Title

Mechanistic insights into the evolution for the differential expression of tandemly arrayed cone opsin genes in zebrafish

Author

Taro Tsujimura

Affiliation

Institute for the Advanced Study of Human Biology (WPI-ASHBi), Kyoto University, Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan. E-mail: tsujimura.taro.4m@kyoto-u.ac.jp

Abstract

The genome of many organisms contains several loci consisting of duplicated genes that are arrayed in tandem. The daughter genes produced by duplication typically exhibit differential expression patterns with each other or otherwise experience pseudogenization. Remarkably, opsin genes in fish are preserved after many duplications in different lineages. This fact indicates that fish opsin genes are characterized by a regulatory mechanism that could intrinsically facilitate the differentiation of the expression patterns. However, little is known about the mechanisms that underlie the differential expression patterns or how they were established during evolution. The loci of green (RH2) and red (LWS) sensitive cone opsin genes in zebrafish have been used as model systems to study the differential regulation of tandemly arrayed opsin genes. Over a decade of studies have uncovered several mechanistic features that might have assisted the differentiation and preservation of duplicated genes. Furthermore, recent progress in the understanding of the transcriptional process in general has added essential insights. In this article, I summarize the current understanding of the transcriptional regulation of differentially expressed tandemly arrayed cone opsin genes in zebrafish and discuss a possible evolutionary scenario that could achieve this differentiation.

Keywords

Zebrafish, opsin, cone photoreceptor, color vision, gene duplication, subfunctionalization, evolution, enhancer, gene expression, cis-regulation, trans-regulation

Expression differentiation of duplicated genes as a course of subfunctionalization

Duplication is a critical source for the emergence of novel genes during genome evolution. When daughter genes after duplication are equivalent to each other, one of them often becomes useless and enters the process of pseudogenization (Zhang, 2003). Thus, for both daughter genes to be stably preserved, with rare exceptions of evolutionary equilibrium (Nowak et al., 1997), they need to functionally diverge. It has been proposed that differentiation of the expression pattern into sub-domains of the original expression area of the ancestral gene should be a major course for the duplicated genes to be fixed (Zhang, 2003). It is nonetheless unclear how the expression differentiation is achieved among seemingly equivalent pairs of duplicated genes. Gene expressions are under the regulation of complex systems, and researchers have scrutinized how evolution has modulated these systems to have duplicated genes expressed differentially.

Most notably, the duplication-degeneration-complementation (DDC) model was proposed as a theoretical framework to explain this process (Force et al., 1999). This model supposes that the daughter genes are equipped modularly with the same repertoire of cis-elements right after the duplication. Since the two genes are redundant with the same expression pattern, one of them may accommodate mutations. If the mutations are introduced to the coding region, the gene will soon be pseudogenized. However, if the mutations are introduced in the cis-elements, the two genes may eventually complement each other in their expression patterns, while keeping the functionality as a whole in the entire expression domain. This scenario leads to the preservation of both duplicates as essential genes, resulting in subfunctionalization (Force et al., 1999). This assumption may explain whole-genome duplication and the duplication of a large genomic segment. For example, the regulation of *pax6a* and *pax6b* in zebrafish represents the model well (Kleinjan et al., 2008). However, in cases of duplications of short segmented regions, particularly tandem duplications, the copied intervals do not necessarily contain the entire repertoire of the cis-regulatory elements. Therefore, understanding the evolutionary scenario for those duplication events requires another conceptual

framework. Lan and Pritchard recently proposed that the down-regulation of gene expressions as an immediate outcome of co-regulation for shared regulatory elements may invoke evolutionary constraints to sustain the survival of the genes for dosage balancing (Lan & Pritchard, 2016). However, to understand the extent this theory applies to individual cases requires studying the regulatory mechanisms in detail.

It is well known that visual opsin genes in fish have experienced extensive rounds of duplication in different lineages (Lin et al., 2017), many of which have resulted in subfunctionalization, as explained below. The high incidence of subfunctionalization events may indicate that the regulation of fish opsin genes holds a mechanistic feature that could facilitate the differential expression of duplicates. Indeed, taking zebrafish as a model organism, critical mechanisms behind the differential expression have been uncovered, providing clues to understanding the evolutionary scenario for subfunctionalization. In this review, I summarize the accumulated knowledge regarding the regulatory mechanism for tandemly arrayed opsin genes in zebrafish and discuss implications of the mechanism in the process of subfunctionalization.

Duplications of opsin genes in fish

Visual opsin genes encode the protein moiety of the visual pigments in the photoreceptor cells in the retina. The sequences of the amino acids largely determine the absorption spectra of the pigments. The chromophore of the visual pigments are either retinal (A1) or 3,4-dehydroretinal (A2), which also affects the sensitivity to light (Allison et al., 2004; Enright et al., 2015; Shichida & Matsuyama, 2009). In vertebrates, visual opsins are phylogenetically classified into five classes, or types (Yokoyama, 2000): rod opsin (rhodopsin, RH1), which is expressed in rod photoreceptor cells and responsible for vision in dim light, and four others that represent cone opsin genes expressed in cone photoreceptor cells for the color vision. The four cone opsin types are called SWS1, SWS2, RH2, and LWS and are sensitive to ultraviolet, blue, green, and red light, respectively. The discrimination of light of different wavelengths requires a comparison of activity between distinct

photoreceptor cells of different spectral sensitivities (Rister & Desplan, 2011). Typically, for spectral distinction at the cellular level, organisms express different types of cone opsin genes in distinct photoreceptor cells arranged in a mosaic manner in the retina (Rister & Desplan, 2011). Comparative genomic studies indicate that the five visual opsin types were generated with two rounds of whole-genome duplication, which most likely occurred before the split of cyclostomes and gnathostomes (Kuraku et al., 2009; Lagman et al., 2013). The system to allocate expression of different opsin genes in distinct photoreceptor types might have emerged in this ancestor (Baden et al., 2020; Lamb et al., 2007). However, the mechanism underlying the differential expression remains unknown, leaving many questions about how the system for expressing one visual pigment in one photoreceptor cell was established. Addressing this question should consider the evolutionary and developmental origin of the four types of cone photoreceptor cells.

Interestingly, some vertebrates possess subtype genes within the opsin types that were evolutionarily produced by gene duplication. Most notably, fish have an extensive repertoire of subtype opsin genes thanks to duplication and preservation events that repeatedly occurred in different lineages (Lin et al., 2017). For example, zebrafish have four green-sensitive RH2 opsin genes, *RH2-1*, *RH2-2*, *RH2-3*, and *RH2-4*, and two red-sensitive LWS opsin genes, *LWS-1* and *LWS-2* (Chinen et al., 2003) (Figure 1). Similarly, medaka fish have two SWS2, three RH2, and two LWS genes (Matsumoto et al., 2006). Importantly, the subtypes of these opsin genes are functionally differentiated in several ways. First, the absorption spectra are different. For example, the reconstituted photopigments from zebrafish *RH2-1*, *-2*, *-3*, and *-4* genes exhibit peak absorption spectra (λmax) of 467 nm, 476 nm, 488 nm, and 505 nm, respectively, and the λmax of zebrafish *LWS-1* and *LWS-2* are 558 nm and 548 nm, respectively (Chinen et al., 2003). The expression patterns are also differentiated in time and space. The subtypes of zebrafish RH2 and LWS genes are all expressed in specific cone photoreceptor types, i.e., the short members of double cones (SDCs) for RH2 and the long members of double cones (LDCs) for LWS (Robinson et al., 1993; Takechi & Kawamura, 2005; Vihtelic et al., 1999). The expression of *SWS1* and *SWS2* is also specific in short single cones (SSCs) and long

single cones (LSCs), respectively. Thus, the developmental rule for only one type of opsin gene being expressed in one photoreceptor cell is always maintained in zebrafish (Figure 1A). However, along the ontogeny of zebrafish development, *RH2-1* and *RH2-2* are expressed earlier than *RH2-3* and *RH2-4* (Takechi & Kawamura, 2005). Spatially, the former are expressed in the central-to-dorsal area in the retina. In contrast, *RH2-3* is expressed in the narrow banded region surrounding the *RH2-1/-2* area. *RH2-4* is expressed in the most dorsal part of the retina, further circumscribing the *RH2-3* area (Takechi & Kawamura, 2005; Tsujimura et al., 2015b) (Figure 1B). Similarly, the expression of *LWS-2* starts earlier and then becomes restricted to the central-to-dorsal area, while that of *LWS-1* starts later and is confined to the ventral part of the juvenile and adult retina (Takechi & Kawamura, 2005; Tsujimura et al., 2010) (Figure 1B). Interestingly, for both RH2 and LWS groups, the early genes expressed in the central-to-dorsal area encode visual pigments sensitive to light of shorter wavelengths. The others are expressed in the ventral retina and show sensitivity to longer wavelength light (Figure 1B). Thus, zebrafish have different spectral sensitivity depending on the visual space. The considerable difference in the light environment of the different depths in shallow water where they live might demand this visual system to be ecologically adaptive. Similarly, the subfunctionalization of opsin genes was described in many other fish species such as medaka, cichlid, guppy, and others (Carleton et al., 2008; Hoffmann et al., 2007; Hofmann & Carleton, 2009; Matsumoto et al., 2006; Owens et al., 2012; Rennison et al., 2011). Such differentiations have presumably occurred independently in different fish lineages (Lin et al., 2017). Probably, the demand for adaptation to the environment in water, which has variable light conditions, might have forced such evolutionary differentiation (Temple, 2011). Still, it is remarkable that the evolutionary reorganization of the complex regulatory mechanism for the differential expression has been achieved in multiple fish species.

