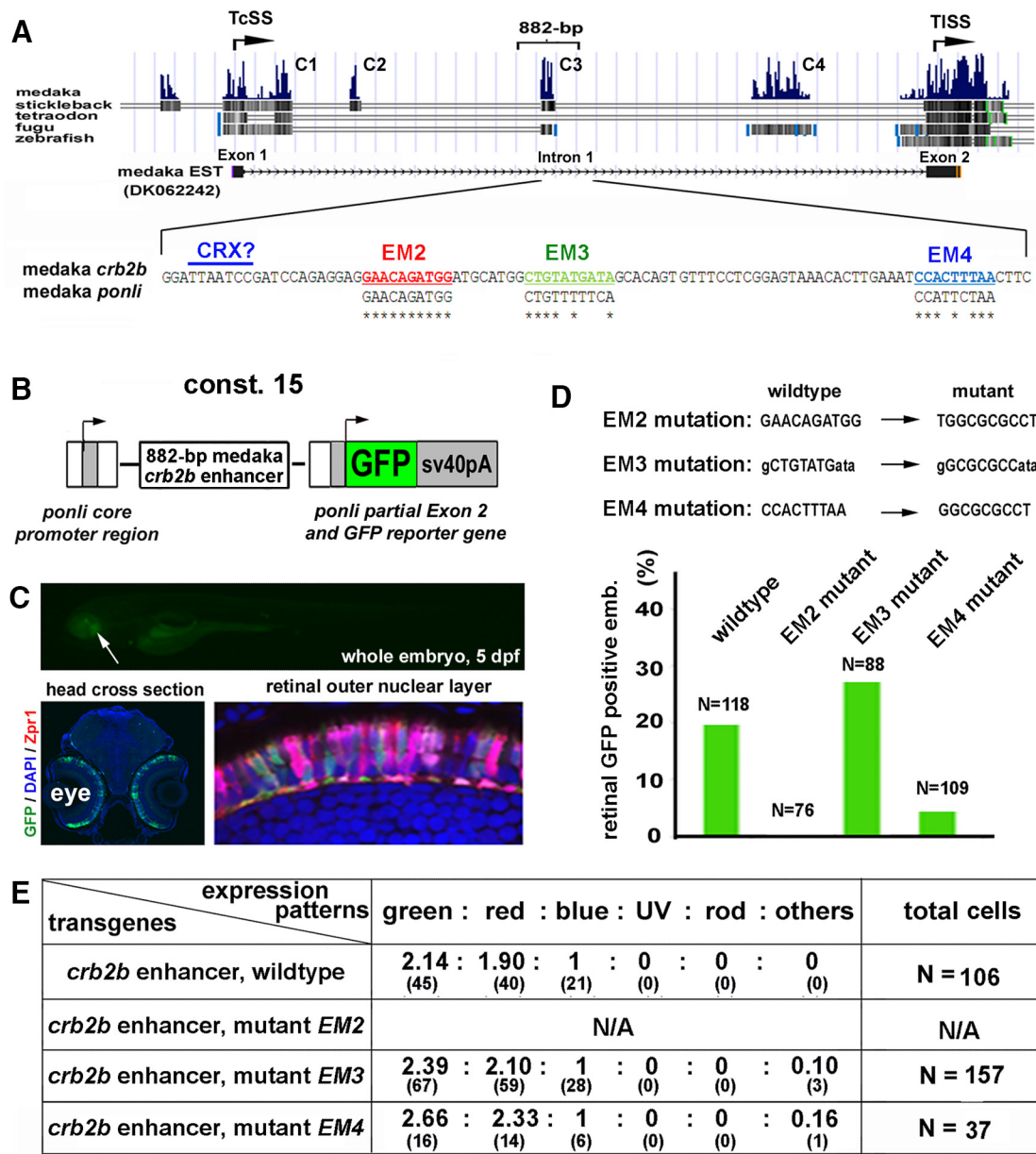


Figure 9. Medaka *crb2b* intronic enhancer drives RGB cone-specific transcription. **A**, Sequence alignment of teleost *crb2b* genes revealed conserved regions around the transcription start site and intron 1. The middle conserved intronic region (C3) contains EM2, EM3, and EM4, which are also present in the medaka *ponli* enhancers. A candidate CRX binding site is present 12 bp upstream of EM2. **B**, A scheme illustrates the structure of chimeric construct 15, in which the *ponli* enhancer of construct 11 was replaced with the 882 bp medaka *crb2b* intronic DNA. **C**, Injection of construct 15 led to GFP expression in the photoreceptor cell layer in 5 dpf retina. **D**, Mutation of EM2 of medaka *crb2b* enhancer abolished retinal expression; however, mutations of EM3 and EM4 did not affect retinal expression. **E**, Larval immunohistological microscopy of 9 dpf larvae revealed that *crb2b* enhancer specifically drove RGB cone-specific expression (compared with the restrictive RGB cone ratios of 2.25:2.07:1 in pt120 wild-type fish; Fig. 7D), and mutation of EM3 and EM4 only slightly affected transcriptional specificity.

ponli and *crb2b* genes are unlikely to be the only genes that define RGB cones' structural and functional characteristics, it is tempting to speculate that other RGB cone-specific genes exist and that their transcription is regulated by a family of rainbow enhancers. Second, like most other identified cone-specific CREs, all rainbow enhancers contain a sequence motif that is either identical or very similar to CRX's (cone rod homeobox) consensus target sequence TTAATCC (Chen et al., 1997). These CRX sites localize 5–12 bp upstream of EM2 and constitute 70% of the EM1 motifs (Figs. 6C, 9A). Because retinal expression was abolished by substitution mutations 4' and d, which removed the tilapia *ponli* CRX site entirely and three sevenths of the zebrafish CRX site, respectively, it is tempting to speculate that CRX recognizes EM1.