Cis-regulation of zebrafish opsin genes

In order to understand gene regulation for tissue- or cell type-specific expression, it is essential to identify the source of the specificity of both the genes

and cell types. While the surrounding cis-regulatory contexts in the genome determine the specificity of the genes, the trans-regulatory factors define the specificity of the cell types in the regulation. Accordingly, the regulation of zebrafish opsin genes has been studied from these two aspects.

Transgenic reporter assays have been employed to understand the cis-regulation of opsin genes in zebrafish. The surrounding genomic regions linked with fluorescent reporter genes such as GFP (green fluorescent protein) and RFP (red fluorescent protein) were introduced into zebrafish to test the regulatory activities that induce gene expressions in the retina.

SWS1 and *SWS2* genes exist as single genes without tandem copies (Chinen et al., 2003) (Figure 1B). The proximal upstream regions of the coding sequences of *SWS1* and *SWS2* are sufficient to define the specific expression in SSCs and LSCs, respectively (Takechi et al., 2003; 2008). Similarly, the rod photoreceptor-specific expression of RH1 could be recapitulated only with the proximal upstream sequences (Hamaoka et al., 2002; Kennedy et al., 2001). Of note, the cell type-specific expression of the reporters under these proximal regions covered the entire retina in the established transgenic lines, just as the expression of the endogenous genes does (Hamaoka et al., 2002; Takechi et al., 2003; 2008).

By contrast, the expression of RH2 and LWS could not be reproduced only by the upstream sequences; instead, the entire locus with the intact context is required to fully recapitulate the differential expression of the genes (Tsujimura et al., 2010; 2007). Those studies utilized large genomic P1-artificial chromosome (PAC) clones encompassing the whole locus, in which fluorescent reporter genes were integrated into the places of the opsin genes. Transgenic zebrafish with the large constructs showed the reporter expression that recapitulated not only the cell-type specificity but also the temporal and spatial expression patterns (Tsujimura et al., 2007; 2010; 2015b). Importantly, the results solidly underlie the idea that the genomic sequences surrounding the opsin genes in cis are the determinant of the regulatory identity of the genes. Based on this finding, different configurations of reporter constructs were further tested to investigate how the collective action of genomic

regions specifies the expression patterns, as explained below first for RH2 and then for LWS.

Cis-regulation for RH2

Among the RH2 subtypes, *RH2-1* is the first gene to be expressed in the retina during embryogenesis (Takechi & Kawamura, 2005). However, the 1.5-kb promoter region of *RH2-1* failed to drive reporter expression in the retina when tested in zebrafish larvae via transgenesis (Tsujimura et al., 2007). The co-injection of genomic fragments around RH2 together with the *RH2-1* reporter construct functionally identified a single enhancer located 15-kb upstream of *RH2-1* as a region capable of inducing the SDC-specific expression (Tsujimura et al., 2007). On the other hand, a PAC clone that had the enhancer deleted could not induce the expression of the reporters for all the RH2 subtype genes in SDCs (Tsujimura et al., 2007). Thus, the identified enhancer regulates all RH2 genes and was named RH2-locus control region (RH2-LCR) (Figure 2). The sharing of the single enhancer by the subtypes critically underlies the differential expression, as it allows coordinated control of the expression from the cluster.

Importantly, RH2-LCR does not have a spatial or temporal preference for gene induction in the retina (Tsujimura et al., 2007; 2015b). It can even induce gene expression heterologously with the promoter of keratin-8 predominantly in the SDCs in the entire area of the retina (Tsujimura et al., 2007; 2015b). Along with the sequence conservation of RH2-LCR with counterparts in the medaka and pufferfish genomes, experimental evidence strongly suggests that RH2-LCR is a functionally conserved descendant of the ancestral enhancer of RH2 (Tsujimura et al., 2015b) (see Figure 3).

On the other hand, the tandemly copied genes have acquired their own identity when determining the expression specificity during evolution. First, the proximal promoter sequences functionally differ among the quadruplicates. When linked with RH2-LCR, the upstream sequences of *RH2-1* and *RH2-2* induced gene expression in the central-to-dorsal area of the retina, while that of *RH2-4* confined the expression to the ventral area (Tsujimura et al., 2015b). On the other hand, the

upstream region of *RH2-3* induced gene expression throughout the retina with the aid of RH2-LCR (Tsujimura et al., 2015b). These results show that *RH2-1*, *-2*, and *-4* have adopted cis-regulatory elements in the immediate upstream regions to roughly specify their differentiated expression patterns, while *RH2-3* has not (Figure 2). Given that ancestral RH2 before the duplications should have been expressed in the entire area of the retina, the *RH2-3* upstream region might represent ancestral regulatory function.

Also, upon the duplication events, the subtype genes obtained their own intrinsic identity, namely, genomic locations, particularly relative to RH2-LCR. Indeed, experiments have shown the importance of the relative positions in the differential expression. Most notably, translocation of RH2-LCR from the original position to immediately downstream of *RH2-3* on the PAC clone drastically altered the gene expression pattern: the expressions of *RH2-1* and *RH2-2* were mostly diminished; the expression of *RH2-3* was increased and extended broadly to the central-to-dorsal area; and the expression of *RH2-4* remained relatively unchanged in the ventral area of the retina (Tsujimura et al., 2007; 2015b). The observed positional effects should be primarily attributable to the position-dependent competition among the subtype genes for the enhancer activity (Figure 2).

Indeed, there is a highly context-dependent interference between the tandemly arrayed genes for the expression induction (Figure 2). For example, placing the upstream region of *RH2-3* between RH2-LCR and the *RH2-4* promoter in a reporter construct completely repressed the gene induction from the *RH2-4* promoter (Tsujimura et al., 2015b). Nonetheless, in the endogenous locus, *RH2-4* is located downstream of *RH2-3*. These facts indicate that the presence of *RH2-1* and *RH2-2* somehow suppresses the activity of the *RH2-3* promoter in the ventral retina (Figure 2). Supporting this idea, insertion of the *RH2-1* promoter between RH2-LCR and the *RH2-3* upstream region in a reporter construct repressed the gene induction from the *RH2-3* promoter in the ventral retina (Tsujimura et al., 2015b). Possibly, this ventral suppression by the presence of *RH2-1* and *RH2-2* helps release *RH2-4* from the blocking effect of *RH2-3*.

Importantly, when only the promoters/genes of *RH2-1* and *RH2-4* were arrayed in this order with RH2-LCR being present at the head, the expression of the two reporters from the promoters of *RH2-1* and *RH2-4* did not cover the entire area of the retina, leaving a void where no gene is expressed from the construct (Tsujimura et al., 2015b). The presence of *RH2-3*, the expression of which is flexibly adjustable depending on the surrounding contexts, might be essential to fill this void. To more profoundly and precisely understand the influence of one gene on the expression of other genes, the deletion and translocation of each gene should be carried out in future experiments.

Cis-regulation for LWS

Similarly, the LWS-activating region (LAR) was identified as a single enhancer that is required to fully activate the expression of both *LWS-1* and *LWS-2* (Tsujimura et al., 2010) (Figure 2). LAR is located upstream of *LWS-1*, and its deletion drastically reduces the expression of both *LWS-1* and *LWS-2* (Tsujimura et al., 2010). Since the deletion of LAR maintained a faint expression of LWS genes specifically in the LDCs, there should be other cis-regulatory regions that specify the cell type-specific expression of LWS outside of LAR (Tsujimura et al., 2010). The spatially differential expression is determined by the proximal cis-elements associated with the subtypes as well as by the competitive regulation for the shared enhancer, as is the case in RH2. While the *LWS-1* upstream sequence specifies the ventral expression, that of *LWS-2* does not have any areal specificity. Instead, it induces expression in the whole area when linked with LAR. However, when *LWS-1* is present between LAR and *LWS-2*, the gene expression from the *LWS-2* promoter is confined to the dorsal-to-central area where *LWS-1* is not expressed (Tsujimura et al., 2010) (Figure 2).

Thus, the cis-regulation of LWS is quite analogous to that of RH2. The shared features of the two systems provide comprehensive insights into the process of the differentiation (Figure 3). First, the sharing of the enhancer, which did not duplicate after the duplication events, triggers the subfunctionalization. The sharing of the enhancer intrinsically generates an asymmetric identity among the tandemly

arrayed genes in terms of the relative positions. Also, the co-regulation may immediately lead to a decrease in the expression levels of the duplicated genes. This down-regulation can help preserve both daughter genes for the dosage balance, as demonstrated by a recent study (Lan & Pritchard, 2016) (Figure 3B). Later during the evolutionary course, the DNA sequences associated with the different subtype genes should have accumulated mutations to enhance further the differentiation of the expression patterns (Figure 3C). One of the subtypes, i.e., *RH2-3* and *LWS-2* in RH2 and LWS, respectively, keeps the cis-elements for the induction in the whole part of the retina. The presence of these genes prevents the tandem array as a whole from leaving gaps where no genes are expressed.

The regulation of red and green opsin genes in humans and primates

The investigation into the cis-regulation of RH2 and LWS still leaves essential questions such as the following. How is the collective and competitive regulation for shared enhancers achieved in the opsin regulation? And, how are the cis-elements modified for the expression differentiation? In order to address these questions, it is critical to understand the enhancer regulation in general as well as the characteristics in the regulation for fish opsin genes. From this perspective, the regulation of human red and green opsin genes should serve as a valuable comparison with that in zebrafish.

Mammals are considered to have a less elaborate system of color vision than many other vertebrates. The ancestor of placental mammals and marsupials lost SWS2 and RH2, perhaps reflecting their nocturnal life (Ahnelt & Kolb, 2000). However, humans and other catarrhines have the LWS opsin gene duplicated (Ibbotson et al., 1992; Nathans et al., 1986). The duplicated LWS genes have undergone subfunctionalization and now encode genes for red- and green-sensitive visual pigments. As a result, catarrhines have trichromatic color vision, which should be beneficial to some aspects of their lives such as foraging for fruits (Melin et al., 2017) and social communications (Hiramatsu et al., 2017). Importantly, to acquire the trichromatic color vision, primates express red and green opsin genes in distinct sets of photoreceptor cells in a mosaic manner (Nathans, 1999). Had

316 photoreceptor cells co-expressed MWS and LWS, the distinction between red and
317 green would not have been achieved at the cellular level.
318 Human red- and green-sensitive cone opsin genes are arrayed in tandem on
319 the X chromosome. A genetic study for the monochromacy condition identified an
320 LCR as an essential enhancer for both opsin genes (Nathans et al., 1989). Other
321 functional genetic studies in mice elegantly showed that the competition for the
322 shared enhancer accounts for the mutually exclusive expression of the two subtype
323 opsin genes in the retina (Smallwood et al., 2002; Wang et al., 1999). Smallwood et
324 al. also showed that the modest mutual expression is only possible by the balanced
325 tension between the LCR activity, the promoter strength of the two genes, and their
326 positional relationship (Smallwood et al., 2002). Intriguingly, the promoter of the
327 green opsin, which is located further away from the LCR, has more potent activity
328 than the red opsin promoter. When the order of the two genes was switched, and
329 the green opsin was placed closer to the LCR, the red opsin gene was almost
330 completely shut down (Smallwood et al., 2002). This regulatory relationship
331 between the green and red opsin promoters is analogous to the RH2 regulation in
332 zebrafish, where the promoter of *RH2-3* has vigorous activity and can block the
333 action of RH2-LCR for other genes when located closest to it (Tsujimura et al., 2007;
334 2015b). Thus, both human and zebrafish systems are characterized by a competitive
335 regulation for shared enhancers, which could have facilitated the preservation and
336 differentiation of the duplicated genes (Figure 3).
337 On the other hand, the way of choosing subtypes for expression is differently
338 organized in zebrafish and primates. While zebrafish have a regional differentiation
339 pattern, primates choose the red or green opsins more or less randomly in the
340 retina to have the mosaic arrangement of the two as a prerequisite for trichromatic
341 vision. This difference should reflect the uniqueness of the cis-elements that
342 zebrafish and primates have accumulated through evolution. Also, physiological
343 demand should have contributed to shaping the ways of the differentiation.
344 Moreover, the difference might be attributable to the uniqueness of the trans-
345 regulatory mechanisms in the different species. In this sense, it is essential to

understand the trans-regulation for the differential expression of RH2 and LWS in zebrafish.

Trans-regulation for RH2 and LWS expression in zebrafish

One of the most important trans-regulators in photoreceptor cells in vertebrates is cone-rod homeobox (CRX) (Chen et al., 1997; Furukawa et al., 1997). CRX is essential for the expression of not only opsin genes but also many other photoreceptor-specific genes (Furukawa et al., 1997; 1999; Livesey et al., 2000; Shen & Raymond, 2004; Yamamoto et al., 2020). In zebrafish, knockdown of one of the *Crx* orthologues resulted in nearly complete loss of the expression of all four cone opsin types as well as rhodopsin in photoreceptor cells (Shen & Raymond, 2004). Consistently, the binding motifs of CRX are found in the SWS2 promoter (Takechi et al., 2008), RH2-LCR (Tsujimura et al., 2007), and LAR (Tsujimura et al., 2010). However, since *Crx* is expressed in all photoreceptor types and activates all cone and rod opsin genes, *Crx* alone does not account for photoreceptor type-specific regulation.

Several trans-regulatory factors have been identified as critical regulators for the cell type-specific expression of opsin genes in zebrafish. Thyroid hormone (TH) receptor B (*thrb*) is required for LWS expression (DuVal & Allison, 2018; Suzuki et al., 2013; Volkov et al., 2020), and the transcription factor Tbx2b is essential for SWS1 expression (Alvarez-Delfin et al., 2009). Also, it was shown that *six6* and *six7* regulate the expression of *SWS2* and RH2 genes (Ogawa et al., 2019; 2015). It is notable that SWS2 and RH2 are regulated at least partially by the same sets of transcription factors since it was reported that RH2-LCR sometimes induces weak gene expression ectopically in LSCs where *SWS2* is normally expressed (Fang et al., 2013; Tsujimura et al., 2007). Chromatin immunoprecipitation assays followed by high-throughput sequencing (ChIP-seq) showed that six6 and six7 bind to RH2-LCR and the *SWS2* promoter (Ogawa et al., 2019), which were functionally validated as important cis-regulators (Takechi et al., 2008; Tsujimura et al., 2007), linking the cis- and trans-regulation. Interestingly, six6 and six7 bind not only to RH2-LCR but also to other regions around the RH2 locus, including the promoter regions of *RH2-1*

and *RH2-2* (Ogawa et al., 2019). It is possible that these regions collectively function to establish the robust expression of RH2 genes.

These transcription factors might also be involved in the differential expression of the subtype choice for the RH2 and LWS expression. Because knockout experiments led to the almost complete loss of the expression of all subtype genes, it is impossible to conclude the role of *six6* and *six7* in the differential expression of RH2 (Ogawa et al., 2015; 2019). In this respect, experiments forcing a more moderate manipulation of the function of these transcription factors are required. Intriguingly, it was shown that the knockout of *six7* upregulated and down-regulated the expression of *LWS-1* and *LWS-2*, respectively, suggesting its role in the differential expression of LWS, but not of RH2 (Ogawa et al., 2015). The ChIP-seq data for six6 and six7 showed a broad but weak binding of the transcription factors around the LWS locus, though no single prominent binding site seems to exist with the exception of the promoter of adjacent *SWS2* (Ogawa et al., 2019). Since the differential expression could be recapitulated without the *SWS2* region (Tsujimura et al., 2010), it is not likely that six6 and six7 binding at the *SWS2* promoter has a significant role in the differential expression of LWS.

A series of recent studies showed that retinoic acid (RA) and TH signaling are critically involved in controlling the choice of the subtypes for RH2 and LWS expression in the retina (Mackin et al., 2019; Mitchell et al., 2015). Both RA (Prabhudesai et al., 2005; Wagner et al., 2000) and TH (Roberts et al., 2006) signaling are known to show a gradient in retinas in several species. These gradients seem to be generated mainly by gradients of catalytic enzymes, transporters, and other binding proteins for the molecules. There is also an accumulation of experimental evidence showing that RA and TH are critical in determining the dorsal-ventral patterning of the retina. For example, in the zebrafish retina, high-dose RA signaling is more associated with the ventral than dorsal identity (Hyatt et al., 1996; Marsh-Armstrong et al., 1994). Similarly, TH is involved in establishing the ventral identity in the mouse retina (Roberts et al., 2006). RH2 and LWS exhibit a similar differential expression pattern while expressed in distinct cone photoreceptor types. Therefore, it makes sense that the determinants for the axial

patterning of the whole retina are involved in the regulation (Figure 3C). Below, I briefly summarize the current knowledge of how LWS and RH2 are regulated by these signaling molecules.