However, it is puzzling that the removal of the 5' four-sevenths of the zebrafish CRX site or 60% of EM1 (substitution mutation 4') did not affect retinal expression (Figs. 6C, 8C). Regardless, because CRX is also expressed in rods and regulates *rod opsin* by interacting with rod-specific Nrl (neural retina-specific leucine zipper protein) (Rehmtulla et al., 1996; Babu et al., 2006; Peng and Chen, 2007; Reks et al., 2014), it is unlikely that CRX itself determines rainbow enhancers' RGB cone-specific transcription. Finally, rainbow enhancers' most critical and conserved motif, EM2 (GAACAGATGG), does not appear to show significant similarity to other cone CRE motifs, including the CPRE-1 motif (cone photoreceptor regulatory element 1) of the α subunit of cone transducin, which is expressed in all four zebrafish cone

types (Smyth et al., 2008). This uniqueness of EM2 highlights the importance of identifying an EM2 transcription factor for understanding *rainbow* enhancer activation. Perhaps as *Nrl* drives photoreceptors to become rods (Oh et al., 2007), the yet to be identified EM2-binding transcription factor may drive photoreceptor to become RGB cones.

Rainbow enhancers as a tool for transgenic research

Two useful expression features of *rainbow* enhancers make them excellent tools for transgenic research of teleost photoreceptors. First, unlike opsin *cis*-regulatory DNAs (Hamaoka et al., 2002; Takechi et al., 2003; Tsujimura et al., 2007), *rainbow* enhancers can be used to drive simultaneous transcription in zebrafish RGB cones; this expression pattern is particularly desirable either for studying cell-nonautonomous gene functions because transgenic manipulation of such genes in one member of the RGB cones may not result in apparent phenotypes due to rescuing effects by neighboring cells or for therapeutic gene expression in multiple types of cones. Second, to avoid potential adverse consequences of transgenic overexpression by *opsin cis*-regulatory DNAs, *rainbow* enhancers can be used to drive much more modest transcription (Figs. 3, 4*D,H*).