Control of LWS expression by RA and TH

The roles of RA and TH in the differential expression were revealed by both gain- and loss-of-function experiments. As a gain-of-function experiment, exogenous administration to zebrafish was carried out to show that both RA and TH increased and expanded the expression of *LWS-1* while reducing the expression of *LWS-2* (Mackin et al., 2019; Mitchell et al., 2015). Remarkably, upon treatment, the expression area of *LWS-1* invades the central-to-dorsal area of the retina, where *LWS-2* is usually expressed. These results are consistent with previous findings that both RA and TH induce ventralization of the retina (Hyatt et al., 1996; Marsh-Armstrong et al., 1994; Roberts et al., 2006). Thus, it is plausible that the *LWS-1* promoter carrying the cis-elements to specify the ventral expression responded to these stimulations.

It should also be noted that the expression change of *LWS-2* does not seem to precede that of *LWS-1* after the stimulations. For example, detecting the decrease of *LWS-2* upon RA administration takes longer than detecting the *LWS-1* upregulation (Mitchell et al., 2015). Also, the induction of a dominant-negative form of the RA receptor RARα did not result in a significant upregulation of *LWS-2* but did repress *LWS-1* (Mitchell et al., 2015). Similarly, upon TH treatment, live imaging experiments revealed that *LWS-2* persists after the expression onset of *LWS-1* (Mackin et al., 2019). Although these experiments could not fully describe the precise order of the expression switch, the results seem very consistent with the fact that *LWS-2* expression is determined secondarily and exclusively from the area of *LWS-1* expression (Tsujimura et al., 2010).

The transcription factors that mediate the action of RA and TH to bind to and regulate LWS directly are mostly unknown. Studies have indicated that the LWS array carries several RA response elements (RAREs) and TH response elements (TREs) (Mackin et al., 2019; Mitchell et al., 2015), some of which are also located

within LAR. Yet the roles of these elements have not been tested. It is also possible that other transcription factors induced by RA and TH affect LWS expression. A recent transcriptomic analysis in zebrafish retinas upon the knockout of *thrb* found only five genes are down-regulated, among which three (LWS-1, LWS-2, and miR-726, a conserved microRNA gene located between SWS2 and LAR) are from the LWS locus (Volkov et al., 2020). These results may indicate that the TH receptor directly regulates LWS expression. However, it should be noted that knockout and knockdown of *thrb* result in nearly complete loss of the red cones, LDCs (DuVal & Allison, 2018; Suzuki et al., 2013; Volkov et al., 2020). Therefore, it remains elusive how the receptor controls LWS expression in LDCs.

Upon the loss-of-function of endogenous TH, the expression area of *LWS-1* was markedly reduced but still remained in the ventral retina, and strong RA signaling was observed (Mackin et al., 2019; Mitchell et al., 2015). This result suggests that RA can induce the ventral expression of *LWS-1* independently of TH signaling. On the other hand, the administration of RA increased the expression of *dio2*, a gene encoding a catalytic enzyme that converts T4 into T3, the active form of TH (Mitchell et al., 2015). Thus, it seems RA and TH impinge on opsin expression through both shared and distinct pathways.

Control of RH2 by TH

Regarding RH2 differential expression, only the effects of TH signaling have been studied so far as trans-regulators (Mackin et al., 2019). Given that the ventral induction by TH administration is key for the activation of *LWS-1* (Mackin et al., 2019), it is expected that the same induction would activate *RH2-4*, which has cis-elements for the ventral expression, while repressing *RH2-1* and *RH2-2*, which have cis-elements that repress ventral expression (Figure 2). However, this was not exactly the case. The expression change of *RH2-2* is different from that of *RH2-1* upon the manipulation of TH signaling. In larvae, the administration of exogenous TH reduced the expression of *RH2-1*, but upregulated *RH2-2*, and the expressions of *RH2-3* and *RH2-4* were slightly upregulated (Mackin et al., 2019). The same study found that in juveniles, *RH2-1* was strongly down-regulated and *RH2-2* was only

slightly down-regulated by the TH induction. Further, when the loss of TH function was forced, *RH2-2* was down-regulated, but *RH2-1* was not. Moreover, administering TH to rescue the loss-of-function resulted in a significant upregulation of *RH2-1*, but not of *RH2-2*. As for the responses of *RH2-3* and *RH2-4*, the inhibition of TH significantly down-regulated *RH2-3* and *RH2-4*, and TH administration either kept or upregulated the two genes.

Thus, the pattern of the expression change of *RH2-2* does not seem to be directly determined by the upstream sequence. Instead, the Mackin study might suggest that together with *RH2-3*, *RH2-2* acts as an intermediate between the expression change of *RH2-1* and *RH2-4*. As described above, one study investigating the cis-regulation of the RH2 locus indicated that the regulatory elements embedded throughout the locus interact and interfere with each other to finally establish the regulatory potential for the subtype genes (Tsujimura et al., 2015b) (Figure 2). Along this line, the observed responses to the TH manipulations could also be an outcome of such collective regulation. Analogously to the responses to TH, the *RH2-3* expression area is between the *RH2-1*/*RH2-2* expression and *RH2-4* expression areas at the normal state. The trends commonly seen in both normal and TH challenged states may indicate that interactions among the cis-elements at the locus inevitably create a graded regulatory potential along the locus from *RH2-1* to *RH2-4*.

Intriguingly, a similar gradation of the regulatory potential is seen at the beta-globin locus in humans. This locus consists of the ε, Gγ, Aγ, δ, and β types of globin genes arranged in this order. At the head of the cluster and upstream of ε, is located the LCR, which controls the collective regulation of the globin genes. The expressed genes switch along the ontogeny of the erythroid cells. In the embryonic yolk sac, ε is expressed. Then in the fetus liver, γ are expressed. In the adult bone marrow, β and δ are strongly and weakly expressed, respectively (Noordermeer & de Laat, 2008).

As with the RH2 and LWS cases, the temporal specificity of the expression is partly encoded in the promoter sequences. The promoter of ε can drive gene expression specifically in yolk-sac-derived erythroid cells, but not later in the fetal liver or adult erythroid cells, when linked with the LCR (Raich et al., 1990). The

promoter of γ linked with the LCR could induce embryonic and fetal expression, but not adult expression (Dillon & Grosveld, 1991). On the other hand, the β globin promoter can induce gene expression throughout all stages (Behringer et al., 1990).

The functional importance of the genomic arrangement of the globin genes with respect to the LCR was investigated by several studies, which altogether showed that reciprocal competition for the LCR activity in a gene-order-dependent manner also underlies the differential expression (Behringer et al., 1990; Hanscombe et al., 1991; Okamura et al., 2009; Tanimoto et al., 1999). The competition and interaction among promoters provide the beta-globin locus with the temporally graded regulatory potential along the genomic coordinate (Foley & Engel, 1992). Thus, the establishment of a graded regulatory potential as the outcome of the collective ensemble of effectively arranged cis-elements might be a common phenomenon to achieve gradually differentiated expression patterns of tandemly arrayed genes.

Enigmas in cis-interactions

Above, I explain that the competitive regulation and collective ensemble of cis-elements specifying regional identities are key to the differential expression of RH2 and LWS genes (Figures 2 and 3). However, the underlying mechanism for the cis-interactions is elusive.

Competitive regulation among tandemly copied genes has also been seen in other loci including olfactory receptor genes (Fuss et al., 2007; Nishizumi et al., 2007; Serizawa et al., 2003). Also, a synthetic configuration with two genes sandwiching a shared enhancer exhibited competitive regulation in *Drosophila melanogaster* (Fukaya et al., 2016). Interestingly, the mutually exclusive regulation in this synthetic system was only observed when the two genes were together asymmetrically placed with the enhancer (Fukaya et al., 2016). When placed at the same distance from the enhancer, the two genes were synchronously activated (Fukaya et al., 2016). Of note, the simultaneous activation led to the idea that transcriptional bursting induced by an enhancer could involve phase separation (Hnisz et al., 2017). In this sense, the competitive regulation might be a result of the

exclusion of one gene from the active transcriptional spot that is preferentially formed with the other genes.

Interestingly, such repulsive regulation was observed even between two genes belonging to different families at the locus of *Tfap2c* and *Bmp7*, where the two genes are located in cis but regulated by distinct enhancers in the mouse forebrain (Tsujimura et al., 2015a). The observed mutual exclusion between *Tfap2c* and *Bmp7* in the forebrain might also be explained by the exclusion of one gene from the neighboring transcriptionally active spot. However, it remains unclear how the competitive regulation emerges.

Also, while a collection of multiple cis-elements such as enhancers seemingly cooperate with each other to make robust and stereotypic expression patterns of genes (Marinić et al., 2013; Montavon et al., 2011), the mechanics underlying the interaction between cis-regulatory elements are missing. Therefore, it is currently impossible to explain how multiple elements in cis interact with each other to produce the differential expression at the RH2 and LWS loci.

To elucidate the regulatory effects of cis-interactions, it is essential to simultaneously capture the epigenetic states of the cis-regulatory regions, the interaction patterns among them, and the transcriptional states of the genes in individual cells. Improving and combining different techniques to study single-cell genomics would be one approach. Imaging-based assays should also be pursued to follow the dynamics of the players in the gene regulations. Further, the roles of cis-regulatory elements should be functionally and comprehensively tested via genome engineering. Applying these diverse and cutting-edge techniques should reveal how multiple cis-regulatory elements interact with each other to establish coordinated gene expressions at various genomic loci, including the RH2 and LWS loci in zebrafish.

Conclusions and outlook

The subfunctionalization of RH2 and LWS has arisen from complex transcriptional regulation (Figures 2 and 3). Notably, the two loci have adopted competitive regulation by a shared enhancer (Figure 3B). This system should have

563 assisted in preserving the duplicated genes for dosage balancing at the early stage
564 and facilitated differentiation among daughters at the later stage (Figure 3B).
565 Further, the proximal sequences around the individual genes acquired cis-elements
566 that collectively produce the differential expression among subtypes (Figure 3C).
567 Importantly, the differentiation seems to involve a pre-existing mechanism that
568 specifies spatial identity across the retina, such as the dorsal-ventral patterning
569 caused by RA and TH (Figure 3C).
570 However, there remain many open questions. First of all, it is uncertain how
571 competition emerges among tandemly arrayed genes for a shared enhancer. Though
572 such regulation is seen at many loci, the mechanism is totally elusive. Therefore, it is
573 difficult to determine whether RH2 and LWS required special conditions to achieve
574 such regulation. Also, the underlying mechanism for the multiple elements in cis to
575 interact with each other for the robustly determined expression patterns is
576 unknown. Since the shuffling of gene orders at the RH2 locus disrupted the current
577 regulation pattern, the arrangement and repertoire of the cis-elements should be a
578 result of evolutionary optimization. However, the grammar behind the cis-
579 interaction needs to be clarified further.
580 Therefore, future studies should take two approaches. On the one hand, it is
581 crucial to further analyze the regulation of RH2 and LWS. So far, only a few players,
582 including RA and TH, have been shown to regulate the differential expression as
583 trans-regulators. Trans-species transgenic assays, in which the genomic clones of
584 zebrafish RH2 and LWS loci are introduced into other fish species such as medaka,
585 would help address how conserved the trans-regulatory mechanism is for the
586 differential expression. Further, cis-elements at both RH2 and LWS loci that directly
587 or indirectly respond to these trans-regulators need to be precisely identified. In
588 this sense, dissecting further the genomic regions around RH2 and LWS is critical
589 for pinpointing the essential sequences that respond to RA and TH. Identifying the
590 cis-elements will help clarify the transcription factors that bind to them.
591 Insights into transcriptional regulation in general are also needed. Recent
592 progress in the field of gene-enhancer interactions has contributed significantly to
593 our understanding of how these interactions are regulated in terms of the 3D

genome (Dekker et al., 2017). However, interactions involving multiple genomic regions are not well understood. Overall, accumulating more knowledge about cis-interactions in the genome should contribute to our understanding of the differential regulation of zebrafish cone opsin genes as well as the evolution of duplicated genes in general.

Figure legends

Figure 1

The repertoire of cone opsin genes in zebrafish.

(A) The four types of cone opsin genes are specifically expressed in distinct types of photoreceptor cells in the retina. (B) The genomic arrangement of the cone opsin genes in zebrafish, together with the λmax and the spatial expression pattern along the dorsal-ventral axis in the adult retina. Note that zebrafish have two and four subtypes of LWS and RH2, respectively. The subtypes are differentiated in both absorption spectra and expression patterns.

Figure 2

A schematic illustration of the cis-regulation for the differential expression of RH2 and LWS.

In both the dorsal and ventral retina, the activity of promoters (halved-ellipses) and the expression states (boxes with arrows) are shown for each gene of RH2 and LWS, as indicated in the bottom table.

Figure 3

A proposed scenario for the subfunctionalization of the duplicated opsin genes.

(A) Before the duplications, the expression of a single ancestral gene was regulated by an ancestral enhancer. (B) Gene duplication outside of the enhancer region resulted in competitive regulation. Although the expression differentiation should not be necessarily prominent at this point, the expression level could be down-regulated due to competition for the shared enhancer. This down-regulation may

favor the preservation of both daughter genes (Lan & Pritchard, 2016). (C) Later during the evolutionary course, some of the subtype genes acquired cis-elements to differentiate the expression pattern. The system may have co-opted the pre-existing mechanism for the dorsal-ventral patterning involving the RA and TH signaling. As a result, the duplicated genes accomplished subfunctionalization.

References

Ahnelt, P.K., & Kolb, H. (2000). The mammalian photoreceptor mosaic-adaptive design. *Progress in Retinal and Eye Research* **19**, 711–777.

Allison, W.T., HAIMBERGER, T.J., Hawryshyn, C.W., & TEMPLE, S.E. (2004). Visual pigment composition in zebrafish: Evidence for a rhodopsin-porphyropsin interchange system. *Vis. Neurosci.* **21**, 945–952.

Alvarez-Delfin, K., Morris, A.C., Snelson, C.D., Gamse, J.T., Gupta, T., Marlow, F.L., Mullins, M.C., Burgess, H.A., Granato, M., & Fadool, J.M. (2009). Tbx2b is required for ultraviolet photoreceptor cell specification during zebrafish retinal development. *Proc. Natl. Acad. Sci. U.S.a.* **106**, 2023–2028.

Baden, T., Euler, T., & Berens, P. (2020). Understanding the retinal basis of vision across species. *Nature Reviews Neuroscience 2015 17:1* **21**, 5–20.

Behringer, R.R., Ryan, T.M., Palmiter, R.D., Brinster, R.L., & Townes, T.M. (1990). Human gamma- to beta-globin gene switching in transgenic mice. *Genes Dev.* **4**, 380–389.

Carleton, K.L., Spady, T.C., Streelman, J.T., Kidd, M.R., McFarland, W.N., & Loew, E.R. (2008). Visual sensitivities tuned by heterochronic shifts in opsin gene expression. *BMC Biol.* **6**, 22–14.

Chen, S., Wang, Q.-L., Nie, Z., Sun, H., Lennon, G., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., & Zack, D.J. (1997). Crx, a Novel Otx-like Paired-Homeodomain Protein, Binds to and Transactivates Photoreceptor Cell-Specific Genes. *Neuron* **19**, 1017–1030.

Chinen, A., Hamaoka, T., Yamada, Y., & Kawamura, S. (2003). Gene Duplication and Spectral Diversification of Cone Visual Pigments of Zebrafish. *Genetics* **163**, 663–675.

Dekker, J., Belmont, A.S., Guttman, M., Leshyk, V.O., Lis, J.T., Lomvardas, S., Mirny, L.A., O'Shea, C.C., Park, P.J., Ren, B., Politz, J.C.R., Shendure, J., Zhong, S., & Network, T.4.N. (2017). The 4D nucleome project. *Nature* **549**, 219–226.

Dillon, N., & Grosveld, F. (1991). Human γ-globin genes silenced independently of other genes in the β-globin locus. *Nature* **350**, 252–254.

DuVal, M.G., & Allison, W.T. (2018). Photoreceptor Progenitors Depend Upon Coordination of gdf6a, thrβ, and tbx2b to Generate Precise Populations of Cone Photoreceptor Subtypes. *Invest. Ophthalmol. Vis. Sci.* **59**, 6089–6101.

Enright, J.M., Toomey, M.B., Sato, S.-Y., et al. (2015). Cyp27c1 Red-Shifts the Spectral Sensitivity of Photoreceptors by Converting Vitamin A1 into A2. *Current Biology* **25**, 3048–3057.

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 Expression Patterns* **13**, 150–159.

Foley, K.P., & Engel, J.D. (1992). Individual stage selector element mutations lead to reciprocal changes in beta- vs. epsilon-globin gene transcription: genetic confirmation of promoter competition during globin gene switching. *Genes Dev.* **6**, 730–744.

Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.L., & Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**, 1531–1545.

Fukaya, T., Lim, B., & Levine, M. (2016). Enhancer Control of Transcriptional Bursting. *Cell* **166**, 358–368.