Conclusion

In conclusion, we have identified two teleost *rainbow* enhancers, the *ponli* and *crb2b* enhancers; to our knowledge, this is the first report of such enhancers with RGB cone-specific transcriptional activity in teleost retina. Our work provides a starting point from which to study *rainbow* enhancer-mediated transcription of *ponli* and *crb2b* and possibly other RGB cone-specific genes; these genes may collectively define RGB cones' distinct transcriptional profile and consequently their functions in initiating color vision.

References

- Arshavsky VY, Lamb TD, Pugh EN Jr (2002) G proteins and phototransduction. *Annu Rev Physiol* 64:153–187. [CrossRef Medline](#)
- Babu S, McIlvain V, Whitaker SL, Knox BE (2006) Conserved cis-elements in the *Xenopus* red opsin promoter necessary for cone-specific expression. *FEBS Lett* 580:1479–1484. [CrossRef Medline](#)
- Carninci P, Sandelin A, Lenhard B, Katayama S, Shimokawa K, Ponjavic J, Sempé CA, Taylor MS, Engström PG, Frith MC, Forrest AR, Alkema WB, Tan SL, Plessy C, Kodzius R, Ravasi T, Kasukawa T, Fukuda S, Kanamori-Katayama M, Kitazume Y, et al. (2006) Genome-wide analysis of mammalian promoter architecture and evolution. *Nat Genet* 38:626–635. [CrossRef Medline](#)
- Chen S, Wang QL, Nie Z, Sun H, Lennon G, Copeland NG, Gilbert DJ, Jenkins NA, Zack DJ (1997) Crx, a novel Otx-like paired-homeodomain protein, binds to and transactivates photoreceptor cell-specific genes. *Neuron* 5:1017–1030. [CrossRef Medline](#)
- Cooper SJ, Trinklein ND, Anton ED, Nguyen L, Myers RM (2006) Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome. *Genome Res* 16:1–10. [CrossRef Medline](#)
- Fang W, Bonaffini S, Zou J, Wang X, Zhang C, Tsujimura T, Kawamura S, Wei X (2013) Characterization of transgenic zebrafish lines that express GFP in the retina, pineal gland, olfactory bulb, hatching gland, and optic tectum. *Gene Expr Patterns* 13:150–159. [CrossRef Medline](#)
- Gosens I, den Hollander AI, Cremers FP, Roepman R (2008) Composition and function of the Crumbs protein complex in the mammalian retina. *Exp Eye Res* 86:713–726. [CrossRef Medline](#)
- Hamaoka T, Takechi M, Chinen A, Nishiwaki Y, Kawamura S (2002) Visualization of rod photoreceptor development using GFP-transgenic zebrafish. *Genesis* 34:215–220. [CrossRef Medline](#)
- Jacobs GH, Nathans J (2009) The evolution of Primate color vision. *Sci Am* 300:56–63. [CrossRef Medline](#)
- Juven-Gershon T, Kadonaga JT (2010) Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Dev Biol* 339:225–229. [CrossRef Medline](#)
- Kulaeva OI, Nizovtseva EV, Polikanov YS, Ulianov SV, Studitsky VM (2012) Distant activation of transcription: mechanisms of enhancer action. *Mol Cell Biol* 32:4892–4897. [CrossRef Medline](#)
- Lenhard B, Sandelin A, Carninci P (2012) Metazoan promoters: emerging characteristics and insights into transcriptional regulation. *Nat Rev Genet* 13:233–245. [CrossRef Medline](#)
- Levine M (2010) Transcriptional enhancers in animal development and evolution. *Curr Biol* 20:R754–R763. [CrossRef Medline](#)
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(−Delta Delta C(T)) Method. *Methods* 25:402–408. [CrossRef Medline](#)
- Luo W, Williams J, Smallwood PM, Touchman JW, Roman LM, Nathans J (2004) Proximal and distal sequences control UV cone pigment gene expression in transgenic zebrafish. *J Biol Chem* 279:19286–19293. [CrossRef Medline](#)