Furukawa, T., Morrow, E.M., & Cepko, C.L. (1997). Crx, a Novel otx-like Homeobox Gene, Shows Photoreceptor-Specific Expression and Regulates Photoreceptor Differentiation. *Cell* **91**, 531–541.

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

Fuss, S.H., Omura, M., & Mombaerts, P. (2007). Local and cis Effects of the H Element on Expression of Odorant Receptor Genes in Mouse. *Cell* **130**, 373–384.

Hamaoka, T., Takechi, M., Chinen, A., Nishiwaki, Y., & Kawamura, S. (2002). Visualization of rod photoreceptor development using GFP-transgenic zebrafish. *Genesis* **34**, 215–220.

Hanscombe, O., Whyatt, D., Fraser, P., Yannoutsos, N., Greaves, D., Dillon, N., & Grosveld, F. (1991). Importance of globin gene order for correct developmental expression. *Genes Dev.* **5**, 1387–1394.

Hiramatsu, C., Melin, A.D., Allen, W.L., Dubuc, C., & Higham, J.P. (2017). Experimental evidence that primate trichromacy is well suited for detecting primate social colour signals. *Proc. Biol. Sci.* **284**, 20162458.

Hnisz, D., Shrinivas, K., Young, R.A., Chakraborty, A.K., & Sharp, P.A. (2017). A Phase Separation Model for Transcriptional Control. *Cell* **169**, 13–23.

Hoffmann, M., Tripathi, N., Henz, S.R., Lindholm, A.K., Weigel, D., Breden, F., & Dreyer, C. (2007). Opsin gene duplication and diversification in the guppy, a model for sexual selection. *Proc. Biol. Sci.* **274**, 33–42.

701 Hofmann, C.M., & Carleton, K.L. (2009). Gene duplication and differential gene
702 expression play an important role in the diversification of visual pigments in fish.
703 *Integr. Comp. Biol.* **49**, 630–643.

704 Hyatt, G.A., Schmitt, E.A., Marsh-Armstrong, N., McCaffery, P., Dräger, U.C., &
705 Dowling, J.E. (1996). Retinoic acid establishes ventral retinal characteristics.
706 *Development* **122**, 195–204.

707 Ibbotson, R.E., Hunt, D.M., Bowmaker, J.K., & Mollon, J.D. (1992). Sequence
708 divergence and copy number of the middle- and long-wave photopigment genes in
709 Old World monkeys. *Proc. Biol. Sci.* **247**, 145–154.

710 Kennedy, B.N., Vihtelic, T.S., Checkley, L., Vaughan, K.T., & Hyde, D.R. (2001).
711 Isolation of a zebrafish rod opsin promoter to generate a transgenic zebrafish line
712 expressing enhanced green fluorescent protein in rod photoreceptors. *J. Biol. Chem.*
713 **276**, 14037–14043.

714 Kleinjan, D.A., Bancewicz, R.M., Gautier, P., Dahm, R., Schonthaler, H.B., Damante, G.,
715 Seawright, A., Hever, A.M., Yeyati, P.L., van Heyningen, V., & Coutinho, P. (2008).
716 Subfunctionalization of Duplicated Zebrafish pax6 Genes by cis-Regulatory
717 Divergence. *PLoS Genet.* **4**, e29.

718 Kuraku, S., Meyer, A., & Kuratani, S. (2009). Timing of genome duplications relative
719 to the origin of the vertebrates: did cyclostomes diverge before or after? *Mol. Biol.*
720 *Evol.* **26**, 47–59.

721 Lagman, D., Ocampo Daza, D., Widmark, J., Abalo, X.M., Sundström, G., & Larhammar,
722 D. (2013). The vertebrate ancestral repertoire of visual opsins, transducin alpha
723 subunits and oxytocin/vasopressin receptors was established by duplication of their
724 shared genomic region in the two rounds of early vertebrate genome duplications.
725 *BMC Evol Biol* **13**, 238.

726 Lamb, T.D., Collin, S.P., & Pugh, E.N. (2007). Evolution of the vertebrate eye: opsins,
727 photoreceptors, retina and eye cup. *Nature Reviews Neuroscience 2015 17:1* **8**, 960–
728 976.

729 Lan, X., & Pritchard, J.K. (2016). Coregulation of tandem duplicate genes slows
730 evolution of subfunctionalization in mammals. *Science* **352**, 1009–1013.

731 Lin, J.-J., Wang, F.-Y., Li, W.-H., & Wang, T.-Y. (2017). The rises and falls of opsin
732 genes in 59 ray-finned fish genomes and their implications for environmental
733 adaptation. *Sci. Rep.* **7**, 1–13.

734 Livesey, F.J., Furukawa, T., Steffen, M.A., Church, G.M., & Cepko, C.L. (2000).
735 Microarray analysis of the transcriptional network controlled by the photoreceptor
736 homeobox gene Crx. *Current Biology* **10**, 301–310.

737 Mackin, R.D., Frey, R.A., Gutierrez, C., Farre, A.A., Kawamura, S., Mitchell, D.M., &
738 Stenkamp, D.L. (2019). Endocrine regulation of multichromatic color vision. *Proc.*
739 *Natl. Acad. Sci. U.S.a.* **116**, 16882–16891.

740 Marinić, M., Aktas, T., Ruf, S., & Spitz, F. (2013). An integrated holo-enhancer unit
741 defines tissue and gene specificity of the Fgf8 regulatory landscape. *Developmental*
742 *Cell* **24**, 530–542.

743 Marsh-Armstrong, N., McCaffery, P., Gilbert, W., Dowling, J.E., & Dräger, U.C. (1994).
744 Retinoic acid is necessary for development of the ventral retina in zebrafish. *Proc.*
745 *Natl. Acad. Sci. U.S.a.* **91**, 7286–7290.

746 Matsumoto, Y., Fukamachi, S., Mitani, H., & Kawamura, S. (2006). Functional
747 characterization of visual opsin repertoire in Medaka (Oryzias latipes). *Gene* **371**,
748 268–278.

749 Melin, A.D., Chiou, K.L., Walco, E.R., Bergstrom, M.L., Kawamura, S., & Fedigan, L.M.
750 (2017). Trichromacy increases fruit intake rates of wild capuchins (Cebus capucinus
751 imitator). *Proc. Natl. Acad. Sci. U.S.a.* **114**, 10402–10407.

752 Mitchell, D.M., Stevens, C.B., Frey, R.A., Hunter, S.S., Ashino, R., Kawamura, S., &
753 Stenkamp, D.L. (2015). Retinoic Acid Signaling Regulates Differential Expression of
754 the Tandemly-Duplicated Long Wavelength-Sensitive Cone Opsin Genes in
755 Zebrafish. *PLoS Genet.* **11**, e1005483.

756 Montavon, T., Soshnikova, N., Mascrez, B., Joye, E., Thevenet, L., Splinter, E., de Laat,
757 W., Spitz, F., & Duboule, D. (2011). A regulatory archipelago controls Hox genes
758 transcription in digits. *Cell* **147**, 1132–1145.

759 Nathans, J. (1999). The evolution and physiology of human color vision: insights
760 from molecular genetic studies of visual pigments. *Neuron* **24**, 299–312.

761 Nathans, J., Davenport, C.M., Maumenee, I.H., Lewis, R.A., Hejtmancik, J.F., Litt, M.,
762 Lovrien, E., Weleber, R., Bachynski, B., Zwas, F., & al, E. (1989). Molecular genetics of
763 human blue cone monochromacy. *Science* **245**, 831–838.

764 Nathans, J., Thomas, D., & Hogness, D.S. (1986). Molecular genetics of human color
765 vision: the genes encoding blue, green, and red pigments. *Science* **232**, 193–202.

766 Nishizumi, H., Kumasaka, K., Inoue, N., Nakashima, A., & Sakano, H. (2007). Deletion
767 of the core-H region in mice abolishes the expression of three proximal odorant
768 receptor genes in cis. *Proc. Natl. Acad. Sci. U.S.a.* **104**, 20067–20072.

769 Noordermeer, D., & de Laat, W. (2008). Joining the loops: β-Globin gene regulation.
770 *IUBMB Life* **60**, 824–833.

771 Nowak, M.A., Boerlijst, M.C., Cooke, J., & Smith, J.M. (1997). Evolution of genetic
772 redundancy. *Nature* **388**, 167–171.

773 Ogawa, Y., Shiraki, T., Asano, Y., Muto, A., Kawakami, K., Suzuki, Y., Kojima, D., &
774 Fukada, Y. (2019). Six6 and Six7 coordinately regulate expression of middle-
775 wavelength opsins in zebrafish. *Proc. Natl. Acad. Sci. U.S.a.* **116**, 4651–4660.

776 Ogawa, Y., Shiraki, T., Kojima, D., & Fukada, Y. (2015). Homeobox transcription
777 factor Six7 governs expression of green opsin genes in zebrafish. *Proc. Biol. Sci.* **282**,
778 20150659.

779 Okamura, E., Matsuzaki, H., Campbell, A.D., Engel, J.D., Fukamizu, A., & Tanimoto, K.
780 (2009). All of the human β-type globin genes compete for LCR enhancer activity in
781 embryonic erythroid cells of yeast artificial chromosome transgenic mice. *The*
782 *FASEB Journal* **23**, 4335–4343.

783 Owens, G.L., Rennison, D.J., Allison, W.T., & Taylor, J.S. (2012). In the four-eyed fish
784 (Anableps anableps), the regions of the retina exposed to aquatic and aerial light do
785 not express the same set of opsin genes. *Biology Letters* **8**, 86–89.

786 Prabhudesai, S.N., Cameron, D.A., & Stenkamp, D.L. (2005). Targeted effects of
787 retinoic acid signaling upon photoreceptor development in zebrafish. *Developmental*
788 *Biology* **287**, 157–167.

789 Raich, N., Enver, T., Nakamoto, B., Josephson, B., Papayannopoulou, T., &
790 Stamatoyannopoulos, G. (1990). Autonomous developmental control of human
791 embryonic globin gene switching in transgenic mice. *Science* **250**, 1147–1149.

792 Rennison, D.J., Owens, G.L., Allison, W.T., & Taylor, J.S. (2011). Intra-retinal variation
793 of opsin gene expression in the guppy (Poecilia reticulata). *J. Exp. Biol.* **214**, 3248–
794 3254.

795 Rister, J., & Desplan, C. (2011). The retinal mosaics of opsin expression in
796 invertebrates and vertebrates. *Developmental Neurobiology* **71**, 1212–1226.

797 Roberts, M.R., Srinivas, M., Forrest, D., de Escobar, G.M., & Reh, T.A. (2006). Making
798 the gradient: Thyroid hormone regulates cone opsin expression in the developing
799 mouse retina. *Proc. Natl. Acad. Sci. U.S.a.* **103**, 6218–6223.

800 Robinson, J., Schmitt, E.A., Hárosi, F.I., Reece, R.J., & Dowling, J.E. (1993). Zebrafish
801 ultraviolet visual pigment: absorption spectrum, sequence, and localization. *Proc.*
802 *Natl. Acad. Sci. U.S.a.* **90**, 6009–6012.

803 Serizawa, S., Miyamichi, K., Nakatani, H., Suzuki, M., Saito, M., Yoshihara, Y., &
804 Sakano, H. (2003). Negative feedback regulation ensures the one receptor-one
805 olfactory neuron rule in mouse. *Science* **302**, 2088–2094.

806 Shen, Y.-C., & Raymond, P.A. (2004). Zebrafish cone-rod (crx) homeobox gene
807 promotes retinogenesis. *Developmental Biology* **269**, 237–251.

808 Shichida, Y., & Matsuyama, T. (2009). Evolution of opsins and phototransduction.
809 *Phil. Trans. R. Soc. B* **364**, 2881–2895.

810 Smallwood, P.M., Wang, Y., & Nathans, J. (2002). Role of a locus control region in the
811 mutually exclusive expression of human red and green cone pigment genes. *Proc.*
812 *Natl. Acad. Sci. U.S.a.* **99**, 1008–1011.

813 Suzuki, S.C., Bleckert, A., Williams, P.R., Takechi, M., Kawamura, S., & Wong, R.O.L.
814 (2013). Cone photoreceptor types in zebrafish are generated by symmetric terminal
815 divisions of dedicated precursors. *Proc. Natl. Acad. Sci. U.S.a.* **110**, 15109–15114.

816 Takechi, M., & Kawamura, S. (2005). Temporal and spatial changes in the expression
817 pattern of multiple red and green subtype opsin genes during zebrafish
818 development. *Journal of Experimental Biology* **208**, 1337–1345.

819 Takechi, M., Hamaoka, T., & Kawamura, S. (2003). Fluorescence visualization of
820 ultraviolet-sensitive cone photoreceptor development in living zebrafish. *FEBS*
821 *Letters* **553**, 90–94.

822 Takechi, M., Seno, S., & Kawamura, S. (2008). Identification of cis-Acting Elements
823 Repressing Blue Opsin Expression in Zebrafish UV Cones and Pineal Cells. *J. Biol.*
824 *Chem.* **283**, 31625–31632.

825 Tanimoto, K., Liu, Q., Bungert, J., & Engel, J.D. (1999). Effects of altered gene order or
826 orientation of the locus control region on human beta-globin gene expression in
827 mice. *Nature* **398**, 344–348.

828 Temple, S.E. (2011). Why different regions of the retina have different spectral
829 sensitivities: a review of mechanisms and functional significance of intraretinal
830 variability in spectral sensitivity in vertebrates. *Vis. Neurosci.* **28**, 281–293.

831 Tsujimura, T., Chinen, A., & Kawamura, S. (2007). Identification of a locus control
832 region for quadruplicated green-sensitive opsin genes in zebrafish. *Proc. Natl. Acad.*
833 *Sci. U.S.a.* **104**, 12813–12818.

834 Tsujimura, T., Hosoya, T., & Kawamura, S. (2010). A single enhancer regulating the
835 differential expression of duplicated red-sensitive opsin genes in zebrafish. *PLoS*
836 *Genet.* **6**, e1001245.

837 Tsujimura, T., Klein, F.A., Langenfeld, K., Glaser, J., Huber, W., & Spitz, F. (2015a). A
838 Discrete Transition Zone Organizes the Topological and Regulatory Autonomy of the
839 Adjacent Tfap2c and Bmp7 Genes. *PLoS Genet.* **11**, e1004897.

840 Tsujimura, T., Masuda, R., Ashino, R., & Kawamura, S. (2015b). Spatially
841 differentiated expression of quadruplicated green-sensitive RH2 opsin genes in
842 zebrafish is determined by proximal regulatory regions and gene order to the locus
843 control region. *BMC Genetics* **16**, 130.

844 Vihtelic, T.S., Doro, C.J., & Hyde, D.R. (1999). Cloning and characterization of six
845 zebrafish photoreceptor opsin cDNAs and immunolocalization of their
846 corresponding proteins. *Vis. Neurosci.* **16**, 571–585.

847 Volkov, L.I., Kim-Han, J.S., Saunders, L.M., Poria, D., Hughes, A.E.O., Kefalov, V.J.,
848 Parichy, D.M., & Corbo, J.C. (2020). Thyroid hormone receptors mediate two distinct
849 mechanisms of long-wavelength vision. *Proc. Natl. Acad. Sci. U.S.a.* **Vol.VII**,
850 201920086.

851 Wagner, E., McCaffery, P., & Dräger, U.C. (2000). Retinoic Acid in the Formation of
852 the Dorsoventral Retina and Its Central Projections. *Developmental Biology* **222**,
853 460–470.

854 Wang, Y., Smallwood, P.M., Cowan, M., Blesh, D., Lawler, A., & Nathans, J. (1999).
855 Mutually exclusive expression of human red and green visual pigment-reporter
856 transgenes occurs at high frequency in murine cone photoreceptors. *Proc. Natl.*
857 *Acad. Sci. U.S.a.* **96**, 5251–5256.

858 Yamamoto, H., Kon, T., Omori, Y., & Furukawa, T. (2020). Functional and
859 Evolutionary Diversification of Otx2 and Crx in Vertebrate Retinal Photoreceptor
860 and Bipolar Cell Development. *Cell Reports* **30**, 658–671.e5.

861 Yokoyama, S. (2000). Molecular evolution of vertebrate visual pigments. *Progress in*
862 *Retinal and Eye Research* **19**, 385–419.