- Oh EC, Khan N, Novelli E, Khanna H, Strettoi E, Swaroop A (2007) Transcription of cone precursors to functional rod photoreceptors by bZIP transcription factor NRL. *Proc Natl Acad Sci U S A* 104:1679–1684. [CrossRef Medline](#)
- Panne D (2008) The enhanceosome. *Curr Opin Struct Biol* 18:236–242. [CrossRef Medline](#)
- Panne D, Maniatis T, Harrison SC (2007) An atomic model of the interferon-beta enhanceosome. *Cell* 129:1111–1123. [CrossRef Medline](#)
- Peng GH, Chen S (2007) Crx activates opsin transcription by recruiting HAT-containing co-activators and promoting histone acetylation. *Hum Mol Genet* 16:2433–2452. [CrossRef Medline](#)
- Pickrell SW, Zhu X, Wang X, Craft CM (2004) Deciphering the contribution of known cis-elements in the mouse cone arrestin gene to its cone-specific expression. *Invest Ophthalmol Vis Sci* 45:3877–3884. [CrossRef Medline](#)
- Pocha SM, Knust E (2013) Complexities of Crumbs function and regulation in tissue morphogenesis. *Curr Biol* 23:R289–R293. [CrossRef Medline](#)
- Pujic Z, Omori Y, Tsujikawa M, Thisse B, Thisse C, Malicki J (2006) Reverse genetic analysis of neurogenesis in the zebrafish retina. *Dev Biol* 293:330–347. [CrossRef Medline](#)
- Raymond PA, Barthel LK, Curran GA (1995) Developmental patterning of rod and cone photoreceptors in embryonic zebrafish. *J Comp Neurol* 359:537–550. [CrossRef Medline](#)
- Rehmtulla A, Warwar R, Kumar R, Ji X, Zack DJ, Swaroop A (1996) The basic motif-leucine zipper transcription factor Nrl can positively regulate rhodopsin gene expression. *Proc Natl Acad Sci U S A* 93:191–195. [CrossRef Medline](#)
- Reks SE, McIlvain V, Zhuo X, Knox BE (2014) Cooperative activation of *Xenopus* rhodopsin transcription by paired-like transcription factors. *BMC Mol Biol* 15:4. [CrossRef Medline](#)
- Robinson J, Schmitt EA, Hárosi FI, Reece RJ, Dowling JE (1993) Zebrafish ultraviolet visual pigment: absorption spectrum, sequence, and localization. *Proc Natl Acad Sci U S A* 90:6009–6012. [CrossRef Medline](#)
- Rouault H, Santolini M, Schweisguth F, Hakim V (2014) Imogene: identification of motifs and cis-regulatory modules underlying gene coregulation. *Nucleic Acids Res* 42:6128–6145. [CrossRef Medline](#)
- Smyth VA, Di Lorenzo D, Kennedy BN (2008) A novel, evolutionarily conserved enhancer of cone photoreceptor-specific expression. *J Biol Chem* 283:10881–10891. [CrossRef Medline](#)
- Solomon SG, Lennie P (2007) The machinery of colour vision. *Nat Rev Neurosci* 8:276–286. [CrossRef Medline](#)
- Swanson CI, Evans NC, Barolo S (2010) Structural rules and complex regulatory circuitry constrain expression of a Notch- and EGFR-regulated eye enhancer. *Dev Cell* 18:359–370. [CrossRef Medline](#)
- Swanson CI, Schimmer DB, Barolo S (2011) Rapid evolutionary rewiring of a structurally constrained eye enhancer. *Curr Biol* 21:1186–1196. [CrossRef Medline](#)
- Takechi M, Hamaoka T, Kawamura S (2003) Fluorescence visualization of ultraviolet-sensitive cone photoreceptor development in living zebrafish. *FEBS Lett* 553:90–94. [CrossRef Medline](#)
- Takechi M, Seno S, Kawamura S (2008) Identification of cis-acting elements repressing blue opsin expression in zebrafish UV cones and pineal cells. *J Biol Chem* 283:31625–31632. [CrossRef Medline](#)
- Thermes V, Grabher C, Ristoratore F, Bourrat F, Choulika A, Wittbrodt J, Joly JS (2002) I-SceI meganuclease mediates highly efficient transgenesis in fish. *Mech Dev* 118:91–98. [CrossRef Medline](#)