863 Zhang, J. (2003). Evolution by gene duplication: an update. *Trends in Ecology &*
864 *Evolution* **18**, 292–298.

865

Fig. 1

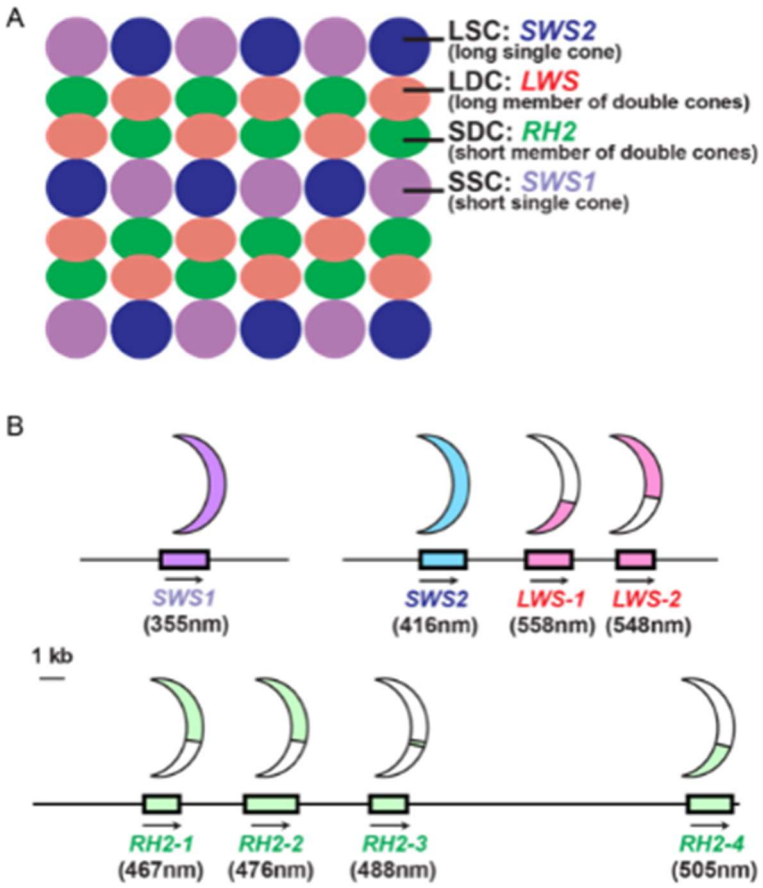

A
LSC: SWS2
(long single cone)
LDC: LWS
(long member of double cones)
SDC: RH2
(short member of double cones)
SSC: SWS1
(short single cone)
B
SWS1
(355nm)
SWS2
(416nm)
LWS-1
(558nm)
LWS-2
(548nm)
1 kb
RH2-1
(467nm)
RH2-2
(476nm)
RH2-3
(488nm)
RH2-4
(505nm)

Fig. 2

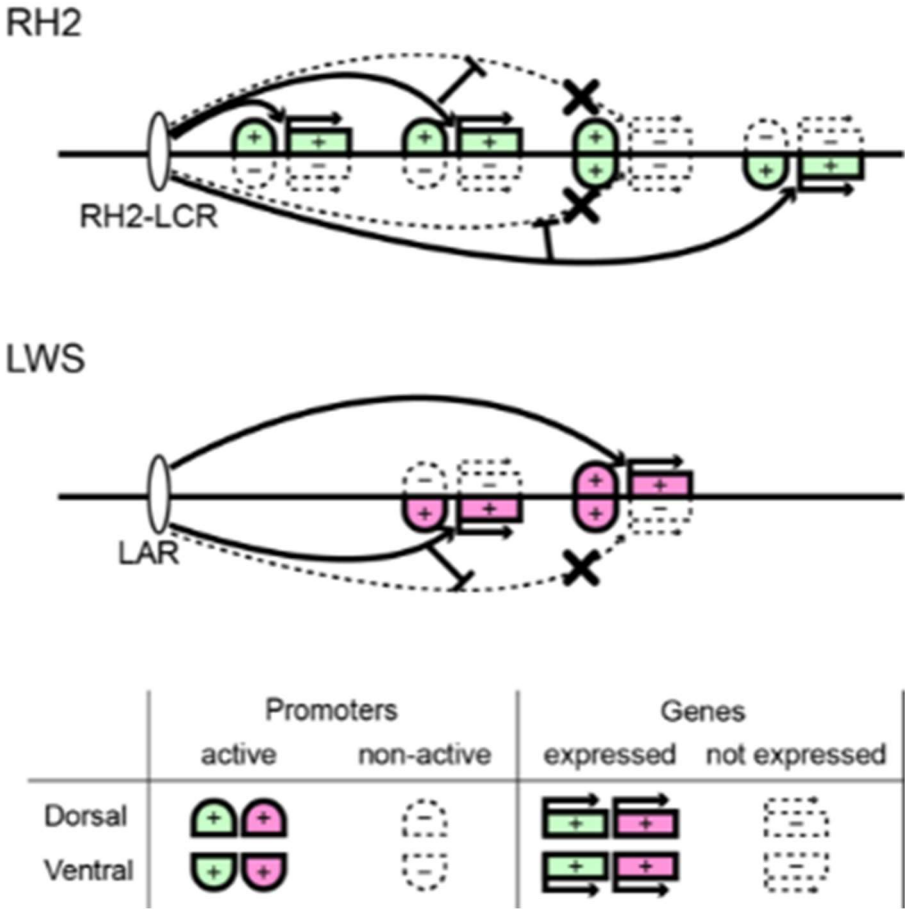
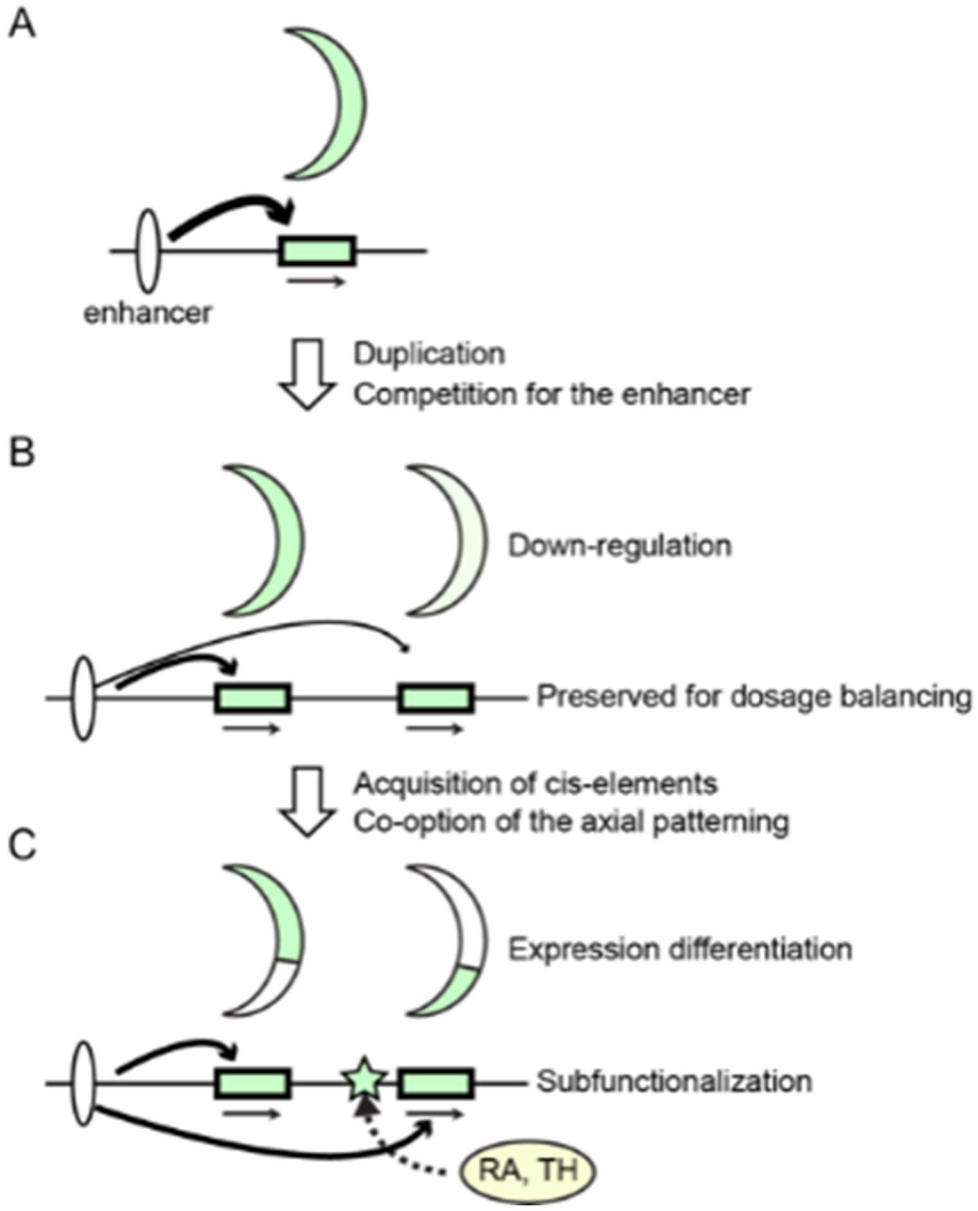

RH2
RH2-LCR
LWS
LAR
Promoters
active
non-active
Genes
expressed
not expressed
Dorsal
Ventral

Fig. 3


A
enhancer
Duplication
Competition for the enhancer
B
Down-regulation
Preserved for dosage balancing
Acquisition of cis-elements
Co-option of the axial patterning
C
Expression differentiation
Subfunctionalization
RA, TH