- Tsujikawa M, Malicki J (2004) Genetics of photoreceptor development and function in zebrafish. *Int J Dev Biol* 48:925–934. [CrossRef Medline](#)
- Tsujimura T, Chinen A, Kawamura S (2007) Identification of a locus control region for quadruplicated green-sensitive opsin genes in zebrafish. *Proc Natl Acad Sci U S A* 104:12813–12818. [CrossRef Medline](#)
- Tsujimura T, Hosoya T, Kawamura S (2010) A single enhancer regulating the differential expression of duplicated red-sensitive opsin genes in zebrafish. *PLoS Genet* 6:e1001245. [CrossRef Medline](#)
- Tsujimura T, Masuda R, Ashino R, Kawamura S (2015) Spatially differentiated expression of quadruplicated green-sensitive RH2 opsin genes in zebrafish is determined by proximal regulatory regions and gene order to the locus control region. *BMC Genet* 16:130. [CrossRef Medline](#)
- Turunen JJ, Niemelä EH, Verma B, Frilander MJ (2013) The significant other: splicing by the minor spliceosome. *Wiley Interdiscip Rev RNA* 4:61–76. [CrossRef Medline](#)
- Warner JB, Philippakis AA, Jaeger SA, He FS, Lin J, Bulyk ML (2008) Systematic identification of mammalian regulatory motifs' target genes and functions. *Nat Methods* 5:347–353. [CrossRef Medline](#)
- Wei X, Malicki J (2002) nagie oko, encoding a MAGUK-family protein, is essential for cellular patterning of the retina. *Nat Genet* 31:150–157. [CrossRef Medline](#)
- Wei X, Zou J, Takechi M, Kawamura S, Li L (2006) Nok plays an essential role in maintaining the integrity of the outer nuclear layer in the zebrafish retina. *Exp Eye Res* 83:31–44. [CrossRef Medline](#)
- Yáñez-Cuna JO, Kvon EZ, Stark A (2013) Deciphering the transcriptional cis-regulatory code. *Trends Genet* 29:11–22. [CrossRef Medline](#)
- Yeo G, Hoon S, Venkatesh B, Burge CB (2004) Variation in sequence and organization of splicing regulatory elements in vertebrate genes. *Proc Natl Acad Sci U S A* 101:15700–15705. [CrossRef Medline](#)
- Zhang SS, Xu X, Liu MG, Zhao H, Soares MB, Barnstable CJ, Fu XY (2006) A biphasic pattern of gene expression during mouse retina development. *BMC Dev Biol* 6:48. [CrossRef Medline](#)
- Zou J, Yang X, Wei X (2010) Restricted localization of ponli, a novel zebrafish MAGUK-family protein, to the inner segment interface areas between green, red, and blue cones. *Invest Ophthalmol Vis Sci* 51:1738–1746. [CrossRef Medline](#)
- Zou J, Wang X, Wei X (2012) Crb apical polarity proteins maintain zebrafish retinal cone mosaics via intercellular binding of their extracellular domains. *Dev Cell* 22:1261–1274. [CrossRef Medline](#)
- Zou J, Wen Y, Yang X, Wei X (2013) Spatial-temporal expressions of Crumbs and Nagie oko and their interdependence in zebrafish CNS during early development. *Int J Dev Neurosci* 31:770–782. [CrossRef Medline](#)